Contents

1 Interest Rates .......................................................... 5
   1.1 Rate of return .................................................. 5
   1.2 Interest rates .................................................. 6
   1.3 Interest rate conventions ...................................... 7
   1.4 Continuous compounding ...................................... 9
   1.5 Effective annual rate ......................................... 11
   1.6 Continuous compounding, a second look ................. 12
   1.7 Zero Rates .................................................... 14
   1.8 Forward Rates ................................................. 15
   1.9 Force of interest (optional) ................................. 16
   1.10 Problems ...................................................... 18

2 Bonds ...................................................................... 21
   2.1 Factors affecting interest rates ............................ 21
   2.2 Risk-free rate ................................................... 22
   2.3 Value of money ................................................ 24
   2.4 Bonds ............................................................. 26
   2.5 LIBOR and Treasuries as measures of the risk-free rate .................................................. 30
   2.6 Zero coupon bonds and more lingo ...................... 31
   2.7 Forward rate relationship ................................. 32
   2.8 Bond duration .................................................. 37
   2.9 Parallel shift in the zero curve ......................... 38
   2.10 Convexity of a bond .................................... 41
   2.11 Yield ............................................................ 42
   2.12 Portfolio of bonds ........................................... 45
   2.13 Problems ...................................................... 46
## CONTENTS

3 Swaps ................................................................. 49

3.1 Bootstrapping ..................................................... 49
3.2 Floating-rate bonds .............................................. 51
3.3 Fixed-floating swaps ............................................ 55
3.4 Deducing swap rates ............................................ 58
3.5 Currency swaps .................................................. 60
3.6 Credit default swaps ............................................ 62
3.7 Problems .......................................................... 62

4 Survey of markets ................................................... 65

4.1 Instruments for raising money ................................. 65
  4.1.1 Bonds ......................................................... 66
  4.1.2 Other loans ................................................ 67
  4.1.3 Equity ......................................................... 68
4.2 Derivatives ......................................................... 69
  4.2.1 Forwards ...................................................... 70
  4.2.2 Futures ....................................................... 70
  4.2.3 Options ....................................................... 71
4.3 Trading on a market versus over-the-counter (OTC) .... 71
4.4 Selling short ...................................................... 72
4.5 Problems .......................................................... 74

5 Forward contracts .................................................. 75

5.1 What is a forward contract? .................................. 75
5.2 The forward price versus the spot price .................. 76
5.3 Computing forward prices .................................... 77
5.4 Valuing a long or short position in a forward contract . 82
5.5 Problems .......................................................... 83

6 Options .............................................................. 85

6.1 Definitions ......................................................... 85
6.2 Some examples .................................................. 87
6.3 Payoff of a portfolio ............................................ 88
6.4 Some common portfolios with options .................... 91
  6.4.1 Straddle ...................................................... 91
  6.4.2 Spreads ....................................................... 92
6.5 The fundamental principle of math finance ............... 93
6.6 Inequalities for option prices ............................... 95
## 11 The Black-Scholes formula

- 11.1 The efficient market hypothesis .................................................. 166
- 11.2 The log-normal model ................................................................. 166
- 11.3 Volatility and expected return of a stock ..................................... 169
- 11.4 Risk neutral pricing ................................................................. 172
- 11.5 Digital options ........................................................................ 175
- 11.6 Problems ................................................................................. 176

## 12 Hedging

- 12.1 Hedging stock positions with forwards ........................................ 180
- 12.2 Hedging bond positions with bonds ......................................... 182
- 12.3 Hedging stock positions with options ........................................ 183
- 12.4 Hedging option positions with stocks ....................................... 184
- 12.5 Problems ................................................................................. 186
In our day to day lives, we are likely to interact with interest rates more than any other topic in this course. The money we deposit in our savings or checking accounts earns interest, whereas our student loans, car loans, and credit card debts require us to pay interest. The basic definitions and conventions about interest rates are the subject of the first chapter.

1.1 Rate of return

An investment will hopefully earn money over time. Suppose an initial investment of $1000 grows to $1250 after 1 month. We have earned $250. Since we expect many investments to earn money in proportion to the amount of initial money invested, we might express our earnings not by quoting the absolute amount of $250, but as a fraction or percentage, of the initial investment. In this case,

\[
\frac{250}{1000} = 0.25 = 25\%
\]

is the rate of return of the investment.

Suppose $P$ dollars is placed in an investment, and the investment becomes worth $Q$ dollars at some later time $T$. Then the rate of return $r$ is given by the formula

\[
r = \frac{Q - P}{P} = \frac{Q}{P} - 1
\]

If $r > 0$, the investment has earned money, whereas if $r < 0$ the investment has lost money.
One of the reasons we use the rate of return, rather than an absolute dollar amount, is that it permits us to compare investments that have a different initial investment. It is not a good comparison to look at two investments that both grew by $250, when one had an initial amount of $1000 and another started with $20,000. A second reason is that for many investments, the money earned is expected to be proportional to the initial amount \( P \), or **principal**, that is invested.

### 1.2 Interest rates

In many situations, we already know at the start what the rate of return will be. The promise of a fixed rate of return is encoded in an **interest rate**, denoted by \( i \).

If we know that we will receive an interest rate equal to \( i \), then we can reverse Equation (1.1) from the previous section. The formula for the final amount \( Q \), given the principal \( P \), becomes

\[
Q = P(1 + i). \tag{1.2}
\]

As a matter of notation, the amount that we earn above the principal amount is referred to as the **interest earned**. In the present case, the interest earned is equal to \( Pi \). The total return \( Q \) is therefore the principal amount \( P \) plus the interest earned \( Pi \).

Next, we consider the situation where every time period \( T \) we receive the interest rate \( i \) applied to all the money we have earned up until that point.

After one time period has elapsed, the principal \( P \) has grown to \( P(1 + i) \) dollars. But then this amount becomes the principal for the second time period. In other words, we start the second time period with \( P(1 + i) \) dollars and earn an interest rate of \( i \). Hence, after 2 time periods, the investment is worth

\[
[P(1 + i)](1 + i) = P(1 + i)^2
\]
dollars. This amount then earns interest in the third time period and becomes worth

\[
[P(1 + i)^2](1 + i) = P(1 + i)^3
\]
dollars.

Continuing in this manner, we see that after \( N \) time periods have elapsed, the investment is worth

\[
Q = P(1 + i)^N \tag{1.3}
\]
dollars. The equation above illustrates the concept of *compounding*. At each time period, we get to earn interest on the amount that the investment is worth at the beginning of the time period.

**Example 1.1.** You deposit $500 in a savings account which will pay a 6% interest rate every year. Compute the amount of money in the account after 1, 2, 3, and 4 years. How much interest is earned in the first year? in the fourth year?

**Solution:** The formula for compounding shows that the amount in the account is

\[ 500(1 + .06)^N \]

at the end of the \( N \)-th year. Plugging in \( N = 1, 2, 3, 4 \) gives 530, 561.8, 595.51, 631.24 dollars, respectively. To compute the interest earned at the end of year \( N \), we can subtract the total amount at the beginning of year \( N \) from the total amount at the end of year \( N \). For \( N = 1 \), we get 530 – 500 = 30 dollars of interest. For \( N = 4 \), we get 631.24 – 595.51 = 35.73 dollars of interest.

### 1.3 Interest rate conventions

It will be helpful to fix some terminology regarding interest rates. In order to quote an interest rate, we must specify the time period over which the rate applies. In most cases, the convention for interest rates is to quote the interest rate for each year and then to introduce compounding during the year at regular intervals.

The convention is as follows. First, we declare the interest rate by stating that an investment pays a rate \( i \) of interest per year or, using the Latin phrase, *per annum*. Note, that we will switch back and forth between writing the interest rate as a percent and as a decimal, so that a rate of \( i \) is the same as a rate of \( 100i \) percent, written \( 100i\% \). Second, we declare how often the investment will be compounded throughout the year by stating that the interest rate will be compounded \( n \) times per year, at equal time intervals.

With these two pieces of information, the convention is that \( n \) times per year, at regular intervals, the investment will be compounded at the interest rate of \( i/n \). This is purely a convention. Once this convention is understood, we can deduce from the previous section, how to value the investment after \( t \) years. In order to do that, we need to figure out how many times in \( t \)
years the investment is compounded. Since we are compounding $n$ times per year and there are $t$ years, the total number of times the investment is compounded is $nt$.

**Formula 1 (Interest Rate Convention).** The final amount of an investment paying an interest rate of $i$ per annum, with compounding $n$ times per year for $t$ years, is

$$Q = P \left( 1 + \frac{i}{n} \right)^{nt}. \quad (1.4)$$

One comment is in order. This formula makes sense only if the number of years $t$ is an multiple of $1/n$. Equivalently, the number $nt$ must be a positive integer. For example, if $n = 2$, so that compounding occurs twice per year, then this formula makes sense whenever $t = 0.5, 1, 1.5, 2, 2.5, \ldots$ years, but not otherwise. We will address a way around this dilemma in the next section.

Finally there are conventions dealing with the number of times per year that compounding happens. These are self-explanatory. See Table 1.1.

<table>
<thead>
<tr>
<th>$n$ times per year</th>
<th>Naming convention</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>annual</td>
</tr>
<tr>
<td>2</td>
<td>semiannual</td>
</tr>
<tr>
<td>4</td>
<td>quarterly</td>
</tr>
<tr>
<td>12</td>
<td>monthly</td>
</tr>
<tr>
<td>52</td>
<td>weekly</td>
</tr>
</tbody>
</table>

**Example 1.2.** A savings account pays 4.5% per annum with quarterly compounding. This means that every quarter the account accrues interest at the rate of $\frac{4.5}{4} = 1.125\%$. After 1 year, if the initial deposit is $600, then the account has a balance (assuming no withdrawals or further deposits) of

$$600 \left( 1 + \frac{.045}{4} \right)^4 = 627.46$$

\(^1\text{Recall that a positive integer is a number of the form 1, 2, 3, 4, \ldots.}\)
dollars. After 1.25 years, the balance is

\[ 600 \left( 1 + \frac{.045}{4} \right)^5 = 634.52 \]

dollars since the number of quarters in 1.25 years is \( 4 \times 1.25 = 5 \). Similarly, if \( t = 2.75 \) years, then the exponent in the formula would be 11.

Note that it is not clear what to do if the amount of time is not a multiple .25 years; for example, if \( t = 1.1 \), we are uncertain how to compute the amount. We resolve this issue in the next section.

You should be able to convert between any two compounding schemes.

**Example 1.3.** If a loan carries an interest rate of 3% per annum with quarterly compounding, this is equivalent to what per annum rate with semiannual compounding?

**Solution:** Solve for \( r \) in the expression

\[ \left( 1 + \frac{r}{2} \right)^2 = \left( 1 + \frac{.03}{4} \right)^4. \]

Then \( r = 0.0301 = 3.01\% \).

### 1.4 Continuous compounding

In the previous section, we introduced the formula for the amount of an investment after \( t \) years with compounding \( n \) times per year at the rate of \( r \) per annum:

\[ Q = P \left( 1 + \frac{r}{n} \right)^{nt}. \quad (1.5) \]

In this section, we study what happens to the formula as \( n \) goes to infinity, so that in a sense compounding takes place at every instant.

If we bring in our knowledge of calculus, then this becomes an application of L'Hopital's rule. Let

\[ y = \left( 1 + \frac{r}{n} \right)^{nt}. \]

We want to compute

\[ \lim_{n \to \infty} y. \]
As \( n \) tends to infinity this expression becomes the indeterminate form \( 1^\infty \). To get around this, we take the natural log of \( y \), yielding

\[
\ln y = nt \ln \left(1 + \frac{r}{n}\right) = \frac{\ln \left(\frac{n+r}{n}\right)}{\frac{1}{nt}}.
\]

Taking the limit as \( n \) goes to infinity now gives the indeterminate form \( 0/0 \), the typical candidate for L’Hopital’s Rule. So we take the derivative of the numerator and divide by the derivative of the denominator (both derivatives are with respect to \( n \)) and get

\[
\frac{n}{n+r} \cdot \frac{r}{n^2} = \frac{nrt}{n + r} = \frac{rt}{1 + \frac{r}{n}}
\]

after simplifying. Taking the limit as \( n \) goes to infinity, this quantity converges to \( rt \). Hence,

\[
\lim_{n \to \infty} \ln y = rt,
\]

and so

\[
\lim_{n \to \infty} y = e^{rt}.
\]

Here, we have used the identity \( y = e^{\ln y} \), along with the fact that \( e^x \) is a continuous function.

In conclusion, we have shown that if we allow the number of times \( n \) that we compound per year to increase to infinity, then Formula (1) becomes \( Q = Pe^{rt} \). This is known as continuous compounding.

**Formula 2 (Continuous compounding).** If an interest rate \( r \) is offered per annum with continuous compounding, then a principal \( P \) will grow to

\[
Q = Pe^{rt}
\]

(1.6)

after \( t \) years, where \( t \) is any number.

If an investment grows according to this formula, we say that it earns interest at the rate of \( r \) per annum with **continuous compounding** or that the investment **compounds continuously** at the rate of \( r \) per annum. A key advantage of this framework is that \( t \) can take on any value of time and the formula will still make sense.
Example 1.4. A bank account pays 3% per annum with continuous compounding. An initial deposit of 1,300 dollars will be worth how much after 2 weeks? after 4 months? after 1 year? How much interest is earned between the sixth month and the ninth month?

Solution: Using Formula (1.6) with \( r = 0.03 \) and \( t = 2/52 = 0.03846 \) gives
\[
1300e^{0.03 \cdot 0.03846} = 1301.50
\]
Similarly, for \( t = 4/12 = 0.3333 \) gives
\[
1300e^{0.03 \cdot 0.3333} = 1313.06
\]
And for \( t = 1 \) gives
\[
1300e^{0.03} = 1339.59
\]
The interest earned between the sixth and ninth month is
\[
1300e^{0.03 \cdot \frac{6}{12}} - 1300e^{0.03 \cdot \frac{6}{12}} = 1329.58 - 1319.65 = 9.94
\]
dollars.

1.5 Effective annual rate

With all these different ways to compound per year, it can be difficult to compare investments. A simple way to compare the returns is to compute the effective annual rate. The effective annual rate, denoted \( r_{\text{eff}} \), is the annual interest rate (with compounding once per year) that would give the equivalent return after \( t = 1 \) years.

For example, to go from continuous compounding at a rate of \( r \) per annum to the effective annual rate \( r_{\text{eff}} \), we find the growth from continuous compounding after 1 year, which is \( Pe^{r} \). Then we solve the equation
\[
P(1 + r_{\text{eff}}) = Pe^{r},
\]
which gives \( r_{\text{eff}} = e^{r} - 1 \). In the previous example, the effective annual rate is
\[
e^{0.03} - 1 = 0.0305 = 3.05%,
\]
just a tad more than 3%.
Example 1.5. If an investment pays 3% per annum with quarterly compounding, what is the effective annual rate?

Solution: We need to solve for \( r_{\text{eff}} \) in the expression

\[
1 + r_{\text{eff}} = \left( 1 + \frac{.03}{4} \right)^4.
\]

Then \( r_{\text{eff}} = 3.034\% \).

It is possible to show for a fixed interest rate \( i \) that as the number \( n \) of times compounding takes place per year increases, the effective annual rate increases. In other words, holding \( i \) fixed, but increasing \( n \), leads to a bigger value in Formula [1]. Furthermore, continuous compounding always leads to a higher effective annual rate than for any value of \( n \) in Formula [1]. See the exercises for a proof. Intuitively, this should make sense: the more frequently the compounding occurs, the more interest is earned on the earlier interest.

1.6 Continuous compounding, a second look

In the previous section, we saw that if continuous compounding occurs at a rate of \( r \) per annum, then the effective annual rate is determined by the equation \( 1 + r_{\text{eff}} = e^r \). Set \( a \) equal to this value.

Then an investment of \( P \) dollars grows to

\[
P e^{rt} = P (e^r)^t = Pa^t
\]

after \( t \) years. Rewriting this gives

\[
P (1 + r_{\text{eff}})^t,
\]

which looks no different than Formula [1] with \( n = 1 \), except that now we can view \( t \) as any number, not just a positive integer. Therefore, in a sense, we probably could have guessed the formula for continuous compounding without having to derive it as we did in Section 1.4.

Formula 3 (Continuous compounding, Version 2). If the effective annual rate is \( r_{\text{eff}} \), then a principal \( P \) will grow to

\[
Q = Pa^t
\]

after \( t \) years, where \( t \) is any number and \( a = 1 + r_{\text{eff}} \).
Example 1.6. A bank offers a one-year Certificate of Deposit (CD) which will earn 4.5% after one year. You sign up for the CD. After one year, let’s say that your investment has grown to $5,100. In general, with a CD, the bank will notify you a few weeks prior to the CD coming due and give you the opportunity to move your money. If you do not act, the money will be rolled over into a new CD with the same time structure, but with the current interest rate. However, you generally have 2 weeks after the new CD begins to remove your money from the new CD without paying a penalty if you decide you want to move the money.

Let’s suppose that the new one-year CD is only offering 3% and that you are rolled over automatically into the new CD. But after the 2 weeks, seeing the poor interest rate, you decide to remove your money. How much money would you receive in interest for the two week period?

Solution: The value of your account would be $5100 \times (1.03)^{2/52} = 5105.80$. The interest earned in the two week period is $5105.80 - 5100 = 5.80$. Note that it is not correct to compute the interest as $5100 \times (1.03)^{2/52}$. The latter value equals 20.28. Here, we have used Formula 1.6 with $t = 2/52 = 0.3846$.

Example 1.7. Suppose you are offered an investment with an effective annual rate of 6%. How long will it take for your principal to double?

Solution: We are looking to find the amount of time $t$ the solves the equation:

$$P(1 + .06)^t = 2P.$$ 

Dividing both sides by $P$ gives

$$1.06^t = 2.$$ 

To solve for $t$, we can take the natural log of both sides, yielding

$$\ln(1.06^t) = \ln(2).$$ 

Using the rules of logarithms leads to the equation: $t \ln(1.06) = \ln(2)$, or

$$t = \frac{\ln(2)}{\ln(1.06)} = 11.9.$$ 

This means that the principal will double after approximately 12 years. Note that the calculation is independent of the principal $P$.

In the remainder of the book, we will use Formula 3 for continuous compounding. Nevertheless, many books do prefer to work with the effective annual rate $r_{\text{eff}}$ and the growth factor $a = 1 + r_{\text{eff}}$, as in Formula 1.6.
1.7 Zero Rates

Up until now, we have implicitly assumed that interest rates are constant over time intervals. In the examples, we have looked at investments and loans where the rate is fixed for whatever time we were interested in. In fact, interest rates do change and unless we have a guarantee of a rate, we cannot make the assumption that an interest rate remains constant.

For example, whereas your savings account may promise an interest rate today, that rate could be slightly different tomorrow. On the other hand, there are many financial instruments where the rate is locked in over a specific time period. The notion of a zero rate is used to specify an interest rate that applies for a specific time period. The zero rate applies only to principal that is invested or borrowed at the beginning of the time period and that is returned (or repaid), with interest, at the end of the time period; nothing can be touched in the intervening time.

**Definition 1.8.** For a time period $T$, a $T$-year zero rate, denoted $r_T$, is an interest rate that applies to a loan or investment beginning today and lasting exactly $T$ years. During the intervening time, nothing can be done with the loan or investment. In particular, no interest is paid or received until the end of the time period. We typically quote zero rates per annum with continuous compounding. Zero rates are also referred to as spot rates.

At every instant, there are zero rates for each time period $T$. A Certificate of Deposit (CD), which banks offer on a regular basis, gives a good illustration of a zero rate. If you take out a 1-year CD, your money is locked up for that year. The bank promises to pay you a set rate for the entire year. Technically, it pays out interest each month, but if you choose to have the interest re-invested, then effectively you are putting your money away for exactly one year; nothing transpires in between. The effective interest rate that you receive on the CD is a 1-year zero rate. This rate is likely to be different from the 2-year zero rate and the 6-month zero rate. On August 7, 2008, E*Trade is offering a 1-year CD with an effective annual rate of 3.3%. The 6-month rate is 2.6% and the 2-year rate is 3.45%.

We often think of zero rates as tied to specific financial instruments. That is, a certain CD or loan at a specific bank implies a $T$-year zero rate.

\[^2\text{You can pull your money out, but you must pay a fee to do that and you may also lose accrued interest. We ignore this feature.}\]
for whatever time period $T$ that the CD or loan covers. Of course, zero rates change. Today a 1-year zero rate might be different than a 1-year zero rate tomorrow. In fact, on E*Trade’s webpage it says explicitly that rates are subject to change daily.

**Example 1.9.** Today the 2-year zero rate implied by a 2-year CD at Washington Mutual is 3.5%. If you invest $100 today for 2 years, what will the CD worth in 2 years? Suppose that tomorrow the rate rises to 3.53%. If you had waited until tomorrow to invest, what will your CD be worth when it comes due?

Solution: In the first scenario, the CD is worth $100e^{0.035(2)} = 107.25$. In the second, it is worth $100e^{0.0353(2)} = 107.32$.

## 1.8 Forward Rates

There is an important generalization of zero rates. Although it is hard to believe, sometimes financial actors want to know what interest rate will govern, for example, an investment that starts in 2 months and lasts for 2 years.

A good example of this is a mortgage rate. If you find a house you would like to buy, you do not acquire the house immediately; rather, you buy the house perhaps 1 or 2 months down the road. You are not exactly interested in today’s zero rates. Rather, today you are concerned about interest rates that start in 2 months (when you close on your house) and that last for 15 or 30 years. This is known as a forward rate.

**Definition 1.10.** For two times $T_1$ and $T_2$, with $T_1 < T_2$, a $(T_1, T_2)$-forward rate, denoted $r_{T_1, T_2}$, is an interest rate that applies to a loan or investment that begins in $T_1$-years and ends in $T_2$-years.

During the intervening time between $T_1$ and $T_2$ years, nothing can be done with the loan or investment. Additionally nothing transpires before $T_1$ years. After $T_1$ years have elapsed, the loan or investment begins. Until $T_2$ years have elapsed, no interest is paid or received. After $T_2$ years, the loan or investment is closed out with principal and interest being paid.

Note that a $T$-year zero rate is a special case of a forward rate. Namely, if $T_1 = 0$ and $T_2 = T$, then we have $r_{0,T} = r_T$. 
Example 1.11. Today, a bank offers a 6-month zero rate of 3.2% and a forward rate of \( r_{0.5,2} = 3.7\% \). Suppose that in six months, the 1.5-year zero rate is 3.9%. If an investor locks in the forward rate, then an initial investment of $100 grows to what amount in 2 years? On the other hand, suppose the investor does not lock in the forward rate. What will be the value of the investment?

Solution: Locking in the forward rate allows the investor to invest for 6-months at 3.2%, followed by investing for 1.5-years at 3.7%. The investment is worth

\[
100e^{0.032(0.5)}e^{0.037(1.5)} = 107.41.
\]

Not locking in the forward rate, leads to

\[
100e^{0.032(0.5)}e^{0.039(1.5)} = 107.73.
\]

Note that by locking in the forward rate, the investor is forced to deposit the money for the remaining 18 months at 3.7% instead of 3.9%.

1.9 Force of interest (optional)

In our discussion of interest rates so far, we have assumed that an interest rate on an investment does not change over time. But in the real world interest rates are constantly changing. A bank does not guarantee the interest rate in a checking account for any period of time. It may change daily, in fact.

Suppose that the interest rate on an investment at a given time \( t \) is given by the function \( r(t) \). We interpret this to mean that if we consider a tiny sliver of time \( \Delta t \), centered around the time \( t \), the interest rate is a constant equal to \( r(t) \), which is given per annum with continuous compounding.

Now imagine we have a principal amount \( P \) and we want to know how it grows over a time interval \( T \). We can break the interval into \( n \) intervals centered at the times

\[
t_1, t_2, \ldots, t_n
\]

with a width of \( \Delta t = T/n \). If we can assume that the interest rate \( r(t) \) is approximately constant over each interval, then it is possible to compute how much \( P \) grows. Over the first interval, \( P \) grows to

\[
P e^{r(t_1)\Delta t},
\]
since we are assuming that \( r(t) \) stays constant over the first small time interval. After the second interval, \( P \) has grown to

\[ Pe^{r(t_1) \Delta t} e^{r(t_2) \Delta t} = Pe^{(r(t_1) + r(t_2)) \Delta t}. \]

Continuing, we get after \( n \) steps

\[ Pe^{\sum_{i=1}^{n} r(t_i) \Delta t}. \]

As we chop up the big interval into more pieces, the expression in the exponent becomes the area under the curve defined by \( r(t) \) between 0 and \( T \), which is expressed as the integral

\[ \int_{0}^{T} r(t)dt = \lim_{n \to \infty} \left( \sum_{i=1}^{n} r(t_i) \cdot \Delta t \right). \]

Summarizing, if \( r(t) \) is the interest rate at time \( t \), measured per annum with continuous compounding, then a principal \( P \) will be worth

\[ Pe^{\int_{0}^{T} r(t)dt} \quad \text{(1.8)} \]

after \( T \) years.

The quantity \( r(t) \) is called the **force of interest**, or sometimes the **short rate**. It can be defined directly for any investment whose value equals \( Q(t) \) at time \( t \) to be

\[ r(t) = \frac{Q'(t)}{Q(t)} \quad \text{(1.9)} \]

To see that Equation (1.8) and Equation (1.9) are the same, we can start with Equation (1.9) and note that the right-hand side is the derivative of \( \ln(Q) \) with respect to \( t \). Thus

\[ r(t) = \frac{d}{dt} \ln(Q), \]

and by the fundamental theorem of calculus

\[ \int_{0}^{T} r(t)dt = \ln(Q(T)) - \ln(Q(0)). \]

The right side equals \( \ln\left(\frac{Q(T)}{Q(0)}\right) \), from which we arrive at Equation (1.8) by applying the exponential function to both sides and using the fact that \( Q(0) = P \).
CHAPTER 1. INTEREST RATES

Example 1.12. A bank account earns interest according to the force of interest \( r(t) = \frac{1}{1+t} \). How much does an initial deposit of $100 grow to after 2 years?

Solution: First, we compute the integral

\[
\int_0^2 r(t) dt = \int_0^2 \frac{1}{1+t} dt = \ln(1+t)|_0^2 = \ln(3) - \ln(1) = \ln(3) - 0.
\]

Plugging into the formula gives \( 100e^{\ln(3)} = 300 \) dollars.

1.10 Problems

Recall that unless otherwise stated, all annual interest rates are assumed to be compounded continuously.

1. An investment of $3000 grows to $4000. What is the rate of return on the investment?

2. A certain investment lasting 2 years guarantees a 9% rate of return on the investment. If the principal for this investment is $20,000, how much is the investment worth at the end of the 2 years? How much interest was earned?

3. A certain investment earns $300 on a principal amount \( P \). How much will it earn on a principal amount of \( 5P \)? \( P/2 \)?

4. A bank account pays 4% interest every year on the amount in the account. An initial deposit of $600 will be worth how much after 1, 2, 3, 4, and 5 years? How much interest is earned each year?

5. An investment paying an annual interest rate of 6% will double in how many years?

6. What annual rate of interest will cause an investment to double in 10 years?

7. A credit card loan charges interest at the rate of 14% per annum compounded monthly. How much does a balance of $300 grow to after 10 months? 1.5 years?
8. Calculate the effective annual rate of the interest rate in the previous example.

9. A bank account pays interest with continuous compounding at the rate of 6% per annum. What is the value of an initial deposit of $700 after 3 months? after 1.7 years?

10. Calculate the effective annual rate in the previous example.

11. Suppose a mortgage has an interest rate of 6.5% per annum with monthly compounding. Find the per annum interest rate with quarterly compounding that would lead to the same effective annual rate.

12. In the previous problem, find the rate with continuous compounding that would lead to the same effective annual rate.

13. A bank account pays 5% interest per annum with continuous compounding. After 1.1 years, an initial deposit \( P \) has become worth $2007. What was \( P \)?

14. As mentioned in Section 1.5 for a fixed interest rate \( i \), the more times per year compounding takes place, the higher the effective annual interest rate. For example, investing \( P \) dollars with compounding once per year gives a total investment of \( P(1 + i) \). But compounding twice per year gives

\[
P \left(1 + \frac{i}{2}\right)^2 = P \left(1 + i + \frac{i^2}{4}\right) = P \left(1 + i\right) + P \left(\frac{i^2}{4}\right),
\]

which is definitely bigger due to the right-hand term which is always positive.

(a) Show that compounding three times per year at the per annum rate \( i > 0 \) is better than compounding twice per year at the same rate.

(b) Show if \( n > m \), then compounding \( n \) times per year at rate \( i > 0 \) is always better than \( m \) times year at rate \( i \). (Hint: we need to show \( (1 + \frac{i}{n})^n > (1 + \frac{i}{m})^m \). This is equivalent to showing \( \log((1 + \frac{i}{n})^n) > \log((1 + \frac{i}{m})^m) \) since \( \log(x) \) is an increasing function. Now consider the function \( F(i) = \log((1 + \frac{i}{n})^n) - \log((1 + \frac{i}{m})^m) \).
CHAPTER 1. INTEREST RATES

Show that $F'(i) > 0$ for $i > 0$ and that $F(0) = 0$. Then use a result from calculus, Rolle’s Theorem, to conclude that $F(i) > 0$ for $i > 0$.

15. A $10,000 business loan that begins today will need to be repaid in 9 months. The balance at that time will be $10,260. What is the implied 9-month zero rate?

16. A bank offers a 3-month zero rate on investments of 3.1% and a forward rate of 3.9% for 3-month investments starting 3 months from now. If you agree to these rates, what will an investment of 500 dollars be worth in 6 months?

17. In the previous problem, what is the 6-month zero rate that leads to the same rate of return?

18. If $r_1 = 5.1\%, r_{1,3} = 6.2\%$, and $r_{3,4} = 5.8\%$, calculate the balance on a $1000 investment after 4 years employing these forward rates.

19. A bank account pays interest with a force of interest equal to $r(t) = t^2 + 1$. How much does an initial deposit of $200 grow to after 4 years?

20. In the previous example what is the effective annual rate during the first year? during the second year?

21. An investment is worth $Q(t) = t^3 + 1$ at time $t$ (measured in years). What is the force of interest?
Chapter 2

Bonds

In this section we introduce the risk-free rate, explain how to value a payment that is paid in the future, and introduce bonds. Bonds are one of the financial instruments used by organizations to raise funds for their operations. We also discuss the connection between forward rates and zero rates.

2.1 Factors affecting interest rates

As a consumer, you probably deal with interest rates in several different contexts. First, you may have student loans, credit card loans, or auto loans. On the other hand, you may have investments, such as a checking or savings account, a money market fund, or certificates of deposit. One quick observation is that the typical interest rate on a loan is higher than on an investment. This is how banks make money. The bank pays interest to gain access to funds, which they then loan out at a higher rate. While it may seem that this is unfair, the bank, besides covering its operating expenses, is taking on risk in extending loans, since the loan recipient may not honor his or her obligations.

Focusing on the interest rate for loans, several factors affect consumer rates. One major factor is whether there is an asset standing behind the loan (such as a house or car) and if there is, the projected value of the asset over time. If there is no asset, then the credit worthiness of the borrower plays a large role. A second major factor is the length of the loan: generally, the longer loan, the higher the interest rate. For example, with a credit card, the loan is for whatever length of time the borrower wishes and there
is no asset backing the loan. Indeed, the credit card company does not much care on what you spend the money. For this reason, credit cards have very high interest rates: a credit card being offered at the time of this writing by Discover carries an interest rate of 11% for holders with the best credit and 19% for those with weaker credit. At the other extreme, a mortgage (that is, a home loan) is backed by the value of the house. In August 2008 a 15-year mortgage carries an interest rate of 6.00% at Florence Savings Bank, while a 30-year mortgage goes for 6.50%.

Another crucial factor that affects interest rates is the underlying currency. The fact is that interest rates in U.S. dollars are not the same as those in Canadian dollars, the Euro, or the Japanese Yen. On July 3, 2008, a major bank in the U.S. seeking an overnight loan from another bank would pay an annual rate of 2.0%. The comparable rate in Canada is 3.0%; in Europe, it is 4.25%; and in Japan, it is 0.5%. These interest rates are set, albeit approximately, by the central bank that manages the respective currency. The central banks take into consideration inflation and growth rates in each currency when they decide what to set these rates at. The overnight rates then have an effect on other interest rates in that currency, although the relationship is complex.

In the next section, we make some assumptions about interest rates that will permit us to analyze the price of financial instruments.

### 2.2 Risk-free rate

As we mentioned in the previous section, different types of loans carry different interest rates. One of the factors mentioned that affects the rate is the possibility that the recipient of the loan will not honor his or her obligations. Perhaps the borrower will fail to make an interest payment, or worse, may completely default on the loan. Hence, a loan carries a risk for the lender.

Conversely, if you invest in a money-market fund, there is a tiny chance that the fund will not only not pay you interest, but may even lose some of your principal. Until 2008, many people thought such an event was rare, if not impossible; however, in September 2008 the Reserve Primary Fund became only the second money-market fund in U.S. history to lose principal. This is referred to as “breaking the buck” since the principal amount of one share in a money-market fund is one dollar. The fund lost principal since it held investments in Lehman Brothers, which filed for bankruptcy in
September 2008.

Throughout these notes, we will need a notion of interest rates that avoids the risk that an investor or lender will not honor his or her obligations. This is not to ignore the importance of analyzing risk; rather, it is a simplification that allows us to answer questions that would otherwise be too difficult to address. In the remainder of the notes, we will assume that at any given instant in time there is a prevailing zero rate $r_T$ that applies to any loan that begins now and lasts for $T$ years and for which there is no possibility of a missed interest payment or default.

To emphasize that this zero rate is special, we refer to $r_T$ as the $T$-year risk-free rate. We make the further assumption that institutional players (such as banks) can both borrow and lend money at this rate. If we do not specify the currency, then the assumption is that the rate applies to dollars.

**Assumption 1 (Risk-free Rate Hypothesis).** For each time period $T$, all institutional players can invest and borrow today at the risk-free rate $r_T$. This rate is a $T$-year zero rate; namely, it is guaranteed for loans and investments that start today and last for exactly $T$ years. There is no possibility that the borrower will default or that the lender will seek the return of the loan. All risk-free rates will be quoted per annum with continuous compounding.

Over time risk-free rates change. That is, today’s 1-month risk-free rate will potentially be different than tomorrow’s. A bit of terminology: the way today’s risk-free rates vary for different values of $T$ is called the **term structure** of interest rates. We usually visualize this by graphing $r_T$ as function of $T$. Sometimes, we may simplify matters further by assuming all risk-free rates are equal; that is, $r_T = r$, independent of $T$. This is referred to as the term structure being flat or constant.

**Example 2.1.** The six-month risk-free rate today is 1.7%. The 1-year risk-free rate is 2.1%. How much will an investment of $100 be worth in 6 months if it invested today at the risk-free rate? If it is borrowed for 6-months? What if it is invested for 12 months at the 12-month risk-free rate? What if it is invested for 9 months?

**Solution:** After six months, the investment will be worth $100e^{0.017(0.5)} = 100.85$. A loan will be worth the same since we are assuming that institutions can both borrow and lend at this rate.
For a 12 month investment, the answer is $100e^{0.021(1)} = 102.12$, using the 12-month risk-free rate.

Since we have not specified a 9-month risk-free rate, we are not in position to answer the last part of the question. We do not have enough information.

### 2.3 Value of money

One of the key concepts of this course is that interest rates determine the current value of a future payment.

Suppose that the risk-free rate now is 6% per annum with continuous compounding, no matter how long you want to borrow or invest your money. A friend offers to sell you a coupon that pays out $100 in exactly one year. How much would you pay today to own such a coupon? Here, we assume that you can borrow and lend at the risk-free rate.

Suppose the coupon costs $P$ dollars and you were to borrow $P$ dollars to pay for the coupon. Then in one year, when you went to pay off your loan, you would owe $Pe^{0.06}$ dollars; on the other hand, you would cash in the coupon and get 100 dollars. Your total gain is $100 - Pe^{0.06}$. Since you wouldn’t want to lose any money on the transaction, you wouldn’t pay $P$ dollars for the coupon unless

\[100 - Pe^{0.06} \geq 0,\]

or

\[Pe^{0.06} \leq 100.\]

On the other hand, if the coupon costs $P$ dollars, your friend could invest the $P$ dollars that you give for the coupon at the same interest rate. After one year, your friend would have $Pe^{0.06}$ from the investment, but then would owe you 100. Your friend’s gain would be $Pe^{0.06} - 100$. Since your friend would not want to lose any money, the inequality

\[Pe^{0.06} \geq 100\]

should hold true.

The two inequalities together show that

\[P = 100e^{-0.06} = 94.18,\]
if neither of you is to have any guaranteed profit from the transaction.

This leads to the equation for the value of a future payment:

**Formula 4 (Present Value).** A payment of $P$ dollars that is to be received $t$ years from now is worth $Pe^{-rt}$ dollars today, where $r = r_t$ is the $t$-year risk-free rate. The amount $Pe^{-rt}$ is called the present value of the future payment $P$.

We say that the future payment $P$ is discounted by the factor $e^{-rt}$ to reflect its value today. Conversely, the future value in $t$ years of a payment today of $Q$ dollars will be worth $Qe^{rt}$, which can be seen by setting $P = Qe^{rt}$ in the previous formula.

**Example 2.2.** You are given the opportunity to receive a signing bonus for a job of $5,000 plus a bonus in six months of $3,000. Alternatively, you could receive a bonus in six months of $8,100 and no signing bonus. If the risk-free rate is 4.5% for all values of $T$, which is the better alternative?

**Solution:** The first opportunity has a present value of $5000 + 3000e^{(-.045)(.5)} = 7,933.25$. The second opportunity has a present value of $8100e^{(-.045)(.5)} = 7,919.79$. Thus the first opportunity is worth more today and should be chosen.

Next, we address the question of how to calculate the present value of a collection of several future payments? The answer is that each payment can be valued separately and the value of the total collection, or portfolio, of payments will be the sum of the individual values. In general we will assume that the value of a collection of financial instruments is the sum of the values of the individual instruments.

**Assumption 2 (Portfolio Valuation).** Unless otherwise specified, the value of a collection of financial instruments (which is called a portfolio) is equal to the sum of the values of each individual financial instrument.

