CFRM 504: Lecture Notes

Matthew Lorig [1]

This version: December 2, 2020

[1]Department of Applied Mathematics, University of Washington, Seattle, WA, USA. e-mail: mlorig@uw.edu

Contents

Preface

These notes are intended to give Masters students in Computational Finance and Risk Management a broad introduction to financial derivatives. Because the focus of this course is on the *applied* aspects of this topic, we will sometimes forgo mathematical rigor, favoring instead a heuristic development. The mathematical statements in these notes should be taken as "true in spirit," but perhaps not always rigorously true mathematically. The hope is that, what the notes lack in rigor, they make up in clarity.

These notes are very much a work in progress. Students are encouraged to e-mail the professor if they find errors.

Acknowledgments

The author of these notes wishes to express his sincere thanks to Weston Barger for checking and writing homework solutions as well making corrections and improvements to the text.

Donations

If you find these notes useful and would like to make a donation to help me develop them further you can donate Bitcoin, Ethereum or Ethereum-based ERC20 tokens to the addresses below.

BTC address: 3QVEVC6QGJ8EmKsnpfXadkshwkDKzCrR4w

ETH address: 0x089d4d033F3E55B80Ac40D5da46812EFeFC8824b

Chapter 1

Review of Probability

In this chapter, we will *very briefly* review some concepts from probability theory. This material is *essential* for understanding the subsequent chapters.

1.1 Sample spaces, events, and probability measures

We begin this section with some important definitions.

Definition 1.1. The set of all possible outcomes of a random experiment is called the *sample space* and is denoted by Ω.

An *outcome* of a random experiment is an *element* of Ω and is denoted by ω.

Definition 1.2. An *event* is any subset of the sample space Ω. We denote events by capital roman letters $\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots$.

Finally, we will denote by $\mathcal{F}$ all of the subsets of a sample space Ω. [1] Note that a single element ω of Ω is also a subset of Ω, as are the empty set $\emptyset$ and Ω itself. Thus, we have $\omega, \emptyset, \Omega \in \mathcal{F}$.

Example 1.3 (Toss a coin). When one tosses a coin, there are two possible outcomes: heads (denoted by H) and tails (denoted by T). Thus, our sample space is $\Omega = \{\mathrm{H}, \mathrm{T}\}$. One possible element or outcome of Ω is to "toss a heads" $\omega = \mathrm{H}$. Two possible events are "toss a heads" $\mathrm{A} = \{\mathrm{H}\}$ and "toss a heads or a tails" $\mathrm{B} = \{\mathrm{H}, \mathrm{T}\}$.

Example 1.4 (Roll a die). A standard die has six sides, labeled one through six. Thus, our sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. One possible element or outcome of Ω is "roll a two" $\omega = 2$. Two possible events are "roll an odd number" $\mathrm{A} = \{1, 3, 5\}$ and "roll a one or two" $\mathrm{B} = \{1, 2\}$.

[1] Rigorously, $\mathcal{F}$ should be a σ-*algebra* of Ω. But, we need not get into such detail in this text.

So far, we have not yet talked about probabilities – only outcomes of a random experiment (elements $\omega \in \Omega$) and events (subsets $A \in \mathcal{F}$). We now introduce the concept of a probability measure.

Definition 1.5. A *probability measure* $\mathbb{P}$ assigns probabilities to events $\mathbb{P} : \mathcal{F} \to [0, 1]$ and satisfies

1. $\mathbb{P}(\Omega) = 1$, and
2. If (A_i) is a countable sequence of disjoint subsets of Ω, (i.e., $A_i \cap A_j = \emptyset$ if $i \neq j$), then

$$\mathbb{P}\Big(\bigcup_i A_i\Big) = \sum_i \mathbb{P}(A_i).$$

The first property in Definition 1.5 is easy to understand. When we perform a random experiment, the probability that *something* happens is one. Thus, in order for a probability measure $\mathbb{P}$ to make any sense at all, we must have $\mathbb{P}(\Omega) = 1$. The second property in Definition 1.5, known as *countable additivity*, ensures that we do not double-count the probability of events. For example, if we toss a coin, the probability of getting a heads or tails should equal the probability of getting a heads plus the probability of getting a tails

$$\mathbb{P}(H \cup T) = \mathbb{P}(H) + \mathbb{P}(T).$$

Note that, together, the two properties in Definition 1.5 imply that $\mathbb{P}(\emptyset) = 0$ as

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) = 1 + \mathbb{P}(\emptyset), \qquad \Rightarrow \qquad \mathbb{P}(\emptyset) = 0,$$

where we have used the fact that Ω and $\emptyset$ are disjoint. As we shall see with the next few examples, a probability measure $\mathbb{P}$ need not assign probabilities to events in a way that is consistent with our real-world experience.

Example 1.6 (Toss a coin). When we toss a coin, our experience tells us that the probability of getting a heads is $1/2$ and the probability of getting tails is $1/2$. Thus, if we want a probability measure $\mathbb{P}$ to reflect our experience, we should have $\mathbb{P}(H) = \mathbb{P}(T) = 1/2$. We can think of $\mathbb{P}$ as the "real world" or "statistical" probability measure. However, we could consider another probability measure $\widetilde{\mathbb{P}}$ that assigns different probabilities to heads and tails. For example, we could have $\widetilde{\mathbb{P}}(H) = 3/4$ and $\widetilde{\mathbb{P}}(T) = 1/4$. The measure $\widetilde{\mathbb{P}}$ does not capture what goes on in the real world. But, it is nonetheless a valid probability measure for the random experiment of tossing a coin. You should check for yourself that $\widetilde{\mathbb{P}}$ satisfies the conditions of Definition 1.5.

Example 1.7 (Roll a die). When we roll a die, our experience tells us that the probability of getting any number $\omega \in \{1, 2, 3, 4, 5, 6\}$ is $1/6$. Thus, if we want a probability measure $\mathbb{P}$ to reflect our experience, we should have $\mathbb{P}(\omega) = 1/6$ for all $\omega \in \{1,2,3,4,5,6\}$. As in the coin toss example, we can think of $\mathbb{P}$ as the "real world" or "statistical" probability measure. However, we could consider another probability measure $\widetilde{\mathbb{P}}$ that assigns different probabilities to each outcome. For example, we could have $\widetilde{\mathbb{P}}(1) = 1/2$ and $\widetilde{\mathbb{P}}(\omega) = 1/10$ for $\omega \in \{2,3,4,5,6\}$. The measure $\widetilde{\mathbb{P}}$ does not capture what goes on in the real world. But, it is nonetheless a valid probability measure for the random experiment of rolling a die. You should check for yourself that $\widetilde{\mathbb{P}}$ satisfies the conditions of Definition 1.5.

It may not be clear at this time why one would ever consider a probability measure $\widetilde{\mathbb{P}}$ that does not reflect our real-world experience. However, when we discuss option pricing, it will become apparent how such a measure can be useful.

Definition 1.8. Two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are said to be *equivalent*, denoted $\mathbb{P} \sim \widetilde{\mathbb{P}}$, if they agree on events that occur with probability one. That is, if

$$\mathbb{P}(A) = 1 \qquad \Leftrightarrow \qquad \widetilde{\mathbb{P}}(A) = 1.$$

Note that if $\mathbb{P} \sim \widetilde{\mathbb{P}}$ we also have

$$\mathbb{P}(A) = 0 \qquad \Leftrightarrow \qquad \widetilde{\mathbb{P}}(A) = 0.$$

In other words, equivalent measures agree on events that occur with probability one and on events that occur with probability zero.

Example 1.9. Let us return to the coin toss example. Consider two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. Suppose that $\mathbb{P}(H) = 1/2$ and $\widetilde{\mathbb{P}}(H) = 3/4$. Then $\mathbb{P} \sim \widetilde{\mathbb{P}}$ because

$$\begin{aligned}
\mathbb{P}(H \cup T) &= \mathbb{P}(H) + \mathbb{P}(T) = 1/2 + 1/2 = 1, \\
\widetilde{\mathbb{P}}(H \cup T) &= \widetilde{\mathbb{P}}(H) + \widetilde{\mathbb{P}}(T) = 3/4 + 1/4 = 1.
\end{aligned}$$

Now, consider a third probability measure $\widehat{\mathbb{P}}$ that satisfies $\widehat{\mathbb{P}}(H) = 1$. Then $\widehat{\mathbb{P}}$ is *not* equivalent to $\mathbb{P}$ (or to $\widetilde{\mathbb{P}}$) because $\mathbb{P}(H) = 1/2 \neq 1$.

1.2 Random variables

Intuitively, we think of a random variables as a quantity whose value is uncertain. For example, the value of a stock 10 days from now is uncertain and, as such, we think of it as being random. We give a precise definition of a random variable below.

Definition 1.10. A random variable X is a function that maps every outcome ω of a random experiment to a number in $\mathbb{R}$

$$X : \Omega \to \mathbb{R}.$$

The set S of values that X can take is called the *state space of* X.

Example 1.11 (Toss a coin). If we toss a coin, we have $\Omega = \{H, T\}$. Two possible random variables on this sample space are

$$X(\omega) = \begin{cases} 1, & \text{if } \omega = H, \\ 0, & \text{if } \omega = T, \end{cases} \qquad Y(\omega) = \begin{cases} 10, & \text{if } \omega = H, \\ 20, & \text{if } \omega = T. \end{cases}$$

Thus, if we toss a heads $\omega = H$, we have $X = 1$ and $Y = 10$. The state space of X is $S = \{0, 1\}$. The state space of Y is $S = \{10, 20\}$.

Example 1.12 (Roll a die). If we roll a die, we have $\Omega = \{1, 2, 3, 4, 5, 6\}$. Two possible random variables on this sample space are

$$X(\omega) = \omega, \qquad Y(\omega) = \omega^2.$$

Thus, if we roll a three $\omega = 3$, we have $X = 3$ and $Y = 3^2 = 9$. The state space of X is $S = \{1, 2, 3, 4, 5, 6\}$. The state space of Y is $S = \{1, 4, 9, 16, 25, 36\}$.

Random variables can be described by their *cumulative distribution function.*

Definition 1.13. The *cumulative distribution function* (CDF) of a random variable X is a function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) := \mathbb{P}(X \leq x), \tag{1.2.1}$$

where $x \in \mathbb{R}$.

Here, we have used the short-hand notation

$$\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}.$$

Now, we make an important note. A random variable X is defined *independently* of a probability measure. However, we see from (1.2.1) that the CDF F_X is defined with respect to a probability measure $\mathbb{P}$. Thus a random variable X can have a CDF F_X under $\mathbb{P}$ and have an entirely different CDF $\widetilde{F}_X$ under $\widetilde{\mathbb{P}}$. We will (hopefully) make this concept clear through some examples in the coming sections.

Sometimes, the value of one random variable X may affect the value of another random variable Y. For example, suppose you roll two 6-sided dice. Let X be the result of the first roll, let Y be the result of the second and let $Z = X + Y$. If I tell you that $Z = 2$, then you know that both $X = 1$ and $Y = 1$. In cases such as these, in order to describe the relationship between two random variables, we will need to consider *joint cumulative distribution function*.

Definition 1.14. The *joint cumulative distribution function* (joint CDF) of two random variables X and Y is a function $F_{X,Y} : \mathbb{R}^2 \to [0, 1]$ defined by

$$F_{X,Y}(x, y) := \mathbb{P}(X \leq x, Y \leq y), \tag{1.2.2}$$

where $x, y \in \mathbb{R}$.

Here, we have used the short-hand notation

$$\{X \leq x, Y \leq y\} = \{\omega \in \Omega : X(\omega) \leq x, Y(\omega) \leq y\}.$$

As in the univariate case, the joint CDF $F_{X,Y}$ is defined with respect to a particular probability measure (in the case of equation (1.2.2), the probability measure is $\mathbb{P}$).

Definition 1.2.2 can easily be extended to any finite number of random variables. For example, the joint CDF of three random variables (X, Y, Z) would be defined as follows

$$F_{X,Y,Z}(x, y, z) := \mathbb{P}(X \leq x, Y \leq y, Z \leq z),$$

where, once again, $F_{X,Y,Z}$ is defined relative to the probability measure $\mathbb{P}$.

Many (but not all) random variables fall in to one of two categories: *discrete* or *continuous*. We will discuss discrete random variables first. Once we understand discrete random variables, we will be able to understand continuous random variables by making some obvious modifications.

1.2.1 Discrete random variables

Definition 1.15. A *discrete random variable* X is any random variable whose state space consists of at most countably many values $S := \{x_1, x_2, \ldots\}$.

Example 1.16. Let us return to the example of rolling a single die. Our sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. Consider a random variable Z defined as follows

$$Z(\omega) = \omega/10.$$

The state space of Z is $S = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$. As the state space S contains only six values, it follows that Z is a discrete random variable.

A discrete random variable can be described by its *probability mass function*.

Definition 1.17. The *probability mass function* (PMF) of a discrete random variable X is a function $f_X : S \to [0,1]$ is defined by

$$f_X(x_i) := \mathbb{P}(X = x_i).$$

Here, as in the previous section, we have introduced short-hand notation

$$\{X = x_i\} = \{\omega \in \Omega : X(\omega) = x_i\}.$$

Just as the CDF F_X of a random variable X is defined with respect to a particular probability measure $\mathbb{P}$, the PMF f_X is also defined with respect to a particular probability measure $\mathbb{P}$. If we wish to speak about the PMF of a discrete random variable X under a probability measure $\widetilde{\mathbb{P}}$ we will write $\widetilde{f}_X$.

Example 1.18 (Roll a die). Let us return to the example of rolling a single die. Our sample space is $\Omega = \{1,2,3,4,5,6\}$. We will consider the following random variable on this space

$$X(\omega) = \begin{cases} 1, & \text{if } \omega \text{ is odd}, \\ 0, & \text{if } \omega \text{ is even}. \end{cases}$$

Now, consider two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ defined as follows

$$\mathbb{P}(\omega) = \frac{1}{6}, \qquad \widetilde{\mathbb{P}}(\omega) = \frac{\omega}{21}, \qquad \omega \in \{1,2,3,4,5,6\}.$$

First, we note that both $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ satisfy Definition 1.5, as they must. Let f_X and $\widetilde{f}_X$ denote the probability mass functions of X under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$, respectively. As the sate space of X is $S = \{0,1\}$ we need only to compute $f_X(x_i)$ and $\widetilde{f}_X(x_i)$ for $x_i = 0$ and $x_i = 1$. We have

$$\begin{aligned} f_X(1) &= \mathbb{P}(X = 1) = \mathbb{P}(\omega \text{ is odd}) = 1/6 + 1/6 + 1/6 = 1/2, \\ \widetilde{f}_X(1) &= \widetilde{\mathbb{P}}(X = 1) = \widetilde{\mathbb{P}}(\omega \text{ is odd}) = 1/21 + 3/21 + 5/21 = 9/21. \end{aligned}$$

As the complement of "ω is odd" is "ω is even," we have that $f_X(0) = 1 - f_X(1) = 1/2$ and $\widetilde{f}_X(0) = 1 - \widetilde{f}_X(1) = 12/21$.

Using the PMF f_X we can compute $\mathbb{P}(X \in A)$ where $A \subseteq \mathbb{R}$ and the *expectation* $\mathbb{E}g(X)$ as follows

$$\mathbb{P}(X \in A) = \sum_{\{i : x_i \in A\}} f_X(x_i), \qquad \mathbb{E}g(X) = \sum_i f_X(x_i) g(x_i). \tag{1.2.3}$$

Note that, as the expectation $\mathbb{E}g(X)$ in (1.2.3) depends on PMF f_X, which depends on a probability measure $\mathbb{P}$, it follows that the expectation $\mathbb{E}g(X)$ also depends on the probability measure $\mathbb{P}$. Here and throughout the text we will use $\mathbb{E}$ to indicate an expectation take with respect to $\mathbb{P}$. If we wish to take an expectation under a different probability measure such as $\widetilde{\mathbb{P}}$ or $\widehat{\mathbb{P}}$ we will use the notation $\widetilde{\mathbb{E}}$ or $\widehat{\mathbb{E}}$. For example

$$\widetilde{\mathbb{P}}(\mathrm{X} \in \mathrm{A}) = \sum_{\{i: x_i \in \mathrm{A}\}} \widetilde{f}_{\mathrm{X}}(x_i), \qquad \widetilde{\mathbb{E}} g(\mathrm{X}) = \sum_i \widetilde{f}_{\mathrm{X}}(x_i) g(x_i),$$

where $\widetilde{f}_{\mathrm{X}}(x_i) = \widetilde{\mathbb{P}}(\mathrm{X} = x_i)$.

Example 1.19. A *binomial random variable* $\mathrm{Bin}(n, p)$ is a discrete random variable X whose probability mass function f_{X} is given by

$$f_{\mathrm{X}}(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}, \qquad q = 1 - p, \qquad k \in \{0, 1, \ldots, n\},$$

where $n \in \mathbb{N}$ and $p \in (0, 1)$. Let us compute $\mathbb{P}(\mathrm{X} \in \{0, n\})$ and $\mathbb{E}\mathrm{X}$. Using (1.2.3) we have

$$\mathbb{P}(\mathrm{X} \in \{0, n\}) = f_{\mathrm{X}}(0) + f_{\mathrm{X}}(n) = p^n + q^n, \qquad \mathbb{E}\mathrm{X} = \sum_{k=1}^{n} f_{\mathrm{X}}(k) k = np.$$

In fact, showing that $\mathbb{E}\mathrm{X} = np$ takes a bit of work. If you cannot show this analytically for general (n, p), check a few simple cases (e.g., $n = 2$ and $p = 1/3$).

Example 1.20. Let us return to Example 1.18. The state space of X was $\mathrm{S} = \{0, 1\}$ and we computed

$$f_{\mathrm{X}}(x_i) = \begin{cases} 1/2 & \text{if } x_i = 0, \\ 1/2 & \text{if } x_i = 1, \end{cases} \qquad \widetilde{f}_{\mathrm{X}}(x_i) = \begin{cases} 12/21 & \text{if } x_i = 0, \\ 9/21 & \text{if } x_i = 1, \end{cases}$$

Let us compute $\mathbb{E}(\mathrm{X} + 3)^2$ and $\widetilde{\mathbb{E}}(\mathrm{X} + 3)^2$. Using (1.2.3), we have

$$\begin{aligned}
\mathbb{E}(\mathrm{X} + 3)^2 &= \sum_{x_i = \{0,1\}} f_{\mathrm{X}}(x_i)(x_i + 3)^2 = f_{\mathrm{X}}(0)(0 + 3)^2 + f_{\mathrm{X}}(1)(1 + 3)^2 \\
&= 1/2 \cdot 9 + 1/2 \cdot 16 = 12.5, \\
\widetilde{\mathbb{E}}(\mathrm{X} + 3)^2 &= \sum_{x_i = \{0,1\}} \widetilde{f}_{\mathrm{X}}(x_i)(x_i + 3)^2 = \widetilde{f}_{\mathrm{X}}(0)(0 + 3)^2 + \widetilde{f}_{\mathrm{X}}(1)(1 + 3)^2 \\
&= 12/21 \cdot 9 + 9/21 \cdot 16 = 12.
\end{aligned}$$

When the value of one discrete random variable X affects the value of another discrete random variable Y, in order to describe how these two random variables interact, we need to consider the *joint probability mass function.*

Definition 1.21. Let X and Y be discrete random variables. The *joint probability mass function of* X *and* Y, written $f_{X,Y}$ is defined by

$$f_{X,Y}(x_i, y_j) := \mathbb{P}(X = x_i, Y = y_j).$$

Using the joint probability mass function $f_{X,Y}$ we can compute $\mathbb{P}((X,Y) \in A)$ where $A \subseteq \mathbb{R}^2$ and $\mathbb{E}g(X,Y)$ as follows

$$\mathbb{P}((X,Y) \in A) = \sum_{\{i,j:(x_i,y_j)\in A\}} f_{X,Y}(x_i, y_j), \qquad \mathbb{E}g(X,Y) = \sum_{i,j} f_{X,Y}(x_i, y_j) g(x_i, y_j).$$

We can obtain the probability mass function f_X from the joint probability mass function $f_{X,Y}$ as follows

$$f_X(x_i) = \sum_j f_{X,Y}(x_i, y_j).$$

The *conditional probability mass function of* X *given* Y, written $f_{X|Y}$ is defined by

$$f_{X|Y}(x_i, y_j) := \mathbb{P}(X = x_i | Y = y_j).$$

We can compute the conditional probability mass function $f_{X|Y}$ from the joint probability mass function $f_{X,Y}$ like so

$$f_{X|Y}(x_i, y_j) = \frac{f_{X,Y}(x_i, y_j)}{f_Y(y_j)}. \tag{1.2.4}$$

Note that (1.2.4) follows from *Bayes' Rule*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \qquad \mathbb{P}(B) > 0,$$

by taking $A = \{X = x_j\}$ and $B = \{Y = y_j\}$. We can compute the conditional probability $\mathbb{P}(X \in A | Y = y_j)$ where $A \subseteq \mathbb{R}$ and the conditional expectation $\mathbb{E}[g(X,Y)|Y = y_j]$ as follows

$$\mathbb{P}(X \in A | Y = y_j) = \sum_{\{i:x_i \in A\}} f_{X|Y}(x_i, y_j), \qquad \mathbb{E}[g(X,Y)|Y = y_j] = \sum_i f_{X|Y}(x_i, y_j) g(x_i, y_j). \tag{1.2.5}$$

Example 1.22. Suppose X and Y have a joint probability mass function given by

$$f_{X,Y}(k, n) = q(1-q)^{n-1} \binom{n}{k} p^k (1-p)^{n-k}, \qquad k \in \{0, 1, 2, \dots, n\}, \qquad n \in \{1, 2, 3, \dots\},$$

where $p, q \in (0, 1)$ and we have introduced the binomial coefficient

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

First, let us compute $f_Y(n)$, the probability mass function of Y. We have

$$\begin{aligned} f_Y(n) &= \sum_{k=0}^{n} f_{X,Y}(k,n) = \sum_{k=0}^{n} q(1-q)^{n-1}\binom{n}{k}p^k(1-p)^{n-k} \\ &= q(1-q)^{n-1}\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} = q(1-q)^{n-1}. \end{aligned}$$

Using the expression for f_Y we can obtain $f_{X|Y}$, the conditional probability mass function of X given $Y = n$. We compute

$$f_{X|Y}(k,n) = \frac{f_{X,Y}(k,n)}{f_Y(n)} = \frac{q(1-q)^{n-1}\binom{n}{k}p^k(1-p)^{n-k}}{q(1-q)^n} = \binom{n}{k}p^k(1-p)^{n-k}.$$

Finally, let us compute $\mathbb{E}[X|Y = n]$, the conditional expectation of X given $Y = n$. We have

$$\mathbb{E}[X|Y = n] = \sum_{k=0}^{n} k\, f_{X|Y}(k,n) = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} = np.$$

Again, if you cannot follow all of the algebra above, check that the above equations are true for some simple cases.

All of the above can be generalized to n discrete random variables.

1.2.2 Continuous random variables

Now that we understand discrete random variables, it will be quite easy to understand continuous random variables. The main difference between discrete and continuous random variables is simply that, where we saw sums $\sum$ in Section 1.2.1, these will become integrals $\int$ in Section 1.2.2.

Definition 1.23. A *continuous random variable* X is a random variable whose CDF F_X can be written as

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\mathrm{d}y, \tag{1.2.6}$$

for some non-negative function $f_X : \mathbb{R} \to \mathbb{R}_+$ call the *probability density function* (PDF).

It follows from (1.2.6) that

$$f_X(x) = \frac{\mathrm{d}F_X}{\mathrm{d}x}(x). \tag{1.2.7}$$

We can think of (1.2.7) as a definition of the PDF f_X. Alternatively, we can think of the PDF f_X as describing the probability that X is in some small interval $\mathrm{d}x$, i.e.,

$$f_X(x)\mathrm{d}x := \mathbb{P}(X \in \mathrm{d}x).$$

Note once again that the density f_X is defined with respect to a particular probability measure $\mathbb{P}$. If we want to express the density of X under a different probability measure $\widetilde{\mathbb{P}}$, we would have $\widetilde{f}_X(x) = d\widetilde{F}_X(x)/dx$ and $\widetilde{f}_X(x)dx = \widetilde{\mathbb{P}}(X \in dx)$.

Example 1.24. Supposed $\Omega = [0,1]$. Let us define a random variable X by $X(\omega) := \omega^2$. Consider two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$, defined as follows

$$\mathbb{P}(d\omega) = d\omega, \qquad \widetilde{\mathbb{P}}(d\omega) = 2\omega d\omega.$$

Note that both $\mathbb{P}(\Omega) = 1$ and $\widetilde{\mathbb{P}}(\Omega) = 1$, as the must, because

$$\begin{aligned}
\mathbb{P}(\Omega) &= \int_0^1 \mathbb{P}(d\omega) = \int_0^1 d\omega = \omega\Big|_0^1 = 1, \\
\widetilde{\mathbb{P}}(\Omega) &= \int_0^1 \widetilde{\mathbb{P}}(d\omega) = \int_0^1 2\omega d\omega = \omega^2\Big|_0^1 = 1.
\end{aligned}$$

Let us find the density of X under both $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. First, under $\mathbb{P}$ for any $x \in [0,1]$ we have

$$\begin{aligned}
F_X(x) &= \mathbb{P}(X(\omega) \le x) = \mathbb{P}(\omega^2 \le x) = \mathbb{P}(\omega \le \sqrt{x}) \\
&= \int_0^{\sqrt{x}} \mathbb{P}(d\omega) = \int_0^{\sqrt{x}} d\omega = \sqrt{x}, \\
f_X(x) &= \frac{dF_X}{dx}(x) = \frac{1}{2\sqrt{x}}.
\end{aligned}$$

Now, under $\widetilde{\mathbb{P}}$ for any $x \in [0,1]$ we have

$$\begin{aligned}
\widetilde{F}_X(x) &= \widetilde{\mathbb{P}}(X(\omega) \le x) = \widetilde{\mathbb{P}}(\omega^2 \le x) = \widetilde{\mathbb{P}}(\omega \le \sqrt{x}) \\
&= \int_0^{\sqrt{x}} \widetilde{\mathbb{P}}(d\omega) = \int_0^{\sqrt{x}} 2\omega d\omega = x, \\
\widetilde{f}_X(x) &= \frac{d\widetilde{F}_X}{dx}(x) = 1.
\end{aligned}$$

Using the probability density function f_X we can compute $\mathbb{P}(X \in A)$ where $A \subseteq \mathbb{R}$ and the *expectation* $\mathbb{E}g(X)$ as follows

$$\mathbb{P}(X \in A) = \int_A f_X(x)dx, \qquad \mathbb{E}g(X) = \int f_X(x)g(x)dx. \tag{1.2.8}$$

Compare (1.2.8) to (1.2.3) and note the similarity. Observe once again that, as $\mathbb{E}g(X)$ depends on f_X, which depends on $\mathbb{P}$, we have that $\mathbb{E}g(X)$ also depends on the probability measure $\mathbb{P}$. If we wish to take an expectation under a different probability measure such as $\widetilde{\mathbb{P}}$, we will use the notation $\widetilde{\mathbb{E}}$. For example

$$\widetilde{\mathbb{P}}(X \in A) = \int_A \widetilde{f}_X(x)dx, \qquad \widetilde{\mathbb{E}}g(X) = \int \widetilde{f}_X(x)g(x)dx.$$

where $\widetilde{f}_X(x)dx \in \widetilde{\mathbb{P}}(X = dx)$.

Example 1.25. An *exponential random variable* $\mathrm{Exp}(\lambda)$ is a continuous random variable X whose probability density function f_X is given by

$$f_X(x) = \mathbb{1}_{\mathbb{R}_+}(x)\lambda e^{-\lambda x},$$

where $\lambda > 0$. Here, we have introduced the *indicator function* $\mathbb{1}_A$, which is defined as follows

$$\mathbb{1}_A(x) := \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

Note that, due to the presence of the indicator function, we have that $f_X(x) = 0$ for all $x < 0$. Let us compute $\mathbb{P}(X \in (a,b))$ with $0 \leq a \leq b$ and $\mathbb{E}X^2$. Using (1.2.8) we have

$$\mathbb{P}(X \in (a,b)) = \int_a^b \lambda e^{-\lambda x} dx = -e^{-\lambda b} + e^{-\lambda a},$$

$$\mathbb{E}X^2 = \int_0^\infty \lambda e^{-\lambda x} x^2 dx = \frac{-e^{-\lambda x}}{\lambda^2}\left(\lambda^2 x^2 + 2\lambda x + 2\right)\Big|_0^\infty = \frac{2}{\lambda^2}.$$

As a check of your understanding, see if you can compute $F_X(x)$ and $\mathbb{E}X^3$ (or at least write the integrals that would need to be computed).

When the value of one continuous random variable X affects the value of another continuous random variable Y, in order to describe this interaction, we will need to consider the *joint probability density function*.

Definition 1.26. Two random variables X and Y are *jointly continuous* if their joint CDF $F_{X,Y}$ can be written as

$$F_{X,Y}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u,v) du dv, \tag{1.2.9}$$

for some non-negative function $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ called the *joint probability density function* (PDF).

It follows from (1.2.9) that

$$f_{X,Y}(x,y) = \frac{\partial^2 F_X}{\partial x \partial y}(x,y). \tag{1.2.10}$$

We can think of (1.2.10) as a definition of the joint PDF $f_{X,Y}$. Alternatively, we can think of the joint PDF $f_{X,Y}$ as describing the probability that X is in some small interval dx and Y is in some interval dy, i.e.,

$$f_{X,Y}(x,y) dx dy := \mathbb{P}(X \in dx, Y \in dy).$$

Using the joint probability density function $f_{X,Y}$ we can compute $\mathbb{P}((X,Y) \in A)$ where $A \subseteq \mathbb{R}^2$ and $\mathbb{E}g(X,Y)$ as follows

$$\mathbb{P}((X,Y) \in A) = \int_A f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y, \qquad \mathbb{E}g(X,Y) = \int f_{X,Y}(x,y)g(x,y)\mathrm{d}x\mathrm{d}y.$$

We can obtain the probability density function f_X from the joint probability density function $f_{X,Y}$ as follows

$$f_X(x) = \int f_{X,Y}(x,y)\mathrm{d}y.$$

The *conditional probability density function* $f_{X|Y}$ is defined in terms of infinitesimal probabilities as follows

$$f_{X|Y}(x,y)\mathrm{d}x := \mathbb{P}(X \in \mathrm{d}x | Y = y).$$

From the joint probability density function $f_{X,Y}$ we can compute the conditional probability mass function $f_{X|Y}$ as follows

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

To compute the conditional probability $\mathbb{P}(X \in A | Y = y)$ where $A \subseteq \mathbb{R}$ and the conditional expectation $\mathbb{E}[g(X,Y)|Y = y]$ we use

$$\mathbb{P}(X \in A|Y = y) = \int_A f_{X|Y}(x,y)\mathrm{d}x, \qquad \mathbb{E}[g(X,Y)|Y = y] = \int f_{X|Y}(x,y)g(x,y)\mathrm{d}x. \quad (1.2.11)$$

Example 1.27. Suppose X and Y have a joint density given by

$$f_{X,Y}(x,y) = \mathbb{1}_{\mathbb{R}^2_+}(x,y)\frac{1}{y}\exp\Big(-y - \frac{x}{y}\Big),$$

where the indicator function $\mathbb{1}_{\mathbb{R}^2_+}$ indicates that $f_{X,Y}(x,y) = 0$ if either $x < 0$ or $y < 0$. Let us compute the probability density f_Y. We have

$$\begin{aligned} f_Y(y) &= \int_0^\infty f_{X,Y}(x,y)\mathrm{d}x = \int_0^\infty \frac{1}{y}\exp\Big(-y - \frac{x}{y}\Big)\mathrm{d}x \\ &= -\exp\Big(-y - \frac{x}{y}\Big)\Big|_{x=0}^{x=\infty} = \mathrm{e}^{-y}. \end{aligned}$$

Next, let us compute $f_{X|Y}(x,y)$, the conditional density of X given $Y = y$. We have

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{1}{y}\exp\Big(-y - \frac{x}{y}\Big)}{\mathrm{e}^{-y}} = \frac{1}{y}\exp\Big(-\frac{x}{y}\Big).$$

We can compute the conditional probability $\mathbb{P}(x_1 < X < x_2 | Y = y)$ where $0 < x_1 < x_2 < \infty$ as follows

$$\begin{aligned}\mathbb{P}(x_1 < X < x_2 | Y = y) &= \int_{x_1}^{x_2} f_{X|Y}(x, y) \mathrm{d}x = \int_{x_1}^{x_2} \frac{1}{y} \exp\Big(-\frac{x}{y}\Big) \mathrm{d}x \\ &= -\exp\Big(-\frac{x}{y}\Big)\Big|_{x=x_1}^{x=x_2} = \exp\Big(-\frac{x_1}{y}\Big) - \exp\Big(-\frac{x_2}{y}\Big).\end{aligned}$$

Lastly, we compute $\mathbb{E}(X|Y = y)$, the conditional expectation of X given $Y = y$. We have

$$\begin{aligned}\mathbb{E}[X|Y = y] &= \int_0^\infty x\, f_{X|Y}(x, y) \mathrm{d}x = \int_0^\infty \frac{x}{y} \exp\Big(-\frac{x}{y}\Big) \mathrm{d}x \\ &= -(x + y) \exp\Big(-\frac{x}{y}\Big)\Big|_{x=0}^{x=\infty} = y.\end{aligned}$$

All of the above can be generalized to n continuous random variables.

1.3 Independence

Intuitively, two events A and B are independent if the outcome of A has no impact on the outcome of B, and vice versa. We can define independence mathematically as follows:

Definition 1.28. Two sets A and B are *independent* if

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

and we write $A \perp\!\!\!\perp B$.

Independent random variables are defined in a similar fashion.

Definition 1.29. Two random variables X and Y are said to be *independent* if

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

for all sets $A, B \subseteq \mathbb{R}$, and we write $X \perp\!\!\!\perp Y$.

It follows from Definition 1.29 that

$$f_{X,Y}(x, y) = f_X(x) f_Y(y), \qquad \text{and} \qquad \mathbb{E}g(X)h(Y) = \mathbb{E}g(X) \cdot \mathbb{E}h(Y), \qquad \text{if} \qquad X \perp\!\!\!\perp Y,$$

where $f_{X,Y}$, f_X and f_Y may be probability mass functions (if X and Y are discrete random variables) or probability density functions (if X and Y are continuous random variables).

1.4 Conditional Expectation revisited

At times, we may wish to compute the conditional expectation of $g(\mathrm{X}, \mathrm{Y})$ given Y without specifying the value of Y. For this, we define

$$\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}] := \psi(\mathrm{Y}), \qquad \text{with} \qquad \psi(y) := \mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}=y], \tag{1.4.1}$$

where expressions for $\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}=y]$ is given in (1.2.5) for discrete random variables and (1.2.11) for continuous random variables, respectively. Note that $\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}]$ is a random variable $\psi(\mathrm{Y})$ whereas $\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}=y]$ is simply a number $\psi(y)$. Conditional expectation, as we have defined it in (1.4.1) has the following useful properties

$$\text{Tower Property :} \qquad \mathbb{E}g(\mathrm{X},\mathrm{Y}) = \mathbb{E}\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}], \tag{1.4.2}$$

$$\text{Taking out what is known :} \qquad \mathbb{E}[g(\mathrm{X})h(\mathrm{Y})|\mathrm{Y}] = h(\mathrm{Y})\mathbb{E}[g(\mathrm{X})|\mathrm{Y}], \tag{1.4.3}$$

$$\text{If } \mathrm{X} \perp\!\!\!\perp \mathrm{Y} \text{ then :} \qquad \mathbb{E}[g(\mathrm{X})|\mathrm{Y}] = \mathbb{E}g(\mathrm{X}). \tag{1.4.4}$$

That *Tower property* (1.4.2) can be further generalized as follows:

$$\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{Y}] = \mathbb{E}[\mathbb{E}[g(\mathrm{X},\mathrm{Y})|\mathrm{X},\mathrm{Y}]|\mathrm{Y}].$$

The basic idea is that you can always condition on additional variables in the inner expectation.

1.5 Stochastic Processes

Throughout this course, we will encounter various processes that evolve randomly in time. Such processes are known as *Stochastic processes*. Let us define exactly what a stochastic process is.

Definition 1.30. A *stochastic process* is a collection of random variables $\mathrm{X} = (\mathrm{X}_t)_{t\in\mathbb{T}}$ indexed by some set $\mathbb{T}$. If $\mathbb{T}$ is countable (e.g., $\mathbb{T} = \mathbb{N}_0 := \{0, 1, 2, \ldots\}$), then the process X is a *discrete time stochastic process*. If $\mathbb{T}$ is uncountable (e.g., $\mathbb{T} = \mathbb{R}_+ := [0, \infty)$), then the process X is a *continuous time stochastic process*. The *state space* S of a stochastic process X is the union of the state spaces of $(\mathrm{X}_t)_{t\in\mathbb{T}}$.

Note that X is a *process* whereas X_t is the *value* of the process X at time t. We can think of a stochastic process $\mathrm{X} : \mathbb{T} \times \Omega \to \mathbb{R}$ in (at least) two ways. First, for any $t \in \mathbb{T}$ we have that $\mathrm{X}_t : \Omega \to \mathbb{R}$ is random variable. Second, for any $\omega \in \Omega$, we have that $\mathrm{X}(\omega) : \mathbb{T} \to \mathbb{R}$ is a function of time. Both interpretations can be useful. Let us take a look at a very simple discrete-time stochastic process.

Example 1.31 (Random walk). Consider a random experiment in which we toss a coin 3 times. The sample space Ω of our experiment is

$$\Omega := \{(\mathrm{HHH}), (\mathrm{HHT}), (\mathrm{HTH}), (\mathrm{THH}), (\mathrm{TTH}), (\mathrm{THT}), (\mathrm{HTT}), (\mathrm{TTT})\},$$

Note that the order of the coin tosses matters here; the sequence (THT) is *not* equivalent to (TTH). A single outcome of Ω will be denoted as $\omega = \omega_1\omega_2\omega_3$, where ω_i is the result of the ith coin toss. For example, one possible element is $\omega = \mathrm{HHT}$. We will assume that each of the coin tosses are independent and that $\mathbb{P}(\omega_i = \mathrm{H}) = p$ and $\mathbb{P}(\omega_i = \mathrm{T}) = 1 - p =: q$. For every $i \in \{1, 2, 3\}$ let us define a random variable B_i as follows

$$\mathrm{B}_i(\omega) = \begin{cases} +1 & \text{if } \omega_i = \mathrm{H}, \\ -1 & \text{if } \omega_i = \mathrm{T}. \end{cases}$$

Now, let us define a stochastic process $\mathrm{X} = (\mathrm{X}_i)_{i\in\{0,1,2,3\}}$ as follows

$$\mathrm{X}_0 = 0, \qquad \mathrm{X}_n(\omega) = \sum_{i=1}^{n} \mathrm{B}_i(\omega), \qquad n \in \{1, 2, 3\}.$$

Clearly, X is a discrete time stochastic process with state space $\mathrm{S} = \{-3, -2, -1, 0, 1, 2, 3\}$. Suppose $\omega = (\mathrm{HTT})$. Then we have $\mathrm{B}_1 = 1$, $\mathrm{B}_2 = -1$ and $\mathrm{B}_3 = -1$. It follows that $\mathrm{X}_1 = 1$, $\mathrm{X}_2 = 0$ and $\mathrm{X}_3 = -1$. Suppose we fix the time $n = 2$. Then we have

$$\begin{aligned}
\mathbb{P}(\mathrm{X}_2 = 2) &= \mathbb{P}(\omega_1 = \mathrm{H} \cap \omega_2 = \mathrm{H}) = p^2, \\
\mathbb{P}(\mathrm{X}_2 = 0) &= \mathbb{P}(\{\omega_1 = \mathrm{H} \cap \omega_2 = \mathrm{T}\} \cup \{\omega_1 = \mathrm{T} \cap \omega_2 = \mathrm{H}\}) = 2qp, \\
\mathbb{P}(\mathrm{X}_2 = -2) &= \mathbb{P}(\omega_1 = \mathrm{T} \cap \omega_2 = \mathrm{T}) = q^2.
\end{aligned}$$

If you are able to understand this simple example very well, you wil be easily able to understand more complicated stochastic processes as we encounter them.

There are a number of properties that a stochastic process X may exhibit. Two properties that are very important are the *Markov property* and the *Martingale property*, which we now define.

Definition 1.32 (Markov). Let $\mathrm{X} = (\mathrm{X}_t)_{t\in\mathbb{T}}$ be a stochastic process. We say that X *is a Markov process* (or X *has the Markov property*) if, for any function g, we have

$$\mathbb{E}[g(\mathrm{X}_\mathrm{T})|\mathcal{F}_t^\mathrm{X}] = \mathbb{E}[g(\mathrm{X}_\mathrm{T})|\mathrm{X}_t], \qquad \text{for all} \qquad 0 \leq t \leq \mathrm{T} < \infty, \tag{1.5.1}$$

where $\mathcal{F}_t^\mathrm{X} := \{\mathrm{X}_s, 0 \leq s \leq t\}$ is all of the information generated by observing the path of X over the time interval $[0, t]$. We call $\mathcal{F}^\mathrm{X} := (\mathcal{F}_t^\mathrm{X})_{t\in\mathbb{T}}$ the *history of* X or *filtration generated by* X. [2]

[2] Strictly speaking, $\mathcal{F}_t^\mathrm{X}$ is the sigma algebra generated by observing the path of X up to time t. But, we need not get into this level of mathematical rigor for this course.

Suppose X is a discrete-time process. In this case the filtration generated by X is simply $\mathcal{F}_n^X = \{X_0, X_1, \ldots, X_n\}$ and the Markov property (1.5.1) becomes

$$\mathbb{E}[g(X_N)|X_0, X_1, \ldots, X_n] = \mathbb{E}[g(X_N)|X_n], \qquad \text{for all} \qquad 0 \leq n \leq N < \infty. \tag{1.5.2}$$

Equation (1.5.2) makes it clear that a Markov process is a process X for which, given the present X_n the future $g(X_N)$ is independent of the past $\{X_0, X_1, \ldots X_{n-1}\}$. Note that *this does not mean the future is independent of the past*! The past obviously affects the present state of a process. But, for a Markov process, only the present state – not how we arrive at the present state – affects the future. Markov processes are useful because they simplify the computation of conditional expectations greatly.

An alternative way to state the Markov property is that, for any $0 \leq t \leq T < \infty$ and any function g, there exists a function h (which depends on t, T and g) such that

$$\mathbb{E}[g(X_T)|\mathcal{F}_t^X] = h(X_t).$$

Example 1.33 (Random Walk). Consider a random experiment in which we toss a coin N times. The sample space Ω of our experiment is all sequences of heads and tails that are of length N. Thus, a generic outcome $\omega = \omega_1\omega_2\ldots\omega_N$ where $\omega_i \in \{H, T\}$. For every $i \in \{1, 2, \ldots, N\}$ let us define a random variable B_i as follows

$$B_i(\omega) = \begin{cases} +1 & \text{if } \omega_i = H, \\ -1 & \text{if } \omega_i = T. \end{cases}$$

Now, let us define a stochastic process $X = (X_n)_{n\in\{0,1,2,\ldots,N\}}$ as follows

$$X_0(\omega) = 0, \qquad X_n(\omega) = \sum_{i=1}^{n} B_i(\omega), \qquad n \in \{1, 2, \ldots, N\}.$$

We can clearly see that X is a Markov process because, for any $0 \leq i \leq j \leq N$ we have

$$\begin{aligned}
\mathbb{E}[g(X_j)|\mathcal{F}_i^X] &= \mathbb{E}[g(X_j)|X_0, \ldots, X_i] \\
&= \mathbb{E}[g(X_j - X_i + X_i)|X_0, \ldots, X_i] \\
&= \mathbb{E}[g(Y + X_i)|X_i], && Y := X_j - X_i = \sum_{k=i+1}^{j} B_k, \\
&= \sum_k f_Y(k) g(k + X_i) =: h(X_i), && f_Y(k) := \mathbb{P}(Y = k).
\end{aligned}$$

where we have used the fact that Y depends only on $B_{i+1}, \ldots, B_j$, each of which are independent of $X_0, \ldots, X_i$.

Definition 1.34 (Martingale). Let $X = (X_t)_{t\in\mathbb{T}}$ and $Y = (Y_t)_{t\in\mathbb{T}}$ be stochastic processes. We say that X *is a martingale with respect to* $\mathcal{F}^Y$ (or X *has the martingale property with respect to* $\mathcal{F}^Y$) if

$$\mathbb{E}[X_T|\mathcal{F}_t^Y] = X_t, \qquad \text{for all} \qquad 0 \le t \le T < \infty.$$

A few notes are in order. First, a stochastic process X can be a martingale with respect to itself. In this case, we have

$$\mathbb{E}[X_T|\mathcal{F}_t^X] = X_t, \qquad \text{where} \qquad 0 \le t \le T < \infty.$$

Second, observe that the martingale property is defined with respect to a particular probability measure $\mathbb{P}$ (because the expectation $\mathbb{E}$ is taken with respect to $\mathbb{P}$). It is possible, for example, that X is a martingale with respect to $\mathcal{F}^Y$ under $\mathbb{P}$ but that X is not a martingale with respect to $\mathcal{F}_Y$ under $\widetilde{\mathbb{P}}$. That is, we may simultaneously have

$$\mathbb{E}[X_T|\mathcal{F}_t^Y] = X_t, \qquad \text{and} \qquad \widetilde{\mathbb{E}}[X_T|\mathcal{F}_t^Y] \neq X_t,$$

Let us look at a simple example.

Example 1.35 (Random Walk). Let us return to the Random walk defined in Example 1.33. Consider two probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ defined as follows

$$\mathbb{P}(\omega_i = H) = p = 1/2, \qquad \widetilde{\mathbb{P}}(\omega_i = H) = \widetilde{p} > 1/2, \qquad \omega_i \perp\!\!\!\perp \omega_j, \qquad i \neq j.$$

From the above specification, we have $\mathbb{P}(B_i = 1) = p = 1/2$ and $\widetilde{\mathbb{P}}(B_i = 1) = \widetilde{p} > 1/2$. The process X is a martingale with respect to $\mathcal{F}^X$ under $\mathbb{P}$ because, for any $0 \le i \le j \le N$ we have

$$\begin{aligned}\mathbb{E}[X_j|\mathcal{F}_i^X] &= \mathbb{E}[B_j + B_{j-1} + \ldots + B_{i+1} + X_i|\mathcal{F}_i^X]\\ &= \mathbb{E}[B_j + B_{j-1} + \ldots + B_{i+1}] + X_i = X_i,\end{aligned}$$

where we have used the fact that, for $k > i$, we have $\mathbb{E}[B_k|\mathcal{F}_i^X] = \mathbb{E}B_k = 0$ and $\mathbb{E}[X_i|\mathcal{F}_i^X] = X_i$. On the other hand, X is not a martingale under $\widetilde{\mathbb{P}}$ because, for any $0 \le i \le j \le N$ we have

$$\begin{aligned}\widetilde{\mathbb{E}}[X_j|\mathcal{F}_i^X] &= \widetilde{\mathbb{E}}[B_j + B_{j-1} + \ldots + B_{i+1} + X_i|\mathcal{F}_i^X]\\ &= \widetilde{\mathbb{E}}[B_j + B_{j-1} + \ldots + B_{i+1}] + X_i = (2p-1)(j-i) + X_i > X_i.\end{aligned}$$

where we have used the fact that, for $k > i$ we have $\widetilde{\mathbb{E}}[B_k|\mathcal{F}_i^X] = \widetilde{\mathbb{E}}B_k = 2p - 1$ and $\widetilde{\mathbb{E}}[X_i|\mathcal{F}_i^X] = X_i$. To check your understanding, try to answer the following question: is X a martingale with respect to $\mathcal{F}^B$ under $\mathbb{P}$ and/or $\widetilde{\mathbb{P}}$?

Please do not confuse the Markov property and the Martingale property; they are entirely separate concepts. It is possible for a process to be a Markov process and not be a martingale. It is possible for a process to be a Martingale but not be a Markov process. It is possible for a process to be both a Markov process and a Martingale. And it is possible for a process to be neither a Markov process nor a Martingale.

Remark 1.36. At times, we use the simplified notation $\mathbb{E}_n$ to indicate an expectation that is conditioned on the a history $\mathcal{F}_n$. That is

$$\mathbb{E}_n \mathrm{Z} := \mathbb{E}[\mathrm{Z}|\mathcal{F}_n]$$

for some random variable Z.

1.6 Exercises

Exercise 1.1. What is the late homework policy?

Exercise 1.2. Consider the random variable X defined in Example 1.18. Compute (a) $\mathbb{E}\mathrm{X}$, (b) $\mathbb{E}e^{\mathrm{X}}$, (c) $\widetilde{\mathbb{E}}\mathrm{X}$ and (d) $\widetilde{\mathbb{E}}e^{\mathrm{X}}$.

Exercise 1.3. Consider a random experiment in which you toss two distinguishable coins. The set of possible outcomes is

$$\Omega = \{(\mathrm{HH}), (\mathrm{HT}), (\mathrm{TH}), (\mathrm{TT})\}.$$

We will denote by $\omega = \omega_1\omega_2$ an element of Ω. For example, if $\omega = (\mathrm{HT})$ then $\omega_1 = \mathrm{H}$ and $\omega_2 = \mathrm{T}$. Now, let us consider two random variables S_1 and S_2, which are defined as follows

$$\mathrm{S}_1 = \begin{cases} u\mathrm{S}_0 & \text{if } \omega_1 = \mathrm{H}, \\ d\mathrm{S}_0 & \text{if } \omega_1 = \mathrm{T}, \end{cases} \qquad \mathrm{S}_2 = \begin{cases} u\mathrm{S}_1 & \text{if } \omega_2 = \mathrm{H}, \\ d\mathrm{S}_1 & \text{if } \omega_2 = \mathrm{T}, \end{cases}$$

where $0 < d < u < \infty$ and $\mathrm{S}_0 > 0$. Suppose that the two coin tosses are independent so that

$$\mathbb{P}(\omega_1 = \mathrm{H}, \omega_2 = \mathrm{H}) = \mathbb{P}(\omega_1 = \mathrm{H}) \cdot \mathbb{P}(\omega_2 = \mathrm{H}).$$

Suppose further that $\mathbb{P}(\omega_i = \mathrm{H}) = p$ for $i = 1, 2$ where $p \in (0, 1)$. For convenience, define $q := 1 - p$. (a) What is the joint probability mass function $f_{\mathrm{S}_1,\mathrm{S}_2}(s_1, s_2)$? (b) What is the conditional probability mass function $f_{\mathrm{S}_1|\mathrm{S}_2}(s_1, s_2)$? (c) What is the conditional probability mass function $f_{\mathrm{S}_2|\mathrm{S}_1}(s_2, s_1)$? (d) What is $\mathbb{E}(\mathrm{S}_2/\mathrm{S}_1)$? (e) What is $\mathbb{E}[(\mathrm{S}_2/\mathrm{S}_1)|\mathrm{S}_1]$?

