Introduction to Mathematical Finance
(Math 537 at UMass, Amherst)

Eric Sommers and Michael Sullivan

August 31, 2016
Contents

1 Interest Rates 7
  1.1 Rate of return 7
  1.2 Interest rates 8
  1.3 Interest rate conventions 9
  1.4 Continuous compounding 11
  1.5 Effective annual rate 13
  1.6 Continuous compounding, a second look 14
  1.7 Time value of money 16
  1.8 Zero Rates 17
  1.9 Force of interest 19
  1.10 Forward Rates 21
  1.11 Problems 22

2 Survey of markets 25
  2.1 Instruments for raising money 25
    2.1.1 Bonds 26
    2.1.2 Other loans 27
    2.1.3 Equity 28
  2.2 Derivatives 29
    2.2.1 Swaps 30
    2.2.2 Credit default swaps 30
    2.2.3 Forwards 31
    2.2.4 Futures 31
    2.2.5 Options 32
  2.3 Trading on a market versus over-the-counter (OTC) 33
  2.4 Selling short 34
  2.5 Problems 36
### CONTENTS

3 Bonds

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Factors affecting interest rates</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>Risk-free rate</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>Portfolio valuation</td>
<td>42</td>
</tr>
<tr>
<td>3.4</td>
<td>Bonds</td>
<td>43</td>
</tr>
<tr>
<td>3.5</td>
<td>LIBOR and Treasuries as measures of the risk-free rate</td>
<td>46</td>
</tr>
<tr>
<td>3.6</td>
<td>Zero coupon bonds and more lingo</td>
<td>48</td>
</tr>
<tr>
<td>3.7</td>
<td>Forward rate relationship</td>
<td>49</td>
</tr>
<tr>
<td>3.8</td>
<td>Bond duration</td>
<td>53</td>
</tr>
<tr>
<td>3.9</td>
<td>Parallel shift in the zero curve</td>
<td>55</td>
</tr>
<tr>
<td>3.10</td>
<td>Convexity of a bond</td>
<td>57</td>
</tr>
<tr>
<td>3.11</td>
<td>Yield</td>
<td>58</td>
</tr>
<tr>
<td>3.12</td>
<td>Portfolio of bonds</td>
<td>61</td>
</tr>
<tr>
<td>3.13</td>
<td>Problems</td>
<td>63</td>
</tr>
</tbody>
</table>

4 Swaps

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Bootstrapping</td>
<td>67</td>
</tr>
<tr>
<td>4.2</td>
<td>Floating-rate bonds</td>
<td>69</td>
</tr>
<tr>
<td>4.3</td>
<td>Fixed-floating swaps</td>
<td>73</td>
</tr>
<tr>
<td>4.4</td>
<td>Deducing swap rates</td>
<td>76</td>
</tr>
<tr>
<td>4.5</td>
<td>Currency swaps</td>
<td>78</td>
</tr>
<tr>
<td>4.6</td>
<td>Credit default swaps</td>
<td>80</td>
</tr>
<tr>
<td>4.7</td>
<td>Problems</td>
<td>80</td>
</tr>
</tbody>
</table>

5 Forward contracts

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>What is a forward contract?</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>The forward price versus the spot price</td>
<td>84</td>
</tr>
<tr>
<td>5.3</td>
<td>Computing forward prices</td>
<td>85</td>
</tr>
<tr>
<td>5.4</td>
<td>Valuing a long or short position in a forward contract</td>
<td>88</td>
</tr>
<tr>
<td>5.5</td>
<td>Problems</td>
<td>89</td>
</tr>
</tbody>
</table>