Since we are making this assumption, we can use Formula 4 to conclude:

**Formula 5 (Valuing Multiple Payments).** Let $P_1, P_2, \ldots, P_n$ be a series of payments occurring at times $t_1, t_2, \ldots, t_n$. Let $r_1, r_2, \ldots, r_n$ be the risk-free
zero rates corresponding to the times $t_1, t_2, \ldots, t_n$, respectively. If we let $B$ be the price today of the totality of the payment stream, then

$$B = \sum_{i=1}^{n} P_i e^{-r_i t_i}$$

(2.1)

**Example 2.3.** Calculate the present value of a payment stream that pays $200 every 3 months for one year, starting 3 months from now. Suppose the risk-free zero rate is 3% for all $T$.

**Solution:** There are 4 payments altogether. We need to find the present value of each payment and add them all up to find the total present value. The final equation is:

$$200(e^{-0.03(0.25)} + e^{-0.03(0.5)} + e^{-0.03(0.75)} + e^{-0.03(1)}) = 785.17.$$ 

**Example 2.4.** Now, calculate the present value of a payment stream that pays $200 every 3 months for one year, starting 3 months from now. Suppose the 3-month risk-free zero rate is 3%, the six-month risk-free rate is 3.2%, the 9-month risk-free rate is 3.25%, and 1-year risk-free rate is 3.3%.

**Solution:** The equation is:

$$200(e^{-0.03(0.25)} + e^{-0.032(0.5)} + e^{-0.0325(0.75)} + e^{-0.033(1)}) = 784.02.$$ 

### 2.4 Bonds

A **bond** is the name for a piece of paper that promises a stream of payments at certain time intervals.

Bonds are issued by governments (federal, state, and local), by corporations, and by other entities (such as utilities, universities, and hospitals). Bonds are a way to raise money. In other words, they are loans, but they are special loans because the entity that wants the loan creates the bond (with the help of an investment bank) and then sells the bonds to many different individuals. The individuals then receive the interest payments at regular intervals. After a certain time period, the entity pays the individuals back the principal amount of the loan. The entity that sells the bond is called the **issuer** of the bond, and an individual that buys the bond is called a **holder** of the bond.
For example, General Electric may issue a 10-year bond that pays $250 every six months for 10 years with a principal amount of $10,000. This means that the holder of the bond receives $250 every six months and at the end of the 10th year, the holder receives the last interest payment of $250 plus the principal amount of $10,000, for a total of $10,250. Since a bond is just a stream of payments it is possible to value a bond by the method of the previous section.

In general bonds are specified by three pieces of information. Bonds have a maturity, which is the length of time the payments will be made. In the previous example, the maturity is 10 years. The amount that the bond pays at maturity, less any final interest payment, is called the par value, face value, or principal of the bond. In the previous example, the face value is $10,000. Finally, the bond has a coupon. This is the interest rate that determines the payments that are made throughout the life of the bond. The reason the interest rate is referred to as the coupon is historical. In earlier days the bond literally was a piece of paper with coupons that could be ripped off and brought to the bank for the interest payment on the appropriate day. The coupon is expressed as an annual interest rate together with the annual frequency of the payments. In the previous example, the coupon is 5% per annum, paid semiannually. This means that over the course of the year you will get 5% of 10,000, or $500, but it will be paid in semiannual installments. Note that there is no compounding that takes place.

At maturity, the bond always pays the last coupon payment plus the principal amount.

Example 2.5. At the end of August 2007, the U.S. government sold 2-year bonds. They had a coupon of 4% per annum, paid semiannually. The face value for each of the bonds was $1,000. Describe the payment stream that a holder of one of the bonds will receive.

Solution: The holder will receive 4 payments. The first is for $20, which comes due 6 months after the bond is issued; the second is for $20, 1 year after the bond is issued; the third is for $20, 1.5 years after the bond is issued; and the last is for $1,020 (the last interest payment, plus the face value), which arrives 2 years after the bond is issued, at maturity.

Example 2.6. How much is the above bond worth when it is issued if the risk-free rate is 5% for all times $T$? is 6% for all times $T$? How much is it
worth 3 months after it is issued if the risk-free rate at that time is 6% for all times $T$?

Solution: In each case, there are 4 payments and we must find their present value. This leads to, in the first case:

$$20e^{-0.05(0.5)} + 20e^{-0.05(1)} + 20e^{-0.05(1.5)} + 1020e^{-0.05(2)} = 980.02.$$ 

In the second case the interest rate is higher:

$$20e^{-0.06(0.5)} + 20e^{-0.06(1)} + 20e^{-0.06(1.5)} + 1020e^{-0.06(2)} = 961.18.$$ 

In the third case, the first payment will be arriving in only 3 months, the second payment in 9 months, etc. The formula is therefore:

$$20e^{-0.06(0.25)} + 20e^{-0.06(0.75)} + 20e^{-0.06(1.25)} + 1020e^{-0.06(1.75)} = 975.71.$$ 

Notice that as the risk-free rate goes up, the present value of each payment goes down, and thus so does the value of bond. Another point worth making is that when a bond is issued its price need not be the same as its face value. Usually the coupon rate is chosen so that the initial price of the bond is close to face or par value, but it need not be the case that the price is exactly the par value. In Section 3.4, we study the coupon rate which makes the initial bond price equal to its par value.

We are often interested in the value of a bond between coupon payments. This can be a little tricky. When analyzing this situation, it is helpful to focus on the maturity of the bond and then work backwards knowing that each coupon payment occurs a fixed amount of time before the maturity date of the bond.

Example 2.7. The Commonwealth of Massachusetts issues a 3-year bond to fund bridge repair and construction. The bond pays coupons semiannually at a rate of 5% per annum and the face value is 10,000 dollars. Ten months after the bond is issued, we wish to value the bond. Which risk-free rates do we need to know?

Solution: The bond has a three year maturity, but 10 months have already elapsed since the bond was issued. Hence there are $3(12) - 10 = 26$ months left before the bond matures and we receive the last payment of 10,250 dollars
at that time. We know that six months before then there will be the second-to-last payment of 250 dollars. This occurs $26 - 6 = 20$ months from now. Continuing in this fashion, we see that payments arrive $26, 20, 14, 8$ and $2$ months from now.

Therefore, in order to value the bond, we need to know the $t$-year risk-free rate for $t = 26/12, 20/12, 14/12, 8/12$ and $2/12$ years. Notice that the bond has already paid out its first coupon payment, but this does not have any effect on the current price of the bond.

A few comments about bonds in the real world. In general, they have face values that are high enough to make them out of reach for a small investor. They also do not trade on an exchange and so they do not change hands as easily as a share of stock would. Instead they are traded through dealers at financial firms. There are electronic exchanges now that handle bond trades, but the market for this is still comparatively small. Also, in the real world, when a bond is traded between coupon payments, the seller would keep a fraction of the next coupon payment, proportional to the time that has elapsed from the last coupon payment. In other words, the price of the bond would not exactly be the future value of the coupon payments. Instead, it would be slightly lower, reflecting the fact that the seller of the bond has kept part of the payment. There are special rules which spell out various conventions for different types of bonds. We do not cover these conventions in these notes.

We also mention that bonds can have added features. The above bonds are all fixed-rate bonds, meaning that coupon payments are fixed throughout the life of the bond. Bonds can also have coupon payments that float, meaning that they vary depending on some well-known interest rate benchmark. A common feature of bonds is that they can be convertible, meaning the holder can at a certain time convert the bond into stock (this applies only to a bond issued by a corporation). A bond can also be callable. This means that if certain conditions are met, the issuer can close out the bond by paying off the bond holder. This feature protects issuers if interest rates were to drop significantly or if they no longer needed to hold so much debt.
2.5 LIBOR and Treasuries as measures of the risk-free rate

There are two standard measures for the risk-free rate in U.S. dollars. The first is U.S. government Treasuries. Treasuries, referring to the U.S. Treasury, are the name for the bonds issued by the U.S. federal government. Although several countries have defaulted on their sovereign bonds (this is a common name for the bonds issued by a national government), the U.S. has never defaulted, and most people would put the chance of a default as very low. Hence, U.S. Treasury bonds are a good measure of the risk-free rate for dollars. For instance, the 3-month Treasury bill that expired on December 6, 2007 was paying a rate of 4%. According to this measure, the risk-free rate for an investment that started on September 6, 2007 and ended on December 6, 2007 would be 4% per annum. That is, the 3-month risk-free rate was $r_{0.25} = 0.04$ on September 6, 2007.

A second measure of the risk-free rate is LIBOR. LIBOR stands for London Interbank Offered Rate. It is an average of the rates offered by banks in London to other banks for their deposits. There are 16 banks that presently determine the LIBOR rate for dollars. The U.S. banks that participate in LIBOR are Citigroup, Bank of America, and J.P. Morgan. Each day between 11 am and 11:10 am (London time), these banks report their rates; the top four and bottom four rates are dropped, and the remaining eight are averaged. This becomes the day’s LIBOR rates. LIBOR rates are quoted for various maturities (that is, they are zero rates for various values of $T$).

Recently, LIBOR has come under attack since there is a belief that some banks are underreporting their rates for fear of showing that they are having difficulty borrowing money (if they offer a high rate for banks to deposit their funds, this is equivalent to only being able to borrow at that high rate). The British Banker’s Association (BBA), which compiles LIBOR, has been studying the accuracy of LIBOR to see if there has been any misrepresentation.

There are LIBOR rates in several currencies. On September 6, 2007 the 3-month LIBOR for dollars was 5.72% per annum and the 1-week LIBOR was 5.76%. On the other hand, the 3-month LIBOR rate for Euros was 4.76%.

The fact that we are assuming that there is one risk-free rate for investments lasting exactly 3 months is certainly undermined by the fact that the 3-month Treasury rate and the 3-month LIBOR can be quite different, as
they were on September 6, 2007. Normally, these rates are much closer to each other. However, the credit markets were in turmoil in September 2007. Investors were buying short-term Treasuries since they were perceived as very safe (most nearly risk-free): this buying drove down the interest rate on the Treasuries. On the other hand, banks were hoarding cash, unwilling to make short-term loans to other banks for fear that some banks were in financial trouble. As a result, LIBOR went up to take into consideration this new risk. The upshot is that the difference, or spread, between LIBOR and U.S. Treasuries was unusually high in September 2007. For example, the spread between 3-month LIBOR and 3-month Treasuries at that time was 1.7%. On the other hand, by July 3, 2008, markets had settled and the spread had diminished to 0.84%. In Fall 2008, the credit market crisis began to bring down some of the large institutional banks. Even after Congress, on October 3, 2008, passed the 750 billion dollar bailout of the financial sector, a week later there was still a significant spread: 1-month LIBOR was at 4.6% and 1-month Treasuries was at 0.07%. Several months later, with some confidence restored in the banks, the spread fluctuated around 0.3%. Despite this disparity, we will use both benchmarks in our analyses since there are many financial instruments that depend on LIBOR and many others that depend on Treasury rates.

When speaking about interest rates, it is convenient to have a finer measure than a single percentage point. For this reason, the word basis point is introduced. A basis point is equal to 0.01 percent. In other words, 100 basis points is equal to 1 percent. For example, the spread between 3-month LIBOR and Treasuries was 84 basis points, or bips, on July 3, 2008.

### 2.6 Zero coupon bonds and more lingo

If the maturity of a bond is very short (around 6 months or less), there might not be any coupons. Instead, the issuer delivers a single payment equal to the face value at maturity and sells the bond at a discount to face value. Interest is earned since the face value will be more than the price that the bond is sold for.

For example, consider a zero coupon bond that is issued by the government. It has a face value of $100 and a maturity of 6 months. The bond sells at a discount of $98. To find the interest rate $r$ that the bond pays (per
annum with continuous compounding), we solve the equation:

$$98 = 100e^{-r(0.5)} \Rightarrow r = \frac{\ln(100/98)}{0.5} = 0.0404.$$  

Notice that the bond implies a 6-month zero rate of 4.04% since there is no exchange of payments during the six month period. This explains the use of the word zero rate: knowing the price of a zero coupon bond determines a zero rate, and vice versa.

**Example 2.8.** A zero coupon Treasury bond that matures in 9 weeks has a face value of $1000 and is trading at $995. What is the 9-week risk-free rate implied by this bond?

**Solution:** The current price of the future payment of 1000 is 995. Hence,

$$995 = 1000e^{-r(9/52)},$$

so $r = \frac{\ln(1000/995)}{9/52} = 0.02896$ is the risk-free rate that applies to 9 weeks.

**Example 2.9.** A zero coupon Treasury bond that matures in 9 weeks has face value of $1000. If the 9-week zero rate is 3.6%, what is the correct price of the bond?

**Solution:** The current price is

$$B = 1000e^{-0.036(9/52)} = 993.79.$$  

Concluding the section, we point out that U.S. Treasuries have different names depending on the maturity of the bond. The word bills or T-bills refers to Treasury bonds that have a maturity of up to 2 years. T-bills are zero coupon bonds. The word note refers to Treasury bonds with a maturity between 2 and 10 years. Finally, bond is used for those bonds with maturities of more than 10 years.

### 2.7 Relationship amongst forward rates and the no arbitrage hypothesis

Recall from Section [1.8](chapter1) that forwards rates are interest rates promised today for investments that start at some time $T_1$ and last until a later time $T_2$.  

2.7. **FORWARD RATE RELATIONSHIP**

For example, a 1-year forward rate of 3.5% that applies to the calendar year 2009 would mean a promise that $100 invested at the beginning of 2009 would grow to $100e^{0.035}$ at the end of 2009. We use the notation $r_{T_1,T_2}$ to denote a forward rate between $T_1$ and $T_2$ that is promised today. If today were June 1, 2008, then the previous example of a forward rate would be $r_{0.5,1.5}$, reflecting the fact that the rate will begin in half a year and last until 1.5 years from now.

There is an important relationship between zero rates and forwards rates if we assume that we are able to both borrow and lend at these rates and that the rates are risk-free.

As an example of the relationship between zero rates and forward rates, suppose you have access to a 1-year zero rate of 4% and a 2-year zero rate of 5%. Then there is an implied forward rate between year 1 and year 2, denoted $r_{1,2}$, which can be calculated as follows.

An investment of $P$ dollars using the 2-year zero rate will grow to $Pe^{0.04 \cdot 2}$ dollars after 2 years.

On the other hand, we can invest $P$ dollars using the zero rate for the first year, and then during the second we can roll over that investment and invest at the forward rate of $r_{1,2}$ between year 1 and year 2, since this forward rate is something that was promised to us today. In this way, the original investment of $P$ dollars becomes

$$Pe^{0.04 \cdot 1}e^{r_{1,2} \cdot 1} = Pe^{0.04 + r_{1,2}}.$$

Since we do not expect either investment to be do better than the other, this leads to the equality:

$$Pe^{0.05 \cdot 2} = Pe^{0.04 + r_{1,2}}.$$

After dividing by $P$ and taking the natural log of both sides, we get the equation

$$0.1 = 0.04 + r_{1,2},$$

or $r_{1,2} = 6\%$. In other words, knowing the 1-year and 2-year zero rates completely determines the forward rate $r_{1,2}$.

Now, let us explain more carefully why the two investments mentioned previously must be equal. Suppose the value of the first investment exceeded the value of the second; that is,

$$Pe^{0.05 \cdot 2} > Pe^{0.04 + r_{1,2}}.$$
Then, instead of investing at the 1-year zero rate, we would borrow \( P \) dollars at that rate. Simultaneously, we would invest that money at the 2-year zero rate. After one year, we need to repay our 1-year loan. Instead, we utilize the forward rate to borrow the amount we owe (we are rolling over our debt); that is, we borrow \( Pe^{0.04} \) at the forward rate we were promised a year ago. We wait another year. At the end of the second year, our investment of \( P \) dollars is worth \( Pe^{0.05(2)} \), while our loan debt is now \((Pe^{0.04})e^{r_{1.2}}\). Since the investment was assumed to be worth more than the amount owed, we have earned a guaranteed profit of

\[
P e^{0.05(2)} - Pe^{0.04+r_{1.2}} > 0.
\]

What is to stop us from increasing the value of \( P \) and making an arbitrary large profit? Notice that this mechanism is completely risk-free since we are assuming are interest rates are risk-free rates, and that we did not even need to put up any money of our own. This seems unbelievable and indeed it essentially is. When imbalances in the financial markets like this exist, they are quickly acted upon and this causes them to disappear. From our theoretical point of view, we will assume that they do not exist at all. That is,

**Assumption 3 (No Arbitrage Hypothesis).** In the financial markets, there does not exist any investment that requires no initial funds and that offers a risk-free profit. Said differently, an investment that offers a guaranteed profit with no principal utilized quickly disappears from the financial markets.

A situation where there is a guaranteed profit with no money (or capital) utilized is called an arbitrage opportunity, or simply an arbitrage. Our assumption above can be restated as saying that arbitrage opportunities do not exist in the markets.

Finishing with our example, we use the assumption of no arbitrage to conclude that the situation we arrived at is impossible. In other words, the inequality \( Pe^{0.05(2)} > Pe^{0.04+r_{1.2}} \) can not be true since otherwise we would be able to earn a risk-free profit with no capital used. A similar argument can be constructed to show that the reverse inequality \( Pe^{0.05(2)} < Pe^{0.04+r_{1.2}} \) can not hold either. Thus, we have given a complete argument to explain the equality

\[
P e^{(.05)(2)} = Pe^{0.04+r_{1.2}},
\]
from which we concluded that \( r_{1,2} = 6\% \).

Now, we will show that a similar relation holds between certain triples of forward rates. Consider the three forward rates

\[ r_{T_1,T_2}, r_{T_2,T_3}, \text{ and } r_{T_1,T_3}, \]

where \( 0 < T_1 < T_2 < T_3 \). Then the relationship between these three rates is given by the following formula.

**Formula 6 (Forward rate relation).** Assuming they are risk-free, the forward rates

\[ r_{T_1,T_2}, r_{T_2,T_3}, \text{ and } r_{T_1,T_3}, \]

are related by the equation

\[ r_{T_1,T_3} = \frac{r_{T_1,T_2}(T_2 - T_1) + r_{T_2,T_3}(T_3 - T_2)}{T_3 - T_1} \]

Here is how the relation is proved. Consider an investment of \( P \) dollars that begins in \( T_1 \) years and lasts until \( T_3 \) years at the rate \( a = r_{T_1,T_3} \). Simultaneously consider a loan of \( P \) dollars that begins in \( T_1 \) years and lasts until \( T_2 \) years at the rate \( b = r_{T_1,T_2} \). The loan is rolled over into another loan at time \( T_2 \), which last until time \( T_3 \) at the forward rate \( c = r_{T_2,T_3} \). Then the investment grows to \( Pe^{a(T_3 - T_1)} \) and the loan grows to \( Pe^{b(T_2 - T_1) + c(T_3 - T_2)} \). Hence, the profit (or loss) for the combined portfolio is \( Pe^{a(T_3 - T_1)} - Pe^{b(T_2 - T_1) + c(T_3 - T_2)} \). Since arbitrage opportunities do not exist, we can not have \( Pe^{a(T_3 - T_1)} > Pe^{b(T_2 - T_1) + c(T_3 - T_2)} \).

On the hand, we could have reversed the investment and the loan, by investing between time \( T_1 \) and \( T_2 \) and then rolling over the investment between \( T_2 \) and \( T_3 \), while simultaneously, borrowing between \( T_1 \) and \( T_3 \). Analyzing this situation would imply that \( Pe^{a(T_3 - T_1)} < Pe^{b(T_2 - T_1) + c(T_3 - T_2)} \) is not possible since arbitrage opportunities do not exist.

We conclude that

\[ Pe^{a(T_3 - T_1)} = Pe^{b(T_2 - T_1) + c(T_3 - T_2)}. \]

Then we divide by \( P \) and take natural logs to get

\[ a(T_3 - T_1) = b(T_2 - T_1) + c(T_3 - T_2). \]

If we solve for \( c = r_{T_1,T_3} \), we get the desired relation between forward rates.
**Example 2.10.** Consider the risk-free forward rates \( r_{1,3} = 2.1\% \) and \( r_{3,4,5} = 3.6\% \). What is the implied rate \( r_{1,4,5} \)?

**Solution:** By the formula, 
\[
r_{1,4,5} = \frac{.021 \times 2 + .036 \times 1.5}{3.5} = 2.74\%
\]

**Example 2.11.** Suppose a bond consists of two payments: one for \( P_1 = $30 \) occurs in 3 months and the other for \( P_2 = $130 \) occurs in 9 months. The 3-month risk-free zero rate is 4% and the 9-month risk-free rate is 5%. If the bond is trading in the market at $152, describe an arbitrage opportunity.

**Solution:** First, we find the price of the bond according to Equation (2.1):
\[
B = 30e^{-0.04(0.25)} + 130e^{-0.05(0.75)} = 154.92,
\]
which is higher than the market price of 152. So we decided that the market price is too low and therefore we should buy the bond. To do so, we can borrow $152 at the 9-month zero rate and buy the bond. Then, we wait 3 months and we receive a payment of $30. At this point, we should invest that money. But at what rate? The answer is that we can invest at the forward rate \( r_{0.25,0.75} \) implied by the 3-month and 9-month zero rates (otherwise, there is an arbitrage opportunity amongst the three rates). Hence, the $30 is invested for 6 months at the rate of
\[
\frac{0.05(0.75) - 0.04(0.25)}{0.5} = 0.055.
\]
After 9 months, we have received a total of \( 30e^{0.055(0.5)} + 130 = 160.84 \) from the bond payments (after investing the first one). On the other hand, we owe our loan principal with interest: \( 152e^{0.05(0.75)} = 157.81 \). We net \( 160.84 - 157.81 = 3.03 \) dollars.

Notice that this net profit is the future value of the initial difference in the two prices. Namely,
\[
3.03 = (154.92 - 152)e^{0.05(0.75)},
\]
where we are using the 9-month risk-free rate in the exponent. This observation leads to an easier way to compute the profit from an arbitrage opportunity: compare the present value of the two portfolios, then, if necessary, compute the future value of this difference.
2.8 Bond duration

In the remainder of the chapter, we will discuss a number of properties of bonds. Our goal is to understand how the price of a bond changes when zero rates change. The first tool to understanding how a given bond will react to interest rate changes is the duration of the bond. Roughly speaking, the duration measures the average time that the payments on the bond are made, discounted by the appropriate zero rates.

**Definition 2.12.** The duration of a bond with payments \( P_1, P_2, \ldots, P_n \) at times \( t_1, t_2, \ldots, t_n \) is defined to be

\[
D = \frac{\sum_{i=1}^{n} t_i P_i e^{-r_i t_i}}{B},
\]

where \( r_i \) is the \( t_i \)-year risk-free zero rate and \( B \) is the price of the bond. Since duration is an average of times, the units of duration are the same as those for time (usually years).

To compare this with a general weighted average, note that if you want the usual average or mean of a set of \( n \) values \( x_1, x_2, \ldots, x_n \), then the mean, denoted \( \bar{x} \), is

\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (2.2)
\]

To get to the weighted average, suppose that we want to weight each value \( x_i \) by the weight \( w_i \). We can think of the weights as saying that certain of the \( x \) values count more than others. The weighted average is defined to be

\[
\bar{x}_{wt} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \quad (2.3)
\]

To get from the weighted average to the usual mean amounts to setting all the weights equal to 1, so that all the \( x \) values count equally. Then the denominator

\[
\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} 1 = n,
\]
giving the familiar denominator in the mean. Frequently, the weights \( w_i \) in the weighted average are assumed to sum up to 1. In that case, the denominator in the weighted average definition simply goes away. This assumption will be standard when we review probability later on. There, the weights will represent the probability of each value occurring. Notice that this assumption is not a problem since we can always choose different weights \( w'_i = \frac{w_i}{d} \), where \( d \) is the denominator \( \sum_{i=1}^{n} w_i \). These modified weights \( w'_i \) sum up to 1 and give the same weighted average as the original weights \( w_i \).

Returning to the definition of duration, we are interested in averaging the times \( t_i \) when the various payments on a bond will occur. The weights we choose are the present value of each payment, namely \( w_i = P_i e^{-r_i t_i} \). Then the weighted average \( D \) is exactly as given in the formula, where we observe that

\[
\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} P_i e^{-r_i t_i} = B.
\]

**Example 2.13.** Calculate the duration of the bond in Example 2.7 when there are 13 months until maturity and the risk-free rates (as percentages) are 1.8, 1.6, and 2.1 for \( t = 1/12, 7/12, 13/12 \) years, respectively.

**Solution:** The price of the bond is

\[
B = 250e^{-0.018(1/12)} + 250e^{-0.016(7/12)} + 10250e^{-0.021(13/12)} = \$10,516.75.
\]

The duration \( D \) of the bond is

\[
\frac{(1/12)250e^{-0.018(1/12)} + (7/12)250e^{-0.016(7/12)} + (13/12)10250e^{-0.021(13/12)}}{B} = 1.0478
\]

years. Note that with bonds with only a few payments left until maturity, their duration will be close to maturity time since the last payment dominates.

### 2.9 Parallel shift in the zero curve

The **zero curve** is the name for the graph of all available (risk-free) zero rates. Over time, the zero curve will change, reflecting the effects of inflation,
fiscal policy on the part of the government, and other factors. To a first approximation, we can study how interest rate affect bond prices by assuming that all zero rates move by the same (small) amount. This is called a parallel shift in the zero curve.

Let $r_1, r_2, \ldots, r_n$ be the zero rates for times $t_1, \ldots, t_n$, respectively. Suppose that, all of a sudden, there is a parallel shift in the zero curve and all rates move by an amount $\Delta r$. In other words, the $t_i$-zero rate is now $r_i + \Delta r$. How does this affect the price of a bond?

Before the shift in the zero curve, a bond would be priced at

$$B = \sum_{i=1}^{n} P_i e^{-r_i t_i},$$

where $P_i$ are the payments at times $t_i$. After the shift, the bond would be priced at $B_{\text{new}}$ where

$$B_{\text{new}} = \sum_{i=1}^{n} P_i e^{-(r_i + \Delta r)t_i} = \sum_{i=1}^{n} P_i e^{-r_i t_i} e^{-\Delta r t_i}. \tag{2.4}$$

To a linear approximation, for small $x$, $e^x \approx 1 + x$.

Since $\Delta r$ is small, we can assume $\Delta r t_i$ is small as well. Using this approximation on $e^{-\Delta r t_i}$, we get that

$$B_{\text{new}} \approx \sum_{i=1}^{n} P_i e^{-r_i t_i} (1 - \Delta r t_i) = \sum_{i=1}^{n} P_i e^{-r_i t_i} - (t_i P_i e^{-r_i t_i}) \Delta r = B - BD \Delta r.$$

Note that we have used the fact that $BD = \sum t_i P_i e^{-r_i t_i}$, the numerator of the duration expression. Hence, this shows that the change in the price of the bond $\Delta B$ is, to a linear approximation, equal to

$$\Delta B = B_{\text{new}} - B \approx -BD \Delta r.$$

Summarizing:
Formula 7 (Bond Price Change Estimate). A parallel shift of $\Delta r$ in the 
zero curve leads to an approximate change $\Delta B$ in bond price of

$$\Delta B \approx -BD\Delta r,$$

where $B$ is the original price of the bond and $D$ is the duration of the bond.

Example 2.14. A certain bond has a duration of 2 years. If the zero curve 
shifts up by 10 bips, by approximately what percent will the bond’s price 
decrease?

Solution: The zero curve shifts by 10 bips, so $\Delta r = .001$. Hence, the percent 
change in price, which is $\frac{\Delta B}{B}$, is equal to $-D\Delta r = -2(.001) = -0.002$. That 
is, the price falls by about 0.2%.

Notice that the presence of the negative sign in Formula 7 means that 
when all zero rates go up, the price of a bond goes down. Conversely, if rates 
go down, the price of a bond goes up. In other words, bond price and interest 
rates move inversely to each other.

Example 2.15. Calculate both the exact change in price and the approxi-
mate change in price of the bond in Example 2.13 if all zero rates move down 
by 12 bips.

Solution: The new relevant zero rates are .018 − .0012 = 0.0168, 0.0148, and 
0.0198, so the new price of the bond is

$$250e^{-0.0168(1/12)} + 250e^{-0.0148(7/12)} + 10250e^{-0.0198(13/12)} = \$10,529.98,$$

and thus the exact change in price is 10529.98 − 10516.75 = 13.23 dollars.

On the other hand, the approximate change is

$$-BD\Delta r = -(10,516.75)(1.0478)(-.0012) = 13.223$$

dollars, which is essentially the same and easier to compute once you know 
the duration.
2.10 Convexity of a bond

The duration is a measure that captures the linear change in bond price as a function of the amount of a parallel shift in the zero curve. The convexity of a bond captures the second order change.

Consider Equation 2.4. If we approximate $e^x$ to second order by

$$e^x \approx 1 + x + \frac{x^2}{2},$$

then we find that

$$B_{\text{new}} \approx \sum P_i e^{-r_i t_i} \left( 1 - \Delta r t_i + \frac{1}{2} t_i^2 (\Delta r)^2 \right),$$

(2.5)

$$= B - BD \Delta r + \frac{1}{2} BC (\Delta r)^2,$$

(2.6)

where we are defining $C$ to be

$$C = \frac{\sum_{i=1}^n (t_i)^2 P_i e^{-r_i t_i}}{B}.$$

This is called the convexity of the bond. Note that the convexity, like the duration, can be interpreted as a weighted average, where now we are averaging the square of the times. In probability (or physics), averaging the squares of a quantity is called a second moment.

Example 2.16. Continuing with the Example 2.15, calculate the convexity of this bond and find a second order approximation to the change in price when the zero rates moved down by 12 bips.

Solution: The convexity is

$$\frac{(1/12)^2 250 e^{-0.018(1/12)}}{B} + \frac{(7/12)^2 250 e^{-0.016(7/12)}}{B} + \frac{(13/12)^2 10250 e^{-0.021(13/12)}}{B} = 1.126.$$

The second order approximation of the change is

$$\Delta B \approx -BD \Delta r + \frac{1}{2} BC (\Delta r)^2 = 13.223 + \frac{1}{2} (10516.75)(1.126)(-0.0012)^2 = 13.232,$$

an even better approximation to the true change in price.
2.11 Yield

The yield of a bond is a measure of the return that the bond gives as an investment. It takes into consideration the fact that payments are occurring at different times.

**Definition 2.17 (Yield).** The yield \( y \) of a bond with a price of \( B \) is the value \( y \) which solves the equation:

\[
B = \sum_{i=1}^{n} P_i e^{-y t_i},
\]

where as before the bond has payments \( P_1, P_2, \ldots, P_n \) at times \( t_1, t_2, \ldots, t_n \).

Notice that the formula for the yield is similar to Equation 2.1. The yield can thus be considered as the risk-free rate that would give the bond price if all the risk-free rates were equal to each other.

The yield and the bond price each determine the other. If we know the bond price \( B \), we can compute the yield \( y \). Conversely, if we know the yield, then we can calculate the price of the bond. This relationship exists without knowing anything about zero rates. For example, we might learn the price of the bond by looking up its market price and then use this value to determine the yield. In a different scenario, we might want to determine the bond price by using the appropriate zero rates. From there, we can determine the yield \( y \) of the bond. To solve for \( y \) in Definition 2.17 given the bond price \( B \), we can either use trial and error, or we can turn to a numerical solver like the one found on most calculators.

**Example 2.18.** A bond pays $10 in 4 months, $10 in 16 months, and $1,010 in 28 months. The bond trades for $1,021 (in other words, the price of the bond is $1021). What is the yield of the bond?

**Solution:** We need to solve the equation:

\[
1021 = 10e^{-y(0.3333)} + 10e^{-y(1.3333)} + 1010e^{-y(2.3333)}
\]

for the yield \( y \). The answer is 0.381% (or 38.1 bips), by using a calculator or MATLAB.

Since the exponential function \( e^x \) is an increasing function (and thus \( e^{-x} = \frac{1}{e^x} \) is a decreasing function), it follows that if all the payments are
the same for two bonds, then the bond with the higher yield will have the lower price. This is because each factor \( e^{-yt_i} \) will decrease when \( y \) increases. Conversely, if the yield decreases, each factor will increase and so will the bond price. This is a fundamental fact, often quoted in the financial press: the bond price and bond yield move inversely to one another. Next, we explore the relationship between changes in bond price and changes in bond yield. We could proceed as we did to derive Formula [7]. Instead, we take a slightly different route.

According to Definition [2.17] we can view the bond price as function of its yield \( y \):

\[
B(y) = \sum_{i=1}^{n} P_i e^{-yt_i}. \tag{2.7}
\]

Taking the derivative of \( B(y) \) with respect to \( y \) gives

\[
\frac{dB}{dy} = \sum_{i=1}^{n} -t_i P_i e^{-yt_i},
\]

using the chain rule. Ignoring the negative sign, this looks very similar to the numerator of the duration expression. In fact, the two quantities will be close enough to warrant approximating the duration by:

**Formula 8.** The duration \( D \) of a bond can be approximated by the expression:

\[
D \approx \frac{\sum_{i=1}^{n} t_i P_i e^{-yt_i}}{B}. \tag{2.8}
\]

With this approximation in mind, we see that

\[
\frac{1}{B} \frac{dB}{dy} = -\frac{\sum_{i=1}^{n} t_i P_i e^{-yt_i}}{B} \approx -D,
\]

or

\[
\frac{dB}{dy} \approx -BD.
\]

Now using linear approximation of the one-variable function \( B(y) \), we get

\[
\Delta B \approx \frac{dB}{dy} \Delta y \approx -BD\Delta y,
\]

which is in the same spirit as Formula [7]. Just like that formula, this is a good approximation for small values of \( \Delta y \).
Formula 9 (Yield Duration Approximation). At a fixed yield $y$, the change in yield $\Delta y$ leads to a change in price $\Delta B$ according to

$$\Delta B \approx -BD\Delta y,$$

where $B$ is the bond price and $D$ is the bond duration.

It is even more useful, to rewrite the approximation formula as:

$$\frac{\Delta B}{B} \approx -D\Delta y. \quad (2.9)$$

This formula expresses the fact that the percent change in bond price (as a function of $\Delta y$) depends only on the duration of the bond. In other words, to a linear approximation, all bonds of the same duration behave the same (in percent terms) when their yield changes.

In the case of a single payment, the duration is just the time until that payment. So the duration of a zero coupon bond is the time until maturity.

Example 2.19. Estimate the duration of the bond in Example 2.18 and use the yield-duration approximation to estimate the change in the bond price when the yield moves to 1%.

Solution: By using Formula 8 with $y = 0.381\%$, the duration is approximately

$$D = \frac{(-0.3333)10e^{-0.00381(1.3333)} + (1.3333)10e^{-0.00381} + (2.3333)1010e^{-0.00381(2.3333)}}{1021},$$

which equals 2.304 years. Notice that this is just short of the time until maturity, which is 2.333 years.

Since $\Delta y = .01 - .038 = .0062$, we get

$$\Delta B = -BD\Delta y = -(1021)(2.304).0062 = -14.58$$
dollars.

We conclude the section by showing why the yield is well-defined (that is, why Equation 2.17 can be solved and there is a unique solution). First, $B(y)$ is a continuous function of $y$. Next, as $y$ get bigger in the positive direction, the value of $B(y)$ tends to zero and as $y$ tends to negative infinity, $B(y)$ becomes bigger and bigger. Note that we are using the fact that all the
$P_i$’s are assumed to be positive. Hence by the Intermediate Value Theorem from calculus, given any positive value of $B$, Equation 2.17 can be solved.

Finally, there is a unique solution since $B(y)$ is a decreasing function of $y$ (as explained earlier using the properties of the exponential function). Thus, two different values $y_1 < y_2$ of $y$ must lead to two different values $B(y_1) > B(y_2)$ of $B$.

### 2.12 Portfolio of Bonds

As discussed in Section 2.3, a portfolio is a collection of positions in various assets (for example, bonds, stocks, cash, etc.). In this section we consider portfolios of bonds and how they behave under interest rate changes.

One nice fact about the duration is that, being a linear quantity, it behaves well with respect to portfolios. Suppose a portfolio consists of two bonds. The first one has payments $P_1, P_2, \ldots, P_n$ at times $t_1, t_2, \ldots, t_n$, and the second one has payments $P'_1, P'_2, \ldots, P'_m$ at times $t'_1, t'_2, \ldots, t'_m$. Then the first bond has a price of $B = \sum P_i e^{-r t_i}$ and duration $D = \sum \frac{P_i t_i e^{-r t_i}}{B}$, while the second bond has a price of $B' = \sum P'_i e^{-r' t'_i}$ and duration $D' = \sum \frac{P'_i t'_i e^{-r' t'_i}}{B'}$, where the $r_i$ and $r'_i$’s are the relevant risk-free rates. Now, the portfolio of the two bonds is worth $B + B'$. If we look at the totality of the payment stream for the two bonds and compute the duration of the stream we find that duration of the portfolio is:

$$
\frac{\sum P_i t_i e^{-r t_i} + \sum P'_i t'_i e^{-r' t'_i}}{B + B'}.
$$

The numerator is just $BD + B'D'$, hence we have shown that the duration of the portfolio is $\frac{BD + B'D'}{B + B'}$, another example of weighted average. This holds more generally for any number of bonds.

**Formula 10.** A portfolio of $k$ bonds with prices $B_1, B_2, \ldots, B_k$ and durations $D_1, D_2, \ldots, D_k$, respectively, has duration equal to

$$
\frac{B_1 D_1 + B_2 D_2 + \cdots + B_k D_k}{B_1 + B_2 + \cdots + B_k},
$$

the weighted average of the individual durations, weighted by the price of the bonds.
The power of this simple formula is that we can quickly compute the duration of a portfolio and then employ Formula 7 or Formula 9 to estimate how a shift in the zero curve or a change in the yield affects the price of the portfolio.

**Example 2.20.** Bond A has a duration of 3 years and a price of $1032. Bond B has a duration of 2 years and a price of $998. Compute the duration of a portfolio of bonds consisting of 1 of Bond A and 2 of Bond B. If the zero curve shifts down by 15 bips, how will the value of the portfolio change?

*Solution:* The price of the portfolio is 1032 + 2(998) = 3028. The duration of the portfolio is:

\[
D = \frac{3(1032) + 2(2 \times 998)}{3028} = 2.34
\]

years. If \(\Delta r = -0.0015\), then \(\Delta B = -(3028)(2.34)(-0.0015) = 10.63\) dollars.