Exercise 1.4. Expectation $\mathbb{E}$ is what known as a *linear operator*, meaning

$$\mathbb{E}[a\mathrm{X} + b\mathrm{Y}] = a\mathbb{E}\mathrm{X} + b\mathbb{E}\mathrm{Y}, \tag{1.6.1}$$

where $a, b \in \mathbb{R}$ are constants. Prove (1.6.1) when X and Y are (a) discrete random variables, and (b) continuous random variables.

Exercise 1.5. The *variance* of a random variable X, denoted $\mathbb{V}\mathrm{X}$ is defined as follows

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

The *covariance* of two random variables X and Y, denoted $\mathrm{Co}\mathbb{V}[\mathrm{X}, \mathrm{Y}]$, is defined as follows

$$\mathrm{Co}\mathbb{V}[\mathrm{X}, \mathrm{Y}] := \mathbb{E}[(\mathrm{X} - \mathbb{E}\mathrm{X})(\mathrm{Y} - \mathbb{E}\mathrm{Y})].$$

Using (1.6.1), show (a) that

$$\mathbb{V}\mathrm{X} = \mathbb{E}\mathrm{X}^2 - (\mathbb{E}\mathrm{X})^2,$$

and (b) that

$$\mathrm{Co}\mathbb{V}[\mathrm{X}, \mathrm{Y}] = \mathbb{E}\mathrm{X}\mathrm{Y} - \mathbb{E}\mathrm{X} \cdot \mathbb{E}\mathrm{Y}.$$

Lastly, (c) what is $\mathrm{Co}\mathbb{V}[\mathrm{X}, \mathrm{Y}]$ when $\mathrm{X} \perp\!\!\!\perp \mathrm{Y}$?

Exercise 1.6. Prove equations (1.4.2), (1.4.3) and (1.4.4) when X and Y are (a) discrete random variables, and (b) continuous random variables.

Exercise 1.7. Let X and Y be independent uniformly distributed continuous random variables on the on the intervals $[a, b]$ and $[c, d]$ respectively. Find $\mathbb{E}[1/\mathrm{X}]$, $\mathbb{E}[1/\mathrm{Y}]$ and $\mathbb{E}[\mathrm{X}/\mathrm{Y}]$. You may assume $0 < a < b < \infty$ and $0 < c < d < \infty$.

Chapter 2

What is a financial derivative?

Definition 2.1. A *derivative* is a financial instrument whose value depends on the value of more basic underlying variables.

Typical *underlyers* are traded assets such as stocks, bonds and currencies. But underlyers can also be non-traded quantities such as interest rates, exchange rates, average temperatures (yes, weather derivatives really exist!) or a financial indices.

The use of derivatives in finance is so wide-spread that everyone who works in the financial sector must know how they are used and how they are valued. To give you an idea of just how wide-spread the use of financial derivatives has become, consider that the Chicago Board Options Exchange (CBOE, `cboe.com`) trades options on 100+ stocks. The options written on certain stocks are even more liquidly traded than the stocks themselves.

2.1 Forward Contracts

The most basic financial derivative is a *forward contract*.

Definition 2.2. A *forward contract* is an agreement between two parties to buy/sell an asset at a certain future date T for a certain price K. The *long side* agrees to buy the asset at time T for price K. The *short side* agrees to sell the asset at time T for price K. The date T is referred to as the *maturity* and the agreed upon price K is referred to as the *delivery price* or *strike*. The delivery price K is fixed at inception so that the initial value of the forward contract is zero.

Forward contracts can be used to *hedge*, i.e., reduce exposure to the future movements of an underlyer.

Example 2.3. An airline wants to sell tickets for a flight from NYC to London in 3 months. In order to know if operating the flight will be profitable the airline needs to know how much gasoline will cost in 3 months. Obviously, we cannot say today what the price of gasoline will be in 3 months. So, the airline is exposed to price fluctuations in the price of gasoline. By taking the long side of a forward contract written on gasoline the airline fixes the price it will pay for gasoline in 3 months times. This allows the airline to figure out today if the flight from NYC to London will be profitable.

Example 2.4. Consider a copper mine in Chile. The mine knows how much it costs to excavate and refine a unit of copper, but it does not know what the price of copper will be in 3 months when the refined copper goes on the market. Thus, the copper mine is exposed to fluctuations in the price of copper. By entering the short side of a forward contract written on copper the mine can fix today the price it will receive for a unit of copper in 3 months. This will allow the mine to decide whether or not they will operate or not (i.e., the mine will operate if they can make a profit, otherwise they will not operate).

Example 2.5 (Hedging with forward contracts). Let $\mathrm{S} = (\mathrm{S}_t)_{t\geq 0}$ denote the value of one unit of copper and suppose it costs c dollars to excavate and refine a unit of copper. Let T denote the time it takes to excavate and refine one unit of copper. If the mine takes the short side of a forward contract on copper, with a maturity date T and delivery price K, then the profit/loss the mine incurs when it sells the copper at time T will be $\mathrm{K} - c$, which is a constant. On the other hand, if the mine does not enter into the short side of a forward contract, then the profit/loss incurred by the mine for selling a unit of copper at time T will be $\mathrm{S}_\mathrm{T} - c$, which is a random number. If $\mathrm{S}_\mathrm{T} < c$ the mine will lose money. If $\mathrm{S}_\mathrm{T} > c$ the mine will make money. Thus, by taking the short side of a forward contract, the copper mine is sacrificing possible future profit in order to reduce risk (i.e., uncertainty).

Note that you do *not* need to own the underlying asset $\mathrm{S} = (\mathrm{S}_t)_{0\leq t}$ before you enter the short side of a forward contract. If you do not own the asset when you enter the short side of a forward contract, you can buy the underlying at market value S_T at the maturity date T and sell it for the strike price K. Thus, the final value of the short side of a forward contract is $\mathrm{K} - \mathrm{S}_\mathrm{T}$. The final value of the long side of a forward contract at maturity date T is $\mathrm{S}_\mathrm{T} - \mathrm{K}$ because the long-side must buy the asset for K and it can immediately sell the asset for the market price S_T. Figure 2.1 plots the final value or *payoff* of the long and short sides of a forward contact, as a function of the underlyer's final value S_T.

As previously mentioned, there is *no cost* to entering a forward contract (i.e., the initial value is zero for both the long and short side by design). This raises an important question: *What is the value of the strike price* K *that makes the initial value of the forward contract zero*? We will answer this question shortly.

Note that, even though one enters a forward contract at no cost, as the underlying asset's price S changes

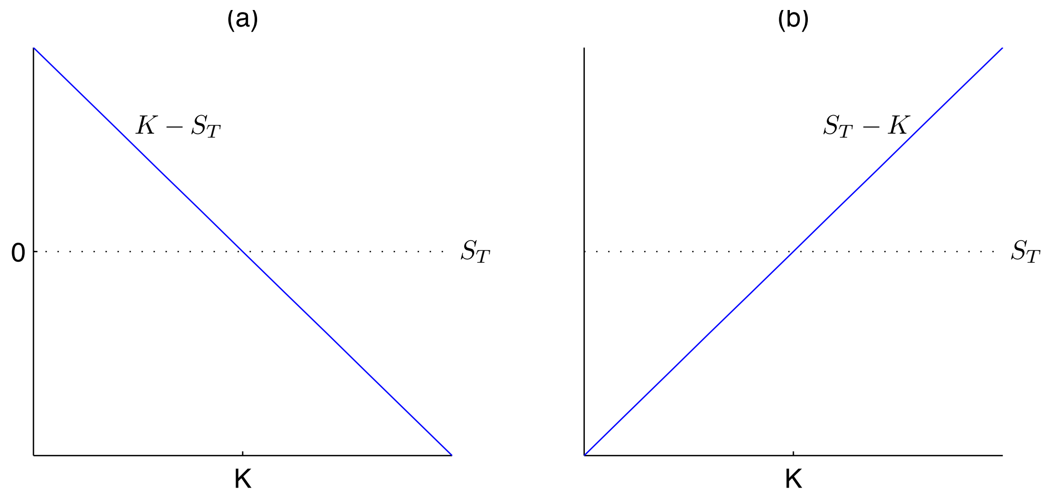

Figure 2.1: (a) Short side payoff. (b) Long side payoff.

over time, the value of the forward contract may become positive of negative. Another question we will seek to answer in this course is: *what the fair value of a forward contract should be at some time in the future*? For example, at time T (the maturity date) we know that the value of the long side of the forward contract is $S_T - K$. But, what is the value of the forward contract at $t \in (0, T)$?

2.2 Options

The two most basic types of options are *Calls* and *Puts*.

- A Call option gives its owner the right, but not the obligation, to *buy* an underlying S for a specified price K at or before a specified date T.
- A Put option gives its owner the right, but not the obligation, to *sell* an underlying S for a specified price K at or before a specified date T.

We call K the *strike* or *exercise price*. We call T the *maturity* or *expiration date*. Calls and Puts can further be classified as either *European* or *American*.

- *European* options can only be exercised at the maturity date T.
- *American* options can be exercised at any time $t \leq T$.

Although not common, there also exist *Bermudan* options, which can be exercised only on fixed dates $t_1 < t_2 < t_3 < t_4 < \cdots \leq T$. The owner of an option is said to be the *buyer* or *on the long side*. The seller of an option is said to be the *writer*, *seller* or *on the short side*.

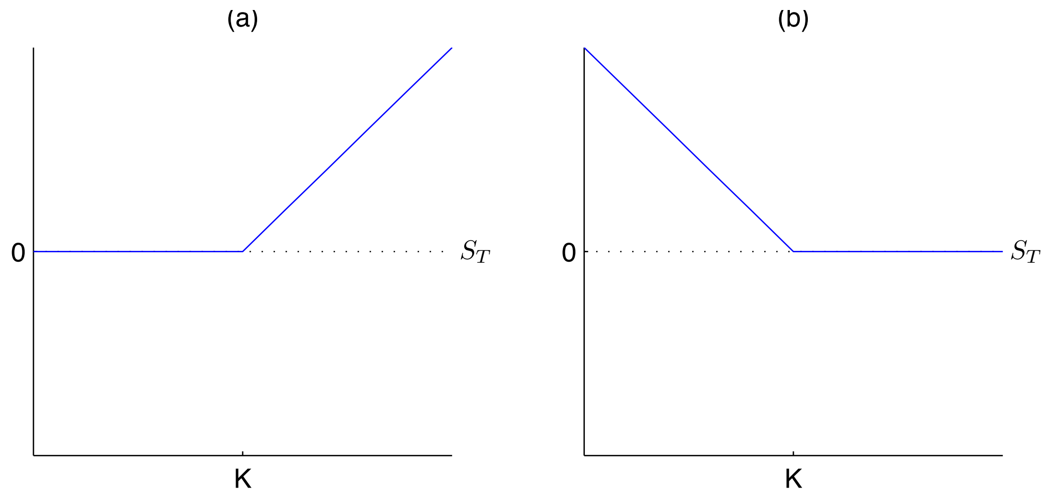

Figure 2.2: (a) Call payoff. (b) Put payoff.

Unlike a forward contract, which is a binding agreement to buy/sell an underlying S at a specified price K, the owner of an option is not required to exercise his right. The owner of an option will only exercise his right if it leads to profit. Consider a European call option. If at the maturity date T, the value of the undelying asset is larger than the strike price (i.e., $S_T > K$), then the owner can (and will) exercise his right to buy the asset for the strike price K and sell the asset at the market price S_T. The profit in this case is $S_T - K$. On the other hand, if the value of the underlying asset is less than the strike price (i.e., $S_T < K$), then the owner will not exercise his right to buy the asset because, rather than buy the asset for K he could buy the asset at the market price S_T. Thus, the value of the call option at time T is $(S_T - K)^+$ where $x^+ := \max\{x, 0\}$. Similarly, the value of a European Put option at time T is $(K - S_T)^+$. Take a moment to reason for yourself why this is the case. In Figure 2.2 we plot the final value or *payoff* of a European Call and Put option as a function of the underlyer's final value S_T.

Note that the payoff of both calls and puts Are non-negative. As such, the values calls and puts must also be non-negative. To own an option, one must *pay* the option seller a *premium* at inception. One important question we seek to answer in this course is: *what is the fair price to pay for an option*?

2.3 Types of traders

There are three basic types of traders:

- Hedgers seek to reduce risk from future movements of underlyings.
- Speculators seek to profit by making bets on future movements of underlyings.

- Arbitrageurs exploit mispricing in markets in order to make profit with little or no risk.

We will give a precise definition of *arbitrage* later in this text. For now, think of an arbitrage as an investment strategy that requires zero initial investment, and yields a non-negative profit at some future date.
Let is see some examples of hedging, speculating and arbitrage.

Example 2.6 (Hedging with options). Consider again the copper mine in Example 2.5. Let $S = (S_t)_{t \geq 0}$ be the price of copper. Suppose c is the cost to produce one unit of copper. In order to hedge against copper prices possibly falling, the copper mine might buy a put contract at time $t = 0$ on copper with strike K and maturity T at a price $P_0(T, K)$. The mine's profit/loss at time T is

$$
\begin{aligned}
S_T + (K - S_T)^+ - (c + P_0(T,K)) &= \begin{cases} K - (c + P_0(T,K)) & \text{if } S_T \leq K, \\ S_T - (c + P_0(T,K)) & \text{if } K < S_T. \end{cases} \\
&\geq K - (c + P_0(T,K)).
\end{aligned}
$$

Thus, the mine limits the downside of its profit/loss to $K-(c+P_0(T,K))$ and leaves the upside unbounded (because there is no limit to how large S_T can be). Had the mine hedged by taking the short side of a forward contract with delivery price K and maturity T, its profit/loss would have been $K - c$, which is a constant larger than $K - (c + P_0(T,K))$ but possibly less than $S_T - (c + P_0(T,K))$. Whether the mine chooses to hedge with a Put option or a forward contract (or whether it wants to hedge at all) is a matter of the mine's attitude towards risk and view of the future.

Example 2.7 (Speculating with options). Let $S = (S_t)_{t \geq 0}$ be the price of IBM stock. At time $t = 0$, the price of IBM stock is S_0. You think the price may go up to $1.2S_0$ by time T. You can speculate on this outcome in the following way:

- Buy a call option with maturity T and strike $K < 1.2S_0$ for a price of $C_0(T, K)$.

- If the price $S_T > K$ then you can buy stock for K by exercising the option and then sell the stock at the market price S_T. In this case, your profit is $S_T - K - C_0(T, K)$

- If $S_T < K$, you do not exercise your option because, rather than buy the stock for K you could buy it for S_T. In this case, you incur a loss of $C_0(T, K)$.

The overall profit/loss of your strategy is the final value of the option minus the price you paid for the call $(S_T - K)^+ - C_0(T, K)$.

Example 2.8 (Arbitrage with a forward). Let $S = (S_t)_{t \geq 0}$ be the price of a stock and let $r > 0$ be the (continuously compounded) risk-free rate of interest (i.e., you can lend or borrow at this rate). Suppose you can enter into a forward (long or short) with maturity T and delivery price $K \neq S_0 e^{rT}$. We will consider two scenarios: $K > S_0 e^{rT}$ and $K < S_0 e^{rT}$.

Case I: If $K > S_0 e^{rT}$, then you can achieve a risk-free profit by doing the following. At time $t = 0$,

- Enter the short side of the forward contract,
- Borrow S_0 from bank,
- Buy stock for S_0.

The total initial cost of this strategy is zero. Then at time $t = T$, you

- Sell stock for the delivery price $K > S_0 e^{rT}$,
- Repay the bank what you owe for the loan $S_0 e^{rT}$.

Your profit is $K - S_0 e^{rT} > 0$. Note that, with zero initial investment, you made a guaranteed profit! This is an arbitrage.

Case II: What would you do if $K < S_0 e^{rT}$? At time $t = 0$, you could

- Enter the long side of a forward contract,
- Sell stock S_0 (you do not need to own stock to do this)
- Put S_0 in bank

The total initial cost of this strategy is zero. At time $t = T$, you close out your position as follows

- Buy back the stock for the delivery price $K < S_0 e^{rT}$,
- Withdraw the money you invested in the bank, which as grown to $S_0 e^{rT}$.

In this case, your profit is $S_0 e^{rT} - K > 0$. Once again, with zero initial investment, you made a guaranteed profit! This is an arbitrage.

We have derived the following result.

Theorem 2.9. *Suppose the risk-free rate of interest is $r \geq 0$. Then, at time t, the no-arbitrage delivery price K of a forward contract written on an underlying $S = (S_t)_{t \geq 0}$ with a maturity date T is $K = F_t^T := S_t e^{r(T-t)}$. If $K \neq F_t^T$ there will be an arbitrage opportunity. We call F_t^T the fair forward price or simply forward price.*

Warning! The phrase *forward price* is very misleading. The forward price $\mathrm{F}_t^{\mathrm{T}}$ is *not* the value of a forward contract. Rather, for a forward contract maturing at time T, the forward price $\mathrm{F}_t^{\mathrm{T}}$ is the unique value that the delivery price K must be in order to preclude arbitrage. Really, we should call $\mathrm{F}_t^{\mathrm{T}}$ the *fair delivery price* of a forward contract.

2.4 Interest Rates and compounding

It will be important throughout this course to understand various notions of interest rates. Let $\mathrm{B} = (\mathrm{B}_t)_{t \geq 0}$ be the value of a bank account. Let $r \geq 0$ be the interest rate offered by the bank per unit time. Below are three different ways that the bank may choose to calculate interest.

- *Simple interest*: $\mathrm{B}_t = (1 + rt)\mathrm{B}_0$
- *Compound interest* (m times): $\mathrm{B}_t = (1 + rt/m)^m \mathrm{B}_0$
- *Continuous interest*: $\mathrm{B}_t = \lim_{m \to \infty}(1 + rt/m)^m \mathrm{B}_0 = \mathrm{e}^{rt}\mathrm{B}_0$.

Unless stated otherwise, *we will always consider continuously compounded interest rates*. Note that

$$(1 + rt) \leq (1 + rt/m)^m \leq \mathrm{e}^{rt},$$

where the inequalities are strict if $r > 0$.

2.5 Present values of certain cash flows

We will often talk about *present values* of *certain* (i.e., non-random) cash flows. Suppose, for example, you will receive a certain cash flow CF^{T} at time T. For example you may know that your employer pays you on the 1st and 15th of every month. Suppose you were offering to sell me your cash flow CF^{T} at time $t < \mathrm{T}$ for a price $\mathrm{P} < \mathrm{e}^{-r(\mathrm{T}-t)}\mathrm{CF}^{\mathrm{T}}$. Then, at time t, I could borrow P from the bank and purchase your cash flow. My initial investment is zero. At time T I would have $\mathrm{CF}^{\mathrm{T}} - \mathrm{P}\mathrm{e}^{r(\mathrm{T}-t)} = (\mathrm{CF}^{\mathrm{T}}\mathrm{e}^{-r(\mathrm{T}-t)} - \mathrm{P})\mathrm{e}^{r(\mathrm{T}-t)} > 0$. This is an arbitrage opportunity. On the other hand, suppose that at time $t < \mathrm{T}$, I offered to purchase your cash-flow of CF^{T} for a price $\mathrm{P} > \mathrm{e}^{-r(\mathrm{T}-t)}\mathrm{CF}^{\mathrm{T}}$. Then you could sell your cash flow to me for P and put your money in the bank. Your initial investment is zero. At time T you would have $\mathrm{P}\mathrm{e}^{r(\mathrm{T}-t)} - \mathrm{CF}^{\mathrm{T}} = \mathrm{e}^{r(\mathrm{T}-t)}(\mathrm{P} - \mathrm{e}^{-r(\mathrm{T}-t)}\mathrm{CF}^{\mathrm{T}}) > 0$. Again, this is an arbitrage opportunity. If we assume there is no arbitrage in the market, then, at time $t \leq \mathrm{T}$, the value of the cash flow CF^{T} must be $\mathrm{e}^{-r(\mathrm{T}-t)}\mathrm{CF}^{\mathrm{T}}$. We have derived the following result.

Theorem 2.10. *Assume the risk-free rate of interest is* $r \geq 0$. *Denote by* $\mathrm{PV} = (\mathrm{PV}_t)_{0\leq t\leq \mathrm{T}}$ *the no-arbitrage* present value *of a certain* cash flow CF^{T}, *to be paid at time* T. *Then we have*

$$\mathrm{PV}_t = \mathrm{e}^{-r(\mathrm{T}-t)}\mathrm{CF}^{\mathrm{T}},$$

for any $t \leq \mathrm{T}$.

Keep in mind, Theorem 2.10 is only valid when the cash-flow CF^{T} is a known, non-random constant. You cannot apply Theorem 2.10 to cash-flows that are random.

Definition 2.11. A *zero-coupon bond* pays one unit of currency at time T.

Theorem 2.12. *Let* $\mathrm{B}^{\mathrm{T}} = (\mathrm{B}^{\mathrm{T}}_t)_{0\leq t\leq \mathrm{T}}$ *denote the value of a zero-coupon bond maturiting at time* T. *Assume the risk free rate of interest is a constant* $r > 0$. *Then we have* $\mathrm{B}^{\mathrm{T}}_t = \mathrm{e}^{-r(\mathrm{T}-t)}$.

Proof. The bond pays a certain cash flow of 1 at time T. The present value of this cash flow $\mathrm{e}^{-r(\mathrm{T}-t)}$ is the bond price. □

2.6 The value of a forward contract after inception

In this section, we will derive the value of the long side of a forward contract. The value of the short-side of a forward contract is minus the value of the long side. Throughout this section, we will use the following notation:

- The maturity date of a forward contract is T.
- The underlying asset of the forward contract is $\mathrm{S} = (\mathrm{S}_t)_{t\geq 0}$.
- The risk-free rate of interest will be r.
- We denote by $\mathrm{F}^{\mathrm{T}}_t$ the fair forward price of a forward contract entered at time t. Recall from Example 2.8 and Theorem 2.9 that $\mathrm{F}^{\mathrm{T}}_t = \mathrm{e}^{r(\mathrm{T}-t)}\mathrm{S}_t$.
- We will denote by $f^{\mathrm{T}} = (f^{\mathrm{T}}_t)_{0\leq t\leq \mathrm{T}}$ the value of the long-side of a forward contract that was entered at time $t = 0$. As there is no cost to entering a forward contract, we know that $f^{\mathrm{T}}_0 = 0$. We also know that the final value of the long-side of a forward contract is $f^{\mathrm{T}}_{\mathrm{T}} = \mathrm{S}_{\mathrm{T}} - \mathrm{F}^{\mathrm{T}}_0$ (because $\mathrm{F}^{\mathrm{T}}_0$ was the fair forward price at time 0 when contract was entered into). Our goals is to find the value of f^{T}_t for $t \in (0, \mathrm{T})$.

We shall make the following assumptions:

- There is no arbitrage in market.
- One can lend and borrow at the risk-free rate of interest r.
- There are no transaction costs.

Theorem 2.13. *Under the assumptions listed above, we have*

$$f_t^T = S_t - F_0^T e^{-r(T-t)} = \left(F_t^T - F_0^T\right) e^{-r(T-t)}.$$

Proof. Suppose you enter the long side of a forward contract at time $t = 0$. At time t, you enter the short side of a forward contract with maturity T. The value of your portfolio at time t is

$$\underbrace{f_t^T}_{\text{value of long side of forward entered at time zero}} + \underbrace{0}_{\text{value of short side of forward entered at time } t}$$

The value of your portfolio at time T is

$$\underbrace{\left(S_T - F_0^T\right)}_{\text{value of long side of forward entered at time zero}} + \underbrace{\left(F_t^T - S_T\right)}_{\text{value of short side of forward entered at time } t} = F_t^T - F_0^T.$$

Note that the right-hand side above is a deterministic cash flow at time T. The present value of this cash flow at time t is

$$\left(F_t^T - F_0^T\right) e^{-r(T-t)}$$

Thus, we conclude that the value of the value at time t of the long side of a forward contract that was entered at time $t = 0$ is

$$f_t^T = \left(F_t^T - F_0^T\right) e^{-r(T-t)},$$

as claimed. □

Example 2.14. Suppose $S_0 = 100$ and $r = 0.10$. Consider a forward contract with a maturity date $T = 1$ year. What is the value at time $t = 6$ months of the long side of a forward contract that was entered at time zero, assuming $S_t = 50$? From Theorem 2.13 we have $f_t^T = S_t - F_0^T e^{-r(T-t)}$. And from Theorem 2.9 we have $F_0^T = e^{rT} S_0$. Thus, we have

$$f_t^T = S_t - (e^{rT} S_0) e^{-r(T-t)} = S_t - S_0 e^{rt} = 50 - 100e^{0.1*(1/2)} \approx -55.13.$$

It makes sense that the value is negative. The long side has agreed to buy the stock at time T for $F_0^T = e^{rT} S_0 \approx 110.52$ and yet the value of the stock at time t is only 50.

2.7 Futures Contracts

Definition 2.15. A *futures* contact written on an underlying S is an agreement between two parties. The contract specifies a series of dates $t_0, t_1, t_2, \ldots, t_n$ where $t_n = \mathrm{T}$. For $i = 1, 2, \ldots, n$, the long side of futures contract agrees to receive $\widehat{\mathrm{F}}^{\mathrm{T}}_{t_i} - \widehat{\mathrm{F}}^{\mathrm{T}}_{t_{i-1}}$ at time t_i while the short side agrees to pay $\widehat{\mathrm{F}}^{\mathrm{T}}_{t_i} - \widehat{\mathrm{F}}^{\mathrm{T}}_{t_{i-1}}$ on these dates (keep in mind $\widehat{\mathrm{F}}^{\mathrm{T}}_{t_i} - \widehat{\mathrm{F}}^{\mathrm{T}}_{t_{i-1}}$ could be negative). The process $\widehat{\mathrm{F}}^{\mathrm{T}} = (\widehat{\mathrm{F}}^{\mathrm{T}}_t)_{0 \le t \le \mathrm{T}}$ is called the *futures price*. The futures price $\widehat{\mathrm{F}}^{\mathrm{T}}_t$ is agreed upon at time t so that the cost to enter the futures contract is zero. The futures price at time T is specified to be $\widehat{\mathrm{F}}^{\mathrm{T}}_{\mathrm{T}} = \mathrm{S}_{\mathrm{T}}$.

Theorem 2.16. *Assume interest rates are a constant r. Then the futures price is equal to the forward price $\widehat{\mathrm{F}}^{\mathrm{T}}_t = \mathrm{F}^{\mathrm{T}}_t$ for all $t \in [0, \mathrm{T}]$.*

Proof. Fix some dates $t_0 = 0 < t_1 < t_2 < \ldots < t_n = \mathrm{T}$. Consider the following strategy: for $i = 0, 1, \ldots, n-1$, at time t_i go long a total of $\mathrm{e}^{r(t_{i+1})}$ futures contracts. Invest/finance all the gains/losses from your strategy in the bank at the risk-free rate of interest. The total profit/loss from this strategy is

$$
\begin{aligned}
\sum_{i=0}^{n-1} (\widehat{\mathrm{F}}^{\mathrm{T}}_{t_{i+1}} - \widehat{\mathrm{F}}^{\mathrm{T}}_{t_i}) \mathrm{e}^{r t_{i+1}} \mathrm{e}^{r(\mathrm{T} - t_{i+1})} &= \mathrm{e}^{r\mathrm{T}} \sum_{i=0}^{n-1} (\widehat{\mathrm{F}}^{\mathrm{T}}_{t_{i+1}} - \widehat{\mathrm{F}}^{\mathrm{T}}_{t_i}) \\
&= \mathrm{e}^{r\mathrm{T}} (\widehat{\mathrm{F}}^{\mathrm{T}}_{\mathrm{T}} - \widehat{\mathrm{F}}^{\mathrm{T}}_0) \\
&= \mathrm{e}^{r\mathrm{T}} (\mathrm{S}_{\mathrm{T}} - \widehat{\mathrm{F}}^{\mathrm{T}}_0)
\end{aligned}
$$

where we have used $\mathrm{S}_{\mathrm{T}} = \widehat{\mathrm{F}}^{\mathrm{T}}_{\mathrm{T}}$. Now consider another strategy: at time $t = 0$ go short a total of $\mathrm{e}^{r\mathrm{T}}$ Forward contracts. The profit/loss from this strategy is

$$
\mathrm{e}^{r\mathrm{T}} (\mathrm{F}^{\mathrm{T}}_0 - \mathrm{S}_{\mathrm{T}}).
$$

If we combine these two strategies into a single portfolio, the final value of our portfolio is

$$
\mathrm{e}^{r\mathrm{T}} (\mathrm{S}_{\mathrm{T}} - \widehat{\mathrm{F}}^{\mathrm{T}}_0) + \mathrm{e}^{r\mathrm{T}} (\mathrm{F}^{\mathrm{T}}_0 - \mathrm{S}_{\mathrm{T}}) = \mathrm{e}^{r\mathrm{T}} (\mathrm{F}^{\mathrm{T}}_0 - \widehat{\mathrm{F}}^{\mathrm{T}}_0).
$$

As there are no costs to enter future and forward contracts, there will be a clear arbitrage opportunity unless the futures at time zero $\widehat{\mathrm{F}}^{\mathrm{T}}_0$ is equal to the forward price at time zero $\mathrm{F}^{\mathrm{T}}_0$. Note that there is nothing special about $t_0 = 0$. The same argument works if we start at some $t_0 > 0$. We therefore have that the futures price is equal to the forward price $\widehat{\mathrm{F}}^{\mathrm{T}}_t = \mathrm{F}^{\mathrm{T}}_t$. □

Example 2.17. Let $r = 0.1$. An investor takes a long position in a futures contract on an underling with an initial value of $\mathrm{S}_0 = 100$. Suppose the futures contract has settlement dates $t_1 = 0.5$ and $t_2 = \mathrm{T} = 1$.

What are the cash flows paid/received by the long side if $S_{t_1} = 80$ and $S_T = 90$? The investor receives $\widehat{F}^T_{t_1} - \widehat{F}^T_0$ at time t_1 and receives $\widehat{F}^T_T - \widehat{F}^T_{t_1}$ at time T. Assuming constant interest rates, the futures price is equal to the forward price. Thus, we need to compute

$$\widehat{F}^T_0 = F^T_0 = S_0 e^{rT} \approx 110.52, \qquad \widehat{F}^T_{t_1} = F^T_{t_1} = S_{t_1} e^{r(T-t_1)} \approx 84.10, \qquad \widehat{F}^T_T = F^T_T = S_T = 90.$$

The the long receives $\widehat{F}^T_{t_1} - \widehat{F}^T_0 = 84.10 - 110.52 = -26.42$ at time t_1 and $\widehat{F}^T_T - \widehat{F}^T_{t_1} = 90 - 84.10 = 5.90$ at time T.

2.8 Factors that affect Call and Put Prices

Throughout this and subsequent sections, we will denote by $C(T,K) = (C_t(T,K))_{0 \le t \le T}$ and $= P(T,K) = (P_t(T,K))_{0 \le t \le T}$, respectively, the prices of a Call and Put options with maturity T and strike K, written on an underlying $S = (S_t)_{0 \le t}$. Various factors affect Call and Put prices. The main factors are

- The value of the underlying asset S_t,
- The time to maturity $T - t$,
- The strike price K,
- The *volatility* of the underlying asset σ
- The risk-free rate of interest r.

We have not yet defined what we mean by *volatility*. In fact, as we will see in this course, there are many different notions of volatility. For now, just think of volatility as a rough measure of how large $\mathbb{V}(\log S_T | S_t)$ (the conditional variance of $\log S_T$ given S_t) is. A larger volatility σ implies a larger variance $\mathbb{V}(\log S_T | S_t)$. Note that $T - t$ and K are specific to the option contract, while σ and S_t are properties of the underlying asset S. The interest rate r generally comes from fixed income market (i.e., the bond marked). Let us examine how each of the above factors affect Call and Put prices.

Increasing $\sigma^2(T-t)$ increase the value both calls and puts. The reason is that, the larger $\sigma^2(T-t)$ is, the larger the variance $\mathbb{V}(\log S_T | S_t)$ becomes. The larger the variance $\mathbb{V}(\log S_T | S_t)$ is, the larger the probability that S_T is *in the money* (i.e., that $S_T > K$ for a Call or that $S_T < K$ for a Put).

Increases the strike K lowers the value of a Call and raises the value of a Put. Suppose for example, that $K_2 > K_1$. Then

$$C_T(T, K_1) = (S_T - K_1)^+ \ge (S_T - K_2)^+ = C_T(T, K_2),$$

$$\mathrm{P_T(T,K_1)} = (\mathrm{K_1} - \mathrm{S_T})^+ \leq (\mathrm{K_2} - \mathrm{S_T})^+ = \mathrm{P_T(T,K_2)}.$$

Because $\mathrm{C_T(T,K_1)} \geq \mathrm{C_T(T,K_2)}$ and $\mathrm{P_T(T,K_1)} \leq \mathrm{P_T(T,K_2)}$ it follows that $\mathrm{C}_t(\mathrm{T,K_1}) \geq \mathrm{C}_t(\mathrm{T,K_2})$ and $\mathrm{P_T(T,K_1)} \leq \mathrm{P_T(T,K_2)}$. If this were not the case, there would be a clear arbitrage opportunity.

Increasing S_t increases the value of a Call and decreases the value of a Put. The reason is that, the larger S_t is, the higher the probability that a Call finishes in the money (i.e., $\mathrm{S_T} > \mathrm{K}$) and the lower the probability that a Put finishes in the money (i.e., $\mathrm{S_T} < \mathrm{K}$).

2.9 Bounds for European options

The core argument that we will use in this and subsequent sections is that, if the value of a portfolio $\mathrm{X^A} = (\mathrm{X}_t^\mathrm{A})_{t\geq 0}$ is greater or equal to the value of another portfolio $\mathrm{X^B} = (\mathrm{X}_t^\mathrm{B})_{t\geq 0}$ at time T, then the value of $\mathrm{X^A}$ must be greater or equal to the value of $\mathrm{X^B}$ at any given time $t \leq \mathrm{T}$. More concisely, we have

$$\mathrm{X_T^A} \geq \mathrm{X_T^B} \qquad \Rightarrow \qquad \mathrm{X}_t^\mathrm{A} \geq \mathrm{X}_t^\mathrm{B} \qquad \forall t \in [0,\mathrm{T}],$$

If the above were not true, there would be an obvious arbitrage opportunity.

Theorem 2.18 (Upper bounds for European Calls and Puts). *We have*

$$\mathrm{C}_t(\mathrm{T,K}) \leq \mathrm{S}_t, \qquad \mathrm{P}_t(\mathrm{T,K}) \leq \mathrm{KB}_t^\mathrm{T}.$$

Proof. We have

$$\begin{aligned}
\mathrm{C_T(T,K)} &= (\mathrm{S_T} - \mathrm{K})^+ \leq \mathrm{S_T} && \Rightarrow && \mathrm{C}_t(\mathrm{T,K}) \leq \mathrm{S}_t, \\
\mathrm{P_T(T,K)} &= (\mathrm{K} - \mathrm{S_T})^+ \leq \mathrm{K} && \Rightarrow && \mathrm{P}_t(\mathrm{T,K}) \leq \mathrm{KB}_t^\mathrm{T}.
\end{aligned}$$

□

Establishing Lower bounds for European options requires a little more work (but not much).

Theorem 2.19 (Lower bound for European Call options).

$$\mathrm{C}_t(\mathrm{T,K}) \geq (\mathrm{S}_t - \mathrm{KB}_t^\mathrm{T})^+. \tag{2.9.1}$$

Proof. First, we claim that

$$\mathrm{C}_t(\mathrm{T,K}) \geq \mathrm{S}_t - \mathrm{KB}_t^\mathrm{T}. \tag{2.9.2}$$

To see this, for $t \in [0, T]$, define

$$X_t^A = C_t(T, K), \qquad X_t^B = S_t - KB_t^T.$$

Then we have

$$X_T^A = C_T(T, K) = (S_T - K)^+, \qquad X_T^B = S_T - K.$$

We also have

$$X_T^A \geq X_T^B \qquad \Rightarrow \qquad X_t^A \geq X_t^B,$$

which is (2.9.2). Next, note that, as the payoff of a call option is non-negative, the value of a call option must be non-negative $C_t(T, K) \geq 0$. Thus, we obtain (2.9.1). □

Theorem 2.20 (Lower bound for Euopean Put Options).

$$P_t(T, K) \geq (KB_t^T - S_t)^+. \tag{2.9.3}$$

Proof. First, we claim that

$$P_t(T, K) \geq KB_t^T - S_t. \tag{2.9.4}$$

To see this, for and $t \in [0, T]$, define

$$X_t^A = P_t(T, K), \qquad X_t^B = KB_t^T - S_t.$$

Then we have

$$X_T^A = P_T(T, K) = (K - S_T)^+, \qquad X_T^B = KB_T^T - S_T = K - S_T.$$

We also have

$$X_T^A \geq X_T^B \qquad \Rightarrow \qquad X_t^A \geq X_t^B,$$

which is (2.9.4). Next, note that, as the payoff of a put option is non-negative, the value of a put option must be non-negative $P_t(T, K) \geq 0$. Thus, we obtain (2.9.3). □

2.10 Put-Call Parity

There exists a very important relation between European calls and European puts with the same strike K and maturity T.

Theorem 2.21 (Put-Call Parity). *The following relation holds between European Call and Put options*

$$C_t(T,K) + KB_t^T = P_t(T,K) + S_t. \tag{2.10.1}$$

Proof. For $t \in [0,T]$, define

$$X_t^A = C_t(T,K) - P_t(T,K), \qquad X_t^B = S_t - KB_t^T$$

Note that

$$X_T^A = (S_T - K)^+ - (K - S_T)^+ = S_T - K, \qquad X_T^B = S_T - KB_T^T = S_T - KB_T^T = S_T - K$$

Consequently, we have

$$X_T^A = X_T^B \qquad \Rightarrow \qquad X_t^A = X_t^B,$$

and the last equality implies equation (2.10.1). □

Note the following special case: if $K = S_t/B_t^T$ then $C_t(T, S_t/B_t^T) = P_t(T, S_t/B_t^T)$. When $K = S_t/B_t^T$ we say the strike K is *at-the-money.*

2.11 American vs European options

In this section, we denote with a superscript "^{A}" American options and with a superscript "^{E}" European options. We claim that $C_t^A(T,K) \geq C_t^E(T,K)$ and $P_t^A(T,K) \geq P_t^E(T,K)$. The reason is that, the owner of an American option has the same right as the owner of a European option (i.e., the ability to exercise the option at time T), but additionally has the right to exercise prior to the maturity date T. The additional optionality makes the American option more valuable than its European counterpart.

Theorem 2.22. *An investor should never exercise an American Call (on a non-dividend paying stock) early. Thus,* $C_t^A(T,K) = C_t^E(T,K)$.

Proof. Using the fact that, for a fixed maturity T and strike K and American option is worth at least as much than a European option, and recalling the lower bound for a European Call, we have

$$C_t^A(T,K) \geq C_t^E(T,K) \geq S_t - KB_t^T. \tag{2.11.1}$$

Now, consider two cases:

Case I: $S_t \leq K$. In this case, one would not exercise the option because the payoff for doing so is zero: $(S_t - K)^+ = 0$.

Case II: $S_t > K$. In this case, if you exercise, you would get a payoff of $S_t - K$. However, note that

$$S_t - K \leq S_t - KB_t^T \overset{(2.11.1)}{\leq} C_t^A(T, K)$$

Thus, it would be better to sell the American Call rather than exercise it. We therefore conclude that it is never optimal to exercise an American call option on a non-dividend paying stock prior to the maturity date T. □

Note that the same argument does *not* apply to an American put. It may (and often is) optimal to exercise an American Put option prior to maturity.

2.12 Examples of arbitrage

From the properties established in the previous sections, you should be able to identify some obvious arbitrage opportunities. In what follows, all options are European-style.

Example 2.23. Suppose $C_t(T, K_1) < C_t(T, K_2)$ with $K_1 < K_2$. This is an obvious arbitrage. Buy one share of $C_t(T, K_1)$, sell one share of $C_t(T, K_2)$ and invest $(C_t(T, K_2) - C_t(T, K_1))$ in a bank account. This strategy requires zero initial investment and at maturity yields a strictly positive value

$$(S_T - K_1)^+ - (S_T - K_2)^+ + (C_t(T, K_2) - C_t(T, K_1))e^{r(T-t)} > 0.$$

Example 2.24. Suppose the risk-free rate of interest is zero $r = 0$ and $C_t(T, S_t) > P_t(T, S_t)$. Again, this is an obvious arbitrage. One can sell one call, buy one put, go long a forward contract, and invest $C_t(T, S_t) - P_t(T, S_t)$ in the bank. At time T this strategy yields

$$\begin{aligned}
&P_T(T, S_t) - C_T(T, S_t) + (S_T - S_t) + (C_t(T, S_t) - P_t(T, S_t)) \\
&= (S_t - S_T)^+ - (S_T - S_t)^+ + (S_T - S_t) + (C_t(T, S_t) - P_t(T, S_t)) \\
&= (C_t(T, S_t) - P_t(T, S_t)) > 0.
\end{aligned}$$

2.13 Constructing Other Payoffs

In this section, we show that put and call payoffs

$$\text{Call Payoffs :} \qquad \phi_C(S_T; K) = (S_T - K)^+, \qquad \text{Put Payoffs :} \qquad \phi_P(S_T; K) = (K - S_T)^+,$$

can be used to construct more general payoffs for the form $\phi(S_T)$, where ϕ is an arbitrary function. Let us begin with two very simple examples.

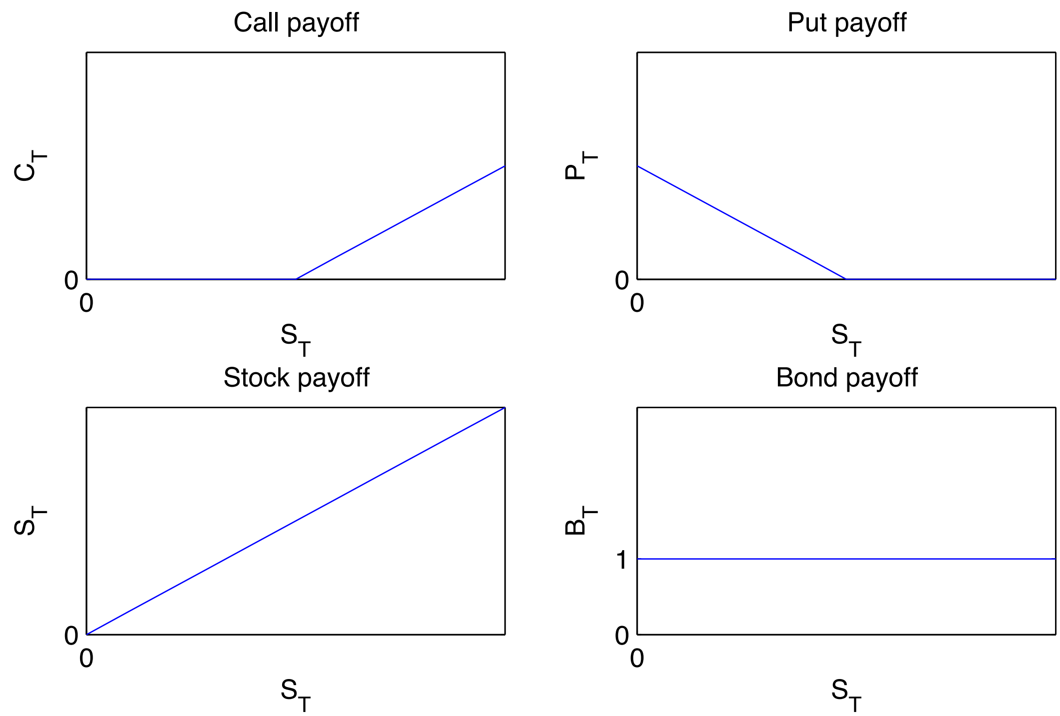

Figure 2.3: The payoffs of calls, puts, stocks and bonds.

Example 2.25. Defining $\varphi(x) := \varphi_{\mathrm{P}}(x;\mathrm{K}_1) + \varphi_{\mathrm{C}}(x;\mathrm{K}_2)$ with $\mathrm{K}_2 > \mathrm{K}_1$, we obtain the payoff depicted in Figure 2.4.

Example 2.26. Defining $\varphi(x) := \varphi_{\mathrm{C}}(x;\mathrm{K}_1) - \varphi_{\mathrm{C}}(x;\mathrm{K}_2)$ with $\mathrm{K}_2 > \mathrm{K}_1$, we obtain the payoff depicted in Figure 2.5.

The following result, from Carr and Madan (1998), shows how to construct very general functions φ from call and put payoffs.

Theorem 2.27. *For any φ that is the difference of convex functions, we have*

$$\begin{aligned}\varphi(x) = \varphi(\lambda) + \varphi'(\lambda)\Big((x-\lambda)^+ - (\lambda - x)^+\Big)\\ + \int_0^\lambda \varphi''(\mathrm{K})(\mathrm{K}-x)^+ \mathrm{dK} + \int_\lambda^\infty \varphi''(\mathrm{K})(x-\mathrm{K})^+\mathrm{dK}.\end{aligned} \tag{2.13.1}$$

where φ' and φ'' denote the first and second derivative of φ and $\lambda \geq 0$ is an arbitrary constant.

Proof. Using integration by parts twice, we have

$$\int_0^\lambda \varphi''(\mathrm{K})(\mathrm{K}-x)^+\mathrm{dK} = (\mathrm{K}-x)^+\varphi'(\mathrm{K})\Big|_0^\lambda - \int_0^\lambda \varphi'(\mathrm{K})\theta(\mathrm{K}-x)\mathrm{dK}$$

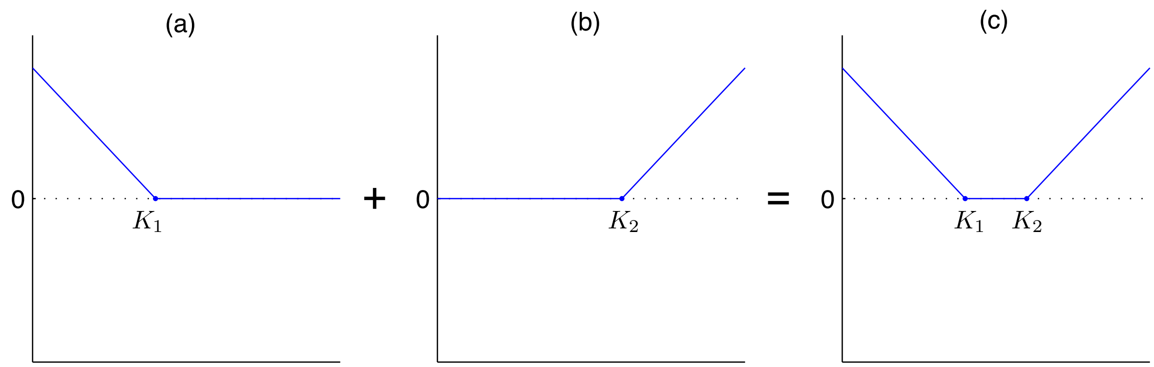

Figure 2.4: A plot of (a) $\varphi_{\mathrm{P}}(\,\cdot\,;\mathrm{K}_1)$, (b) $\varphi_{\mathrm{C}}(\,\cdot\,;\mathrm{K}_2)$, and (c) $\varphi = \varphi_{\mathrm{P}}(\,\cdot\,;\mathrm{K}_1) + \varphi_{\mathrm{C}}(\,\cdot\,;\mathrm{K}_2)$.