6 Options

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Definitions</td>
<td>91</td>
</tr>
<tr>
<td>6.2</td>
<td>Some examples</td>
<td>93</td>
</tr>
<tr>
<td>6.3</td>
<td>Payoff of a portfolio</td>
<td>95</td>
</tr>
<tr>
<td>6.4</td>
<td>Some common portfolios with options</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>6.4.1 Straddle</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>6.4.2 Spreads</td>
<td>97</td>
</tr>
</tbody>
</table>
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5</td>
<td>The fundamental principle of math finance</td>
<td>99</td>
</tr>
<tr>
<td>6.6</td>
<td>Inequalities for option prices</td>
<td>100</td>
</tr>
<tr>
<td>6.7</td>
<td>Put-call parity</td>
<td>103</td>
</tr>
<tr>
<td>6.8</td>
<td>Dividends and carrying costs</td>
<td>105</td>
</tr>
<tr>
<td>6.9</td>
<td>Problems</td>
<td>106</td>
</tr>
<tr>
<td>7</td>
<td>Probability and Statistics I</td>
<td>113</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>113</td>
</tr>
<tr>
<td>7.2</td>
<td>Definition of probability</td>
<td>114</td>
</tr>
<tr>
<td>7.3</td>
<td>Random variables</td>
<td>117</td>
</tr>
<tr>
<td>7.4</td>
<td>Expectation for a finite sample space</td>
<td>118</td>
</tr>
<tr>
<td>7.5</td>
<td>Variance and standard deviation</td>
<td>120</td>
</tr>
<tr>
<td>7.6</td>
<td>Covariance and correlation</td>
<td>121</td>
</tr>
<tr>
<td>7.7</td>
<td>Gathering statistics on data sets</td>
<td>122</td>
</tr>
<tr>
<td>7.8</td>
<td>Problems</td>
<td>124</td>
</tr>
<tr>
<td>8</td>
<td>Binomial trees</td>
<td>127</td>
</tr>
<tr>
<td>8.1</td>
<td>One-step model</td>
<td>128</td>
</tr>
<tr>
<td>8.1.1</td>
<td>A call option on a one-step tree</td>
<td>128</td>
</tr>
<tr>
<td>8.1.2</td>
<td>An example of an exotic derivative on a stock</td>
<td>130</td>
</tr>
<tr>
<td>8.1.3</td>
<td>Pricing formula for any derivative on a 1-step tree</td>
<td>131</td>
</tr>
<tr>
<td>8.2</td>
<td>Risk-neutral world</td>
<td>134</td>
</tr>
<tr>
<td>8.3</td>
<td>More than one time step</td>
<td>136</td>
</tr>
<tr>
<td>8.4</td>
<td>American options</td>
<td>140</td>
</tr>
<tr>
<td>8.5</td>
<td>Options with dividends</td>
<td>142</td>
</tr>
<tr>
<td>8.6</td>
<td>Incomplete markets</td>
<td>143</td>
</tr>
<tr>
<td>8.7</td>
<td>Problems</td>
<td>143</td>
</tr>
<tr>
<td>9</td>
<td>Probability II</td>
<td>149</td>
</tr>
<tr>
<td>9.1</td>
<td>Probability Density</td>
<td>149</td>
</tr>
<tr>
<td>9.2</td>
<td>Expectation</td>
<td>153</td>
</tr>
<tr>
<td>9.3</td>
<td>The normal variable</td>
<td>155</td>
</tr>
<tr>
<td>9.4</td>
<td>Conditional probability and independence</td>
<td>160</td>
</tr>
<tr>
<td>9.5</td>
<td>Central Limit Theorem</td>
<td>162</td>
</tr>
<tr>
<td>9.6</td>
<td>Problems</td>
<td>164</td>
</tr>
</tbody>
</table>
10 The Black-Scholes formula
10.1 The efficient market hypothesis ........................................ 168
10.2 The log-normal model ................................................... 168
10.3 Volatility and expected rate of return of a stock .................... 172
10.4 Risk neutral pricing ...................................................... 175
10.5 Proof of the Black-Scholes formula .................................... 179
10.6 Digital options .......................................................... 180
10.7 Problems .................................................................... 181

11 The Greeks
11.1 Delta ........................................................................ 185
   11.1.1 Delta for a call option and for a put option .................... 186
   11.1.2 Properties of Delta .................................................. 188
   11.1.3 Delta hedging ........................................................ 189
11.2 The other Greeks ......................................................... 190
11.3 Greeks for binomial trees ............................................... 192
11.4 Problems .................................................................... 193

12 Stochastic calculus
12.1 Brownian motion .......................................................... 198
12.2 A review of the chain rule ............................................... 202
12.3 Itô’s Chain Rule: a first look .......................................... 205
12.4 Stochastic differential equations and revisiting Itô’s rule ......... 208
12.5 Problems .................................................................... 213