### 2.13 Problems

1. A 2-year T-note has a par value of $1000. It pays a coupon of 4.5% per annum, paid semiannually. If the bond is issued today, calculate the price of the bond if we assume all risk-free zero rates are 4%.

2. In the previous problem, suppose that today is now 1 month after the T-note has been issued. Now, all risk-free zero rates are 5% per annum. What is the new price of the bond? What if all the risk-free zero rates are 3%, what would the new price of the bond be?

3. A 2-year T-note was issued 9 months ago with a face value of $1000. It pays a 5% per annum coupon, paid semiannually. Suppose that the 3-month zero rate is 6%; the 9-month zero rate is 6.1%; the 15-month zero rate is 6.2%; and the 21-month zero rate is 6.3%, where all of these rates are per annum with continuous compounding. What is the price for the bond today?

4. What is the zero rate implied by a zero coupon bond that has a face value of $1000 that comes due in 4 months and that trades at $992?
5. What is the zero rate implied by a zero coupon bond that has a face value of $1000 that comes due in 7 months and that trades at $983?

6. Calculate the forward rate between 4 months from now and 7 months from now, based on the zero rates of the previous two problems.

7. In Formula [\[ T_1 = 0 \] (so that there are two zero rates). Solve for the remaining forward rate (the one which is not a zero rate).

8. Describe an arbitrage opportunity if \( r_{2.5} = 3.7\% \), \( r_3 = 3.8\% \), and \( r_{2.5,3} = 3.4\% \). That is, describe a series of investments and loans that would yield an unlimited, risk-free profit.

9. Suppose that you can borrow at a rate of 5\% per annum and invest at a rate of 4\% per annum. Consider a coupon that pays $200 in 2 years. Using a no arbitrage argument, show that there is a range of possible prices for the coupon. In other words, show that if the price of the coupon is below \( 200e^{-0.05(2)} \), then it is possible to borrow money and buy the coupon, leading to a guaranteed risk-free profit. Conversely, show that if the price is above \( 200e^{-0.04(2)} \), then it is possible to sell the coupon and invest the proceeds, also leading to a guaranteed risk-free profit.

10. Suppose that 3-month LIBOR is 5.6\% and that 6-month LIBOR is 5.4\%. You are offered a forward rate today \( r_{3,6} \) starting in 3 months and ending in 6-months of 5.8\%. Suppose you are allowed to borrow or invest up to $1000 at the forward rate. Describe a way to make a guaranteed profit. How much is your profit?

11. A 3-year bond has a face value of $1,000. It pays an 8\% per annum coupon, with annual coupon payments. The 1-year zero rate is 5\%, the 2-year zero rate is 6\%; and the 3-year zero rate is 7\%, where all of these rates are per annum with continuous compounding. What is the price of the bond today (3 years before maturity)?

12. In the previous problem, describe an arbitrage opportunity if the bond trades for $1,005. Explain each step of the arbitrage. Show that the guaranteed gain is equal to the future value of the difference between the theoretical price and the observed market price.
13. Using the same bond from Problem 11, what is the value of the bond after it matures?

14. Using the same bond from Problem 11, compute the yield of the bond (assuming its theoretical price). Compute its duration using the zero rates given. Compute its convexity. Also estimate its duration by using the yield.

15. Continuing with the previous problem: if the zero curve shifts down by 50 bips, calculate the exact change in the bond price. Also, calculate the approximate change in price using one of the duration approximation formulas (Formula 7 or Formula 9). Finally, calculate the change in price using the second order approximation formula (Equation 2.6).

16. Bond A has a duration of 2 years. Bond B has a duration of 10 years. A pension fund owns $100,000 in each bond. If zero rates move down by 50 bips, which bond fares better?

17. Consider a portfolio which contains $25,000 of Bond A (above) and $10,000 of Bond B. Compute the duration of the portfolio. How does the value of the portfolio change (in percent terms), when the zero curve moves up by 25 bips?

18. If you expect inflation to rise dramatically in the coming months, is it better to own bonds with a long duration or bonds with a shorter duration?

19. Notice how in Example 2.15, the first order approximation was an underestimate for the new bond price.

Consider now any coupon-bearing bond. Suppose there is a small parallel shift in either direction in the zero curve. Assume that the second order approximation in Equation 2.6 is the true new price. That is, assume the shift is so small that a third order, fourth order, etc, is no more accurate than the second order approximation. Prove that the first order approximation is always an underestimate.

Hint: Why is Section 2.10 called “Convexity?”
Chapter 3

Swaps and more on bonds

This chapter introduces the swap, the name for agreements that involve two parties that agree to exchange, or swap, payments on a regular basis. Each payment is based on an interest rate that may be fixed, or it may float, meaning it may vary according to an interest rate benchmark such as LIBOR. The typical swap involves an exchange of one set of payments that is based on a fixed interest rate for another set that is based on a floating interest rate. These fixed-floating swaps are used by companies and governments to manage interest rate exposure. In addition, swaps can involve payments that are based on two different currencies. These are known as currency swaps. Currency swaps are used, for example, by companies to manage the cash-flows of their operations in different countries.

We also explore bootstrapping, a term that refers in this context to using the market price of several bonds of varying maturity to determine zero rates, one at a time.

3.1 Bootstrapping

As we saw in Section 2.6 knowing the price of a zero coupon bond allows us to determine an implied zero rate for the length of time remaining until the maturity of the bond. Now, suppose we know the price of two bonds: one is a zero coupon bond and the other has two coupon payments, where the first payment occurs at the maturity of the zero coupon bond. Then we are in the position to deduce two zero rates: the zero rates that correspond to the times of the two payments of the second bond.
**Example 3.1.** Bond A is a zero coupon bond that matures in 4 months. Its face value is $1000 and its market price is $992. Bond B is a bond that matures in 10 months, with semiannual payments and a coupon of 8% per annum. Its face value if $1000 and its market price is $1062. What are the 4-month and 10-month zero rates implied by these bonds?

*Solution:* First, we can find the 4-month rate from Bond A, which is a zero coupon bond:

\[ 992 = 1000e^{-r_{4}(4/12)} \Rightarrow r_{4} = 0.02410 = 2.41\%. \]

The equation for Bond B’s price is:

\[ 1062 = 40e^{-r_{4}(4/12)} + 1040e^{-r_{10}(10/12)}, \]

except that we now know that \( r_{4} = 0.02410 \). Plugging that in and solving gives \( r_{10} = 0.02058 = 2.058\% \), the 10-month zero rate.

The process can be repeated for any number of bonds. Consider the case of 3 bonds:

**Example 3.2.** Bond A is a zero coupon bond maturing in 6 months. Bond B and Bond C are bonds that were issued in the past and that pay semiannual coupons. Bond B matures in 1 year and Bond C matures in 1.5 years. The face value of all bonds is $1000. The table shows the current market prices and coupon rates for the bonds (the rates are semiannual). Find the 6-month, 1-year, and 1.5-year zero rates implied by these bonds. Assume they have just paid out their most recent coupons (i.e., Bond B has two coupons remaining).

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Market price</th>
<th>Coupon rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-month</td>
<td>$994</td>
<td></td>
</tr>
<tr>
<td>1-year</td>
<td>$988</td>
<td>5.3%</td>
</tr>
<tr>
<td>1.5-year</td>
<td>$1001</td>
<td>5.5%</td>
</tr>
</tbody>
</table>

*Solution:* The 6-month zero rate (from Bond A) is 1.204%. This goes into the 1-year bond’s price formula

\[ 988 = 26.5e^{-0.01204(5)} + 1026.5e^{-r_{1}(1)}, \]
and determines the 1-year zero rate, which is 6.525%. Finally, these two zero rates go into the 1.5-year bond’s price formula

\[ 1001 = 27.5e^{-0.01204(0.5)} + 27.5 \cdot 5e^{-0.06525(1)} + 1027.5e^{r_{1.5}(1.5)}, \]

which determines the missing 1.5-year zero rate of 5.376%.

3.2 Floating-rate bonds

Up until now, we have considered bonds that pay a fixed amount of money every fixed time period, say annually or semiannually. In addition to these types of bonds, there are bonds whose coupon payments depend on an interest rate which changes over time. Typically the bonds are pegged to LIBOR. Such bonds are called floating-rate bonds since their coupon payments fluctuate, or float, at a prevailing interest rate. Floating-rate bonds are also known as variable-rate bonds.

Floating-rate bonds have the advantage for the borrower that their coupon interest rates are usually lower than fixed-rate bonds when they are issued; however, they have the disadvantage to the borrower that the interest rate can reset to higher rates over time. From the investor’s point of view, a floating rate bond is a hedge against interest rates increasing.

In the mortgage market, the difference between fixed- and floating-rate bonds has its analog in the difference between a fixed-rate mortgage and an adjustable-rate mortgage (ARM). The latter have interest rates that reset over time. The fact that many ARM’s are resetting at much higher rates than their initial rate is one of the many factors underlying the 2007-2009 credit crisis.

As an example of a floating-rate bond, consider a bond that has coupons every 6 months, in which each coupon payment is determined by the 6-month LIBOR from six months before the coupon is paid. Suppose that the bond has a 2-year maturity and its face value is $1,000. Suppose the bond was issued on September 1, 2008. On that day, 6-month LIBOR is recorded. Then the first coupon payment of the floating-rate bond, which takes place on March 1, 2009 (6 months after the bond is issued), will be for

\[ 1000 \times (.031/2) = 15.50 \]
dollars, where we are quoting LIBOR per annum with semiannual compounding. In other words, between September 1, 2008 and March 1, 2009, the holder and the issuer of the bond are aware of the amount of the first coupon payment; however, they do not yet know any of the future coupon payments. Next, on March 1, 2009, the 6-month LIBOR is recorded; let us say it was 3.7%. This new rate then determines the amount of the second coupon payment, which will occur on September 1, 2009. It will be for

$$1000 \times \left(\frac{0.037}{2}\right) = 18.50$$

dollars. The pattern continues for the last two coupon payments: 6-month LIBOR on September 1, 2009 will determine the coupon payment on March 1, 2010; and 6-month LIBOR on March 1, 2010 will determine the last coupon payment on September 1, 2010, when the bond matures. Note that, just as for fixed-rate bonds, the face value is paid at maturity. See Table 3.1.

<table>
<thead>
<tr>
<th>Date</th>
<th>6-month LIBOR (per annum, semiannual compound)</th>
<th>Coupon paid out($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/1/08</td>
<td>3.1%</td>
<td>None</td>
</tr>
<tr>
<td>3/1/09</td>
<td>3.7%</td>
<td>15.50</td>
</tr>
<tr>
<td>9/1/09</td>
<td>3.8%</td>
<td>18.50</td>
</tr>
<tr>
<td>3/1/10</td>
<td>4.1%</td>
<td>19.00</td>
</tr>
<tr>
<td>9/1/10</td>
<td>3.9%</td>
<td>20.50 + 1000</td>
</tr>
</tbody>
</table>

Table 3.1: A 2-year floating-rate bond, with semiannual coupons

In the example from Table 3.1, we used an interest rate with semiannual compounding to express the coupon payments. In order to value a floating-rate bond, it will be convenient to quote the rate using continuous compounding. Recall that if interest is earned at a rate of $r$ with continuous compounding, then after $T$ years, the interest earned will be

$$P e^{rT} - P = P(e^{rT} - 1),$$

since we need to subtract out the principal $P$ to find the interest (see Section 1.4). For example, if we quote the 6-month LIBOR above with continuous
3.2. FLOATING-RATE BONDS

compounding, then the next coupon will be $1000(e^{r \cdot 0.5} - 1)$, where $r$ is the per annum rate with continuous compounding.

The determination of the price of floating-rate bond is rather interesting. To determine it, we need to work backwards.

Here is an obvious fact: just after a bond matures and its last coupon payment and principal have been paid out, it is worth nothing. But what is the value of the bond immediately before the last payment? It will be worth exactly the last coupon payment plus the principal since the time until maturity is negligible (meaning that we can ignore the present value discount). Taking the previous floating-rate bond as an example, just before September 1, 2010, the bond is worth exactly $1,020.50.

What about one month earlier, on August 1, 2010? At that time, the holder of the bond knows that in one month, the bond will pay $1,020.50. Indeed, the holder has known this ever since March 1, 2010 when the 6-month LIBOR was recorded. Therefore, the bond price can be determined by valuing a payment of $1,020.50 that will occur in one month. To do that, the 1-month risk-free rate, $r_1$, is needed. If we know $r_1$, then we know the bond is worth

$$1020.50e^{-r_1(1/12)}.$$ 

Similarly, we can value the bond at any time during the last six months before maturity if we know the appropriate risk-free rate.

In particular, there is a very simple answer for the value of the bond on March 1, 2010. At that moment, the 6-month LIBOR which determines the last coupon payment is known. At the same time, we can use this same LIBOR as the 6-month risk-free rate. Denote this rate by $r_6$, quoted with continuous compounding. Then the final payment of interest and principal can be expressed as $1000e^{r_6(6/12)}$, since there is the principal of $1000$ plus the coupon of $1000(e^{r_6(6/12)} - 1)$. To value this payment on March 1, 2010, we discount it using $r_6$ since $r_6$ is the 6-month risk-free rate. This leads to a value of

$$[1000e^{r_6(6/12)}] \times e^{-r_6(6/12)} = 1000,$$

since the exponential factors cancel each other. In other words, we arrive at a key fact: just after the second-to-last coupon payment the bond is worth its face value of $1000$!

This process can be repeated. Since just after the second-to-last coupon payment the bond is worth its face value of $1000$, we know that just before
that payment it is worth
\[1000 + 1000(e^{r_6} - 1) = 1000e^{r_6},\]
where \(r_6\) is the 6-month LIBOR that was recorded on September 1, 2009.

Why? The bond is worth its face value of 1000 just after the second-to-last payment as we just argued; just before the payment, we know that we are going to get a payment of \(1000(e^{r_6} - 1)\) and then we will be left with a bond worth 1000. The total sum is \(1000e^{r_6}\). We can repeat the above argument and deduce that the price of the bond is
\[
[1000e^{r_6}] \times e^{-r_6} = 1000
\]
just after the third-to-last payment (on September 1, 2009). Repeating these arguments, we see that

**Formula 11 (Valuing a floating-rate bond).** A floating-rate bond with face value \(P\) is worth exactly \(P\) when it is first issued and also immediately after every coupon payment (except the last coupon payment). After the last coupon payment, the bond is worth nothing.

Immediately before the coupon payment times, the bond is worth
\[P + I,\]
where \(I\) is the amount of the upcoming coupon payment.

At any other time \(t\), the bond is worth
\[(P + I)e^{-rT},\]
where \(T\) is the time until the next coupon payment, \(r\) is the \(T\)-year LIBOR rate at time \(t\), and \(I\) is the amount of the next coupon payment.

**Example 3.3.** A floating-rate bond pays coupons every year based on 1-year LIBOR. The bond has a 5-year maturity and a face value of $1000. What is the bond worth just after the third coupon payment? What is the bond worth just after it is issued?

**Solution:** In both cases, the bond is worth its face value of $1000 by Formula 11.
Example 3.4. Continuing with the previous example, suppose that at year 2 of the bond, 1-year LIBOR is 4.2% per annum with continuous compounding. What is the bond worth just before the third coupon payment? What if the 1-year LIBOR at year 2 is 4.3% per annum with annual compounding?

Solution: With continuous compounding of 4.2%, the next coupon payment equals

\[ I = 1000(e^{0.042(1)} - 1) = 42.89. \]

Hence the bond is worth $1042.89 since the bond is worth its principal amount plus the next coupon payment \( I \).

With an annual interest rate of 4.3%, the next coupon payment is

\[ I = 1000 \times 0.043 = 43, \]

and so the bond would be worth 1000 + 43 = $1043.

Example 3.5. Continuing with the previous example, what is the bond worth at \( t = 2.5 \) if 6-month LIBOR at \( t = 2.5 \) is 3.9% and if 1-year LIBOR at \( t = 2.5 \) is 4.6% (both per annum with continuous compounding)? Assume that at \( t = 2 \), the 1-year LIBOR was 4.3% per annum with annual compounding.

Solution: The 1-year LIBOR at time \( t = 2.5 \) is not relevant. Instead, the two relevant rates are the 1-year LIBOR from year 2 (which determines the third coupon) and the 6-month LIBOR now at \( t = 2.5 \) (which allows us to find the present value of the bond’s worth in 6 months). The next coupon amount is \( I = 43 \) using the 1-year rate from six months prior. Hence by Formula 11, the bond is currently worth

\[ (1000 + 43)e^{-0.039(6/12)} = 1022.86. \]

3.3 Fixed-floating swaps

A swap refers, in a broad sense, to any exchange of payments. The main variety (referred to as a plain-vanilla swap or a fixed-floating swap) is an
exchange of payments between two entities A and B. The structure is as follows: entity A sends entity B a payment stream based on a fixed interest rate determined at the beginning of the swap, while B sends A a payment stream based on a floating interest rate determined at the time of the last exchange (or swap) of payments. The fixed interest rate and floating interest rate are a percentage of the same principal amount, or face value. At the end of the swap, the parties exchange the last payment. Since the principal amount is the same for both entities, it does not make sense to exchange the principal, as one would do for bonds, since the principal amounts cancel each other out.

Consider the following swap from the news. Jefferson County (Alabama) and J. P. Morgan Chase enter into a swap agreement on January 1, 2007. The two entities agree to swap payments every 3 months. Chase will pay Jefferson County a quarterly payment based on 4% per annum, while Jefferson County will pay Chase the 3-month LIBOR rate from three months ago. The face value of the swap is 1 million dollars and the swap will last for 2 years.

Suppose that on January 1, 2007, the 3-month LIBOR rate is 2.8% (per annum, quarterly compounding). Then on April 1, 2007, Chase pays out

\[1,000,000 \times \frac{0.04}{4} = 10,000\]
dollars, while Jefferson County pays out

\[1,000,000 \times \frac{0.028}{4} = 7000\]
dollars. On that day, Jefferson County nets $3000, which it probably needs to settle the coupon payments on its fixed-rate municipal bonds.

Unfortunately for Jefferson County, 3-month LIBOR rose dramatically since it entered into this swap. Suppose that on April 1, 2008, 3-month LIBOR was 5.6%. Then on July 1, 2008, Chase will still pay out $10,000, but Jefferson County now needs to pay out

\[1,000,000 \times \frac{0.056}{4} = 14,000\]
dollars. Jefferson County (or more specifically, its sewer authority in the real-world version) has lost $4,000 on the arrangement that quarter. While the exact numbers used in this example are fictional, the swap arrangement itself is

---

\(^1\)We have modified the details.
real and the change in interest rates is leading Jefferson County’s sewer authority to the brink of bankruptcy as of this writing since the amount of the outstanding swaps (and related obligations) are actually closer to 3.2 billion dollars! While swaps can be used to effectively manage interest rate exposure by municipalities, this is an example of a situation where the municipality (or its counterparty, the bank) misjudged the risks in the hunt for higher returns.

Next, we wish to price a swap. The price of a swap is designed to be zero when the swap begins. In other words, when two entities enter into a swap, they choose the payment streams so that they do not need to exchange any payments at the outset. However, as interest rates fluctuate the swap will have positive value for one entity and negative value for the other.

It is not hard to price a swap, since for one party it looks just like a portfolio consisting of a long position in a fixed-rate bond and a short position in a floating-rate bond (see Section 4.4 for a discussion of a short position). Consider the previous example. From Jefferson County’s point of view, they can pretend that they own a fixed-rate bond issued by Chase and that they have sold Chase a floating-rate bond. Hence the value of the swap to Jefferson County at anytime is

\[ V_{\text{swap}} = B_{\text{fixed}} - B_{\text{float}}, \]

where \( B_{\text{fixed}} \) is the value of a fixed-rate bond with 3 month coupons with a coupon rate of 4% and a face value 1 million dollars, and \( B_{\text{float}} \) is the value of a floating bond based on 3-month LIBOR and a face value of 1 million dollars. Notice that the net payment stream of this portfolio of bonds (one long, one short) is the same as the payment stream of the swap. Indeed, even as the bond matures, the exchange of the principal amounts of 1 million is an offsetting transaction in the portfolio; that is, one payment of a million goes from Chase to Jefferson County and an offsetting one goes from Jefferson County to Chase.

**Example 3.6.** Suppose Cisco and Citigroup enter into a 2-year fixed-floating swap on June 1, 2006. The face value of the swap is 1 million dollars and the payments are exchanged semiannually. Citigroup will pay Cisco based on a fixed rate of 5% (per annum, semiannual compounding) and Cisco will pay Citigroup based on 6-month LIBOR. Consider Table 3.2 below consisting of 6-month LIBOR rates and use it to determine the swap payments.

**Solution:** We have listed the answers in the table itself.
A fixed-floating swap

<table>
<thead>
<tr>
<th>Date</th>
<th>6-mo LIBOR (semiannual)</th>
<th>Citi pays ($)</th>
<th>Cisco pays ($)</th>
<th>Cisco’s net ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6/1/06</td>
<td>2.8%</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>12/1/06</td>
<td>3.2%</td>
<td>25K</td>
<td>14K</td>
<td>11K</td>
</tr>
<tr>
<td>6/1/07</td>
<td>5.1%</td>
<td>25K</td>
<td>16K</td>
<td>9K</td>
</tr>
<tr>
<td>12/1/07</td>
<td>4.8%</td>
<td>25K</td>
<td>25.5K</td>
<td>-0.5 K</td>
</tr>
<tr>
<td>6/1/08</td>
<td>4.9%</td>
<td>25K</td>
<td>24 K</td>
<td>1K</td>
</tr>
</tbody>
</table>

Table 3.2: A swap based on 6-month LIBOR

**Example 3.7.** Suppose in the previous swap, the 4-month LIBOR rate is 2.7% two months after the swap begins. Suppose that 10-month LIBOR is 2.9%, 16-month is 3.1%, and 22-month LIBOR is 3.3% (all of these rates are per annum with continuous compounding). What is the value of the swap to Cisco? to Citigroup?

**Solution:** The value of the swap to Cisco is $B_{fixed} - B_{float}$, where $B_{fixed}$ is a fixed-rate bond with 22 months until maturity and $B_{float}$ is a floating-rate bond with 22 months until maturity. To find the price of the fixed-rate bond, we need to use all four zero rates. The fixed-rate bond price (in thousands of dollars) is

$$B_{fix} = 25e^{-0.027(4/12)} + 25e^{-0.029(10/12)} + 25e^{-0.031(16/12)} + 1025e^{-0.033(22/12)} = 1037.99,$$

while the floating-rate bond price (in thousands of dollars) is

$$B_{float} = 1014e^{-0.027(4/12)} = 1004.91,$$

since the upcoming coupon payment is set at $14,000. Hence, the value of the swap to Cisco is $33080, while to Citi it is worth $−33080. In other words, the swap is a liability for Citi, and Citi would record a loss of $33,080 when it reports its earning.

### 3.4 Deducing swap rates

An interesting problem for a swap is to figure out the correct fixed rate to be swapped against the floating rate. This is not hard. Since the swap is
3.4. DEDUCING SWAP RATES

not worth anything when it is negotiated, we have that $V_{swap} = 0$. In other words,

$$B_{fix} = B_{float}$$

when the swap begins. Next, we know that a floating bond is worth its face value $P$ when it is issued. Therefore,

$$B_{float} = P,$$

when the swap begins. Finally, let $c$ denote the coupon rate that determines the payments of the fixed-rate bond. In our present context, this is the swap rate that we would like to know. For the sake of an example, assume that the payments are exchanged every year for three years. Then the fixed-rate bond is worth

$$B_{fix} = (Pc)e^{-r_1(1)} + (Pc)e^{-r_2(2)} + (Pc + P)e^{-r_3(3)},$$

where $r_i$ are the relevant zero rates.

Hence, to determine the swap rate $c$, we need to solve the identity

$$P = (Pc)e^{-r_1(1)} + (Pc)e^{-r_2(2)} + (Pc + P)e^{-r_3(3)},$$

which leads to

$$1 = c(e^{-r_1(1)} + e^{-r_2(2)} + e^{-r_3(3)}) + e^{-r_3(3)},$$

which is easily solved for $c$.

We note that solving this equation amounts to finding the coupon rate on the fixed-bond which makes its initial price equal to its face value $P$. This coupon rate is referred to as the par rate, since it makes the bond price equal to its par value at issuance.

Example 3.8. Determine the swap rate on a one-year swap that exchanges semiannual payments. The current 6-month and 1-year LIBOR are 4% and 4.5%, respectively.

Solution: Let $c$ be the swap rate, i.e., the amount of the coupons of the fixed-rate bond, quoted per annum. Each of the two coupons will equal $P \times \frac{c}{2}$. Hence we need to solve the equation:

$$P = \frac{Pc}{2}(e^{-0.04(5)} + e^{-0.045(1)}) + Pe^{-0.045(1)}.$$
The $P$ cancels out and we get

$$c = 4.55\%.$$  

### 3.5 Currency swaps

There are other types of swaps in addition to fixed-floating swaps. One example is a **currency swap** where a company exchanges an amount in one currency (based on a fixed or floating rate) for an amount in a second currency (based on either a fixed or floating rate). Currency swaps are used by companies to manage their revenue streams in different countries. They are also used by central banks to manage their currency reserves and perhaps the exchange rates of their currencies.

For example, a dollar-yen currency swap might involve exchanging payments based on 1 million dollars and 105 million yen. A US company might send yen that it earns in Japan to a Japanese bank in exchange for dollars (in which it reports its earnings in). The exchange might be set up as follows: the US company sends a 4% coupon on the yen and receives a 5% coupon on the dollars. In other words, the US company pays 4,000K in JPY (Japanese Yen) at every annual exchange and receives 50K in USD (US Dollars). The interest rates in the exchange have to do with interest rates on US dollars as compared to interest rates on Japanese Yen. This particular example indicates that interest rates on Yen are lower than those for US dollars. The principals are usually chosen to match the currency exchange rate at the beginning of the swap. Unlike the fixed-floating swap earlier, the principals are exchanged at the end of the swap, because moving currency exchanging rates may have made one principal worth more than the other.

Let $B_J$ be the present value of the Japanese bond in JPY. Note that this is computed using the Japanese risk-free rate set in Japan. Let $B_{US}$ be the present value of the US bond in USD. Of course this is computed using the US risk-free rate set in the US. Let $S_0$ be the price of 1 JPY in dollars. Then the value of the swap to the US company is

$$V_{\text{swap}} = B_{US} - S_0 B_J.$$
3.5. **CURRENCY SWAPS**

Note that it is important to discount the yen coupons using the Japanese rate (compute $B_J$), and then convert their sum into dollars (multiply by $S_0$). If we converted them into dollars and then used the US rate to discount them, we would be making the (probably false) assumption that the JPY-USD exchange rate is the same each time coupons are exchanged.

**Example 3.9.** On September 1, 2010, the price of one euro (EUR) is 1.20 dollars. The risk free rates are 4% in the US and 3% in Europe (both rates per annum with semi-annual compounding). A swap is created between Boeing based in the US (paying coupons in EUR) and Airbus based in Europe (paying coupons in USD), where the principal exchanged in equivalent to 10 million EUR. The swap expires in 5 years. Determine the payments exchanged for the swap.

**Solution:** Since the swap is set up to match the exchange rate and risk-free rates on September 1, 2010, the terms of the swap are as follows: on September 1, 2010 Boeing pays Airbus 10 million EUR and receives 12 million USD; every 6 months thereafter, until September 1, 2015, Boeing pays $0.04/2 \times 12 = 0.24$ million USD and receives $0.03/2 \times 10 = 0.15$ million EUR; on September 1, 2015, in addition to a last coupon, the original principals are exchanged back.

**Example 3.10.** Consider the swap constructed in the previous example. Suppose September 1, 2013, the price of one EUR has increased to 1.30 dollars. Both US and European risk free rates are now 3% per annum with semi-annual compounding. Compute the value of the swap to Boeing in USD, and to Airbus in EUR, assuming a coupon has just been exchanged.

**Solution:** In the millions of USD, the value of the US bond is

\[
B_{US} = 0.24 \left(1 + \frac{0.03}{2}\right)^{-1} + 0.24 \left(1 + \frac{0.03}{2}\right)^{-2} + 0.24 \left(1 + \frac{0.03}{2}\right)^{-3} + 12.24 \left(1 + \frac{0.03}{2}\right)^{-4}
\]

\[= 12.23.\]

In millions of EUR, the value of the European bond is

\[
B_E = 0.15 \left(1 + \frac{0.03}{2}\right)^{-1} + 0.15 \left(1 + \frac{0.03}{2}\right)^{-2} + 0.15 \left(1 + \frac{0.03}{2}\right)^{-3} + 10.15 \left(1 + \frac{0.03}{2}\right)^{-4}
\]

\[= 10.\]
Let $S = 1.30$ be the price of one EUR in dollars. The value of the swap to Boeing in millions of USD is

$$V_{\text{swap}} = B_{US} - SB_E = -0.67.$$ 

Let $S' = 1.3^{-1} = 0.77$ be the price of one USD in euros. The value of the swap to Airbus in millions of EUR is

$$V_{\text{swap}} = B_E - S'B_{US} = 0.52.$$ 

### 3.6 Credit default swaps

Another kind of swap that has been in the news lately is a credit-default swap (CDS). CDS are instruments that allow one company to insure against the risk of default by a second company. For example, say Goldman Sachs (GS) wants to insure itself against Lehman Brothers (LEH) defaulting on its bonds. GS wants to insure 10 million dollars of LEH bonds. It enters into a CDS with the insurer AIG. It pays AIG an initial amount and then an annual premium. The premium is determined by a market for such swaps. If LEH defaults (which it did!), then AIG would have to pay GS the 10 million dollars. Actually the terms of the swap are more complicated, requiring AIG to post collateral if the premium on the CDS increases. In other words, as the marketplace views LEH as more likely to default, it requires the insuring company on the CDS to deliver funds to cover some of the payout in the event that the default does occur. This posting of collateral was one of the factors that brought about the downfall of AIG in September 2008.

### 3.7 Problems

1. Consider two bonds, Bond A and Bond B. Bond A is a zero coupon bond maturing in 5 months. Its face value is $1000 and it trades for $991. Bond B matures in 11 months. It will pay coupons of $30 in 5 and 11 months plus face value of $1000 at maturity. Bond B trades for $1015. Compute the zero rates implied by these two bonds.
2. The treasury has just issued a 6-month zero coupon bond, a 1-year note, and a 2-year note. Six months ago, the treasury issued a 2-year note denoted T2. The face value of all bonds is $1,000. The three notes pay a coupon twice a year (the T2 note has just paid its first payment, so you can ignore that). The table shows the current market price and coupon rate for the four bonds. Find the 6-month, 1 year, 1.5 year, and 2 year zero rates implied by these bonds.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Market price</th>
<th>Coupon rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-month</td>
<td>$996</td>
<td></td>
</tr>
<tr>
<td>1-year</td>
<td>$987</td>
<td>5.25%</td>
</tr>
<tr>
<td>2-year</td>
<td>$976</td>
<td>5.5%</td>
</tr>
<tr>
<td>T2</td>
<td>$982</td>
<td>5%</td>
</tr>
</tbody>
</table>

3. A floating-rate bond has face value $10,000. What is the bond worth when it is issued?

4. A floating-rate bond has face value $1,000. It pays coupons every year based on 1-year LIBOR and has a maturity of 5 years. When the bond is issued, 1-year LIBOR is 4% (with an annual convention). What is the amount of the first coupon of the bond? the second coupon of the bond? What is the value of the bond just before the first coupon? just after the first coupon? just after the second coupon?

5. Continuing with the previous example. Suppose that 3 months after the bond is issued, the 9-month LIBOR rate is 3.8% (with per annum continuous compounding). What is the price of the bond 3 months after the bond is issued?

6. You hold a floating-rate bond in your investment portfolio. It pays 6-month LIBOR every 6 months. The face value is $10,000 and you are expecting the next coupon in 2 months for 250 dollars. Your broker calls you up and says the company backing the bond has gone bankrupt and the bond is now worthless. How much have you lost on the bankruptcy if 2-month LIBOR at that time is 3.5% (with per annum continuous compounding)?
7. A two-year swap has a face value of 1 million dollars and begins on January 1, 2009. A fixed rate of 7% (per annum, semiannual compounding) is exchange for 6-month LIBOR every six months. Consider the table of LIBOR rates and determine when the payments are exchanged and for what amounts.

<table>
<thead>
<tr>
<th>Date</th>
<th>6-mo LIBOR (continuous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1/09</td>
<td>4.7%</td>
</tr>
<tr>
<td>7/1/09</td>
<td>4.3%</td>
</tr>
<tr>
<td>1/1/10</td>
<td>4.5%</td>
</tr>
<tr>
<td>7/1/10</td>
<td>5.0%</td>
</tr>
<tr>
<td>1/1/11</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

8. In the previous problem, what is the value of the swap on January 1, 2009?

9. Continuing with the previous problem, what is the value of the swap on March 1, 2009 to the party receiving the fixed payments if all LIBOR rates at that time are equal to 4.6% (with per annum continuous compounding)?

10. IBM and UBS are negotiating a 3-year swap with an annual exchange of payments based on 1-year LIBOR. What is the swap rate if LIBOR rates obey the formula

   \[ r_t = \sqrt{t} + 2 \]

   percent, with per annum continuous compounding where \( t \) is measured in years? In other words, the 9-month LIBOR rate is \( \sqrt{.75} + 2 = 2.87 \) percent.

11. European-based Airbus and US-based Boeing have two years remaining on a currency swap agreement, with payments made annually. One is about to occur today. The principals exchanged are 10 million USD and 8 million EUR, with rates 4% and 5%, respectively. Airbus is making USD coupon payments. Current risk-free rates are 4.2% in US and 4.6% in Europe (per annum, with continuous compounding). The current exchange rate is 1.3 dollars per euro. Compute the value of the swap to Airbus, in EUR.
Chapter 4

A survey of financial instruments and markets

So far we have encountered bonds and swaps and learned how to determine their price. In this chapter, we introduce the major financial instruments and describe how they are traded. We also discuss the concepts of margin and selling short.

4.1 Instruments for raising money

The main purpose of financial instruments is to raise funds on behalf of corporations, governments, non-profits and other entities to accomplish a variety of goals. These goals may include: (1) making investments in labor or equipment; (2) building or repairing infrastructure such as bridges, buildings, utility networks, or sewer systems; (3) maintaining a level of capital reserves that is required by law (for a financial entity); (4) re-loaning the funds to other institutions (for a financial entity); (5) paying off existing debts that are about to mature; (6) acquiring or merging with another corporation.

The holders or owners of these instruments, that is, the investors, use the instruments to earn a profit (hopefully) that exceeds the risk-free rate. They are willing to take on the risk of underperforming the risk-free rate (or losing money) in exchange for the possibility of earning more.

Financial instruments are also used to manage risks that institutions face. These risks include inflationary risks, interest-rate risks, and currency risks. For example, a food company that needs to purchase corn throughout the
year will employ a strategy using futures contracts (or perhaps forward contracts) to reduce its exposure to the change in the price of corn. As another example, a multinational computer company that receives income in many different foreign currencies, but reports its earnings in its home currency, will use currency swaps or forward contracts to manage its exposure to the changes in currency exchange rates. Instruments can also be used to manage tax liability.

Finally, they can be used for pure speculation. Speculation generally refers to a bet in the financial markets that does not involve an actual investment in an entity. For example, if you buy stock in a company, you are investing. If you bet that the price of oil will go up faster than the inflation rate, then you are speculating.

4.1.1 Bonds

Governments, corporations, and non-profit institutions use bonds are the primary way to raise money. The bond is essentially a loan that these entities take out. But instead of their being one person, or counterparty, making the loan, the bonds are issued and then sold to individual investors, pension funds, mutual funds, etc. In this way, the risk of the loan is widely distributed. Bond holders (those buying the bond) have a right to recover the assets of the company if the company fails to make coupon payments (or return the principal). Such an event is known as a default. Of course, the bond holder might not recover very much, as appears to be the case in the bankruptcy and default of Lehman Brothers where bondholders are set to receive about 9 cents for each dollar of bonds that they hold as of the time of this writing.

The logistics of issuing a bond are complicated and involve a bond-rating company (e.g., S&P, Fitch, or Moody) rating the credit-worthiness of the bond and an investment bank (e.g., Goldman Sachs) underwriting the bond. This means that the investment bank determines the coupon rate and prices the bond and then looks for investors to buy the bond. The investment bank is also involved in determining how much money should be raised by the issuer. It is also entirely possible that the investment bank plans to hold onto some of the bonds.

Bonds issued by local and state governments and non-profits such as universities, transit authorities, or hospitals are known as municipal bonds. These bonds are mainly used to raise capital to repair or build new infras-
4.1. INSTRUMENTS FOR RAISING MONEY

Tax receipts or fees collected by the municipality are used to pay off the bonds. Because bond holders of municipal debt wouldn’t be allowed to take over the assets of a hospital or city that declared bankruptcy, it has until recently been the practice for municipal bonds to be insured by a third party. However, during the credit crisis of 2007-8 these third-party issuers, known as monoline insurers, such as Ambac, got into financial difficulties by insuring collateralized debt obligations (CDO’s) and now this whole model is under attack. Indeed, many big municipalities including the State of California have been increasing the amount of their bonds that are issued without insurance. Municipal bonds are generally exempt from federal taxes and state taxes for residents of the state where the bond is issued. Treasury bonds (those issued by the U.S. government) are only exempt from federal taxes.

4.1.2 Other loans

Institutions use other types of loans to fund their operations. They may borrow directly from a bank; this is called a bank loan. These loans may be secured, that is, backed by specific assets that the company owns (for example, a specific factory or property), or they may be unsecured, that is, backed solely by the credit-worthiness of the institution. In the event that the company declares bankruptcy, secured loan holders are first in line to claim the assets promised to them; next in line are the unsecured loan holders. After them come the bond holders.

There are markets that companies use to raise short-term capital. These include repurchase agreements (known as repos) where a company gives a bank a financial instrument such as a Treasury bond in exchange for cash. The company agrees to buy back the instrument at a later time (usually one day or a few days) for an amount higher than the loaned cash.