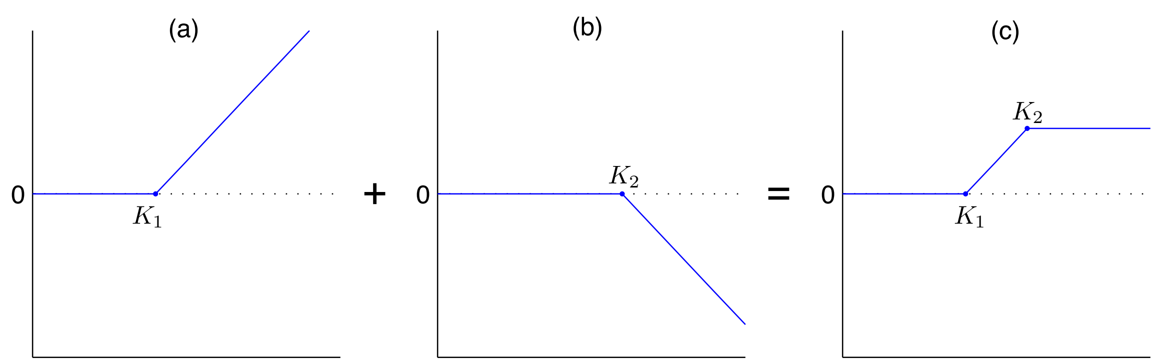

Figure 2.5: A plot of (a) $\varphi_{\mathrm{C}}(\,\cdot\,;\mathrm{K}_1)$, (b) $-\varphi_{\mathrm{C}}(\,\cdot\,;\mathrm{K}_2)$, and (c) $\varphi = \varphi_{\mathrm{C}}(\,\cdot\,;\mathrm{K}_1) - \varphi_{\mathrm{C}}(\,\cdot\,;\mathrm{K}_2)$.

$$
\begin{aligned}
&= (\mathrm{K}-x)^+\varphi'(\mathrm{K})\Big|_0^\lambda - \varphi(\mathrm{K})\theta(\mathrm{K}-x)\Big|_0^\lambda + \int_0^\lambda \varphi(\mathrm{K})\delta(\mathrm{K}-x)\mathrm{dK} \\
&= (\lambda-x)^+\varphi'(\lambda) - \varphi(\lambda)\theta(\lambda-x) + \int_0^\lambda \varphi(\mathrm{K})\delta(\mathrm{K}-x)\mathrm{dK} \qquad (2.13.2)
\end{aligned}
$$

where we have used

$$
(\mathrm{K}-x)^+ = \int_{-\infty}^{\mathrm{K}} \theta(\lambda-x)\mathrm{d}\lambda, \qquad \theta(\mathrm{K}-x) = \int_{-\infty}^{\mathrm{K}} \delta(\lambda-x)\mathrm{d}\lambda,
$$

which implies

$$
\partial_{\mathrm{K}}(\mathrm{K}-x)^+ = \theta(\mathrm{K}-x), \qquad \partial_{\mathrm{K}}\theta(\mathrm{K}-x) = \delta(\mathrm{K}-x).
$$

Next, once again using integration by parts, we have

$$
\begin{aligned}
\int_\lambda^\infty \varphi''(\mathrm{K})(x-\mathrm{K})^+\mathrm{dK} &= (x-\mathrm{K})^+\varphi'(\mathrm{K})\Big|_\lambda^\infty + \int_\lambda^\infty \varphi'(\mathrm{K})\theta(x-\mathrm{K})\mathrm{dK} \\
&= (x-\mathrm{K})^+\varphi'(\mathrm{K})\Big|_\lambda^\infty + \varphi(\mathrm{K})\theta(x-\mathrm{K})\Big|_\lambda^\infty + \int_\lambda^\infty \varphi(\mathrm{K})\delta(x-\mathrm{K})\mathrm{dK} \\
&= -(x-\lambda)^+\varphi'(\lambda) - \varphi(\lambda)\theta(x-\lambda) + \int_\lambda^\infty \varphi(\mathrm{K})\delta(x-\mathrm{K})\mathrm{dK} \qquad (2.13.3)
\end{aligned}
$$

where we have used

$$
(x-\mathrm{K})^+ = \int_{\mathrm{K}}^\infty \theta(x-\lambda)\mathrm{d}\lambda, \qquad \theta(x-\mathrm{K}) = \int_{\mathrm{K}}^\infty \delta(x-\lambda)\mathrm{d}\lambda,
$$

which implies

$$
\partial_{\mathrm{K}}(x-\mathrm{K})^+ = -\theta(x-\mathrm{K}), \qquad \partial_{\mathrm{K}}\theta(x-\mathrm{K}) = -\delta(x-\mathrm{K}).
$$

Inserting (2.13.2) and (2.13.3) into (2.13.1) and using

$$
\int_0^\infty \varphi(\mathrm{K})\delta(x-\mathrm{K})\mathrm{dK} = \varphi(x),
$$

we find that the right-hand side simplifies to $\varphi(x)$, as claimed. □

Now, consider a European option whose payoff is $\varphi(\mathrm{S_T})$. Let V_t be the value of this option at time t. Suppose we can observe call and put prices at every strike $\mathrm{K} \in [0,\infty)$. Denote the time t price of a European Call option with strike K and maturity T by $\mathrm{C}_t(\mathrm{T},\mathrm{K})$. Likewise, denote by $\mathrm{P}_t(\mathrm{T},\mathrm{K})$ the price of a European Put option with strike K and maturity T. Finally, let $\mathrm{B}_t^{\mathrm{T}} = \mathrm{e}^{-r(\mathrm{T}-t)}$ be the price of a zero-coupon bond that pays 1 at time T. We know from the above result that

$$
\varphi(\mathrm{S_T}) = \varphi(\lambda) + \varphi'(\lambda)\Big((\mathrm{S_T}-\lambda)^+ - (\lambda-\mathrm{S_T})^+\Big)
$$

$$
\begin{aligned}
&+\int_0^\lambda \varphi''(\mathrm{K})(\mathrm{K}-\mathrm{S_T})^+ d\mathrm{K} + \int_\lambda^\infty \varphi''(\mathrm{K})(\mathrm{S_T}-\mathrm{K})^+ d\mathrm{K} \\
&= \varphi(\lambda)\mathrm{B_T^T} + \varphi'(\lambda)\Big(\mathrm{C_T}(\mathrm{T},\lambda) - \mathrm{P_T}(\mathrm{T},\lambda)\Big) \\
&+\int_0^\lambda \varphi''(\mathrm{K})\mathrm{P_T}(\mathrm{T},\mathrm{K}) d\mathrm{K} + \int_\lambda^\infty \varphi''(\mathrm{K})\mathrm{C_T}(\mathrm{T},\mathrm{K}) d\mathrm{K}
\end{aligned}
$$

where we have used $(\mathrm{S_T}-\mathrm{K})^+ = \mathrm{C_T}(\mathrm{T},\mathrm{K})$ and $(\mathrm{K}-\mathrm{S_T})^+ = \mathrm{P_T}(\mathrm{T},\mathrm{K})$. Therefore, we must have

$$
\begin{aligned}
\mathrm{V}_t &= \mathrm{B}_t^\mathrm{T}\varphi(\lambda) + \varphi'(\lambda)\Big(\mathrm{C}_t(\mathrm{T},\lambda) - \mathrm{P}_t(\mathrm{T},\lambda)\Big) \\
&+\int_0^\lambda \varphi''(\mathrm{K})\mathrm{P}_t(\mathrm{T},\mathrm{K}) d\mathrm{K} + \int_\lambda^\infty \varphi''(\mathrm{K})\mathrm{C}_t(\mathrm{T},\mathrm{K}) d\mathrm{K}. \qquad (2.13.4)
\end{aligned}
$$

Note that, everywhere we had a call payoff $\mathrm{C_T}(\mathrm{T},\mathrm{K})$ we have replaced it with the corresponding call price $\mathrm{C}_t(\mathrm{T},\mathrm{K})$ and likewise for puts and bonds. What equation (2.13.4) tells us is that the value V_t of a European derivative with payoff $\varphi(\mathrm{S_T})$ can be obtained by observing call and put prices $(\mathrm{C}_t(\mathrm{T},\mathrm{K}))_{\mathrm{K}\geq 0}$ and $(\mathrm{P}_t(\mathrm{T},\mathrm{K}))_{\mathrm{K}\geq 0}$, respectively). This is a *model free* result. We do not need to assume any model for S in order to derive the value V_t of the option. We only need to observe European call and put prices on the market.

Example 2.28. Assume for simplicity that $r = 0$. In terms of Calls and Puts, what is the time $t = 0$ value of a contract that pays $\varphi(\mathrm{S_T}) = -2\log(\mathrm{S_T}/\mathrm{S_0})$ at time T? We have

$$
\varphi(s) = -2\log(s/\mathrm{S_0}), \qquad \varphi'(s) = -2/s, \qquad \varphi''(s) = 2/s^2.
$$

Note that (2.13.4) holds for any $\lambda \geq 0$. Let's choose $\lambda = \mathrm{S_0}$. Note that

$$
\varphi(\mathrm{S_0}) = -2\log(\mathrm{S_0}/\mathrm{S_0}) = -2\log 1 = 0,
$$

so the first term in (2.13.4) is zero. We also have, by put-call parity that $\mathrm{C_0}(\mathrm{T},\mathrm{S_0}) - \mathrm{P_0}(\mathrm{T},\mathrm{S_0}) = 0$ (make sure you understand why), so the second term in (2.13.4) is zero. Finally, we are left with

$$
\mathrm{V_0} = \int_0^{\mathrm{S_0}} \frac{2}{\mathrm{K}^2}\mathrm{P_0}(\mathrm{T},\mathrm{K}) d\mathrm{K} + \int_{\mathrm{S_0}}^\infty \frac{2}{\mathrm{K}^2}\mathrm{C_0}(\mathrm{T},\mathrm{K}) d\mathrm{K}.
$$

We will come back to this example when we talk about variance swaps.

Example 2.29. Suppose $\varphi(\mathrm{S_T})$ is given by

$$
\varphi(\mathrm{S_T}) = \begin{cases} 0 & \mathrm{S_T} < \mathrm{K_1}, \\ \mathrm{S_T} - \mathrm{K_1}, & \mathrm{K_1} \leq \mathrm{S_T} < \mathrm{K_2}, \\ \mathrm{K_2} - \mathrm{K_1}, & \mathrm{K_2} \leq \mathrm{S_T}. \end{cases}
$$

We can write the time t value of an option with this payoff in terms of calls and puts a follows. First, we have

$$\varphi'(s) = \begin{cases} 0 & s < K_1, \\ 1, & K_1 \leq s < K_2, , \\ 0, & K_2 \leq s, \end{cases} \qquad \varphi''(s) = \delta(s - K_1) - \delta(s - K_2).$$

Let's choose $\lambda < K_1$. Then, using (2.13.4) we have

$$\begin{aligned} V_t &= \int_\lambda^\infty \varphi''(K) C_t(T, K) dK \\ &= \int_\lambda^\infty \delta(K - K_1) C_t(T, K) dK - \int_\lambda^\infty \delta(K - K_1) C_t(T, K) dK. \\ &= C_t(T, K_1) - C_t(T, K_2). \end{aligned}$$

2.14 Exercises

Exercise 2.1. Suppose that if you hold an asset S then, for $i = 1, 2, \ldots, n$ you will receive known dividend payments of D_i at time $t_i < T$ (if you short the asset S, then you will owe the dividend payments to the person you sold the asset to). (a) Use no-arbitrage arguments to derive the fair forward price F_t^T. It may help to first find F_0^T. (b) Suppose you enter a forward contract written on S at time zero. What is the *value* f_t^T of this forward contract at time $t < T$?

Exercise 2.2. Suppose I offer to pay you a certain cash flow CF_1 at time t_1 if you pay me a certain cash flow of CF_2 at time t_2. You think about it for a moment and tell me: "I'll make that trade if you also pay me V_0 today at time zero." I think about your offer and say: "that sounds fair to me." (a) What is CF_2 in terms of V_0, r, t_1, t_2 and CF_1? (b) Suppose $CF_1 > 0$ and $CF_2 > 0$. Define $R := CF_1/CF_2$. Give a condition on R which will guarantee that $V_0 > 0$. (c) What is the condition on R that guarantees $V_0 > 0$ when $CF_1 < 0$ and $CF_2 > 0$?

Exercise 2.3. Suppose there is a European contract that at time T pays

$$\text{Payoff} = \begin{cases} 0, & S_T < K_1, \\ S_T - K_1 & K_1 \leq S_T < K_2, \\ K_2 - K_1 & K_2 \leq S_T. \end{cases}$$

Use the upper and lower bounds we derived for European call options to give upper and lower bounds on the price of this contract at time $t < T$.

Exercise 2.4. We used the argument many times that, if $X_T^A \geq X_T^B$ we must also have $X_t^A \geq X_t^B$ to preclude arbitrage. Suppose that $X_T^A \geq X_T^B$ and $X_t^A < X_t^B$. Construct a trading strategy X that satisfies $X_t = 0$ and $X_T \geq 0$.

Chapter 3

Binomial model

Up to now we have not put *any* model on the underlyer S, although we did need to make some assumptions on the risk-free rate of interest. We deduced forward and futures prices, bounds for calls and puts, as well as the Put-Call Parity relation using no arbitrage arguments. Model-free results are fantastic when available because they are robust to model misspecification. Unfortunately, we have gone as far as we can with model-free results. In order to price Calls and Puts, we need to assume a model for the dynamics of S. We will begin with a very simple model. This simple model will help us understand the fundamental concepts of option pricing. Once we understand this simple model, we will be able to easily understand the more realistic models that will be presented later in this text.

3.1 One-Step Binomial Model

Consider an experiment in which we toss a coin. Our sample space is $\Omega = \{\mathrm{H}, \mathrm{T}\}$. Obviously, there are only two possible outcomes $\omega = \mathrm{H}$ and $\omega = \mathrm{T}$. The probability of these outcomes is $\mathbb{P}(\mathrm{H}) = p$ and $\mathbb{P}(\mathrm{T}) = 1 - p =: q$. Now, consider a market on the sample space that includes a stock $\mathrm{S} = (\mathrm{S}_n)_{n \in \{0,1\}}$ and a bond $\mathrm{B} = (\mathrm{B}_n)_{n \in \{0,1\}}$. The value of the stock S_0 at time $n = 0$ is known. At time $n = 1$, the value of the stock is given by

$$\mathrm{S}_1 = \begin{cases} u\mathrm{S}_0, & \text{if } \omega = \mathrm{H}, \\ d\mathrm{S}_0, & \text{if } \omega = \mathrm{T}. \end{cases}$$

Figure 3.1 depicts how the stock S behaves in the one-step binomial model. In case it is not obvious from Figure 3.1, the variable u stands of "up" and the variable d stands for "down." The bond has an initial value of B_0 and grows to a value of $\mathrm{B}_1 = (1 + r)\mathrm{B}_0$. Figure 3.2 shows how the bond behaves in

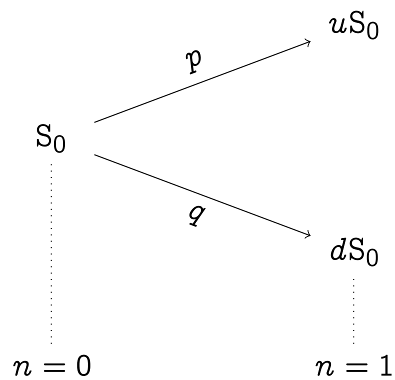

Figure 3.1: A one-step binomial tree for a stock S.

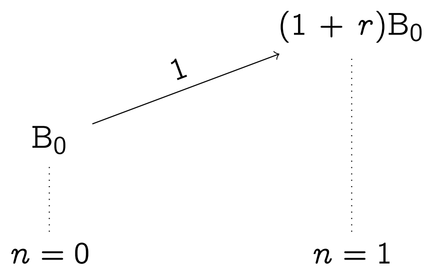

Figure 3.2: The behavior of a bond B in the one-step binomial model.

the one-step binomial model. The no-arbitrage assumption implies that

$$0 < d < 1 + r < u.$$

Take a moment to derive why this must be the case!

3.1.1 Pricing European options by replication

We now derive the value of a European option with payoff $\phi(S_1)$. We will do this by finding a portfolio $X = (X_n)_{n\in\{0,1\}}$ that has the same value at time $n = 1$ as the option (i.e., $X_1 = \phi(S_1)$). This approach is called *pricing by replication*. Let Δ_0 be the number of shares of S in the portfolio at time $n = 0$. If the remaining value of the portfolio is invested in bonds, then the portfolio would contain $(X_0 - \Delta_0 S_0)/B_0$ bonds. We have

$$X_0 = \underbrace{\Delta_0}_{\#\text{ shares}} S_0 + \underbrace{\frac{X_0 - \Delta_0 S_0}{B_0}}_{\#\text{ of bonds}} B_0.$$

Note that both Δ_0 and $(X_0 - \Delta_0 S_0)/B_0$ can be negative. At time $n = 1$ the value of the portfolio is

$$X_1 = \Delta_0 S_1 + (X_0 - \Delta_0 S_0)\frac{1}{B_0}B_1 = \begin{cases} \Delta_0 u S_0 + (X_0 - \Delta_0 S_0)(1+r) & \text{if } \omega = H, \\ \Delta_0 d S_0 + (X_0 - \Delta_0 S_0)(1+r) & \text{if } \omega = T. \end{cases}$$

Let $V = (V_n)_{n \in \{0,1\}}$ denote the value of the option. The value of the option at time $n = 1$ is

$$V_1 = \phi(S_1) = \begin{cases} \phi(uS_0) & \text{if } \omega = H, \\ \phi(dS_0) & \text{if } \omega = T. \end{cases}$$

Thus, to create a portfolio for which $X_1 = V_1$ we must have

$$\Delta_0 u S_0 + (X_0 - \Delta_0 S_0)(1+r) = \phi(uS_0), \qquad \Delta_0 d S_0 + (X_0 - \Delta_0 S_0)(1+r) = \phi(dS_0).$$

Solving for X_0 and Δ_0, we obtain

$$\Delta_0 = \frac{\phi(uS_0) - \phi(dS_0)}{uS_0 - dS_0}, \qquad X_0 = \frac{1}{1+r}\Big(\widetilde{p}\phi(uS_0) + \widetilde{q}\phi(dS_0)\Big), \tag{3.1.1}$$

where we have introduced $\widetilde{p}$ and $\widetilde{q}$, which are given by

$$\widetilde{p} = \frac{(1+r) - d}{u - d}, \qquad \widetilde{q} = 1 - \widetilde{p}. \tag{3.1.2}$$

Because we have created a portfolio X for which $V_1 = X_1$, we must also have

$$V_0 = X_0, \tag{3.1.3}$$

because, if (3.1.3) did not hold, there would be a clear arbitrage opportunity. Suppose, for example, that $V_0 < X_0$, and try to construct an investment strategy that has zero initial cost and guarantees a profit at time $n = 1$.

Remark 3.1. Note that value of the option V_0, the initial value of the replication portfolio X_0 and the number of shares Δ_0 of the stock S do *not* depend on the probabilities p and q! Therefore, we could have changed the probabilities p and q and this would not affect the price of the option. This is quite an astonishing and unexpected result. We will use this in the next section.

Example 3.2. Suppose S evolves as described in the binomial model with $u = 1.1$, $d = 0.9$ and $S_0 = 100$ and $r = 0$. What is the price of a Call option with maturity at time T with strike $K = 100$? The payoff of the Call is $\phi(S_1) = (S_1 - 100)^+$. In the event the stock goes up we have $S_1 = uS_0 = 110$ and the call payoff is $\phi(uS_0) = (110 - 100)^+ = 10$. In the event the stock price goes down, we have $S_1 = dS_0 = 90$ and the call payoff is $\phi(dS_0) = (90 - 100)^+ = 0$. Thus, we must find X_0 and Δ_0 so that

$$\Delta_0 u S_0 + (X_0 - \Delta_0 S_0)(1+r) = \phi(uS_0),$$

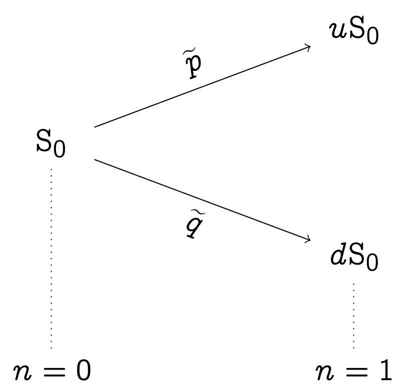

Figure 3.3: A risk-neutral one-step binomial tree for a stock.

$$\Delta_0 110 + (X_0 - \Delta_0 100) = 10$$

and

$$\Delta_0 dS_0 + (X_0 - \Delta_0 S_0)(1 + r) = \phi(dS_0),$$
$$\Delta_0 90 + (X_0 - \Delta_0 100) = 0.$$

Solving for X_0, we find $X_0 = 5$. We could find the put price in a similar fashion. But, a faster method to obtain the put price is to use Put-Call parity. We have $C_0(1, K) - P_0(1, K) = S_0 - K$. So, for $K = S_0 = 100$ we have $C_0(1, S_0) = P_0(1, S_0) = 5$.

3.1.2 Risk-Neutral Valuation of European options

We now provide an alternative method of valuing derivatives known as *risk-neutral valuation*. Let us consider the binomial model under a new probability measure $\widetilde{\mathbb{P}}$. Let us denote by $\widetilde{p}$ and $\widetilde{q}$, respectively, the probabilities of heads and tails

$$\widetilde{\mathbb{P}}(H) = \widetilde{p}, \qquad\qquad \widetilde{\mathbb{P}}(T) = \widetilde{q}.$$

Forget that we have previously assigned values to $\widetilde{p}$ and $\widetilde{q}$. Assume for the moment that the values of $\widetilde{p}$ and $\widetilde{q}$ are unknown. The dynamics of S under $\widetilde{\mathbb{P}}$ are depicted in Figure 3.3. Observe in Figure 3.3 that $\widetilde{\mathbb{P}}(S_1 = uS_0) = \widetilde{p}$ and $\widetilde{\mathbb{P}}(S_1 = dS_0) = \widetilde{q}$ whereas in Figure 3.1 we have $\mathbb{P}(S_1 = uS_0) = p$ and $\mathbb{P}(S_1 = dS_0) = q$. In other words, in Figures 3.1 and 3.3 the *same* events can occur (i.e., the stock goes up or the stock goes down), but the *probabilities* of these events have changed.

Now, let us denote by $\widetilde{\mathbb{E}}$ an expectation taken under a probability measure $\widetilde{\mathbb{P}}$. Let us try to find the values of $\widetilde{p}$ and $\widetilde{q}$ such that $\widetilde{\mathbb{E}}(X_1/B_1) = X_0/B_0$. Stated another way, let us search for a probability measure $\widetilde{\mathbb{P}}$ under which the process X/B is a martingale. You will see why this is useful in a moment. For any choice of Δ_0, we have

$$\begin{aligned}\widetilde{\mathbb{E}}\left(\frac{X_1}{B_1}\right) &= \Delta_0 \widetilde{\mathbb{E}}\left(\frac{S_1}{B_1}\right) + \frac{(X_0 - \Delta_0 S_0)}{B_0}\frac{B_1}{B_1} \\ &= \Delta_0\left(\widetilde{p}\frac{uS_0}{B_0(1+r)} + \widetilde{q}\frac{dS_0}{B_0(1+r)}\right) + \frac{(X_0 - \Delta_0 S_0)}{B_0}.\end{aligned}$$

Thus, in order for $\widetilde{\mathbb{E}}(X_1/B_1) = X_0/B_0$, we must have

$$\frac{X_0}{B_0} = \Delta_0\left(\widetilde{p}\frac{uS_0}{B_0(1+r)} + \widetilde{q}\frac{dS_0}{B_0(1+r)}\right) + \frac{(X_0 - \Delta_0 S_0)}{B_0}.$$

Using the above equation, and the fact that $\widetilde{p} + \widetilde{q} = 1$, we find that $\widetilde{p}$ and $\widetilde{q}$ are given by (3.1.2).

Now, recall from Section 3.1.1 that, by choosing Δ_0 and X_0 according to (3.1.1), we can set up a replicating portfolio X such that $X_1 = \phi(S_1) = V_1$. Recall further that neither X_0 nor Δ_0 depended on the probability measure $\mathbb{P}$ (see Remark 3.1). Thus, changing from $\mathbb{P}$ to $\widetilde{\mathbb{P}}$ does not affect the value of an option. Suppose we did not already know the initial value X_0 of the replicating portfolio. We can use the fact that X/B is a martingale under $\widetilde{\mathbb{P}}$ to find X_0 rather easily. We have

$$\frac{X_0}{B_0} = \widetilde{\mathbb{E}}\left(\frac{X_1}{B_1}\right) = \widetilde{\mathbb{E}}\left(\frac{\phi(S_1)}{B_1}\right) = \widetilde{p}\frac{\phi(uS_0)}{B_0(1+r)} + \widetilde{q}\frac{\phi(dS_0)}{B_0(1+r)}.$$

Multiplying both side by B_0 we obtain

$$X_0 = \frac{1}{(1+r)}\Big(\widetilde{p}\phi(uS_0) + \widetilde{q}\phi(dS_0)\Big),$$

which agrees with (3.1.1). Because we know that $X_1 = \phi(S_1) = V_1$, it must be the case that the initial value of the option must be $V_0 = X_0$. Thus, we have stumbled across a very simple two-step method to price options in the binomial model.

Step 1: Find a probability measure $\widetilde{\mathbb{P}}$ (meaning find the values of $\widetilde{p}$ and $\widetilde{q}$) so that X/B is a martingale

$$\frac{X_0}{B_0} = \widetilde{\mathbb{E}}\left(\frac{X_1}{B_1}\right). \tag{3.1.4}$$

Step 2: Compute the initial value V_0 of an option that pays $\phi(S_1)$ using

$$\frac{V_0}{B_0} = \widetilde{\mathbb{E}}\left(\frac{V_1}{B_1}\right) = \widetilde{\mathbb{E}}\left(\frac{\phi(S_1)}{B_1}\right). \tag{3.1.5}$$

Computing option prices using the above two-step process is known as *risk-neutral pricing.*

We call $\widetilde{\mathbb{P}}$ the *risk neutral probability measure* or *martingale measure*. We call B the *numéraire*. If we denominate the value of a portfolio X in units of the numéraire B, then under the risk-neutral measure $\widetilde{\mathbb{P}}$, equation (3.1.4) tells us that, in expectation, the value of X does not change.

In fact, there is nothing special about using the bond B as the numéraire. We can use any strictly positive portfolio as numéraire. For example, if we use S as the numéraire, then the above two-step procedure would become

Step 1: Find $\widehat{\mathbb{P}}$ (meaning assign probabilities to events) so that X/S is a martingale

$$\frac{X_0}{S_0} = \widehat{\mathbb{E}}\left(\frac{X_1}{S_1}\right).$$

Step 2: Compute the initial value V_0 of an option that pays $\phi(S_1)$ using

$$\frac{V_0}{S_0} = \widehat{\mathbb{E}}\left(\frac{V_1}{S_1}\right) = \widehat{\mathbb{E}}\left(\frac{\phi(S_1)}{S_1}\right).$$

Try this for yourself and see that you get the same price V_0 of an option when you use S as numéraire as when you use B as numéraire.

Example 3.3. Let's compute the price of a call option assuming $S_0 = 100$, $K = 100$, $u = 1.1$, $d = 0.9$ and $r = 0$. First, we must find $\widetilde{p}$ and $\widetilde{q}$. Using (3.1.4) we have

$$\begin{aligned}
\frac{S_0}{B_0} &= \widetilde{\mathbb{E}}\left[\frac{S_1}{B_1}\right] \\
&= \widetilde{p}\frac{uS_0}{B_0 e^{rT}} + \widetilde{q}\frac{dS_0}{B_0 e^{rT}} \\
&= \frac{S_0}{B_0}(1.1\widetilde{p} + 0.9\widetilde{q}).
\end{aligned}$$

Using $\widetilde{q} = 1 - \widetilde{p}$, we find $\widetilde{p} = 0.5$ and $\widetilde{q} = 0.5$. Next, using (3.1.5) we have

$$\begin{aligned}
\frac{V_0}{B_0} &= \widetilde{\mathbb{E}}\left[\frac{V_1}{B_1}\right] = \widetilde{\mathbb{E}}\left[\frac{(S_1 - K)^+}{B_1}\right] \\
&= \widetilde{p}\left(\frac{(uS_0 - K)^+}{B_0(1+r)}\right) + \widetilde{q}\left(\frac{(dS_0 - K)^+}{B_0(1+r)}\right) \\
&= 0.5\left(\frac{10}{B_0}\right) + 0.5\left(\frac{0}{B_0}\right) = \frac{5}{B_0}.
\end{aligned}$$

Thus, we have $V_0 = 5$.

3.2 N-step binomial model

We are now going to extend the one-period binomial model to an N-step binomial model. Consider and experiment in which we toss a coin N times. Our sample space is

$$\Omega = \{\text{all sequences of H's and T's of length N}\}.$$

We will denote a sample outcome ω of this experiment as follows

$$\omega = \omega_1\omega_2\dots\omega_N, \qquad\qquad \omega_i \in \{H, T\}$$

For example, if $N = 5$, a possible outcome is $\omega = HTTHH$. We will assume that coin tosses are independent and that $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = 1 - p =: q$.

Now, consider a market on the sample space that includes a stock $S = (S_n)_{n\in\{0,1\dots,N\}}$ and a bond $B = (B_n)_{n\in\{0,1,\dots,N\}}$. The value of the stock S_0 at time $n = 0$ is known. At time $n+1$, the value of the stock is given by

$$S_{n+1} = \begin{cases} uS_n, & \text{if } \omega_{n+1} = H, \\ dS_n, & \text{if } \omega_{n+1} = T. \end{cases}$$

For example, if $\omega = HHT\dots$ then we would have $S_1 = uS_0$, $S_2 = u^2S_0$ and $S_3 = u^2dS_0$. Figure 3.4 depicts how the stock S behaves in the 3-step binomial model. Note that tree recombines. Also note in the figure that

$$S_n^{(i)} = u^i d^{n-i} S_0, \qquad\qquad i \in \{0, 1, \dots, n\}. \tag{3.2.1}$$

The bond in the N-step binomial model has an initial value of B_0. At time $n+1$, the value of the bond is given by

$$B_{n+1} = (1+r)B_n.$$

Thus, we have $B_n = (1+r)^n B_0$. As in the one-step binomial model, we will assume $0 < d < 1+r < u < \infty$, which guarantees that there is no arbitrage in the market.

3.2.1 Pricing European options by replication

Consider a European option that pays $\phi(S_N)$ at time N. We would like to find a portfolio $X = (X_n)_{n\in\{0,1,\dots,N\}}$ that replications the payoff $X_N = \phi(S_N)$. Figure 3.5 depicts how the portfolio X behaves in the 3-step binomial model. Assuming replication is possible, then at time N we know that $X_N^{(i)} = \phi(S_N^{(i)})$ for $i = 0, 1, 2, \dots, N$. We would like to find $X_n^{(i)}$ and $\Delta_n^{(i)}$ for $n = 0, 1, \dots, N-1$ and $i = 0, 1, \dots, n$.

Suppose the value of the stock and replicating portfolio at time n are $S_n^{(i)}$ and $X_n^{(i)}$, respectively. If the portfolio contains $\Delta_n^{(i)}$ of the stock S, then we can write the value $X_n^{(i)}$ as follows

$$X_n^{(i)} = \Delta_n^{(i)} S_n^{(i)} + (X_n^{(i)} - \Delta_n^{(i)} S_n^{(i)}) \frac{1}{B_n} B_n.$$

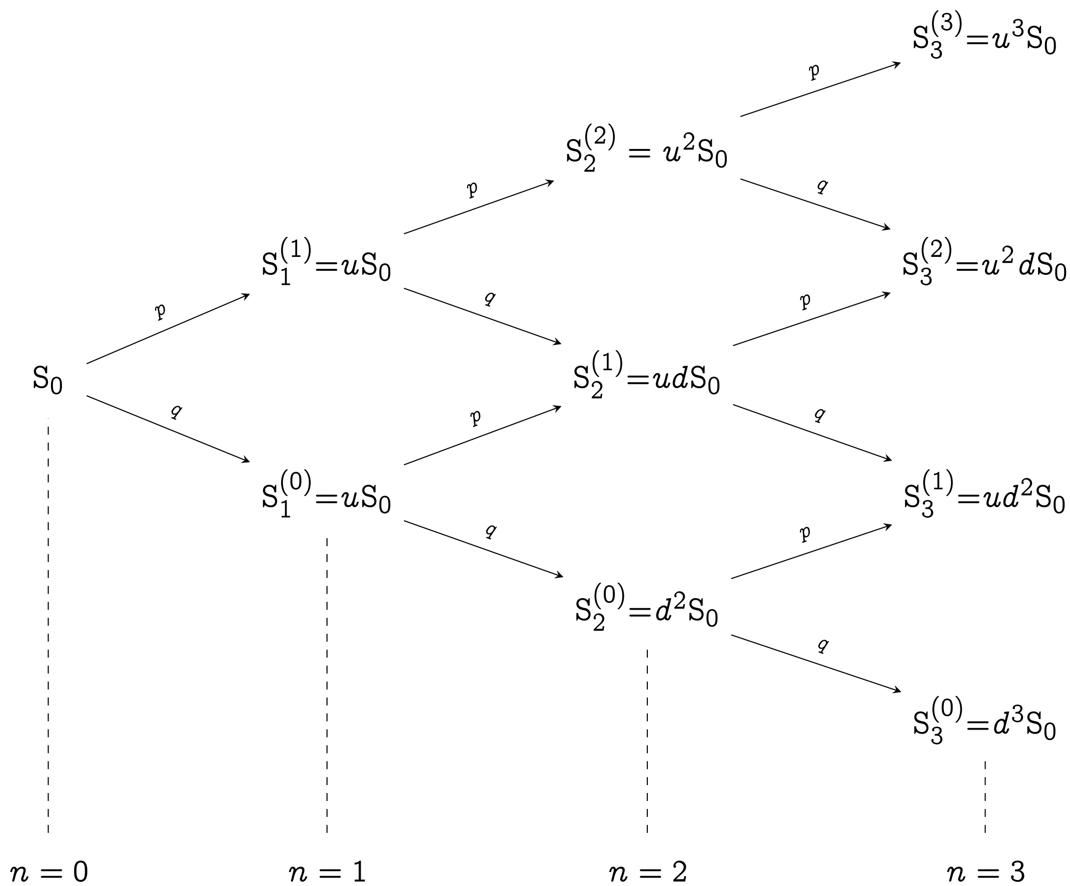

Figure 3.4: A 3-step binomial tree for a stock S.

At time $n+1$, that value of the portfolio becomes

$$X_{n+1} = \begin{cases} \Delta_n^{(i)} S_{n+1}^{(i+1)} + (X_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r), & \text{if } \omega_{n+1} = H, \\ \Delta_n^{(i)} S_{n+1}^{(i)} + (X_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r), & \text{if } \omega_{n+1} = T. \end{cases}$$

We would like to have

$$X_{n+1} = \begin{cases} X_{n+1}^{(i+1)}, & \text{if } \omega_{n+1} = H, \\ X_{n+1}^{(i)}, & \text{if } \omega_{n+1} = T. \end{cases}$$

In order for this to occur, the following two equations must be satisfied

$$X_{n+1}^{(i+1)} = \Delta_n^{(i)} S_{n+1}^{(i+1)} + (X_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r),$$
$$X_{n+1}^{(i)} = \Delta_n^{(i)} S_{n+1}^{(i)} + (X_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r).$$

Solving the above system of equations for $X_n^{(i)}$ and $\Delta_n^{(i)}$, we obtain

$$X_n^{(i)} = \frac{1}{1+r}\Big(\widetilde{p}X_{n+1}^{(i+1)} + \widetilde{q}X_{n+1}^{(i)}\Big), \qquad \Delta_n^{(i)} = \frac{X_{n+1}^{(i+1)} - X_{n+1}^{(i)}}{S_{n+1}^{(i+1)} - S_{n+1}^{(i)}}, \tag{3.2.2}$$

where $\widetilde{p}$ and $\widetilde{q}$ are given by (3.1.2). Using the fact that $X_N^{(i)} = \phi(S_N^{(i)})$, where $S_n^{(i)}$ is given by (3.2.1), we can now use backwards recursion to find $X_n^{(i)}$ and $\Delta_n^{(i)}$ at every node. First we find $X_{N-1}^{(i)}$ and $\Delta_{N-1}^{(i)}$ for $i = 1, 2, \ldots, N-1$. Then we find $X_{N-2}^{(i)}$ and $\Delta_{N-2}^{(i)}$ for $i = 1, 2, \ldots, N-2$, etc..

Because, by constriction, the portfolio X replicates the option payoff $X_N = \phi(S_N)$, we conclude that the value $V = (V_n)_{n\in\{1,2,\ldots,N\}}$ of the derivative asset is given by $V_n = X_n$ for all n.

Remark 3.4. Note once again that, as in the one-step binomial model, in the N-step binomial model, the probabilities $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = q$ do not affect the values of X_n, Δ_n or V_n! Thus, we can change the probabilities p and q and this will not affect the values of the option or the replication strategy. We will use this in the next section.

3.2.2 Risk-neutral pricing of European options

In this section, we will demonstrate how to value a derivative asset in the N-period binomial model using the risk-neutral pricing method. To begin let us consider the the N-step binomial model under a probability measure $\widetilde{\mathbb{P}}$. Note that we are *not* changing how the stock S behaves as a function of ω (i.e., if $\omega = HHT\ldots$, then under both $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ we have $S_1 = uS_0$, $S_2 = u^2S_0$ and $S_3 = u^2dS_0$). We are only changing the *probabilities* of getting a H or T at every step. Suppose that, under $\widetilde{\mathbb{P}}$, the probabilities of heads and tails are given by

$$\widetilde{\mathbb{P}}(H) = \widetilde{p}, \qquad \widetilde{\mathbb{P}}(T) = \widetilde{q},$$

where $\widetilde{p}$ and $\widetilde{q}$ are, for the moment, unknown. Let $\widetilde{\mathbb{E}}_n$ denote an expectation under $\widetilde{\mathbb{P}}$ conditioned on the history of the stock $\mathcal{F}_n^S = \{S_0, S_1, \ldots, S_n\}$. That is, for any random variable Z, we have

$$\widetilde{\mathbb{E}}_n Z = \widetilde{\mathbb{E}}(Z|\mathcal{F}_n^S) = \widetilde{\mathbb{E}}(Z|S_0, S_1, \ldots, S_n).$$

Let us search for the values of $\widetilde{p}$ and $\widetilde{q}$ such that $\mathbb{E}_n(X_{n+1}/B_{n+1}) = X_n/B_n$. You will see why this is useful in a bit. Assuming that Δ_n and X_n can be determined from $\mathcal{F}_n^S$ (which is true if one considers only investment strategies that depend on past movements of S), we have

$$\widetilde{\mathbb{E}}_n\left(\frac{X_{n+1}}{B_{n+1}}\right) = \Delta_n\widetilde{\mathbb{E}}_n\left(\frac{S_{n+1}}{B_{n+1}}\right) + \frac{(X_n - \Delta_n S_n)}{B_0}\frac{B_1}{B_1}$$

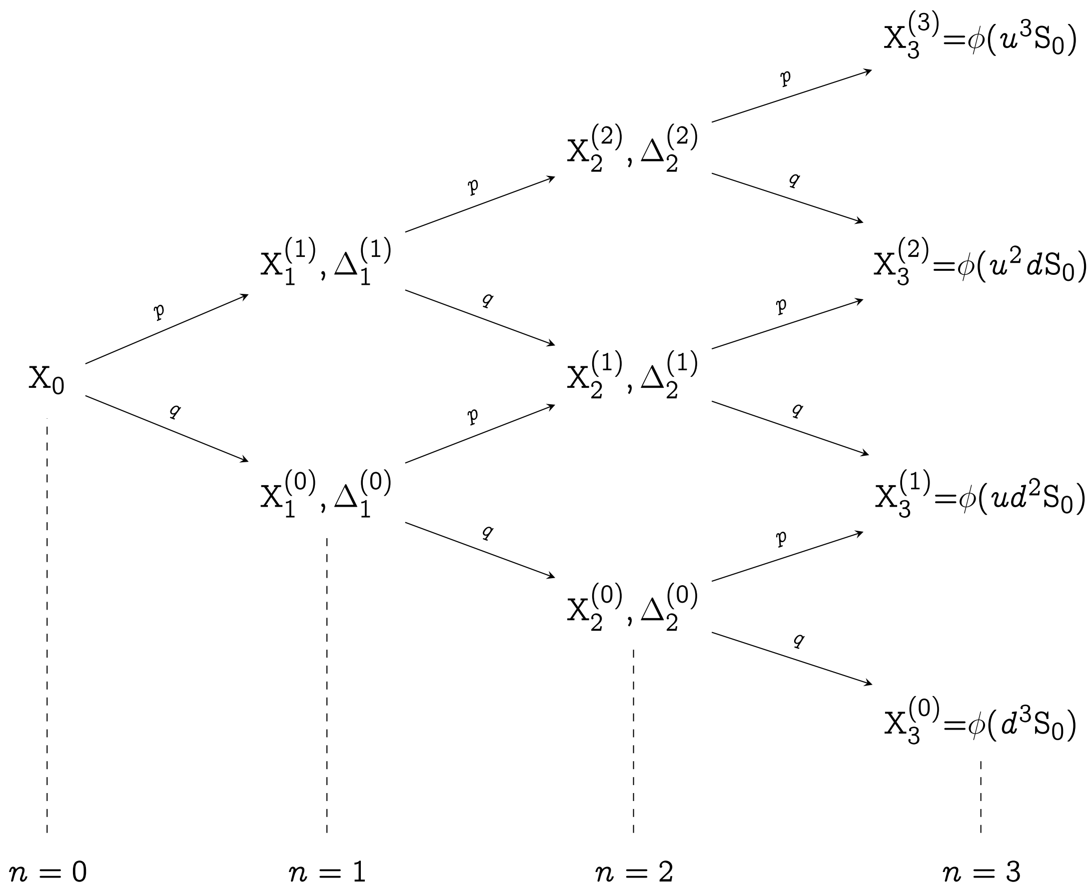

Figure 3.5: A 3-step binomial tree for a portfolio X.

$$= \Delta_n\Big(\widetilde{p}\frac{uS_n}{B_0(1+r)} + \widetilde{q}\frac{dS_n}{B_0(1+r)}\Big) + \frac{(X_0 - \Delta_0 S_0)}{B_0}.$$

Thus, in order for $\mathbb{E}_n(X_{n+1}/B_{n+1}) = X_n/B_n$ we must have

$$\frac{X_n}{B_n} = \Delta_n\Big(\widetilde{p}\frac{uS_n}{B_0(1+r)} + \widetilde{q}\frac{dS_n}{B_0(1+r)}\Big) + \frac{(X_0 - \Delta_0 S_0)}{B_0}.$$

Using the above equation, and the fact that $\widetilde{p} + \widetilde{q} = 1$, we find that $\widetilde{p}$ and $\widetilde{q}$ are given by (3.1.2). Now, with $\widetilde{p}$ and $\widetilde{q}$ given by (3.1.2), we have

$$\mathbb{E}_n\left(\frac{X_{n+m}}{B_{n+m}}\right) = \mathbb{E}_n\mathbb{E}_{n+m-1}\left(\frac{X_{n+m}}{B_{n+m}}\right) = \mathbb{E}_n\left(\frac{X_{n+m-1}}{B_{n+m-1}}\right) = \ldots = \mathbb{E}_n\left(\frac{X_{n+1}}{B_{n+1}}\right) = \frac{X_n}{B_n}. \quad (3.2.3)$$

Thus, for any investment strategy (X, Δ) we have that the process X/B is a martingale under $\widetilde{\mathbb{P}}$.

Now, recall from Section 3.2.1 that, by choosing Δ_n and X_n according to (3.2.2), we can set up a replicating portfolio X such that $X_N = \phi(S_N) = V_N$. Recall further that neither X_n nor Δ_n depended on the probability measure $\mathbb{P}$ (see Remark 3.4). Thus, changing from $\mathbb{P}$ to $\widetilde{\mathbb{P}}$ does *not* affect the value of an option V_n, the value of the replicating portfolio X_n or the trading strategy Δ_n. Suppose we did not already know the value X_n of the replicating portfolio. We can use the fact that X/B is a martingale under $\widetilde{\mathbb{P}}$ to find X_n rather easily. We have

$$\frac{X_n^{(i)}}{B_n} = \widetilde{\mathbb{E}}_n\left(\frac{X_{n+1}}{B_{n+1}}\right) = \widetilde{p}\frac{X_{n+1}^{(i+1)}}{B_n(1+r)} + \widetilde{q}\frac{X_{n+1}^{(i)}}{B_n(1+r)}.$$

Mutiplying both side by B_n we obtain

$$X_n^{(i)} = \frac{1}{(1+r)}\left(\widetilde{p}X_{n+1}^{(i+1)} + \widetilde{q}X_{n+1}^{(i)}\right),$$

which agrees with (3.2.2). We can now use backwards recursion to find X_n at any node. Alternatively, we could simply compute

$$\begin{aligned}\frac{X_n}{B_n} &= \widetilde{\mathbb{E}}_n\left(\frac{X_N}{B_N}\right) = \widetilde{\mathbb{E}}_n\left(\frac{\phi(S_N)}{B_N}\right)\\ &= \sum_{k=0}^{N-n}\frac{\phi(u^{N-n-k}d^kS_n)}{B_n(1+r)^{N-n}}\mathbb{P}(S_N = u^{N-n-k}d^kS_n|S_n),\\ \mathbb{P}(S_N = u^{N-n-k}d^kS_n|S_n) &= \binom{N-n}{k}\widetilde{p}^{N-n-k}\widetilde{q}^k.\end{aligned}$$

Though, for large $N-n$ is typically faster computationally to use backwards recursion. Because we know that $X_N = \phi(S_N) = V_N$, it must be the case that $V_n = X_n$ for all n. Thus, we have stumbled across a very simple two-step method to price options in the binomial model.

Step 1: Find a probability measure $\widetilde{\mathbb{P}}$ (meaning find the values of $\widetilde{p}$ and $\widetilde{q}$) so that X/B is a martingale

$$\frac{X_n}{B_n} = \widetilde{\mathbb{E}}_n\left(\frac{X_{n+1}}{B_{n+1}}\right). \tag{3.2.4}$$

Note that, from (3.2.3), equation (3.2.4) is sufficient to show that X/B is a martingale under $\widetilde{\mathbb{P}}$.

Step 2: Compute the initial value V_n of an option that pays $\phi(S_N)$ using

$$\frac{V_n}{B_n} = \widetilde{\mathbb{E}}_n\left(\frac{V_{n+1}}{B_{n+1}}\right), \qquad n = 0, 1, \ldots, N-1, \qquad V_N = \phi(S_N). \tag{3.2.5}$$

Computing option prices using the above two-step process is known as *risk-neutral pricing*.

As in the one-step binomial model, we call $\widetilde{\mathbb{P}}$ the *risk neutral probability measure* or *martingale measure*. We call B the *numéraire*. If we denominate the value of a portfolio X in units of the numéraire B, then under the risk-neutral measure $\widetilde{\mathbb{P}}$, equation (3.2.4) tells us that, in expectation, the value of X

does not change.

Also, as in the one-period case, there is nothing special about using the bond B as the numéraire. We can use any strictly positive portfolio as numéraire. For example, if we use S as the numéraire, then the above two-step procedure would become

Step 1: Find $\widehat{\mathbb{P}}$ (meaning assign probabilities to events) so that X/S is a martingale

$$\frac{X_n}{S_n} = \widehat{\mathbb{E}}\left(\frac{X_{n+1}}{S_{n+1}}\right).$$

Step 2: Compute the initial value V_n of an option that pays $\phi(S_N)$ using

$$\frac{V_n}{S_n} = \widehat{\mathbb{E}}_n\left(\frac{V_{n+1}}{S_{n+1}}\right), \qquad n = 0, 1, \ldots, N-1, \qquad V_N = \phi(S_N). \tag{3.2.6}$$

Example 3.5. Consider the two-step binomial model $N = 2$. Assume $r = 0$, $S_0 = 100$, $u = 1.1$ and $d = 0.9$. What is the price of a call option with strike $K = 100$? At time $n = 2$ the stock S can be one of three values $u^2 S_0 = 121$, $udS_0 = 99$, and $d^2 S_0 = 81$. The payoff $\phi(S_2) = (S_2 - K)^+$ in these three states is $(u^2 S_0 - K)^+ = 21$, $(udS_0 - K)^+ = 0$ and $(d^2 S_0 - K)^+ = 0$. With u and d as given, we have $\widetilde{p} = \widetilde{q} = 0.5$. Thus, defining

$$\begin{aligned} V_1^{(1)} &:= \widetilde{\mathbb{E}}[\phi(S_2)|S_1 = uS_0] = \widetilde{p} \cdot 21 + \widetilde{q} \cdot 0 = 10.5, \\ V_1^{(0)} &:= \widetilde{\mathbb{E}}[\phi(S_2)|S_1 = dS_0] = \widetilde{p} \cdot 0 + \widetilde{q} \cdot 0 = 0, \end{aligned}$$

the value of the option V_0 is given by

$$V_0 = \widetilde{\mathbb{E}}[V_1] = \widetilde{p} V_1^{(1)} + \widetilde{q} V_1^{(0)} = \widetilde{p} \cdot 10.5 + \widetilde{q} \cdot 0 = 5.25.$$

3.2.3 Risk-neutral pricing of American Options

Recall that an American option gives its holder the right, but not the obligation, to exercise the option at any time before or at maturity. It will be optimal for the owner of an American option to exercise his right only if the value of immediately exercising the option is greater than the value of holding the option. From (3.2.6), we have that

$$\{\text{value of holding the option at time } n\} = B_n \widetilde{\mathbb{E}}_n\left(\frac{V_{n+1}}{B_{n+1}}\right). \tag{3.2.7}$$

On the other hand, we also have that

$$\{\text{value of exercising the option at time } n\} = \phi(S_n). \tag{3.2.8}$$

The value of the American option at time n is the maximum of (3.2.7) and (3.2.8). Thus, we have

$$\mathrm{V}_n = \max\left\{\phi(\mathrm{S}_n), \mathrm{B}_n\widetilde{\mathbb{E}}_n\left(\frac{\mathrm{V}_{n+1}}{\mathrm{B}_{n+1}}\right)\right\}, \qquad n = 0, 1, \ldots, \mathrm{N}-1, \qquad \mathrm{V}_\mathrm{N} = \phi(\mathrm{S}_\mathrm{N}). \quad (3.2.9)$$

Using (3.2.9), we can value an American option using backwards recursion. The following example illustrates this procedure.

Example 3.6. Consider an American Put option in the two-step binomial model. We fix the following parameters

$$u = 1.2, \qquad d = 0.8, \qquad r = 0.051, \qquad \mathrm{S}_0 = 50, \qquad \mathrm{K} = 52.$$

With u, d and r as given, we find

$$\widetilde{p} = \frac{(1+r)-d}{u-d} \approx 0.6282, \qquad \widetilde{q} = 1-\widetilde{p} \approx 0.3718.$$

Next, let is find the value of S at each of the nodes. Using the notation of Figure 3.4, we have

$$\begin{aligned} \mathrm{S}_0 &= 50, & \mathrm{S}_1^{(1)} &= u\mathrm{S}_0 = 60, & \mathrm{S}_1^{(0)} &= d\mathrm{S}_0 = 40, \\ \mathrm{S}_2^{(2)} &= u^2\mathrm{S}_0 = 72, & \mathrm{S}_2^{(1)} &= ud\mathrm{S}_0 = 48, & \mathrm{S}_2^{(0)} &= d^2\mathrm{S}_0 = 32, \end{aligned}$$

Setting $\phi(s) = (\mathrm{K}-s)^+$, the value of exercising the option at each node is

$$\begin{aligned} \phi(\mathrm{S}_0) &= 2, & \phi(\mathrm{S}_1^{(1)}) &= \phi(u\mathrm{S}_0) = 0, & \phi(\mathrm{S}_1^{(0)}) &= \phi(d\mathrm{S}_0) = 12, \\ \phi(\mathrm{S}_2^{(2)}) &= \phi(u^2\mathrm{S}_0) = 0, & \phi(\mathrm{S}_2^{(1)}) &= \phi(ud\mathrm{S}_0) = 4, & \phi(\mathrm{S}_2^{(0)}) &= \phi(d^2\mathrm{S}_0) = 20, \end{aligned}$$

Now, let us figure out the value of the option at time $n = 1$. At the top node we have

$$\begin{aligned} \mathrm{V}_1^{(1)} &= \max\{\phi(u\mathrm{S}_0), \frac{1}{1+r}\widetilde{\mathbb{E}}[\phi(\mathrm{S}_2)|\mathrm{S}_1 = u\mathrm{S}_0 = 60]\} \\ &= \max\{\phi(u\mathrm{S}_0), \frac{1}{1+r}\left(\widetilde{p}\phi(\mathrm{S}_2^{(2)}) + \widetilde{q}\phi(\mathrm{S}_2^{(1)})\right)\} \\ &= \max\{0, 1.4147\} = 1.4147. \end{aligned}$$

So, if $\mathrm{S}_1 = u\mathrm{S}_0$, then it is better to hold the option to maturity (this should be obvious, since the payoff from exercises is zero). At the bottom node we have

$$\begin{aligned} \mathrm{V}_1^{(0)} &= \max\{\phi(d\mathrm{S}_0), \frac{1}{1+r}\widetilde{\mathbb{E}}[\phi(\mathrm{S}_2)|\mathrm{S}_1 = d\mathrm{S}_0 = 40]\} \\ &= \max\{\phi(d\mathrm{S}_0), \frac{1}{1+r}\left(\widetilde{p}\phi(\mathrm{S}_2^{(1)}) + \widetilde{q}\phi(\mathrm{S}_2^{(0)})\right)\} \\ &= \max\{12, 9.4636\} = 12. \end{aligned}$$

As the exercise value is greater than the value of holding the option, if $S_1 = dS_0$, then it is better to exercise the option. Finally, let us compute the value of the option at time zero. We have

$$\begin{aligned}
V_0 &= \max\{\phi(S_0), \frac{1}{1+r}\widetilde{\mathbb{E}}[V_1|S_0 = 50]\} \\
&= \max\{\phi(S_0), \frac{1}{1+r}\left(\tilde{p}V_1^{(1)} + \tilde{q}V_1^{(0)}\right)\} \\
&= \max\{2, 5.0894\} = 5.0894.
\end{aligned}$$

As the value of holding the option is greater than the value of exercising the option, it is best to hold the option at time $n = 0$.

We have shown how to value an American option using risk-neutral valuation. A natural question to ask is: how do we replicate the option? We would like for $X_n = V_n$ for every n. To accomplish this, just as in the European case, we try to set up a replicating portfolio. Suppose the value of the stock and replicating portfolio at time n are $S_n^{(i)}$ and $X_n^{(i)} = V_n^{(i)}$, respectively. If the portfolio contains $\Delta_n^{(i)}$ of the stock S, then we can write the value $X_n^{(i)}$ as follows

$$X_n^{(i)} = \Delta_n^{(i)} S_n^{(i)} + (V_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})\frac{1}{B_n}B_n.$$

At time $n+1$, that value of the portfolio becomes

$$X_{n+1} = \begin{cases} \Delta_n^{(i)} S_{n+1}^{(i+1)} + (V_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r), & \text{if } \omega_{n+1} = H, \\ \Delta_n^{(i)} S_{n+1}^{(i)} + (V_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r), & \text{if } \omega_{n+1} = T. \end{cases}$$

We would like to have

$$X_{n+1} = \begin{cases} V_{n+1}^{(i+1)}, & \text{if } \omega_{n+1} = H, \\ V_{n+1}^{(i)}, & \text{if } \omega_{n+1} = T. \end{cases}$$

In order for this to occur, the following two equations must be satisfied

$$\begin{aligned}
V_{n+1}^{(i+1)} &= \Delta_n^{(i)} S_{n+1}^{(i+1)} + (V_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r), \\
V_{n+1}^{(i)} &= \Delta_n^{(i)} S_{n+1}^{(i)} + (V_n^{(i)} - \Delta_n^{(i)} S_n^{(i)})(1+r).
\end{aligned}$$

Solving either one of these equations for $\Delta_n^{(i)}$, we obtain

$$\Delta_n^{(i)} = \frac{V_{n+1}^{(i+1)} - V_{n+1}^{(i)}}{S_{n+1}^{(i+1)} - S_{n+1}^{(i)}},$$

where $V_n * (i)$ denotes the value of the American option at time n assuming $S_n = S_n^{(i)}$.

Example 3.7. Let us compute the Δ_0 of the put option discussed in Example 3.6 above. We have

$$\Delta_0 = \frac{V_1^{(1)} - V_1^{(0)}}{S_1^{(1)} - S_1^{(0)}} = \frac{V_1^{(1)} - V_1^{(0)}}{uS_0 - dS_0} = \frac{1.4147 - 12}{60 - 40} = -0.5292.$$

Notice that Δ_0 is negative. The Δ at the other nodes can be computed in a similar fashion.

3.2.4 Calibration to data

Consider an derivative written on $S = (S_t)_{0 \leq t \leq T}$ that matures at time T. In order to value this derivative using the N-step binomial model, we need to relate model parameters u, d to data. To begin, let us define

$$\delta t := T/N, \qquad S_{t_n} \equiv S_n := S_{n\delta t}, \qquad n = 0, 2, \ldots, N-1, N,$$

where we have abused notation slightly by writing $S_{t_n} \equiv S_n$. Suppose, from time series data, we have an estimate for expected rate of return of a stock μ and the expected variance σ^2 per unit time δt. Then, under the real world measure $\mathbb{P}$ (because we observe S under $\mathbb{P}$ – *not* under $\widetilde{\mathbb{P}}$) we have

$$\mathbb{E}\left[\frac{S_{t_n+\delta t}}{S_{t_n}}\middle| S_{t_n}\right] = e^{\mu\delta t}, \qquad \mathbb{V}\left[\frac{S_{t_n+\delta t}}{S_{t_n}}\middle| S_{t_n}\right] = \sigma^2\delta t.$$

We want to create a binomial tree that has these same statistical properties. Thus, we have

$$\mathbb{E}\left[\frac{S_{t_n+\delta t}}{S_{t_n}}\middle| S_{t_n}\right] = \mathbb{E}\left[\frac{S_{n+1}}{S_n}\middle| S_n\right] = pu + qd = e^{\mu\delta t}.$$

Using $q + p = 1$ and solving the above equation for p, we obtain

$$p = \frac{e^{\mu\delta t} - d}{u - d}. \tag{3.2.10}$$

Next, we focus on the variance. We have

$$\begin{aligned}
\mathbb{V}\left[\frac{S_{t_n+\delta t}}{S_{t_n}}\middle| S_{t_n}\right] &= \mathbb{V}\left[\frac{S_{n+1}}{S_n}\middle| S_n\right] \\
&= \mathbb{E}\left[\left(\frac{S_{n+1}}{S_n}\right)^2\middle| S_n\right] - \left(\mathbb{E}\left[\frac{S_{n+1}}{S_n}\middle| S_n\right]\right)^2 \\
&= (pu^2 + qd^2) - (pu + dq)^2 = \sigma^2\delta t.
\end{aligned}$$

Using (3.2.10) and the fact that $p + q = 1$, we obtain

$$e^{\mu\delta t}(u + d) - ud - e^{2\mu\delta t} = \sigma^2\delta t, \tag{3.2.11}$$

We would like to find u and d so that (3.2.11) is satisfied. We guess that

$$u = \mathrm{e}^{\sigma\sqrt{\delta t}}, \qquad\qquad d = \mathrm{e}^{-\sigma\sqrt{\delta t}}. \tag{3.2.12}$$

Let us suppose that $\delta t \ll 1$. This is reasonable because, for any T we can make δt as small as we like by choosing N sufficiently large. Expanding the exponential terms, we have

$$\mathrm{e}^{\mu\delta t} = 1 + \mu\delta t + \mathcal{O}(\delta t^2), \tag{3.2.13}$$

$$u = 1 + \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t + \mathcal{O}(\delta t^{3/2}), \tag{3.2.14}$$

$$d = 1 - \sigma\sqrt{\delta t} + \frac{1}{2}\sigma^2\delta t + \mathcal{O}(\delta t^{3/2}). \tag{3.2.15}$$

Inserting (3.2.13), (3.2.14) and (3.2.15) into the left-hand side of (3.2.11) yields

$$\mathrm{e}^{\mu\delta t}(u + d) - ud - \mathrm{e}^{2\mu\delta t} = \sigma^2\delta t = \sigma^2\delta t + \mathcal{O}(\delta t^{3/2}).$$

Thus, by choosing p and q according to (3.2.10) and by choosing u and d according to (3.2.12), we have by construction that

$$\mathbb{E}\left[\frac{\mathrm{S}_{t_n+\delta t}}{\mathrm{S}_{t_n}}\middle|\mathrm{S}_{t_n}\right] = \mathrm{e}^{\mu\delta t}, \qquad\qquad \mathbb{V}\left[\frac{\mathrm{S}_{t_n+\delta t}}{\mathrm{S}_{t_n}}\middle|\mathrm{S}_{t_n}\right] = \sigma^2\delta t + \mathcal{O}(\delta t^{3/2}).$$

In other words, the N-step binomial tree has, to a high degree of accuracy, the correct statistical properties under $\mathbb{P}$.

We now consider the N-step binomial model under $\widetilde{\mathbb{P}}$ with u and d as chosen in (3.2.12). Recall that $\widetilde{p}$ and $\widetilde{q}$ are given by (3.1.2). It is interesting to note that

$$\mathbb{V}\left[\frac{\mathrm{S}_{n+1}}{\mathrm{S}_n}\middle|\mathrm{S}_n\right] = \widetilde{\mathbb{V}}\left[\frac{\mathrm{S}_{n+1}}{\mathrm{S}_n}\middle|\mathrm{S}_n\right].$$

In other words, variance does not change when moving between $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ — only the expected return changes.

3.3 Exercises

Exercise 3.1. Consider the one-step binomial model, depicted in Figure 3.6. Show that if

$$0 < d < u < 1 + r < \infty, \qquad \text{or} \qquad 0 < 1 + r < d < u < \infty,$$

then you can construct a portfolio X that satisfies $\mathrm{X}_0 = 0$ and $\mathrm{X}_1 > 0$.

Exercise 3.2. Consider the one-step binomial model, depicted in Figure 3.6. Suppose that, if you own a share of S, you will receive a dividend payment of $a\mathrm{S}_0$ at time T (and if you short the stock, you owe this dividend payment to the owner). Assume

$$0 < d + a < 1 + r < u + a < \infty.$$

(a) Construct a portfolio X (meaning find the initial value X_0 and the number of shares Δ_0 invested in S at time zero) so that $\mathrm{X}_\mathrm{T} = \phi(\mathrm{S}_\mathrm{T})$, where ϕ is an arbitrary payoff function. (b) Find a probability measure $\widetilde{\mathbb{P}}$ (meaning find $\widetilde{p}$ and $\widetilde{q}$) such that $\widetilde{\mathbb{E}}(\mathrm{X}_\mathrm{T}/\mathrm{B}_\mathrm{T}) = \mathrm{X}_0/\mathrm{B}_0$. (c) Using your answer to part (b), compute $\mathrm{B}_0\widetilde{\mathbb{E}}(\phi(\mathrm{S}_\mathrm{T})/\mathrm{B}_\mathrm{T})$. How does this compare to the value of X_0 you found in part (a)? (d) Is it true that $\widetilde{\mathbb{E}}(\mathrm{S}_\mathrm{T}/\mathrm{B}_\mathrm{T}) = \mathrm{S}_0/\mathrm{B}_0$? Try to explain why or why not.

Exercise 3.3. Consider the one-step binomial model, depicted in Figure 3.6. Assume no dividends are paid (i.e., $a = 0$). (a) In class, we introduced a two-step procedure to find the time-zero value V_0 of an option that pays $\phi(\mathrm{S}_\mathrm{T})$ at time T. Step 1: find $\widetilde{\mathbb{P}}$ (meaning find $\widetilde{p}$ and $\widetilde{q}$) such that $\widetilde{\mathbb{E}}(\mathrm{S}_\mathrm{T}/\mathrm{B}_\mathrm{T}) = \mathrm{S}_0/\mathrm{B}_0$. Step 2: Using your answer from Step 1, compute V_0 using $\mathrm{V}_0/\mathrm{B}_0 = \widetilde{\mathbb{E}}(\mathrm{V}_\mathrm{T}/\mathrm{B}_\mathrm{T})$ where $\mathrm{V}_\mathrm{T} = \phi(\mathrm{S}_\mathrm{T})$. Use this method to compute V_0. (b) We are now going to introduce a different (but similar) two-step procedure to find the time-zero value V_0 of an option that pays $\phi(\mathrm{S}_\mathrm{T})$ at time T. Step 1: find $\widehat{\mathbb{P}}$ (meaning find $\widehat{p}$ and $\widehat{q}$) such that $\widehat{\mathbb{E}}(\mathrm{B}_\mathrm{T}/\mathrm{S}_\mathrm{T}) = \mathrm{B}_0/\mathrm{S}_0$. Step 2: Using your answer from Step 1, compute V_0 using $\mathrm{V}_0/\mathrm{S}_0 = \widehat{\mathbb{E}}(\mathrm{V}_\mathrm{T}/\mathrm{S}_\mathrm{T})$ where $\mathrm{V}_\mathrm{T} = \phi(\mathrm{S}_\mathrm{T})$. Use this method to compute V_0. The expressions you obtain for V_0 in parts (a) and (b) should be the same. We will re-visit this when we introduce the concept of a *numéraire*.

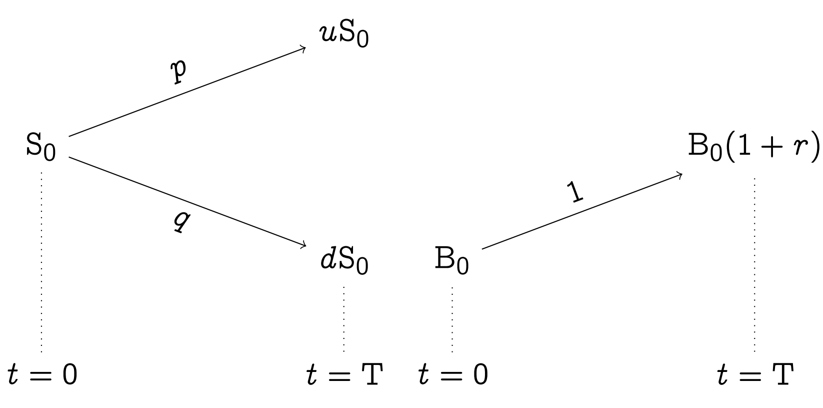

Figure 3.6: One-period binomial model

Please put your code for each exercise in a separate file. Your final submission should be a single ZIP file uploaded to Canvas.

Exercise 3.4. Consider the multi-period binomial model for a stock S and bond B. At each step we have

$$\mathrm{S}_{(n+1)\delta t} = \begin{cases} u\mathrm{S}_{n\delta t} & \text{with probability } p \\ d\mathrm{S}_{n\delta t} & \text{with probability } q \end{cases}, \qquad \mathrm{B}_{(n+1)\delta t} = \mathrm{e}^{r\delta t}\mathrm{B}_{n\delta t}.$$

We will set

$$\delta t = \mathrm{T}/\mathrm{N}, \qquad u = \mathrm{e}^{\sigma\sqrt{\delta t}}, \qquad d = \mathrm{e}^{-\sigma\sqrt{\delta t}}.$$

We call σ the *volatility* of S. Write a program to price and hedge a European put. Your inputs should be the initial stock price S_0, the maturity date T, the strike K, the risk free rate of interest r, the volatility σ, and the number of steps in your binomial tree N. Your outputs should be the value of the option (or replicating portfolio) at time zero X_0, and the number of shares of S you should have in the replicating portfolio at time zero Δ_0.

Your program should read the input from a text file called input.txt. The input file input.txt will consist of a single line listing S_0, T, K, r, σ, N in that order, separated by commas. An example input file will be posted to Canvas. Your program should write the solution to a text file called output.txt. Your output file should list X_0 and Δ_0 on a single line, separated by a comma. You may use either MATLAB, Python 2/Python 3, C++, or R.

Exercise 3.5. Modify your program from Exercise 3.4 to price and hedge an American put. Your inputs should be the initial stock price S_0, the maturity date T, the strike K, the risk free rate of interest r, the volatility σ, and the number of steps in your binomial tree N. Your outputs should be the value of the option (or replicating portfolio) at time zero X_0, and the number of shares of S you should have in the replicating portfolio at time zero Δ_0.

Your program should read the input from a text file called input.txt. The input file input.txt will consist of a single line listing S_0, T, K, r, σ, N in that order, separated by commas. An example input file will be posted to Canvas. Your program should write the solution to a text file called output.txt. Your output file should list X_0 and Δ_0 on a single line, separated by a comma. You may use either MATLAB, Python 2/Python 3, C++, or R.

Chapter 4

Fundamental Theorems of Asset Pricing

In this chapter, we will discuss the 1st and 2nd Fundamental Theorems of asset pricing.

4.1 1st Fundamental Theorem of Asset Pricing

Although we have often referred to arbitrage and the lack of it in this text, we have yet to give a mathematical definition of exactly what an arbitrage is. Below, we provide a precise mathematical definition of what we mean by an *arbitrage*.

Definition 4.1 (Arbitrage). An *arbitrage* is a portfolio X satisfying the following three points:

1. $X_0 = 0$ (starting from nothing),
2. $\mathbb{P}(X_T \geq 0) = 1$ (probability of losing money is zero),
3. $\mathbb{P}(X_T > 0) > 0$ (probability of making something is greater than zero),

for some $T > 0$.

Note that if X is an arbitrage under $\mathbb{P}$ then it is also an arbitrage under any equivalent measure $\widetilde{\mathbb{P}}$ because equivalent measures agree on events that occur with probability one.

Suppose that the no-arbitrage condition in the binomial model $d < 1 + r < u$ is not satisfied. Then we can show that there is an arbitrage. For example, assume $0 < 1 + r < d < u$. Let us construct a portfolio X that satisfies the three points in Definition 4.1 above. At time zero, we buy $\Delta_0 > 0$ shares of S and sell $\Delta_0 S_0 / B_0$ bonds. The value of this portfolio at time zero is the

$$X_0 = \Delta_0 S_0 - \frac{\Delta_0 S_0}{B_0} B_0 = 0.$$

At time $n = 1$, the value of the portfolio is

$$X_1 = \Delta_0 S_1 - \frac{\Delta_0 S_0}{B_0} B_1 \geq \Delta_0 d S_0 - \Delta_0 S_0 (1+r) = \Delta_0 S_0 (d - (1+r)) > 0.$$

As, by assumption, $1 + r < d$, we see that $\mathbb{P}(X_T \geq 0) = 1$ and $\mathbb{P}(X_t > 0) > 0$. As the three points in Definition 4.1 are fulfilled, we see that X is an arbitrage. The case $0 < d < u < 1 + r$ is treated similarly.

We also showed that, in the binomial model, we could price assets by changing to a probability measure $\widetilde{\mathbb{P}}$ under which X/B is a martingale for all portfolios X. In searching for $\widetilde{\mathbb{P}}$, we found that

$$\widetilde{\mathbb{E}}\left(\frac{X_1}{B_1}\right) = \frac{X_0}{B_0} \qquad \Leftrightarrow \qquad \widetilde{p} = \frac{(1+r) - d}{u - d}. \tag{4.1.1}$$

Note now that, if $0 < 1 + r < d < u$ and therefore there is an arbitrage, we have that from (4.1.1) that $\widetilde{p} < 0$. As probabilities cannot be negative, we conclude that the martingale measure $\widetilde{\mathbb{P}}$ does not exist! In fact, there is a connection between the existence of $\widetilde{\mathbb{P}}$ and the absence of arbitrage.

Theorem 4.2 (1st fundamental theorem of asset pricing). *A market defined under a probability measure $\mathbb{P}$ is free of arbitrage if and only if there exists a measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, under which all traded portfolios, denominated in units of a common numéraire N, are martingales.*

Note that Theorem 4.2 applies to *all* market models – not only the binomial model.

Theorem 4.2 gives us a very simple way to check if a market is free of arbitrage.

Step 1: Choose a numéraire asset N. (Typically, we choose the numéraire asset N to be the bond B, but we do not have to make this choice).

Step 2: Search for a probability measure $\widetilde{\mathbb{P}}$ under which, for *all* portfolios X, we have the martingale property for X/N

$$\frac{X_t}{N_t} = \widetilde{\mathbb{E}}\left[\frac{X_T}{N_T}\middle|\mathcal{F}_t\right]. \tag{4.1.2}$$

Here, the filtration $\mathcal{F}_t$ represents all of the information in the market that is available at time t. For example, in the binomial model, $\mathcal{F}_t = \mathcal{F}_t^S$. But, in a more complicated market $\mathcal{F}_t$ could include a the history of various stocks, economic indicators, interest rates, etc..

Again, by "search for a probability measure $\widetilde{\mathbb{P}}$" we mean "assign probabilities to events". For example, in the binomial model, in searching for $\widetilde{\mathbb{P}}$ we may assign the probabilities $\widetilde{p}$ and $\widetilde{q}$ in Figure 3.3. Note, we are *not* allowed to change u, d, r, etc., as these are not probabilities of events. If we cannot find a probability measure $\widetilde{\mathbb{P}}$ under which (4.1.2) holds, then the market contains an arbitrage.

Theorem 4.2 also gives us a way to price derivative assets in a way that guarantees we do not introduce arbitrage into the market. Suppose V_t is the value of a derivative asset with payoff $V_T = \phi(A_T)$. If we assign a value V_t using

$$\frac{V_t}{N_t} = \widetilde{\mathbb{E}}\left[\frac{V_T}{N_T}\middle|\mathcal{F}_t\right] = \widetilde{\mathbb{E}}\left[\frac{\phi(A_T)}{N_T}\middle|\mathcal{F}_t\right] \tag{4.1.3}$$

then we are guaranteed to not introduce arbitrage into the market.

Example 4.3 (Binomial model with dividends). Consider, the one-period binomial model. Suppose that, at time $n = 1$, the owner of one share of S receives dividend payment of aS_0. If we assume the market contains only a stock and a bond B, then the value of any portfolio X at time $n = 0$ is given by

$$X_0 = \Delta_0 S_0 + (X_0 - \Delta_0 S_0)\frac{1}{B_0}B_0,$$

and at time $n = 1$ the value of X becomes

$$X_1 = \Delta_0(S_1 + aS_0) + (X_0 - \Delta_0 S_0)\frac{1}{B_0}B_1$$

Let us choose the bond B_0 as numéraire (as we typically do) and search for a probability measure $\widetilde{\mathbb{P}}$ under which X/B is a martingale. We would like

$$\begin{aligned}\frac{X_0}{B_0} &= \widetilde{\mathbb{E}}\frac{X_1}{B_1}\\ &= \Delta_0\left(\widetilde{\mathbb{E}}\frac{S_1}{B_1} + a\frac{S_0}{B_1}\right) + (X_0 - \Delta_0 S_0)\frac{1}{B_0}\frac{B_1}{B_1}\\ &= \Delta_0\left(\widetilde{p}\frac{uS_0}{B_0(1+r)} + \widetilde{q}\frac{dS_0}{B_0(1+r)} + a\frac{S_0}{B_0(1+r)}\right) + (X_0 - \Delta_0 S_0)\frac{1}{B_0}.\end{aligned}$$

For the above equation to hold, we must have

$$\widetilde{p}u + \widetilde{q}d + a = 1 + r$$

Solving the above equation for $\widetilde{p}$ and using $\widetilde{p} + \widetilde{q} = 1$ we obtain

$$\widetilde{p} = \frac{(1 + r - a) - d}{u - d}, \qquad \widetilde{q} = 1 - \widetilde{p}. \tag{4.1.4}$$

What are the no-arbitrage conditions? In order for $\widetilde{\mathbb{P}}$ to exist, we must have $\widetilde{p}, \widetilde{q} \in (0, 1)$. Thus the no arbitrage condition is

$$d < 1 + r - a < u.$$

Now, assume the no-arbitrage conditions are satisfied. Suppose we wish to price a derivative asset that pays $V_1 = \phi(S_1)$ at time $n = 1$. In order for the market to be free of arbitrage, the initial value V_0 of this asset must satisfy

$$\frac{V_0}{B_0} = \widetilde{\mathbb{E}}\frac{V_1}{B_1} = \widetilde{\mathbb{E}}\frac{\phi(S_1)}{B_1} = \widetilde{p}\frac{\phi(uS_0)}{B_0(1+r)} + \widetilde{q}\frac{\phi(dS_0)}{B_0(1+r)},$$

where $\widetilde{p}$ and $\widetilde{q}$ are given by (4.1.4).

Example 4.4 (Forward contracts revisited). Let is find the fair forward price and the fair value of a forward contract after inception using our pricing formula (4.1.3). First, we choose as numéraire, a bond, whose value is given by

$$\mathrm{B}_t = \mathrm{B}_0 \mathrm{e}^{rt}$$

Let V_t be the value of a forward contract with payoff $\mathrm{S}_\mathrm{T} - \mathrm{K}$. We know that, in order for there to be no arbitrage in the market, there must be some probability measure $\widetilde{\mathbb{P}}$, under which $(\mathrm{V}_t/\mathrm{B}_t)$ is a martingale. Thus, for all $t \leq \mathrm{T}$ we have

$$\begin{aligned}
\frac{\mathrm{V}_t}{\mathrm{B}_t} &= \widetilde{\mathbb{E}}\left[\frac{\mathrm{V}_\mathrm{T}}{\mathrm{B}_\mathrm{T}}\middle|\mathcal{F}_t\right] = \widetilde{\mathbb{E}}\left[\frac{\mathrm{S}_\mathrm{T} - \mathrm{K}}{\mathrm{B}_\mathrm{T}}\middle|\mathcal{F}_t\right] \\
&= \widetilde{\mathbb{E}}\left[\frac{\mathrm{S}_\mathrm{T}}{\mathrm{B}_\mathrm{T}}\middle|\mathcal{F}_t\right] - \frac{\mathrm{K}}{\mathrm{B}_\mathrm{T}} = \frac{\mathrm{S}_t}{\mathrm{B}_t} - \frac{\mathrm{K}}{\mathrm{B}_\mathrm{T}}, \\
\mathrm{V}_t &= \mathrm{S}_t - \frac{\mathrm{K}\mathrm{B}_t}{\mathrm{B}_\mathrm{T}} = \mathrm{S}_t - \mathrm{K}\mathrm{e}^{-r(\mathrm{T}-t)}.
\end{aligned}$$

So, we have found the fair value of a forward contract with strike K at time t. Recall that the fair forward price F_t^T (not to be confused with the value of a forward contract) is the strike K that makes the value of the forward contract equal to zero. Setting $\mathrm{V}_t = 0$ in the last equation above, we have

$$0 = \mathrm{S}_t - \mathrm{K}\mathrm{e}^{-r(\mathrm{T}-t)} \qquad \Rightarrow \qquad \mathrm{K} = \mathrm{S}_t \mathrm{e}^{r(\mathrm{T}-t)} = \mathrm{F}_t^\mathrm{T}.$$

In particular, at time zero, we have $\mathrm{K} = \mathrm{e}^{r\mathrm{T}}\mathrm{S}_0$. Note that we did not need a model for S in order to determine the value of a forward contract or the fair forward price.

4.2 2nd Fundamental Theorem of Asset Pricing

In the one-period binomial model, the risk-neutral measure $\widetilde{\mathbb{P}}$ was unique; there was one and *only* one way to assign risk-neutral probabilities

$$\widetilde{p} = \frac{(1+r) - d}{u - d}, \qquad \widetilde{q} = 1 - \widetilde{p}.$$

However, for market models in general, when we search for a risk-neutral measure $\widetilde{\mathbb{P}}$, we may find that it is not unique. Let us see a simple example.

4.2.1 The trinomial model as an example

In the trinomial model, the stock can take one of three different paths at each node. This is illustrated in Figure 4.1. With reference to the numbers in the Figure 4.1, we make the following assumptions

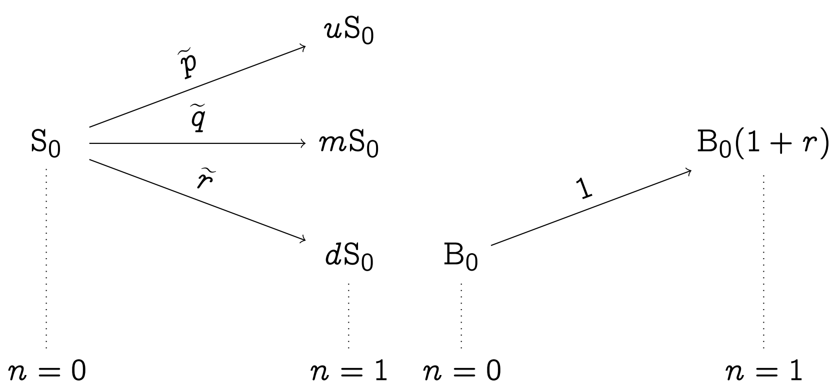

Figure 4.1: A one-step trinomial tree for a stock S and bond B under $\widetilde{\mathbb{P}}$.

$$0 < d < 1 + r < u, \qquad\qquad 0 < d < m < u.$$

We will now try to find a risk-neutral measure $\widetilde{\mathbb{P}}$ (i.e., we search for probabilities $\widetilde{p}, \widetilde{q}, \widetilde{r}$ such that X/B is a martingale under $\widetilde{\mathbb{P}}$). We would like

$$\frac{\mathrm{X}_0}{\mathrm{B}_0} = \widetilde{\mathbb{E}}\left[\frac{\mathrm{X}_1}{\mathrm{B}_1}\right] = \Delta_0 \frac{(\widetilde{p}u + \widetilde{q}m + \widetilde{r}d)\mathrm{S}_0}{\mathrm{B}_0(1+r)} + \frac{\mathrm{X}_0}{\mathrm{B}_0} - \frac{\Delta_0 \mathrm{S}_0}{\mathrm{B}_0},$$

where we have used $\mathrm{X}_1 = \Delta_0 \mathrm{S}_1 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r)$. Thus, we must have

$$\widetilde{p}u + \widetilde{q}m + \widetilde{r}d = 1 + r, \qquad\qquad \widetilde{p} + \widetilde{q} + \widetilde{r} = 1.$$

We have two equations and three unknowns. In general, we cannot solve this system uniquely for $\widetilde{p}, \widetilde{q}, \widetilde{r}$. However, if we fix $\widetilde{r}$, then we can solve for $\widetilde{p}(\widetilde{r})$ and $\widetilde{q}(\widetilde{r})$ as functions of $\widetilde{r}$. Doing this, we obtain

$$\widetilde{p}(\widetilde{r}) = \frac{(1+r) - m - \widetilde{r}d + m\widetilde{r}}{u - m}, \qquad\qquad \widetilde{q}(\widetilde{r}) = 1 - \widetilde{p}(\widetilde{r}) - \widetilde{r}. \tag{4.2.1}$$

Let us denote by $\widetilde{\mathbb{P}}(\widetilde{r})$ the risk-neutral measure in which $\widetilde{p}(\widetilde{r})$ and $\widetilde{q}(\widetilde{r})$ are given by (4.2.1). As $\widetilde{r}$ can take any value as long as $\widetilde{r}$, $\widetilde{p}(\widetilde{r})$ and $\widetilde{q}(\widetilde{r})$ are all in $(0,1)$ (so that $\widetilde{\mathbb{P}}(\widetilde{r}) \sim \mathbb{P}$) there are *infinitely many* probability measures $\widetilde{\mathbb{P}}(\widetilde{r})$!

Now, what happens if we try to replicate a European option with payoff $h(\mathrm{S}_\mathrm{T})$ in the trinomial model? At time $n = 0$ the value of a portfolio X is given by

$$\mathrm{X}_0 = \Delta_0 \mathrm{S}_0 + \frac{\mathrm{X}_0 - \Delta_0 \mathrm{S}_0}{\mathrm{B}_0} \mathrm{B}_0,$$

and at time $n = 1$ the value becomes

$$\mathrm{X}_1 = \Delta_0 \mathrm{S}_1 + \frac{\mathrm{X}_0 - \Delta_0 \mathrm{S}_0}{\mathrm{B}_0} \mathrm{B}_1$$

$$= \begin{cases} \Delta_0 u \mathrm{S}_0 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r) & \text{if } \mathrm{S}_1 = u\mathrm{S}_0, \\ \Delta_0 m \mathrm{S}_0 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r) & \text{if } \mathrm{S}_1 = m\mathrm{S}_0, \\ \Delta_0 d \mathrm{S}_0 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r) & \text{if } \mathrm{S}_1 = d\mathrm{S}_0. \end{cases}$$

At time $n = 1$ the value of the derivative asset is

$$h(\mathrm{S}_\mathrm{T}) = \begin{cases} h(u\mathrm{S}_0) & \text{if } \mathrm{S}_1 = u\mathrm{S}_0, \\ h(m\mathrm{S}_0) & \text{if } \mathrm{S}_1 = m\mathrm{S}_0, \\ h(d\mathrm{S}_0) & \text{if } \mathrm{S}_1 = d\mathrm{S}_0. \end{cases}$$

In order for $\mathrm{X}_1 = h(\mathrm{S}_1)$ we must have

$$\begin{aligned} h(u\mathrm{S}_0) &= \Delta_0 u \mathrm{S}_0 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r), \\ h(m\mathrm{S}_0) &= \Delta_0 m \mathrm{S}_0 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r), \\ h(d\mathrm{S}_0) &= \Delta_0 d \mathrm{S}_0 + (\mathrm{X}_0 - \Delta_0 \mathrm{S}_0)(1+r). \end{aligned}$$

We have three equations and two unknowns (Δ_0 and X_0). The above system of equations is over-determined. Unless we are very lucky, there will not be any combination of X_0 and Δ_0 for which the above three equations will be satisfied. Thus, we will not be able find a portfolio X that replicates the derivative payoff $h(\mathrm{S}_\mathrm{T})$.

There is a connection between the uniqueness of risk-neutral measures $\widetilde{\mathbb{P}}$ and the ability to perfectly hedge. Before describing this connection, let us first introduce the notion of a complete market.

Definition 4.5. A market that is free of arbitrage is said to be *complete* if every derivative asset can be perfectly replicated. If a market is not complete, we say it is *incomplete*.

The binomial model is an example of a complete market. The trinomial model is an example of an incomplete market.

Theorem 4.6 (Second fundamental Theorem of Asset Pricing). *A market, defined under a probability measure $\mathbb{P}$, which is free of arbitrage, is complete if and only if there exists a* <u>*unique*</u> *equivalent probability measure $\widetilde{\mathbb{P}}$ under which all portfolios, denominated in units of a common numéraire, are martingales.*

Theorem 4.6 applies to *all* market models – not just the binomial and trinomial models.

Theorem 4.6 gives us an easy way to check if a market is complete. We start with a market, defined under some probability measure $\mathbb{P}$. We then choose a numéraire asset N (typically, we choose the numéraire to be the bond $N = B$). We now search for a probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, under which *all* portfolios X are martingales

$$\frac{X_t}{N_t} = \widetilde{\mathbb{E}}\left[\frac{X_T}{N_T}|\mathcal{F}_t\right]. \tag{4.2.2}$$

If there is a unique way to choose $\widetilde{\mathbb{P}}$ (as there is in the binomial model), then the market is complete and every derivative asset can be hedged perfectly. If there are multiple ways to choose $\widetilde{\mathbb{P}}$ such that (4.2.2) is satisfied (as in the trinomial model), then the market is incomplete, and derivative assets cannot be hedged perfectly. This brings up the following question:

Question: How do we price derivative assets in incomplete markets?

In fact, there is no unique answer to this question. Below, using the trinomial model as an example, we will discuss four possible answers.

4.2.2 Answer 1: the Market chooses a pricing measure

Consider the trinomial model depicted in Figure 4.1. Suppose we observe from market quotes the price V_0 of a derivative asset with payoff $h(S_1)$. If the market is free of arbitrage, then we know that

$$\frac{V_0}{B_0} = \widetilde{\mathbb{E}}^{(\widetilde{r})}\left[\frac{V_1}{B_1}\right] = \widetilde{\mathbb{E}}^{(\widetilde{r})}\left[\frac{h(S_1)}{B_1}\right],$$

for some choice of $\widetilde{r}$. Solving for V_0, we obtain

$$V_0 = \frac{1}{1+r}\Big(\widetilde{p}(\widetilde{r})h(uS_0) + \widetilde{q}(\widetilde{r})h(mS_0) + \widetilde{r}h(dS_0)\Big), \tag{4.2.3}$$

where $\widetilde{p}(\widetilde{r})$ and $\widetilde{q}(\widetilde{r})$ are given by (4.2.1). Using (4.2.3) we can solve for $\widetilde{r}$ in terms of V_0. This defines uniquely the pricing measure $\widetilde{\mathbb{P}}^{(\widetilde{r})}$. In a sense, the market has *chosen* the pricing measure.

Suppose now, that a new derivative with payoff $g(S_1)$ is introduced to the market. If there is no arbitrage in the market, then the initial price P_0 of this asset *must* be given by

$$\frac{P_0}{B_0} = \widetilde{\mathbb{E}}^{(\widetilde{r})}\left[\frac{P_1}{B_1}\right] = \widetilde{\mathbb{E}}^{(\widetilde{r})}\left[\frac{g(S_1)}{B_1}\right].$$

Solving for P_0, we obtain

$$P_0 = \frac{1}{1+r}\Big(\widetilde{p}(\widetilde{r})g(uS_0) + \widetilde{q}(\widetilde{r})g(mS_0) + \widetilde{r}g(dS_0)\Big), \tag{4.2.4}$$

As we already know the value of $\widetilde{r}$ in terms of V_0, we have the value of P_0 explicitly.

The market (S, B) is *incomplete* because we cannot replicate the payoff $h(S_1)$ by trading only the stock

S and the bond B. However, The market (S, B, V) is *complete* because we can replicate the payoff $g(S_1)$ by trading the stock S, the bond B and the option V with payoff $h(S_1)$. To see this, observe that a hedging portfolio at time $n = 0$ is given by:

$$X_0 = \Delta_0 S_0 + \Gamma_0 V_0 + \frac{X_0 - \Delta_0 S_0 - \Gamma_0 V_0}{B_0} B_0.$$

At time $n = 1$, the portfolio's value becomes

$$\begin{aligned} X_1 &= \Delta_0 S_1 + \Gamma_0 V_1 + \frac{X_0 - \Delta_0 S_0 - \Gamma_0 V_0}{B_0} B_1 \\ &= \Delta_0 S_1 + \Gamma_0 h(S_1) + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0)(1 + r). \end{aligned}$$

In order for $X_1 = g(S_1)$ the following three equations must be satisfied

$$\begin{aligned} g(uS_0) &= \Delta_0 uS_0 + \Gamma_0 h(uS_0) + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0)(1 + r), \\ g(mS_0) &= \Delta_0 mS_0 + \Gamma_0 h(mS_0) + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0)(1 + r), \\ g(dS_0) &= \Delta_0 dS_0 + \Gamma_0 h(dS_0) + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0)(1 + r), \end{aligned}$$

With three equations, we can now solve for the three unknowns in our hedging portfolio: Δ_0, X_0 and Γ_0. If we do this, we will find that $X_0 = P_0$ where P_0 is given by (4.2.4).

Example 4.7. Consider the trinomial model with the following parameters

$$u = 1.3, \qquad m = 1.0, \qquad d = 0.7, \qquad r = 0, \qquad S_0 = 100, \qquad B_0 = 1.$$

What are the possible no-arbitrage values of a call option with strike $K_1 = 100$? As $r = 0$ we have that $B_T = B_0 = 1$. As such, all traded assets are martingales under some risk-neutral measure $\widetilde{\mathbb{P}}(\widetilde{r})$. Thus, we have

$$\begin{aligned} V_0(\widetilde{r}) &= \widetilde{\mathbb{E}}^{(\widetilde{r})}[V_1] = \widetilde{\mathbb{E}}^{(\widetilde{r})}[(S_1 - K_1)^+] \\ &= \widetilde{p}(\widetilde{r})(uS_0 - K_1)^+ + \widetilde{q}(\widetilde{r})(mS_0 - K_1)^+ + \widetilde{r}(dS_0 - K_1)^+ \end{aligned}$$

where $\widetilde{p}(\widetilde{r})$ and $\widetilde{q}(\widetilde{r})$ are given by (4.2.1). Note that we have indicated the dependence of V_0 on $\widetilde{r}$. With u, m and d as given, we have

$$\widetilde{p}(\widetilde{r}) = \widetilde{r}, \qquad \widetilde{q}(\widetilde{r}) = 1 - 2\widetilde{r}.$$

In order for $\widetilde{\mathbb{P}}^{(\widetilde{r})}$ to be an probability measure equivalent to $\mathbb{P}$, we need that $\widetilde{p}, \widetilde{r}, \widetilde{q} \in (0, 1)$. Thus, we must have $\widetilde{r} \in (0, 1/2)$. Noting that $(S_T - K_1)^+ \neq 0$ only when $S_T = uS_0$ we find that

$$V_0(\widetilde{r}) = \widetilde{p}(\widetilde{r})(uS_0 - K_1)^+ = \widetilde{p}(\widetilde{r}) \cdot 30.$$

As $\widetilde{p}(\widetilde{r}) = \widetilde{r}$ and $\widetilde{r} \in (0, 1/2)$ we find the following possible no-arbitrage values of the call option

$$0 < V_0 < 15.$$

Now, suppose we observe from market quotes that the price of a call option with strike $K_1 = 100$ is $V_0 = 10$. Let us find the price P_0 of a put option with strike $K_2 = 90$, and let us construct a portfolio that replicates the payoff of the put. To begin, we know that

$$\begin{aligned} V_0(\widetilde{r}) &= \widetilde{\mathbb{E}}^{(\widetilde{r})}[V_1] = \widetilde{\mathbb{E}}^{(\widetilde{r})}[(S_1 - K_1)^+] \\ &= \widetilde{p}(\widetilde{r})(uS_0 - K_1)^+ + \widetilde{q}(\widetilde{r})(mS_0 - K_1)^+ + \widetilde{r}(dS_0 - K_1)^+. \end{aligned} \tag{4.2.5}$$

Solving (4.2.5) for $\widetilde{r}$ we find

$$\widetilde{r} = 1/3, \qquad \text{and thus, from (4.2.1), we have} \qquad \widetilde{p} = 1/3, \qquad \widetilde{q} = 1/3.$$

To find the price P_0 of the put option with strike K_2 we compute

$$\begin{aligned} P_0 &= \widetilde{\mathbb{E}}^{(\widetilde{r})}[P_1] = \widetilde{\mathbb{E}}^{(\widetilde{r})}[(S_1 - K_2)^+] \\ &= \widetilde{p}(\widetilde{r})(K_2 - uS_0)^+ + \widetilde{q}(\widetilde{r})(K_2 - mS_0)^+ + \widetilde{r}(K_2 - dS_0)^+ \\ &= 20/3. \end{aligned}$$

Now, we want to construct a portfolio X consisting of (S, B, V) such that $X_1 = P_1 = (K_2 - S_1)^+$. At time zero we have

$$X_0 = \Delta_0 S_0 + \Gamma_0 V_0 + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0) B_0.$$

At time $n = 1$ the value of this portfolio is

$$\begin{aligned} X_1 &= \Delta_0 S_T + \Gamma_0 V_1 + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0) B_1 \\ &= \Delta_0 S_T + \Gamma_0 (S_1 - K_1)^+ + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0). \end{aligned}$$

Setting $X_1 = (K_2 - S_1)^+$ gives us three equations with three unknowns (X_0, Δ_0, Γ_0). We have

$$\begin{aligned} \Delta_0 uS_0 + \Gamma_0(uS_0 - K_1)^+ + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0) &= (K_2 - uS_0)^+, \\ \Delta_0 mS_0 + \Gamma_0(mS_0 - K_1)^+ + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0) &= (K_2 - mS_0)^+, \\ \Delta_0 dS_0 + \Gamma_0(dS_0 - K_1)^+ + (X_0 - \Delta_0 S_0 - \Gamma_0 V_0) &= (K_2 - dS_0)^+. \end{aligned}$$

Solving this system of equations, we obtain

$$X_0 = 20/3, \qquad \Delta_0 = -2/3, \qquad \Gamma_0 = 2/3.$$

Note that $X_0 = P_0$. This should not be surprising, as the price of any claim is equal to the price of the replicating portfolio in a complete market (recall, the market (S, V, B) *is* complete).

4.2.3 Answer 2: Super-hedging

Definition 4.8. Consider a derivative asset that pays $h(S_T)$ at time T. A portfolio strategy X is a *super-hedge* for $h(S_T)$ if $\mathbb{P}[X_T \geq h(S_T)] = 1$.

If the bank that sells a derivative with payoff $h(S_T)$ wishes to guarantee that it does not lose money, then it should charge at least the smallest initial value X_0 of super-hedging portfolio for the derivative. That is, the bank should charge at least the *minimum super-hedging price* $\overline{V}_0$, which is given mathematically by

$$\overline{V}_0 := \inf\{X_0 : X_T \geq h(S_T)\}. \tag{4.2.6}$$

How does one find $\overline{V}_0$? The following theorem addresses this question.

Theorem 4.9 (Cost of super-hedging). *Consider a market defined under a probability measure $\mathbb{P}$. The minimum super-hedging price $\overline{V}_0$ of a derivative that pays $h(S_T)$ at time T is given by*

$$\overline{V}_0 = \sup_{\widetilde{\mathbb{P}} \sim \mathbb{P}} B_0 \widetilde{\mathbb{E}}\Big(\frac{h(S_T)}{B_T}\Big),$$

where the supremum is taken over all risk-neutral measures $\widetilde{\mathbb{P}}$ that are equivalent to $\mathbb{P}$, with the bond B as numéraire.

We will not prove Theorem 4.9. Rather, we will simply comment on the intuition behind the theorem. We know from the first fundamental Theorem of asset pricing that, if $\widetilde{\mathbb{P}}$ is a risk-neutral measure equivalent to $\mathbb{P}$, then if we define V_0 via

$$V_0 = B_0 \widetilde{\mathbb{E}} \frac{h(S_T)}{B_T},$$

then V_0 is a no-arbitrage price for a derivative that pays $h(S_T)$. The range of no-arbitrage prices for the derivative that pays $h(S_T)$ is then given by

$$\{\text{no-arbitrage range}\} = \Big(\inf_{\widetilde{\mathbb{P}} \sim \mathbb{P}} B_0 \widetilde{\mathbb{E}} \frac{h(S_T)}{B_T}, \sup_{\widetilde{\mathbb{P}} \sim \mathbb{P}} B_0 \widetilde{\mathbb{E}} \frac{h(S_T)}{B_T}\Big),$$

where the $\inf$ and $\sup$ are taken over all risk-neutral measures $\widetilde{\mathbb{P}}$ equivalent to $\mathbb{P}$. If a bank sells a derivative for any amount $X_0 > \sup_{\widetilde{\mathbb{P}} \sim \mathbb{P}} B_0 \widetilde{\mathbb{E}}[h(S_T)/B_T]$ then there is an arbitrage in the market. In particular, there is a trading strategy that satisfies $\mathbb{P}(X_T - h(S_T) > 0) > 0$ and $\mathbb{P}(X_T - h(S_T) \geq 0) = 1$. Thus, we conclude the super-hedging price (4.2.6) must be given by the upper end of the no-arbitrage range, which is (4.9).

In the context of the trinomial model depicted in Figure 4.1, the minimum super-hedging price $\overline{V}_0$ of a derivative that pays $h(S_1)$ at time $n = 1$ is given by

$$\begin{aligned}\overline{V}_0 &= \sup_{\widetilde{\mathbb{P}}(\widetilde{r})\sim\mathbb{P}} \left\{ B_0 \widetilde{\mathbb{E}}^{(\widetilde{r})} \left[\frac{h(S_1)}{B_1} \right] \right\} \\ &= \sup_{\widetilde{r}} \left\{ \frac{1}{1+r} \Big(\widetilde{p}(\widetilde{r}) h(uS_0) + \widetilde{q}(\widetilde{r}) h(mS_0) + \widetilde{r} h(dS_0) \Big) \right\}. \end{aligned} \tag{4.2.7}$$

where $\widetilde{p}(\widetilde{r})$ and $\widetilde{q}(\widetilde{r})$ are given by (4.2.1). Note that the supremum is likely to occur at the end-point of the valid interval for $\widetilde{r}$, which is why we take a sup rather than a max.

Example 4.10. Consider the trinomial model with the following parameters

$$u = 1.3, \qquad m = 1.0, \qquad d = 0.7, \qquad r = 0, \qquad S_0 = 100, \qquad B_0 = 1.$$

Let us find the minimal cost to super-hedge a call option with strike $K_1 = 100$. From (4.2.7), we have

$$\begin{aligned}\overline{V}_0 &= \sup_{\widetilde{\mathbb{P}}(\widetilde{r})} \left\{ \widetilde{p}(\widetilde{r})(uS_0 - K_1)^+ + \widetilde{q}(\widetilde{r})(mS_0 - K_1)^+ + \widetilde{r}(dS_0 - K_1)^+ \right\} \\ &= \sup_{\widetilde{\mathbb{P}}(\widetilde{r})} \left\{ \widetilde{p}(\widetilde{r})(uS_0 - K_1)^+ \right\}, \end{aligned}$$

where we have used the fact that $(mS_0 - K_1)^+ = 0$ and $(dS_0 - K_1)^+ = 0$. With (u, m, d, r) as given we know from Example 4.7 that

$$\widetilde{p}(\widetilde{r}) = \widetilde{r}, \qquad \widetilde{r} \in (0, 1/2).$$

As such, the minimum super-hedging cost is

$$\overline{V}_0 = \sup_{\widetilde{\mathbb{P}}(\widetilde{r})} \left\{ \widetilde{p}(\widetilde{r})(uS_0 - K_1)^+ \right\} = \sup_{\widetilde{r}\in(0,1/2)} \{ \widetilde{r} \cdot 30 \} = 15.$$

Thus, if a bank sells the claim for $\overline{V}_0 = 15$ it is guaranteed to not lose money (with a proper hedging strategy).