Commercial paper refers to a market for short-term loans that last from one day to a few months. These are the primary way that financial companies fund their short-term obligations. Commercial paper may be backed by assets or not. The main buyers of commercial paper are money-market funds. During the crisis of 2007-8, the commercial paper market was frozen at several junctures. In the beginning, commercial paper backed by mortgages was a problem (this led to a small crisis in Canada in 2007). In October 2008 several money-market funds lost money on the bankruptcy of Lehman, which made its commercial paper worthless. This led money-market funds
to be very conservative with whom they lent to and to demand high interest rates. In one of many unprecedented moves, the Federal Reserve intervened in the commercial paper market and bought commercial paper directly from companies, including General Electric, the biggest issuer of commercial paper in the U.S.

4.1.3 Equity

Equity refers to ownership in a company. When a company is founded, it is a privately-held company, owned by one or more people who started the company. They make the decisions about the company. Later, as the company gets big, the owners seek to bring in new expertise, to decrease their exposure if the company should struggle, and to reward their workers with an ownership stake. At this point, the company may go public. It will issue stock to the public on a stock exchange in an initial public offering (IPO). The owners retain stock for themselves and they may give some to their employees. The remaining shares of stock are sold at a price determined by the investment bank underwriting the offering. The funds raised are used by the company to fund its growth.

After the initial offering, the shares trade on an exchange and the price fluctuates. Owners of the shares are owners of the company and have a right to vote in annual meetings on issues related to the running of the company. Usually this amounts to selecting the people who will oversee the company (the Board of Directors).

Stock shareholders may receive a dividend, or cash payout, on a quarterly basis. Generally, companies that are fast-growing do not pay out dividends, whereas those that are established and have excess cash do. For example, for many years Microsoft paid no dividend. Now a share of Microsoft stock pays a dividend of 13 cents a quarter. This amounts to a return of about 2.2% given the price of its stock on October 16, 2008. On the other hand, General Motors pays a dividend of 25 cents per quarter. This amounts to a return of 15% since its stock has fallen dramatically in recent days. Dividend payouts are set by companies when they report their quarterly earnings. Recently, many financial companies have slashed their dividends in order to conserve cash.

Stock can be issued as common shares or preferred shares. Preferred

\[1\] As of this writing in October 2008.
shares normally carry a much higher dividend amount. They function almost like a bond except that they are easier to trade than bonds since they come in smaller units and they trade on a stock exchange. Bonds tend to be harder to trade since they have higher prices and they must be traded through a broker. Preferred shares are treated differently than bonds by regulators with respect to capital ratios, debt ratios, and taxes. This is one of the reasons that many financial companies issued preferred shares in the credit crisis rather than bonds. For example, bonds are viewed as debt (a negative), while preferred shares are viewed as equity (a positive). The amount of debt to equity that a company carries affects its credit rating. Preferred shareholders are above common shareholders when it comes to dividing up the assets of a bankrupt corporation (but both are below holders of debt, such as bond holders). Preferred shares often come with no voting rights.

Another example of a hybrid instrument (which preferred shares are) is a convertible bond. It begins as bond and then can be converted into stock after some time has elapsed.

4.2 Derivatives

Companies use derivatives to manage their exposure to risks, namely interest-rate risks, currency risks, and inflationary risks. Derivatives refer to a financial instrument whose value is derived from another instrument or asset. If we think of bonds (and their implied interest rates), stocks, and tangible assets such as commodities (gold, oil, wheat) and currencies as the basic building blocks of the economy, then derivatives are the objects whose values depend on these.

The main derivatives are forwards (and futures), options, and swaps. Most traditional derivatives, such as futures and options are tightly regulated. When you hear people blaming derivatives for the credit crisis, this refers to unregulated derivatives such as credit default swaps.

We have already discussed swaps. We now briefly outline the other main derivatives.

\footnote{Depending on the type of swap, it may not be considered a derivative.}
4.2.1 Forwards

A forward contract (or forward) is an agreement between two specific parties. One party agrees to buy an asset (such as oil or wheat) at a later time; the other party agrees to sell that asset. The two parties agree on a date for the exchange of the asset and a price. This becomes a binding contract. No money is exchanged until the agreed upon time (the delivery date). Regardless of what the asset costs on the delivery date, the parties exchange the asset for the previously agreed upon price.

In summary a forward is just a contract that is signed today to buy something in the future. See Chapter 5 for more details on forwards.

4.2.2 Futures

A futures contract (or future) is a more advanced way to handle the idea behind a forward. Futures are contracts on assets that trade on an exchange, like a stock would. However, unlike a stock, the investor does not buy anything. Instead, the investor leaves money in a margin account with his broker. At the end of each day, the futures contract is marked-to-market; that is, money is taken from the margin account to cover the gain or loss on the contract for that day. It is as if you sold the contract each day, and then immediately bought it back for the next day. The actual futures contract itself stops trading at a well-defined time. For a certain period up until that time, the holder of a long position could be assigned by the exchange to buy the asset at the price at which the contract is trading if there is a holder of a short position who wants to sell. However, most futures contracts are closed out before this happen, meaning that if an investor is long a contract, he goes to the futures exchange and sells a contract (and the two positions cancel each other out).

There are some other distinctions between futures and forwards contracts: a forward can have any delivery date, whereas the regulated futures only have several dates in a year; when the asset is a financial instrument, the futures contract is often settled in cash instead of an actual transfer of the asset. Mathematically, however, the most important difference is the futures’ margin account described in the previous paragraph. Because the gain or loss of the futures position is realized frequently, there is little risk in defaulting. If the trader defaults on a margin payment, he is closed out and his broker is responsible to meet the last margin call. Another mathematical repercussion of
4.3. TRADING ON A MARKET VERSUS OVER-THE-COUNTER (OTC)

the margin account appears in the argument for computing the futures price. In Chapter 5, we shall see how to compute a forward price by constructing a portfolio with the underlying asset and cash. A no-arbitrage argument will then determine the forward price in terms of the portfolio. This argument will not work with the futures: if you own the asset and cash, you are never forced to meet margin calls, and so the cash flows do not match those of the futures contract.

4.2.3 Options

A call option is a contract that gives the owner of the contract the option to buy an asset at a certain time \( T \) for a certain price \( K \), called the strike price. The owner does not have to exercise this option, but may if she chooses. The investor who sells, or writes, the contract is obliged to sell the asset at \( K \) dollars if the owner wants to buy the asset. As with futures, options typically trade through an exchange, so buyer and seller are matched through the exchange if an exercise takes place.

A put option is a contract that gives the owner of the contract the option to sell an asset at a certain time \( T \) for a certain price \( K \). The owner does not have to exercise this option, but may if she chooses. The investor who sells, or writes, the contract is obliged to buy the asset at \( K \) dollars if the owner wants to sell it.

Options are mostly constructed where the underlying asset is a stock. You can look up option prices on Yahoo! finance and other websites. Options are also written on futures contracts.

Options come in two flavors, European and American. See Chapter 6 for more details on options.

4.3 Trading on a market versus over-the-counter (OTC)

Financial instruments may trade on an exchange or they may be exchanged between two parties. The latter is referred to as trading over-the-counter (or OTC). Bonds, swaps, and forwards generally trade over-the-counter, whereas stocks and futures trade on exchanges. Stocks may also be traded OTC, as when a big investor buys up a large block of shares from another big investor or in a special issuance of stock. Stock exchanges are companies themselves
(and may even have their own stock listed on an exchange!). The big stock markets in the U.S. are Nasdaq-OMX (OMX is the Nordic stock exchange) and NYSE (the New York Stock Exchange). Surprisingly, stocks that are listed on one exchange can trade on other exchanges. There are several smaller exchanges where stocks also trade. Futures in the U.S. trade on the CME (Chicago Mercantile Exchange), NYMEX (New York Mercantile Exchange), or the ICE (Intercontinental Exchange). CME acquired NYMEX in 2008, with complete integration expected in 2009. In 2007, CME acquired CBOT, its Chicago competitor since the 1800’s.

4.4 Selling short

Selling something you do not own is usually called stealing, but there is a formal process in the financial markets called **selling short**. When you sell short, you are allowed to borrow a financial instrument from a brokerage firm and sell it in the marketplace. At a later date, you have to buy back the instrument and return it to the brokerage firm (except in the case where the instrument has no value and no obligations attached to it; for example, if you sell short a bond, then after the bond matures, it has no value).

A major assumption we will make is that it costs nothing to sell short. For example, suppose you want to sell short one share of Google stock. If Google is trading at $540, then you would receive $540. In our idealized framework, you are free to invest these proceeds as you see fit. Of course, at some later time, you have to buy the stock back and return the borrowed shares to the brokerage firm; you are obliged to do so.

In the real world there are costs associated with selling short. Generally you are required to put up **margin**. This means that you leave a certain amount of cash or other assets as collateral in your brokerage account. At ETrade, for instance, the rule is that you should post 50% margin initially. In other words, in order to sell short one share of Google at $540, you need to have $270 in your account to cover the possibility that you will lose money on the transaction. Although in our framework you are free to invest the $540 as you please, in the real world most brokerages hold that money and you do not earn interest on it. A big bank or hedge fund can sell short and do what they want with the proceeds; however, they still must post margin and they also pay a fee to borrow the shares. All of these real world situations notwithstanding, we will assume that you can freely sell short any instrument
and invest the proceeds as you like. Recently the SEC (Security Exchange Commission) banned new short selling in certain stocks for a period of time. As of this writing (October 16, 2008), that ban has now been lifted.

There is another important feature about selling short: you become responsible for any payments that the original owner is expecting. In other words, if the original owner is expecting a 30 dollar payment in 3 months on a bond, and you have sold short the bond, you must make this payment to the original owner. In actuality, you make the payment to your brokerage firm and the firm handles the logistics of assigning the payment to the appropriate person.

Example 4.1. Bond A is trading today at $1,019 in the market. The bond matures in 1 year and has a coupon of 5% per annum with semianual payments and a face value of $1000. Donna is long the bond (she owns it) in an account that pays no interest on cash assets; her account contains no other positions. William also has a brokerage account that pays no interest on cash assets. William has just sold short 2 of Bond A.

For both Donna and William, describe the positions in their accounts today and in 1 year after the bond matures. Describe the payment streams between Donna and her brokerage firm and also between William and his brokerage firm over the course of the year. Mark the book of both Donna and William today and just after the bond matures.

Solution: For Donna: today, she is long one Bond A and has no cash. She will receive $25 in 6 months and $1025 in 1 year from her brokerage firm. After 1 year, she will have $1,050 and nothing else. Today Donna’s account is valued at $1019 and in 1 year it will be valued at $1,050.

For William: he is short two of Bond A today and has $2,038 in his account. He needs to pay out $50 in 6 months and $2050 in one year to his broker. After 1 year, he will be in debt $62 and have no other positions. Today his account is valued at zero (his short position cancels out the cash position in his account); in 1 year, his account will be valued at $-62 dollars. He will certainly be contacted by his broker to add (or post) more money to his account.
4.5 Problems

1. A stock is trading in the marketplace for $50. The stock pays a dividend of $2 every quarter to the owner of each share of stock (starting in 3 months time).

Sue has an account at the Banks of Leeds which contains $100 in cash. She sells short one share of the stock and keeps the proceeds in her account (where there is no interest earned). The stock trades at $52 three months later and $49 six months later. She buys the stock back 7 months later at $45.

Describe the value of her account right before she sells short; right after she sells short; 3 months later; 6 months later; and 7 months later. Keep track of the value of each position in her account. Generally, a short position in an asset has negative value since, like a loan, it requires the holder of the short position to spend money to free the holder of an obligation.

2. A bond maturing in 5 years with a 6% per annum coupon, paid semi-annually, currently sells for $95. A second bond maturing in 5 years has an 8% per annum coupon, paid semiannually, and currently sells for $97. Both bonds have a face value of $100. What is the 5-year zero rate implied by the two bonds? Hint: construct a portfolio using the two bonds that receives no payments until five years from now.

3. Go to Yahoo finance (or another financial website). Look up the options on Apple stock (AAPL). Write down the price of Apple stock and the price of the November call option with strike price $110. Record the day and time that you are recording the prices.

4. Go to the CME website. List your favorite contract from each of the following areas: commodities, equities, and currencies.

5. Look up the following terms related to futures contracts: first intent date, open interest, and Eurodollars.
Chapter 5

Forward contracts

This chapter discusses forward contracts and how to price them.

5.1 What is a forward contract?

A forward contract (or forward) is an agreement between two parties, where one party agrees to deliver something, called the asset, to a second party for a predetermined price, called the forward price, at a predetermined date, called the delivery date. When the contract is entered, or begins, no money is exchanged. An agreement is simply made.

For example, suppose on February 28, a farmer agrees to sell 20,000 kilograms (kg) of grain for $15,000 to a grain distributor, to be delivered on May 15. The farmer and distributor have entered a forward contract, where the forward price is $15,000 and the delivery date is May 15.

In any forward contract, there are two parties. The party agreeing to sell the asset (for example, the farmer), has, or enters into, the short position. The party agreeing to buy the asset (for example, the distributor), has, or enters into, the long position. In financial lingo, we also say that the distributor is going long the grain and the farmer is going short the grain.

It is important to understand that the agreed upon forward price has been locked. While the value of the asset could change after the two parties have entered the contract, the delivery price cannot change.

Suppose on May 15, grain is trading at $1/kg. The farmer is committed to delivering the grain to the distributor for $15,000 even though the current price of the grain is $20000 \times 1 = 20,000$ dollars. Thus the farmer has incurred a
$5,000 loss by entering the forward contract. In other words, the value of the farmer’s short position in the contract on May 15 is $-5,000. Similarly, the value of the grain distributor’s long position is $+5,000 on May 15; in effect, the distributor is receiving the grain at a discount to the current market price, also known as the spot price. To realize this profit, the distributor simply pays $15,000 to the farmer for the grain, and resells it immediately for $20,000 on the market. (We ignore any associated transaction costs incurred by this strategy.) Naturally, neither party knew of this outcome on February 28 when the forward contract was entered.

In general, consider a forward contract on an asset with delivery date $T$ and forward price $K$. Let $S_T$ be the spot price of the asset at time $T$, that is, the price to buy or sell the asset at time $T$. Then the value of the long position in the contract just before the asset is exchanged is $S_T - K$. The reason for this is that the contract forces the buyer to purchase the asset at $K$, whereas the asset is actually worth $S_T$, resulting in a net gain of $S_T - K$, which may be positive or negative. Similarly, the value of the short position is $K - S_T$. In Section 5.4, we discuss how to value the long and short positions of a forward contract at any time before the delivery date.

5.2 The forward price versus the spot price

The spot price of an asset is the price at which the asset is currently trading. That is, today, you can buy the asset at the spot price from someone else selling that asset on the market.

The spot price differs from the forward price. The forward price refers to the delivery price attached to a forward contract negotiated today for delivery at time $T$. The forward price does not make sense unless a delivery date is specified. If today is September 1, one bushel of apples might have a spot price of $50, a forward price of $45 for an October 1 delivery, and a forward price of $65 for a January 1 delivery. (Perhaps the January apples have to be harvested from some distant, warmer orchard, or they need to be stored longer, which incurs a cost). We might refer to the October 1 forward price as the one-month forward price, and the January 1 forward price as the 4-month forward price, since October 1 and January 1 are one and four months in the future relative to September 1.

We now introduce some notation. Fix an asset and let $S_0$ denote its spot price, today, at time 0. Let $F_{0,T}$ refer to the $T$-year forward price today of
the asset. Often we will just write $F_0$, when the context is clear that the
delivery is taking place at time $T$. The relationship between $S_0$, the spot
price today, and the various forward prices $F_{0,T}$ is complex in the real world,
taking into account expectations of future weather, political, or financial
events. However, if we make certain assumptions about the asset and the
forward contract, then we can establish a relationship between the spot price
and the forward prices. A key assumption that we make is that when the
forward price is established in a forward contract, the contract has no value.
In other words, both parties choose the forward price in such a way that the
contract has no value.

5.3 Computing forward prices

Consider a forward contract on an asset for delivery at time $T$. The delivery
price attached to the contract, which we also called the forward price in
the previous section, is $K = F_0$. We now study how $F_0$ relates to the spot
price $S_0$. Recall from the previous section that when the forward price is
established, it is done in a way that ensures that the forward contract has no
value.

Example 5.1. Assume there are no arbitrage opportunities. The spot price
of a barrel of oil on October 15 is

$$S_0 = \$75.$$

The six-month risk free rate is 4%. Assume that it costs nothing to store
oil over the next six months. What should the 1/2-year forward price of one
barrel of oil be?

Consider a portfolio where we have:

• borrowed $75,
• bought one barrel of oil for $75,
• taken a short position in one forward contract for oil, with delivery
date $T = .5$ years.

What is the value of this portfolio on October 15? It consists of three po-
sitions: a loan for $75, a barrel of oil which costs $75, and a short position
in a new forward contract. These positions each have value $-75, 75,$ and $0$
dollars, respectively. Hence, adding these three numbers up, we find that the portfolio has no value.

But what is the value of the portfolio on the delivery date? To calculate that, we study what happens to the portfolio on that day:

- we deliver the oil to the party who went long the forward contract.
- we receive $F_0$ from the party who went long (the delivery price)
- we repay the loan of $75e^{0.04(\frac{6}{12})} = 76.52$

Our net cash flow (that is, the value of the portfolio) is guaranteed to be

$$-76.52 + F_0.$$ 

Since there is no risk in this investment plan and no money wagered at the beginning, we could not have made a profit without violating the no-arbitrage assumption; hence,

$$-76.52 + F_0 \leq 0.$$ 

A similar argument with the opposite positions shows that $-76.52 + F_0 \geq 0$. Thus,

$$F_0 = 76.52 = 75e^{0.04\times \frac{6}{12}}.$$

Following this example, we can derive the general formula for the forward price of an asset. First we suppose the asset pays no dividends and requires no costs to store.

**Formula 12 (The forward price with no costs or income).** Let $r$ be the $T$-year risk-free rate with per annum continuous compounding. Consider a forward contract for delivery of the asset at time $T$ with forward price $F_0$. Then

$$F_0 = S_0e^{rT}.$$ 

We next derive the forward price when the barrel of oil costs money to store.

**Example 5.2.** Consider the same example as above, but this time assume the barrel of oil costs $5 to store for 6 months, paid upfront.

Now we have to construct a portfolio consisting of a barrel of oil, the fee for the storage, a loan to cover the barrel and the fee (for $80), and a short
position in the forward contract. As we have set it up, the portfolio has no value today.

At time \( T = 0.5 \) years later, the portfolio is worth

\[
-80e^{-0.04(0.5)} + F_0,
\]

since the loan was for $80, leading to the conclusion that

\[
F_0 \leq 80e^{-0.04(0.5)} = 81.62.
\]

What happens if we want to prove the opposite inequality? We would construct a portfolio by taking a long position in the forward contract, selling short one barrel of oil, and investing the proceeds of the short sale. We intend to hold this portfolio for 6 months, so by selling short the oil, we have removed the need for the original owner to pay storage to a storage facility. Instead, we get the $5 storage payment for playing the same role as a storage facility. In other words, our proceeds from the short sale together with the agreement to hold the short position for 6 months totals $75 + $5 = $80. Carrying through the analysis for this new portfolio, we deduce the opposite inequality and hence that

\[
F_0 = 80e^{-0.04(0.5)} = 81.62.
\]

We would like a formula which applies to this example as well as the previous one. To achieve this, we introduce the notion of the effective spot price, which we use again in Sections 6.8 and 11.4 when discussing option prices. The effective spot price, \( S_{\text{eff}}^0 \), is the value of the asset at time 0 given that you will not actually possess the asset between times 0 and \( T \), but will own it at time \( T \).

For example, if the asset is a stock, the effective spot price accounts for the fact that you will not receive any dividends between times 0 and \( T \); thus, the effective spot price is the price of the stock minus the present value of the dividends to be received

\[
S_{\text{eff}}^0 = S_0 - I.
\]

Naturally, the effective spot price is lower than the spot price since you are deprived of these dividends.
Conversely, if the asset is a commodity, then the effective spot price is higher since you do not have to store the asset. Thus, the effective spot price of a commodity is the market price plus the costs for storing it

\[ S_{0}^{\text{eff}} = S_0 + C. \]

The effective stock price is a little more subtle when there is a continuous effect on the asset. Suppose the asset is an index fund based on many stocks paying dividends at different times. Since one “index fund asset” is made of many small fractional assets, the index-asset receives many small payments, which is essentially equivalent to receiving a continuous dividend stream. Recall the present-value and continuous compounding discussion in Chapter 1. When we compute the present value of $100 in \( T \) years, we “stripped” the money of its interest which it continuously earned to conclude the present value was \( 100e^{-rT} \). So suppose we say that the asset earns a continuous (fixed) dividend at a rate of 100\( q\)% per annum, then the price of the asset stripped of this dividend is

\[ S_{0}^{\text{eff}} = S_0 e^{-qT}. \]

Another common example is when the asset is one unit of foreign currency. The asset, deposited in the central bank of the foreign government, earns interest at the foreign risk-free rate, say 100\( r_f\)% per annum compounded continuously. Since this can be viewed as a continuous dividend

\[ S_{0}^{\text{eff}} = S_0 e^{-r_f T} \]

where \( S_0 \) is the spot price of one unit of foreign currency in dollars. We present an example of this below.

In the previous example, we saw that

\[ F_0 = 81.62 = (75 + 5)e^{0.4 \times 0.5} = S_0^{\text{eff}} e^{0.4 \times 0.5}. \]

**Formula 13. (Forward price with costs and/or dividends)** Let \( S_0^{\text{eff}} \) be the effective spot price of an asset. Then, the forward price is

\[ F_0 = S_0^{\text{eff}} e^{rT}. \]

Note that this formula includes the previous one, since if the asset has no carrying costs or dividends, then \( S_0^{\text{eff}} = S_0 \).
Example 5.3. The current exchange rate is 1 dollar (USD) to 0.8 euros (EUR). Assume the risk-free rate on all maturities in the US is 2% per annum and the risk-free rate in Europe is 3% per annum. The forward price for one EUR to be delivered in 9 months is $1.35. Construct an arbitrage opportunity.

Solution:

Because we consider the euro to be the “asset”, we need to compute the price of one EUR in terms of dollars

\[ S_0 = \frac{1}{0.8} = 1.25 \]

USD. Then we use Formula [13] to compute what should be the forward price

\[ F_0 = S_0^{eff} e^{rT} = 1.25e^{-0.03(0.75)}e^{0.02(0.75)} = 1.24 \]

USD.

The forward contract is overvalued by the market so there is an arbitrage opportunity.

- Take a short position in the forward contract. That is, promise to sell one EUR in 9 months for 1.35 USD.
- Borrow (in the United States)

\[ S_0^{eff} = 1.25e^{-0.03(0.75)} = 1.22 \]

USD to buy \( e^{-0.03(0.75)} = 0.98 \) EUR.
- Deposit the 0.98 EUR in a European bank.

In 9 months, you will have

\[ 0.98e^{-0.03(0.75)} = e^{-0.03(0.75)}e^{0.03(0.75)} = 1 \]

EUR which you deliver for 1.35 USD. You repay your (United States) loan of \( 1.22e^{0.02(0.75)} = 1.24 \) USD. Your net profit is \( 1.35 - 1.24 = 0.11 \) USD.

We make two remarks about this example. First, we did not need to know the exchange rate in 9 months, just like we do not need to know the price of any underlying asset in the future when computing its forward price. Second, we only needed to buy 0.98 EUR instead of 1 EUR, in order to have 1 EUR available for delivery in 9 months.
5.4 Valuing a long or short position in a forward contract

In deriving the forward price in the previous section, we made the assumption that a forward contract has no value when it is first negotiated. After some time has elapsed, however, the contract will likely have nonzero value. It will have positive value for one party, and the negative of that value for the other party. The change in value is due mostly to the change in spot price of the asset, but it is also affected by changes in interest rates and the shrinking time until delivery.

Let us return to the example of the barrel of oil in Example 5.1 (so we are ignoring storage costs). Suppose that on November 15, the spot price of oil is $90 per barrel and that the risk-free rate is \( r = 5\% \) for all times. Consider a new portfolio constructed on November 15 consisting of two positions: one barrel of oil and a short position in the forward contract that was negotiated on October 15 with a delivery price \( K = 76.52 \). What is the value of this portfolio? It is equal to \( S_0 = 90 \) (for the oil) plus \( f_s \), where \( f_s \) is the value of the short position of the contract, which we are trying to figure out.

What is the value of the portfolio at the delivery date 5 months later? The oil gets exchanged for \( K = 76.52 \) dollars under the contract. In other words, the portfolio consists only of that $76.52 in cash. Now we know the value today of a portfolio that pays cash at a future time: it is the present value of the cash. Thus, the value of the portfolio on November 15 is

\[
76.52e^{-0.05\left(\frac{5}{12}\right)} = 74.94,
\]

and so we get the equality:

\[
90 + f_s = 76.52e^{-0.05\left(\frac{5}{12}\right)} = 74.94,
\]

or \( f_s = -15.05 \) dollars. Since the long and short positions have opposite values, the value of the long position \( f \) is $15.05.

In general, if we take a forward contract (perhaps negotiated in the past) with a delivery time of \( T \) and a delivery price of \( K \), then

**Formula 14.** The value \( f \) of the long position of the forward contract is

\[
f = S_{0}^{\text{eff}} - Ke^{-rT},
\]

where \( r \) is the \( T \)-year risk-free rate and where \( S_{0}^{\text{eff}} \) is the effective spot price.
5.5. PROBLEMS

Note that \( f = 0 \) if the delivery price \( K \) equals the current forward price \( S_0^{\text{eff}}e^{rT} \), which is consistent with the claim that the initial value of a long (and short) forward position is 0.

5.5 Problems

1. A trader enters into a short forward cotton contract with delivery in one year. The forward price is $0.50 per pound. The delivery size is 50,000 pounds. How much does the trader gain or lose if the cotton price in one year from now is (a) $0.482 per pound? (b) $0.513 per pound?

2. The price of heating oil today (October 16) is $2.10 per gallon. A heating oil company offers customers the opportunity to lock in heating oil for January 16 delivery at the forward price. The risk-free rate is 5% for all times (continuous compounding). What is the price they offer their customers? What if it costs the company 3 cents to store a gallon of heating oil for 3 months starting today?

3. Continuing with the previous example, suppose that on November 16, heating oil has dropped to $1.80 per gallon and the risk-free rate has risen to 6% for all times. How much has the customer gained or lost by locking in the price? How much has the company gained or lost? First compute your answer ignoring the storage costs and then compute your answer if it costs 2 cents on November 16 to store a gallon of heating oil for 2 months.

4. The price of oil is currently $100 per barrel. The contract size is one barrel. The forward price for delivery in one year is $130. You can borrow money at 7% per annum with annual compounding. (Credit is rather tight.) Assume the cost of storing one barrel of oil is nothing (you have some free room in the garage), nor does it provide any income. Describe an arbitrage opportunity.

5. Suppose the risk free rate is 5% per annum with continuous compounding. Suppose the forward price on a (non-dividend paying) stock for delivery in 1 year is $60. Let \( F \) be the forward price on the same stock
but with delivery in 18 months. Prove using no-arbitrage that

\[ F \leq 60e^{0.05 \times (\frac{18}{12} - 1)}. \]

6. A trader can buy gold at $850 per ounce and sell it at $849 per ounce. The trader can borrow money at 6% per year and invest money at 5.5% per year. Both rates are with annual compounding. For what range of one-year forward prices of gold does the trader have no arbitrage opportunity? Assume no storage costs. (Hint: this problem has two parts, the upper bound and the lower bound.)

7. Suppose the price of one EUR is 1.15 USD. The European risk-free rate is 4% per annum with annual compounding, and the American risk-free rate is 4% per annum with continuous compounding. What is the forward price of one EUR to be delivered in 9 months?

8. A stock is expected to pay a dividend of $1 per share in two months and again in five months. The stock price is $50 and the risk free rate is 4% for all times. An investor has taken a short position in a six-month forward contract on the stock. What is the forward price? What is the initial value of this forward contract?

9. This is a continuation of the previous problem. Three months later the price of the stock is $48 and the risk free rate is still 4%. What is the 3-month forward price for the stock at this time? What is the value of the investor’s position in the forward contract from the previous problem?

10. Derive the general formula for the value of a short position in a forward contract if the asset yields a continuous dividend of \( r \) per annum, the spot price is \( S_0 \), and the delivery date is in \( T \) years.
Chapter 6

Options

In this chapter, we introduce options and discuss their basic properties. We also further develop the No Arbitrage Hypothesis and recast it as the Principle of Math Finance.

6.1 Definitions

In Section 4.2.3 we gave a brief introduction to options. An option or option contract is like a forward contract except that one party gets to decide whether the exchange of the underlying asset will take place.

There are two types of options, call options and put options, and each of these has two flavors, American or European. In all these variations there are always two parties. There is the buyer of the option and the seller of the option. The buyer pays the seller a fee, or premium, to acquire the option. Note that this differs from the case of a forward contract where there is no exchange of payments when the forward contract is negotiated.

First, we discuss call options. In an American call option, the buyer of the call option pays the seller for the right to buy an asset at a fixed price $K$ (the strike price) at any time up until a given date (the expiration date). The buyer of the option has the right to buy the asset at $K$, but the buyer has no obligation to do so. On the other hand, the seller of the contract is obligated to sell the asset to the buyer at $K$ if the buyer chooses to buy the asset. When the buyer decides to buy the asset, we say the buyer exercises the option and the seller is assigned the option. It is possible that the buyer may choose not to exercise the option; this will happen, for example, when
buying the asset at $K$ is a losing proposition, i.e., the asset is trading for less than $K$ on the market.

In a European call option the buyer can only exercise the option at the expiration of the option, not sooner.

Next, we discuss put options. In a put option, the buyer of the put option pays the seller for the right to sell an asset at a fixed price $K$. In the case of an American put option, the buyer can exercise this right to sell the asset at any time up until expiration, whereas in the case of a European put option, the buyer gets to decide whether to exercise the put option only on the expiration date.

Generally, options trade on an exchange and buyers and sellers of options are matched by the exchange. The primary exchange in the US is the Chicago Board of Options Exchange (CBOE). The exchange (or its clearinghouse, which may or may not be part of the exchange itself) decides how to match buyers and sellers when the buyer of an option decides to exercise an option: the exchange will randomly assign the option to an investor who has sold the option.

The buyer of an option is said to have a long position in the option; the seller is said to have a short position. The seller is also called the writer of the option, because in order for an option to be created, there needs to be an investor who creates, or writes, the contract.

Options are usually written on stock, and options that trade on exchanges are American options. For example, according to Yahoo! Finance on October 28, 2008, the call option on IBM with expiration November 21, 2008 and strike price $K = 90$ traded for $4.70. IBM stock traded for $89.29. In other words, the right to buy one share of IBM at $90$ on or before November 21 was selling for $4.70. There were 15,222 such contracts in existence (this amount is called the open interest in the contract). In order for a new contract to be created, a seller would have to approach the market with an offer to sell the $15,233^{rd}$ contract. Options are also commonly written on equity indices and on futures contracts.

At the current time $t < T$, a call (put) option is said to be in-the-money if $S_t > K$ ($S_t < K$). The call (put) is said to be out-of-the-money if $S_t < K$ ($S_t > K$). Both the call and put are said to be at-the-money if $S_t = K$. The terminology is self-explanatory. In the example above, the call option with $K = 90$ is out-of-the-money, but a call option on IBM with $K = 85$ is in-the-money.

Options contracts are normally written on 100 shares of stock and the
price of the contract is quoted per share. In most of our examples, we will pretend that it is possible to transact in options on a single share in order to simplify the discussion.

6.2 Some examples

Suppose you buy a call option on IBM stock with expiration on November 21 and strike price 90. With the premium quoted in the last section, you would pay $4.70 to buy this call option.

Let us suppose that this option is a European option. Then the option cannot be exercised until November 21, when the option expires (options are structured to expire on the third Friday of the month). If on that Friday, IBM is trading at 90.01 or higher (say $96), you would definitely exercise the option: you call up your broker, your broker notifies the exchange, the exchange assigns your exercise to someone who has a short position in the call option, that person sells you the share of IBM at 90, and then you immediately sell the share on a stock exchange at 96 (if you do not want to hold it). The amount you gain by converting the option to cash is 6 dollars; in other words, this particular call option at expiration is worth $6. Your total profit, ignoring transaction costs, is $6 − $4.70 = $1.30.

On the other hand, if IBM were trading at $89 at expiration, it would make no sense to exercise the option. Why would you buy IBM at $K = 90 as permitted in the option contract when you could buy it more cheaply in the marketplace? In this case the option is worthless at expiration. The profit on buying the option is $0 − $4.70 = −$4.70.

It is worth noting that if you buy or sell an option, you are not required to hold the contract until expiration. You are always free to go to the market and sell the option (or if you have a short position, you can buy an option to close out your position).

Next, suppose you already own 1 share of Morgan Stanley (Stock symbol: MS). You decide to write a call on 1 share of MS with strike price 60 and expiration on the third Friday in March. The option exchange finds someone who wants to own a call option on MS and you get paid 5 dollars from this person (this amount is determined by trading in the existing call option contracts on MS stock). Now suppose that the option is an American option and MS pays a dividend of 2 dollars in February. It is likely that the owner of the option will choose to exercise the call option early, right before the
dividend is paid by Morgan Stanley. If this occurs, then you could be assigned to sell your 1 share of MS. If so, you are obligated to sell the MS share for $60 at that time. You will not receive the dividend. A portfolio consisting of a long position in a stock and a short position in a call option on that stock, such as the one just discussed, is called a **covered call**.

Finally, suppose you forecast that the price of Google stock (GOOG) will fall dramatically in the short term. One way to speculate on this hunch is to buy an American put option on GOOG. On October 28, suppose GOOG trades at $357. A December put option on GOOG with strike price $340 trades for $15. You pay the $15 and buy the put. Now suppose that GOOG drops to $310 during the first week in November. This would value the put option at a minimum of $30 since you can always exercise the put and make $340 - 310 = 30$ dollars. On the one hand, you could exercise the put and receive $30. But it is very likely that the put option in the market is worth more than $30. So instead of exercising the put, you could go ahead and sell the put, making a nice profit on your investment.

Or if you prefer, you can hold onto the option for a longer time, any time until the third Friday of December. If you do wait until the expiration date and if GOOG is trading below $340 on the third Friday of December, then you would definitely exercise the option and if it is trading above $340, you would definitely not exercise the option. For example, if GOOG is trading at $333, then exercising the put option allows you to sell GOOG at $340. If you do not already own the shares, then you would first have to buy a share at $333. The amount you gain by exercising the option is $340 - 333 = 7$. In this scenario, you lost money overall (namely $8), even though the option was worth $7 at expiration. Note that if GOOG had risen dramatically, your maximum possible loss is the $15 that you paid for the put option.

### 6.3 Payoff of a portfolio

The **payoff** of a portfolio is the value of the portfolio after the portfolio is liquidated and all positions are turned into cash. Of course the price, or value, of a portfolio is the same whether we liquidate it or not (at least in theory), so usually the term payoff is reserved for some interesting time when the portfolio undergoes some change, such as when an option expires or a forward contract matures. If no specific time is given, the payoff always refers to this *interesting* time. Unless otherwise specified, the payoff of a contract
or portfolio refers to the long position in the contract or portfolio.

**Example 6.1.** The payoff of a forward contract which matures at time $T$ and has delivery price $K$ is $S_T - K$. The owner of the contract must buy the asset at $K$ and can immediately sell the asset at $S_T$, even if this is a losing proposition.

**Example 6.2.** Consider a long position in a call option. A person with this portfolio has the option to buy the stock for $K$. If $S_T < K$, this option is worthless since the option holder can buy the stock for the cheaper price $S_T$ from the market. Hence in that case the payoff of the option is 0 since the owner does nothing (we say the option “expires worthless”). It $S_T > K$, the holder exercises the option to buy the stock for $K$, and can resell it in the market for $S_T$, netting a cash flow of $S_T - K$. We sometimes call this the **exercise region** of the option. If $S_T = K$, then both approaches net the option holder $0. These two scenarios can be encapsulated in one mathematical expression: the payoff (or value) of the long call position is

$$
\max(S_T - K, 0) = \begin{cases} 
0 & \text{if } S_T < K \\
S_T - K & \text{if } K \leq S_T 
\end{cases}
$$

See the solid graph in Figure 6.1. Note that the writer of the call has the exact opposite payoff

$$
- \max(0, S_T - K) = \min(0, K - S_T).
$$

**Example 6.3.** The payoff of a put option with strike price $K$ and expiration at time $T$ is

$$
\max(K - S_T, 0)
$$

To see this, note that if the stock price $S_T$ is bigger than $K$, there is no point in exercising the put option since this would require the owner to buy the stock at $S_T$ (in the marketplace) and sell it at $K$ (according to the put option agreement), a losing proposition. On the other hand, if $S_T$ is less than $K$, the owner of the option will earn $K - S_T$ by buying the stock at $S_T$ and selling at $K$. 

Example 6.4. If you build a portfolio which is long a call and short a put, both with strike price \( K \) and expiration \( T \), then if \( S_T > K \) the call will have a payoff of \( S_T - K \) and the put will have payoff 0. On the other hand, if \( S_T \leq K \), then the call will have a payoff of 0 and the position in the put option will have a payoff \( S_T - K \) (since we are short the put). Hence the payoff will always be \( S_T - K \). This shows that this portfolio is equivalent to a portfolio consisting of one single forward with delivery price \( K \) and expiration at \( T \).

The profit of a portfolio is the payoff of the portfolio minus the cost to acquire the portfolio. Be careful not to confuse payoff and profit.

Example 6.5. In Figure 6.1 we see the graph of the payoff function of the call from the earlier example, \( \max(0, S_T - K) \). If the price (premium) of the option is \( c \) today, then the future value of the cost when the option
expires is $ce^{rT}$. Thus the (future value of the) profit is the pay-off less cost:

$$\text{profit} = \text{payoff} - \text{future value of cost} = \max(0, S_T - K) - ce^{rT}.$$ 

The call option’s profit region is where the profit function is positive

$$\{S_T \geq K + ce^{rT}\}.$$ 

This is not the same as the exercise region, which is $\{S_T \geq K\}$. When $K + ce^{rT} \geq S_T \geq K$, the option holder still exercises the call to reduce the loss.

We will sometimes simplify these computations by assuming $r = 0$ in order to ignore the time value of money when comparing cost and pay-off.