4.2.4 Answer 3: Indifference pricing

According to standard economic theory, every individual has a *utility function* U that, roughly speaking, maps one's wealth to one's happiness. All utility functions are increasing, because more money makes an individual happier (again, according to economic theory). Additionally, all utility functions are concave, because the marginal happiness one gains from making an additional unit of currency goes down the

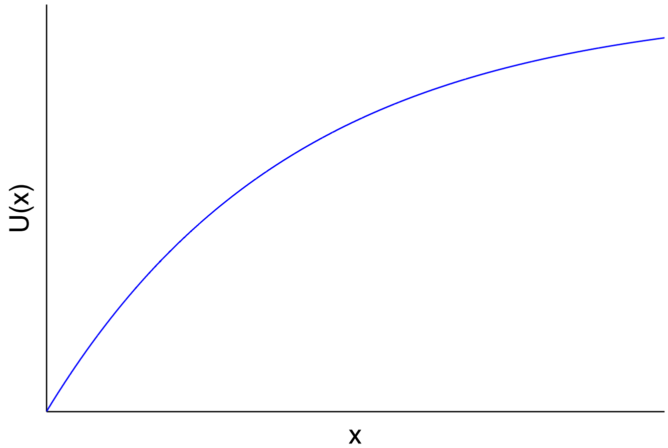

Figure 4.2: Example of a utility function U.

more money one already has. For example, increasing one's wealth from \$10 to \$20 makes one much happier than increasing one's wealth from \$10000 to \$10010. Figure 4.2 plots a generic utility function. An example of a utility function is $U(x) = -e^{-\gamma x}$.

The idea behind *indifference pricing* is that an investor will only buy an option if it increases his expected future utility. Let x be the initial wealth of an investor and consider a claim with payoff $h(S_T)$. We define an investor's *value function* by

$$v(x,k) := \sup_{\Delta} \mathbb{E}\left[U(X_T^{\Delta} + kh(S_T)|X_0^{\Delta} = x\right], \tag{4.2.8}$$

where the sup is taken over all investment strategies Δ. For example, in a one-period model, if an investor purchases Δ_0 shares of an asset S at time $n = 0$ and puts the rest of his money in a bond, then the final value of this portfolio will be

$$X_1^{\Delta} = \Delta_0 S_1 + (X_0^{\Delta} - \Delta_0 S_0)\frac{B_1}{B_0}.$$

Thus, an investment strategy corresponds merely to choosing a value for Δ_0. But, in an N-period model, an investment strategy Δ would involve choosing Δ_n for $n = 0, 1, \dots, N-1$. In a continuous time model, an investment strategy Δ would involving choosing Δ_t for all $t \in [0, T]$ where T is some target date in the future.

Note that the expectation in (4.2.8) is taken with respect to the real-world probability measure $\mathbb{P}$ – *not* a risk-neutral measure.

Definition 4.11 (Indifference price). The *indifference price* for k claims with payoff $h(S_T)$ as the unique solution p of the following equation

$$v(x,0) = v(x - kp, k). \tag{4.2.9}$$

where the function v is defined in (4.2.8).

The left-hand side of (4.2.9) is the maximum expected utility that an investor with initial wealth x can achieve by dynamically trading only in a stock S and a bond B. The right-hand side of (4.2.9) is the maximum expected utility that an investor with initial wealth x can achieve by purchasing k claims with payoff $h(S_T)$ for a unit price p at time zero (leaving him with $x - kp$), and dynamically trading the stock S and the bond B. Thus, equation (4.2.9) gives the price p of the claim that pays $h(S_T)$ such that the investor with initial wealth x can achieve the same expected utility either with or without buying the claim. In other words, the above equation gives the price p that makes the investor *indifferent* to the existence of the claim. Let us investigate how to solve (4.2.9) through an example.

Example 4.12. Consider the trinomial model with the following parameters

$$\begin{aligned} u &= 1.3, & m &= 1.1, & d &= 0.9, \\ \mathbb{P}[S_1 = uS_0] &= 1/3, & \mathbb{P}[S_1 = mS_0] &= 1/3, & \mathbb{P}[S_1 = dS_0] &= 1/3, \\ S_0 &= 100, & B_0 &= 100, & B_1 &= 110. \end{aligned}$$

Assume the utility function of an investor is given by $U(x) = -\exp(-x/10)$. Let us find the indifference price for one Call option with payoff $h(S_1) = (S_1 - K)^+$ where $K = 110$ for an investor with initial wealth x. First we find the value function of the investor assuming he cannot invest in the call option. We have

$$v(x,0) = \sup_{\Delta_0} \mathbb{E}[U(X_1^\Delta) | X_0^\Delta = x], \qquad X_1^\Delta = \Delta_0 S_1 + (x - \Delta_0 S_0)\frac{B_1}{B_0}.$$

Computing the expectation and simplifying, we obtain

$$\begin{aligned} v(x,0) &= \sup_{\Delta_0} \frac{-e^{-1.1x/10}}{3} \left(e^{-20\Delta_0/10} + 1 + e^{20\Delta_0/10} \right) \\ &= \sup_{\Delta_0} \frac{-e^{-1.1x/10}}{3} \left(2\cosh(20\Delta_0/10) + 1 \right). \end{aligned}$$

To find the optimal Δ_0, denoted Δ_0^*, we differentiate the argument of the $\sup$ and set it equal to zero. We have

$$0 = \frac{\partial}{\partial \Delta_0} \frac{-e^{-1.1x/10}}{3} \left(2\cosh(20\Delta_0/10) + 1 \right) \Big|_{\Delta_0 = \Delta_0^*}$$

$$= \frac{-\mathrm{e}^{-1.1x/10}}{3} \sinh(20\Delta_0^*/10) \qquad \Rightarrow \qquad \Delta_0^* = 0.$$

Thus, we have

$$v(x,0) = \frac{-\mathrm{e}^{-1.1x/10}}{3} \left(2\cosh(20\Delta_0/10) + 1\right)\Big|_{\Delta_0 = \Delta_0^*} = -\mathrm{e}^{-1.1x/10}.$$

Next, we compute the value function of the investor assuming he purchases one call option for price p. We have

$$v(x-p,1) = \sup_{\Delta_0} \mathbb{E}[\mathrm{U}(\mathrm{X}_1^\Delta + h(\mathrm{S}_1))|\mathrm{X}_0^\Delta = x-p], \qquad \mathrm{X}_1^\Delta = \Delta_0 \mathrm{S}_1 + (x - p - \Delta_0 \mathrm{S}_0)\frac{\mathrm{B}_1}{\mathrm{B}_0}.$$

Computing the expectation and simplifying, we obtain

$$v(x-p,1) = \sup_{\Delta_0} \frac{-\mathrm{e}^{-1.1(x-p)/10}}{3} \left(\mathrm{e}^{20\Delta_0/10} + 1 + \mathrm{e}^{-20(\Delta_0+1)/10}\right).$$

To find the optimal Δ_0, once again denoted Δ_0^*, we differentiate the argument of the sup and set it equal to zero. We have

$$\begin{aligned}
0 &= \frac{\partial}{\partial \Delta_0} \frac{-e^{-1.1(x-p)/10}}{3} \left(\mathrm{e}^{20\Delta_0/10} + 1 + \mathrm{e}^{-20(\Delta_0+1)/10}\right)\Big|_{\Delta_0 = \Delta_0^*} \\
&= \frac{-\mathrm{e}^{-1.1(x-p)/10}}{3} \frac{20}{10} \left(\mathrm{e}^{20\Delta_0^*/10} - \mathrm{e}^{-20(\Delta_0^*+1)/10}\right) \qquad \Rightarrow \qquad \Delta_0^* = -1/2.
\end{aligned}$$

Thus, we have

$$v(x-p,1) = \frac{-\mathrm{e}^{-1.1(x-p)/10}}{3} \left(\mathrm{e}^{20\Delta_0/10} + 1 + \mathrm{e}^{-20(\Delta_0+1)/10}\right)\Big|_{\Delta_0 = \Delta_0^*} = \frac{-\mathrm{e}^{-1.1(x-p)/10}}{3}(1 + 2/\mathrm{e}).$$

Finally, to find the indifference price we set $v(x,0) = v(x-p,1)$ and solve for p. We have

$$-\mathrm{e}^{-1.1x/10} = \frac{-\mathrm{e}^{-1.1(x-p)/10}}{3}(1+2/\mathrm{e}) \qquad \Rightarrow \qquad p = (10/1.1)\log\left(\tfrac{3}{1+2/\mathrm{e}}\right) \approx 4.97.$$

This examples was taken directly from: https://en.wikipedia.org/wiki/Indifference_price#Example.

A few final remarks about indifference prices. First, the indifference prices for any claim is guaranteed to exist. In particular, if one is in a complete market, the indifference price is equal to the unique no-arbitrage price. Second, indifference pricing is a *nonlinear* pricing mechanism. In other words, if $p(k)$ is the price per claim for k claims that each pay $h(\mathrm{S}_\mathrm{T})$ at time T, we have that that $p(k) \neq p(k')$ when $k \neq k'$. In other words, the indifference price for k claims is not equal (in general) to k times the price of one claim. Lastly, indifference prices do not introduce arbitrage into the market.

4.2.5 Answer 4: Minimum-variance hedging

That last method (among those we will discuss) for pricing derivatives in incomplete markets is to charge the initial value x^* of the portfolio X^{Δ^*} that minimizes the variance (under the real-world measure $\mathbb{P}$) of X_T^Δ minus the derivative payoff $h(S_T)$. That is, the bank charges x^* for the derivative that pays $h(S_T)$ where

$$(x^*, \Delta^*) = \underset{x,\Delta}{\operatorname{argmin}}\, \mathbb{V}(X_T^\Delta - h(S_T)|X_0^\Delta = x\Big),$$

Note that the maximum is taken over all values of initial wealth x and all investment strategies Δ.

In a one-period model, such as the trainman model, the investment strategy Δ corresponds to choosing only Δ_0. Thus, we can find x^* and Δ^* rather easily. Define

$$J(x, \Delta_0) = \mathbb{V}\Big(X_1^\Delta - h(S_T)|X_0^\Delta = x\Big), \qquad X_1^\Delta = \Delta_0 S_1 + (x - \Delta_0 S_0)\frac{B_1}{B_0}.$$

To find the optimal values of x and Δ_0, simply solve the following equations for (x^*, Δ_0^*)

$$\left.\frac{\partial J(x, \Delta_0)}{\partial x}\right|_{(x,\Delta_0)=(x^*,\Delta_0^*)} = 0, \qquad \left.\frac{\partial J(x, \Delta_0)}{\partial \Delta_0}\right|_{(x,\Delta_0)=(x^*,\Delta_0^*)} = 0.$$

The down-side of this approach is that the resulting derivative price x^* may be arbitrageable.

4.3 Exercises

Exercise 4.1. Consider the following model for a stock market with a stock S and a bond B. At time zero, the stock price is worth S_0 and the bond is worth B_0. At time T, under the real world probability measure $\mathbb{P}$ we have

$$\mathbb{P}(S_T = uS_0) = p, \qquad \mathbb{P}(S_T = dS_0) = 1 - p, \qquad \mathbb{P}(B_T = RB_0) = q, \qquad \mathbb{P}(B_T = rB_0) = 1 - q.$$

Here, we assume that the stock and bond price move *independently*, meaning, for example, that

$$\mathbb{P}(S_T = uS_0, B_T = RB_0) = \mathbb{P}(S_T = uS_0)\mathbb{P}(B_T = RB_0),$$

and likewise for other possible outcomes. We shall assume that $0 < d < r < R < u < \infty$ and that p and q are strictly positive.

(a) Find $\widetilde{p}$ as a function of $\widetilde{q}$ so that $\widetilde{\mathbb{E}}^{(\widetilde{q})}(S_T/B_T) = S_0/B_0$. Clearly show what equation you are solving in order to find $\widetilde{p}(\widetilde{q})$. Note that events that are independent under $\mathbb{P}$ should remain independent under $\widetilde{\mathbb{P}}(\widetilde{q})$.

(b) Is the market (S, B) complete or incomplete? Explain your answer.

(c) Let $u = 12/10$, $R = 11/10$, $r = 1$, $d = 8/10$, $S_0 = 100$ and $B_0 = 1$. What are upper and lower bounds for the price of a call option with payoff $(S_T - 100)^+$?

(d) Suppose a call the call option with payoff $h(S_T) = (S_T - 100)^+$ is selling for $V_0 = 12$. What is value X_0 of a put option that pays $g(S_T) = (90 - S_T)^+$? For part (d) you should find the value X_0 of the put by computing an expectation off the form $\widetilde{\mathbb{E}}^{(\widetilde{q})}(g(S_T)/B_T) = X_0/B_0$. Clearly write the terms involved in computing the expectation and state the value you are using for $\widetilde{q}$.

(e) Suppose you want to replicate the put option that pays $g(S_T)$ by investing in the Stock S, the bond B and the the call option that pays $h(S_T) = (S_T - 100)^+$ (which, as in part (d) has initial value $V_0 = 12$). What is the initial value of the portfolio X_0, the number of shares Δ_0 you should put in S_0 and the number of shares Γ_0 you should invest in the call option?

Exercise 4.2. Consider the following model for a stock market with a stock S and a bond B. At time zero, the stock price is worth S_0 and the bond is worth B_0. At time T, under the real world probability measure $\mathbb{P}$ we have

$$\mathbb{P}(S_T = uS_0) = p, \qquad \mathbb{P}(S_T = dS_0) = q, \qquad \mathbb{P}(B_T = RB_0) = p, \qquad \mathbb{P}(B_T = rB_0) = q.$$

Here, we assume that the stock and bond price move *together*, meaning that

$$S_T = uS_0 \quad \Leftrightarrow \quad B_T = RB_0,$$

(a) Find a probability measure $\widetilde{\mathbb{P}}$ (meaning assign probabilities to $\widetilde{p}$ and $\widetilde{q}$) such that S/B is a martingale.

(b) Is the market (S, B) complete? Explain your answer.

(c) Let $u = 12/10$, $R = 11/10$, $r = 1$, $d = 8/10$, $S_0 = 100$ and $B_0 = 1$. What are upper and lower bounds for the price of a call option with payoff $(S_T - 100)^+$? If the upper and lower bounds are the same, explain why this is the case.

(d) State the definition of an *arbitrage*. Show that if the price V_0 falls above the upper bound for the call price (which you obtain in part (c)), then there is an arbitrage. State exactly what the arbitrage strategy is (i.e., state what should you be buying and selling at time zero and show that this strategy gives you a guaranteed profit.)?

Chapter 5

Brownian motion, Itô processes and SDEs

5.1 Brownian motion

Louis Bachelier is credited as being the first person to study the process that we call *Brownian motion*. He used the process as a model for stock price movements in his thesis *The Theory of Speculation*, which was published in 1900. Brownian motion is also referred to as a *Weiner process* in honer of Norbert Wiener, who proved a number of important mathematical properties of Brownian motion.

Definition 5.1 (Brownian Motion). A *Brownian motion* or *Wiener Process*, defined under a probability measure $\mathbb{P}$, is a process $\mathrm{W} = (\mathrm{W}_t)_{t \geq 0}$ that satisfies the following

1. The process starts from zero: $\mathrm{W}_0 = 0$.

2. Non-overlapping increments are independent: $\mathrm{W}_\mathrm{T} - \mathrm{W}_t \perp\!\!\!\perp \mathrm{W}_u - \mathrm{W}_s$ for all $0 \leq s \leq u \leq t \leq \mathrm{T}$.

3. Increments are stationary and normally distributed: $\mathrm{W}_\mathrm{T} - \mathrm{W}_t \sim \mathrm{W}_{\mathrm{T}-t} \sim \mathrm{N}(0, \mathrm{T} - t)$.

4. The process W has continuous sample paths.

Recall that

$$\mathrm{Z} \sim \mathrm{N}(\mu, \sigma^2) \qquad \Rightarrow \qquad \mathbb{P}(\mathrm{Z} \in \mathrm{d}z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right] \mathrm{d}z,$$

Before discussing some important properties of Brownian motion, we need to introduce some important concepts.

Definition 5.2. We say that a random process $\Delta = (\Delta_t)_{t \geq 0}$ is *adapted* to a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$, if, for all $t \geq 0$, the information in $\mathcal{F}_t$ is sufficient to determine the value of Δ_t. In this case, we write $\Delta_t \in \mathcal{F}_t$.

Some examples should help clarify the concept of adaptedness.

Example 5.3. Suppose $\Delta_t = f(X_t)$ for some function f. Then we clearly have $\Delta_t \in \mathcal{F}_t^X$ because, by observing $\{X_s, 0 \leq s \leq t\}$ we can determine $\Delta_t = f(X_t)$ uniquely.

Example 5.4. Suppose $\Delta_t = \max_{s \in [0,t]} Y_s$. Then we clearly have $\Delta_t \in \mathcal{F}_t^Y$ because, by observing $\{Y_s, 0 \leq s \leq t\}$ we can determine $\Delta_t = \max_{s \in [0,t]} Y_s$.

Definition 5.5 (Filtration for W). We say that $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ is a *filtration for* W if (i) W is adapted to $\mathcal{F}$ and (ii) for all $0 \leq s \leq t \leq T$ we have that $W_T - W_t \perp\!\!\!\perp \mathcal{F}_s$.

Roughly speaking, $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ is a filtration for W if the information in $\mathcal{F}_t$ is sufficient to determine $\{W_s, 0 \leq s \leq t\}$ and if the information in $\mathcal{F}_t$ gives us no information about $W_T - W_t$ for any $T \geq t$. Again, some example will help clarify things a bit.

Example 5.6. The natural filtration $\mathcal{F}^W = (\mathcal{F}_t^W)_{t \geq 0}$ of a Brownian motion W, is a filtration for W because (i) the information in $\mathcal{F}_t^W = \{W_s, 0 \leq s \leq t\}$ is sufficient to determine W_t (i.e., W is adapted to $\mathcal{F}^W$) and (ii) the increment $W_T - W_t$ is independent of $\mathcal{F}_t^W = \{W_s, 0 \leq s \leq t\}$.

Example 5.7. Suppose $\mathcal{F}^{\pm} = (\mathcal{F}_t^{\pm})_{t \geq 0}$ is defined by $\mathcal{F}_t^{\pm} = \{W_s, 0 \leq s \leq t \pm \delta\}$ where $\delta > 0$ is a constant. Then $\mathcal{F}_t^+$ is not a filtration for W because $W_{t+\delta} - W_t$ is *not* independent of $\mathcal{F}_t^+ = \{W_s, 0 \leq s \leq t + \delta\}$. Likewise, $\mathcal{F}_t^-$ is *not* a filtration for W because the information in $\mathcal{F}_t^- = \{W_s, 0 \leq s \leq t - \delta\}$ is not sufficient to determine W_t (thus, W is not adapted to $\mathcal{F}^-$).

Throughout this chapter, $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ will always represent a filtration for W.

With the important technicalities out of the way, let discuss two important properties of Brownian motion.

1. Brownian motion is a Markov process. That is, for any function g and for any $0 \leq t \leq T < \infty$ we have $\mathbb{E}[g(W_T)|\mathcal{F}_t^W] = \mathbb{E}[g(W_T)|W_t]$.
2. Brownian motion is a martingale under $\mathbb{P}$ with respect to $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$. That is, for any $0 \leq t \leq T < \infty$ we have $\mathbb{E}[W_T|\mathcal{F}_t] = \mathbb{E}[W_T|W_t]$.

We can easily derive the martingale property as follows

$$\mathbb{E}[W_T|\mathcal{F}_t] = \mathbb{E}[W_T|W_t] = \mathbb{E}[(W_T - W_t) + W_t|W_t] = \mathbb{E}[W_T - W_t] + W_t = W_t.$$

The third equality follows from the fact that $W_T - W_t \perp\!\!\!\perp W_t$. The fourth equality follows from the fact that $W_T - W_t \sim N(0, T - t)$ and thus $\mathbb{E}[W_T - W_t] = 0$.

Example 5.8. Let $X = (X_t)_{t \geq 0}$ be defined by $X_t := W_t^2 - t$. Let us show that X is a martingale under $\mathbb{P}$ with respect to $\mathcal{F}^W$. We have

$$\begin{aligned}
\mathbb{E}[X_T|\mathcal{F}_t] &= \mathbb{E}[W_T^2 - T|W_t] \\
&= \mathbb{E}[(W_T - W_t)^2 + 2W_T W_t - W_t^2 - T|W_t] \\
&= \mathbb{E}[(W_T - W_t)^2|W_t] + 2W_t\mathbb{E}[W_T|W_t] - W_t^2 - T \\
&= T - t + 2W_t W_t - W_t^2 - T \\
&= W_t^2 - t = X_t.
\end{aligned}$$

Make sure you understand how each equality follows from the properties of Brownian motion!

Example 5.9 (Generalized Brownian motion). The *Generalized Brownian motion* is a process $X = (X_t)_{t \geq 0}$ that is given by

$$X_t := \mu t + \sigma W_t,$$

for some constants $\mu \in \mathbb{R}$ and $\sigma > 0$. Note that X_t is normally distributed because

$$Z \sim N(0,1) \qquad \Rightarrow \qquad a + bZ \sim N(a, b^2),$$

and therefore

$$W_t \sim N(0,t) \qquad \Rightarrow \qquad W_t \sim \sqrt{t}Z \qquad \Rightarrow \qquad X_t = \mu t + \sigma\sqrt{t}Z \sim N(\mu t, \sigma^2 t).$$

Note that we can compute the mean and variance of X_t directly. We have

$$\begin{aligned}
\mathbb{E}[X_t] &= \mu t + \sigma\mathbb{E}[W_t] = \mu t, \\
\mathbb{V}[X_t] &= \mathbb{E}[(X_t - \mathbb{E}[X_t])^2] = \sigma^2\mathbb{E}[W_t^2] = \sigma^2 t.
\end{aligned}$$

5.2 Itô integrals and Itô processes

Brownian motion can be used as a building block to create more general processes known as Itô integrals. Presumably, you are familiar with a Riemann integral

$$\begin{aligned}
\int_0^T f(t)\mathrm{d}t &:= \lim_{\delta t \to 0} \sum_{j=0}^{n-1} f(t_j^*)(t_{j+1} - t_j), & t_j^* &\in [t_j, t_{j+1}], \\
\delta t &:= \max_j |t_{j+1} - t_j|, & 0 &= t_0 < t_1 < \ldots < t_n = T. \qquad (5.2.1)
\end{aligned}$$

where the particular choice of t_j^* does not matter. You may be somewhat less familiar with the Riemann-Stieltjes integral

$$\int_0^T f(t)\mathrm{d}g(t) := \lim_{\delta t \to 0} \sum_{j=0}^{n-1} f(t_j^*)(g(t_{j+1}) - g(t_j)), \qquad t_j^* \in [t_j, t_{j+1}].$$

where, as long as g is piecewise differentiable, the particular choice of t_j^* does not matter. Similar to the Riemann-Stieltjes integral, we can define an *Itô integral*.

Definition 5.10 (Itô Integral). Suppose that $\sigma = (\sigma_t)_{t\geq 0}$ is a adapted to a filtration $\mathcal{F}$ of a Brownian motion W. We define

$$\int_0^T \sigma_t \mathrm{d}W_t := \lim_{\delta t \to 0} \sum_{j=0}^{n-1} \sigma_{t_j}(W_{t_{j+1}} - W_{t_j}),$$

where δt is as described in (5.2.1).

Believe it or not, it is actually important that we have written σ_{t_j} in and *not* $\sigma_{t_j^*}$ for some $t_j^* \in (t_j, t_{j+1}]$. The reason, without getting too much in the weeds, is that paths of Brownian motion W are not differentiable (*anywhere*!).

It is not terribly important that you remember the definition of an Itô integral. But, it is important that you remember a few properties. In what follows, let $J = (J_t)_{t\geq 0}$ be the following Itô integral

$$J_t := \int_0^t \sigma_s \mathrm{d}W_s,$$

and let $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ be a filtration for W. Then the following holds:

1. The process J is adapted to $\mathcal{F}$. That is, $J_t \in \mathcal{F}_t$ for all $t \geq 0$.
2. The process J is a martingale with respect to $\mathcal{F}$ (i.e., $\mathbb{E}[J_T|\mathcal{F}_t] = J_t$ for any $0 \leq t \leq T < \infty$).
3. The *Itô isometry* holds: $\mathbb{E}J_t^2 = \mathbb{E}\int_0^T \sigma_t^2 \mathrm{d}t$.
4. If the process $\sigma = (\sigma_t)_{t\geq 0}$ is a deterministic (i.e., not random) function $\sigma_t = \sigma(t)$, then the Itô integral is normally distributed $J_T \sim N(0, \Sigma^2(T))$ where $\Sigma^2(T) = \int_0^T \sigma^2(t)\mathrm{d}t$.

Having defined Itô integrals and discussed some of their properties, we are now in a position to describe what we mean by an Itô process.

Definition 5.11 (Itô Prosess). An *Itô process* is a process of the form

$$X_T = X_0 + \int_0^T \mu_t \mathrm{d}t + \int_0^T \sigma_t \mathrm{d}W_t, \tag{5.2.2}$$

where $\mu = (\mu_t)_{t\geq 0}$ and $\sigma = (\sigma_t)_{t\geq 0}$ are adapted to a filtration $\mathcal{F}$ for a Brownian motion W.

Very often, we will write the process X in differential form:

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t. \tag{5.2.3}$$

Equation (5.2.3) gives us an intuitive idea of how X behaves in a small increment $\mathrm{d}t$ of time. Roughly speaking, we have

$$X_{t+\mathrm{d}t} - X_t \approx \mu_t \mathrm{d}t + \sigma_t \underbrace{(W_{t+\mathrm{d}t} - W_t)}_{\sim N(0,\mathrm{d}t)}.$$

But, keep in mind that, when we write "$\mathrm{d}X_t = \ldots$" as we have in (5.2.3), what we really mean is that X_T is given by (5.2.2).

It will be important in subsequent chapters for us to be able to identify when an Itô process is a martingale. We know that

$$\mathrm{d}X_t = \sigma_t \mathrm{d}W_t \qquad \Rightarrow \qquad X \text{ is a martingale}, \tag{5.2.4}$$

because Itô integrals are martingales. Equation (5.2.4) gives us a *sufficient condition* for any Itô process X to be a martingale. Although we will not prove it, it turns out that (5.2.4) is also a *necessary condition* for X to be a martingale. In other words, we have

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t, \qquad \mu \neq 0 \qquad \Rightarrow \qquad X \text{ is a not martingale}.$$

For an Itô process X to be a martingale, the "$\mathrm{d}t$"-term *must* be zero. To see why this is the case, note that

$$\mathbb{E}(X_{t+\mathrm{d}t}|\mathcal{F}_t) \approx X_t + \mu_t \mathrm{d}t + \sigma_t \underbrace{\mathbb{E}(W_{t+\mathrm{d}t} - W_t|\mathcal{F}_t)}_{=0} = X_t + \mu_t \mathrm{d}t.$$

If $\mu_t \neq 0$ then $\mathbb{E}(X_{t+\mathrm{d}t}|\mathcal{F}_t) \neq X_t$ and thus X is not a martingale.

5.3 Itô's Lemma

Suppose $X = (X_t)_{t\geq 0}$ is an Itô process that, in differential form, is given by (5.2.3). We can define another process $f(X) = (f(X_t))_{t\geq 0}$. We expect that $\mathrm{d}f(X_t)$ is something like

$$\mathrm{d}f(X_t) = g_t \mathrm{d}t + h_t \mathrm{d}W_t.$$

for some processes $g = (g_t)_{t\geq 0}$ and $h = (h_t)_{t\geq 0}$. *Itô's Lemma*, presented below, tells us how to find the processes g and h.

Consider first a deterministic (i.e., not random) process that satisfies the following Ordinary Differential Equation (ODE)

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \mu(t), \qquad \text{or, in differential notation,} \qquad \mathrm{d}x(t) = \mu(t)\mathrm{d}t.$$

Then, we have by chain rule that

$$\mathrm{d}f(x(t)) = f'(x(t))\mathrm{d}x(t) = f'(x(t))\mu(t)\mathrm{d}t.$$

The above computation *suggests* that, when X is given by (5.2.3), we have

$$\mathrm{d}f(\mathrm{X}_t) \stackrel{?}{=} f'(\mathrm{X}_t)\mathrm{dX}_t \stackrel{?}{=} f'(\mathrm{X}_t)\Big(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t\Big). \tag{5.3.1}$$

Unfortunately, this is not quite correct. It turns out that we are missing a term of order $\mathrm{d}t$ in (5.3.1). To understand why this is the case, consider what happens to the process $f(\mathrm{X})$ in a infinitesimally small increment of time. We have

$$\begin{aligned}
\mathrm{d}f(\mathrm{X}_t) &= f(\mathrm{X}_t + \mathrm{dX}_t) - f(\mathrm{X}_t) \\
&= f(\mathrm{X}_t) + f'(\mathrm{X}_t)\mathrm{dX}_t + \tfrac{1}{2}f''(\mathrm{X}_t)(\mathrm{dX}_t)^2 + \ldots - f(\mathrm{X}_t) \\
&= f'(\mathrm{X}_t)(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t) + \tfrac{1}{2}f''(\mathrm{X}_t)(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t)^2 + \ldots \\
&= f'(\mathrm{X}_t)(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t) + \tfrac{1}{2}f''(\mathrm{X}_t)(\mu_t^2(\mathrm{d}t)^2 + 2\mu_t\sigma_t\mathrm{d}t\mathrm{dW}_t + \sigma_t^2(\mathrm{dW}_t)^2) + \ldots.
\end{aligned} \tag{5.3.2}$$

Now, recall that

$$\mathbb{E}(\mathrm{dW}_t)^2 = \mathbb{E}(\mathrm{W}_{t+\mathrm{d}t} - \mathrm{W}_t)^2 = \mathrm{d}t.$$

This implies that dW_t is $\mathcal{O}(\sqrt{\mathrm{d}t})$ and thus

$$(\mathrm{dW}_t)^2 \text{ is } \mathcal{O}(\mathrm{d}t), \qquad \mathrm{d}t\mathrm{dW}_t \text{ is } \mathcal{O}(\mathrm{d}t^{3/2}), \qquad (\mathrm{d}t)^2 \text{ is } \mathcal{O}(\mathrm{d}t^2).$$

Because the $f'(\mathrm{X}_t)$-term in (5.3.2) has a $\mathcal{O}(\mathrm{d}t)$ term, we need to keep the dW_t^2 term, which is $\mathcal{O}(\mathrm{d}t)$. On the other hand, we can ignore the $\mathrm{d}t\mathrm{dW}_t$ and $\mathrm{d}t^2$ terms, which are of $\mathcal{O}(\mathrm{d}t^{3/2})$ and $\mathcal{O}(\mathrm{d}t^2)$, respectively. Thus, we have from (5.3.2) that

$$\mathrm{d}f(\mathrm{X}_t) = f'(\mathrm{X}_t)(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t) + \tfrac{1}{2}f''(\mathrm{X}_t)\sigma_t^2(\mathrm{dW}_t)^2.$$

Now, roughly speaking, for a infinitesimally small time step $\mathrm{d}t$ we have

$$\mathbb{P}(\mathrm{dW}_t = \sqrt{\mathrm{d}t}) = 1/2, \qquad \mathbb{P}(\mathrm{dW}_t = -\sqrt{\mathrm{d}t}) = 1/2. \tag{5.3.3}$$

One can check that (5.3.3) yields the correct mean $\mathbb{E}\mathrm{dW}_t = 0$ and variance $\mathbb{V}\mathrm{dW}_t = \mathrm{d}t$. As we have $\mathrm{dW}_t = \pm\sqrt{\mathrm{d}t}$ it follows that $\mathrm{dW}_t^2 = \mathrm{d}t$. We have derived (albeit in a rather hand-waiving fashion), the following result.

Theorem 5.12 (Itô's Lemma, 1 dimension). *Let μ_t, σ_t be stochastic processes adapted to a filtration $\mathcal{F}$ for a Brownian motion* W*. Suppose* X *is the following Itô process given by* (5.2.3)*. Then we have*

$$\mathrm{d}f(\mathrm{X}_t) = \Big(\mu_t f'(\mathrm{X}_t) + \frac{1}{2}\sigma_t^2 f''(\mathrm{X}_t)\Big)\mathrm{d}t + f'(\mathrm{X}_t)\sigma_t \mathrm{dW}_t, \tag{5.3.4}$$

for any function f that is twice differentiable.

An easy way to remember Itô's Lemma is to use the following two-step procedure.

Step 1: Define $\mathrm{d}f(\mathrm{X}_t) := f(\mathrm{X}_t + \mathrm{dX}_t) - f(\mathrm{X}_t)$ and expand $f(\mathrm{X}_t + \mathrm{dX}_t)$ as a Taylor series about X_t to second order:

$$\mathrm{d}f(\mathrm{X}_t) = f'(\mathrm{X}_t)\mathrm{dX}_t + \frac{1}{2}f''(\mathrm{X}_t)(\mathrm{dX}_t)^2. \tag{5.3.5}$$

Step 2: Insert the expression (5.2.3) for dX_t and use the rules

$$(\mathrm{d}t)^2 = 0, \qquad \mathrm{d}t\mathrm{dW}_t = 0, \qquad (\mathrm{dW}_t)^2 = \mathrm{d}t. \tag{5.3.6}$$

Example 5.13. Suppose that X is given by (5.2.3). What is $\mathrm{d}f(\mathrm{X}_t)$ when $f(x) = x^p$? We have $f'(x) = px^{p-1}$ and $f''(x) = p(p-1)x^{p-2}$. Thus, from (5.3.5), we obtain

$$\mathrm{d}(\mathrm{X}_t^p) = p\mathrm{X}_t^{p-1}\mathrm{dX}_t + \frac{1}{2}p(p-1)\mathrm{X}_t^{p-2}(\mathrm{dX}_t)^2.$$

Now, we insert the expression (5.2.3) for dX_t and use the rules in (5.3.6). We have

$$\begin{aligned}
\mathrm{d}(\mathrm{X}_t^p) &= p\mathrm{X}_t^{p-1}(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t) + \frac{1}{2}p(p-1)\mathrm{X}_t^{p-2}(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t)^2 \\
&= p\mathrm{X}_t^{p-1}(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t) + \frac{1}{2}p(p-1)\mathrm{X}_t^{p-2}(\mu_t^2(\mathrm{d}t)^2 + 2\mu_t\sigma_t\mathrm{d}t\mathrm{dW}_t + \sigma_t^2(\mathrm{dW}_t)^2) \\
&= p\mathrm{X}_t^{p-1}(\mu_t\mathrm{d}t + \sigma_t\mathrm{dW}_t) + \frac{1}{2}p(p-1)\mathrm{X}_t^{p-2}\sigma_t^2\mathrm{d}t \\
&= \Big(p\mathrm{X}_t^{p-1}\mu_t + \frac{1}{2}p(p-1)\mathrm{X}_t^{p-2}\sigma_t^2\Big)\mathrm{d}t + p\mathrm{X}_t^{p-1}\sigma_t\mathrm{dW}_t.
\end{aligned}$$

Alternatively, we could have derived this expression using (5.3.4). Whether you choose to use (5.3.4) or the two-step method described above is up to you.

5.4 Extension to multiple dimensions

Itô's Lemma generalizes in a straightforward manner to multiple dimensions. We provide the main result without proof below.

Theorem 5.14 (d-dimensional Itô Formula). *Consider a d-dimensional process*

$$\mathrm{X} = (\mathrm{X}_t^{(1)}, \mathrm{X}_t^{(2)}, \dots, \mathrm{X}_t^{(d)})_{t \geq 0},$$

whose ith component is given by

$$\mathrm{dX}_t^{(i)} = \mu_t^{(i)} \mathrm{d}t + \sum_{j=1}^{n} \sigma_t^{(i,j)} \mathrm{dW}_t^{(j)}, \qquad i = 1, 2, \dots, d, \tag{5.4.1}$$

where $\mathrm{W} = (\mathrm{W}_t^{(1)}, \mathrm{W}_t^{(2)}, \dots, \mathrm{W}_t^{(n)})_{t \geq 0}$ is a n-dimensional Brownian motion with independent components (i.e., $\mathrm{W}^{(j)}$ is a one-dimensional Brownian motion for all j and $\mathrm{W}^{(j)} \perp\!\!\!\perp \mathrm{W}^{(k)}$ if $j \neq k$). Suppose $f : \mathbb{R}^d \to \mathbb{R}$. Then we have

$$\begin{aligned} \mathrm{d}f(\mathrm{X}_t) &= \Big(\sum_{i=1}^{d} \mu_t^{(i)} \partial_{x_i} f(\mathrm{X}_t) + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} (\sigma_t \sigma_t^{\mathrm{T}})^{(i,j)} \partial_{x_i} \partial_{x_j} f(\mathrm{X}_t) \Big) \mathrm{d}t \\ &\quad + \sum_{i=1}^{d} \sum_{j=1}^{n} \sigma_t^{(i,j)} \partial_{x_i} f(\mathrm{X}_t) \mathrm{dW}_t^{(j)}. \end{aligned} \tag{5.4.2}$$

where we have used the notation

$$\partial_{x_i} := \frac{\partial}{\partial x_i}, \qquad (\sigma_t \sigma_t^{\mathrm{T}})^{(i,j)} = \sum_{k=1}^{n} \sigma_t^{(i,k)} \sigma_t^{(j,k)}.$$

As in the one-dimensional case, rather than remember the multi-dimensional Itô formula (5.4.2) we can simply use the following two-step procedure

Step 1: Define $\mathrm{d}f(\mathrm{X}_t) := f(\mathrm{X}_t + \mathrm{dX}_t) - f(\mathrm{X}_t)$ and expand $f(\mathrm{X}_t + \mathrm{dX}_t)$ as a Taylor series about X_t to second order:

$$\mathrm{d}f(\mathrm{X}_t) = \sum_{i=1}^{d} \partial_{x_i} f(\mathrm{X}_t) \mathrm{dX}_t^{(i)} + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \partial_{x_i} \partial_{x_j} f(\mathrm{X}_t) \mathrm{dX}_t^{(i)} \mathrm{dX}_t^{(j)}. \tag{5.4.3}$$

Step 2: Insert the expression (5.4.1) for $\mathrm{dX}_t^{(i)}$ and use the rules

$$(\mathrm{d}t)^2 = 0, \qquad \mathrm{d}t \mathrm{dW}_t^{(i)} = 0, \qquad \mathrm{dW}_t^{(i)} \mathrm{dW}_t^{(i)} = \delta_{i,j} \mathrm{d}t, \tag{5.4.4}$$

where $\delta_{i,j}$ is a Kronecker delta (i.e., $\delta_{i,j} = 0$ if $i \neq j$ and $\delta_{i,i} = 1$).

Example 5.15 (Product Rule). Suppose that X and Y are Itô processes of the form

$$\mathrm{dX}_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{dW}_t^{(1)}, \qquad \mathrm{dY}_t = \alpha_t \mathrm{d}t + \beta_t \mathrm{dW}_t^{(1)} + \gamma_t \mathrm{dW}_t^{(2)}, \tag{5.4.5}$$

where $W^{(1)}$ and $W^{(2)}$ independent Brownian motions. What is $d(X_tY_t)$? Setting $f(x, y) = xy$, we have

$$\partial_x(xy) = y, \qquad \partial_y(xy) = x, \qquad \partial_x^2(xy) = 0, \qquad \partial_y^2(xy) = 0, \qquad \partial_x\partial_y(xy) = 1.$$

Thus, from (5.4.3) we have

$$d(X_tY_t) = Y_t dX_t + X_t dY_t + dX_t dY_t. \tag{5.4.6}$$

Equation (5.4.6) is known as the *product rule* for Itô processes. Next, using the expressions for dX_t and dY_t given in (5.4.5) and the rules provided in (5.4.4), we obtain

$$\begin{aligned} d(X_tY_t) &= Y_t(\mu_t dt + \sigma_t dW_t^{(1)}) + X_t(\alpha_t dt + \beta_t dW_t^{(1)} + \gamma_t dW_t^{(2)}) \\ &\quad + (\mu_t dt + \sigma_t dW_t^{(1)})(\alpha_t dt + \beta_t dW_t^{(1)} + \gamma_t dW_t^{(2)}) \\ &= Y_t(\mu_t dt + \sigma_t dW_t^{(1)}) + X_t(\alpha_t dt + \beta_t dW_t^{(1)} + \gamma_t dW_t^{(2)}) + \sigma_t\beta_t dt. \end{aligned}$$

Of course, we could have obtained the above expression by using (5.4.2) directly.

Example 5.16. Suppose X is given by (5.2.3). Let us compute $df(t, X_t)$. Note that f is a function of two variables. We Taylor expand $df(t, X_t) := f(t + dt, X_t + dX_t) - f(t, X_t)$ to second order about (t, X_t) to obtain

$$df(t, X_t) = \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial t} dt + \frac{1}{2}\Big(\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{\partial^2 f}{\partial x \partial t} dX_t dt + \frac{\partial^2 f}{\partial t^2} dt^2\Big).$$

Now, we insert the expression for dX_t and use the rules $(dt)^2 = 0$, $dW_t dt = 0$ and $(dW_t)^2 = dt$ and obtain

$$df(t, X_t) = \left[\frac{\partial f}{\partial t}(t, X_t) + \mu_t \frac{\partial f}{\partial x}(t, X_t) + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}(t, X_t)\right] dt + \sigma_t \frac{\partial f}{\partial x}(t, X_t) dW_t.$$

5.5 Stochastic Differential Equations

Consider an Itô process $X = (X_t)_{t\geq 0}$ whose coefficients μ and σ are deterministic functions of t and X_t. Then the dynamics of X are of the form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t. \tag{5.5.1}$$

In the simplest case, X lives in $\mathbb{R}$ and $W = (W_t)_{t\geq 0}$ is a one-dimensional Brownian motion. But, more generally, we could have $X \in \mathbb{R}^d$ and $W \in \mathbb{R}^n$, in which case we would have $\mu : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n}$. The ith component of X would then given by

$$dX_t^{(i)} = \mu^{(i)}(t, X_t)dt + \sum_{j=1}^{n} \sigma^{(i,j)}(t, X_t)dW_t^{(j)}, \qquad i = 1, 2, \dots, d.$$

We call an equation of the form (5.5.1) a *Stochastic Differential Equation* (SDE). The solution of an SDE is a functional of a the Brownian motion W. That is $X_T = F[\{W_t, 0 \le t \le t\}]$. We will not prove the following very important theorem.

Theorem 5.17. *If an SDE of the form* (5.5.1) *has a unique solution* X, *then* X *is a Markov process.*

What theorem 5.17 tell us is that, for any function g, and for any $T > t$, there exists a function u such that

$$\mathbb{E}[g(X_T)|\mathcal{F}_t^X] = \mathbb{E}[g(X_T)|X_t] = u(t, X_t),$$

where $\mathcal{F}_t^X = \{X_s, 0 \le s \le t\}$ is the history of X up to time t (i.e., $\mathcal{F}^X$ is the filtration generated by observing the path of X). Let us see some examples of SDEs.

Example 5.18 (Geometric Brownian Motion). Suppose $S = (S_t)_{t \ge 0}$ is the solution of the following SDE

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

For reasons that will become apparent later, let us compute $d(\log S_t)$. We have

$$f(s) = \log s, \qquad f'(s) = \frac{1}{s}, \qquad f''s) = \frac{-1}{s^2}.$$

Thus, from Itô's Lemma, we have

$$\begin{aligned} d \log S_t &= \frac{1}{S_t} dS_t + \frac{1}{2}\frac{-1}{S_t^2}\sigma^2 S_t^2 dt \\ &= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t. \end{aligned}$$

Integrating from 0 to T we obtain

$$\log S_T - \log S_0 = \log(S_T/S_0) = \left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W_T.$$

Taking the exponential both sides, and multiplying both by S_0, we obtain

$$S_T = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right).$$

We now have an explicit expression for S_T in terms of T and W_T. As we shall see, people often model the price of a stock as a geometric Brownian motion.

Example 5.19 (Ornstein-Uhlenbeck Process). An Ornstein Uhlenbek (OU) process is the solution X of an SDE of the following form

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t.$$

Let us define $Y_t := X_t - \theta$. Then we have $dX_t = dY_t$ and thus

$$dY_t = \kappa(\theta - X_t)dt + \sigma dW_t = -\kappa Y_t dt + \sigma dW_t.$$

Now define $Z_t := e^{\kappa t} Y_t$. Noting that $d(e^{\kappa t}) = \kappa e^{\kappa t} dt$, we have

$$dZ_t = \kappa e^{\kappa t} Y_t dt + e^{\kappa t} dY_t = e^{\kappa t}\sigma\, dW_t.$$

Integrating from 0 to T, we obtain

$$Z_T - Z_0 = \int_0^T e^{\kappa t}\sigma dW_t.$$

Next, using $Z_t := e^{\kappa t} Y_t$, we have

$$e^{\kappa T} Y_T - Y_0 = \int_0^T e^{\kappa t}\sigma dW_t.$$

Using $Y_t := X_t - \theta$, we obtain

$$e^{\kappa T}(X_T - \theta) - (X_0 - \theta) = \int_0^T e^{\kappa t}\sigma dW_t.$$

Finally, solving for X_T, we find

$$X_T = \theta + e^{-\kappa T}(X_0 - \theta) + \int_0^T e^{-\kappa(T-t)}\sigma dW_t.$$

We now have an explicit expression for X_T in terms of T and $(W_s)_{0\le s\le t}$. Note that X_T is a normally distributed random variable because the Itô integral has a deterministic integrand. Let us find the mean and variance of X_T. We have

$$\mathbb{E}X_T = \theta + e^{-\kappa T}(X_0 - \theta),$$

where we have used the fact the the expected value of an Itô integral is zero. Next, we compute the variance of X_T. We have

$$\mathbb{V}X_T = \mathbb{E}[(X_T - \mathbb{E}[X_T])^2] = \mathbb{E}\left[\int_0^T e^{-\kappa(T-t)}\sigma dW_t\right] = \int_0^T e^{-2\kappa(T-t)}\sigma^2 dt = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa T}),$$

where the third equality follows from the Itô isometry.

5.6 Simulating a SDE

Many SDEs do not have explicit solutions. In such cases, it is often useful to simulate approximate sample paths of the solution. Consider the following SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t.$$

In a small interval of time $\mathrm{d}t$, the process X behaves approximately as follows

$$X_{t+\delta t} - X_t \approx \mu(t, X_t)\delta t + \sigma(t, X_t)(W_{t+\delta t} - W_t),$$

where $W_{t+\delta t} - W_t \sim N(0, \mathrm{d}t)$. The above equation provides a mechanism for simulating (at least approximately) sample paths of X.

Step 1: Divide the the time interval of interest $[0, T]$ into intervals of size $\delta t := T/N$, where N is a fixed positive integer.

Step 2: For $n = 0, 1, \ldots, N-1$, set

$$\widehat{X}_{(n+1)\delta t} = \widehat{X}_{n\delta t} + \mu(n\delta t, \widehat{X}_{n\delta t})\delta t + \sigma(n\delta t, \widehat{X}_{n\delta t})\sqrt{\delta t} Z_{n+1},$$

where the $(Z_n)_{n\in\{1,2,\ldots,N\}}$ are iid standard normal random variables $Z_n \sim N(0, 1)$.

The resulting discrete time process $\widehat{X} = (\widehat{X}_{n\delta t})_{n\in\{0,1,\ldots,N\}}$ has approximately the same distribution as the solution X of the SDE

$$(\widehat{X}_0, \widehat{X}_{\delta t}, \ldots, \widehat{X}_{N\delta t}) \overset{\mathcal{D}}{\approx} (X_0, X_{\delta t}, \ldots, X_{N\delta t}).$$

Example 5.20. See Mathematica Notebook `MonteCarlo.nb` for an example.

In some situations, it may be better to make a variable transformation before running the simulation. For example, consider a Geometric Brownian motion

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t.$$

If we simulate this process directly, we are making a small error at every step because, for every n we are making approximation

$$\int_{n\delta t}^{(n+1)\delta t} \mu S_u \mathrm{d}u + \int_{n\delta t}^{(n+1)\delta t} \sigma S_u \mathrm{d}W_u \approx \mu S_{n\delta t}\delta t + \sigma S_{n\delta t}(W_{(n+1)\delta t} - W_{n\delta t}).$$

Now, suppose we define a new process X as follows $X_t = \log S_t$. Then the dynamics of X are

$$\mathrm{d}X_t = (\mu - \frac{1}{2}\sigma^2)\mathrm{d}t + \sigma \mathrm{d}W.$$

Now, if we simulate X we make no error because, for every n we have exactly

$$\int_{n\delta t}^{(n+1)\delta t} (\mu - \tfrac{1}{2}\sigma^2)\mathrm{d}u + \int_{n\delta t}^{(n+1)\delta t} \sigma \mathrm{d}W_u = (\mu - \tfrac{1}{2}\sigma^2)\delta t + \sigma(W_{(n+1)\delta t} - W_{n\delta t}).$$

Thus, in this case, is better to simulate sample paths of X rather than sample paths of S. Once we have generated a sample path of X, we can obtain a sample path of S by setting $S_t = e^{X_t}$.

5.7 Exercises

Exercise 5.1. Consider a model for a stock and a money market account

$$\begin{aligned}
dS_t &= \mu S_t dt + \sigma S_t dW_t, \\
dM_t &= r M_t dt.
\end{aligned}$$

Define $X_t = S_t/M_t$. Compute dX_t. You should be able to write your answer as an SDE involving only X_t (neither S nor M should appear in your answer).

Exercise 5.2. Consider a model for two stocks

$$\begin{aligned}
dS_t &= \mu S_t dt + \sigma S_t dW_t, \\
dY_t &= \alpha Y_t dt + \beta Y_t(\rho dW_t + \sqrt{1-\rho^2} dB_t),
\end{aligned}$$

where $\rho \in (-1, 1)$ and W and B are independent Brownian motions. Define $Z_t = S_t/Y_t$. Compute dZ_t. You should be able to write your answer as an SDE involving only Z_t (neither S nor Y should appear in your answer).

Exercise 5.3. Consider a model for a stock S.

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

Define $P_t = S_t^\beta$. Compute dP_t. You should be able to write your answer as an SDE involving only P_t (i.e., S should not appear in your answer). Find all values of β for which the process P is a martingale.

Exercise 5.4. Consider the following two-dimensional SDE

$$dX_t = -\tfrac{1}{2} X_t dt - Y_t dW_t, \qquad\qquad Y_t = -\tfrac{1}{2} Y_t dt + X_t dW_t.$$

(a) Compute $d(X_t^2 + Y_t^2)$.
(b) Show that $X_t = \cos W_t$ and $Y_t = \sin W_t$ is a solution of the above SDE (i.e., compute dX_t and dY_t and show that you obtain the SDE above.
(c) Using your result from part (b), explain your answer to part (a).

Exercise 5.5. Suppose X satisfies

$$dX_t = \mu(X_t) dt + \sigma(X_t) dW_t.$$

Define $M_t = e^{-\lambda t} \phi(X_t)$. Derive an ordinary differential equation that ϕ must satisfy in order for M to be a martingale.

Exercise 5.6. Suppose X satisfies

$$dX_t = \mu dt + \sigma dW_t.$$

Define Y, the exponential moving average of X, by $Y_t = \int_{-\infty}^{t} e^{-\lambda(t-s)} X_s ds$. Next, define $Z_t = X_t - Y_t$. Derive an SDE for Z.

Exercise 5.7. Suppose Z is given by

$$Z_t = \exp\Big(-\frac{1}{2} \int_0^t \Delta_s^2 ds + \int_0^t \Delta_s dW_s \Big).$$

where Δ is bounded and adapted to the filtration generated by observing the Brownian motion W. What is $\mathbb{E}Z_t$?

Chapter 6

Black-Scholes model

Consider a financial market in which a stock $S = (S_t)_{0 \le t \le T}$ and a bond $B = (B_t)_{0 \le t \le T}$, whose dynamics under the real-world probability measure $\mathbb{P}$ are given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \qquad dB_t = r B_t dt, \tag{6.0.1}$$

where $W = (W_t)_{0 \le t \le T}$ is a Brownian motion. Expression (6.0.1) is the well-known Black-Scholes model. In this chapter, we derive the price of a European option in the Black-Scholes setting. Just as in the Binomial model, we will present two methods (i) pricing by replication, and (ii) risk-neutral pricing.

6.1 Pricing by replication

Consider a European option that pays $\phi(S_T)$ at time T. We will try to construct a portfolio $X = (X_t)_{0 \le t \le T}$ that replicates the option payoff: $X_T = \phi(S_T)$. Let us denote by $\Delta = (\Delta_t)_{0 \le t \le T}$ the number of shares of S we hold in our portfolio. If changes in the value of X are due only to the changes in the value of S and B (i.e., the portfolio is self-financing), then the dynamics of X must be as follows

$$\begin{aligned} dX_t &= \Delta_t dS_t + (X_t - \Delta_t S_t) \frac{1}{B_t} dB_t \\ &= \Delta_t (\mu S_t dt + \sigma S_t dW_t) + (X_t - \Delta_t S_t) r dt \\ &= \Delta_t S_t (\mu - r) dt + \Delta_t S_t \sigma dW_t + r X_t dt, \end{aligned} \tag{6.1.1}$$

where, in the second equality, we have used (6.0.1). It will simplify the computations that follow if we derive the dynamics of (X/B). Using the product rule, we compute

$$d\Big(\frac{X_t}{B_t}\Big) = \frac{1}{B_t} dX_t + X_t d\Big(\frac{1}{B_t}\Big) + dX_t d\Big(\frac{1}{B_t}\Big)$$

$$
\begin{aligned}
&= \frac{1}{\mathrm{B}_t}\mathrm{dX}_t + \frac{-\mathrm{X}_t}{\mathrm{B}_t^2}\mathrm{dB}_t \\
&= \frac{\Delta_t \mathrm{S}_t}{\mathrm{B}_t}(\mu - r)\mathrm{d}t + \frac{\Delta_t \mathrm{S}_t}{\mathrm{B}_t}\sigma \mathrm{dW}_t. \qquad (6.1.2)
\end{aligned}
$$

where we have used (6.0.1) and (6.1.1).

Let $\mathrm{V} = (\mathrm{V}_t)_{0 \le t \le \mathrm{T}}$ denote the value of the option with payoff $\phi(\mathrm{S_T})$. The option value V will fluctuate as S moves randomly. Thus, it is reasonable to expect that the option value V is a function v of time and the stock price: $\mathrm{V}_t = v(t, \mathrm{S}_t)$. Obviously, V should depend on other parameters as well (e.g., r and σ). But these other parameters are constant whereas t and S evolve in time. Using the multidimensional version of Itô's Lemma, we compute the dynamics of V as follows

$$
\begin{aligned}
\mathrm{dV}_t &= \mathrm{d}v(t, \mathrm{S}_t) \\
&= \frac{\partial v}{\partial t}(t, \mathrm{S}_t)\mathrm{d}t + \frac{\partial v}{\partial s}(t, \mathrm{S}_t)\mathrm{dS}_t + \frac{1}{2}\frac{\partial^2 v}{\partial s^2}(t, \mathrm{S}_t)(\mathrm{dS}_t)^2 \\
&= \left[\frac{\partial v}{\partial t}(t, \mathrm{S}_t) + \mu \mathrm{S}_t \frac{\partial v}{\partial s}(t, \mathrm{S}_t) + \frac{1}{2}\sigma^2 \mathrm{S}_t^2 \frac{\partial^2 v}{\partial s^2}(t, \mathrm{S}_t)\right]\mathrm{d}t + \frac{\partial v}{\partial s}(t, \mathrm{S}_t)\sigma \mathrm{S}_t \mathrm{dW}_t, \qquad (6.1.3)
\end{aligned}
$$

where we have used (6.0.1). Once again, it will simplify the computations that follow if we derive the dynamics of (V/B). Using the product rule, we obtain

$$
\begin{aligned}
\mathrm{d}\left(\frac{\mathrm{V}_t}{\mathrm{B}_t}\right) &= \left(\frac{1}{\mathrm{B}_t}\right)\mathrm{dV}_t + \mathrm{V}_t \mathrm{d}\left(\frac{1}{\mathrm{B}_t}\right) + \mathrm{dV}_t \mathrm{d}\left(\frac{1}{\mathrm{B}_t}\right) \\
&= \frac{1}{\mathrm{B}_t}\mathrm{dV}_t + \frac{-\mathrm{V}_t}{\mathrm{B}_t^2}\mathrm{dB}_t \\
&= \frac{1}{\mathrm{B}_t}\left[\frac{\partial v}{\partial t}(t, \mathrm{S}_t) + \mu \mathrm{S}_t \frac{\partial v}{\partial s}(t, \mathrm{S}_t) + \frac{1}{2}\sigma^2 \mathrm{S}_t^2 \frac{\partial^2 v}{\partial s^2}(t, \mathrm{S}_t) - r v(t, \mathrm{S}_t)\right]\mathrm{d}t \\
&\quad + \frac{1}{\mathrm{B}_t}\frac{\partial v}{\partial s}(t, \mathrm{S}_t)\sigma \mathrm{S}_t \mathrm{dW}_t.
\end{aligned}
$$

where we have used (6.0.1) and (6.1.3), as well as $\mathrm{V}_t = v(t, \mathrm{S}_t)$.

Recall, we are seeking a portfolio X that replicates the option value V. Suppose that

$$
\mathrm{X}_0 = \mathrm{V}_0 \qquad \text{and} \qquad \mathrm{d}\left(\frac{\mathrm{X}_t}{\mathrm{B}_t}\right) = \mathrm{d}\left(\frac{\mathrm{V}_t}{\mathrm{B}_t}\right). \qquad (6.1.4)
$$

Then we would have $\mathrm{X}_t = \mathrm{V}_t$ for all $t \in [0, \mathrm{T}]$ because

$$
\frac{\mathrm{X}_t}{\mathrm{B}_t} = \frac{\mathrm{X}_0}{\mathrm{B}_0} + \int_0^t \mathrm{d}\left(\frac{\mathrm{X}_s}{\mathrm{B}_s}\right) = \frac{\mathrm{V}_0}{\mathrm{B}_0} + \int_0^t \mathrm{d}\left(\frac{\mathrm{V}_s}{\mathrm{B}_s}\right) = \frac{\mathrm{V}_t}{\mathrm{B}_t}.
$$

In order for (6.1.4) to hold, we must have

$$
0 = \mathrm{d}\left(\frac{\mathrm{V}_t}{\mathrm{B}_t}\right) - \mathrm{d}\left(\frac{\mathrm{X}_t}{\mathrm{B}_t}\right)
$$

$$= \frac{1}{B_t}\left[\frac{\partial v}{\partial t}(t,S_t) + \mu S_t \frac{\partial v}{\partial s}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 v}{\partial s^2}(t,S_t) - rv(t,S_t) - (\mu - r)\Delta_t S_t\right] dt$$
$$+ \frac{1}{B_t}\left(\frac{\partial v}{\partial s}(t,S_t) - \Delta_t\right)\sigma S_t dW_t. \qquad (6.1.5)$$

The dW_t-term in (6.1.5) will equal zero if and only if we set

$$\Delta_t = \frac{\partial v}{\partial s}(t,S_t). \qquad (6.1.6)$$

Inserting expression (6.1.6) for Δ_t into (6.1.5), we obtain

$$0 = \frac{1}{B_t}\left[\frac{\partial v}{\partial t}(t,S_t) + rS_t \frac{\partial v}{\partial s}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 v}{\partial s^2}(t,S_t) - rv(t,S_t)\right] dt.$$

The dt-term will equal zero if and only if the function v satisfies

$$0 = \frac{\partial v}{\partial t} + rs\frac{\partial v}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} - rv. \qquad (6.1.7)$$

Equation (6.1.7) is the famous *Black-Scholes Partial Differential Equation* (PDE). In order to find a solution v of (6.1.7) we need a terminal condition for the function v. This can be obtained as follows

$$v(T,S_T) = V_T = \phi(S_T), \qquad \Rightarrow \qquad v(T,s) = \phi(s). \qquad (6.1.8)$$

The terminal condition (6.1.8) together with the PDE (6.1.7) are sufficient to determine the function v uniquely. We will solve the PDE (6.1.7) in a subsequent section. For now, let us review what we have done.

1. We have supposed that the dynamics of a stock S and a bond B are given by (6.0.1).

2. We have considered a European option that pays $\phi(S_T)$ at time T. The value $V = (V_t)_{0\leq t\leq T}$ of this option is given by $V_t = v(t,S_t)$ where the function v satisfies PDE (6.1.7) and terminal condition (6.1.8).

3. If the initial value of a portfolio X is given by $X_0 = V_0 = v(0,S_0)$ and if the number of shares Δ of S one holds in this portfolio is given by (6.1.6), then the portfolio X replicates the value V of the option. In particular, we have $X_T = V_T = \phi(S_T)$.

6.2 Risk-Neutral Pricing

The 1st fundamental theorem of asset pricing (Theorem 4.2) states that a market defined under a probability measure $\mathbb{P}$ is free of arbitrage if and only if there exists a probability measure $\widetilde{\mathbb{P}}$, equivalent to

$\mathbb{P}$, under which all portfolios X, denominated in units of a common numéraire N are martingales. In the Black-Scholes setting (6.0.1) this leads to a very simple procedure for finding the value $V = (V_t)_{0\leq t\leq T}$ of an option that pays $\phi(S_T)$ at time T.

1. Choose a numèraire. We will choose the bond $B = (B_t)_{0\leq t\leq T}$, as this choice leads to the easiest computations.

2. Search for a probability measure $\widetilde{\mathbb{P}}$, under which (X/B) is a martingale for all portfolios X. That is, find $\widetilde{\mathbb{P}}$ such that
$$\frac{X_t}{B_t} = \widetilde{\mathbb{E}}\left(\frac{X_T}{B_T}\middle|\mathcal{F}_t^S\right),$$
where X is given by (6.1.1) and $\mathcal{F}^S$ is the filtration generated by observing S.

3. The value V of the option can then be computed using
$$\frac{V_t}{B_t} = \widetilde{\mathbb{E}}\left(\frac{V_T}{B_T}\middle|\mathcal{F}_t^S\right) = \widetilde{\mathbb{E}}\left(\frac{\phi(S_T)}{B_T}\middle|\mathcal{F}_t^S\right). \tag{6.2.1}$$

In the binomial model, "search for a probability measure $\widetilde{\mathbb{P}}$" meant that we should look for probabilities $\widetilde{p}$ and $\widetilde{q}$ that the stock would go up and down. But, what does it mean to "search for a probability measure $\widetilde{\mathbb{P}}$" when the market is described by SDEs (6.0.1)? The following Theorem provides an answer to this question.

Theorem 6.1 (Girsanov's Theorem). *Suppose* $W = (W_t^{(1)}, \ldots, W_t^{(d)})_{0\leq t\leq T}$ *is a d-dimensional Brownian motion with independent componentes under a probability measure* $\mathbb{P}$*. Define a process* $\widetilde{W} = (\widetilde{W}_t^{(1)}, \ldots, \widetilde{W}_t^{(d)})_{0\leq t\leq T}$ *by*
$$\widetilde{W}_t = \int_0^t \gamma_s \mathrm{d}s + W_t, \tag{6.2.2}$$
where $\gamma = (\gamma_t^{(1)}, \ldots, \gamma_t^{(d)})_{0\leq t\leq T}$ *is adapted to a filtration* $\mathcal{F}$ *for* W*. Then there exists a probability measure* $\widetilde{\mathbb{P}}$*, equivalent to* $\mathbb{P}$*, under which* $\widetilde{W}$ *is a d-dimensional Brownian motion with independent components.*

Let us see how to use Theorem 6.1 to write the dynamics of (X/B) under a risk-neutral measure $\widetilde{\mathbb{P}}$. Using equation (6.1.2), the dynamics of X/B are given by
$$\begin{aligned}\mathrm{d}\left(\frac{X_t}{B_t}\right) &= \frac{\Delta_t S_t}{B_t}(\mu - r)\mathrm{d}t + \frac{\Delta_t S_t}{B_t}\sigma \mathrm{d}W_t \\ &= \frac{\Delta_t S_t}{B_t}(\mu - r)\mathrm{d}t + \frac{\Delta_t S_t}{B_t}\sigma(\mathrm{d}\widetilde{W}_t - \gamma_t \mathrm{d}t)\end{aligned}$$

$$= \frac{\Delta_t S_t}{B_t}(\mu - r - \sigma\gamma_t)dt + \frac{\Delta_t S_t}{B_t}\sigma d\widetilde{W}_t,$$

where, in the second equality, we have used (6.2.2). Now, we know from Theorem 6.1 that, for any choice of γ, there exists a probability $\widetilde{\mathbb{P}}$ under which $\widetilde{W}$ is a Brownian motion. If we choose γ such that the dt-term equals zero, then X/B will be a martingale under $\widetilde{\mathbb{P}}$. Thus, we choose

$$\gamma_t = \frac{\mu - r}{\sigma}. \tag{6.2.3}$$

Now, what are the dynamics of S under $\widetilde{\mathbb{P}}$? We have

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t \\ &= \mu S_t dt + \sigma S_t(d\widetilde{W}_t - \gamma_t dt) \\ &= (\mu - \sigma\gamma_t)S_t dt + \sigma S_t d\widetilde{W}_t \\ &= rS_t dt + \sigma S_t d\widetilde{W}_t, \end{aligned}$$

where, in the last line, we have used (6.2.3). Under $\widetilde{\mathbb{P}}$, the stock price S is a geometric Brownian motion with drift r and volatility σ. We found an explicit expression for S in Example 5.18. We have

$$S_T = S_t \exp\left[(r - \tfrac{1}{2}\sigma^2)(T - t) + \sigma(\widetilde{W}_T - \widetilde{W}_t)\right].$$

We can now compute the value V of the option. From (6.2.1), we have

$$\begin{aligned} V_t &= B_t\widetilde{\mathbb{E}}\left(\frac{\phi(S_T)}{B_T}\middle|\mathcal{F}_t^S\right) \\ &= e^{-r(T-t)}\widetilde{\mathbb{E}}\left(\phi\left(S_t \exp\left[(r - \tfrac{1}{2}\sigma^2)(T - t) + \sigma(\widetilde{W}_T - \widetilde{W}_t)\right]\right)\middle|S_t\right) \\ &= \int_{\mathbb{R}} \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(\frac{-y^2}{2\sigma^2(T-t)}\right)\phi\left(S_t \exp\left[(r - \tfrac{1}{2}\sigma^2)(T - t) + y\right]\right)dy =: v(t, S_t), \end{aligned} \tag{6.2.4}$$

where, in the first equality, we have used the fact that S is a Markov process, and in the second equality, we have used the fact that $\sigma(\widetilde{W}_T - \widetilde{W}_t) \sim N(0, \sigma^2(T - t))$ under $\widetilde{\mathbb{P}}$. For all but a few particular choices of ϕ, the integral in (6.2.4) must be computed numerically. The function v defined in (6.2.4) is, in fact, the unique solution of the Black-Scholes PDE (6.1.7) and TC (6.1.8). In the next section, we will show how to find this function using PDE methods.

Remark 6.2. Observe that *the drift* μ *of* S *does not appear* in the price (6.2.4) of the option nor does it appear in the Black-Scholes PDE (6.1.7). In fact, although we have assumed that the drift of S was a constant μ in (6.0.1), we could have assumed that the drift of S was a stochastic process $\mu = (\mu_t)_{0\leq t\leq T}$, and this would not have had any effect at all on the option price. The fact that μ plays no role in the pricing of options is analogous to the real-world probabilities p and q playing no role in the binomial setting.

6.3 Solving Black–Scholes PDE

Recall that the Black–Scholes PDE is and TC are

$$0 = \left(\partial_t - r + rs\partial_s + \frac{1}{2}\sigma^2 s^2 \partial_s^2\right) v, \qquad v(\mathrm{T}, s) = \phi(s), \tag{6.3.1}$$

where, we have re-introduced the notation $\partial_s^n := \frac{\partial^n}{\partial s^n}$ and likewise for other partial derivatives. We will solve this PDE in a series of steps.

Step 1: Change of Variables. Suppose that

$$v(t, s) = \mathrm{e}^{-r(\mathrm{T}-t)} u(\tau(t), x(t, s)), \tag{6.3.2}$$

where the functions τ and x are defined as follows

$$\tau(t) := \mathrm{T} - t, \qquad x(t, s) := \frac{1}{\sigma}\left[\log s + \left(r - \frac{1}{2}\sigma^2\right)(\mathrm{T} - t)\right].$$

Let us see what the function u must satisfy. Using the chain rule, we find

$$\begin{aligned}
\frac{\partial v}{\partial t} &= \mathrm{e}^{-r(\mathrm{T}-t)}\left[ru + \partial_\tau u \cdot \partial_t \tau + \partial_x u \cdot \partial_t x\right], \\
\frac{\partial v}{\partial s} &= \mathrm{e}^{-r(\mathrm{T}-t)} \partial_x u \cdot \partial_s x, \\
\frac{\partial^2 v}{\partial s^2} &= \mathrm{e}^{-r(\mathrm{T}-t)}\left[\partial_x^2 u \cdot (\partial_s x)^2 + \partial_x u \cdot \partial_s^2 x\right],
\end{aligned}$$

where

$$\partial_t \tau(t) = -1, \qquad \partial_t x(t, s) = \frac{-1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right), \qquad \partial_s x(t, s) = \frac{1}{\sigma s}, \qquad \partial_s^2 x(t, s) = \frac{-1}{\sigma s^2}.$$

Plugging the above expressions in to the Black-Scholes PDE (6.3.1) yields

$$0 = (-\partial_\tau + \frac{1}{2}\partial_x^2)u, \qquad u(0, \sigma^{-1}\log s) = \phi(s) \qquad \Rightarrow \qquad u(0, x) = \phi(\mathrm{e}^{\sigma x}). \tag{6.3.3}$$

We have obtained a PDE and initial condition (IC) for the function u.

Step 2: Fundamental solution. We introduce a function Γ, which is given by

$$\Gamma(\tau, x, y) := \frac{1}{\sqrt{2\pi\tau}} \mathrm{e}^{\frac{-(x-y)^2}{2\tau}}.$$

One can check by direct computation that Γ satisfies

$$0 = (-\partial_\tau + \frac{1}{2}\partial_x^2)\Gamma, \qquad \Gamma(0, x, y) = \delta(x - y).$$

Thus, Γ satisfies the PDE in (6.3.3), but not the IC. We call Γ the *fundamental solution* of $(-\partial_\tau + \frac{1}{2}\partial_x^2)$.

Step 3: Solution of the PDE We claim that the solution u of (6.3.3) is given by

$$u(\tau, x) = \int_{\mathbb{R}} dy\, \Gamma(\tau, x, y)\phi(e^{\sigma y}).$$

To see this, note that

$$u(0, x) = \int_{\mathbb{R}} \underbrace{\Gamma(0, x, y)}_{\delta(x-y)} \phi(e^{\sigma y}) = \phi(\mathrm{e}^{\sigma x})\mathrm{d}y.$$

Thus, u satisfies the IC $u(0, x) = \phi(\mathrm{e}^{\sigma x})$. Next, note that

$$(-\partial_\tau + \frac{1}{2}\partial_x^2)u(\tau, x) = \int_{\mathbb{R}} \underbrace{(-\partial_\tau + \frac{1}{2}\partial_x^2)\Gamma(\tau, x, y)}_{=0} \phi(e^{\sigma y})\mathrm{d}y = 0.$$

Hence, the function u also satisfies the PDE in (6.3.3).

Step 4: Undo the change of variables. Finally, we need to write the function v in terms of t and s. Using (6.3.2), we have

$$\begin{aligned} v(t, s) &= \mathrm{e}^{-r(\mathrm{T}-t)} u(\tau(t), x(t, s)) \\ &= \mathrm{e}^{-r(\mathrm{T}-t)} \int_{\mathbb{R}} \Gamma(\tau(t), x(t, s), y)\phi(e^{\sigma y})\mathrm{d}y \\ &= \int_{\mathbb{R}} \frac{\mathrm{e}^{-r(\mathrm{T}-t)}}{\sqrt{2\pi(\mathrm{T}-t)}} \exp\left(\frac{-(\log s + (r - \sigma^2/2)(\mathrm{T}-t) - y)^2}{2\sigma^2(\mathrm{T}-t)}\right) \phi(\mathrm{e}^{\sigma y})\mathrm{d}y. \end{aligned} \tag{6.3.4}$$

One can show, via a change of variables, that the expression (6.3.4) agrees with the expression given for v in (6.2.4).

6.4 Black-Scholes formulas for European Calls and Puts

For call and put options we have $\phi(s) = (s - \mathrm{K})^+$ and $\phi(s) = (\mathrm{K} - s)^+$. In these cases the above integral (6.3.4) can be computed explicitly, leading to the following expressions for call and put prices

$$\mathrm{C}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \mathrm{S}_t \Phi(d_+) - \mathrm{K}\Phi(d_-)\mathrm{e}^{-r(\mathrm{T}-t)}, \tag{6.4.1}$$

$$\mathrm{P}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \mathrm{K}\mathrm{e}^{-r(\mathrm{T}-t)}\Phi(-d_-) - \mathrm{S}_t\Phi(-d_+), \tag{6.4.2}$$

where Φ and $d_\pm = d_\pm(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})$ are given by

$$d_\pm(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) := \frac{1}{\sigma\sqrt{\mathrm{T}-t}}\left[\log\left(\frac{\mathrm{S}_t}{\mathrm{K}}\right) + \left(r \pm \frac{1}{2}\sigma^2\right)(\mathrm{T}-t)\right], \qquad \Phi(x) := \int_{-\infty}^{x} \frac{\mathrm{e}^{\frac{-z^2}{2}}}{\sqrt{2\pi}}\mathrm{d}z. \tag{6.4.3}$$

Note that the function Φ is the *cumulative distribution function of a standard normal random variable*. When it causes no confusion, we may omit the arguments $(\sigma, \mathrm{T}, \mathrm{K})$ from C^{BS} and P^{BS}. Note that

$$\Delta_t^{\text{BS-Call}} = \frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial s}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \Phi(d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})) \in (0,1),$$
$$\Delta_t^{\text{BS-Put}} = \frac{\partial \mathrm{P}^{\mathrm{BS}}}{\partial s}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \Phi(d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})) - 1 \in (-1,0).$$

These are the deltas of calls and puts, respectively, (i.e. the number of shares Δ of the stock S one should holds in a portfolio X in order to replicate a Call or Put). Thus, to replicate a Call, one always holds a positive number of shares of S and to replicate a Put, one always holds a negative number of shares of S. Take a moment to think about why this is the case intuitively (*hint*: think about what happens to the value of a Call or Put when the underlying stock S goes up).

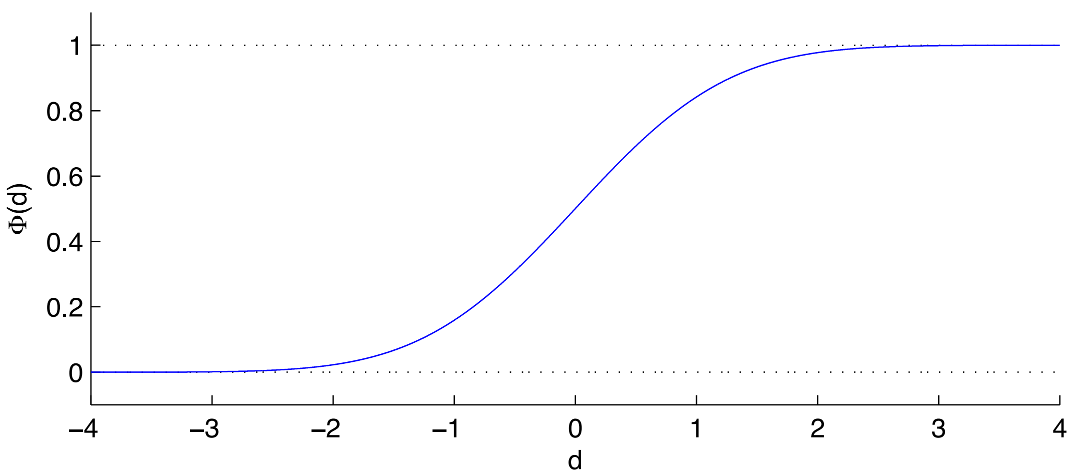

Figure 6.1: The cumulative distribution function Φ of a standard normal random variable.

6.5 Implied Volatility

Note from (6.4.1) and (6.4.2), that the Black-Scholes price of a Call or Put depends on five observable variables $(t, \mathrm{S}_t, \mathrm{T}, \mathrm{K}, r)$ and one *unobservable* variable σ. This permits us to make the following definition.

Definition 6.3 (Implied volatility). Let $\mathrm{C}_t(\mathrm{T}, \mathrm{K})$ be the price of a European Call option, written on an underlyer S with strike K and maturity T, as quoted on the market. We define the *Implied volatility* $\sigma(\mathrm{T}, \mathrm{K})$ of this option as the unique positive solution of

$$\mathrm{C}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma(\mathrm{T}, \mathrm{K}), \mathrm{T}, \mathrm{K}) = \mathrm{C}_t(\mathrm{T}, \mathrm{K}), \tag{6.5.1}$$

where C^{BS} is given by (6.4.1).

The implied volatility of a Put with price $P_t(T,K)$ is defined analogously, i.e., as the unique positive solution $\sigma(T,K)$ of

$$P^{BS}(t,S_t;\sigma(T,K),T,K) = P_t(T,K).$$

Observe that the a Put and Call with the same strike and maturity must have the same implied volatility. To see why this is the case, suppose that we have $C_t(T,K) = C^{BS}(t,S_t;\sigma_c,T,K)$ and $P_t(T,K) = P^{BS}(t,S_t;\sigma_p,T,K)$. Then using Put-Call parity we have

$$\begin{aligned}
S_t - B_t^T K = C_t(T,K) - P_t(T,K) &= C^{BS}(t,S_t;\sigma_c,T,K) - P^{BS}(t,S_t;\sigma_p,T,K) \\
&= C^{BS}(t,S_t;\sigma_c,T,K) - P^{BS}(t,S_t;\sigma_c,T,K) + P^{BS}(t,S_t;\sigma_c,T,K) - P^{BS}(t,S_t;\sigma_p,T,K) \\
&= S_t - B_t^T K + P^{BS}(t,S_t;\sigma_c,T,K) - P^{BS}(t,S_t;\sigma_p,T,K).
\end{aligned}$$

Thus, we have

$$P^{BS}(t,S_t;\sigma_c,T,K) = P^{BS}(t,S_t;\sigma_p,T,K),$$

which is true if and only if $\sigma_c = \sigma_p$.

We can think of the implied volatility $\sigma(T,K)$ as the volatility that one must insert into the Black–Scholes Call formula $C^{BS}(t,S_t;\sigma(T,K),T,K)$ in order for it to equal the market price $C_t(T,K)$ of a Call. There

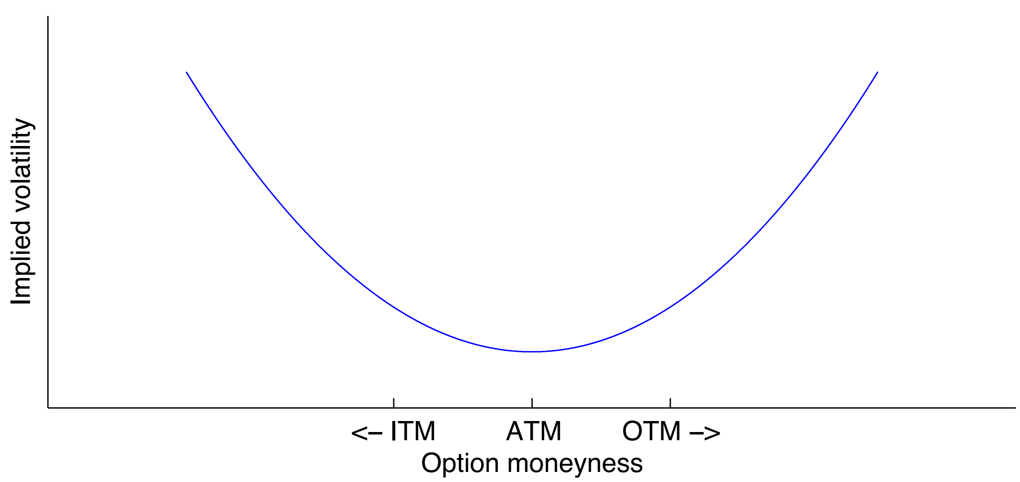

Figure 6.2: This is an illustration of the so called volatility smile. The implied volatility tends to be lower for at-the-money (ATM) options, and then increase the further into the money (ITM) or out of the money (OTM) the option gets.

is no analytical solution $\sigma(T, K)$ of (6.5.1). Thus, one must find $\sigma(T, K)$ numerically.

It is important to recognize that the Black-Scholes Call and Put prices are derived assuming a very particular *model* (6.0.1) for S, namely, that S is a geometric Brownian motion. *If* this model captured the true dynamics of S, the implied volatility $\sigma(T, K)$ for all (T, K) would equal the volatility σ of S in (6.0.1). However, when we use market options data to find the implied volatilities $\sigma(T, K)$ for different maturities T and strikes K, we observe that implied volatility $\sigma(T, K)$ is *not* constant. For a fixed T, what we typically observe is that the implied volatility $\sigma(T, K)$ is a convex function of K, as illustrated in Figure 6.2

Thus, we have strong evidence that the the Black-Scholes does *not* correctly capture the dynamics S. Nevertheless, the Black-Scholes model is a good first approximation for the dynamics of S. And, having a solid understanding of how to price and replicate options in a Black-Scholes setting helps us to understand how to price and replicate options in more complicated settings. Moreover, option prices are typically quoted in units of Black-Scholes implied volatility rather than in units of dollars and cents.

The implied volatility surface $(T, K) \to \sigma(T, K)$ is not completely arbitrary. $\sigma(T, K)$ must satisfy certain constraints. For example, using the fact that Call prices a strictly decreasing in K and Put prices are strictly increasing in K, we can find bounds of the slope of the implied volatility. We have

$$\frac{\partial C_t(T,K)}{\partial K} = \underbrace{\frac{\partial C^{BS}}{\partial K}}_{<0} + \underbrace{\frac{\partial C^{BS}}{\partial \sigma}}_{>0} \frac{\partial \sigma}{\partial K} \leq 0 \qquad \Rightarrow \qquad \frac{\partial \sigma}{\partial K} \leq \frac{-\partial C^{BS}/\partial K}{\partial C^{BS}/\partial \sigma} \tag{6.5.2}$$

where we have used the short-hand $C^{BS} = C^{BS}(t, S_t; \sigma, T, K)$. We can derive a lower bound for $\partial\sigma(T, K)/\partial K$ using Puts. We have

$$\frac{\partial P_t(T,K)}{\partial K} = \underbrace{\frac{\partial P^{BS}}{\partial K}}_{>0} + \underbrace{\frac{\partial P^{BS}}{\partial \sigma}}_{>0} \frac{\partial \sigma}{\partial K} \geq 0, \qquad \Rightarrow \qquad \frac{\partial \sigma}{\partial K} \geq \frac{-\partial P^{BS}/\partial K}{\partial P^{BS}/\partial \sigma}, \tag{6.5.3}$$

where we have used the short-hand $P^{BS} = P^{BS}(t, S_t; \sigma, T, K)$. From (6.5.2), (6.5.3) and the Black-Scholes formulas for Calls (6.4.1) and Puts (6.4.2), we obtain

$$\frac{-\sqrt{2\pi}}{S_t\sqrt{T-t}}(1 - \Phi(d_-))e^{\frac{1}{2}d_+^2} \leq \frac{\partial \sigma}{\partial K} \leq \frac{\sqrt{2\pi}}{S_t\sqrt{T-t}}\Phi(d_-)e^{\frac{1}{2}d_+^2}.$$

where we have take $r = 0$ for simplicity. Keep in mind that $d_\pm$ depends on (t, S_t, σ, T, K), as indicated in (6.4.3). Thus, the bound on the slope of implied volatility depends on the level of implied volatility.

6.6 Properties of the Black-Scholes Formula

Independent of the model we use for the underlying stock S, as $S_0/K \to \infty$, we expect $\widetilde{\mathbb{P}}(S_T \geq K) \approx 1$. And thus, we have

$$\lim_{S_0/K \to \infty} C_0(T,K) = e^{-rT}\widetilde{\mathbb{E}}(S_T - K)^+ \approx e^{-rT}\widetilde{\mathbb{E}}\,(S_T - K) = S_0 - Ke^{-rT}.$$

Now, consider the Black-Scholes price of a Call option when $t = 0$:

$$C^{BS}(0, S_0; \sigma, T, K) = S_0\Phi(d_+) - e^{-rT}K\Phi(d_-), \qquad d_\pm = \frac{\log(S_0/K) - (r \pm \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}.$$

As $S_0/K \to \infty$, we have

$$\begin{aligned}
\lim_{S_0/K \to \infty} C_0 &= S_0 \lim_{S_0/K \to \infty} \Phi(d_+) - e^{-rT}K \lim_{S_0/K \to \infty} \Phi(d_-) \\
&= S_0\Phi(\lim_{S_0/K \to \infty} d_+) - e^{-rT}K\Phi(\lim_{S_0/K \to \infty} d_-) \\
&= S_0 - e^{-rT}K,
\end{aligned}$$

where we have used $\Phi(\lim_{S_0/K\to\infty} d_\pm) = \Phi(\infty) = 1$. Thus, the Black-Scholes Call price agrees with our intuition.

Next, as the volatility of a stock goes to zero $\sigma \to 0$, we expect the drift of the stock to go to the risk-free rate of interest $\mu \to r$. Thus, we have

$$\lim_{\sigma \to 0} S_T = S_0 e^{rT} \qquad \Rightarrow \qquad \lim_{\sigma \to 0} C_0(T,K) = e^{-rT}\widetilde{\mathbb{E}} \lim_{\sigma \to 0}(S_T - K)^+ = (S_0 - K^{-rT})^+.$$

Now, noting that

$$\lim_{\sigma \to 0} d_\pm(0, S_0; \sigma, T, K) = \lim_{\sigma \to 0} \frac{\log(S_0/K) + (r \pm \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = \begin{cases} \infty & \text{if } S_0 e^{rT} > K, \\ -\infty & \text{if } S_0 e^{rT} < K. \end{cases}$$

It follows that

$$\begin{aligned}
\lim_{\sigma \to 0} C^{BS}(0, S_0; \sigma, T, K) &= \begin{cases} S_0 - e^{-rT}K & \text{if } S_0 e^{rT} > K, \\ 0 & \text{if } S_0 e^{rT} < K, \end{cases} \\
&= (S_0 - e^{-rT}K)^+.
\end{aligned}$$

Thus, we since once again that the Black-Scholes Call price agrees with our intuition.

We can derive the corresponding limits for Black-Scholes Put prices using Put/Call parity:

$$e^{-rT}K + C^{BS}(0, S_0; \sigma, T, K) = S_0 + P^{BS}(0, S_0; \sigma, T, K). \tag{6.6.1}$$

Using (6.6.1), as $S_0/K \to \infty$ we find

$$\begin{aligned}\lim_{S_0/K\to\infty} P^{BS}(0,S_0;\sigma,T,K) &= \lim_{S_0/K\to\infty} C^{BS}(0,S_0;\sigma,T,K) - S_0 + e^{-rT}K \\ &= S_0 - Ke^{-rT} - S_0 + e^{-rT}K = 0,\end{aligned}$$

and as $\sigma \to 0$, we obtain

$$\begin{aligned}\lim_{\sigma\to 0} P^{BS}(0,S_0;\sigma,T,K) &= \lim_{\sigma\to 0} C^{BS}(0,S_0;\sigma,T,K) + e^{-rT}K - S_0 \\ &= \begin{cases} S_0 - e^{rT}K + e^{-rT}K - S_0 = 0 & \text{if } S_0 e^{rT} > K, \\ 0 + e^{-rT}K - S_0 = e^{-rT}K - S_0 & \text{if } S_0 e^{rT} < K, \end{cases} \\ &= (e^{-rT}K - S_0)^+.\end{aligned}$$

You should check to see if these limits agree with your intuition.

6.7 Exercises

Exercise 6.1. Suppose that the stock price S and the bond price B under the physical measure $\mathbb{P}$ is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \qquad dB_t = rB_t dt.$$

where W is a Brownian motion under $\mathbb{P}$.

(a) Consider a claim that pays $\phi(S_T) = S_T^p$ at time T. Derive an explicit expression for $V_t = v(t, S_t)$, the value of the claim at time $t < T$.

(b) Let X be a portfolio that replicates the option payoff

$$dX_t = \Delta_t dS_t + (X_t - \Delta_t S_t)\frac{1}{B_t}dB_t.$$

Derive an explicit expression for Δ_t as a function of (t, S_t). Also, derive an expression for X_0 as a function of S_0.

(c) Write a Monte Carlo algorithm to plot sample paths of B and S. Your inputs should be $(S_0, B_0, \mu, \sigma, r, T, N)$ where N is the number of time steps in your Monte Carlo simulation (i.e., your time step should be $\delta t = T/N$). The output of your program should be a plot of a paths of S and B. *Hint*: you may wish to simulate $Z := \log S$ and $Y = \log B$ rather than S and B and then set $S_t = e^{Z_t}$ and $B_t = e^{Y_t}$.

(d) Modify your Monte Carlo algorithm to plot sample paths of V and X. Your inputs should be $(S_0, B_0, \mu, \sigma, r, T, p, N)$. The output of your program should be a plot of a paths of V and X. *Hint*: if

the paths of V and X are not almost identical, then there is an error in your program.
(e) Explain why the paths of V and X do not overlap exactly.
(f) Using your program from part (d), run M sample paths and plot a histogram of $X_T - V_T$.
Instructions: Parts (c), (d), and (f) ask you to perform simulations and plot the results. For these parts, use the following parameters

$$S_0 = 1, \quad B_0 = 1, \quad \mu = 0.1, \quad \sigma = 0.2, \quad r = 0.03, \quad T = 1. \quad p = -2, \quad N = 100, \quad M = 1000.$$

You should turn in your homework pdf, code and plots in a single zip file. The code for each part should be in a separate file i.e. you should turn in three distinct source code files. The filenames of your source code and figures should clearly indicate to which part (c, d or f) they correspond. Your code should run in batch mode called from the command line and produce your plots. For example, if you write your code using R language then it should run using the commands Rscript yourRfile.R or R CMD BATCH yourRfile.R, Python code should run with python yourPYfile.py etc. Your programs should read the input from a text file called input.txt which will be located in the directory with your code. The input file input.txt will consist of a single line listing $S_0, B_0, \mu, \sigma, r, T, p, N, M$ in that order, separated by commas. Your output should be appropriately named plots. Note that some of your programs will not use all of the inputs (e.g., only part (f) requires M. Nevertheless, you should only have *one* input.txt file.

Exercise 6.2. Consider the Black-Scholes model for a stock S and a bond B

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \qquad\qquad dB_t = r B_t dt.$$

Suppose that, if you own one share of S then you receive a cash flow of $q S_t dt$ during the time interval dt.
(a) Write the dynamics of a portfolio X of an investor who puts Δ_t shares of S_t at time t.
(b) Consider a claim that pays $\phi(S_T)$ at time T. Let $V_t = v(t, S_t)$ be the price of this claim at time t. Using a replication argument, derive a PDE for the function $v(t, s)$. Also state the terminal condition for v at time T.
(c) When dividends are involved, the no-arbitrage condition is that (X/B) is a martingale for *all* portfolios X under the risk-neutral measure $\widetilde{\mathbb{P}}$. If we have

$$dS_t = \widetilde{\mu} S_t dt + \sigma S_t d\widetilde{W}_t,$$

under $\widetilde{\mathbb{P}}$, then what should $\widetilde{\mu}$ be so that (X/B) is a martingale? Again, do not forget the S pays dividends.

Exercise 6.3. *Challenge Question*: Consider a global market consisting of a domestic bond B (in Dollars) and foreign bond C (in Euros) and an exchange rate Q (in Dollars per Euro), whose dynamics under the physics measure $\mathbb{P}$ are given by

$$dB_t = r B_t dt,$$

$$\mathrm{d}C_t = RC_t\mathrm{d}t,$$
$$\mathrm{d}Q_t = \mu Q_t\mathrm{d}t + \sigma Q_t\mathrm{d}W_t.$$

Derive an expression for the value (in Dollars) of an option that pays $\phi(Q_T C_T)$ in Dollars at time T. Derive a PDE for the option value.

Exercise 6.4. Derive the Black-Scholes price of a Call option assuming that the underlying stock S pays a dividend of $qS_t\mathrm{d}t$ during the interval $[t, t+\mathrm{d}t)$. *Hint*: If a trader holds a single share of the stock S at all times, then the change in the value of his portfolio is $\mathrm{d}X_t = \mathrm{d}S_t + qS_t\mathrm{d}t$.

Chapter 7

Greeks

The *Greeks* are (in addition to being an ancient civilization and the inventors of democracy), the sensitivities of portfolios or options with respect to objects that affect the portfolios' or options' values such as, e.g., the value of the underlying S, the time t, or the volatility σ. If this seems a bit abstract at the moment, things will become more clear as we go through some examples.

Throughout this chapter, we will consider a stock (or some other underlyer) $\mathrm{S} = (\mathrm{S}_t)_{t \geq 0}$ whose dynamics under the physical measure $\mathbb{P}$ are of the form

$$\mathrm{dS}_t = \mu_t \mathrm{S}_t \mathrm{d}t + \sigma_t \mathrm{S}_t \mathrm{dW}_t,$$

where $\mu = (\mu_t)_{t \geq 0}$ and $\sigma = (\sigma_t)_{t \geq 0}$ are some stochastic Markov processes. The value $\mathrm{V} = (\mathrm{V}_t)_{t \geq 0}$ of a derivative in this setting would be a function of time, the stock's price, and the instantaneous level of volatility: $\mathrm{V}_t = v(t, \mathrm{S}_t, \sigma_t)$. Calls and Puts would additionally be parametrized by their strikes and maturities dates. Thus, we have

$$\mathrm{C}_t(\mathrm{T}, \mathrm{K}) = \mathrm{C}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}), \qquad \mathrm{P}_t(\mathrm{T}, \mathrm{K}) = \mathrm{P}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}).$$

7.1 Delta

Consider a portfolio $\mathrm{X} = (\mathrm{X}(t, \mathrm{S}_t, \sigma_t))_{t \geq 0}$ of traded assets. For example, a portfolio X could consist of α shares of a stock S, β Call options with maturity T and strike K, and γ put options with the same strike and maturity. In this case, we have

$$\mathrm{X}(t, \mathrm{S}_t, \sigma_t) = \alpha \mathrm{S}_t + \beta \mathrm{C}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}) + \gamma \mathrm{P}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}). \tag{7.1.1}$$

We define the *Delta* $\Delta = (\Delta_t)_{t\geq 0}$ of a portfolio to be the derivative of X with respect to the stock price S. That is

$$\Delta_t := \frac{\partial \mathrm{X}}{\partial s}(t, \mathrm{S}_t, \sigma_t).$$

For example, if X is given by (7.1.1), then we have

$$\begin{aligned}
\Delta_t &= \frac{\partial}{\partial s}\mathrm{X}(t, \mathrm{S}_t, \sigma_t) \\
&= \frac{\partial}{\partial s}\left(\alpha \mathrm{S}_t + \beta \mathrm{C}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}) + \gamma \mathrm{P}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K})\right) \\
&= \alpha + \beta \frac{\partial}{\partial s}\mathrm{C}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}) + \gamma \frac{\partial}{\partial s}\mathrm{P}(t, \mathrm{S}_t, \sigma_t; \mathrm{T}, \mathrm{K}).
\end{aligned}$$

As previously mentioned, for a portfolio consisting of a single Call option, we have in the Black-Scholes setting that

$$\Delta_t^{\text{BS-Call}} = \frac{\partial \mathrm{C}^{\text{BS}}}{\partial s}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \Phi(d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})).$$

7.2 Gamma

We define the *Gamma* $\Gamma = (\Gamma_t)_{t\geq 0}$ of a portfolio to be the second derivative of X with respect to the stock price S. That is

$$\Gamma_t := \frac{\partial^2 \mathrm{X}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t) = \frac{\partial \Delta_t}{\partial s},$$

where, in the second equality, we have used $\Delta_t := \partial \mathrm{X}(t, \mathrm{S}_t, \sigma_t)/\partial s$. The Gamma of a single Call option in the Black-Scholes framework is given by

$$\Gamma_t^{\text{BS-Call}} = \frac{\partial^2 \mathrm{C}^{\text{BS-Call}}}{\partial s^2}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \frac{\phi(d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}))}{\mathrm{S}_t \sigma \sqrt{\mathrm{T} - t}}, \qquad \phi(d) = \frac{\mathrm{e}^{-d^2/2}}{\sqrt{2\pi}},$$

In order to hedge a call option, one has to adjust the Δ as S moves randomly. The Γ tells us how much Δ changes for a given change in S. Note from (6.4.3) that

$$\begin{aligned}
&\text{if} \quad \log^2\left(\frac{\mathrm{S}_t}{\mathrm{K}}\right) < \sigma^2(\mathrm{T} - t), \quad &&\text{then} \quad &&\lim_{t\to\mathrm{T}} d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) \to 0, \\
& &&\text{and thus} \quad &&\lim_{t\to\mathrm{T}} \phi(d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})) \to 1/\sqrt{2\pi},
\end{aligned}$$

which implies that

$$\lim_{t\to\mathrm{T}} \Gamma^{\text{BS-Call}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}) = \lim_{t\to\mathrm{T}} \frac{\phi(d_+(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K}))}{\mathrm{S}_0 \sigma \sqrt{\mathrm{T} - t}} = \infty.$$

This presents a *huge* problem when trying to replicate a Call option. If $S_t \approx K$, then as $t \to T$ a small change in S results in a large change in $\Delta^{\text{BS-Call}}(t, S_t; \sigma, T, K)$. On the the other hand, if $\log^2 S_t/K > \sigma^2(T-t)$, then as $t \to T$ we have $\Gamma^{\text{BS-Call}}(t, S_t; \sigma, T, K) \to 0$, and thus, replication becomes nearly static.

7.3 Theta

Finally, we define the *Theta* $\Theta = (\Theta_t)_{t\geq 0}$ of a portfolio to be the derivative of X with respect to time t:

$$\Theta_t := \frac{\partial X}{\partial t}(t, S_t, \sigma_t).$$

In the Black-Scholes setting, the Θ of a single call options is given by

$$\begin{aligned}
\Theta_t^{\text{BS-Call}} &= \frac{\partial C^{\text{BS}}}{\partial t}(t, S_t; \sigma, T, K) \\
&= -S_t \phi(d_+(t, S_t; \sigma, T, K)) \frac{\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(d_-(t, S_t; \sigma, T, K)).
\end{aligned}$$

7.4 Greek-neutral Hedging

A portfolio X is said to be *Greek-neutral* if a given Greek of a portfolio is zero. For example

$$\begin{aligned}
&\Delta\text{-neutral}: && \Delta_t = \frac{\partial X}{\partial s}(t, S_t, \sigma) = 0, \\
&\Gamma\text{-neutral}: && \Gamma_t = \frac{\partial^2 X}{\partial s^2}(t, S_t, \sigma) = 0.
\end{aligned}$$

Let us see why it can be advantageous to have a portfolio that is neutral to certain Greeks.