### 6.4 Some common portfolios with options

There are a number of common option portfolios. These are also referred to as option strategies.

#### 6.4.1 Straddle

A straddle is a portfolio consisting of a long position in a call and a long position in a put, both with the same strike price $K$ and expiration at time $T$. Usually the options are at-the-money or close to it.

To find the payoff of a straddle, note that at expiration, if $S_T \geq K$, the call position is worth $S_T - K$ and the put position is worthless; whereas if $S_T < K$, the put position is worth $K - S_T$ and the call position is worthless. Hence the payoff is

$$\text{Straddle}(S_T) = |S_T - K|,$$

the absolute value of $S_T - K$. If you buy a straddle, then you are betting that the stock price will move a lot, but you are not sure in which direction. We say that you are going long volatility (or buying volatility) because you hope that stock volatility will be high. We introduce the notion of volatility more precisely in Chapter 11.
Example 6.6. Consider a straddle with strike price 50 expiring in $T = 0.75$ years, whose underlying call price is 2 dollars and put price is 3 dollars. Assume the risk-free rate is (essentially) 0%. Find the profit region of the straddle.

Solution: The straddle costs $2 + 3 = 5$ dollars, so its profit is

$$|S_T - 50| - 5e^{0.05 \times 0.25} = |S_T - 50| - 5.$$ 

The profit region is therefore

$$\{S_T \geq 55 \text{ or } S_T \leq 45\}.$$ 

For example, if the current spot price were 49 (so the higher priced put was in-the-money and the lower priced call was out-of-the-money), then the stock would in 9 months need to decrease by 4 dollars or increase by 6 dollars for the straddle to be profitable. This is not an unreasonable goal for “normal” stock volatility.

For numerical examples involving non-zero risk-free rates, see the exercises.

6.4.2 Spreads

A spread option is created from either calls or puts, all with the same expiration time, possibly different strike prices.

A bull spread is a portfolio consisting of two options. In the call version of the bull spread, you are long one call at $K_1$ and short one call at $K_2$, with $K_1 < K_2$. Usually, $S_0 < K_1$, and the holder of a bull spread will benefit from a rise in the stock price, although the total profit is capped. The payoff of the bull spread is

$$\text{Bull}_c(S_T) = \begin{cases} 0 & \text{if } S_T < K_1 \\ S_T - K_1 & \text{if } K_1 \leq S_T \leq K_2 \\ K_2 - K_1 & \text{if } K_2 < S_T \end{cases}$$

A bull spread can also be created with two put options, in place of the call options.

In a bear spread, the investor takes the opposite position of a bull spread. In the call version of the bear spread, you are short one call at $K_1$
and long one call at $K_2$, with $K_1 < K_2$ and $K_2 < S_0$. The holder of a bear spread will benefit from a drop in the stock price, although the total profit is capped. A bear spread can also be created with two put options, in place of the call options.

**Example 6.7.** In one line, write the payoff function of a bear spread involving two puts with strike prices $20$ and $30$.

**Solution:** Going long a bear spread using puts is going short the put with strike $20$ and long the put with strike $30$, so we can write the payoff function as a difference

$$\text{Bear}_p(S_T) = \max(0, 30 - S_T) - \max(0, 20 - S_T)$$

If we graph the payoff function for this bear spread, we see that it is non-negative. So this bear spread must sell at a positive premium, like the call or put option. At first glance this is not obvious since the bear spread involves selling one put (or call) option and buying another. We will revisit the implied relationships among put (or call) prices later.

In a **butterfly spread**, there are three strike prices: $K_1 < K_2 < K_3$, where $K_2 = \frac{1}{2}(K_1 + K_3)$. It can be created with calls or puts. In the call version, the investor is long one call at $K_1$, short two calls at $K_2$, and long one call at $K_3$. In the put version, the investor is long one put at $K_1$, short two puts at $K_2$, and long one put at $K_3$. An investor might go long a butterfly spread if the stock price is near $K_2$ and the investor feels that the stock will move only moderately before the expiration time of the options. The exercises has several examples involving butterflies.

**6.5 The fundamental principle of math finance**

The fundamental principle of math finance is a very simple but useful idea for estimating, bounding and comparing the value of derivatives. It follows from the No Arbitrage Hypothesis. The principle applies to two portfolios, say Portfolio A and Portfolio B. Usually we cannot say how much a certain portfolio will be worth at a later time, but we might be able to say that Portfolio A will definitely be worth at least as much as Portfolio B.

**Example 6.8.** Portfolio A today consists of $100$ cash invested at 5%. Portfolio B consists of a zero coupon bond with principal $100$ maturing in 1
year. Certainly the bond will be worth $100 in 1 year (assuming no defaults) and certainly the cash in 1 year will be worth more than $100 (since it is earning interest). We can say for sure that Portfolio A will be worth more than Portfolio B in 1 year.

In the above example, it is intuitively clear that you would pay more now for Portfolio A since it is the one that will be worth more later— that is all there is to the fundamental principle.

Here is the precise statement of the fundamental principle:

**Fundamental Principle:** Take two portfolios, A and B. If the value of portfolio A can be guaranteed to be greater than or equal to the value of portfolio B at time \( T \), then the value of portfolio A is greater than or equal to the value of portfolio B at anytime before time \( T \), including right now.

The proof is easy. There is an arbitrage if the value of portfolio A turned out to be less than the value of portfolio B right now. You could sell portfolio B now and buy portfolio A now. Then wait until time \( T \). At that time, you would sell portfolio A and buy portfolio B. You would be guaranteed a profit. Since we do not believe arbitrage opportunities exist, this can not happen. \(^1\)

Applying this principle twice provides a special case we often use which we restate as: if the value of portfolio A exactly equals that of portfolio B at time \( T \), then their current values are equal.

**Example 6.9.** Suppose it costs \( C \) dollars to store oil for \( T \) years paid upfront. Let \( r \) be the \( T \)-year zero rate. Consider Portfolio A: 1 short position in a forward contract on oil with delivery price \( K \) at time \( T \) and one barrel of oil in storage until time \( T \). Portfolio A will definitely be worth \( K \) dollars at time \( T \) because the payoff from our position in the forward is \( K - S_T \) (we are short) and the payoff from the oil is \( S_T \). Said differently, we will sell the oil for \( K \) and that is all we will have left at time \( T \).

Next, consider Portfolio B consisting of the present value of \( K \) dollars, \( Ke^{-rT} \). If we invest this money at the \( T \)-year zero rate, then we are guaranteed that Portfolio B is worth \( K \) dollars at time \( T \).

Then the fundamental principle says that each of these portfolios is worth at least as much as the other now; in other words, they are worth the same.

\(^1\)Technically, we only need to guarantee that if we are long portfolio A and short portfolio B, then the value of this combined position is at least zero at time \( T \).
amount now. This allows us to prove our formula for the price of the forward now. Recall that \( f \) refers to the price of the long position in the forward.

Then Portfolio A is now worth \(-f + S_0 + C\), since it will cost \( S_0 \) to buy the oil now and \( C \) dollars to store it for \( T \) years. Portfolio B is now worth \( Ke^{-rT} \) (it is just cash). We conclude that \(-f + S_0 + C = Ke^{-rT}\), or

\[
f = S_0 + C - Ke^{-rT} = S_{0\text{eff}} - Ke^{-rT},
\]

our formula for the value of a long position in a forward contract.

6.6  Inequalities for option prices

The principle is very useful for getting inequalities for option prices. However, we have to be careful if we have a short position in an American option– since the option can be assigned before expiration, this might make it difficult to guarantee the payoff at expiration.

Example 6.10. Show that for a put option with strike \( K \) and expiration at time \( T \) that

\[
p \geq Ke^{-rT} - S_0
\]

where \( p \) is value now of the put option.

Solution: Build a portfolio with a long position in one put and one share of stock. Then this portfolio is guaranteed to be worth \( \max(K, S_T) \) at time \( T \). Why? It is worth at least \( K \) since we can exercise the option and sell the stock for \( K \). But if \( S_T > K \), we might as well forget about the put and just sell the stock in the market for \( S_T \).

Next consider a second portfolio with just cash equal to \( Ke^{-rT} \) (\( r \) as usual). Then this portfolio will be worth \( K \) at time \( T \). Since the maximum of \( K \) and \( S_T \) is always bigger than or equal to \( K \), we conclude that the value of the first portfolio is at least as big as the value of the second portfolio at time \( T \).

The fundamental principle then says that the value now of the first portfolio is at least as big as the value of the second portfolio. In other words,

\[
p + S_0 \geq Ke^{-rT},
\]

which is what we wanted.
Note that this inequality holds whether the option is European or American since we are going to go long the option if we need to arbitrage. However, there is a better inequality if the put is American. Since an American put can be exercised at any time, it is always worth at least \( K - S_0 \) (if not, buy the option and exercise it immediately). And this is a better inequality assuming that \( r > 0 \), which is typical. In other words, for an American put

\[
P \geq K - S_0 > Ke^{-rT} - S_0.
\]

We can improve the inequalities. Neither put is ever negatively priced, \( p, P \geq 0 \). Also neither put cannot be worth less than \( K \) when exercised. Such a maximum payoff occurs if the stock price is zero. Combining these observations with the previous inequalities, we get

\[
Ke^{-rT} \geq p \geq \max(Ke^{-rT} - S_0, 0)
\]

\[
K \geq P \geq \max(K - S_0, 0).
\]

For European call options, the analog of the put inequality from the previous example is

\[
c \geq S_0 - Ke^{-rT}.
\]

However, when the call is American, this inequality cannot be easily improved because

\[
S_0 - Ke^{-rT} > S_0 - K
\]

if \( r > 0 \).

This has an important consequence: it is never advantageous to exercise an American call on a stock with no dividends before expiration. Why? If you did, you would have something that is worth exactly \( S_0 - K \) at the time that you exercise (let us just pretend that time is now, hence the subscript zero). However, the lower bound says that the call is actually worth slightly more; it is worth \( S_0 - Ke^{-rT} \) which is bigger than \( S_0 - K \). Consequently, there is no point in exercising the call option since you will be giving it a lower value than it is really worth. Since the American call is worth at least as much as the European call, another way to say this result is that their values are exactly the same.

So comparing European and American options

\[
P \geq p \quad \text{while} \quad C = c
\]
when the stock pays no dividends. If the stock pays dividends, then we can only conclude that \( C \geq c \) instead of \( C = c \) since by exercising the American call early you earn the dividends.

**Example 6.11.** IBM stock trades at $100. The 1-year risk free rate is 3%. A European call option on IBM with strike $90 expiring in one year, costs $10. Construct an arbitrage opportunity.

**Solution:** Since 
\[
100 - 90e^{-0.03 \times 1} = 12.66 > 10,
\]
short-sell a stock, buy the call and invest the remaining \( 100 - 10 = 90 \). In one year, with draw the cash and buy the stock (back) either using the call or on the market to net a risk-free profit of
\[
90e^{0.03 \times 1} - \min(S_1, 90) \geq 92.74 - 90 = 2.74
\]
where here \( S_1 \) is the unknown price of IBM in one year.

We can also use the Fundamental Principle to compare European options with different strike prices.

**Example 6.12.** Compare the price of a European call \( c_1 \) with strike price \( K_1 \) and expiration \( T \) with the price of a European call \( c_2 \) with strike price \( K_2 > K_1 \) and the same expiration \( T \).

**Solution:** Consider the portfolio which is long the first call and short call. This is exactly the bull-spread described in Section 6.4. Note that the payoff of this portfolio Bull\(_c\)(\( S_T \)) is always greater than or equal to 0; thus, the present value of the portfolio, \( c_1 - c_2 \), must be also greater than or equal to 0:
\[
c_1 \geq c_2.
\]

A similar argument using either a bull spread or bear spread with European puts shows that
\[
p_2 \geq p_1.
\]
In the exercises, we derive some more inequalities using butterfly spreads.
6.7 Put-call parity

It turns out that owning a European put on a stock together with the stock itself is the same as owning a European call on the stock and some cash. In fact, this identity is often used by market-markers to balance out the number of put and call contracts. In this section $r$ always refers to the $T$-year zero rate.

Consider two portfolios:

- $A$ has 1 European put option on a stock with strike $K$ and expiration $T$ and 1 share of the stock
- $B$ has 1 European call option on a stock with strike $K$ and expiration $T$ and $K e^{-rT}$ in cash.

Then the value of portfolio $A$ at $T$ is

$$\max(K, S_T),$$

and the value of portfolio $B$ at $T$ is the same, as you can check. We can conclude from the fundamental principle that the value of portfolio $A$ is worth the same as portfolio $B$ right now. Hence, we get the formula for **put-call parity** for European options:

$$p + S_0 = c + K e^{-rT},$$

where $p$ is the value now of the put and $c$ is the value now of the call.

The reason we must assume that the options are European is that if they are American we can only get the inequality to go one way since we can be assigned the option. For American options, we can only conclude that:

$$K \geq P - C + S_0 \geq K e^{-rT}$$

(see the exercises).

**Example 6.13.** Suppose that the current price of a stock is $31. The risk-free rate is 10%. The price of a 3-month European call with strike $30$ is $3$. The price of a 3-month European put with strike $30$ is $2.25$. Construct a risk-free profit involving one call and one put.
6.8. DIVIDENDS AND CARRYING COSTS

Solution: Let Portfolio A be the call and $Ke^{-rT}$ in cash. Let Portfolio B be the put and one share of stock. As we saw above, these must be worth the same. So we check to see if put-call parity holds:

\[ c + Ke^{-rT} = 3 + 30e^{-1 \times 3/12} = 32.26 \]
\[ p + S_0 = 2.25 + 31 = 33.25 \]

Portfolio B is overpriced compared to Portfolio A. So buy Portfolio A and sell Portfolio B.\(^2\)

At time 0, our cash flow is $33.25 - 32.26 = $0.96. In 3 months, no matter what happens, our cash-flow is \(\max(S_T, K) - \max(S_T, K) = 0\).

We can also combine inequalities from Section 6.6 with put-call parity to derive other inequalities among the call or put prices.

**Example 6.14.** Let \(p_1\) and \(p_2\) be the values of two European put options, with strike prices \(K_1 < K_2\), both expiring at \(T\), and both with the same underlying asset whose spot price is \(S_0\). Derive an upper and lower bound for \(p_1\) in terms of \(p_2\).

**Solution:** As mentioned in Section 6.6, we can use a bull or bear spread to conclude that

\[ c_1 \geq c_2, \quad p_1 \leq p_2. \]

Apply put-call parity to both sides of the first inequality

\[ p_1 + S_0 - K_1e^{-rT} \geq p_2 + S_0 - K_2e^{-rT}. \]

Simplifying this and combining it with the second inequality, we get

\[ p_2 \geq p_1 \geq p_2 - (K_2 - K_1)e^{-rT}. \]

6.8 Dividends and carrying costs

Most options are written on individual stocks. Many stocks pay dividends. This affects the behavior of call options and put options in different ways.

\(^2\)We do not know if B is overpriced or A underpriced, or both. For example, selling B and doing nothing with A is not an arbitrage opportunity, just a speculative gamble.
As we saw above, there is no reason to exercise an American call option early when the stock pays no dividend. But if there are dividends, this might not be the case.

Roughly speaking, when a stock pays a dividend, its price drops a bit to reflect the payout of the dividend (see the discussion on effective spot price in Section 5.3). So it might be better to exercise right before a dividend payment, especially close to expiration time, since otherwise you would lose this small amount by which the stock price drops.

The formula for put-call parity for a European option also changes. For a stock paying a dividend, let $D$ be the present value of all of its dividends between now and expiration. Then,

$$p + S_0 - D = c + Ke^{-rT}.$$ 

See the exercises for a derivation of this and an example.

In general, put-call parity uses the effective stock price

$$p + S_{0}^{\text{eff}} = c + Ke^{-rT}$$

where the effective spot price differs from $S_0$ due to any or the mentioned reasons, including carrying costs and dividend yields. Other inequalities for European options are also modified by replacing $S_0$ with $S_{0}^{\text{eff}}$. For example, the bounds on the European put become

$$Ke^{-rT} \geq p \geq \max(Ke^{-rT} - S_{0}^{\text{eff}}, 0).$$

6.9 Problems

1. Consider the following three European call options, all with expiration at time $T = 1$: option A has strike $10$; option B has strike $15$; and option C has strike $20$.

   (a) Create a bull spread from options A and B, and graph its payoff as a function of $S_T$.

   (b) Create a bear spread from options B and C, and graph its payoff as a function of $S_T$.

   (c) Create a butterfly spread from options A, B, and C, and graph its payoff as a function of $S_T$. 
2. Consider the three calls from the previous problem expiring in $T = 1$ year. Suppose that the premium of the three options are $7, 3, 1$ for $A, B, \text{ and } C$, respectively, today (time 0). Suppose the spot price on the underlying asset today is $14$.

(a) Which call(s) is out-of-the-money?
(b) Suppose the risk-free rate is constant at 0%. Graph the profit function for each of the previous strategies.
(c) Suppose the risk-free rate is constant at 5%. Graph the profit function for the butterfly strategy. By how much would the asset have to change over the year for your butterfly portfolio to be a loss (negative profit)?

3. Consider the set-up in the previous problem, where the risk-free rate is 0%.

(a) What are the prices of the corresponding European puts, $A', B', C'$?
(b) Graph the pay-off and profit functions for the straddle using call $B$ and put $B'$.

4. Consider a certain butterfly spread on American International Group stock (AIG): this is a portfolio that is long one call at $50$, long one call at $70$, and short 2 calls at $60$. Assume expiration of all options is at the same time $t = T$.

(a) Graph the payoff of this portfolio at expiration $T$ as a function of the stock price $S_T$ of AIG.
(b) If today the calls cost $13.10, 5.00, \text{ and } 1.00$ for the strikes at $50, 60, \text{ and } 70$, respectively, what will be the profit or loss (PnL) from buying this spread if the stock turns out to be trading at $55$ at time $T$? at $35$?
(c) Explain, using the fundamental principle, why the spread must have a positive value now.
(d) Deduce that the price of the call with $K = 60$ is at most the average of the prices of the other two calls.

5. Consider Portfolio A: long 1 American put on a stock with strike price $K$ and expiration at $T$; short 1 American call on the same stock with the same strike and expiration date; and long 1 share of stock.
(a) If you are long portfolio A and the call is exercised, show that you will have at least $K$ dollars plus the value of the put at time $T$, no matter when the call is exercised.

(b) If you are long portfolio A, show that you are guaranteed a payoff of at least $K$ at time $T$. Conclude that

$$P - C + S_0 \geq K e^{-rT}.$$

6. Next, consider Portfolio B: short 1 American put, long 1 American call, short 1 share of stock, and $K$ dollars in cash. Assume the options have the same strike price $K$ and expiration $T$ and that they are options on the stock.

(a) Show that if you are long portfolio $B$ and the put is exercised, then you are guaranteed a payoff of at least zero at time $T$, no matter when the put option is exercised.

(b) Now show that whether the put is exercised or not, you are guaranteed a payoff of at least zero. Conclude that

$$C - P - S_0 + K \geq 0.$$

Combining with the previous problem, this gives the put-call inequality for American options:

$$K \geq P - C + S_0 \geq K e^{-rT}.$$

7. Suppose that the current price of an asset is $31. The asset is expected to pay a dividend of $3 in one month. The asset also has an upfront carrying cost associated to it of $1. The risk-free rate is 5%. The price of a 3-month European call on this asset with strike $30 is $4. The price of a 3-month European put with strike $30 is $3.25. Construct a risk-free profit involving one call and one put.

8. Let $c$ and $p$ be the price today (at time $t = 0$) of a European call and put, respectively, with strike $K$ and expiration date $T$. Let $S_0$ be the spot price of the underlying asset, which pays a dividend. Let $D$ be the present value of the dividend. Assume for simplicity that the risk-free zero curve is flat and fixed at $r$ from $t = 0$ until $t = T$. Show that

$$p + S_0 = c + D + K e^{-rT}$$

using the fundamental principal.
9. Consider a call option on IBM with strike price $100 and expiration in 5 months. Suppose IBM is trading at $112. The risk-free rate is 4.5\% for all maturities. IBM will pay a dividend in 3 months of 40 cents per share.

(a) What is a lower bound for the value of the option?

(b) Suppose this call option is trading on the Philadelphia Stock Exchange for $18.50. What is the price of the corresponding put option if both options are European?

10. Consider a European put option on Google with strike price $400 and expiration in 5 months. Suppose Google is trading at $380. The risk-free rate is 4.5\% for all maturities.

(a) If the put trades for $15 is there an arbitrage opportunity? Construct one if there is.

(b) If the put trades for $15 and Google pays a dividend of $10 per share in 1 month is there an arbitrage opportunity? Construct one if there is.

11. Consider an American put and a European put with the same strike price \( K \) and expiration \( T \). In this problem, we investigate the relationship between the price \( P \) of the American put and the price \( p \) of the European put. We already know that \( P - p \geq 0 \) since an American put is always worth at least as much as a European put.

Consider a portfolio where you sell one American put and buy one European put and invest the proceeds \( P - p \). Note that the initial value of the portfolio is zero. In this problem, assume that risk-free zero curve stays flat and fixed at \( r \) until time \( T \).

(a) Show that if \( r = 0 \), then the portfolio is guaranteed to be worth at least \( P - p \) at time \( T \). Deduce that \( P = p \).

(b) Show that if \( r > 0 \), then the portfolio at time \( T \) will be worth at least

\[
(P - p)e^{rT} + K - Ke^{r(T-t')}
\]

where \( t' \) is either the time when the American put is exercised or it equals \( T \). Since \( e^{r(T-t')} \leq e^{rT} \), conclude that \( P \leq p + K(1 - e^{-rT}) \), or else there is an arbitrage opportunity.
12. Let $p_K$ be the current price of a European put expiring at time $T$ with strike price $K$. Let $S_0$ be the spot price of the underlying asset. Compare the following quantities if possible. If not enough information is available to make a definitive comparison, be sure to indicate that that is the case.

(a) Compare $p_{50}, p_{55}, p_{60}$.
(b) Compare $p_{50}$ and 50.
(c) Compare $p_{50}$ and $S_0$.
(d) Compare $2p_{54}$ and $p_{50} + p_{60}$.
(e) Compare $2p_{56}$ and $p_{50} + p_{60}$.
(f) Compare $p_{55} - p_{50}$ and $p_{60} - p_{55}$. 
Chapter 7

Probability and Statistics I

In this chapter we review some of the basics of probability. We define random variables, expectation (for finite sample spaces), variance, covariance, standard deviation, and correlation.

7.1 Introduction

Suppose you toss a coin. It can come up heads or tails. If you toss the coin many times, you expect to see an equal number of heads and tails (assuming the coin is fair). The idea behind probability is to assign a number to each outcome of the coin toss to reflect the fact that if you toss the coin more and more times, the number of heads that will show up gets closer and closer to 50% of the total number of flips, and similarly the number of tails gets closer and closer to 50% of the total number of flips. Formally, we assign a probability of 0.5 to each of the two outcomes.

Next, suppose you roll a pair of dice, one blue die and one red die. The possible outcomes are all pairs \((i, j)\) where \(i = 1, 2, \ldots, 5,\) or 6 is number on the blue die and \(j = 1, 2, \ldots, 5,\) or 6 is the number on the red die. There are 36 possible outcomes in total. Each one of these outcomes should be equally likely and so we assign a probability of \(1/36\) to each. By running this experiment (i.e., rolling the pair of dice) many times, we would be able to verify that our probabilities are correct.

For many experiments, we are not in a position to run them over and over. Consider the example of the price of a stock and let the experiment this time be the price of the stock at 4 p.m. on November 14. Now, we have only one
look at the price. We could try to estimate the probability that the stock price is at $45.11 or $13.23, but we would not be able to verify easily how accurate our probabilities are. If the stock is trading at $14 on November 13, then it seems more likely that the stock will be at $13.23 than at $45.11. But how much more likely is a difficult question. An important topic in finance is developing models for the price of a stock. We will introduce a model of stock price; while the model will be rather rough, it will be a good place to start for estimating the probability that a stock attains a certain price at a later time.

Even though we cannot always know for sure what the probabilities of certain outcomes are, we can still develop the formal language of probability.

### 7.2 Definition of probability

Consider an experiment with various possible outcomes. The experiment may be able to be performed repeatedly or it may not. The set of all outcomes, called the **sample space**, is denoted by Ω (the Greek letter omega). In the coin toss example from the previous section,

\[ Ω = \{\text{Heads, Tails}\}, \]

and in the roll of a pair of dice,

\[ Ω = \{(i, j) \mid i \in \{1, 2, \ldots, 6\} \text{ and } j \in \{1, 2, \ldots, 6\}\}. \]

For the experiment recording the price of a stock at 4 p.m. on November 14, Ω consists of all nonnegative numbers. Since stock price is usually quoted in cents, we might want to restrict to nonnegative decimals with two decimal places. Then the price of the stock can (in theory) take on an infinite number of values; hence, Ω is an infinite set in this case.

In probability theory we assign a number, called the **probability**, to each outcome of the experiment and also to every combination of outcomes. The reason behind this becomes clearer when Ω is infinite. For example, suppose you throw a dart at a dartboard. The chance that you will hit a specific point on the dartboard is zero. In other words, each individual outcome has no chance of occurring! However, if you divide the board in half and pick the right half, then you will likely hit the right half roughly 50% of the time. The probability that would be assigned to the combination of all of
the outcomes on the right half is 0.5. Similarly any combination of outcomes would have a probability assigned to it. For example, you might want to know the probability that you hit a bulls-eye.

We refer to combinations of outcomes as **events**. In the language of set theory, an event is a subset \( A \) of \( \Omega \). Recall, that we write this as 

\[ A \subset \Omega. \]

It is helpful to understand how the three basic operations of set theory relate to this interpretation of events. Let \( A, B \subset \Omega \) be two events. Then the union of \( A \) and \( B \), denoted \( A \cup B \), is the event corresponding to the outcomes that are either in \( A \) or in \( B \). The intersection of \( A \) and \( B \), denoted \( A \cap B \), is the event corresponding to the outcomes that are both in \( A \) and in \( B \). The complement of \( A \), denoted \( A^c \), is the set of all outcomes in \( \Omega \) which are not in \( A \). Notice that \( A \cap A^c = \emptyset \), the empty set, and that \( A \cup A^c = \Omega \). In general, if \( A \cap B = \emptyset \) for two events \( A \) and \( B \), then we say the events are **disjoint**.

We are now ready to define a probability on the set of outcomes \( \Omega \).

**Definition 7.1.** Let \( \Omega \) be the sample space of an experiment. To each event \( A \subset \Omega \), we assign a number \( p(A) \), called the **probability** of \( A \). Roughly speaking, \( p(A) \) measures the percent of occurrences where the outcome of the experiment lies in \( A \) when the experiment is performed more and more times.

The probability function \( p \) satisfies some conditions:

1. \( 0 \leq p(A) \leq 1 \) for each subset \( A \subset \Omega \).
2. \( p(\Omega) = 1 \) and \( p(\emptyset) = 0 \).
3. For \( A, B \subset \Omega \), we have \( p(A \cup B) = p(A) + p(B) \) if \( A \cap B = \emptyset \).

In words, the first condition is saying every event will occur anywhere from 0% to 100% of the time. The requirement that \( p(\Omega) = 1 \) is saying that there is a 100% chance that the outcome will lie in the set of all outcomes (!). The requirement that \( p(\emptyset) = 0 \) is saying that there is no chance that no outcome occurs. The third condition is the crucial one: it is saying that the chance that an outcome lies in \( A \) or \( B \) is equal to the chance it lies in \( A \) plus the chance it lies in \( B \), provided that \( A \) and \( B \) are disjoint events.

The third condition can be replaced with the more general requirement

\[ p(A \cup B) = p(A) + p(B) - p(A \cap B), \]
which translates to: *the likelihood that the outcome is in $A$ or $B$ equals the likelihood that the outcome is in $A$ plus the likelihood that the outcome is in $B$ minus the likelihood that the outcome is in both $A$ and $B$, since that event has been accounted for twice.*

Another important identity that follows from the second and third conditions is

$$p(A^c) + p(A) = 1$$

for any event $A$. This reflects the fact that 100% of the time an outcome will either be in $A$ or it won’t.

If $\Omega$ is a finite set, then specifying $p$ on each outcome is enough to determine the value of $p$ on any combination of outcomes by the third condition of the definition. In the example of rolling the pair of dice, let $A$ be the event where the the blue die is 2. Then

$$A = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\}.$$  

To compute $p(A)$ note that the event $\{(2, 6)\}$ and the event

$$A' = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5)\}$$

are disjoint. Therefore,

$$p(A) = p(\{(2, 6)\}) + p(A') = \frac{1}{36} + p(A').$$

Next, notice that the events $\{(2, 5)\}$ and $A'' = \{(2, 1), (2, 2), (2, 3), (2, 4)\}$ are disjoint and their union is $A'$. Thus,

$$p(A') = \frac{1}{36} + p(A''),$$

and so

$$p(A) = \frac{2}{36} + p(A'').$$

We can repeat this line of reasoning several more times, to conclude that $p(A) = \frac{6}{36} = \frac{1}{6}$. Indeed, a similar argument shows that since all outcomes of the experiment of rolling two dice are equally likely (with probability 1/36), the probability of an event $B$ is equal to the number of elements in $B$ times 1/36. For example, let $B$ be the event where the sum of the two dice is 8. Then

$$B = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}.$$  

There are 5 outcomes in $B$ and so $p(B) = \frac{5}{36}$. 

Example 7.2. What is the probability that either the blue die is 2 or that the sum of the two dice is 8?

Solution: Let $A$ be the event where the blue die is 2 and $B$ be the event where the two dice add to 8. The event $A \cup B$ represents the set of outcomes that are either in $A$ or in $B$. The question is thus asking us to compute $P(A \cup B)$. Notice that $A$ and $B$ have a nonempty intersection: $A \cap B = \{(2,6)\}$. Therefore by the modified formulation of condition 3,

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = 6/36 + 5/36 - 1/36 = \frac{10}{36}.$$ 

Alternatively, we could have listed the number of elements in $A \cup B$ and seen that are 10 elements.

7.3 Random variables

In this section, we discuss random variables.

Definition 7.3. A random variable $X$ on a sample space $\Omega$ is a function from $\Omega$ to the set of real numbers $\mathbb{R}$. Mathematically, we write this as

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

In other words, $X$ outputs a real number $X(\omega)$ for each outcome $\omega \in \Omega$.

Example 7.4. Consider an experiment where a fair coin is flipped two times. An example of a random variable $X$ is the number of heads flipped. Then

$$X(\{(H,H)\}) = 2, \quad X(\{(H,T)\}) = 1, \quad X(\{(T,H)\}) = 1, \quad X(\{(T,T)\}) = 0.$$ 

Example 7.5. Consider the sample space $\Omega$ consisting of the possible prices of a stock at some time $t$. Let $S_t$ denote this price. Then $X = S_t$ is a random variable on $\Omega$. Another random variable on $\Omega$ is

$$Y = \max (S_t - K, 0),$$

where $K$ is a positive number. Then $Y$ is a random variable which represents the amount that a call option would be worth if $t$ is the time that the option expires.
Example 7.6. Consider the sample $\Omega$ consisting of the possible annual returns of an investment one year from now. Then $\Omega$ consists of all possible percentages; for example, $-10\% \in \Omega$ and $12.5\% \in \Omega$. Suppose the 1-year risk-free rate is 6\%. An example of a random variable on $\Omega$ is the random variable $X$ representing the amount that this investment beats the risk-free rate.

7.4 Expectation for a finite sample space

Let $\Omega$ be a sample space equipped with a probability $p$. Let $X$ be a random variable on $\Omega$. Assume that $\Omega$ is a finite set; this is an assumption we will make for the remainder of the chapter. We wish to know, on average, what the value of $X(\omega)$ is if the likelihood of an outcome $\omega$ occurring is $p(\omega)$. This average is called the expectation or expected value of the random variable $X$.

Definition 7.7. Assume that $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ is a finite set, then the expectation of $X$, denoted $E[X]$ is

$$E[X] = \sum_{i=1}^{n} p(\omega_i)X(\omega_i).$$

Notice that this is just a weighted average of all the possible values of $X$, where the weights are the probabilities of each outcome.

Example 7.8. Compute the expected value of the random variable from Example 7.4.

Solution: Each of the four outcomes is equally likely. So

$$E[X] = \left(\frac{1}{4}\right)(2) + \left(\frac{1}{4}\right)(1) + \left(\frac{1}{4}\right)(1) + \left(\frac{1}{4}\right)(0) = 1.$$

In other words, on average we expect to see 1 head when we flip a coin twice, which makes intuitive sense.

Example 7.9. Consider a very simple model for a stock. Suppose that each day the stock either goes up by 10\% or down by 8\%. If the probability that the stock goes up is 40\% and down is 60\%, what is the expected daily return of the stock?
7.4. EXPECTATION FOR A FINITE SAMPLE SPACE

Solution: The expected return is \( .4(10) + .6(-8) = - .8 \). That is, we expect the stock to lose \(-0.8\%\) on average each day. After 1 week, we would expect to lose \(4\%\) on our initial investment.

If \(X\) and \(Y\) are two random variables, then \(X + Y\) is also a random variable and so is \(aX\), where \(a \in \mathbb{R}\). The constant function \(a\) is also a random variable: it assigns to every outcome the number \(a\). The expectation satisfies the following properties:

1. \(E[X + Y] = E[X] + E[Y]\)
2. \(E[aX] = aE[X]\)
3. \(E[a] = a\).

Example 7.10. Consider the simple stock model from Example 7.9. Assume the spot price today is \$100. You own a bull spread expiring tomorrow, made from calls with strike prices 95 and 105. That is, you are long the call with strike 95 and short the call with strike 105. What is the expected payoff of the bull spread tomorrow?

Solution: Let \(S_1\) be the random variable representing the stock price tomorrow, then

\[
E[\text{payoff}] = E[\max(S_1 - 95, 0) - \max(S_1 - 105, 0)]
= E[\max(S_1 - 95, 0)] - E[\max(S_1 - 105, 0)].
\]

Note that \(\max(S_1 - 105, 0)\) is another random variable. Since there is a 40\% chance that \(S_1\) will be \((1 + 0.1) \times 100 = 110\) and a 60\% chance that \(S_1\) will be \((1 - 0.08) \times 100 = 92\), we compute

\[
E[\max(S_1 - 95, 0)] = 0.4 \times \max(110 - 95, 0) + 0.6 \times \max(92 - 95, 0)
= 0.4 \times 15 + 0.6 \times 0 = 6
\]

Similarly,

\[
E[\max(S_1 - 105, 0)] = 2.
\]

Thus the expected payoff is \(6 - 2 = 4\).

Later, we will define the expectation for an infinite sample space. The sum in Definition 7.7 will be replaced by an integral.
CHAPTER 7. PROBABILITY AND STATISTICS I

7.5 Variance and standard deviation

We now define the variance and standard deviation of a random variable $X$. It relies only on the definition of expectation, so when we extend the definition of $E[X]$ to infinite sample spaces, all of the material in this section and the next one will still be valid. It will improve notation to denote the expectation $E[X]$ by $\bar{X}$.

**Definition 7.11.** The variance of $X$, denoted $\text{Var}(X)$, is

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

Using the properties of expectation, it is possible to show that

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

We leave this as an exercise.

Since the random variable $(X - \bar{X})^2$ is always nonnegative, its expectation is also nonnegative. That is, $\text{Var}(X)$ is a nonnegative number. This allows us to define:

**Definition 7.12.** The standard deviation of $X$, denoted $\text{SD}(X)$, is

$$\text{SD}(X) = \sqrt{\text{Var}(X)}.$$

The variance and standard deviations are measures of how much the random variable wanders from the expected value $\bar{X}$. One advantage of standard deviation over the variance is that it has the same units as the underlying random variable. So if $X$ measures the price of a stock in dollars, both the expectation $E[X]$ and the standard deviation $\text{SD}(X)$ are given in dollars. By contrast, the variance is given in dollars-squared, which is not a very natural set of units.

**Example 7.13.** Compute the variation and standard deviation of the random variable in Example 7.4.

**Solution:** The variance $\text{Var}(X)$ is given by

$$\frac{1}{4}(2 - 1)^2 + \frac{1}{4}(1 - 1)^2 + \frac{1}{4}(1 - 1)^2 + \frac{1}{4}(0 - 1)^2 = \frac{1}{2}.$$ 

Then $\text{SD}(X) = \frac{1}{\sqrt{2}}$.

Both the variance and standard deviation are measures of risk. The next example illustrates why.
Example 7.14. A new game-show called Take or No-Take offers contestants the chance either to open one of three doors or take a fixed amount of money offered by a banker. Behind one door is $1, behind another door is $1,000, and behind the last door is $10,000. The amounts are randomly placed behind the doors. If you are a contestant and you are offered $5,000 guaranteed by the banker or the chance to open one door, what would you do? What if the banker offers $3,000?

Solution: The expected gain from opening one door is
\[
\frac{1}{3}(1) + \frac{1}{3}(1000) + \frac{1}{3}(10000) = 3667
\]
dollars. The standard deviation from this scenario is 4496.6 dollars. On the other hand, the banker’s offers have zero standard deviation, i.e., no risk. If you are offered $5,000, you should probably take it since you can walk away with an amount higher than the expected amount of opening a door, and the latter scenario carries much higher risk. On the other hand, if you are offered $3,000, this is less than the expected gain from opening a door. Some people would be happy to accept the lower amount, with no risk. Others, would be willing to take a chance and hold out for an amount closer to the expected value. This trade-off between expected gain and risk is the basis of mean-variance portfolio theory, the topic of the next chapter.

7.6 Covariance and correlation

The covariance and correlation measure how closely two random variables are related. Let \( X \) and \( Y \) be two random variables. Suppose we want to compute \( E[X + Y] \). That is easy; the answer is \( E[X] + E[Y] \), or using our short-hand, it is \( \bar{X} + \bar{Y} \). On the other hand, what if we want to compute \( Var(X + Y) \)? This is more complicated:

\[
Var(X+Y) = E[((X+Y)-(\bar{X}+\bar{Y}))^2] = Var(X) + Var(Y) + 2E[(X-\bar{X})(Y-\bar{Y})],
\]
after some computations. The obstacle to having the variance of \( X + Y \) be equal to the variance of \( X \) plus the variance of \( Y \) is given a name.