Suppose a bank has sold a Call option C and owns α shares of S. Then, at time t, the value of the bank's portfolio is

$$X(t, S_t, \sigma_t) = -C(t, S_t, \sigma_t) + \alpha S_t,$$

where, to ease notation, we have omitted the dependence of the Call price on the strike T and maturity K. What would happen to the value of this portfolio if the stock price jumps from S_t to $S_t+\delta$ instantaneously? The change in the value of the portfolio is

$$\begin{aligned}
X(t, S_t+\delta, \sigma_t) - X(t, S_t, \sigma_t) &= -\left(C(t, S_t+\delta, \sigma_t) - C(t, S_t, \sigma_t)\right) + \alpha\left((S_t+\delta) - S_t\right) \\
&= -\left(\frac{\partial C}{\partial s}(t, S_t, \sigma_t)\delta + \frac{1}{2}\frac{\partial^2 C}{\partial s^2}(t, S_t, \sigma_t)\delta^2 + \mathcal{O}(\delta^3)\right) + \alpha\delta,
\end{aligned}$$

where, in the second equality, we have expanded $\mathrm{C}(t, \mathrm{S}_t + \delta, \sigma_t)$ as a Taylor series about S_t. Suppose now, that the bank had chosen α so that the portfolio X was Δ-neutral

$$0 = \Delta_t = \frac{\partial \mathrm{X}}{\partial s}(t, \mathrm{S}_t, \sigma_t) = -\frac{\partial \mathrm{C}}{\partial s}(t, \mathrm{S}_t, \sigma_t) + \alpha, \qquad \Rightarrow \qquad \alpha = \frac{\partial \mathrm{C}}{\partial s}(t, \mathrm{S}_t, \sigma_t).$$

In this case, the change in the value of the portfolio would have been

$$\mathrm{X}(t, \mathrm{S}_t + \delta, \sigma_t) - \mathrm{X}(t, \mathrm{S}_t, \sigma_t) = -\frac{1}{2}\frac{\partial^2 \mathrm{C}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t)\delta^2 + \mathcal{O}(\delta^3),$$

because the order $\mathcal{O}(\delta)$ terms have canceled each other out. Thus, a small change in the stock price δ only affect the portfolio value at order $\mathcal{O}(\delta^2)$. Remember for $\delta < 1$ we have $\delta^2 < \delta$. Unfortunately, for the bank, call prices are convex as a function of S (this is true in general, not just in the Black-Scholes framework). Thus, for a sufficiently small δ we have

$$\frac{\partial^2 \mathrm{C}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t) > 0, \qquad \Rightarrow \qquad \mathrm{X}(t, \mathrm{S}_t + \delta, \sigma_t) - \mathrm{X}(t, \mathrm{S}_t, \sigma_t) = -\frac{1}{2}\frac{\partial^2 \mathrm{C}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t)\delta^2 + \mathcal{O}(\delta^3) < 0.$$

So, if the stock jumps in from S_t to $\mathrm{S}_t + \delta$ the bank loses money. If the bank had constructed a portfolio that was both Δ-neutral *and* Γ-neutral, the gains or losses of the portfolio would have been pushed to $\mathcal{O}(\delta^3)$.

Example 7.1. Suppose a bank sells a Call, owns α shares of S and owns β of a put options. Omitting the dependence of the Call and Puts on their maturity dates and strikes, we have

$$\mathrm{X}(t, \mathrm{S}_t, \sigma_t) = -\mathrm{C}(t, \mathrm{S}_t, \sigma_t) + \alpha \mathrm{S}_t + \beta \mathrm{P}(t, \mathrm{S}_t, \sigma_t).$$

How should the bank choose α and β so that the portfolio is both Δ-neutral and Γ-neutral? We would like

$$\begin{aligned}
0 = \Delta_t &= \frac{\partial \mathrm{X}}{\partial s}(t, \mathrm{S}_t, \sigma_t) = -\frac{\partial \mathrm{C}}{\partial s}(t, \mathrm{S}_t, \sigma_t) + \alpha + \beta\frac{\partial \mathrm{P}}{\partial s}(t, \mathrm{S}_t, \sigma_t), \\
0 = \Gamma_t &= \frac{\partial^2 \mathrm{X}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t) = -\frac{\partial^2 \mathrm{C}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t) + \beta\frac{\partial^2 \mathrm{P}}{\partial s^2}(t, \mathrm{S}_t, \sigma_t).
\end{aligned}$$

To make the portfolio Γ-neutral, the bank should choose

$$\beta = \frac{\partial^2 \mathrm{C}/\partial s^2(t, \mathrm{S}_t, \sigma_t)}{\partial^2 \mathrm{P}/\partial s^2(t, \mathrm{S}_t, \sigma_t)}. \tag{7.4.1}$$

And, to make the portfolio Δ-neutral, the bank should choose

$$\alpha = \frac{\partial \mathrm{C}}{\partial s}(t, \mathrm{S}_t, \sigma_t) - \beta\frac{\partial \mathrm{P}}{\partial s}(t, \mathrm{S}_t, \sigma_t),$$

where β is given by (7.4.1) If the bank does this, then as the stock moves from S_t to $\mathrm{S}_t + \delta$, the change in the value of the portfolio will be $\mathcal{O}(\delta^3)$.

Example 7.2. Suppose a bank owns a call $\mathrm{C}(\mathrm{T},\mathrm{K})$, sells a put $\mathrm{P}(\mathrm{T},\mathrm{K})$ and sells one share of the underlying S. The value of the portfolio at time t is

$$\mathrm{X}(t,\mathrm{S}_t,\sigma) = \mathrm{C}(t,\mathrm{S}_t,\sigma_t;\mathrm{T},\mathrm{K}) - \mathrm{P}(t,\mathrm{S}_t,\sigma_t;\mathrm{T},\mathrm{K}) - \mathrm{S}_t.$$

We claim that this portfolio is Δ- and Γ-neutral. Why? Independent of the model for S, we have from put-call parity that

$$\mathrm{C}(t,\mathrm{S}_t,\sigma_t;\mathrm{T},\mathrm{K}) - \mathrm{P}(t,\mathrm{S}_t,\sigma_t;\mathrm{T},\mathrm{K}) = \mathrm{S}_t - \mathrm{K}\mathrm{e}^{-r(\mathrm{T}-t)}$$

Hence, we have

$$\mathrm{X}(t,\mathrm{S}_t,\sigma_t) = -\mathrm{K}\mathrm{e}^{-r(\mathrm{T}-t)}, \qquad \Rightarrow \qquad \frac{\partial^n \mathrm{X}}{\partial s^n}(t,\mathrm{S}_t,\sigma_t) = 0, \qquad \text{for all } n \geq 1.$$

7.5 Relationship between Δ, Θ and Γ

Suppose S is a geometric Brownian motion (as in the Black-Scholes framework). Let X be a portfolio of derivatives written on S. The value of the portfolio at time t is

$$\mathrm{X}(t,\mathrm{S}_t) = \sum_i \alpha_i u_i(t,\mathrm{S}_t),$$

where $u_i(t,\mathrm{S}_t)$ is the price of the ith derivative and α_i is the number of shares of the ith derivative in the portfolio. Each u_i must satisfy the Black–Scholes PDE:

$$0 = \left(\partial_t + rs\partial_s + \frac{1}{2}\sigma^2 s^2 \partial_s^2 - r\right) u_i(t,s).$$

Note that the price of a bond $\mathrm{B}_t = \mathrm{e}^{-r(\mathrm{T}-t)}$ also satisfies this PDE, as can easily be checked by direct computation

$$0 = \left(\partial_t + rs\partial_s + \frac{1}{2}\sigma^2 s^2 \partial_s^2 - r\right) \mathrm{e}^{-r(\mathrm{T}-t)}.$$

Let us define the Black-Scholes operator

$$\mathcal{A}^{\mathrm{BS}} := \left(\partial_t + rs\partial_s + \frac{1}{2}\sigma^2 s^2 \partial_s^2 - r\right).$$

By linearity we have

$$\mathcal{A}^{\mathrm{BS}}\mathrm{X}(t,\mathrm{S}_t) = \sum_i \alpha_i \mathcal{A}^{\mathrm{BS}} u_i(t,\mathrm{S}_t) = 0.$$

Thus, we see that X satisfies the Black-Scholes PDE:

$$0 = \left(\partial_t + r\mathrm{S}_t\partial_s + \frac{1}{2}\sigma^2\mathrm{S}_t^2\partial_s^2 - r\right)\mathrm{X}(t,\mathrm{S}_t) = 0.$$

Writing the partial derivatives in terms of the Greeks, we have

$$0 = \Theta_t^{\mathrm{BS}} + r\mathrm{S}_t\Delta_t^{\mathrm{BS}} + \frac{1}{2}\sigma^2\mathrm{S}_t^2\Gamma_t^{\mathrm{BS}} - r\mathrm{X}^{\mathrm{BS}}(t,\mathrm{S}_t).$$

If the portfolio is Δ-neutral (which can be achieved by choosing the α_i's so that $\sum_i \alpha_i\partial_s u_i(t,\mathrm{S}_t) = 0$), then

$$0 = \Theta_t^{\mathrm{BS}} + \frac{1}{2}\sigma^2\mathrm{S}_t^2\Gamma_t^{\mathrm{BS}} - r\mathrm{X}^{\mathrm{BS}}(t,\mathrm{S}_t).$$

If $r \approx 0$ it follows that

$$\Theta_t^{\mathrm{BS}} \approx -\frac{1}{2}\sigma^2\mathrm{S}_t^2\Gamma_t^{\mathrm{BS}}.$$

Thus, if the portfolio is Δ-neutral, then the change in the value of the portfolio is opposite the sign of Γ. Although we have derived the above in the Black-Scholes setting, the result is roughly true in any diffusion model for S.

7.6 Vega

Let us now consider a single Call or Put option with the same strike T and maturit K. We can always express the value of a Call or Put through their implied volatility $\sigma_t(\mathrm{T},\mathrm{K})$. We have

$$\mathrm{C}_t(\mathrm{T},\mathrm{K}) = \mathrm{C}^{\mathrm{BS}}(t,\mathrm{S}_t;\sigma_t(\mathrm{T},\mathrm{K}),\mathrm{T},\mathrm{K}), \qquad \mathrm{P}_t(\mathrm{T},\mathrm{K}) = \mathrm{P}^{\mathrm{BS}}(t,\mathrm{S}_t;\sigma_t(\mathrm{T},\mathrm{K}),\mathrm{T},\mathrm{K}).$$

We define the *Vega* $\mathcal{V} = (\mathcal{V}_t)_{t\geq 0}$ of Call or Put to be the derivative of the Call or Put price with respect to the implied volatility:

$$\mathcal{V} := \frac{\partial\mathrm{C}^{\mathrm{BS}}}{\partial\sigma}(t,\mathrm{S}_t;\sigma_t(\mathrm{T},\mathrm{K}),\mathrm{T},\mathrm{K}) = \frac{\partial\mathrm{P}^{\mathrm{BS}}}{\partial\sigma}(t,\mathrm{S}_t;\sigma_t(\mathrm{T},\mathrm{K}),\mathrm{T},\mathrm{K}).$$

Note that "Vega" is not actually a Greek letter. But is sounds Greek! Because implied volatility is a stochastic process, a trader may wish to know how the value of a Call or Put in his portfolio changes when $\sigma_t(\mathrm{T},\mathrm{K})$ moves. In the Black–Scholes setting, the $\mathcal{V}$ of a single Call option satisfies

$$\mathcal{V}_t^{\text{BS-Call}} = \frac{\partial\mathrm{C}^{\mathrm{BS}}}{\partial\sigma}(t,\mathrm{S}_t;\sigma,\mathrm{T},\mathrm{K}) = (\mathrm{T}-t)\sigma\mathrm{S}_t^2\frac{\partial\mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t,\mathrm{S}_t;\sigma,\mathrm{T},\mathrm{K}) = (\mathrm{T}-t)\sigma\mathrm{S}_t^2\Gamma_t^{\text{BS-Call}}.$$

Thus, there is a relationship between the $\mathcal{V}^{\text{BS-Call}}$ and the $\Gamma^{\text{BS-Call}}$. In fact, the above relation holds for *all* European options – not just Calls.

7.7 Exercises

Exercise 7.1. This exercise is meant to give you an intuitive feel for the Greeks. Assume for simplicity that the risk-free rate of interest is zero $r = 0$. Suppose S is modeled as in the Black-Scholes model $\mathrm{dS}_t = \mu \mathrm{S}_t \mathrm{d}t + \sigma \mathrm{S}_t \mathrm{dW}_t$. The value of a call option at time t given $\mathrm{S}_t = s$ is given by

$$\mathrm{C}^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K}) = s\Phi(d_+) - \mathrm{K}\Phi(d_-), \qquad d_\pm = \frac{1}{\sigma\sqrt{\mathrm{T}-t}}\left(\log\frac{s}{\mathrm{K}} \pm \frac{1}{2}\sigma^2(\mathrm{T}-t)\right).$$

For all of the plots in this problem, assume $\sigma = 0.2$, $\mathrm{K} = 1$, $\mathrm{T} = 1$. For computational work, leave σ, K, and T as unknown constants.

(a) On a single graph, plot $\mathrm{C}^{\mathrm{BS}}(t, s; \mathrm{T}, \mathrm{K})$ as a function of s for $t = 0$, $t = \mathrm{T}/2$ and $t = 3\mathrm{T}/4$. Also plot the call payoff $(s - \mathrm{K})^+$ as a function of s on the same graph.

(b) Derive an expression for the $\Delta^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K})$ by direct computation (do *not* just copy formulas from the notes). On a single graph, plot $\Delta^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K})$ as a function of s for $t = 0$, $t = \mathrm{T}/2$ and $t = 3\mathrm{T}/4$.

(c) Derive an expression for the $\Gamma^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K})$ by direct computation (do *not* just copy formulas from the notes). On a single graph, plot $\Gamma^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K})$ as a function of s for $t = 0$, $t = \mathrm{T}/2$ and $t = 3\mathrm{T}/4$.

(d) Derive an expression for the the Vega $\mathcal{V}^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K})$ by direct computation (do *not* just copy formulas from the notes). Show that $\mathcal{V}^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K}) = (\mathrm{T} - t)\sigma s^2 \Gamma^{\mathrm{BS}}(t, s; \sigma, \mathrm{T}, \mathrm{K})$.

Exercise 7.2. Assume for simplicity that the risk-free rate of interest is zero $r = 0$. Suppose that the price of call options are given by $\mathrm{C}(t, \mathrm{S}_t; \mathrm{T}, \mathrm{K})$ for some function C. Suppose further that you can observe the implied volatility smile $\sigma(\mathrm{T}, \mathrm{K})$ at every T and K. In particular, assume that $\sigma(\mathrm{T}, \mathrm{K}) = \mathrm{I}(\mathrm{T}-t, \log(\mathrm{K}/\mathrm{S}_t))$ for some function $\mathrm{I}(\tau, x)$. Let $\mathrm{C}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})$ denote the Black-Scholes price of a call option at time t with maturity T and strike K, assuming a fixed volatility σ. And let $\Delta^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})$, $\Theta^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})$ and $\mathcal{V}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma, \mathrm{T}, \mathrm{K})$ denote the Black-Scholes Delta, Theta and Vega respectively.

(a) Derive an expression for the Δ of a call option. Your answer should involve Black-Scholes Greeks and derivatives of I.

(b) Derive an expression for the Θ of a call option. Your answer should involve Black-Scholes Greeks and derivatives of I.

Exercise 7.3. Assume that $\mathrm{dS}_t = \mu \mathrm{S}_t \mathrm{d}t + \sigma \mathrm{S}_t \mathrm{dW}_t$ and $\mathrm{dB}_t = r\mathrm{B}_t \mathrm{d}t$. Suppose you have a portfolio X that consists of a α shares of a stock S, β call options C^{BS}, γ put options P^{BS} and δ bonds B. Do *not* assume $r = 0$. Note that the call and put options are written on S.

(a) Suppose the Call and the Put option have the same strike and maturity. Is it possible to have a portfolio X that is Γ-neutral but is not $\mathcal{V}$-neutral?

(b) Consider a portfolio Y that has the following greeks: Δ^{Y}, Γ^{Y}, $\mathcal{V}^{\mathrm{Y}}$ and Θ^{Y}. Find a system of equations for α, β, γ and δ so that the portfolio $(\mathrm{X} + \mathrm{Y})$ is Δ-neutral, Γ-neutral, Θ-neutral and $\mathcal{V}$-neutral. Note that the call and put may have different strikes and maturities. BONUS (not required): Find α, β, γ and δ. Your answer should involve partial derivatives of P and C with respect to s, σ and t.

Chapter 8

Stochastic volatility

If we assume that stock prices are strictly positive processes and have continuous sample paths, then the most general model for S is

$$\mathrm{d}S_t = \mu_t S_t \mathrm{d}t + \sigma_t S_t \mathrm{d}W_t, \tag{8.0.1}$$

where $W = (W_t)_{t \geq 0}$ is a Brownian motion under the real-world probability measure $\mathbb{P}$ and $\mu = (\mu_t)_{t \geq 0}$ and $\sigma = (\sigma_t)_{t \geq 0}$ are stochastic processes adapted to a filtration $\mathcal{F}$ of W. In the Black-Scholes framework, the volatility process σ was constant. In this chapter, we will explore some of the consequences of allowing for the possibility that σ is stochastic.

8.1 Δ-hedging when volatility is stochastic

Suppose that, under $\mathbb{P}$, a stock S has the following dynamics given by (8.0.1). Suppose, further, that the market prices a call option *as if* volatility were a constant

$$\mathrm{d}S_t = \mu_t S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t. \qquad \text{(assumption of Market participants)} \tag{8.1.1}$$

Note that (8.1.1) implies that the price of a Call option is the Black-Scholes price $C_t(T, K) = C^{BS}(t, S_t; \sigma, T, K)$. If a banker sells a call option $C^{BS}(t, S_t, \sigma)$ and Δ-hedges, what will the profit or loss (P&L) of his portfolio be at maturity? For simplicity, let us assume that $r = 0$. The dynamics of the banker's portfolio X are

$$\begin{aligned}
\mathrm{d}X_t &= \Delta_t \mathrm{d}S_t + (X_t - \Delta_t S_t)\frac{1}{B_t}\mathrm{d}B_t \\
&= \frac{\partial C^{BS}}{\partial s}(t, S_t; \sigma)\mathrm{d}S_t
\end{aligned}$$

where, in the second equality, we have used $\Delta_t = \partial_s \mathrm{C}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma)$ and $\mathrm{dB}_t = 0$. The dynamics of the Call option are

$$\begin{aligned}
\mathrm{dC}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma) &= \frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial t}(t, \mathrm{S}_t; \sigma)\mathrm{d}t + \frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial s}(t, \mathrm{S}_t; \sigma)\mathrm{dS}_t + \frac{1}{2}\frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)(\mathrm{dS}_t)^2 \\
&= \frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial t}(t, \mathrm{S}_t; \sigma)\mathrm{d}t + \frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial s}(t, \mathrm{S}_t; \sigma)\mathrm{dS}_t + \frac{1}{2}\frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)\sigma_t^2 \mathrm{S}_t^2 \mathrm{d}t,
\end{aligned}$$

where we have omitted the (T, K)-dependence, which is not important for the computation we are carrying out. Combining the three above equations, we obtain

$$\begin{aligned}
\mathrm{dX}_t - \mathrm{dC}^{\mathrm{BS}}(t, \mathrm{S}_t; \sigma) &= -\left[\frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial t}(t, \mathrm{S}_t; \sigma) + \frac{1}{2}\frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)\sigma_t^2 \mathrm{S}_t^2\right]\mathrm{d}t \\
&= -\left[\frac{\partial \mathrm{C}^{\mathrm{BS}}}{\partial t}(t, \mathrm{S}_t; \sigma) + \frac{1}{2}\frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)\sigma^2 \mathrm{S}_t^2\right]\mathrm{d}t \\
&\quad + \frac{1}{2}(\sigma^2 - \sigma_t^2)\mathrm{S}_t^2 \frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)\mathrm{S}_t^2 \mathrm{d}t \\
&= \frac{1}{2}(\sigma^2 - \sigma_t^2)\mathrm{S}_t^2 \frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)\mathrm{S}_t^2 \mathrm{d}t,
\end{aligned}$$

where, in the 2nd equality, the term in square brackets is equal to zero $[\ldots] = 0$ because C^{BS} satisfies the Black-Scholes PDE. Finally, we obtain

$$\begin{aligned}
\mathrm{P\&L}_\mathrm{T} &= \mathrm{X}_\mathrm{T} - \mathrm{C}^{\mathrm{BS}}(\mathrm{T}, \mathrm{S}_\mathrm{T}; \sigma) \\
&= \mathrm{X}_0 - \mathrm{C}^{\mathrm{BS}}(0, \mathrm{S}_0; \sigma) + \int_0^\mathrm{T} \left(\mathrm{dX}_t - \mathrm{dC}(t, \mathrm{S}_t; \sigma)\right) \\
&= \int_0^\mathrm{T} \left(\sigma^2 - \sigma_t^2\right) \frac{1}{2}\mathrm{S}_t^2 \frac{\partial^2 \mathrm{C}^{\mathrm{BS}}}{\partial s^2}(t, \mathrm{S}_t; \sigma)\mathrm{d}t,
\end{aligned}$$

where we have used the fact that $\mathrm{X}_0 = \mathrm{C}^{\mathrm{BS}}(0, \mathrm{S}_0; \sigma)$. Thus the final profit or loss depends on the fixed volatility σ that the market assumes, and the path that the real-world volatility process $(\sigma_t)_{0 \le t \le \mathrm{T}}$ actually takes. Note that $\partial_s^2 \mathrm{C}^{\mathrm{BS}} > 0$. Thus, for a banker that sells an option and Δ-hedges, he would prefer that the market prices options with a large σ.

8.2 Variance Swaps

As we showed in the previous section, Δ-hedging in an environment when volatility is stochastic leaves one exposed to fluctuations in volatility. The desire to reduce volatility risk has led to the development of various volatility derivatives. The simplest of these derivatives is the *variance swap*. Let us denote by

$\mathrm{VS} = (\mathrm{VS}_t)_{0 \le t \le T}$ the value of the long side of a variance swap. The payoff of the long side of a variance swap VS_T is

$$\mathrm{VS}_T = \sum_{i=0}^{n-1} \left(\log S_{t_{i+1}} - \log S_{t_i} \right)^2 - K, \tag{8.2.1}$$

where $0 = t_0 < t_1 < t_2 < \ldots < t_n = T$ are some pre-specified dates, and K is a constant. Typically $\delta t := t_{i+1} - t_i$ is one trading day. The sum in (8.2.1) is referred to as the *floating leg* and the constant K is referred to as the *fixed leg*. Note that as $\delta t \to 0$ the floating leg converges to an integral

$$\lim_{\delta t \to 0} \sum_{i=0}^{n-1} \left(\log S_{t_{i+1}} - \log S_{t_i} \right)^2 = \int_0^T (\mathrm{d} \log S_t)^2. \tag{8.2.2}$$

Assuming S has dynamics of the form (8.0.1), we have from Itô's Lemma that

$$\mathrm{d} \log S_t = \left(\mu_t - \frac{1}{2}\sigma_t^2 \right) \mathrm{d}t + \sigma_t \mathrm{d}W_t \qquad \Rightarrow \qquad (\mathrm{d} \log S_t)^2 = \sigma_t^2 \mathrm{d}t. \tag{8.2.3}$$

Inserting (8.2.3) into (8.2.2) yields

$$\lim_{\delta t \to 0} \mathrm{VS}_T = \int_0^T \sigma_t^2 \mathrm{d}t - K. \tag{8.2.4}$$

Henceforth, we shall assume that the payoff of the long side of a variance swap is given by the right-hand side of (8.2.4). As with any swap, a variance swap has no cost to enter. The fair strike K is given by setting the initial value of the variance swap to zero and solving for K. We have

$$0 = \mathrm{VS}_0 = \frac{1}{B_T} \widetilde{\mathbb{E}} \left[\int_0^T \sigma_t^2 \mathrm{d}t - K \right] \qquad \Rightarrow \qquad K = \widetilde{\mathbb{E}} \left[\int_0^T \sigma_t^2 \mathrm{d}t \right], \tag{8.2.5}$$

where $\widetilde{\mathbb{E}}$ denotes an expectation taking with respect to the market's chosen risk-neutral measure $\widetilde{\mathbb{P}}$.

Let us see how we can replicate the floating leg $\int_0^T \sigma_t^2 \mathrm{d}t$ of a variance swap. For simplicity, we will once again assume that $r = 0$. Under the pricing measure $\widetilde{\mathbb{P}}$, the stock S has dynamics of the form

$$\mathrm{d}S_t = \sigma_t S_t \mathrm{d}\widetilde{W}_t,$$

where $\widetilde{W} = (\widetilde{W}_t)_{0 \le t \le T}$ is a Brownian motion under $\widetilde{\mathbb{P}}$. Using Itô's Lemma, we have

$$\mathrm{d} \log S_t = -\frac{1}{2}\sigma_t^2 \mathrm{d}t + \sigma_t \mathrm{d}\widetilde{W}_t \qquad \Rightarrow \qquad \sigma_t^2 \mathrm{d}t = 2\sigma_t \mathrm{d}\widetilde{W}_t - 2\mathrm{d} \log S_t = \frac{2}{S_t} \mathrm{d}S_t - 2\mathrm{d} \log S_t.$$

Integrating from 0 to T, we obtain

$$\int_0^T \sigma_t^2 \mathrm{d}t = \int_0^T \frac{2}{S_t} \mathrm{d}S_t - 2 \log \left(\frac{S_T}{S_0} \right). \tag{8.2.6}$$

Thus, to replicate the floating leg of a variance swap, one can do the following

1. At time $t = 0$, buy a European *log contract* with payoff

$$\phi(S_t) = -2\log\left(\frac{S_T}{S_0}\right),$$

2. For all $0 \le t \le T$ hold $\Delta_t = \frac{2}{S_t}$ shares of S.

3. Lend and borrow from the bank as needed to finance the position.

At maturity T, the above portfolio will have a value equal to the floating leg of the variance swap $\int_0^T \sigma_t^2 dt$. Note that, because $\widetilde{\mathbb{P}} \sim \mathbb{P}$, the replication strategy above holds under $\mathbb{P}$ as well.

Now, let us determine the fair strike K of a VS. From (8.2.5) and (8.2.6), we have

$$\begin{aligned} K = \widetilde{\mathbb{E}}\left[\int_0^T \sigma_t^2 dt\right] &= \widetilde{\mathbb{E}}\left[\int_0^T \frac{2}{S_t} dS_t\right] - 2\widetilde{\mathbb{E}}\left[\log\left(\frac{S_T}{S_0}\right)\right] \\ &= \widetilde{\mathbb{E}}\left[\int_0^T \frac{2}{S_t}\sigma_t S_t d\widetilde{W}_t\right] - 2\widetilde{\mathbb{E}}\left[\log\left(\frac{S_T}{S_0}\right)\right] = -2\widetilde{\mathbb{E}}\left[\log\left(\frac{S_T}{S_0}\right)\right]. \end{aligned}$$

The above result says that the fair strike K of a variance swap is equal to the price $-2\widetilde{\mathbb{E}}[\log(S_T/S_0)]$ of a (path-independent) European log contract payoff. In reality, one cannot buy log contracts in the market. However, one can create a synthetic log contract from European Calls and Puts.

Recall from Section 2.13 that, for any $\kappa > 0$ and any reasonable function ϕ we have

$$\begin{aligned} \phi(S_T) &= \phi(\kappa) + \phi'(\kappa)\Big((S_T - \kappa)^+ - (\kappa - S_T)^+\Big) \\ &\quad + \int_0^\kappa \phi''(K)(K - S_T)^+ dK + \int_\kappa^\infty \phi''(K)(S_T - K)dK. \end{aligned}$$

By taking the risk-neutral expectation of both sides of the above equation, we obtain

$$\begin{aligned} \widetilde{\mathbb{E}}[\phi(S_T)] &= \phi(\kappa) + \phi'(\kappa)\Big(C_0(T,\kappa) - P_0(T,\kappa)\Big) \\ &\quad + \int_0^\kappa \phi''(K)P_0(T,K)dK + \int_\kappa^\infty \phi''(K)C_0(T,K)dK, \end{aligned}$$

where we have used $\widetilde{\mathbb{E}}[(K-S_T)^+] = P_0(T,K)$ and $\widetilde{\mathbb{E}}[(S_T-K)^+] = C_0(T,K)$. Setting $\phi(S_T) = -2\log(S_T/S_0)$ and choosing $\kappa = S_0$ yields

$$\begin{aligned} \widetilde{\mathbb{E}}\left[-2\log\frac{S_0}{S_0}\right] &= -2\log\frac{S_T}{S_0} + \frac{-2}{S_0}\Big(C_0(T,S_0) - P_0(T,S_0)\Big) \\ &\quad + \int_0^{S_0} \frac{2}{K^2}P_0(T,K)dK + \int_{S_0}^\infty \frac{2}{K^2}C_0(T,K)dK \\ &= \int_0^{S_0} \frac{2}{K^2}P_0(T,K)dK + \int_{S_0}^\infty \frac{2}{K^2}C_0(T,K)dK, \end{aligned} \tag{8.2.7}$$

where we have used

$$\log(S_0/S_0) = 0 \qquad \text{and} \qquad C_0(T, S_0) - P_0(T, S_0) = \widetilde{\mathbb{E}}(S_T - S_0) = 0.$$

As $K = \widetilde{\mathbb{E}}[\int_0^T \sigma_t^2 dt] = \widetilde{\mathbb{E}}[-2\log(S_T/S_0)]$, we see from (8.2.7) that the market's forward looking expectation of realized volatility over the interval $[0, T]$ can be deduced by observing European Call and Put prices that mature at time T.

8.3 VIX

You may have heard of the CBOE's *volatility index* or VIX. The VIX is a discretized approximation of the integral given in (8.2.7). Roughly, we have

$$\mathrm{VIX}_t^2 = \sum_{K_i < S_0} \frac{2}{K_i^2} P_t(t + T, K_i) + \sum_{K_i > S_0} \frac{2}{K_i^2} C_t(t + T, K)\Delta K_i, \qquad \Delta K_i = K_{i+1} - K_i.$$

Thus, the VIX_t^2 is an approximate measure of the market's forward-looking risk-neutral expectation of integrated volatility $\widetilde{\mathbb{E}}[\int_t^{t+T} \sigma_s^2 ds | \mathcal{F}_t]$.

8.4 Exercises

Instructions: Exercise 8.1 asks you to write code that produces a plot. You should turn in your homework pdf, code and plots in a single zip file. Your code should run in batch mode called from the command line and produce your plots. For example, if you write your code using R language then it should run using the commands Rscript yourRfile.R or R CMD BATCH yourRfile.R, Python code should run with python yourPYfile.py etc. Your programs should read the input from a text file called input.txt which will be located in the directory with your code. The input file input.txt will consist of a single line listing $S_0, K, T, \sigma_1, \sigma_2, \mu, N, M$ in that order, separated by commas. Your output should be appropriately named plots.

Exercise 8.1. Assume the risk free rate of interest is zero $r = 0$. Suppose that the dynamics of S are given by $dS_t = \mu S_t dt + \sigma_1 S_t dW_t$. Suppose the the market prices options assuming that the dynamics of S are given by $dS_t = \mu S_t dt + \sigma_2 S_t dW_t$ where $\sigma_2 > \sigma_1$. Assume you *sell* a call option for the black-Scholes price $C^{BS}(t, S_t; T, K, \sigma_2)$ and you Δ-hedge by holding $\Delta^{BS}(t, S_t; T, K, \sigma_2)$ shares of S at all times $t \in [0, T]$. Thus, the dynamics of your replicating portfolio X are

$$dX_t = \Delta^{BS}(t, S_t; T, K, \sigma_2)dS_t + (X_t - \Delta^{BS}(t, S_t; T, K, \sigma_2)S_t)\frac{1}{B_t}dB_t, \quad X_0 = C^{BS}(0, S_0; T, K, \sigma_2).$$

The value of the option, according to the market is $C^{BS}(t, S_t; T, K, \sigma_2)$ for all $t \in [0, T]$.
Write Monte Carlo algorithm that plots the value of your hedging portfolio X_t and the value of the option $C^{BS}(t, S_t; T, K, \sigma_2)$ as a function of time. Plot a histogram of $(X_T - (S_T - K)^+)$ for M sample paths. Remember to simulate the paths of S using the true dynamics (i.e., with σ_1). Use the following parameters:

$$S_0 = K = T = 1, \qquad \sigma_1 = 0.2, \qquad \sigma_2 = 0.3, \qquad \mu = 0.1, \qquad N = 1000, \qquad M = 1000,$$

where N is the number of time-steps ($\delta T = T/N$) and M is the number of sample paths.

Exercise 8.2. Let the risk-free rate of interest be zero $r = 0$ so that the value of a bond is a constant $B_t = 1$. Assume that, under the risk-neutral pricing measure $\widetilde{\mathbb{P}}$, the dynamics of a stock S are given by

$$dS_t = \sigma_t S_t d\widetilde{W}_t,$$

where $(\sigma_t)_{t \geq 0}$ is a stochastic process. Consider a contract that pays $\int_0^T \sigma_t^2 dt$ at time T (i.e., the floating leg of a variance swap). In class, we showed that

$$\int_0^T \sigma_t^2 dt = \int_0^T \frac{2}{S_t} dS_t - 2 \log\left(\frac{S_T}{S_0}\right).$$

And thus, to replicated the payoff $\int_0^T \sigma_t^2 dt$ one should buy a European contract that pays $-2\log(S_T/S_0)$, hold $\Delta_t = 2/S_t$ shares of S_t at all times t, and lend or borrow from the bank as needed. What is the number of bonds Γ_t one should hold at time t? Your answer should involve a stochastic integral with respect to S.

Chapter 9

Dupire's Local Volatility

It would be preferable to have some model for S that results in option prices that agree with market quotes. As previously mentioned, the Black-Scholes model does not accomplish this task. In this chapter, we will introduce a model for S that induces option prices that are consistent with market quotes.

Consider a model in which the volatility of a stock is a deterministic function of time and the asset's price $\sigma_t = \sigma(t, \mathrm{S}_t)$. If we assume the risk-free rate of interest is zero ($r = 0$), then the dynamics of S under the risk-neutral probability measrue $\widetilde{\mathbb{P}}$ are of the form

$$d\mathrm{S}_t = \sigma(t, \mathrm{S}_t)\mathrm{S}_t d\widetilde{\mathrm{W}}_t, \tag{9.0.1}$$

where $\widetilde{\mathrm{W}}$ is a $\widetilde{\mathbb{P}}$ Brownian motion. The price of a call option in this setting is a function C of time, the stock price, the maturity and strike $\mathrm{C}(t, \mathrm{S}_t; \mathrm{T}, \mathrm{K})$. Assume that, at time t, we can observe the price of a Call option $\mathrm{C}_t(\mathrm{T}, \mathrm{K})$ for every $\mathrm{T} \geq t$ and $\mathrm{K} \geq 0$. We would like to find a function σ such that the model-induced option prices $\mathrm{C}(t, \mathrm{S}_t; \mathrm{T}, \mathrm{K})$ match the observed market prices $\mathrm{C}_t(\mathrm{T}, \mathrm{K})$. That is, we would like

$$\mathrm{C}(t, \mathrm{S}_t; \mathrm{T}, \mathrm{K}) = \mathrm{C}_t(\mathrm{T}, \mathrm{K}), \qquad \forall\, \mathrm{T} \geq t,\ \mathrm{K} \geq 0.$$

Let us consider a European option with payoff $\phi(\mathrm{S}_\mathrm{T})$ and let us denote by $\mathrm{V} = (\mathrm{V}_t)$ the value of this option. We know from the first fundamental Theorem of asset pricing that V/B is a martingale under $\widetilde{\mathbb{P}}$ with respect to $\mathcal{F}^\mathrm{S}$. As we have assumed $r = 0$, this implies the bond price B is a constant, and thus V is a martingale. Explicitly, we have

$$\mathrm{V}_t = \widetilde{\mathbb{E}}[\mathrm{V}_\mathrm{T}|\mathcal{F}_t^\mathrm{S}] = \widetilde{\mathbb{E}}[\phi(\mathrm{S}_\mathrm{T})|\mathcal{F}_t^\mathrm{S}].$$

Now, the stock price S, as the solution of SDE (9.0.1), is a Markov process. It follows that there exists a function v such that $V_t = v(t, S_t)$. Taking the differential of $v(t, S_t)$ we find that

$$\begin{aligned} dv(t, S_t) &= \frac{\partial v}{\partial t}(t, S_t)dt + \frac{\partial v}{\partial s}(t, S_t)dS_t + \frac{1}{2}\frac{\partial^2 v}{\partial s^2}(t, S_t)(dS_t)^2 \\ &= \left(\partial_t + \frac{1}{2}\sigma^2(t, S_t)S_t^2\partial_s^2\right) v(t, S_t)dt + \sigma(t, S_t)S_t\partial_s v(t, S_t)d\widetilde{W}_t. \end{aligned}$$

Because $(v(t, S_t))_{0\leq t\leq T}$ is a martingale, the dt-term above must equal zero. Thus, we obtain the a pricing PDE for the function v

$$0 = \left(\partial_t + \frac{1}{2}\sigma^2(t, s)s^2\partial_s^2\right) v, \qquad v(T, s) = \phi(s), \tag{9.0.2}$$

where the terminal condition $v(T, s) = \phi(s)$ comes from the fact that $V_T = v(T, S_T) = \mathbb{E}[\phi(S_T)|\mathcal{F}_T^S] = \phi(S_T)$. When $\phi(s) = (s - K)^+$ we have

$$0 = \left(\partial_t + \frac{1}{2}\sigma^2(t, s)s^2\partial_s^2\right) C(t, s; T, K), \qquad C(T, s; T, K) = (s - K)^+.$$

Differentiating the above equations twice with respect to K yields

$$0 = \left(\partial_t + \frac{1}{2}\sigma^2(t, s)s^2\partial_s^2\right) \partial_K^2 C(t, s; T, K), \qquad C(T, s; T, K) = \delta(s - K). \tag{9.0.3}$$

where we have used that $\partial_K^2(s - K)^+ = \delta(s - K)$. Now, note that

$$C(t, s; T, K) = \widetilde{\mathbb{E}}[(S_T - K)^+|S_t = s] = \int_0^\infty \widetilde{p}(t, s; T, y)(y - K)^+ dy,$$

where $\widetilde{p}(t, s; T, y)dy = \widetilde{\mathbb{P}}(S_T \in dy|S_t = s)$ is the (unknown) transition density of S. Differentiating the above equation twice with respect to K, we obtain

$$\begin{aligned} \partial_K^2 C(t, s; T, K) &= \int_0^\infty \widetilde{p}(t, s; T, y)\partial_K^2(y - K)^+ dy \\ &= \int_0^\infty \widetilde{p}(t, s; T, y)\delta(y - K)dy \\ &= \widetilde{p}(t, s; T, K). \end{aligned} \tag{9.0.4}$$

Combining (9.0.3) and (9.0.4), we have

$$0 = \left(\partial_t + \frac{1}{2}s^2\sigma^2(t, s)\partial_s^2\right) \widetilde{p}(t, s; T, y), \qquad \widetilde{p}(T, s; T, y) = \delta(s - y). \tag{9.0.5}$$

Equation (9.0.5) is known as the *Kolmogorov Backward equation* (KBE). Note that the KBE is an PDE and terminal condition for $\widetilde{p}$ in the *backward variables* (t, s). The forward variables (T, y) are merely parameters. For reasons that will become apparently later, we will need to know what $\widetilde{p}$ satisfies in *forward variables* (T, y). To this end, let us fix a $\bar{\mathrm{T}} > t$ and let us consider a general function u of time and the stock price. Using the short-hand notation $\widetilde{\mathbb{E}}_{t,s} \cdot := \widetilde{\mathbb{E}}[\,\cdot\,|\mathrm{S}_t = s]$ we have

$$\begin{aligned}
\mathbb{E}_{t,s} u(\bar{\mathrm{T}}, \mathrm{S}_{\bar{\mathrm{T}}}) &= u(t,s) + \widetilde{\mathbb{E}}_{t,s} \int_t^{\bar{\mathrm{T}}} \mathrm{d}u(\mathrm{T}, \mathrm{S}_\mathrm{T}) \\
&= u(t,s) + \widetilde{\mathbb{E}}_{t,s} \int_t^{\bar{\mathrm{T}}} \left(\partial_\mathrm{T} + \frac{1}{2}\sigma^2(\mathrm{T}, \mathrm{S}_\mathrm{T})\mathrm{S}_\mathrm{T}^2 \partial_s^2 \right) u(\mathrm{T}, \mathrm{S}_\mathrm{T})\mathrm{dT} + \widetilde{\mathbb{E}}_{t,s} \int_t^{\bar{\mathrm{T}}} (\ldots)\mathrm{d}\widetilde{\mathrm{W}}_\mathrm{T} \\
&= u(t,s) + \int_0^\infty \int_t^{\bar{\mathrm{T}}} \left\{ \widetilde{p}(t,s,\mathrm{T},y) \left[\partial_\mathrm{T} + \frac{1}{2}\sigma^2(\mathrm{T}, y^2) y^2 \partial_y^2 \right] u(\mathrm{T}, y) \right\} \mathrm{dT}\mathrm{d}y,
\end{aligned}$$

where the precise form of $(\ldots)$ is not important, as the expected value of an Itô integral is zero. Integrating the first term in the integral by parts once with respect to T, and integrating the second term in the integral by parts twice with respect to y, we obtain

$$\begin{aligned}
\widetilde{\mathbb{E}}_{t,s} u(\bar{\mathrm{T}}, \mathrm{S}_{\bar{\mathrm{T}}}) &= u(t,s) + \int_0^\infty \int_t^{\bar{\mathrm{T}}} u(\mathrm{T}, y) \left[-\partial_\mathrm{T} + \frac{1}{2}\partial_y^2 \sigma^2(\mathrm{T}, y) y^2 \right] \widetilde{p}(t,s,\mathrm{T},y)\mathrm{dT}\mathrm{d}y \\
&\quad + \int_0^\infty u(\mathrm{T}, y)\widetilde{p}(t,s,\mathrm{T},y)\mathrm{d}y \Big|_{\mathrm{T}=t}^{\mathrm{T}=\bar{\mathrm{T}}} \\
&= u(t,s) + \int_0^\infty \int_t^{\bar{\mathrm{T}}} u(\mathrm{T}, y) \left[-\partial_\mathrm{T} + \frac{1}{2}\partial_y^2 \sigma^2(\mathrm{T}, y) y^2 \right] \widetilde{p}(t,s,\mathrm{T},y)\mathrm{dT}\mathrm{d}y \\
&\quad + \widetilde{\mathbb{E}}_{t,s} u(\bar{\mathrm{T}}, \mathrm{S}_{\bar{\mathrm{T}}}) - u(t,s),
\end{aligned}$$

where we have used $\lim_{y \to 0} \widetilde{p}(t,s;\mathrm{T},y) = 0$ and $\lim_{y \to \infty} \widetilde{p}(t,s;\mathrm{T},y) = 0$. Canceling the terms $\widetilde{\mathbb{E}}_{t,s} u(\bar{\mathrm{T}}, \mathrm{S}_{\bar{\mathrm{T}}})$ and $u(t,s)$, which both appear twice above with opposite signs, we obtain

$$0 = \int_0^\infty \int_t^{\bar{\mathrm{T}}} u(\mathrm{T}, y) \left[-\partial_\mathrm{T} + \frac{1}{2}\partial_y^2 \sigma^2(\mathrm{T}, y) y^2 \right] \widetilde{p}(t,s,\mathrm{T},y)\mathrm{dT}\mathrm{d}y.$$

The above equation must hold for *all* functions u. The only way for this to happen is if $\widetilde{p}(t,s;\mathrm{T},y)$ satisfies

$$0 = -\partial_\mathrm{T}\widetilde{p}(t,s,\mathrm{T},y) + \frac{1}{2}\partial_y^2 \Big(\sigma^2(\mathrm{T}, y) y^2 \widetilde{p}(t,s,\mathrm{T},y) \Big), \qquad \widetilde{p}(t,s;t,y) = \delta(s-y). \tag{9.0.6}$$

The above equation is known as the *Kolmogorov Forward Equation* (KFE). The KFE is a PDE and initial condition in forward variables (T, y). The backward variables (t,s) serve merely as parameters.

Now we return to the call option. If we differentiate the value of a Call option with respect to the maturity date T, we obtain

$$\begin{aligned}
\partial_\mathrm{T}\mathrm{C}(t,s;\mathrm{T},\mathrm{K}) &= \int_0^\infty \partial_\mathrm{T}\widetilde{p}(t,s,\mathrm{T},y)(y-\mathrm{K})^+\mathrm{d}y \\
&= \int_0^\infty (y-\mathrm{K})^+ \partial_y^2 \left(\frac{1}{2}\sigma^2(\mathrm{T}, y) y^2 \widetilde{p}(t,s;\mathrm{T},y) \right) \mathrm{d}y \qquad \text{(by (9.0.6))}
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{2}\sigma^2(\mathrm{T},y)y^2\widetilde{p}(t,s;\mathrm{T},y)\partial_y^2(y-\mathrm{K})^+\mathrm{d}y && \text{(integration by parts)}\\
&= \int_0^\infty \frac{1}{2}\sigma^2(\mathrm{T},y)y^2\widetilde{p}(t,s;\mathrm{T},y)\delta(y-\mathrm{K})\mathrm{d}y\\
&= \frac{1}{2}\sigma^2(\mathrm{T},\mathrm{K})\mathrm{K}^2\widetilde{p}(t,s;\mathrm{T},\mathrm{K})\\
&= \frac{1}{2}\sigma^2(\mathrm{T},\mathrm{K})\mathrm{K}^2\partial_\mathrm{K}^2\mathrm{C}(t,s,\mathrm{T},\mathrm{K}). && \text{(by (9.0.4))}
\end{aligned}
$$

Solving the above equation for $\sigma^2(\mathrm{T},\mathrm{K})$ yields

$$
\sigma^2(\mathrm{T},\mathrm{K}) = \frac{\partial_\mathrm{T}\mathrm{C}(t,s,\mathrm{T},\mathrm{K})}{\frac{1}{2}\mathrm{K}^2\partial_\mathrm{K}^2\mathrm{C}(t,s,\mathrm{T},\mathrm{K})}. \tag{9.0.7}
$$

Equation (9.0.7) is *Dupire's Formula*. It also holds if we replace replacing Call prices with Put prices because $\partial_\mathrm{K}^2(\mathrm{K}-\mathrm{S_T})^+ = \delta(\mathrm{K}-\mathrm{S_T})$. The formula tells us how to construct the function $\sigma(t,s)$ from (observed Market) Call and Put prices so that the model (9.0.1) produces option prices that are consistent with market quotes.

9.1 Practical considerations

Implementation of Dupire's Formula requires one to differentiate Call and/or Put prices with respect to T and K. This, in turn, requires a smooth surface of call prices $(\mathrm{T},\mathrm{K}) \mapsto \mathrm{C}(\mathrm{T},\mathrm{K})$. In practice, we only have access to Market quotes at discrete maturities and strikes: $(\mathrm{T}_i,\mathrm{K}_j) \mapsto \mathrm{C}(\mathrm{T}_i,\mathrm{K}_j)$. We will discuss two possible ways to handle this.

Finite Difference Approximations. For simplicity, let us assume that Call prices are quoted at uniformly spaced intervals in T and K. We will denote the sizes of the maturity intervals as $\Delta\mathrm{T}$ and the sizes of the strike intervals as $\Delta\mathrm{K}$. In order to obtain an approximation for $\partial_\mathrm{T}\mathrm{C}(\mathrm{T},\mathrm{K})$ we expand $\mathrm{C}(\mathrm{T}+\Delta\mathrm{T},\mathrm{K})$ as a Taylor series about (T,K). We have

$$
\mathrm{C}(\mathrm{T}\pm\Delta\mathrm{T},\mathrm{K}) = \mathrm{C}(\mathrm{T},\mathrm{K}) \pm \partial_\mathrm{T}\mathrm{C}(\mathrm{T},\mathrm{K})\Delta\mathrm{T} + \frac{1}{2!}\partial_\mathrm{T}^2\mathrm{C}(\mathrm{T},\mathrm{K})\Delta\mathrm{T}^2 + \mathcal{O}(\Delta\mathrm{T}^3).
$$

From the above equation, we obtain

$$
\partial_\mathrm{T}\mathrm{C}(\mathrm{T},\mathrm{K}) = \frac{\mathrm{C}(\mathrm{T}+\Delta\mathrm{T},\mathrm{K})-\mathrm{C}(\mathrm{T}-\Delta\mathrm{T},\mathrm{K})}{2\Delta\mathrm{T}} + \mathcal{O}(\Delta\mathrm{T}^2). \tag{9.1.1}
$$

Equation (9.1.1) is known as the *central difference approximation* of $\partial_\mathrm{T}\mathrm{C}(\mathrm{T},\mathrm{K})$. To obtain an approximation for $\partial_\mathrm{K}^2\mathrm{C}(\mathrm{T},\mathrm{K})$ we again use a Taylor series. We have

$$
\mathrm{C}(\mathrm{T},\mathrm{K}\pm\Delta\mathrm{K}) = \mathrm{C}(\mathrm{T},\mathrm{K}) \pm \partial_\mathrm{K}\mathrm{C}(\mathrm{T},\mathrm{K})\Delta\mathrm{K} + \frac{1}{2!}\partial_\mathrm{K}^2\mathrm{C}(\mathrm{T},\mathrm{K})\Delta\mathrm{K}^2 \pm \frac{1}{3!}\partial_\mathrm{K}^3\mathrm{C}(\mathrm{T},\mathrm{K})\Delta\mathrm{K}^3 + \mathcal{O}(\Delta\mathrm{K}^4)
$$

From the above equation, we obtain

$$\partial_K^2 C(T,K) = \frac{C(T,K+\Delta K) - 2C(T,K) + C(T,K-\Delta K)}{\Delta K^2} + \mathcal{O}(\Delta K^2). \tag{9.1.2}$$

Equation (9.1.2) is known as the *central difference approximation* of $\partial_K^2 C(T,K)$.