Definition 7.15. The \textit{covariance} of \( X \) and \( Y \), denoted \( Cov(X,Y) \), is

\[
Cov(X,Y) = E[(X-\bar{X})(Y-\bar{Y})].
\]
CHAPTER 7. PROBABILITY AND STATISTICS I

Notice that \( \text{Cov}(X, X) = \text{Var}(X) \). Roughly speaking the covariance is measuring to what extent \( Y \) differs from its expected value when \( X \) differs from its expected value. An equivalent way to write the covariance, using the properties of expectation, is \( E[XY] - \bar{X}\bar{Y} \).

The correlation is closely related to the covariance.

**Definition 7.16.** The correlation of \( X \) and \( Y \), denoted by \( \rho_{X,Y} \), is

\[
\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\text{SD}(X)\text{SD}(Y)}.
\]

The correlation turns out always to be a number between \(-1\) and 1. It can be shown that a correlation of 1 implies that \( X = aY + b \) for two numbers \( a \) and \( b \) where \( a > 0 \). A correlation of \(-1\) means that \( X = aY + b \) with \( a < 0 \). In these cases, \( X \) and \( Y \) are related exactly in a linear fashion\(^1\).

As the correlation moves away toward 0, the variable \( X \) has a weaker linear relationship with \( Y \). In other words, trying to approximate \( X \) by \( aY + b \) becomes worse and worse. When we reach a correlation of 0, the linear approximation is meaningless.

### 7.7 Different probabilities

In Chapters 9 and 11 we will need the notion of computing the expectation of a random variable using two different probabilities. Although changing probabilities is a simple idea in this context, it underlies the important concept of risk-neutral pricing in Chapters 9 and 11, and thus warrants its own section.

**Example 7.17.** Consider the same bull spread from Example 7.10. The underlying stock's spot price is $100, only now it has a 50% chance of going up 10% and a 50% chance of going down 8% by tomorrow. Compute the expected payoff of the spread.

**Solution:**

\[
E[\text{payoff}] = E[\max(S_t - 95, 0)] - E[\max(S_t - 105, 0)]
\]

\[
= (0.5 \times \max(110 - 95, 0) + 0.5 \times \max(92 - 95, 0))
- (0.5 \times \max(110 - 105, 0) + 0.5 \times \max(92 - 105, 0))
= 0.5 \times 15 - 0.5 \times 5 = 5.
\]

\(^1\)Technically, we should call this affine.
So the same random variable (the bull spread payoff) has a different expectation. The higher expectation is not surprising. A derivative which is bullish on the stock should expect to pay more if the stock is now more likely to go up than before.

We could also compute the variance of the payoff under this new probability, however, Chapters 9 and 11 are concerned only with expectation computations under different probabilities.

### 7.8 Gathering statistics on data sets

Suppose we have an experiment with possible outcomes Ω and a probability \( p \) on Ω. Suppose that we can run the experiment repeatedly. Let \( X \) be a random variable on Ω. In the real world, we may not actually know \( p \), but we can seek to estimate the expectation and standard deviation of \( X \) by running the experiment many times and recording the value of \( X \) on the outcomes. Let \( x_1, \ldots, x_n \) be the values of \( X \) on the outcomes of running the experiment \( n \) times. If we make the assumption that each run of the experiment is independent of the other runs (and some other typical assumptions about \( X \)), then the best estimate for \( E[X] \) is just the usual average

\[
\bar{x} = \frac{\sum x_i}{n},
\]

and a good estimate for \( SD(X) \) is

\[
s_x = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}.
\]

Notice that these coincide with the definition of the expectation and standard deviation of a random variable on a finite sample space of size \( n \) where every outcome is equally likely. In many cases, people also like another estimate for \( SD(X) \), namely:

\[
\sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}.
\]

We will use the former; the two estimates are extremely close when there is a large number of data points. Similarly, we can estimate the correlation between two random variables by calculating the correlation of the respective data sets of the two random variables. That is, if the outcomes for \( Y \) are
\{y_1, \ldots, y_n\}$ corresponding to the outcomes for $X$, then we can estimate $\rho_{X,Y}$ by

$$\frac{1}{n} \left( \sum x_i y_i \right) - \bar{x} \bar{y} \overline{s_x s_y}.$$

**Example 7.18.** The returns of stock $A$ over a week are 2, 3, $-5$, $-4$, 3 (as a percentage). The returns of stock $B$ are 4, 1, 0, $-4$, $-1$ over the same time period. What is the correlation of the returns over this period?

**Solution:** We have $\bar{x} = -.2$ and $\bar{y} = .4$ and $s_x = 3.554$ and $s_y = 2.6077$. The correlation is

$$\frac{(2)(4)+(3)(1)+(-5)(0)+(-4)(-3)+(3)(-1)}{5} - (-.2)(.4) \overline{3.554 \cdot 2.6077} = .519.$$
7.9 Problems

1. Let \( \Omega \) be a sample space and \( p \) a probability on \( \Omega \).

   (a) Using the definition of probability show that if three events
   
   \[ A, B, C \subset \Omega \]

   are mutually disjoint, i.e., that \( A \cap B = \emptyset \), \( B \cap C = \emptyset \), and \( A \cap C = \emptyset \), then
   
   \[ p(A \cup B \cup C) = p(A) + p(B) + p(C). \]

   (b) Given two sets \( C \subset A \), we write \( A - C \) for the set of elements in \( A \) that are not in \( C \). Notice that in this case that \( A = C \cup (A - C) \) and \( C \cap (A - C) = \emptyset \) and use this to establish an identity involving \( p(A) \), \( p(C) \) and \( p(A - C) \). What does your identity say when \( A = \Omega \)?

   (c) Establish the identity mentioned in Section 7.2
   
   \[ p(A \cup B) = p(A) + p(B) - p(A \cap B) \]

   for any two events \( A \) and \( B \). (Hint: consider the set \( C = A \cap B \)
   and also the two sets \( A - C \) and \( B - C \).)

2. Consider the experiment where a nickel and a dime are flipped and a six-sided die is rolled.

   (a) How many outcomes are there in this experiment? How many outcomes are there where the nickel is a head?

   (b) Assuming each outcome is equally likely, what is the probability of the event \( A \) where the nickel is a head? of the event \( B \) where the dime is a head? of \( A \cap B \)? of \( A \cup B \)? of \( A^c \)?

3. Consider the experiment where a six-sided die is rolled. Suppose that the probability that the number \( i \) shows up is \( i/21 \).

   (a) What is the probability that the die shows a number greater than or equal to 4?
(b) Let $X$ be the random variable equal to the number on the die. Let $Y = X^3$. What is $E[X]$? What is $E[X - 3]$? What is $E[X^2]$? What is $E[Y]$?

4. Show that $E[(X - \bar{X})^2] = E[X^2] - \bar{X}^2$, using just the three properties of expectation at the end of Section 7.4. Either of these formulas can be taken as the definition for the variance of $X$.

5. Show that $Var(aX + b) = a^2Var(X)$, using the properties of expectation.

6. Show that if $X = aY + b$, then $\rho_{X,Y} = 1$ if $a > 0$ and $\rho_{X,Y} = -1$ if $a < 0$. What happens if $a = 0$?

7. Consider the model of a stock where it either goes up by 7% or down by 4% each month. Let $p_{up}$ be the probability that stock goes up and $p_{down}$ the probability that stock goes down.

   (a) Suppose that $p_{up} = 52\%$ and $p_{down} = 48\%$. What is the expected monthly return of the stock?

   (b) If the 1-month risk-free rate is 5%, what probabilities $p_{up}$ and $p_{down}$ lead to an expected monthly return equal to the risk-free rate?

8. Let the percent daily returns of IBM stock over a week equal $-1, -3, 0, 1, -2$ and the percent returns of JNJ stock be $0.5, 1, -1, -2, 3$. Calculate the correlation of these returns.
Chapter 8

Mean-Variance Portfolio Theory

This chapter discusses the theory of portfolio selection that relies on the expected return of each investment and the variance of the return. In addition, the correlation between investments plays a central role. The theory has its origins in the work of Harry Markowitz, an economist at UCSD.

8.1 Introduction

The main idea behind mean-variance portfolio theory (MV theory) is that investors will choose between two investments based upon the expected return of each investment and the standard deviation (or variance) of the return.

Consider two investments $A$ and $B$. Let $A$ and $B$ also denote the random variable of the returns of each investment over a future time period. Typically the time period would be a day, a month, or a year, but any time horizon will work. As in the previous chapter, let $\bar{A}$ and $\sigma_A$ be the expected return and standard deviation of the return for investment $A$, and $\bar{B}$ and $\sigma_B$ be the expected return and standard deviation of the return for investment $B$. Then mean-variance theory posits that an investor will:

1. Choose investment $A$ over $B$ if $\bar{A} = \bar{B}$ and $\sigma_A < \sigma_B$.

2. Choose investment $A$ over $B$ if $\sigma_A = \sigma_B$ and $\bar{A} > \bar{B}$.

In other words, if the expected returns are the same, the investor will choose the investment with a smaller standard deviation, i.e., with less risk. And if
CHAPTER 8. MEAN-VARIANCE PORTFOLIO THEORY

the amount of risk is the same, the investor will choose the investment with the greatest expected return.

The goal of MV theory is to find those investments for which there does not exist any investment that outperforms them according to the above two conditions. Such investments are said to lie on the efficient frontier. The reason for this terminology is that it is convenient to represent investments in mean-standard deviation space, which is a graph where the x-axis is labeled by standard deviation and the y-axis is labeled by the expected return. Each investment A is plotted as a point \((\sigma_A, \bar{A})\). The efficient frontier turns out to be a piece of a convex curve.

Of course, in order to carry out the determination of the efficient frontier, we must know the expected return and standard deviation of all possible investments. It is not possible to know these values definitively, but we can use historical data to estimate them. More sophisticated modeling techniques can also be used. For the purposes of these notes, we assume that we know \((\sigma_A, \bar{A})\) for each investment A.

8.2 A combination of two investments

We want to build a portfolio using two investments A and B. Let \(x_A\) be the amount allocated to investment A and \(x_B\) be the amount allocated to B. Since we assume our investment capital is constant, we should assume

\[x_A + x_B = 1.\]

In other words, \(x_A\) and \(x_B\) are the fraction of our capital allocated to each investment. If there are no short-sales allowed, then \(0 \leq x_A \leq 1\) and \(0 \leq x_B \leq 1\). If short-sales are permitted, then these constraints are lifted. The theory is easiest to analyze when short-sales are allowed. However, in practice, there are usually constraints on short-selling. For example, a mutual fund cannot engage in short-selling. Another popular investment strategy limits short-selling to 30% of capital; this strategy is known as 130/30 since the portfolio is long 130% of capital and short 30%. For the case of two investments, this strategy would imply

\[-.3 \leq x_A, x_B \leq 1.3.\]

Let \(P\) denote the new portfolio consisting of \(x_A\) of A and \(x_B = 1 - x_A\) of
8.2. A COMBINATION OF TWO INVESTMENTS

B. The expected return of $P$ is given by

$$\bar{P} = x_A \bar{A} + (1 - x_A) \bar{B},$$

by the properties of expectation from the previous chapter. The standard deviation of the returns of $P$ is more complicated. We have

$$\sigma^2_P = x_A^2 \sigma_A^2 + (1 - x_A)^2 \sigma_B^2 + 2x_A(1 - x_A)\sigma_A \sigma_B \rho_{A,B}$$

(8.1)

where $\rho_{A,B}$ is the correlation between $A$ and $B$. This follows from the definition of correlation and the properties of expectation (see Exercise 5 in Chapter 7). The first question to answer is what curve does $P$ trace out in mean-standard deviation space as we vary $x_A$. We re-write Equation 8.1 gathering like terms:

$$\sigma^2_P = x_A^2 (\sigma_A^2 + \sigma_B^2 - 2\sigma_A \sigma_B \rho_{A,B}) - 2x_A(\sigma_B (\sigma_B - \sigma_A \rho_{A,B})) + \sigma_B^2. $$

Let

$$u = \sigma_A^2 + \sigma_B^2 - 2\sigma_A \sigma_B \rho_{A,B}$$

and

$$v = \sigma_B (\sigma_B - \sigma_A \rho_{A,B}).$$

A computation shows that $u$ is the variance of the random variable $A - B$. Hence $u \geq 0$. Completing the square, we get

$$\sigma^2_P = u(x_A - \frac{v}{u})^2 + (\sigma_B^2 - \frac{v^2}{u}).$$

(8.2)

In other words,

$$\sigma^2_P - u(x_A - \frac{v}{u})^2$$

is a constant. Since $u \geq 0$, this shows that the graph of $(\sigma_P, x_A)$ will be a hyperbola. Now, since $x_A$ and $\bar{P}$ are related by the linear equation

$$x_A = \frac{\bar{P} - \bar{B}}{A - B},$$

it follows that the graph of $(\sigma_P, \bar{P})$ is also a hyperbola. The hyperbola opens to the right with asymptotes equal to the lines

$$\sigma_P = \pm \sqrt{u}[\left(\frac{1}{A - B}\right)\bar{P} - (\frac{\bar{B}}{A - B} + \frac{v}{u})].$$
The hyperbola will pass through a point of minimum standard deviation. That occurs when
\[ x_A = \frac{v}{u}, \]
since \( u \geq 0 \). The minimum standard deviation will be
\[ \sigma_P = \sqrt{\sigma_B^2 - \frac{v^2}{u}}, \]
noting that the term inside the square root is always nonnegative. If we fix the convention that \( \sigma_A > \sigma_B \) (and so \( \bar{A} > \bar{B} \) ), then the portfolios \( P \) that lie on the efficient frontier correspond to \( x_A \geq \frac{v}{u} \).

**Formula 15.** Let \( A \) and \( B \) represent two investments with \( \sigma_A > \sigma_B \) and \( \bar{A} > \bar{B} \). Let
\[ P = x_A A + (1 - x_A) B \]
be a portfolio built from \( A \) and \( B \). If short-selling is allowed, then the efficient frontier through \( A \) and \( B \) is the set of portfolios \( P \) where \( x_A \geq \frac{v}{u} \). The graph of the frontier in mean-standard deviation space is the upper part of a hyperbola that opens to the right. The minimal variance occurs when \( x_A = \frac{v}{u} \).

There are two special cases worth pointing out:

1. If \( \sigma_B = 0 \) so that \( B \) is a risk-free asset, then \( v = 0 \) and the hyperbola degenerates into a line. The frontier is a line connecting \( B \) with \( A \). The minimal variance occurs when \( x_A = 0 \), meaning that all of the portfolio is invested in \( B \), which makes intuitive sense since the variance of \( B \) is zero.

2. If \( \rho_{A,B} = \pm 1 \), then \( \sigma_B^2 = \frac{v^2}{u} \). Again, the frontier degenerates from a hyperbola and it lies on one of the asymptotes of the hyperbola.

### 8.3 Inserting a risk-free asset

If we now assume that in the set of all investments, there is a risk-free investment, then we will show that the efficient frontier is a line and there is a single risky portfolio that we should invest in.
In the previous section, we noted that if the investment $B$ is risk-free, then $\sigma_B = 0$ and $v = 0$. Also, $u = \sigma_A^2$. Therefore the equation of the efficient frontier is given by

$$\sigma_P = \sigma_A \left[ \frac{1}{A-B} \bar{P} - \left( \frac{B}{A-B} \right) \right].$$

Writing this as a function of $\sigma_P$ gives

$$\bar{P} = \frac{\bar{A} - \bar{B}}{\sigma_A} \sigma_P + \bar{B}.$$  

This is the equation for line with $y$-intercept $\bar{B}$ and slope

$$m = \frac{\bar{A} - \bar{B}}{\sigma_A}.$$  

Let $r$ denote the risk-free rate for the time period of the investment horizon. Then if we can borrow and lend at $r$, this provides a risk-free investment. Any risky investment can be combined with this risk-free investment to produce efficient portfolios along the line:

$$\bar{P} = \frac{\bar{A} - r}{\sigma_A} \sigma_P + r.$$  

There is an important consequence of this simple observation.

Let

$$s = \frac{\bar{A} - r}{\sigma_A},$$

the slope of the efficient frontier. This is amount of expected return that the risky investment provides above the risk-free rate divided by its risk. It is called the information ratio of $A$ or the Sharpe ratio of $A$.

The analysis of the previous section shows that the efficient frontier of two risky assets is a hyperbola. As we add more and more risky assets, the frontier might shift, but it is still a piece of a hyperbola. Now, when we add a risk-free asset, the efficient frontier becomes a line, which is exactly tangent to the frontier consisting only of risky assets. In other words, the new frontier touches the risky frontier at exactly one point. This point is the investment $A$ on the risky frontier with maximal Sharpe ratio $s$. Assuming no constraints on short-selling, it is possible to compute $A$ precisely using linear algebra, but this would take us a bit beyond the scope of these notes.
**Chapter 8. Mean-Variance Portfolio Theory**

**Formula 16.** In the presence of a risk-free asset, the efficient frontier is a line passing through a unique portfolio $P$ consisting only of risky assets. This risky portfolio is the portfolio with largest Sharpe ratio $s = \frac{p - r}{\sigma_P}$.

Sharpe and others took this observation as the basis of the Capital Asset Pricing Model (CAPM). It says that if all investors assign the same $\bar{A}$ and $\sigma_A$ to each investment $A$, then there is a preferred portfolio $P$ consisting of risky assets, the one with largest Sharpe ratio. All investors will hold a portion of their investments in $P$ and a portion in the risk-free asset; in other words, each investor places his optimal portfolio somewhere on the efficient frontier joining the risk-free asset and $P$, depending on how much risk the investor wants to take. Each investor has a utility function which guides them in this decision. Nevertheless, since all investors invest in $P$, that implies that $P$ is exactly the market portfolio. Working backwards, it says that the optimal portfolio of risky assets is the one that assigns a fraction of capital in each investment according to the fraction of the market value of the investment.

### 8.4 Problems

1. Show that if $\rho_{A,B} = 1$, then $\sigma_B^2 = \frac{\nu^2}{w}$ in Equation 8.2. Conclude that the efficient frontier lies on a line connecting $A$ and $B$ in mean-standard deviation space. Show that the minimal variance portfolio occurs when the portfolio is short $A$ (that is, $x_A < 0$), where we are keeping the assumption that $\sigma_A > \sigma_B$. What is the minimal variance?

2. Show that if $\rho_{A,B} = -1$, then again $\sigma_B^2 = \frac{\nu^2}{w}$ in Equation 8.2. Show that the minimal variance portfolio occurs when

$$x_A = \frac{\sigma_B}{\sigma_A + \sigma_B}.$$ 

Explain why this number is between 0 and 1. Conclude that the efficient frontier lies on a line emanating from a risk-free portfolio to $A$ in mean-standard deviation space. What is the expected return of the minimal variance portfolio? Again, keep the assumption that $\sigma_A > \sigma_B$.

3. Suppose you can invest in two risky investments $A$ and $B$ with $\sigma_B = 3$ and $\bar{B} = 4$ and with $\sigma_A = 7$ and $\bar{A} = 9$ (these are percentages). The correlation between $A$ and $B$ is 0.4. What is the portfolio of minimal variance containing these two assets? What is this minimal variance?
4. Continuing with the previous problem, suppose that the risk-free rate is 3%. What is the portfolio built from $A$ and $B$ with the largest Sharpe ratio? (Hint: set up the formula for the Sharpe ratio as a function of $x_A$ and take the derivative).

5. Continuing with the previous problem, an investor seeks an investment with a return of 6%. According to CAPM, what is the amount of risk they should expect from such an investment?
Chapter 9

Binomial trees and risk-neutral pricing

The goal of this chapter is to value options and other derivatives, assuming simple models for the stock price, known as binomial trees. Binomial tree models have three important attributes: they illustrate simple but fundamental ideas behind derivative pricing; they provide “discrete” versions of the so-called log-normal model used to derive the celebrated Black-Scholes formula, which we discuss in a later chapter; and, they can be used to approximate the prices of complex derivatives, such as American options, which have no known pricing formula in the log-normal model.

We begin in Section 9.1 with the most basic of trees, known as the one-step tree. We generalize this to the n-step tree in Section 9.3. In Sections 9.4 and 9.5, we describe how to change the models when the options are American, or when the stocks have dividends or carrying costs.

In Section 9.2, we demonstrate how trees are an example of the important concept of risk-neutral pricing. This crucial idea states that the price of the option (like the forward price) depends on the risk-free rate and the spot price of the stock (and some other parameters), but it does not depend on the expected rate of return of the stock. Most financial models, even complicated ones, possess this feature.

A crucial step in justifying the option price formulas is the application of the fundamental principal from Chapter 6: two portfolios with the same guaranteed payoff at time $T$ must have the same price at the initial time.
9.1 One step model

Previously, the stock could take on any non-negative value $S_t$. In reality, the stock price has only incremental values, in units of cents, like $S_t = \$31.67$. The one-step binomial tree simplifies things further with the following assumptions:

- There are only two times: the present time $t = 0$ and the expiration date of the option at $t = T$.

- If the spot price of the stock is $S_0$, then the stock price at time $T$ is either $S_0u$ or $S_0d$. We assume $u > d$, and call the $u$ and $d$ factors the up- and down-factors, respectively (even though in some models $d$ may be greater than 1).

These are not very realistic assumptions, so in Section 9.3 we increase the number of time steps and possible stock values. Such an increase leads to a better approximation of reality, but makes the model more difficult to implement.

**Example 9.1.** The spot price of Cisco stock (symbol: CSCO) is $S_0 = \$20$. In three months, the stock price $S_{3/12}$ will be either $\$18$ or $\$22$. Assume the risk-free rate is 12% per annum with continuous compounding. Compute the price $c$ of a 3-month European call on one share with strike $\$21$.

**Solution:** Note that we did not mention the probability of the stock going up or down. So we cannot determine the expected return on the stock.

Consider the following two portfolios:

- **Portfolio A:** long $N$ shares of CSCO and short one 3-month European call option on CSCO with strike $\$21$.

- **Portfolio B:** long $P$ dollars in cash.

We want to choose $N$ and $P$ such that the two portfolios have the same payoff at the expiration date $T = 3/12$. This may not seem possible since Portfolio $B$ is risk-free: it is guaranteed to be worth $Pe^{0.12\times3/12}$ in three months. In contrast, Portfolio $A$ seems risky: it can take on 2 different values. Nonetheless, such a choice of $N$ and $P$ can be made.
Let us look at Portfolio A more closely. Should the stock go up, the \( N \) shares will be worth \( 22N \), and the call will be exercised by the holder of the long position in the call, making the call worth
\[-(22 - 21) = -1\]
in Portfolio A. Therefore the total value of Portfolio A will be worth \( 22N - 1 \). On the other hand, if the stock price goes down to 18, the call will expire worthless, and so the portfolio will be worth \( 18N \). We can choose \( N \) to make these two possible values equal:

\[22N - 1 = 18N, \quad \text{or} \quad N = 0.25.\]

In both cases Portfolio A will be worth \( 18 \times 0.25 = 22 \times 0.25 - 1 = 4.50 \) at \( T = 3/12 \).

Now our task seems more achievable. Choose Portfolio B so that it is also worth $4.50 at time \( T = 3/12 \). This means Portfolio B contains

\[4.37 = 4.50e^{-0.12 \times 3/12}\]
in cash today. Since both portfolios are guaranteed to be worth $4.50 at time \( T = 3/12 \), the fundamental principle says they are worth the same today. Since Portfolio A consists of \( N = 0.25 \) shares of stock and a short position in the call option, we know that

\[\text{value of Portfolio A today} = 20 \times 0.25 - c,\]

where \( c \) is the price of the call option. The value of Portfolio B today is $4.37. Setting these values equal to each other, we get

\[20 \times 0.25 - c = 4.37,\]

and thus \( c = 0.63 \).

It may seem surprising that we did not need to know the probability of the stock going up or down to compute the price of the call. Suppose the up and down probabilities for Cisco stock were 90% and 10%, respectively. Now suppose that another stock, for example PepsiCo (symbol: PEP), was also trading at $20 today, and had a 15% chance of rising to $22 in 3 months and an 85% chance of dropping to $18 in 3 months. The above example
would imply that the Cisco and PepsiCo European calls with strike $21 and expiration in 3 months would both cost $0.63. But certainly the Cisco option is more appealing. This apparent paradox is resolved by inaccurate choices in the binomial trees. Who would invest $20 today in PepsiCo for 85% chance of it going to $18, when that same $20 invested in Cisco would have a 90% chance of going to $22? Not many, and so the (relative) lack of demand for a $20 PepsiCo stock would drive the price of PepsiCo stock to be less than Cisco.

Let us extend the discussion of the single call option in the above example to more general derivatives.

Example 9.2. Google (symbol: GOOG) has a spot price $S_0 = 400$ which will change to $S_0u = 450$ or $S_0d = 410$ in six-months, $T = 6/12$. Consider a derivative made of the following financial securities

- One six-month European put on Google with strike $420$.
- One six-month straddle on Google with strike $425$. Recall that a straddle consists of a long position in a call and a long position in a put, both with the same strike price and the same expiration (see Chapter 6).
- Two shares of Google stock.

Compute the value $h_u$ of this derivative in 6 months should the stock price of Google in 6 months be $450$. Compute the value $h_d$ of this derivative in 6 months should the stock price of Google in 6 months be $410$.

Solution: Let $S_T$ denote the price of Google in six months. Then at time $T$ the put will be worth $\max(420 - S_T, 0)$, the two shares of stock will be worth $2 \times S_T$, and the straddle will be worth

$$\max(S_T - 425, 0) + \max(425 - S_T, 0) = |S_T - 425|.$$ 

So if $S_T = S_0u = 450$, then

$$h_u = \max(420 - 450, 0) + 2 \times 450 + |450 - 425| = 925.$$ 

If $S_T = S_0d = 410$, then

$$h_d = \max(420 - 410, 0) + 2 \times 410 + |410 - 425| = 845.$$
9.1. ONE STEP MODEL

Figure 9.1: A one-step tree. On the left, at the present time 0, there is one node with the spot price of the stock $S_0$ and the (still unknown) current value of the derivative $h$. On the right, at the expiration time $T$, there are two nodes. The upper one is realized should the stock price change by a factor of $u$, $S_T = S_0u$. The lower node is realized should the stock price change by a factor of $d$, $S_T = S_0d$. Next to each possible stock price is the corresponding value of the derivative, which we assume we know.

Next, we consider the most general situation where the stock has a spot price $S_0$ and the price will change to $S_0u$ or $S_0d$ at time $T$. Consider any derivative whose value at time $T$ we can compute, as in the above example. We label the derivative value $h_u$ if the stock changes to $S_0u$ and $h_d$ if the stock changes to $S_0d$. We wish to compute $h$, the current price of this derivative. See Figure 9.1.

Chose Portfolio $A$ to be long $N$ shares and short one such derivative.

\[
\text{value of Portfolio } A \text{ if stock goes up} = S_0uN - h_u \\
\text{value of Portfolio } A \text{ if stock goes down} = S_0dN - h_d
\]
To make Portfolio $A$ risk-free, we need to make the value of the portfolio the same whether the stock goes up or goes down:

$$S_0uN - h_u = S_0dN - h_d.$$  \hfill (9.1)

Solving for $N$, we get

$$N = \frac{h_u - h_d}{S_0u - S_0d}.$$  

With this value of $N$, Portfolio $A$ is guaranteed to be worth

$$S_0uN - h_u = S_0u\frac{h_u - h_d}{S_0u - S_0d} - h_u,$$

at time $T$, using the first expression for the value in Equation 9.1. Next, by the fundamental principal, the current value of Portfolio $A$ must be worth the present value of its guaranteed value at time $T$, namely

$$\text{current value of Portfolio } A = e^{-rT} \left( S_0u\frac{h_u - h_d}{S_0u - S_0d} - h_u \right).$$

On the other hand, we also know that the current value of Portfolio $A$ is

$$S_0N - h.$$

If we also substitute $N = \frac{h_u - h_d}{S_0u - S_0d}$ into this formula, we get

$$\text{current value of Portfolio } A = S_0\frac{h_u - h_d}{S_0u - S_0d} - h.$$  

Setting the two different expressions for the current value of Portfolio $A$ equal to each other, and solving for $h$, yields

$$h = S_0\frac{h_u - h_d}{S_0u - S_0d} + e^{-rT} \left( S_0u\frac{h_u - h_d}{S_0u - S_0d} - h_u \right).$$

After some algebra, we can rewrite the price of the derivative as

$$h = e^{-rT}(ph_u + (1 - p)h_d)$$

where we define

$$p = \frac{e^{rT} - d}{u - d}.$$  

We explain the benefit of doing this in the next section. To summarize,
9.2. RISK-NEUTRAL WORLD

**Formula 17.** In the one-step binomial tree model of a stock, if a derivative on the stock has a known value at \( t = T \) of \( h_u \) when the stock price moves to \( S_0 u \) and a known value of \( h_d \) when the stock price moves to \( S_0 d \). Then the value \( h \) of the derivative at \( t = 0 \) is given by the formula

\[
h = e^{-rT}(ph_u + (1-p)h_d),
\]

where

\[
p = \frac{e^{rT} - d}{u - d}.
\]

**Example 9.3.** Find the current value of the derivative in Example 9.2 if the 6-month risk-free rate is 7%.

*Solution:* We showed that \( h_u = 925 \) and \( h_d = 845 \). The down-factor is \( d = 410/400 = 1.025 \) and the up-factor is \( u = 450/400 = 1.125 \). We have

\[
p = e^{0.07 \times 0.5} - 1.025 \over 1.125 - 1.025 = 0.106,
\]

Thus

\[
h = e^{-0.07 \times 0.5}(0.106 \times 925 + 0.894 \times 845) = 824.12
\]

dollars.

9.2 Risk-neutral world

We now discuss one of the most important features in modern day finance: risk-neutral derivative pricing.

Let us revisit Example 9.1. Translate the numbers into the notation for the general case: \( u = 22/20 = 1.1, d = 18/20 = 0.9, r = 0.12, S_0 = 20, T = 0.25, h_u = 1, h_d = 0 \)

\[
p = e^{0.12 \times 0.25} - 0.9 \over 1.1 - 0.9 = 0.65, \quad 1 - p = 0.35.
\]

Suppose \( p \) represents the probability of the stock increasing to $22, and hence \( 1 - p \) represents the probability of the stock decreasing to $18. Then,
just by manipulating the numbers in the second line below we see that

\[
E[S_T] = pS_0u + (1 - p)S_0d
\]

\[
= 0.65 \times 22 + 0.35 \times 18 = 20.60 = 20e^{0.12\times0.25}
\]

\[
= S_0e^{rT}.
\]

So if \((p, 1 - p)\) represented the real world probability, then the expected (not guaranteed) growth rate of the stock would be the risk-free rate.

Is it reasonable to assume that in the real world, the risky stock is expected to grow at the risk-free rate? We look to a simpler more intuitive example for the answer. Suppose you just inherited one billion dollars, and you could either keep it, or else bet it on an immediate coin flip, with a 50% chance of losing it all and a 50% chance of ending with two billion dollars. Both strategies have an expected return of 0%: the risk-free strategy of holding the money has an (instantaneous) return of 0%, as does the risky gamble in that on average, you end up with what you started with. However you, like most other real world investors, would forego the gamble. This is because you are risk-averse. Note that this behavior is central to the mean-variance theory discussed in Chapter 8.

In a risk-neutral world, investors are indifferent to risk. Presented with the billion dollar coin flip, both choices are equally attractive to them. They expect the same return on all assets regardless of risk, which, as we will see below, is the risk-free rate. In particular, the \((p, 1 - p) = (0.65, 0.35)\) probability mentioned in the above example is not the real world probability, but the risk-neutral one.

Let us analyze mathematically the key feature of the risk-neutral world: that the expected returns on all assets, regardless of riskiness, equals the risk-free return. We denote the expectation (of risky assets) in this world by \(E_{RNW}[\cdot]\), and expectations in the real world by \(E_{RW}[\cdot]\). It is important to understand that these two worlds may have different probabilities assigned to the same random events (like a future stock price rising or falling). See Section 7.7.

Let \(B_t\) denote some cash at time \(t\), sitting in an account and earning the risk-free rate \(r\). Note that \(B_t\) is deterministic: we know exactly what the
balance $B_t$ will be at any future time $t$. Thus, in both worlds:

$$E_{RW}[B_T] = B_T = B_0e^{rT}, \quad E_{RNW}[B_T] = B_T = B_0e^{rT}$$

In the risk-neutral world any (risky) stock is expected to grow at this rate as well

$$E_{RNW}[S_T] = S_0e^{rT}.$$  

Whereas in the real world, the risky stock return $\mu$ is defined by

$$E_{RW}[S_T] = S_0e^{\mu T}.$$  

Most people believe that $\mu > r$. This means that stocks are expected (but not guaranteed) to grow at a rate greater than the risk-free one. This is the stock-analogue of the risk-aversity exhibited in the coin flip scenario. In order to invest in a risky asset, investors must be compensated for the higher risk.

Let $h_T$ denote the payoff function for a derivative on the stock, that is

$$h_T(S_0u) = h_u, \quad h_T(S_0d) = h_d.$$  

Translating Formula 17 into the language of this section, we see that the current price $h$ of the derivative is the discounted expected payoff of the derivative in the risk-neutral world:

$$h = e^{-rT}(ph_u + (1-p)h_d) = e^{-rT}E_{RNW}[h_T].$$

To re-emphasize the point, recall that for $(p, 1-p)$ to be the risk-neutral probability of the stock going (up, down), we require that $p$ solves the equation

$$pS_0u + (1-p)S_0d = S_0e^{rT} \quad (= E_{RNW}[S_T]).$$

And in our one-step model, this is true if and only if

$$p = \frac{e^{rT} - d}{u - d},$$

1Recall that if a random variable $Y$ is in fact deterministic, then $E[Y] = Y$, where on the left hand side of the equation $Y$ is the name of the random variable, and on the right hand side $Y$ is the value that it achieves no matter the outcome.

2Many gamblers on the other hand are risk prone. They exchange their risk-free cash for a risky gamble even when the expected return rate is negative.
as a quick calculation will show.

Applying this formula to our example
\[ c = e^{-0.12 \times 0.25} E_{RNW}[\max(S_{0.25} - 21, 0)] = e^{-0.12 \times 0.25} (0.65 \times 1 + 0.35 \times 0) = 0.63 \]
as we computed before. Recall that since \( S_{0.25} \) is a random variable, then so too is \( S_{0.25} - 21 \), and hence \( \max(S_{0.25} - 21, 0) \); thus, it makes sense to be able to take an expectation of \( \max(S_{0.25} - 21, 0) \).

The upshot of this discussion is the following: to compute the price of a derivative, go to the risk-neutral world, compute its expected payoff there, discount it, and you have today’s real world price. This idea is known as risk-neutral valuation. Although we only verified this for the one-step model, this principle works in a very broad-context, including all the ones we will discuss in this text (several time steps, continuous model, etc). In a more advanced math finance text, one can prove the following deep mathematical theorem

**Theorem 9.4.** If the market has no arbitrage opportunities, then the price of any derivative that can be replicated with traded assets such as stock and cash, is the discounted expected derivative payoff in the risk-neutral world.

We end with one remark. In order for \((p, 1 - p)\) to be a legitimate probability, we need non-negativity of both \( p \) and \( 1 - p \). This is true if and only if \( 0 \leq p \leq 1 \). From the definition of \( p \), this is equivalent to requiring

\[ d \leq e^{rT} \leq u. \]

These inequalities follow from the No-Arbitrage assumption, one of which we leave for the exercises. Note that we do not require \( d < 1 \), but we call \( d \) the down factor as it represents the lower branch of the tree in Figure 9.1.

### 9.3 More than one time step

We consider what happens as we partition time into more steps between the present, \( t = 0 \), and the expiration date \( t = T \). We describe in detail the two-step model, and provide a formula for the \( n \)-step model.

**Example 9.5.** Consider a European put option on Amazon.com Inc stock (symbol: AMZN) which expires in one year, with strike $52. Amazon’s spot
9.3. **MORE THAN ONE TIME STEP**

Figure 9.2: In this two-step tree, there are three nodes at the expiration $t = 1$, two nodes at the intermediary time $t = 0.5$, and one node at the present time $t = 0$. Each node is labeled with the associated stock price, as well as the (sometimes unknown) value of the put option at that time and that asset price.

The price $S_0$ is $50$. The risk-free rate $r$ is $5\%$. Assume the stock price can go up or down by $20\%$ every 6 months. Compute the price $h$ of the put using a two-step binomial tree.

**Solution:** There are two time steps, each of length $\Delta t = T/2 = 6/12 = 0.5$ years. The up and down factors for each step are $u = 1.2$ and $d = 0.8$. An important built-in feature of this tree is that it is recombining (see Figure 9.2):

$$S_0ud = S_0du$$

If the stock first goes up and then down, it reaches the same node if it were first to go down and then up. This may not seem so significant in the two-step model, but for the $n$-step model, a non-recombining tree, which could happen with the up-factors (or down-factors) differing at different time steps, could have $2^n$ instead of just $n + 1$ different possible final stock prices.
In Figure 9.2 above, $h_u$ and $h_d$ denote the two (not yet computed) values of the put at $t = 0.5$ and $h_{uu}, h_{ud} = h_{du}, h_{dd}$ denote the three possible (easily computed) values of the put at the expiration time $t = 1$.

We tackle this problem by computing the value of the put option at the different nodes, working backwards-in-time. We compute the payoffs

$$h_{uu} = \max(0, 52 - 72) = 0, \quad h_{ud} = \max(0, 52 - 48) = 4.$$ 

Consider those two nodes, branching off of the $h_u$ node, as forming part of their own one-step tree within the two-step tree. So we pretend the time is $t = 6/12$ and the spot price on Amazon stock is $S_{0u} = \$60$. What is a derivative worth, that in six months will pay $0$ if Amazon rises to $\$72$, and $\$4$ if Amazon drops to $\$48$? We have already considered this sort of problem in Section 9.1. Let

$$p = \frac{e^{rt} - d}{u - d} = \frac{e^{0.05 \times 0.5} - 0.8}{1.2 - 0.8} = 0.56, \quad 1 - p = 0.44.$$ 

Then our formula from Section 9.1 says the derivative is worth

$$h_u = e^{-0.05 \times 0.5} (0.56 \times 0 + 0.44 \times 4) = \$1.72.$$ 

Similarly, just below we have another one-time step sub-tree with derivative values at the nodes given by $h_{ud} = \$4$ is the stock rises and $h_{dd} = \$20$ if the stock drops. Applying similar reasoning as above, we compute

$$h_d = e^{-0.05 \times 0.5} (0.56 \times 4 + 0.44 \times 20) = \$10.77.$$ 

We apply this idea a third time, where now we suppose we are at today’s node with Amazon trading at $S_0 = \$50$.

Since we know $h_u$ and $h_d$, we can forget that they came from their own subtrees. What would we pay today for a derivative worth $\$1.72$ if Amazon
9.3. MORE THAN ONE TIME STEP

rises to $60 and $10.77 if Amazon drops to $40?  

\[
h = e^{-0.05 \times 0.5} (0.56 \times 1.72 + 0.44 \times 10.77) = $5.56.
\]

We rewrite the Amazon example in general notation to derive a general formula. The stock price is \( S_0 \), and will change by a factor of \( u \) or \( d \) over each of two time periods of length \( \Delta t \). Let \( r \) be the risk-free rate. Let \( h \) be the current value of a derivative paying \( h_{uu}, h_{ud}, h_{dd} \) at the three expiration date nodes.

If we embed the formulas for \( h_u \) and \( h_d \) in terms of \( h_{uu}, h_{ud}, h_{dd} \), into the formula for \( h \), we get that

\[
h = e^{-r \Delta t} (p h_u + (1-p) h_d)
\]

\[
= e^{-r \Delta t} \left( p \left( e^{-r \Delta t} (ph_{uu} + (1-p)h_{ud}) \right) + (1-p) \left( e^{-r \Delta t} (ph_{ud} + (1-p)h_{dd}) \right) \right)
\]

\[
= e^{-r \Delta t} \left( p^2 h_{uu} + 2p(1-p)h_{ud} + (1-p)^2 h_{dd} \right), \quad \text{where}
\]

\[
p = \frac{e^{r \Delta t} - d}{u - d}.
\]

We now have a single formula for the two-step model, and no longer need to compute three subproblems.

We claim that this is consistent with the risk-neutral pricing concept discussed in Section 9.2. Let \( S_T \) denote the price at time \( T = 2\Delta t \). Then with quite a bit of algebra for the first equality, we see that

\[
p^2(S_0u^2) + 2p(1-p)(S_0ud) + (1-p)^2(S_0d^2) = S_0e^{2r \Delta t} = E_{RNW}[S_T].
\]

This implies we can interpret \( p \) and \( 1-p \) as the risk-neutral probabilities of the stock changing by a factor of \( u \) and \( d \), respectively, at each of the two time steps. For example, in the risk-neutral world, the probability of

\[\text{footnote}{There is one subtlety in the logic behind the third subtree. Although the European put is worth $1.72 or $10.77 in six months, it does not actually provide such pay-offs at those nodes. Strictly speaking then, we cannot simply apply the fundamental principle to this third subtree. Technically, we need dynamic hedging which generalizes the fundamental principle. We do not cover this in the text. Roughly, with dynamic hedging, we can construct a portfolio of Amazon stocks and cash, which has the same value as the European put at all nodes. In the end, our logic behind the third subtree is still sound, and the } h = $5.56 \text{ answer still holds.} \]
$S_T = S_0ud$ is $2p(1-p)$ since there are 2 ways of reaching that node, each with a probability of $p(1-p)$ of occurring.

In general, suppose there are $n$ time steps between $t = 0$ and $t = T$. Let $h_{u^k d^{n-k}}$ be the payoff of the derivative when the stock price goes up $k$ times by the factor of $u$, and down $(n-k)$ times by a factor of $d$, to end at $S_0 u^k d^{n-k}$. Since the tree is recombining, the precise branch (order of ups and downs) that the stock price follows is not important. Then the price of the derivative today, using the binomial theorem, is

$$h = e^{-rT} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} h_{u^k d^{n-k}}$$

(9.2)

The proof is similar to the two-step model. Use a backwards-in-time induction step to reduce to the $(n-1)$-step model. Again we can interpret this as risk-neutral valuation, since

$$e^{-rT} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} h_{u^k d^{n-k}} = e^{-rT} E_{RNW}[h_T].$$

As mentioned before, since we have derived the pricing formula for the general $n$-step model, we no longer need to break up the tree into many one-step subtrees. However, in Section 9.4 we show how to price American options using binomial trees. Since American options can be exercised at any time, their prices do not have simple one line formulas like their European counterpart. We still have to break up the tree into many one-step trees, and consider them individually.

### 9.4 American options

Recall from Chapter 6 that without dividends, a European call and American call have the same price. Thus, to illustrate the new level of complexity with the binomial tree method when the early exercise feature is present, we will price an American put instead of an American call. We consider the two-step model. The one-step model is simpler (and the many-step model is harder).

Let us revisit the two-step example from Section 9.3.

To distinguish the American put from its European counterpart, we denote its value at the different nodes with capital letter: $H, H_u, H_d, H_{uu}, H_{ud}, H_{dd}$. Recall for our example, the risk-neutral probability is $(p, 1-p) = (0.56, 0.44)$,
the risk-free rate is 5%, $u = 1.2$, $d = 0.8$, $S_0 = 50$, $K = 52$ and the time step $\Delta t = 0.5$.

The goal is to compute $H$. We know that

$$H_{uu} = h_{uu} = 0, \quad H_{ud} = h_{ud} = 4, \quad H_{dd} = h_{dd} = 20$$

because by time $T = 1$, we have missed any chance of early exercise and so the European and American puts have the same value. See Figure 9.3.

At an earlier node, as a holder of an American option, we can either exercise at that node, receiving the node’s stock price when we sell our share, or decide not to exercise. If we choose not to exercise, we assume the value of our put, like the European put, is the discounted risk-neutral expectation of what the American put will be at the two subsequent nodes. We will choose the maximum of those two values. This determines the value of the American put at that earlier node.

Consider the node when the stock price is $S_0d = 40$. In that case an early exercise of the put has a payoff of $52 - 40 = 12$. Delaying the exercise
makes it European. We saw in Section 9.3 that the value of the European put at this node is $h_d = 10.77$. So

$$H_d = \max(52 - 40, 10.77) = \$12$$

as we decide to exercise early.

At the node when the stock price is $S_0u = 60$,

$$H_u = \max(52 - 60, 1.72) = \$1.72.$$ 

Clearly it is not advantageous to exercise early at this node.

At the first node

$$H = \max \left( e^{-0.05 \times 0.5} (0.56 \times 1.72 + 0.44 \times 12), 52 - 50 \right) = \max(6.09, 2) = \$6.09.$$ 

where here we are taking the maximum of the following: (1) early (immediate) exercise, with payoff $(52 - 50)$; and (2) waiting another 6 months. To compute the value of waiting, note that we use $H_d = 12, H_u = 1.72$ in the expectation, and not $h_d = 10.77, h_u = 1.72$. This is because even though we decided not to exercise early at $t = 0$, we still have the option to exercise early at $t = 0.5$.

### 9.5 Options with dividends and carrying costs

In the one-step tree model without dividends, recall the risk-neutral probability $(p, 1 - p)$ was characterized by

$$pS_0u + (1 - p)S_0d = S_0e^{rT} = E_{RNW}[S_T].$$

If the stock had dividends or carrying costs, this would affect the risk-free growth in exactly the same way as it affects the forward prices: the stock price in the formula needs to be replaced by the effective stock price. For example, if the stock has a carrying cost whose present value is $\$U$, then for any risk-neutral investor to invest in this stock, the expected return must compensate for the carrying cost; thus

$$pS_0u + (1 - p)S_0d = E_{RNW}[S_T] = (S_0 + U)e^{rT},$$

Solving for $p$ we get

$$p = \frac{S_0 + U}{S_0} e^{rT} - d \quad \frac{u - d}{u - d}.$$
If the stock pays a continuous dividend yield of \( q \), the result is slightly neater
\[
p = \frac{e^{(r-q)T} - d}{u - d}.
\]

For all the different variations, once we compute \((p, 1-p)\) we can use risk-neutral pricing as before to get
\[
h = e^{-rT}(ph_u + (1-p)h_d).
\]
The \(n\)-step model is similar.

### 9.6 Incomplete markets

Why a binomial tree? Why not a trinomial tree? Of course an \(n\)-step recombining binomial tree has fewer nodes than an \(n\)-step recombining trinomial tree. But both only grow linearly in \(n\). The difference is not so significant given that we can implement either on a computer.

The main advantage of the binomial tree is that it we can replicate the derivatives using stocks and cash. This allows us to use Theorem 9.4 and move to the risk-neutral world to price our derivatives. We say that the market is incomplete if we are unable to replicate the options we wish to price (and buy and sell) using the underlying stock and cash. In reality, many experts believe that markets are incomplete, which makes for interesting but complicated mathematical adjustments to the current financial model.

Back to the trinomial tree. Consider a simple one-step example. The current price of Nokia (symbol: NOK) is $20. The stock can move to $18, $21 or $23 in one year. The risk-free rate is 5%. We would like to use the technique in Section 9.1 to price a European call on Nokia with strike $19, expiring in one year. We need to construct Portfolio \(A\), which is long \(N\) stocks, short the call, and which is risk-free. This means \(N\) must solve
\[
18N - 0 = 21N - (21 - 19) = 23N - (23 - 19).
\]
This is an over-determined system of equations which, unless we were extremely lucky, has no solution. Hence we cannot hedge the risk in the call using only the stock and cash.
9.7 Problems

1. A stock is currently priced at $30. In 3 months it will be either $27 or $36. The risk-free rate is 5% per annum with continuous compounding. Compute the price of a 3-month European call with strike price $31.

2. A stock is currently priced at $30. In 3 months it will either go down by a factor of 10% or up by a factor of 20%. The risk-free rate is 5% per annum with continuous compounding.
   (a) Compute the price of a 3-month European put with strike price $31.
   (b) Use put-call parity to price a 3-month European call with strike price $31.
   (c) Compare (with explanation) your answer in part (b) with your answer to the previous problem.

3. Yahoo stock is trading today at $26. Suppose that in one month, the stock will be either $30 or $25. Using a one-step binomial tree, value the following options which expire in one month:
   (a) A European call option with $K = 20$, first with a risk-free rate $r = 0$ and then again with a risk-free rate of $r = 5\%$.
   (b) A European call option with $K = 28$ with risk-free rate $r = 0$ and $r = 5\%$.
   (c) A European call option with $K = 35$ with risk-free rate $r = 0$ and $r = 5\%$.
   (d) Repeat the above calculations for European puts with the same strike prices and risk-free rates.

4. A stock is currently priced at $25. In 4 months it will be either $22 or $29. The risk-free rate is 6% per annum with continuous compounding. Let $S_{4/12}$ be the price of the stock in 4 months. Compute the price of a derivative that pays you $(S_{4/12})^3$ dollars in 4 months.

5. A stock is currently priced at $50. In 4 months it will be either $48 or $56. The risk-free rate is 5% per annum with continuous compounding.
9.7. PROBLEMS

(a) Compute the price of a derivative consisting of one European call with strike price $50 and one European put with strike $52. Both expire in 4 months.

(b) Now assume that the above options are American and price the derivative.

6. A stock is currently priced at $25. In 6 months it will be either $26 or $30. The risk-free rate is 12% per annum with continuous compounding.

(a) Verify that the No-Arbitrage assumption holds. (Hint: see the discussion following Theorem 9.4.)

(b) Use the risk-neutral pricing method to price a European put option expiring in 6 months with strike price $28.

(c) If your portfolio is long this European put, how many stocks should it have to be risk-free?

7. Let $S_0$ denote the spot price of a stock, which after one-time step $T$ will be either $S_0^u$ or $S_0^d$ with $u > d$. Let $r$ be the risk-free rate. Suppose $d > e^{rT}$. Construct an arbitrage opportunity involving one share of stock and some cash. What is the minimum guaranteed profit of this strategy?

8. A stock price is currently $30. Over the next two 6 month periods it is expected to up by 12% or down by 8%. The risk-free rate is 4% per annum with continuous compounding.

(a) Compute the price of a one-year European put option with strike price $32.

(b) Compute the price of a one-year American put option with strike price $32.

9. Consider the set-up from the previous problem. The Meirkiec is a derivative made up of the following: the American put from the previous problem; the European put from the previous problem; a straddle combination with strike $35$; a long position in a forward contract for delivery of the underlying stock in one year with delivery price $33$; a zero-coupon bond maturing in 1 year with principal $100$. Compute the price of the Meirkiec.
10. A stock price is currently priced at $25. In 1 year it will either be $26 or $30. The risk-free rate is 5% per annum with continuous compounding.

(a) Suppose the stock pays a continuous dividend yield of 3% per annum. Construct an arbitrage opportunity.

(b) Suppose the stock pays a continuous dividend yield of 1% per annum. In this case there is no arbitrage opportunity. So go ahead and price a 1-year at-the-money European call option on the stock.

11. Consider a one-step binomial model for a stock. Now, the stock is worth $S_0$, but will be worth either $S_0u$ or $S_0d$ at time $T$ where $d < 1 < u$. For an American put option on this stock, one is allowed to either exercise the option now or wait until expiration at time $T$.

(a) Consider an American put on the stock with strike price $K$ where $K > S_0u$. Show that it is always better to exercise the put early when $r > 0$.

(b) Consider an American put with strike price $K$ where $K < S_0d$. What is the value of the put now?

(c) Find the number $\theta$ such that when the strike price $K$ is bigger than $\theta$, it is better to exercise the option early and when the strike price is less than $\theta$, it is better to wait until expiration.

12. Find the range of values for which the two-step American put from Section 9.4 has an early exercise at the intermediate time $t = 0.5$ but not at the initial time $t = 0$.

13. A stock price is currently priced at $25. Every week it will either go up or down a dollar. The risk-free rate is 4% per annum with continuous compounding.

(a) Compute the price of an at-the-money European put expiring in one year.

(b) Write a program to compute the price of an at-the-money American put expiring in one year.

14. Recall that in the risk-averse world, the expected rate of return of a stock $\mu$ is greater than the risk-free rate $r$. Consider the one-step model.
If \((q, 1 - q)\) represent the up and down probability in the risk-averse real world while \((p, 1 - p)\) represent the up and down probability in the risk-neutral world, prove that \(\mu > r\) implies \(q > p\).
Chapter 10

Probability II: Infinite sample spaces

In this chapter we continue developing the probability tools needed for option pricing. We introduce the probability density function for an infinite sample space and extend the definition of the expectation of a random variable to this setting. The Gaussian or normal distribution is introduced and studied. We also discuss what it means for two events to be independent.

10.1 Cumulative distribution function and probability density function

Recall our setup from Chapter 7. The sample space $\Omega$ is the set of all possible outcomes of an experiment. The probability function $p$ assigns to each event $A \subset \Omega$ a number $p(A)$, called the probability of the event. In this chapter, $\Omega$ may be finite or infinite.

Let $Z$ be a random variable on $\Omega$. In other words, $Z$ is a function from $\Omega$ to the real numbers $\mathbb{R}$. To simplify notation, if $x$ is a number, we will write $Z \leq x$ for the event consisting of all $\omega \in \Omega$ where $Z(\omega) \leq x$.

**Example 10.1.** A coin is flipped four times. Let $Z$ be the random variable counting the number of heads. What is $Z \leq 1$? If the coin is fair, what is $p(Z \leq 1)$?

**Solution:** We are looking for the event

$$A = \{\omega \in \Omega \mid Z(\omega) \leq 1\}.$$
Then

\[ A = \{TTTT, HTTT, THTT, TTHT, TTTH\}. \]

If the coin is fair then \( p(Z \leq 1) = p(A) = \frac{5}{16} \) since all sixteen outcomes are equally likely.

**Definition 10.2.** The cumulative distribution function or c.d.f of the random variable \( Z \) is the function

\[ F_Z : R \rightarrow R \]

defined by

\[ F_Z(x) = p(Z \leq x). \]

In the previous example of tossing the coin four times, \( F_Z(1) = \frac{5}{16} \), whereas \( F_Z(-1) = 0 \) and \( F_Z(4) = 1 \). Notice that we are suppressing the probability function in the notation for the c.d.f. The probability function needs to be specified in order to define the cumulative distribution function.

**Example 10.3.** A circular dartboard of radius one is centered at the origin. Consider the experiment of throwing a dart at the board and assume that the dart must hit the board. Also assume that every outcome is equally likely; this means that the probability of hitting a certain region on the board is proportional to the area of the region. Let \( Z \) be the random variable of the \( x \) coordinate of the dart when it hits the board. What is \( F_Z(x) \)?

**Solution:** The value of \( F_Z(x) \) is the area of the circle to the left of \( x \). Using some calculus, this can be shown to be

\[ F_Z(x) = \pi - (\arccos(x) + x\sqrt{1-x^2}), \]

when \(-1 \leq x \leq 1\). If \( x < -1 \), then no area is enclosed and so \( F_Z(x) = 0 \). And if \( x > 1 \), then the whole circle is enclosed and \( F_Z(x) = 1 \).

The cumulative distribution function makes sense whether \( \Omega \) is finite or infinite. Using the c.d.f., we can define the probability density function of a random variable \( Z \).\footnote{Strictly speaking certain hypotheses on \( Z \) must be satisfied.}
Definition 10.4. Let $Z$ be a random variable on $\Omega$. The probability density function of $Z$, denoted by $f_Z(x)$, has the defining property that

$$F_Z(x) = \int_{-\infty}^{x} f_Z(y)dy.$$ 

The probability density function or p.d.f. has the following property: the integral

$$\int_{a}^{b} f_Z(y)dy$$

measures the probability that $Z$ is between $a$ and $b$. In other words,

$$\int_{a}^{b} f_Z(y)dy = p(a \leq Z \leq b). \tag{10.1}$$

Said differently, if we graph the p.d.f. of $Z$, the area under the curve between $a$ and $b$ is the probability of the event where $Z$ is between $a$ and $b$. If the c.d.f. $F_Z(x)$ is nicely behaved, then by the Fundamental Theorem of Calculus, the p.d.f. will be the derivative of $F_Z(x)$

$$\frac{dF_Z(x)}{dx} = f_Z(x).$$

Example 10.5. Consider a square dartboard with endpoints $(0, 0), (1, 0), (0, 1)$, and $(1, 1)$. Let $Z$ be the $x$-coordinate of a dart thrown at this board (which we assume always hits the board). Find $f_Z(x)$.

Solution: Clearly, $f_Z(x) = 0$ if $x > 1$ or $x < 0$ since the dart must hit the board. Now if $0 \leq x \leq 1$, the c.d.f. $F_Z(x)$ is equal to the area of the square to the left of $x$. In other words, $F_Z(x) = x \times 1 = x$. Taking the derivative reveals that $f_Z(x) = 1$ when $0 \leq x \leq 1$. A random variable whose p.d.f. is constant is said to have a uniform distribution.

There are two important properties of $f_Z(x)$:

1. $f_Z(x) \geq 0$.
2. $\int_{-\infty}^{\infty} f_Z(y)dy = 1$
The first property reflects the fact that the probability of any event is always nonnegative. The second property reflects the fact that the event where \( Z \) is between negative infinity and positive infinity is all of \( \Omega \) and hence has probability 1. Any reasonable function that satisfies the above two properties is a candidate to be a p.d.f. for some random variable on some sample space.

Typically we are more interested in the p.d.f. than the c.d.f. of a random variable. Generally we will specify what the p.d.f. of a random variable is, rather than deduce it from the context.

**Example 10.6.** Let \( Y \) be the price of Pepsi stock in 17 days from today. Suppose that the our best financial analyst crunches the numbers and forecasts that the p.d.f. is given by

\[
f_Y(x) = \begin{cases} 
\frac{1}{(x - 50)^4} & \text{if } x \leq 49 \text{ or } x \geq 51, \\
\frac{1}{6} & \text{if } 49 < x < 51.
\end{cases}
\]

Note that \( f_X(y) \geq 0 \). Calculus confirms that \( \int_{-\infty}^{\infty} f_Y(x)dx = 1 \). So \( f_Y(x) \) is a legitimate p.d.f. What is the probability that Pepsi stock in 17 days will be between $48 and $52? What is the probability that Pepsi will trade between $58 and $62?

**Solution:** We use Formula [10.1]

\[
p(48 \leq Y \leq 52) = \int_{48}^{52} f_Y(x)dx = .916 = 91.6\% \\
p(58 \leq Y \leq 62) = \int_{58}^{62} f_Y(x)dx = 0.0004 = 0.04\%
\]

So Pepsi will probably hover around $50.

### 10.2 Expectation

We can now extend the definition of expectation to handle all possible sample spaces \( \Omega \) whether finite or infinite. If \( \Omega \) is finite, the definition from Chapter 7 is still valid and remains the preferred definition.

**Definition 10.7.** The expected value, or expectation, \( E[X] \) of a random variable \( X \) is defined by

\[
E[X] = \int_{-\infty}^{\infty} y f_X(y)dy.
\]
Note that $E[X]$ is just a number. It is often referred to as the mean of $X$.

**Example 10.8.** Consider the Pepsi stock from the previous section.

$$E[Y] = \int_{-\infty}^{\infty} x f_Y(x)dx$$

$$= \int_{-\infty}^{49} \frac{x}{(x-50)^4}dx + \int_{49}^{51} \frac{x}{6}dx + \int_{51}^{\infty} \frac{x}{(x-50)^4}dx$$

$$= \$50.$$ So the expected price of Pepsi stock in 17 days is $\$50$.

If $g : \mathbb{R} \to \mathbb{R}$ is a function and $X : \Omega \to \mathbb{R}$ is a random variable, then the composition

$$g \circ X = g(X) : \Omega \to \mathbb{R} \to \mathbb{R}$$

goes from sample points to numbers. Recall that a composition of functions is defined by using the output of the first function as the input of the second. Hence $g(X)$ is another random variable. As noted in Chapter 7, we can also add, subtract, and multiply random variables. Basically anything we can do to a function, we can do to random variables, since random variables are functions.

The expectation has some useful properties, some of which we already mentioned in Chapter 7.

**Theorem 10.9.** Let $X$ and $Y$ be random variables. Let $g : \mathbb{R} \to \mathbb{R}$. Let $a$ be any constant. Then

- $E[g(X)] = \int_{-\infty}^{\infty} g(z)f_X(z)dz$

$$E[aX + Y] = aE[X] + E[Y].$$

- $E[a] = a$

**Example 10.10.** Continuing with the example, suppose the price of Pepsi stock today is 49 dollars. Recall that in 17 days, Pepsi shares trade at $Y$ with $E[Y] = 50$. Suppose the spot price of Coca-Cola is 80 dollars. Denote
by $W$ its price in 17 days. Assume that $W$ has a uniform distribution with p.d.f.

$$f_W(x) = 0.05 \text{ if } 72 \leq x \leq 92 \text{ and, } f_W(x) = 0 \text{ otherwise.}$$

You are long 30 Pepsi shares and 60 Coca-Cola shares. Compute the expected 17-day return of your portfolio.

**Solution:** Recall that the return over the next 17 days is the value of the portfolio in 17 days, less its value today. We do not adjust for time-value of money. Hence the expected return is

$$E[30(Y - 49) + 60(W - 80)]$$

$$= 30(E[Y] - 49) + 60(E[W] - 80)$$

$$= 30(50 - 49) + 60 \int_{72}^{92} 0.05x \, dx - 4800$$

$$= 30 + 4920 - 4800 = 150$$

dollars.

The definition of variance, standard deviation, covariance, and correlation all remain valid for a general sample space $\Omega$ since these definitions were given in terms of expectation. For example, in this setting we have

**Definition 10.11.** The variance of $X$ is equal to

$$Var(X) = E[(X - E[X])^2]$$

$$= E[X^2] - E[X]^2$$

$$= \int_{-\infty}^{\infty} y^2 f_X(y) \, dy - \left( \int_{-\infty}^{\infty} y f_X(y) \, dy \right)^2.$$

### 10.3 The normal variable

A **standard normal variable** $Z$ is defined as any random variable with the following probability density function

$$f_Z(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

Graphically, this function is “bell-shaped.”
A random variable with this p.d.f. is said to have a **standard normal distribution**. It is important to realize that there can be many random variables with the same probability density function. In particular, there can be many random variables that have a standard normal distribution.

Let us check that \( f_Z \) is a legitimate probability density function. First, the function is always nonnegative since the exponential is always positive

\[
\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \geq 0.
\]

Second, we need to check that the area under its graph is 1

\[
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1.
\]

The integral can be done using polar coordinates and a little multivariable calculus.

We compute its expectation

\[
E[Z] = \int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 0
\]

where we either use integration by substitution, letting \( u = y^2/2 \), or note that the integrand is an odd function, and so integrates to 0. (Recall that a function \( f \) is **odd** if \( f(-x) = -f(x) \) and **even** if \( f(-x) = f(x) \).)

To compute the variance, we need to integrate by parts, setting

\[
u = \frac{1}{\sqrt{2\pi}} y, \quad dv = ye^{-\frac{y^2}{2}} dy
\]

to get from the second to third line below.

\[
\begin{align*}
Var(X) &= E[Z^2] - E[Z]^2 \\
&= E[Z^2] - 0^2 \\
&= \frac{1}{\sqrt{2\pi}} y \cdot (-e^{-\frac{y^2}{2}}) \bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} (-e^{-\frac{y^2}{2}}) dy \\
&= 0 - 0 - (-1) = 1.
\end{align*}
\]

The second term in the third line above is 1 by our previous discussion. Note \( SD(Z) = \sqrt{1} = 1 \).
Notationally, we write
\[ Z \sim N(0, 1) \]
and say that \( Z \) is \textit{normally distributed with mean 0 and standard deviation 1}.

Since normally distributed random variables play such a prominent role in math finance, it is useful to know their cumulative distribution function
\[
F_Z(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.
\]
Unfortunately, there are no known methods to compute the indefinite integral algebraically. Hence the c.d.f. is usually computed with a computer or a calculator. It can also be looked up in a table.

The cumulative distribution function \( F_Z(x) \) gives the probability that \( Z \) will be less than \( x \), \( p(Z \leq x) \). However, in finance we are also interested in finding the probability that \( Z \) lies in a certain range of values, \( p(w \leq Z \leq x) \).

\[
p(w \leq Z \leq x) = p(Z \leq x) - p(Z \leq w) = F_Z(x) - F_Z(w).
\]

When \( w = -x \), it turns out that a table or calculator need only be accessed once instead of twice as above. We first use the fact that \( f_Z \) is even to derive the following.
\[
p(Z \leq -x) = \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy
= \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy
= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy
= 1 - p(Z \leq x);
\]
therefore,
\[
p(-x \leq Z \leq x) = F_Z(x) - F_Z(-x) = 2F_Z(x) - 1.
\]

**Example 10.12.** Find a range of values about 0 in which \( Z \) has a 90\% chance of being in that range. This is called a \textit{90\% confidence interval} of \( Z \) about its mean.
Solution: The question asks us to find $x$ such that

$$p(-x \leq Z - 0 \leq x) = 0.90$$

The “−0” in the above expression is the emphasize that we are interested in a range around the mean of $Z$, which happens to be 0.

Using the previous identity, we need an $x$ that satisfies

$$0.90 = 2p(Z \leq x) - 1, \text{ or } F_Z(x) = 1.9/2 = 0.95.$$ 

Using a table or a computer, we find $x = 1.645$

$$0.90 = p(-1.645 \leq Z \leq 1.645)$$

Note that we chose an upper and lower bound equidistant from the mean (1.645 and −1.645 are the same distance from 0). There are other 90% confidence intervals without this “symmetric” property; however, we will not consider them here.

Now we define the general normal variable. Pick real values $\sigma > 0$ and $\mu$. Let $X = \sigma Z + \mu$. From Section 10.2 and Chapter 7,

$$E[X] = E[\sigma Z + \mu] = \sigma E[Z] + \mu = \mu$$

$$SD(X) = SD(\sigma Z + \mu) = \sigma SD(X) = \sigma.$$ 

We write

$$X \sim N(\mu, \sigma)$$

and say that $X$ is normally distributed with mean $\mu$ and standard deviation $\sigma$. We can further check that the probability density function of $X$ is

$$f_X(y) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$$ 

Example 10.13. Compute a 90% confidence interval of $X$ about its mean.

Solution: Note that for any random variable $Y$ and real numbers $a, b, c$ with $b > 0$

$$p(a \leq bY + c) = p(a - c \leq bY) = p \left( \frac{a - c}{b} \leq Y \right)$$

If $b < 0$ we have to flip “<” to “>” when $b$ switches sides.
Since we do not have a table for \( X \), we will try to use the above identity to extract the answer from our information on \( Z \).

\[
0.90 = p(-1.645 \leq Z \leq 1.645) \\
= p(-1.645\sigma \leq \sigma Z \leq 1.645\sigma) \\
= p(-1.645\sigma + \mu \leq \sigma Z + \mu \leq 1.645\sigma + \mu) \\
= p(-1.645\sigma + \mu \leq X \leq 1.645\sigma + \mu).
\]

So our 90%-confidence interval is \((-1.645\sigma + \mu, 1.645\sigma + \mu)\). In other words, we are 90% confident that \( X \) will be within 1.645 standard deviations of its mean.

The **moment generating function** for any random variable \( Y \), \( \Phi_Y(t) \), is defined by

\[
\Phi_Y(t) = E[e^{tY}].
\]

This expression may seem a bit confusing. On the left hand side, \( t \) is a variable, but on the right hand side, we treat it as a constant when we multiply it with the random variable \( Y \) and compute the expectation.

**Example 10.14.** First compute the moment generating function of a standard normal variable \( Z \sim N(0,1) \). Then compute the moment generating function of any normal variable \( X \sim N(\mu, \sigma) \).

**Solution:** Remember, when computing the integrals in the expectation, treat \( t \) as a constant. In the first step, we use Theorem \[10.9\]. In the third step we complete the squares. In the fourth step we substitute \( w \) for \( y - t \).

\[
\Phi_Z(t) = \int_{-\infty}^{\infty} e^{ty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-2t)^2}{2}} dy \\
= e^{t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}} dy \\
= e^{t^2} \cdot 1 = e^{t^2}.
\]

We could repeat the above integral calculus when computing \( \Phi_X(t) \), or we could take a short-cut and use Theorem \[10.9\] to reduce the work to the
10.3. THE NORMAL VARIABLE

previous calculation.

\[
\Phi_X(t) = E[e^{tX}] = E[e^{t(\sigma Z + \mu)}] = e^{\mu} E[e^{t\sigma Z}] = e^{\mu} \Phi_Z(t\sigma) = e^{\mu} e^{\frac{(t\sigma)^2}{2}} = e^{\frac{(t\sigma)^2}{2} + t\mu}.
\]

**Example 10.15.** Let \( X \) be normally distributed with mean \( \mu \) and standard deviation \( \sigma \). Suppose your personal wealth at time \( t \geq 0 \) is \( 10000e^{tX} \). How long until you expect your wealth to double from its initial \((t = 0)\) value of \( 10000e^0 = 10000 \) dollars? At that time, what is the probability that your wealth has indeed doubled?

**Solution:** We would like to find \( t \) such that

\[
E[10000e^{tX}] = 10000e^{\frac{(t\sigma)^2}{2} + t\mu} = 10000 \times 2, \text{ that is}
\]

\[
\frac{(t\sigma)^2}{2} + t\mu = \ln 2
\]

This has a (positive) solution of

\[
t_* = \frac{-\mu + \sqrt{\mu^2 + 2\sigma^2 \ln 2}}{\sigma^2}.
\]

Write \( X = \sigma Z + \mu \) where \( Z \) is a standard normal variable. We next need to compute

\[
p(10000e^{t_*X} \geq 20000) = p(t_*X \geq \ln 2) = p(t_*(\sigma Z + \mu) \geq \ln 2) = p \left( Z \geq \frac{\ln 2}{t_*} - \frac{\mu}{\sigma} \right) = p \left( Z \leq -\frac{\ln 2}{t_*} - \frac{\mu}{\sigma} \right) = F_{\mathcal{N}} \left( \frac{\mu - \ln 2}{\sigma} \right)
\]
where \( t_\ast \) is expressed in terms of \( \mu \) and \( \sigma \) in the previous paragraph.

We return to the moment generating function in Chapter 11.

### 10.4 Conditional probability and independence

Let \( A, B \subset \Omega \) be two events in the sample space \( \Omega \). Intuitively, the conditional probability, \( p(A|B) \), is the probability that \( A \) will occur given that \( B \) occurs.

**Example 10.16.** Consider the experiment of flipping a nickel and a dime. Let \( A \) be the event that both are heads, let \( B \) be the event that the nickel is heads. Then

\[
p(A|B) = \text{probability that both are heads given that the nickel is heads} = 0.5
\]

\[
p(B|A) = \text{probability that the nickel is heads given that both are heads} = 1.
\]

Formally, the definition of **conditional probability** is

\[
p(A|B) = \frac{p(A \cap B)}{p(B)}.
\]

Let us check the consistency of definitions in the example above

\[
p(A \cap B) = p(A) = 0.25, \quad p(B) = 0.50, \quad \text{so} \quad p(A|B) = \frac{0.25}{0.5} = 0.5.
\]

Two events \( A, B \subset \Omega \) are said to be **independent** if

\[
p(A \cap B) = p(A)p(B).
\]

(10.2)

The reason for this terminology comes from the following observation. If \( A \) and \( B \) are independent then

\[
p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{p(A)p(B)}{p(B)} = p(A).
\]

So the probability that \( A \) occurs is not altered by knowing what happens to \( B \). Similarly, \( p(B|A) = p(B) \).

Two random variables are **independent** if for any numbers \( a, b \)

\[
p(X \leq a \text{ and } Y \leq b) = p(X \leq a)p(Y \leq b).
\]

(10.3)
Example 10.17. Let us continue with the coin-flipping experiment.

Let $C$ be the event that the dime flips head, and $A$ and $B$ be the events from the previous example.

Then

$$p(B) = 0.5, \quad p(C) = 0.5, \quad \text{and}$$

$$p(B \cap C) = p(\{h, h\}) = 0.25.$$ 

So we see that

$$p(B \cap C) = 0.5 \times 0.5 = p(B)p(C)$$

confirming our intuition that the nickel flipping heads and the dime flipping heads are independent events.

In contrast,

$$p(A) = 0.25, \quad p(C) = 0.5, \quad \text{and}$$

$$p(A \cap C) = p(\{h, h\}) = 0.25.$$ 

So $A$ and $C$ are not independent events given that

$$p(A \cap C) \neq p(A)p(C).$$

The other way to see this is that the probability that the dime flips heads, $p(C)$, is 0.5, but the probability that the dime flips heads given both coins flip heads, $p(C|A)$, is 1. These probabilities would need to be equal if the events were independent.

Let $X_1$ be 3 if nickel flips heads and 1 otherwise. Let $X_2$ be 7 if dime flips heads and 0 otherwise. Let $X_3$ be the face value of the coins that flipped head. Then $X_1$ and $X_2$ are independent while $X_1$ and $X_3$ are not.

10.5 Problems

1. Let $T$ denote the time at which the shoes you purchased today at $t = 0$ go out of fashion. Suppose the p.d.f. of $T$ is given by $f_T(t) = ke^{-0.1t}$ for $t \geq 0$ and $f_T(t) = 0$ for $t < 0$. Here $t$ is measured in years and $k$ is some constant.

   (a) Find the value of $k$ that makes $f_T$ an actual p.d.f.
(b) What is the probability of your shoes going out of fashion within one month of buying them?

(c) What is the probability that they go out fashion between years 1 and 2? (That is, they become uncool in their second year.)

(d) What is the probability that they are still in fashion after 10 years?

(e) Is it possible that you purchased out-of-fashion shoes? Why or why not?

2. Fix two numbers $a < b$. Consider the random variable $X$ whose p.d.f. $f_X(x)$ is given by

$$f_X(x) = c \text{ if } a \leq x \leq b \text{ and, } f_X(x) = 0 \text{ otherwise.}$$

(a) Find the value of $c$ that makes $f_X$ an actual p.d.f.

(b) Graph the c.d.f. $F_X(x)$.

3. Consider the random variable $X$ from the previous problem.

(a) Compute $E[X]$.

(b) Compute a 90% confidence interval about its mean.

(c) Compute $Var(X)$.

4. Consider the nickel-dime coin flip and the random variables $X_1, X_2$ from the end of Section 10.4. Compute $p(X_1 + X_2 = 8 | X_1 + X_2 > 1)$. Note “$X_1 + X_2 = 8$” is the event that the random variables $X_1$ and $X_2$ sum to 8.

5. The risk-free rate (per annum with semi-annual compounding) is 5%. You have a portfolio worth 10000 dollars today. Let $X$ be the return of your portfolio in 6 months. Suppose the p.d.f. of $X$ is given by

$$f_X(x) = \frac{x + 50}{80000} \text{ if } -50 \leq x \leq 350, \text{ and } 0 \text{ otherwise.}$$

(a) Compute the expected return.

(b) What is the probability that your portfolio will lose money in 6 months?

(c) Are you risk-neutral, risk-averse, or risk-prone? Justify your answer.
6. Let $X$ be a normally distributed variable with mean $\mu$ and standard deviation $\sigma$.

(a) Compute $E[e^{2X}]$.

(b) Compute $E[X^2 + e^{2X} + 17]$.

(c) Compute $Var(e^{2X} + 17)$.

7. For time $t > 0$, let $X_t$ be a random variable normally distributed with mean $0.08t$ and standard deviation $0.2\sqrt{t}$. Let $S_0$ denote the known spot price of an asset today and let the random variable $S_t$ denote the unknown price of the asset in time $t$ years. Assume

$$S_t = S_0 e^{X_t}.$$

(a) If you invest all your personal money today in the asset, by when do you expect your personal worth to triple?

(b) What is the probability that your worth triples after two years?

8. You manage a mutual fund worth 100 million dollars today. The return on your portfolio in 10 days is normally distributed with mean 0 and standard deviation 2 million. Find the number $N$ such that you can say the following: “I am 99% sure that my fund will not lose more than $N$ dollars after 10 days.”

$^2$N is known as the 10-day Value at Risk, a tool developed by JP Morgan in the 1990s for measuring investment risk. It is supposed to approximate a “worst-case” scenario, by ignoring the worst 1%. This approximation some blame as a partial cause of the many bank failures in 2008-2009.
Chapter 11

The Black-Scholes formula

In Section 9.3, Equation 9.2 gives a formula for derivative prices when the stock price is modeled by a binomial tree, making an arbitrary number of up and down jumps. What happens as those jumps become smaller and more frequent? It turns out that the limiting stock price behaves according to the so-called “log-normal” model, which we describe in Section 11.2. This model can also be derived from some basic economic assumptions, including the efficient market hypothesis of Section 11.1.