Smooth Interpolation. Rather than use finite difference methods, one might instead choose to fit the volatility surface with a parametric curve in K for every fixed maturity T_i. Consider the *Stochastic Volatility Inspired* (SVI) parametrization of implied volatility

$$\sigma^{SVI}(T,k) := \sqrt{\left[a(T) + b(T)\left(\rho(T)(k-x-m(T)) + \sqrt{(k-x-m(T))^2 + \xi^2(T)}\right)\right]},$$

where $k := \log K$, $x = \log S_0$ and $a(T)$, $b(T)$, $\rho(T)$, $m(T)$ and $\xi(T)$ are parameters to be obtained via calibration. Now, consider the following procedure

1. For each maturity T_i, find the parameters $a(T)$, $b(T_i)$, $\rho(T_i)$, $m(T_i)$ and $\xi(T_i)$ that give the best fit (in the least squares sense) to the market's implied volatility smile. This yields a call price

$$C(T_i,K) = C^{BS}(0,S_0;\sigma^{SVI}(T_i,\log K),T_i,K)$$

   for each maturity T_i and every $K \geq 0$.

2. Use central difference (9.1.1) to obtain an approximation of $\partial_T C(T_i,K)$.

3. To obtain an approximation of $\partial_K^2 C(T_i,K)$ simply differentiate call prices analytically

$$\partial_K^2 C(T_i,K) = \partial_K^2 C^{BS}(0,S_0;\sigma^{SVI}(T_i,\log K)),T_i,K).$$

9.2 Advantages and Disadvantages of local volatility

One *advantage* of the local volatility framework (aside from the fact it induces option prices that fit market quotes) the local volatility model is a *complete market model* (i.e., derivatives can be perfectly replicated). Assuming that $r = 0$ for simplicity the dynamics of a portfolio X and an option price $v(t,S_t)$ are given by

$$dX_t = \Delta_t dS_t + (X_t - \Delta_t S_t)\frac{1}{B_t}dB_t = \Delta_t dS_t,$$
$$dv(t,S_t) = \left(\partial_t + \frac{1}{2}\sigma^2(t,S_t)S_t^2\partial_s^2\right)v(t,S_t)dt + \partial_s v(t,S_t)dS_t = \partial_s v(t,S_t)dS_t,$$

where we have used $\mathrm{dB}_t = 0$ and the fact that the $\mathrm{d}t$-term is zero by (9.0.2). Thus, if we choose

$$\Delta_t = \partial_s v(t, \mathrm{S}_t), \qquad \text{then we have} \qquad \mathrm{dX}_t = \mathrm{d}v(t, \mathrm{S}_t).$$

If we further have $\mathrm{X}_0 = v(0, \mathrm{S}_0)$ then we will have $\mathrm{X}_t = v(t, \mathrm{S}_t)$ for all $t \in [0, \mathrm{T}]$ where T is the maturity date of the option.

The main *disadvantage* of the local volatility framework is that the volatility function $\sigma(t, s)$ must be recalibrated daily! That is, if we find a local volatility function $\sigma(t, s) = f_1(t, s)$ on day 1, then we very often find a different local volatility function $\sigma(t, s) = f_2(t, s)$ on day two, and still a different volatility function $\sigma(t, s) = f_3(t, s)$ on day three and so on. This, of course, brings into question the use of the local volatility framework.

9.3 Exercises

Exercise 9.1. Assume for simplicity that the risk-free rate of interest is zero $r = 0$. Suppose we model the stock price S under the risk-neutral pricing measure as

$$\mathrm{dS}_t = f(t, \mathrm{S}_t)\mathrm{S}_t \mathrm{d}\widetilde{\mathrm{W}}_t.$$

In order for the dynamics of S to be consistent with option prices on the market, Dupire's formula must be satisfied

$$f^2(\mathrm{T}, \mathrm{K}) = \frac{\partial_\mathrm{T} \mathrm{C}(t, s, \mathrm{T}, \mathrm{K})}{\frac{1}{2}\mathrm{K}^2 \partial_\mathrm{K}^2 \mathrm{C}(t, s, \mathrm{T}, \mathrm{K})},$$

where $\mathrm{C}(t, s, \mathrm{T}, \mathrm{K})$ denotes the value of a call option with maturity T and strike K given $\mathrm{S}_t = s$. Suppose that, for every T and K, the implied volatility of an option is a function of T only: $\sigma(\mathrm{T})$ (i.e., there is no K-dependence).

(a) Derive an expression for $f(\mathrm{T}, \mathrm{K})$. In particular, show that $f(\mathrm{T}, \mathrm{K})$ has no dependence on K.

(b) Show that

$$\sigma^2(\mathrm{T}) = \frac{1}{\mathrm{T} - t} \int_t^\mathrm{T} f^2(u, \mathrm{K})\mathrm{d}u.$$

Chapter 10

One-factor Stochastic Volatility Models

In a standard (one-factor) *stochastic volatility* model the dynamics of a stock price S under the physical measure $\mathbb{P}$ are of the form

$$\mathrm{dS}_t = \mu(\mathrm{Y}_t)\mathrm{S}_t\mathrm{d}t + \sigma(\mathrm{Y}_t)\mathrm{S}_t\mathrm{dW}_t^{\mathrm{S}}, \qquad \mathrm{dY}_t = \alpha(\mathrm{Y}_t)\mathrm{d}t + \beta(\mathrm{Y}_t)\mathrm{dW}_t^{\mathrm{Y}}, \tag{10.0.1}$$

$$\mathrm{dW}_t^{\mathrm{Y}} = \rho\mathrm{dW}_t^{\mathrm{S}} + \bar{\rho}\mathrm{dW}_t, \qquad \bar{\rho} = \sqrt{1-\rho^2}, \tag{10.0.2}$$

where $\mathrm{W^S} = (\mathrm{W}_t^{\mathrm{S}})_{t\geq 0}$ and $\mathrm{W} = (\mathrm{W}_t)_{t\geq 0}$ are independent, one-dimensional Brownian motions. One can check that $\mathrm{W^Y} = (\mathrm{W}_t^{\mathrm{Y}})_{t\geq 0}$, defined by (10.0.2), is a Brownian motion (i.e., it satisfies the properties given in the definition of a Brownian motion). Furthermore, we have $\mathrm{dW}_t^{\mathrm{S}}\mathrm{dW}_t^{\mathrm{Y}} = \rho\mathrm{d}t$. Thus, we say that $\mathrm{W^S}$ and $\mathrm{W^Y}$ are *correlated* and we call $\rho \in [-1, 1]$ the *correlation coefficient*. We shall refer to Y as the *driver of volatility* or volatility driving process because $\sigma(\mathrm{Y})$ the volatility of S. In most (but not all) stochastic volatility models, the Y process is *mean-reverting* meaning that Y fluctuates around a long-run mean (like the OU process) that we discussed in Example 5.19. As in the Black-Scholes setting, the drift μ plays no role in the pricing of options. Let us see a few example of one-factor stochastic volatility models.

Example 10.1 (Heston). In the the stochastic volatility model of Heston (1993), the coefficient functions α, β and σ in (10.0.1) are given by

$$\alpha(y) = \kappa(\theta - y), \qquad \beta(y) = \delta\sqrt{y}, \qquad \sigma(y) = \sqrt{y}.$$

The Y process in the Heston model is commonly known as a *Cox-Ingersoll-Ross* (CIR) processes. The Heston model is well-known in industry and in academia due to the fact that there exists a closed-form formula for European option prices.

Example 10.2 (Exponential OU). The Exponential Ornstein-Uhlenbeck (ExpOU) stochastic volatility model, the coefficient functions α, β and σ in (10.0.1) are given by

$$\alpha(y) = m(\theta - y), \qquad \beta(y) = \delta, \qquad \sigma(y) = \exp(y).$$

European option prices in this setting can be computed in closed-form only S and Y are uncorrelated: $\rho = 0$.

Example 10.3 (3/2's). In the 3/2's stochastic volatility model, the coefficient functions α, β and σ in (10.0.1) are given by

$$\alpha(y) = y\kappa(\theta - y), \qquad \beta(y) = \delta y\sqrt{y}, \qquad \sigma(y) = \sqrt{y}.$$

European option prices in this setting can be computed in closed-form involving special functions.

10.1 Pricing by replication

We would like to derive a pricing formula and replication strategy for European options in the stochastic volatility setting (10.0.1). In the Black-Scholes and local volatility settings, we were able to replicate a derivative with a self-financing portfolio that contained only two assets – the underlying stock S and a bond B. Perfect replication of the derivative asset was possible because the derivative price depended on a single Brownian motion W, which we could hedge by trading the stock. In the stochastic volatility setting (10.0.1), the derivative asset's price depends two sources of randomness W^S and W^Y. This means that we can no longer replicate the derivative asset by trading only the stock and the bond. We will consider a replicating portfolio containing the stock, the bond, and another option.

in the stochastic volatility setting (10.0.1), we expect the value $V = (V_t)_{0 \le t \le T}$ of any option to be of the form

$$V_t = v(t, S_t, Y_t),$$

for some function v. Suppose we have *sold* an option with payoff $h^{(1)}(S_{T_1})$ expiring at time T_1. We will denote the value of this option as $V_t^{(1)} = v^{(1)}(t, S_t, Y_t)$. We will attempt to replicate this option with a portfolio consisting of the underlying stock S, the bond B, and an option expiring at time $T_2 > T_1$ with payoff $h^{(2)}(S_{T_2})$. We will denote the value of this option as $V_t^{(2)} = v^{(2)}(t, S_t, Y_t)$. As usual, we will assume $r = 0$ for simplicity.

We set up our hedging portfolio X so that we have Δ shares of S at time and Γ shares of $V^{(2)}$. The dynamics of X are then given by

$$dX_t = \Delta_t dS_t + \Gamma_t dV_t^{(2)} + (X_t - \Delta_t S_t - \Gamma_t V_t^{(2)})\frac{1}{B_t}dB_t$$

$$= \Delta_t \mathrm{d}\mathrm{S}_t + \Gamma_t \partial_s v^{(2)} \mathrm{d}\mathrm{S}_t + \Gamma_t \partial_y v^{(2)} \mathrm{d}\mathrm{Y}_t + \Gamma_t (\partial_t + \mathcal{M}) v^{(2)} \mathrm{d}t, \tag{10.1.1}$$

where we have used the shorthand $v^{(2)} = v^{(2)}(t, \mathrm{S}_t, \mathrm{Y}_t)$ for notational clarity and we have introduced an operator $\mathcal{M}$, which is defined as follows

$$\mathcal{M} := \frac{1}{2}\sigma^2(y) s^2 \partial_s^2 + \frac{1}{2}\beta^2(y)\partial_y^2 + \rho\sigma(y)\beta(y) s \partial_s \partial_y.$$

We would like that change in the value of our portfolio $\mathrm{d}\mathrm{X}_t$ to equal the change in the value of the option $\mathrm{d}\mathrm{V}_t^{(1)}$. Thus, we compute

$$\mathrm{d}\mathrm{V}_t^{(1)} = \mathrm{d}v_t^{(1)} = \partial_s v^{(1)} \mathrm{d}\mathrm{S}_t + \partial_y v^{(1)} \mathrm{d}\mathrm{Y}_t + (\partial_t + \mathcal{M}) v^{(1)} \mathrm{d}t. \tag{10.1.2}$$

In order to have $\mathrm{d}\mathrm{X}_t = \mathrm{d}\mathrm{V}_t^{(1)}$ we must match the $\mathrm{d}\mathrm{S}_t$, the $\mathrm{d}\mathrm{Y}_t$ and the $\mathrm{d}t$ terms in (10.1.1) and (10.1.2). From

$$\begin{aligned} \Delta_t + \Gamma_t \partial_s v^{(2)} &= \partial_s v^{(1)}, && \text{(matching the } \mathrm{d}\mathrm{S}_t \text{ terms)} \\ \Gamma_t \partial_y v^{(2)} &= \partial_y v^{(1)}, && \text{(matching the } \mathrm{d}\mathrm{Y}_t \text{ terms)} \end{aligned}$$

we deduce that Γ and Δ must satisfy

$$\Gamma_t = \frac{\partial_y v^{(1)}}{\partial_y v^{(2)}}, \qquad \Delta_t = \partial_s v^{(1)} - \Gamma_t \partial_s v^{(2)}. \tag{10.1.3}$$

We are left with $\mathrm{d}t$ terms. Comparing (10.1.1) and (10.1.2), and using the expression for Γ_t in (10.1.3), we must have

$$\underbrace{\left[\partial_{\mathrm{Y}} v^{(1)}\right]^{-1} (\partial_t + \mathcal{M}) v^{(1)}}_{\text{independent of } \mathrm{T}_2} = \underbrace{\left[\partial_{\mathrm{Y}} v^{(2)}\right]^{-1} (\partial_t + \mathcal{M}) v^{(2)}}_{\text{independent of } \mathrm{T}_1} =: f(t, s, y). \tag{10.1.4}$$

As the left-hand side in independent of T_1 and the right-hand side is independent of T_1 we deduce that both sides must be equal to a function f that is independent of both T_1 and T_2. For reasons that will be come apparent later, we define a function Λ as follow

$$\Lambda(t, s, y) := \frac{\alpha(y) - f(t, s, y)}{\beta(y)} \qquad \Rightarrow \qquad f(t, s, y) = \alpha(y) - \beta(y)\Lambda(t, s, y). \tag{10.1.5}$$

Substituting (10.1.5) into (10.1.4) we find that $v^{(1)}$ and $v^{(2)}$ must satisfy

$$(\partial_t + \widetilde{\mathcal{A}}) v^{(i)} = 0, \qquad v^{(i)}(\mathrm{T}_i, s, y) = h^{(i)}(s), \tag{10.1.6}$$

where the operator $\widetilde{\mathcal{A}}$ is defined as follows

$$\widetilde{\mathcal{A}} := \frac{1}{2}\sigma^2(y) s^2 \partial_s^2 + \rho\sigma(y)\beta(y) s \partial_s \partial_y + \frac{1}{2}\beta^2(y)\partial_y^2 + \left[\alpha(y) - \beta(y)\Lambda(t, s, y)\right] \partial_y, \tag{10.1.7}$$

and we have added the obvious terminal condition $v^{(i)}(\mathrm{T}_i, s, v) = h^{(i)}(s)$, which is needed to ensure that $\mathrm{V}^{(i)}_{\mathrm{T}_i} = v^{(i)}(\mathrm{T}_i, \mathrm{S}_{\mathrm{T}_i}, \mathrm{Y}_{\mathrm{T}_i}) = h^{(i)}(\mathrm{S}_{\mathrm{T}_i})$. Thus, we have derived a pricing PDE (10.1.6) for European options in a stochastic volatility setting.

The function Λ is called the *market price of volatility risk*. Note that we have *not* uniquely determined Λ yet! A good student will be asking him/herself: *what should* Λ *be*? The situation is similar situation when we encountered in the trinomial model. Recall that, in the trinomial model, there was no unique risk-neutral probability measure $\widetilde{\mathbb{P}} = (\widetilde{p}, \widetilde{q}, \widetilde{r})$ to price options in a no-arbitrage fashion. However, we could parametrize $\widetilde{\mathbb{P}}(\widetilde{r}) = (\widetilde{p}(\widetilde{r}), \widetilde{q}(\widetilde{r}), \widetilde{r})$ in terms of $\widetilde{r}$. Given the price of a single option, we could determine $\widetilde{r}$, and therefore $(\widetilde{p}(\widetilde{r}), \widetilde{q}(\widetilde{r}), \widetilde{r})$ uniquely. And then we could and price other options in a unique way.

In the stochastic volatility setting the role of $\widetilde{r}$ is played by the function Λ. In other words, different choices of Λ results in different risk-neutral probability measures $\widetilde{\mathbb{P}}(\Lambda)$. We generally assume that the market has chosen a unique Λ, under which all options are priced. Unlike the trinomial model, we cannot determine Λ from a single option price. To find Λ we specify a particular form (e.g., $\Lambda(t, s, y) = \lambda\sqrt{y}$. We then *calibrate* our model to the observed option prices (meaning we search for the parameters that result in options prices that most closely match the option prices on the market in some sense). We will discuss calibration in detail in Section 10.4.

10.2 Risk neutral pricing

The 1st fundamental theorem of asset pricing (Theorem 4.2) states that a market defined under a probability measure $\mathbb{P}$ is free of arbitrage if and only if there exists a probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, under which all portfolios X, denominated in units of a common numéraire N are martingales. In the Stochastic volatility setting (10.0.1) this leads to a very simple procedure for finding the value $\mathrm{V} = (\mathrm{V}_t)_{0 \leq t \leq \mathrm{T}}$ of an option that pays $\phi(\mathrm{S}_\mathrm{T})$ at time T.

1. Choose a numèraire. We will choose the bond $\mathrm{B} = (\mathrm{B}_t)_{0 \leq t \leq \mathrm{T}}$, as this choice leads to the easiest computations.

2. Search for a probability measure $\widetilde{\mathbb{P}}$, under which (X/B) is a martingale for all portfolios X. That is, find $\widetilde{\mathbb{P}}$ such that

$$\frac{\mathrm{X}_t}{\mathrm{B}_t} = \widetilde{\mathbb{E}}\left(\frac{\mathrm{X}_\mathrm{T}}{\mathrm{B}_\mathrm{T}}\middle|\mathcal{F}^{\mathrm{S},\mathrm{Y}}_t\right),$$

   where X is given by (6.1.1) and $\mathcal{F}^{\mathrm{S},\mathrm{Y}}$ is the filtration generated by observing S and Y.

3. The value V of the option can then be computed using

$$\frac{V_t}{B_t} = \widetilde{\mathbb{E}}\Big(\frac{V_T}{B_T}\Big|\mathcal{F}_t^{S,Y}\Big) = \widetilde{\mathbb{E}}\Big(\frac{\phi(S_T)}{B_T}\Big|\mathcal{F}_t^{S,Y}\Big).$$

Let us carry this procedure out. As usual, we will assume for simplicity that $r = 0$, which implies that $dB_t = 0$. Because B is a constant, X/B will be a martingale if and only if X is a martingale. Thus, we look for a probability measure $\widetilde{\mathbb{P}}$ under which X is a martingale. To begin, let us define $\widetilde{W}^S = (\widetilde{W}_t^S)_{0\le t\le T}$ and $\widetilde{W} = (\widetilde{W}_t)_{0\le t\le T}$ as follows

$$\widetilde{W}_t^S = \int_0^t \gamma_s ds, \qquad \widetilde{W}_t = \int_0^t \lambda_s ds,$$

where $\gamma = (\gamma_t)_{0\le t\le T}$ and $\lambda = (\lambda_t)_{0\le t\le T}$ are, for the moment, unspecified stochastic processes. We know from Girsanov's Theorem 6.1 that there exists a probability measure $\widetilde{\mathbb{P}}$ under which $\widetilde{W}^S$ and $\widetilde{W}$ are Brownian motions. Let us write the dynamics of X under this probability measure. We have

$$\begin{aligned} dX_t &= \Delta_t dS_t + (X_t - \Delta_t S_t)\frac{1}{B_t}dB_t \\ &= \Delta_t \mu(Y_t)S_t dt + \sigma(Y_t)S_t dW_t^S \\ &= \Delta_t \mu(Y_t)S_t dt + \Delta_t\sigma(Y_t)S_t(d\widetilde{W}_t^S - \gamma_t dt) \\ &= \Delta_t\Big(\mu(Y_t) - \sigma(Y_t)\gamma_t\Big)dt + \Delta_t\sigma(Y_t)d\widetilde{W}_t^S. \end{aligned} \tag{10.2.1}$$

Examining (10.2.1) we see that X will be a martingale under $\widetilde{\mathbb{P}}$ if we choose

$$\gamma_t = \frac{\mu(Y_t)}{\sigma(Y_t)}. \tag{10.2.2}$$

Thus, the dynamics of S under $\widetilde{\mathbb{P}}$ must be given by

$$\begin{aligned} dS_t &= \mu(Y_t)S_t dt + \sigma(Y_t)S_t dW_t^S \\ &= \mu(Y_t)S_t dt + \sigma(Y_t)S_t\Big(d\widetilde{W}_t^S - \gamma_t dt\Big) \\ &= \sigma(Y_t)S_t d\widetilde{W}_t^S, \end{aligned} \tag{10.2.3}$$

where we have used (10.2.2). What about the dynamics of Y? Well, they must be of the form

$$\begin{aligned} dY_t &= \alpha(Y_t)dt + \beta(Y_t)dW_t^Y \\ &= \alpha(Y_t)dt + \beta(Y_t)(\rho dW_t^S + \bar{\rho}dW_t) \\ &= \alpha(Y_t)dt + \beta(Y_t)\rho(d\widetilde{W}_t^S - \gamma_t dt) + \beta(Y_t)\bar{\rho}(d\widetilde{W}_t - \lambda_t dt) \\ &= \Big(\alpha(Y_t) - \rho\beta(Y_t)\gamma_t - \bar{\rho}\beta(Y_t)\lambda_t\Big)dt + \beta(Y_t)(\rho d\widetilde{W}_t^S + \bar{\rho}d\widetilde{W}_t) \end{aligned}$$

$$= \Big(\alpha(Y_t) - \rho\beta(Y_t)\frac{\mu(Y_t)}{\sigma(Y_t)} - \bar{\rho}\beta(Y_t)\lambda_t\Big)dt + \beta(Y_t)d\widetilde{W}_t^Y,$$

where we have introduced $\widetilde{W}^Y = (\widetilde{W}_t^Y)_{0\le t\le T}$, which is defined as follows

$$\widetilde{W}_t^Y := \rho\widetilde{W}_t^S + \bar{\rho}\widetilde{W}_t.$$

One can check that $\widetilde{W}^Y$ is a Brownian motion under $\widetilde{\mathbb{P}}$ and that $d\widetilde{W}_t^S d\widetilde{W}_t^Y = \rho dt$. As Y is not a traded asset, there is no reason why it should be a martingale. As such, we, have the freedom to choose λ as we like. It is reasonable to assume that λ is of the form $\lambda_t = \lambda(t, S_t, Y_t)$. Defining

Does this mean

$$\Lambda(t,s,y) := \rho\frac{\mu(y)}{\sigma(y)} + \bar{\rho}\lambda(t,s,y),$$

we can express the dynamics of Y under $\widetilde{\mathbb{P}}$ as follows

$$dY_t = (\alpha(Y_t) - \beta(Y_t)\Lambda(t,S_t,Y_t))dt + \beta(Y_t)d\widetilde{W}_t^Y. \tag{10.2.4}$$

Finally, as the bond price B is a constant, the price of the option V is given by

$$V_t = \widetilde{\mathbb{E}}[\phi(S_T)|\mathcal{F}_t^{S,Y}] = \widetilde{\mathbb{E}}[\phi(S_T)|S_t, Y_t] =: v(t, S_t, Y_t), \tag{10.2.5}$$

where the dynamics of (S, Y) under $\widetilde{\mathbb{P}}$ are given by (10.2.3) and (10.2.4), respectively. Note that, in the second equality above, we have used the fact that (S, Y), as the solution of an SDE, is a Markov process. If we knew the conditional density of S_T given (S_t, Y_t) under $\widetilde{\mathbb{P}}$, we could compute the option price V_t using (10.2.5). Unfortunately, it is typically not possible to find $\mathbb{P}(S_T \in ds|S_t, Y_t)$ in a stochastic volatility setting. Nevertheless, we can derive a PDE and terminal condition for the option price $V_t = v(t, S_t, Y_t)$.

The option price V must be a martingale under $\widetilde{\mathbb{P}}$ (otherwise there would be an arbitrage). Thus, if we take the differential of dV_t, the dt-term must equal zero. By Itô's Lemma, we have from (10.2.3) and (10.2.4) that

$$\begin{aligned} dv(t,S_t,Y_t) &= \partial_t v dt + \partial_s v dS_t + \partial_y v dY_t + \tfrac{1}{2}\partial_s^2 (dS_t)^2 + \tfrac{1}{2}\partial_y^2 (dY_t)^2 + \partial_s\partial_y v dS_t dY_t \\ &= (\partial_t + \widetilde{\mathcal{A}})v(t,S_t,Y_t)dt + (\ldots)d\widetilde{W}_t^S + (\ldots)d\widetilde{W}_t^Y, \end{aligned}$$

where the operator $\widetilde{\mathcal{A}}$ is given by (10.1.7) and the particular form of the $(\ldots)$ terms is not particularly important. Setting the dt-term equal to zero, we find that v must satisfy

$$0 = (\partial_t + \widetilde{\mathcal{A}})v, \qquad v(T,s,y) = \phi(s), \tag{10.2.6}$$

where the terminal condition comes from the fact that $v(T, S_T, Y_T) = \phi(S_T)$. Comparing (10.2.6) with (10.1.6), we see that we have obtained the same PDE for the option price using risk-neutral pricing as we did using a replication argument.

10.3 The Heston model

In this section, we will show how to find a the solution v of the pricing PDE (10.2.6) when the underlying stock price S has Heston dynamics. Under the physical measure $\mathbb{P}$, the dynamics of S are described in Example 10.1. The operator $\widetilde{\mathcal{A}}$ defined in (10.1.7), is given explicitly by

$$\widetilde{\mathcal{A}} = \frac{1}{2} ys\partial_s^2 + \rho\delta sy\partial_{sy}^2 + \frac{1}{2}\delta^2 y\partial_{yy}^2 + \left[\kappa(\theta - y) - \delta\sqrt{y}\Lambda(t,s,y)\right]\partial_y .$$

One typically guesses that Λ is of the form $\Lambda(t,s,y) = \lambda\sqrt{y}$. Inserting this form into the above expression for $\mathcal{A}$ and performing some algebra, we obtain

$$\widetilde{\mathcal{A}} = \frac{1}{2} ys^2\partial_s^2 + \rho\delta ys\partial_{sy}^2 + \frac{1}{2}\delta y\partial_{yy}^2 + \kappa^*(\theta^* - y)\partial_y ,$$

where $\kappa^* = \kappa + \delta\lambda$, $\theta^* = \kappa\theta/(\kappa\delta\lambda)$. To ease notation, we will make the notational change $\kappa^* \to \kappa$ and $\theta^* \to \theta$. With this change, pricing PDE (10.2.6) then becomes

$$0 = \left(\partial_t + \frac{1}{2}ys^2\partial_s^2 + \rho\delta ys\partial_x\partial_y + \frac{1}{2}\delta y\partial_y^2 + \kappa(\theta - y)\partial_y\right) v(t,s,y), \qquad v(\mathrm{T},s,y) = \phi(s). \tag{10.3.1}$$

In order to solve this PDE, let us define

$$v(t,s,y) := u(t,x(s),y), \qquad x(s) := \log s. \tag{10.3.2}$$

Inserting (10.3.2) into (10.3.1) and using

$$\begin{aligned} \partial_s v &= \frac{1}{s}\partial_x u, & \partial_s^2 v - \frac{1}{s^2}(\partial_x^2 - \partial_x)u, && \partial_y v &= \partial_y u, \\ \partial_y^2 v &= \partial_y^2 u, & \partial_t v = \partial_t u, && \partial_s\partial_y v &= \frac{1}{s}\partial_x\partial_y v, \end{aligned}$$

we find that u satisfies

$$0 = \left(\partial_t + \tfrac{1}{2}y(\partial_x^2 - \partial_x) + \rho\delta y\partial_x\partial_y + \tfrac{1}{2}\delta y\partial_y^2 + \kappa(\theta - y)\partial_y\right) u(t,x,y), \tag{10.3.3}$$

$$u(\mathrm{T},x,y) = \phi(e^x) =: \varphi(x). \tag{10.3.4}$$

Now, let us introduce the *Fourier Transform*. For any function f we define its Fourier transform $\widehat{f}$ by

$$\widehat{f}(\xi) := \int_{\mathbb{R}} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}\xi x} f(x).$$

It is well known that one can recover the function f from its Fourier transform $\widehat{f}$ using the *Inverse Fourier transform*

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\xi\, \mathrm{e}^{\mathrm{i}\xi x}\widehat{f}(\xi). \tag{10.3.5}$$

Strictly speaking, we should place some conditions on the function f above (e.g., that it is integrable and continuous). But, for our purposes, we do not need to get into too much technical detail here. We will use Fourier transforms to solve PDE (10.3.3). To begin, we multiply (10.3.3) and (10.3.4) by $\mathrm{e}^{\mathrm{i}\xi x}$ and integrate with respect to x. Defining the Fourier transform of u

$$\widehat{u}(t,\xi,y) := \int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}\xi x} u(t,x,y),$$

and noting that

$$\begin{aligned}
\int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}\xi x} \partial_t u(t,x,y) &= \partial_t \int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}\xi x} u(t,x,y) = \partial_t \widehat{u}(t,\xi,y), \\
\int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}\xi x} \partial_x u(t,x,y) &= -\int_{\mathbb{R}} \mathrm{d}x \, u(t,x,y) \partial_x \mathrm{e}^{-\mathrm{i}\xi x} = \mathrm{i}\xi \widehat{u}(t,\xi,y), \\
\int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}\xi x} \partial_x^2 u(t,x,y) &= \int_{\mathbb{R}} \mathrm{d}x \, u(t,x,y) \partial_x^2 \mathrm{e}^{-\mathrm{i}\xi x} = -\xi^2 \widehat{u}(t,\xi,y),
\end{aligned}$$

we obtain the following PDE for $\widehat{u}$

$$0 = \Big(\partial_t + \tfrac{1}{2} y(-\xi^2 - \mathrm{i}\xi) + \rho \mathrm{i}\xi \delta y \partial_y + \tfrac{1}{2}\delta y \partial_y^2 + \kappa(\theta - y)\partial_y \Big) \widehat{u}(t,\xi,y), \tag{10.3.6}$$

$$\widehat{u}(\mathrm{T},\xi,y) = \widehat{\varphi}(x) := \int_{\mathbb{R}} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}\xi x} \varphi(x). \tag{10.3.7}$$

To solve PDE (10.3.6) we guess that

$$\widehat{u}(t,\xi,y) = \widehat{\varphi}(\xi) \mathrm{e}^{\mathrm{C}(t,\xi) + y\mathrm{D}(t,\xi)}. \tag{10.3.8}$$

where we impose that

$$\mathrm{C}(\mathrm{T},\xi) = 0, \qquad \mathrm{D}(\mathrm{T},\xi) = 0, \qquad \text{so that} \qquad \widehat{u}(\mathrm{T},x,y) = \widehat{\varphi}(\xi).$$

Inserting our guess (10.3.8) into (10.3.7), using

$$\begin{aligned}
\partial_t \widehat{u}(t,\xi,y) &= \Big(\partial_t \mathrm{C}(t,\xi) + y \partial_t \mathrm{D}(t,\xi) \Big) \widehat{u}(t,\xi,y), \\
\partial_y \widehat{u}(t,\xi,y) &= \mathrm{D}(t,\xi) \widehat{u}(t,\xi,y), \\
\partial_y^2 \widehat{u}(t,\xi,y) &= \mathrm{D}^2(t,\xi) \widehat{u}(t,\xi,y),
\end{aligned}$$

and then dividing the result by $\widehat{u}(t,\xi,y)$, we obtain

$$0 = \partial_t \mathrm{C}(t,\xi) + y \partial_t \mathrm{D}(t,\xi) + \tfrac{1}{2} y(-\xi^2 - \mathrm{i}\xi) + \rho \mathrm{i}\xi \delta y \mathrm{D}(t,\xi) + \tfrac{1}{2}\delta y \mathrm{D}^2(t,\xi) + \kappa(\theta - y)\mathrm{D}(t,\xi).$$

The above equation must hold for *every* possible value of y. The only way for this to hold is for the terms that contain y and the terms that do not contain y to sum to zero. That is, we must have

$$0 = \partial_t \mathrm{C}(t,\xi) + \kappa\theta \mathrm{D}(t,\xi),$$

$$0 = \partial_t \mathrm{D}(t,\xi) + \tfrac{1}{2}(-\xi^2 - \mathrm{i}\xi) + \rho \mathrm{i}\xi\delta \mathrm{D}(t,\xi) + \tfrac{1}{2}\delta \mathrm{D}^2(t,\xi) - \kappa \mathrm{D}(t,\xi).$$

The second equation above is an Ordinary Differential Equation (ODE) in t for the function $\mathrm{D}(t,\xi)$. Specifically, it is an *Riccati* ODE, which can be solved explicitly. Once we have found the function $\mathrm{D}(t,\xi)$ we can obtain $\mathrm{C}(t,\xi)$ by integrating. From $-\mathrm{dC}(t,\xi) = \kappa\theta \mathrm{D}(t,\xi)\mathrm{d}t$ we obtain

$$-\int_t^{\mathrm{T}} \mathrm{dC}(t,\xi) = -\Big(\mathrm{C}(\mathrm{T},\xi) - \mathrm{C}(t,\xi)\Big) = \mathrm{C}(t,\xi) = \kappa\theta \int_t^{\mathrm{T}} \mathrm{D}(s,\xi)\mathrm{d}s.$$

We will not give the expressions for C and D here, as the formulas are rather complicated and we would not learn much by staring at them. A curious reader can fend expressions for C and D in any paper written on the Heston model (of which there are *hundreds*). To obtain $u(t,x,y)$ from $\widehat{u}(t,\xi,y)$ we make use of the inverse Fourier transform (10.3.5). We have

$$\begin{aligned} u(t,x,y) &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\xi\, \mathrm{e}^{\mathrm{i}\xi x}\widehat{u}(t,\xi,y) \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}\xi\, \mathrm{e}^{\mathrm{i}\xi x}\widehat{\varphi}(\xi)\mathrm{e}^{\mathrm{C}(t,\xi)+y\mathrm{D}(t,\xi)}. \end{aligned}$$

Finally, to obtain the option price $v(t,s,y)$, simply use $v(t,s,y) = u(t,x(s),y) = u(t,\log s,y)$.

10.4 Calibration of the Heston model

Let us denote the price of a Call option in the Heston model with Strike K and maturity T as follows

$$\mathrm{C}^{\mathrm{H}}(t,\mathrm{S}_t;\mathrm{T},\mathrm{K},\Phi), \qquad\qquad \Phi := (\mathrm{Y}_t,\kappa,\theta,\delta,\rho).$$

Note that, while $(t,\mathrm{S}_t,\mathrm{T},\mathrm{K})$ are all observable, the parameters $\Phi := (\mathrm{Y}_t,\kappa,\theta,\delta,\rho)$ are *not* directly observable and must be found via calibration. Roughly speaking, *calibration* is the process of finding the unobservable parameters Φ that minimize the difference (in some sense) between model-induced Call prices and market quotes. Because Call prices with differing strikes and maturities have wildly different prices, it is common to perform a calibration procedure with to implied volatilities, which, as unitless quantities, do not vary too much across strikes and maturities. In this way, Call options with different strikes and maturities are more equally weighted in the calibration procedure.

With the above in mind, we define the *Heston implied volatility* $\sigma^{\mathrm{H}}(\mathrm{T},\mathrm{K};\Phi)$ and the *Market implied volatility* $\sigma^{\mathrm{M}}(\mathrm{T},\mathrm{K})$ as the unit positive solutions of

$$\begin{aligned} \mathrm{C}^{\mathrm{BS}}(t,\mathrm{S}_t;\sigma^{\mathrm{H}}(\mathrm{T},\mathrm{K};\Phi),\mathrm{T},\mathrm{K}) &= \mathrm{C}^{\mathrm{H}}(t,\mathrm{S}_t;\mathrm{T},\mathrm{K},\Phi), \\ \mathrm{C}^{\mathrm{BS}}(t,\mathrm{S}_t;\sigma^{\mathrm{M}}(\mathrm{T},\mathrm{K}),\mathrm{T},\mathrm{K}) &= \mathrm{C}_t^{\mathrm{M}}(\mathrm{T},\mathrm{K}). \end{aligned}$$

where, as usual, $C^{BS}(t, S_t; \sigma, T, K)$ denotes the price of a call option in the Black-Scholes setting with maturity T, strike K and volatility σ. In a typical calibration procedure, the optimal unobservable parameters Φ^* are defined as follows

$$\Phi^* := \arg\min_\Phi = \sum_{(T,K)} \left(\sigma^M(T,K) - \sigma^H(T,K,\Phi)\right)^2 .$$

If the Heston model truly captured the risk-neutral dynamics of the underlying S, once the optimal parameters $\Phi^* = (Y_t^*, \kappa^*, \theta^*, \delta^*, \rho^*)$ are obtained, all of them except Y_t^* (which fluctuates randomly over time) would remain fixed forever. In practice, however, banks will re-calibrate on a daily basis. Of course, this means that the Heston model does not truly capture the risk-neutral dynamics of the underlyer. But, recalibration is a standard industry practice.

At this point, one might ask the following: *if we used market prices to calibrate a model, why do we even need a model? Would it not be better to just buy and sell option at the market price?* There are various reasons to need a model for an underlying. First, a model is required in order to derive replication strategies for options. It is entirely possible that model A and model B result in identical European option prices but result in different replication strategies. Second, a model is needed to price path-dependent (i.e., not European) derivatives.

10.5 Exercises

Exercise 10.1. To do.

Chapter 11

Barrier Options

So far, we have focused mainly on European derivatives. These are the simplest type of derivatives because their final values depend only on the final value of the underlying asset. There are, however, may derivatives whose values depend on the entire path of the underlying from inception to maturity. In this chapter we will learn how to price and replicate a variety of *Barrier options*, which are *path-dependent* derivatives.

Let τ_m be the first time S hits a level m. Mathematically, we can write this as

$$\tau_m := \inf\{t \geq 0 : \mathrm{S}_t = m\}.$$

Note that τ_m is a random time. The payoff of a variety of Barrier options can be expressed in terms of τ_m. In what follows, assume that $\mathrm{L} < \mathrm{S}_0 < \mathrm{U}$.

$$\begin{aligned} \text{Knock-out Payoff} &= \phi(\mathrm{S}_\mathrm{T})\mathbb{1}_{\{\tau_m > \mathrm{T}\}}, \\ \text{Knock-in Payoff} &= \phi(\mathrm{S}_\mathrm{T})\mathbb{1}_{\{\tau_m \leq \mathrm{T}\}}. \end{aligned}$$

Recall that an *indicator function* $\mathbb{1}$ is defined as follows

$$\mathbb{1}_\mathrm{A} = \begin{cases} 1 & \text{if event A occurs,} \\ 0 & \text{if event A does not occur.} \end{cases}$$

Thus, a knock-out option pays the option holder $\phi(\mathrm{S}_\mathrm{T})$ only if S does not hit a level m prior to the maturity date T. A knock-in option, on the other hand, pays the option holder $\phi(\mathrm{S}_\mathrm{T})$ only if S hits a level m prior to the maturity date T.

11.1 Parity Relations

Let $\mathrm{V}^{\mathrm{E}} = (\mathrm{V}_t^{\mathrm{E}})_{0\le t\le \mathrm{T}}$, $\mathrm{V}^{\mathrm{KO}} = (\mathrm{V}_t^{\mathrm{KO}})_{0\le t\le \mathrm{T}}$ and $\mathrm{V}^{\mathrm{KI}} = (\mathrm{V}_t^{\mathrm{KI}})_{0\le t\le \mathrm{T}}$ denote, respectively, the values of a European, Knock-in and Knock-out options, each with same payoff function ϕ. Noting that $1 = \mathbb{1}_{\{\tau_m > \mathrm{T}\}} + \mathbb{1}_{\{\tau_m \le \mathrm{T}\}}$, we have

$$\phi(\mathrm{S_T}) = \phi(\mathrm{S_T})\mathbb{1}_{\{\tau_m > \mathrm{T}\}} + \phi(\mathrm{S_T})\mathbb{1}_{\{\tau_m \le \mathrm{T}\}}.$$

Dividing both sides of the above equation by $\mathrm{B_T}$, and taking conditional expectations under the risk-neutral pricing measure $\widetilde{\mathbb{P}}$, we obtain

$$\begin{aligned}
& & \widetilde{\mathbb{E}}\left[\frac{\phi(\mathrm{S_T})}{\mathrm{B_T}}\middle|\mathscr{F}_t^{\mathrm{S}}\right] &= \widetilde{\mathbb{E}}\left[\frac{\phi(\mathrm{S_T})\mathbb{1}_{\{\tau_m > \mathrm{T}\}}}{\mathrm{B_T}}\middle|\mathscr{F}_t^{\mathrm{S}}\right] + \widetilde{\mathbb{E}}\left[\frac{\phi(\mathrm{S_T})\mathbb{1}_{\{\tau_m \le \mathrm{T}\}}}{\mathrm{B_T}}\middle|\mathscr{F}_t^{\mathrm{S}}\right] \\
\Rightarrow & & \frac{\mathrm{V}_t^{\mathrm{E}}}{\mathrm{B}_t} &= \frac{\mathrm{V}_t^{\mathrm{KO}}}{\mathrm{B}_t} + \frac{\mathrm{V}_t^{\mathrm{KI}}}{\mathrm{B}_t} \\
\Rightarrow & & \mathrm{V}_t^{\mathrm{E}} &= \mathrm{V}_t^{\mathrm{KO}} + \mathrm{V}_t^{\mathrm{KI}}.
\end{aligned} \tag{11.1.1}$$

Equation (11.1.1) is a *knock-in knock-out parity* relation. If we can compute or observe the values of a European and knock-out options, then we can use (11.1.1) to deduce the value of a knock-in option. Note that the above result is a model-free result, in the sense that we have not had to assume anything about the dynamics of the underlying asset S.

11.2 Static Hedging of Barrier options

In this section, we will show how to price and replicate a knock-out barrier option under the assumption that $r = 0$ and that S follows a geometric Brownian motion (i.e., the Black-Scholes model). The price a replication strategy for a knock-in option can be deduced from (11.1.1). Under the risk-neutral pricing measure, the dynamics of S are given by

$$\mathrm{dS}_t = \sigma \mathrm{S}_t d\widetilde{\mathrm{W}}_t, \tag{11.2.1}$$

where $\widetilde{\mathrm{W}}$ is a Brownian motion under $\widetilde{\mathbb{P}}$.

Lemma 11.1 (Put-Call Symmetry). *Let* S *be given by* (11.2.1)*. Then, for any function* $\mathrm{G} : \mathbb{R}_+ \to \mathbb{R}$ *we have*

$$\widetilde{\mathbb{E}}[\mathrm{G}(\mathrm{S_T})|\mathrm{S}_t] = \widetilde{\mathbb{E}}\left[\frac{\mathrm{S_T}}{\mathrm{S}_t}\mathrm{G}\left(\frac{\mathrm{S}_t^2}{\mathrm{S_T}}\right)\middle|\mathrm{S}_t\right]. \tag{11.2.2}$$

Proof. Let $X = (X_t)_{0 \le t \le T}$ be defined as follows $X_t = \log S_t$. Then we have

$$X_T | X_t = X_t - \tfrac{1}{2}\sigma^2(T-t) + \sigma(\widetilde{W}_T - \widetilde{W}_t) \sim N\Big(X_t - \tfrac{1}{2}\sigma^2(T-t), \sigma^2(T-t)\Big).$$

We compute

$$\begin{aligned}
\widetilde{\mathbb{E}}[G(S_T)|S_t] &= \widetilde{\mathbb{E}}[G(e^{X_T})|X_t] \\
&= \int_{\mathbb{R}} dx\, G(e^x) \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(\frac{-(x - X_t + \frac{1}{2}\sigma^2(T-t))^2}{2\sigma^2(T-t)}\right) \\
&= \int_{\mathbb{R}} dy\, G(e^{2X_t - y}) \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(\frac{-(-y + X_t + \frac{1}{2}\sigma^2(T-t))^2}{2\sigma^2(T-t)}\right) && (x = 2X_t - y) \\
&= \int_{\mathbb{R}} dy\, e^{y - X_t} G(e^{2X_t - y}) \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(\frac{-(y - X_t + \frac{1}{2}\sigma^2(T-t))^2}{2\sigma^2(T-t)}\right) \\
&= \widetilde{\mathbb{E}}[e^{X_T - X_t} G(e^{2X_t - X_T})|X_t] \\
&= \widetilde{\mathbb{E}}\left[\frac{S_T}{S_t} G\Big(\frac{S_t^2}{S_T}\Big) \Big| S_t\right],
\end{aligned}$$

which establishes (11.2.2). □

We can use Put-Call symmetry in order to statically replicated a knock-out barrier option.

Theorem 11.2. *Let* S *be given by* (11.2.1). *Suppose* $L < S_0$ *and consider a knock-out option with maturity* T *and payoff* $\phi(S_T)\mathbb{1}_{\{\tau_L > T\}}$. *To replicate this claim, purchase at time* $t = 0$ *a European claim with payoff*

$$\phi^{KO}(S_T) = \mathbb{1}_{\{S_T > L\}}\phi(S_T) - \mathbb{1}_{\{S_T < L\}} \frac{S_T}{L} \phi\Big(\frac{L^2}{S_T}\Big).$$

If and when the barrier claim knocks out, clear the position in $\phi^{KO}(S_T)$ *at no cost.*

Proof. Suppose that $\tau_L > T$. This implies that $S_T > L$ and thus

$$\phi^{KO}(S_T) = \phi(S_T), \qquad \text{(when } \tau_L > T)$$

The knock-out claim also pays $\phi(S_T)$ when $\tau_L > T$. Suppose, on the other hand, that $\tau_L \le T$. Then, at time τ_L we have

$$\begin{aligned}
\widetilde{\mathbb{E}}[\mathbb{1}_{\{S_T > L\}}\phi(S_T)|S_{\tau_L}] &= \widetilde{\mathbb{E}}\left[\mathbb{1}_{\{S_{\tau_L}^2/S_T > L\}} \frac{S_T}{S_{\tau_L}} \phi\Big(\frac{S_{\tau_L}^2}{S_T}\Big)\Big| S_{\tau_L}\right] && \text{(by (11.2.2))} \\
&= \widetilde{\mathbb{E}}\left[\mathbb{1}_{\{L^2/S_T > L\}} \frac{S_T}{L} \phi\Big(\frac{L^2}{S_T}\Big)\Big| S_{\tau_L}\right] && \text{(because } S_{\tau_L} = L)
\end{aligned}$$

$$= \widetilde{\mathbb{E}}\left[\mathbb{1}_{\{L>S_T\}}\frac{S_T}{L}\phi\left(\frac{L^2}{S_T}\right)\Big|S_{\tau_L}\right],$$

and thus $\widetilde{\mathbb{E}}[\phi^{KO}(S_T)|S_{\tau_L}] = 0$. □

Example 11.3. Consider a knock-out call option with payoff

$$(S_T - K)^+\mathbb{1}_{\{\tau_H>T\}}, \qquad 0 < L < S_0 < K. \tag{11.2.3}$$

Let us find the European option that statically replicates this claim. First we identify

$$\phi(S_T) = (S_T - K)^+.$$

Then, according to Theorem 11.2, at time zero we should buy a European claim that pays

$$\begin{aligned}
\phi^{KO}(S_T) &= \mathbb{1}_{\{S_T>L\}}(S_T - K)^+ - \mathbb{1}_{\{S_T<L\}}\frac{S_T}{L}\left(\frac{L^2}{S_T} - K\right)^+ \\
&= (S_T - K)^+ - \mathbb{1}_{\{S_T<L\}}\frac{K}{L}\left(\frac{L^2}{K} - S_T\right)^+ \\
&= (S_T - K)^+ - \frac{K}{L}\left(\frac{L^2}{K} - S_T\right)^+,
\end{aligned}$$

where, in the last line we have used that $S_T < L^2/K$ implies $S_T < L$. Thus, to statically replicate the knock-out call option (11.2.3), at time $t = 0$ one should buy a single European call option with Strike K and sell K/L European put options with strike L^2/K. Now, consider a knock-in call option with payoff

$$(S_T - K)^+\mathbb{1}_{\{\tau_H\le T\}}, \qquad 0 < L < S_0 < K.$$

How would you statically replicate this claim? *Hint*: use (11.1.1).

11.3 Exercises

Exercise 11.1. Assume that the risk-free rate of interest is zero $r = 0$ and the dynamics of S under the risk-neutral pricing measure $\widetilde{\mathbb{P}}$ are given by $dS_t = \sigma S_t d\widetilde{W}_t$. Recall the put-call symmetry (PCS) result

$$\widetilde{\mathbb{E}}[G(S_T)|S_t] = \widetilde{\mathbb{E}}[(S_T/S_t)G(S_t^2/S_T)|S_t].$$

Let $\tau_L := \inf\{t \ge 0 : S_t = L\}$ be the first hitting time of S to level $L < S_0$.
(a) Consider a knock-out option with a payoff $\mathbb{1}_{\{\tau_L>T\}}\theta(S_T - K)$ where θ is a Heaviside function (i.e. $\theta(x) = 1$ if $x \ge 0$ and $\theta(x) = 0$ if $x < 0$). Assume $K > L$. Find the payoff function $\phi(S_T)$ of a European claim that has the same value of the knock-out option at any time $t \le T \wedge \tau_L$ (i.e., find the payoff

function of a European claim that statically hedges the knock-out claim).
(b) Consider a knock-in option with a payoff $\mathbb{1}_{\{\tau_{\mathrm{L}} \leq \mathrm{T}\}}\theta(\mathrm{S}_{\mathrm{T}} - \mathrm{K})$. Assume $\mathrm{K} > \mathrm{L}$. Find the payoff function $\psi(\mathrm{S}_{\mathrm{T}})$ of a European claim that has the same value of the knock-in option at any time $t \leq \mathrm{T} \wedge \tau_{\mathrm{L}}$ (i.e., find the payoff function of a European claim that statically hedges the knock-in claim).
(c) Plot the functions $\phi(\mathrm{S}_{\mathrm{T}})$ and $\psi(\mathrm{S}_{\mathrm{T}})$ obtained in parts (a) and (b) for $\mathrm{H} = 1$ and $\mathrm{K} = 2$.

Bibliography

Carr, P. and D. Madan (1998). Towards a theory of volatility trading. *Volatility: new estimation techniques for pricing derivatives*, 417.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud. 6*(2), 327–343.