As in Chapter 9, we use the risk-neutral pricing principle to value European-style derivatives. Again we see that the (real world) expected return of the stock is not relevant in the final answer; however, the stock volatility (see Section 11.3) plays an important role for pricing options. Recall that risk-neutral pricing requires us to compute the expectation of the price of the derivative in the risk-neutral world. In Section 11.4 we make such computations for some fundamental examples. In particular, when the derivative is a European call or put, the computation leads to the Black-Scholes formula, a main result in modern finance.

To simplify the exposition, we assume that all investors can borrow and save money at the same rate, and that the yield curve remains constant and flat for all time. We will denote this risk-free rate with continuous compounding by \( r \).


11.1 The efficient market hypothesis

Let $\Omega$ be the sample space consisting of the closing price of a stock on all future days. Suppose today is $t = 0$ so the spot price $S_0$ is a known quantity. Suppose $S_0 = 400.00$. Let $S_3$ and $S_4$ be the price of the stock in 3 and 4 days, respectively. These are random variables. Recalling the definition of independence from Section 10.4, one could ask if these two random variables are independent. If they are, then

$$P(S_4 \leq 29 | S_3 \leq 30) = \frac{P(S_4 \leq 29 \cap S_3 \leq 30)}{P(S_3 \leq 30)} = \frac{P(S_4 \leq 29)P(S_3 \leq 30)}{P(S_3 \leq 30)} = P(S_4 \leq 29).$$

Could this be possible? If today’s price is 400 then $P(S_4 \leq 29)$ should be tiny. It is unlikely that the stock will lose more than 90% its face value in the next 4 days, dropping from 400 to below 29. On the other hand $P(S_4 \leq 29 | S_3 \leq 30)$ is probably a bit less than 50%. There is a good chance that a 30-dollar stock on day 3 will stick around that price for a day, possibly dropping below 29 by day 4. The upshot is that the stock prices on different days are not independent.

Next we consider the percent changes in the stock price. Let $X = \frac{S_4 - S_3}{S_3}$, $Y = \frac{S_7 - S_6}{S_6}$ represent the daily percent change in stock prices on the fourth and seventh days, respectively.

There is a hypothesis, known as the efficient market hypothesis which states that $X$ and $Y$ are independent. The reason for this hypothesis is the following: when the market determines the price of the stock to be $S_3$, the market participants do so knowing all worldwide events that have occurred up until day 3. Thus, $X$ is only affected by information that occurs on day 4. Similarly, $Y$ is only affected by information on day 7. So $X$ and $Y$ are independent random variables.

11.2 The log-normal model

Under the following mild assumptions on the stock price $S_t$ we can get a nice mathematical model of the stock price:
1. Assume the efficient market hypothesis, that is, returns over different time periods are independent.

2. Assume that the percent changes in stock prices for different non-overlapping time periods with the same interval of time are identically distributed. This means that the p.d.f.’s for the random variables representing these time intervals are the same. For example, if $X$ is the percent change in stock price in the third week of August and $Y$ is the percent change in stock price in the fourth week of November then

\[
P(0.9 \leq X \leq 1.7) = P(0.9 \leq Y \leq 1.7)
\]

3. Assume the stock price is continuous (no jumps).

Using some mathematical techniques which lies outside of the scope of this text, we can show that these three assumptions imply that the returns are log-normal; that is, for $t \leq T$

\[
\ln \left( \frac{S_T}{S_t} \right) \sim N(a(T - t), \sigma \sqrt{T - t})
\]

for some constants $a$ and $\sigma$.

This model is intricately related to the binomial model. One can prove that as the number of steps approaches infinity, the binomial tree model for the discrete stock price process converges to the log-normal model for the continuous stock price process. This proof uses the Central Limit Theorem, which is a fundamental theorem in probability, and which also lies outside the scope of this text.

Log-normality can be restated two other ways. Let $Z \sim N(0, 1)$ be a standard normal variable. Then

\[
\ln \left( \frac{S_T}{S_t} \right) = \sigma \sqrt{T - t}Z + a(T - t) \quad \text{or} \quad S_T = S_t e^{\sigma \sqrt{T - t}Z + a(T - t)}.
\]

It would be more correct to write $Z(t, T)$ instead of $Z$, since as random variables, $\ln \frac{S_T}{S_t}$ and $\ln \frac{S_{T'}}{S_{t'}}$ are “driven” by different normal variables. To simplify notation, however, we suppress the $t$ and $T$ dependence.
Suppose that today is time \( t \); thus, \( S_t \) is a known quantity and no longer random. Suppose that in the real world, see Section 9.2, the stock is expected to grow at a rate \( \mu \), by which we mean,

\[
E_{RW}[S_T] = S_t e^{\mu(T-t)}.
\]

Recall that the stock in the risk-neutral world, like any other asset, is expected to grow at the risk free rate

\[
E_{RNW}[S_T] = S_t e^{r(T-t)}.
\]

As mentioned before, most people believe that investors are risk-averse and must be compensated for investing in risky assets; thus, like in the binomial tree model, this averseness implies \( \mu > r \). In the real world (or risk-neutral world), \( \mu \) (or \( r \)) is called the expected return of the stock.

The parameter \( a \) in the log-normal model is different in the real world than in the risk-neutral world. In the real world

\[
S_t e^{\mu(T-t)} = E_{RW}[S_T] = E_{RW}[S_t e^{\sigma \sqrt{T-t}Z + a(T-t)}] = S_t e^{a(T-t)} E_{RW}[e^{\sigma \sqrt{T-t}Z}] = S_t e^{a(T-t)} \Phi_Z(\sigma \sqrt{T-t}) = S_t e^{a(T-t)} e^{\frac{\sigma^2(T-t)}{2}} = S_t e^{a(T-t)} e^{\frac{\sigma^2(T-t)}{2}} = S_t e^{a(T-t)} e^{\frac{\sigma^2(T-t)}{2}},
\]

where we have used the moment generating formula (see Equation 10.3). So \( a = \mu - \frac{\sigma^2}{2} \) and the log-normal model for the stock in the real world is

\[
\ln \left( \frac{S_T}{S_t} \right) \sim N \left( (\mu - \frac{\sigma^2}{2})(T-t), \sigma \sqrt{T-t} \right).
\]

A similar computation, replacing everywhere \( \mu \) with \( r \), proves that the log-normal model for the stock in the risk neutral world is

\[
S_T = S_t e^{\sigma \sqrt{T-t}Z + (r - \frac{\sigma^2}{2})(T-t)}.
\]

To summarize:
Formula 18 (Log-normal model). The stock price $S_t$ follows the log-normal model. That is,

$$S_T = S_t e^{\sigma \sqrt{T-t} Z + \left(\mu - \frac{\sigma^2}{2}\right)(T-t)}$$

where $\mu$ is the expected return of the stock and $t \leq T$ and $\sigma$ is a parameter called the volatility of the stock. Here, $Z$ is a standard normal variable.

11.3 Volatility and expected return of a stock

Let $T = t + 1$ be one unit of time into the future. Then

$$\ln \left( \frac{S_{t+1}}{S_t} \right) \sim N \left( \mu - \frac{\sigma^2}{2}, \sigma \right).$$

Thus $\sigma$ is the standard deviation of $\ln \frac{S_{t+1}}{S_t}$. Recall that throughout the text we have been using years for our units of time. If we continue to do so, then $\sigma$ is called the annual volatility. If the units of times are days it is called daily volatility. Since the default is years, annual volatility is often simply called volatility.

Informally, volatility $\sigma$ represents how much the percent change in stock price fluctuates in one unit of time. Let $T = t + \Delta t$ where $\Delta t$ is small (like a couple of days, if the units are in years). Using Taylors series for the exponential function $e^x$,

$$\frac{S_{t+\Delta t}}{S_t} = e^{\sigma \sqrt{\Delta t} Z + \left(\mu - \frac{\sigma^2}{2}\right) \Delta t}$$

$$= 1 + \sigma \sqrt{\Delta t} Z + \left(\mu - \frac{\sigma^2}{2}\right) \Delta t + \frac{1}{2} \left( \sigma \sqrt{\Delta t} Z + \left(\mu - \frac{\sigma^2}{2}\right) \Delta t \right)^2 + \ldots$$

$$= 1 + \sigma \sqrt{\Delta t} Z + O(\Delta t)$$

where $O(\Delta t)$ represents the terms involving higher powers of $\Delta t$.

On the other hand, letting $\Delta S = S_{t+\Delta t} - S_t$,

$$\frac{S_{t+\Delta t}}{S_t} = \frac{S_t + \Delta S}{S_t} = 1 + \frac{\Delta S}{S_t}$$
Combining these two expressions of $\frac{S_{t+\Delta t}}{S_t}$ and ignoring the small $O(\Delta t)$ term, we get

$$\frac{\Delta S}{S_t} \approx \sigma \sqrt{\Delta t} Z.$$ 

This approximation implies

$$SD\left(\frac{\Delta S}{S_t}\right) \approx \sigma \sqrt{\Delta t}.$$ 

This computation shows that the volatility coefficient, scaled by $\sqrt{\Delta t}$, measures fluctuations in the percent change over small time increments $\Delta t$. This role of volatility is more intuitive and frequently used than its formal definition as the standard deviation of $\ln \frac{S_{t+1}}{S_t}$.

As a final remark, we mention a useful coincidence. There are approximately $256 = 16^2$ business days (during which stock prices can change) in the year. So if $\sigma$ measures the standard deviation in percent change of stock price over a year, then

$$\frac{\sigma}{\sqrt{256}} = \frac{\sigma}{16}$$

measures the standard deviation in percent change of stock price over a day. Alternatively, the annual volatility is 16 times the daily volatility. We note that volatility is usually quoted as a percentage.

**Example 11.1.** Suppose the current stock price of a stock is $20$, the expected return per annum in $12\%$, the volatility is $20\%$.

1. Compute the expected stock price in $4$ (business) days, as well as the expected percent change over this period.
2. Approximate the standard deviation of the stock price over the $4$ days.
3. Provide a $90\%$ confidence interval for the stock price in $4$ days.  

**Solution:**

1. Let $t = 0$ be the present and set $T = \frac{4}{256}$, then

$$E_{RW}[S_T] = 20e^{0.12\times\frac{4}{256}} = 20.04.$$ 

Such a computation is useful for fund managers to estimate the “Value at Risk” (VaR) of their portfolio. See Exercise 8 from Chapter 10.
Therefore the expected percent change is exactly

\[
E \left[ \frac{S_T - S_0}{S_0} \right] = \frac{E[S_T] - S_0}{S_0} = \frac{20.04}{20} - 1 = 0.002 = 0.2\
\]

2. Since \( S_0 = 20 \) is known

\[
SD(S_T) = SD(S_T - S_0) = SD(\Delta S) \\
\approx S_0 \sigma \sqrt{T} \\
= 20 \times 0.2 \times \sqrt{\frac{4}{256}} = 0.50
\]

Note how the expected change \( E[\Delta S] = 0.04 \) is much less than the fluctuation \( SD(\Delta S) \approx 0.50 \). This is because \( \Delta t = \frac{4}{256} \) is smaller than \( \sqrt{\Delta t} \) by a factor of 8.

To compute the standard deviation exactly, we could use the following formula for log-normal random variables which we leave as an exercise

\[
SD(S_T) = S_0 e^{\mu T} \sqrt{e^{\sigma^2 T} - 1}.
\]

3. Recall that if \( X \sim N(\alpha, \beta) \), that is \( X \) is normally distributed with mean \( \alpha \) and standard deviation \( \beta \), then

\[
0.9 = P(-1.645\beta + \alpha \leq X \leq 1.645\beta + \alpha)
\]

Choose

\[
\alpha = (\mu - \sigma^2/2)T = \left(0.12 - \frac{0.2^2}{2}\right) \times \frac{4}{256} = 0.0016
\]

\[
\beta = \sigma \sqrt{T} = 0.20 \times \sqrt{\frac{4}{256}} = 0.025
\]

then

\[
S_T = S_0 e^{\sigma \sqrt{T} Z + (\mu - \sigma^2/2)T} \\
= 20 e^{0.0016 + 0.025Z} \\
= 20 e^X
\]
To get a 90% confidence interval for $S_T$ from the confidence interval for $X$,

$$0.9 = P(-1.645 \times 0.025 + 0.0016 \leq X \leq 1.645 \times 0.025 + 0.0016) = P(20e^{-1.645\times0.025+0.0016} \leq 20e^X \leq 20e^{1.645\times0.025+0.0016}) = P(20e^{-1.645\times0.025+0.0016} \leq S_T \leq 20e^{1.645\times0.025+0.0016})$$

So a 90%-confidence interval for the future stock price $S_T$ is

$$(20e^{-1.645\times0.025+0.0016}, 20e^{1.645\times0.025+0.0016}) = (19.225, 20.873).$$

### 11.4 Risk neutral pricing

Recall the risk-neutral pricing principle in the binomial tree model from Section 9.2. For example, in the one-step model, the price of a derivative is given by

$$h = e^{-rT}(ph_u + (1-p)h_d) = e^{-rT}E_{RNW}[\text{payoff of derivative}]$$

where $E_{RNW}[\cdot]$ denotes the expectation in the risk-neutral world. This means you go to the risk neutral world to compute the real world price of the derivative.

To prove this formulas, we only assumed that the stock price behaved like a binomial tree and that there was no arbitrage. Black and Scholes (and Merton) in 1973 proved that if the stock price follows the log-normal model, and if there is no arbitrage, then the price of the derivative is again

$$h = e^{-rT}E_{RNW}[\text{payoff of derivative}].$$

We continue to call this risk neutral valuation or pricing. We will see below that we have already verified this result for simple “derivatives” such as bonds and forwards.

**Example 11.2.** Consider the derivative that pays you $100, regardless of the outcome, in $T$ years. This derivative is just a $T$-year zero-coupon bond. Suppose today is time 0. Let $r$ be the risk-free rate. Recall the yield curve...
11.4. RISK NEUTRAL PRICING

is flat and constant. Then according to risk neutral valuation, the price $h$ of the derivative is

$$h = e^{-rT}E_{RNW}[100] = \$100e^{-rT}$$

as we already knew.

**Example 11.3.** Consider a derivative on an asset whose payoff is $S_T - K$ dollars where $K$ is some pre-specified number and $S_T$ is the spot price of the asset at time $T$. Suppose today is time 0 and the spot price is $S_0$. Then the price of this derivative is

$$h = e^{-rT}E_{RNW}[S_T - K] = e^{-rT}(S_0e^{rT} - K) = S_0 - e^{-rT}K.$$ 

If you call $K$ the “delivery price” then this derivative is a long position in a forward contract with delivery price $\$K$ and $h$, as we have seen before, is the value of the position.

**Example 11.4.** We could make the previous example more complicated. A derivative pays you 3 times the amount of the stock price in 9 months, as well as $\$75$. The current spot price is $\$20$. The expected return of the stock is 10%, the risk-free rate is 6%. The price of the derivative $h$ is then

$$h = e^{-0.06\times9/12}E_{RNW}[3S_{9/12} + 75] = 3 \times 20 + 75e^{-0.06\times9/12}.$$ 

Note that the expected return is irrelevant to this problem. In the calculation, we have used the fact that

$$E_{RNW}[3S_{9/12}] = 3E_{RNW}[S_{9/12}] = 3S_0e^{0.06\times9/12}.$$ 

**Example 11.5.** A “power derivative” has a pay-off given by a power of the stock at some future time. Suppose the stock has the specifications in the previous example and the stock volatility is 20%. Consider a power derivative that pays you the square of the stock price in 9 months. There is no “delivery” price associated to this derivative.

$$h = e^{-rT}E_{RNW}[S_T^2]$$

$$= e^{-rT}E_{RNW}[S_0^2e^{2(r-\sigma^2/2)T+2\sigma\sqrt{T}Z}]$$

$$= e^{-0.06\times0.75}E_{RNW}[20^2e^{2(0.06-(0.2^2/2))0.75+2\times0.2\times\sqrt{0.75}Z}]$$

$$= e^{-0.06\times0.75}20^2e^{2(0.06-(0.2^2/2))0.75e^{(2\times0.2\sqrt{0.75})^2/2}} = 431.15$$
Example 11.6. Consider the stock from the previous example. Consider a European call expiring in 9 months with strike price $19. The price is

\[ h = e^{-0.06 \times 9/12} E_{RNW}[\max(S_{9/12} - 19, 0)]. \]

This expectation is surprisingly hard to compute. As with the power derivative, we need to know the stock volatility to compute it.

The previous example (as well as the idea of risk-neutral pricing) earned Scholes and Merton the Nobel prize in economics in 1997. The answer bears the name the Black-Scholes formula after the original two authors who wrote the paper proving it. (Black died before 1997.)

Consider a stock with a spot price of \( S_0 \) and volatility equal to \( \sigma \). Let \( r \) be the risk-free rate.

**Formula 19 (Black-Scholes Formula for a Call).** The price \( h \) of a European call option with expiration date \( T \) and strike price \( K \) is given by

\[
    h = e^{-rT}(S_0 e^{rT} F_Z(d_1) - K F_Z(d_2)) = S_0 F_Z(d_1) - K F_Z(d_2) e^{-rT}
\]

where

\[
    d_1 = \frac{\ln S_0 + (r + \frac{\sigma^2}{2}) T}{\sigma \sqrt{T}}, \quad d_2 = \frac{\ln S_0 + (r - \frac{\sigma^2}{2}) T}{\sigma \sqrt{T}}.
\]

The case of the European put with the same specifications follows from put-call parity.

**Formula 20 (Black-Scholes Formula for a Put).** The price of a European put with strike \( K \) and expiration \( T \) is

\[
    h = K F_Z(-d_2) e^{-rT} - S_0 F_Z(-d_1).
\]

Recall that \( F_Z(x) \) is the cumulative distribution function for the standard normal variable \( Z \).

If the option is on a stock which pays dividends or has a carrying cost, then as before we replace the stock price with the effective stock price, \( S_0^{eff} \), introduced in Section 5.3. For example, if the present value of the carrying cost is \$U, then \( S_0^{eff} = S_0 + U \) and the Black-Scholes formula for a European call with the same specifications as above becomes

\[
    h = (S_0 + U) F_Z(d_1) - K F_Z(d_2) e^{-rT}
\]
where
\[ d_1 = \frac{\ln S_0 + U + (r + \frac{\sigma^2}{2})T}{\sigma \sqrt{T}}, \quad d_2 = \frac{\ln S_0 + U - (r - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}}. \]

Note that one must replace \( S_0 \) with \( S_0^{\text{eff}} = S_0 + U \) where it appears in \( d_1 \) and \( d_2 \).

### 11.5 Digital options

In this section, we begin the process of verifying the Black-Scholes formula. Although we do not complete the verification, we derive a formula for yet another derivative along the way.

Consider any derivative whose pay-off at time \( T \) is given by \( g(S_T) \). From Chapter 10, the price \( h \) is
\[
h = e^{-rT} \mathbb{E}_{RNW}[g(S_T)] = e^{-rT} \mathbb{E}_{RNW}[g(S_0 e^{(r-\sigma^2/2)T + \sigma \sqrt{T} Z})] = e^{-rT} \int_{-\infty}^{\infty} g(S_0 e^{(r-\sigma^2/2)T + \sigma \sqrt{T} z}) e^{-z^2/2} \sqrt{2\pi} dz
\]

We have already seen how to compute this price for some examples of payoffs \( g \), but in general, the integral can be quite hard to compute. Rather than carrying the computation out for the European call, where \( g(z) = \max(z - K, 0) \), we consider a simpler yet similar derivative

**Example 11.7.** A digital option pays the holder $1 at some future time if the stock at that time lies above a certain threshold, and nothing otherwise. This can be considered as some “pared down” version of a European call with strike \( K \), as the option pays money when the stock price is above \( K \), and is worthless otherwise. Consider the stock whose spot price is \( S_0 \) and volatility is \( \sigma \). Let \( r \) be the risk-free rate. Compute the price of a digital option which pays $1 at time \( T \) if \( S_T \) is above the threshold \( K \) and nothing otherwise.

What must our standard normal variable be for the stock price to pass the threshold?

\[
S_T > K \iff S_0 e^{(r-\sigma^2/2)T + \sqrt{T} \sigma Z} > K \iff Z > \frac{\ln \frac{S_0}{K} - (r - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}} = -d_2
\]
So our payoff function is
\[ g(S_0 e^{(r-s^2/2)T+\sigma \sqrt{T}Z}) = 0 \] if \( Z \leq -d_2 \) and
\[ g(S_0 e^{(r-s^2/2)T+\sigma \sqrt{T}Z}) = 1 \] if \( Z > -d_2 \).

Plugging this into the integral for the price of a general derivative,
\[
h = e^{-rT} \int_{-\infty}^{\infty} g(S_0 e^{(r-s^2/2)T+\sigma \sqrt{T}Z}) \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz
\]
\[
= e^{-rT} \left( \int_{-\infty}^{-d_2} 0 \times \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz + \int_{-d_2}^{\infty} 1 \times \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \right)
\]
\[
= e^{-rT} \int_{-\infty}^{d_2} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz
\]
\[
= e^{-rT} F_Z(d_2).
\]

Note how this is similar to the second term in the Black-Scholes formula for a European call, \(-KF_Z(d_2)e^{-rT}\), which represents the \$K the call holder pays (instead of \$1 the digital option holder receives) to exercise the call should the stock price \( S_T \) pass the threshold \( K \).

### 11.6 Problems

1. The current spot price of a stock is $50, the expected return of the stock is 7%, and the volatility of the stock is 20%. The risk-free rate is 5%. Find an 80%-confidence interval for the stock price in 3 months. Compute the expected percent change in stock between months 0 and 6.

2. The current spot price of a stock is $34, the expected return of the stock is 8%, and the volatility of the stock is 20%. The risk-free rate is 3%. Compute the price of a European call option on the stock with strike price $35 expiring in 4 months.

3. With the same parameters as the previous problem, compute the price of a straddle on the stock with strike price $35 expiring in 4 months.
4. Verify that the Black-Scholes formulas for the call and put satisfy put-call parity.

5. The current spot price of a stock is $21, the expected return of the stock is 11%, and the volatility is 12%. The risk-free rate is 5%. Compute the price of a derivative whose payoff in 5 months is

\[ \ln(S_{5/12}) + 32 \]

where \( S_{5/12} \) is the stock price in 5 months.

6. The current spot price of a stock is $20, the expected return of the stock is 10%, and the volatility of the stock is 25%. The risk-free rate is 4%. Compute the price of a derivative whose payoff in 4 months is

\[ \ln((S_{4/12})^{5}) + (S_{4/12})^{0.441} + 32 \]

where \( S_{4/12} \) is the stock price in 4 months.

7. If the example in Section 11.5 is a pared down version of a European call, describe a similar digital option which represents a pared down put. Price it, and formulate a “digital version” of put-call parity comparing the price of these two derivatives.

8. The current spot price of a stock is $34, the expected return of the stock is 8%, and the volatility of the stock is 20%. The risk-free rate is 3%. Compute the price of a derivative whose payoff in 6 months is

- $8 if the stock price in 6 months, \( S_{6/12} \), is below $35,
- $5 if \( 35 \leq S_{6/12} \leq 55 \), and
- nothing otherwise.

9. Assuming the log-normal price for a stock

\[ S_T = S_te^{\sigma \sqrt{T-t}Z + (\mu - \frac{\sigma^2}{2})(T-t)} \]

verify the formula

\[ SD(S_T) = S_0 e^{\mu T} \sqrt{e^{\sigma^2 T} - 1} \]
10. Note that the Black-Scholes formula gives the price of European call $c$ given the time to expiration $T$, the strike price $K$, the stock’s spot price $S_0$, the stock’s volatility $\sigma$, and the risk-free rate of return $r$:

$$c = c(T, K, S_0, \sigma, r).$$

All the variables but one are “observable,” because an investor can quickly observe $T, K, S_0, r$. The stock volatility, however, is not observable. Rather it relies on the choice of models the investor uses. The price of the option, $c$, if traded, is observable.

So we can flip the problem around. Given observables $T, K, S_0, r$ and $c$, what volatility $\sigma$ should the stock have in order for the Black-Scholes formula to be correct. This is called the \textit{implied volatility}, $\sigma_{BS}$. Some calculus, shows that $\sigma_{BS}$ exists and is unique.

The current spot price is $40$, the expected return of the stock is 8%, the risk-free rate is 3%. A European call option on the stock with strike price $40$ expiring in 4 months is currently trading for $2$. Estimate by trial and error the implied volatility of the stock.

Hint: Start with a guess of 20%. If the formula gives a price that is lower (higher) than the market price, increase (decrease) your guess.
Chapter 12
Hedging

The notes so far have focused on pricing financial instruments such as bonds, forwards, and options. People borrow needed funds or invest surplus money, which explains the necessity for why bonds. But aside from speculative gambling, why does the market have forwards and options?

Financial mathematics would be easier if everyone purchased/sold what they wanted/had when they wanted/had it. In the example of Section 5.1, why did the grain farmer enter into a short position in February with May delivery, instead of waiting until May to sell the grain? The farmer did this in order to lock in the price. Essentially, the farmer hedged, or reduced, some of the risk, or uncertainty, associated to fluctuating grain prices. We shall see in Section 12.1 how forwards can be used to hedge against price fluctuations in a more elaborate way.

We also present several hedges available to the financial investor, which vary based on the investor’s portfolio and the type of risk the investor wishes to hedge. In Section 12.2, we show how a bond investor can use other bonds to hedge against the risk of changing interest rates. In Section 12.3, we show how an equity investor can use put options to hedge against the risk of changing prices values. And in Section 12.4, we show how an option writer can use stocks (and cash) to hedge against the option risk.

In each of these investment cases, the hedge is only temporary, and the strategy has to be adjusted at some other time if the investor wants to continue to reduce the exposure to risk. Of course, the investor can simply sell all of his or her assets to eliminate all risk now and forever. However, this might incur too many transaction costs. The idea is to provide a temporary hedge against risk (perhaps the some period of excess uncertainty is approaching)
without completely “unwinding” all of the positions.

Aside from the hedge in Section 12.3, all of these hedges are imperfect in that they only provide a model for partially offsetting a certain risk. A main cause of the credit crisis in 2008 was that the risk associated to many of the investments was inadequately hedged.

12.1 Hedging stock positions with forwards

In Section 5.3, we saw that

\[ F_0 = S_0 e^{rT} \]

where \( F_0 \) is the forward price of an asset to be delivered on \( T \geq 0 \), and \( S_0 \) is the asset’s spot price. According to this formula, the spot price and forward price moves are completely synchronized. The spot price (instantaneously) increases by $1 if and only the futures price increases by $1e^{rT} at the exact same time.

In reality, although the change in spot price and forward price are often assumed to be very correlated, they are not perfectly so. If the spot price increases, economists believe that the corresponding forward prices often rise as well, but not always.

We illustrate with an example how a consumer can take advantage of this correlation to reduce unwanted exposure to fluctuating prices.

**Example 12.1.** On June 10th, ABC Co knows that it needs 1000 barrels of oil in late October or early November. The forward contract size of oil is 100 barrels. The earliest delivery date following the latest time ABC needs oil, is December 1. ABC lacks the facilities to store that much oil for several months. Suppose on June 10th the spot price of oil is $110 a barrel and the futures price of 100 barrels of oil to be delivered December 1 is $12000

\[ S_{\text{June 10}} = 110, \quad F_{\text{June 10}} = 12000. \]

ABC could wait until it needs the oil, but with such fluctuations in oil spot prices, it does not want to run the risk that oil may be too expensive in October or November. So ABC preforms the following hedging strategy.

- On June 10, ABC takes a long position in 10 December oil contracts (1000 barrels).
• On November 3, ABC decides it wants the oil. Suppose the spot price of oil went up, and, not too surprisingly, the forward price of December oil went up as well (not necessarily by the same percentage)

\[ S_{Nov \ 3} = 135, \quad F_{Nov \ 3} = 14000. \]

• On November 3, ABC closes its contract by taking a short position in 10 forward contracts.

• On November 3, ABC buys the 1000 barrels of oil it needed at the spot price.

What was the overall price for the 1000 barrels of oil? To see this, we must compute all the cash flows. On November 3, ABC paid \(1000 \times 135 = 135000\). On December 1, ABC received

\[ 10 \times (14000 - 12000) = 20000 \]

dollars. If we ignore the time value of money, which is a reasonable approximation since November 3 is close to December 1, the overall cash flow for the barrels was

\[ -135000 + 20000 = -115000 \]

dollars.

So ABC paid $115000, which is an amount closer to the June 10 prices, \(1000 \times 110 = 110000\), than the November 3 prices, \(1000 \times 135 = 135000\).

There are two points to make in the above example.

The first is that this was not a perfect hedge. ABC had to pay a bit more for its oil in November than it could have paid had it bought the oil in June (assuming no storage costs). ABC was not fully hedged against changing oil prices, it had only reduced its exposure. This imperfect hedge is due to the fact that the forward and spot prices are not perfectly correlated. In the example above, the December forward price did not increase by the same percent as did the spot price. If the forward price had increased more, a computation would show that ABC would have paid less in November than it could have in June.

The second point is an important one. A hedge does not necessarily mean a price reduction. In the above example, ABC benefitted with a positive net
cash flow of $20000 from its short and long forward positions. This $20000 partially offset the $25000 increase in the price of required oil. If the oil spot price had gone down, then most likely ABC would have lost money from its short and long forward positions. The upshot is that by hedging with forward contracts, ABC mitigated the impact of the changing oil spot prices, for better or for worse. This mitigation is the point of a hedge.

12.2 Hedging bond positions with bonds

In Section 2.9, Formula 7 gave an estimate for the change in the bond price if there was a parallel shift in zero curve. If $B$ is the original bond price (or price of a portfolio of bonds), $D_B$ its duration, $\Delta r$ the (assumed) uniform shift in all zero rates, and $\Delta B$ the resulting change in price, then

$$\Delta B \approx -BD_B \Delta r.$$ 

Suppose $A$ denotes the value of another investment based on bonds with duration $D_A$, where again

$$\Delta A \approx -AD_A \Delta r.$$ 

Suppose now an investor already has portfolio $B$, does not want to be exposed to the risk of changing interest rates for some short time, but does not want to off-load his or her position in Portfolio $B$. The investor takes a position in $N$ of portfolio $A$. The investor’s new duration, by Formula 10 is

$$\frac{BD_B + NAD_A}{B + NA}.$$ 

Choose

$$N = \frac{-BD_B}{AD_A}$$

and the new duration is 0, which by Formula 7 implies the new portfolio (approximately) does not change when the zero rates undergo a parallel shift. This is called duration hedging.

We defer the numerical examples to the Problems.

---

1. $A$ can also be an interest rate futures, which is not discussed in these notes.
Clearly this is an imperfect hedge. Formula \[7\] is only an approximation. We already saw a better approximation using the bond’s convexity; thus, duration and convexity hedging would provide a better hedge. However, there is an infinite sequence of such corrections, one for each term in the Taylor series. An “infinite hedge” would be too costly in terms of transaction costs.

Another shortcoming is that we assume all the zero rates shift by the same amount, which certainly need not be the case. For example, sometimes the zero curve increases with maturity dates (“normal”) while sometimes it decreases (“inverted”). The curve cannot change from one to the other using only parallel shifts.

12.3 Hedging stock positions with options

One of the original purposes of options was to provide a hedge, or insurance, against future negative events. We give an explicit example of how to do this. The Black-Scholes formula then provides a price on this “insurance policy.” Unlike with previous examples, this is a perfect hedge of the potential downside associated to the upcoming uncertainty.

**Example 12.2.** Suppose a fund manager has a portfolio which mirrors the S&P 500 Index. The fund is worth $600 million. The current S&P Index is priced at $1000. S&P volatility is estimated to be 15% per annum. The risk free rate is 5%.

The manager wants to insure against the fund dropping below $480 million in the next three months by buying European puts on S&P. What can the manager do? What is the price of this insurance?

**Solution:** Pretend the fund is
\[
\frac{600,000,000}{1000} = 600,000
\]
S&P shares. So the fund drops below $480 million if and only if 600,0000 shares of S&P drop below $480 million, if and only if one share of S&P drops below
\[
\frac{480,000,000}{600,000} = 800
\]
dollars.
If the manager buys 600,000 European puts, each with strike $800 and expiring in 3 months, then the value of the portfolio (not including the expense of buying the puts) cannot drop below $480,000,000.

For example, suppose the S&P dropped to $750 in 3 months. The fund is worth

$$600,000 \times 750 = 450,000,000 < 480,000,000$$

dollars, but the puts can be exercised for a payoff of

$$600,000 \times (800 - 750) = 30,000,000$$

dollars, bringing the fund total back up to the $480,000,000 minimum.

To compute the price of this insurance, plug in the values into the Black-Scholes formula above and multiply the answer by 600,000.

### 12.4 Hedging option positions with stocks

Suppose a financial institution writes a call option on a stock. The institution faces a significant, arbitrarily high, downside if the stock should significantly increase in value. How can the institution hedge this risk, without buying another option?

Let \( h \) denote the price of the option and \( S \) denote the price of the stock. It is reasonable to assume \( h \) is a function of \( S \), \( h = h(S) \). For example, if the call option is European, we can use the Black-Scholes formula in Section 11.4

\[
h = h(S) = SF_Z(d_1) - KF_Z(d_2)e^{-rT}
\]

treating the other factors, \( r, K, T, \sigma \), as constants.

Assume the stock price changes by \( \Delta S \), and no other factors change. In particular, this means that the stock price rapidly changed, since we are assuming that the time-to-maturity, and hence real time, has not changed. Let \( \Delta h \) denote the resulting change in the option price. Using the first Taylor polynomial, we get

\[
\Delta h \approx \frac{dh}{dS} \Delta S.
\]

The factor \( \frac{dh}{dS} \) is known as the delta of the option. It measures the sensitivity of the option price to the stock. It is the first of several such “measurements,” commonly known as the Greeks, which we do not discuss in these notes.
Note that we do not assume the Black-Scholes model in this discussion. However, if we were to assume the Black-Scholes formula for \( h(S) \), we could compute the delta \( \frac{dh}{dS} \), taking care to remember that \( S \) appears in the formula for \( d_1 \) and \( d_2 \). Assuming the Black-Scholes formula,

\[
\Delta = F_Z(d_1).
\]

Once the bank knows the delta of its call, it can hedge against the risk associated with changing stock prices by trading the underlying stock. We illustrate such delta hedging with a numerical example.

**Example 12.3.** A bank writes an option whose price \( h = 5 \). The underlying asset has a spot price 50. Bank analysts estimate the option to have a delta of \(-0.2\).

1. Estimate the price of the option if the stock price immediately changes to 51.

2. What should the bank do if it wants to hedge such a fluctuation over the short term?

**Solution:**

The price of the option changes to

\[
h_{\text{new}} \approx h + \Delta = h + \frac{dh}{dS} \Delta S = 5 + (-0.2) \times (51 - 50) = 4.80.
\]

Suppose the bank takes a position in \(-\Delta = 0.2\) shares of the asset. Then the bank’s overall portfolio has two changes:

- A change of \( 4.80 - 5 = -0.20 \) due to the decrease in the option value.
- A change of \( 0.2 \times (51 - 50) = 0.20 \) due to the recent increase in its asset holding.

We assume the bank can buy fractions of the asset. In reality, the bank’s position in options is sufficiently large that this hedging strategy can be approximated by an integer value of shares.
CHAPTER 12. HEDGING

The net affect is no change in value for the bank. The delta of its new portfolio is zero; we say that it has become **delta-neutral**.

To recap delta-hedging, for each option with delta $\Delta$, take a position in $-\Delta$ of the underlying share.

Just as we compute the delta assuming the Black-Scholes model (see the exercises), we can also compute the delta of an option assuming the binomial tree model. We illustrate how to do this in the one-time step case. Recall the notation from Chapter 9. The current price is $S_0$, the price in one-time step will be $S_0u$ or $S_0d$. The option in those two cases will be $h_u$ or $h_d$. Then the delta is

$$\Delta = \frac{dh}{dS} \approx \frac{h(S_0u) - h(S_0d)}{S_0u - S_0d} = \frac{h_u - h_d}{S_0u - S_0d}$$

where here we estimate the derivative (sensitivity to stock price change) using a quotient difference. Note that this is exactly the $N$ we computed in Section 9.1. Thus we were already delta-hedging when turning the option into a risk-free portfolio by adding $-N$ shares. Examples appear in the exercises.

12.5 Problems

1. Sometimes a forward hedge actually exacerbates instead of mitigate price fluctuations. Suppose a potato farmer will harvest 10000 pounds of potatoes in August and September, and can store them (essentially) cost free. Suppose the delivery dates of potato contracts are Jan 1, May 1, July 1 and Oct 1 of each year. The spot price of potatoes is $800 per 1000 pounds. All forward prices of potatoes are $825/1000 pounds.

   (a) Describe a hedging strategy for the farmer which best follows the hedging example in this chapter.

   (b) Suppose all 10000 pounds are picked by September 15. The spot price has risen to $900/1000 pounds. The forward price has changed to $x/1000 pounds. What is the total revenue that the farmer receives for the potatoes? Assume all interested rates are 0.

   (c) For what range of $x$ does the hedge actually exacerbate instead of mitigate the price fluctuations?

2. In the example of Section 12.3, compute how much the find manager would have to pay for this insurance made of puts.
3. The spot price of a stock is $40. Over the next six months, the stock will go up to $45 or down to $36. The risk-free rate is 5%. Compute the deltas of call options with the following strike prices: 37, 38, 39, 40, 41, 42, 43, 44.

4. Let $\Delta_c$ be the delta of a European call. Let $\Delta_p$ be delta of a European put with the same specifications.

   (a) Without assuming the Black-Scholes model, show that
   \[ \Delta_p = \Delta_c - 1. \]

   (b) Explain in words why the delta of any European call is positive and the delta of any European put is negative.

   (c) Given (b), explain algebraically why the delta of any European call is less than 1.

5. Verify that if we assume the Black-Scholes formula for the price of a European call $h(S)$, then delta is $\Delta = F_Z(d_1)$. 