13 The Black-Scholes partial differential equation
13.1 Deriving the Black-Scholes PDE ...................................... 218
13.2 More on the Black-Scholes PDE ...................................... 222
   13.2.1 Black-Scholes and the Greeks .................................. 222
   13.2.2 Dividends ........................................................... 223
   13.2.3 Hedging generalizations ......................................... 224
13.3 Problems .................................................................... 225
Chapter 1

Interest Rates

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 called 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

$$r = \frac{Q - P}{P} = \frac{Q}{P} - 1 \quad (1.1)$$

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

$$Q = P(1 + i).$$

(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

$$Q = P(1 + i)^N$$

(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 ac-
count is

\[ 500(1 + 0.06)^N \]

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

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

(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.\(^1\) 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

\[
600 \left( 1 + \frac{0.045}{4} \right)^4 = 627.46
\]

\(^1\)Recall that a positive integer is a number of the form 1, 2, 3, 4, \ldots
1.4. CONTINUOUS COMPOUNDING

dollars. After 1.25 years, the balance is

\[ 600 \left( 1 + \frac{0.045}{4} \right)^5 = 634.52 \]

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 of .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

\[ \left( 1 + \frac{r}{2} \right)^2 = \left( 1 + \frac{0.03}{4} \right)^4. \]

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'Hôpital'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

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

Taking the limit as \( n \) goes to infinity now gives the indeterminate form \( 0/0 \), the typical candidate for L'Hôpital's Rule. To make life easier we set \( u = 1/n \). Then as \( n \) goes to infinity, \( u \) goes to zero and our limit becomes

\[
\lim_{n \to \infty} \ln y = \lim_{u \to 0} \frac{t \ln(1 + ur)}{u}.
\]

So we apply L'Hôpital's Rule by taking the derivative of the numerator and dividing by the derivative of the denominator (both derivatives are with respect to \( u \)) and we get

\[
t \cdot \frac{1}{1+ur} \cdot r
\]

after simplifying. Taking the limit as \( u \) goes to zero, 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

\[
Q = Pe^{rt}
\]

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

\[
1300e^{0.03 \cdot \frac{9}{12}} - 1300e^{0.03 \cdot \frac{6}{12}} = 1329.58 - 1319.65 = 9.94
\] 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 \( r_{eff} \), 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 \( r_{eff} \), we find the growth from continuous compounding after 1 year, which is \( Pe^r \). Then we solve the equation

\[
P(1 + r_{eff}) = Pe^r,
\]

which gives \( r_{eff} = e^r - 1 \). In the previous example, the effective annual rate

\[
e^{0.03} - 1 = 0.0305 = 3.05\%
\]

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

$$1 + r_{\text{eff}} = \left(1 + \frac{0.03}{4}\right)^4.$$

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.

One piece of terminology is commonly used. In these situations where the per annum interest rate $i$ is different from the effective annual rate, we call the rate $i$ the nominal interest rate. This use of the word nominal alerts us to the distinction between the true annual rate (the effective annual rate) and the rate $i$ which is being used according to some compounding convention (whether discrete as in Section 1.3 or continuous as in Section 1.4).

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 common value.

Then an investment of $P$ dollars grows to

$$Pe^{rt} = P(e^r)^t = Pa^t$$

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.
1.6. CONTINUOUS COMPOUNDING, A SECOND LOOK

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

$$Q = Pa^t$$  \hspace{1cm} (1.7)

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 will only earn 3% after one year 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 $5100(1.03)^{2/52} = 5105.80$. The interest earned in the two week period is $5.80$. Note that it is not correct to compute the interest as $5100(0.03)^{2/52}$. The latter value equals 20.28. Here, we have used Formula 1.6 with $t = 2/52 = 0.3846$ and we have assumed that interest is computed continuously over the week period.

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 Time value of money

In Section 3 we discuss why different borrowers might expect different interest rates based on their credit-worthiness, which measures how likely they are to pay back their loan as oppose to fully or partially defaulting on it. In this section, however, we assume that no such risk is present. And as such we called the (universal to all) rate the risk-free (interest) rate.

One of the key concepts of this course is that the existence of a risk-free rate determine the current value of a future payment.

Suppose that the risk-free rate now and forever is 6% per annum with continuous compounding. You are given a choice between \( P \) and $100 in a year. What should \( P \) to make them mutually attractive? You could put the \( P \) in a bank and let it accrue interest over one-year. You end up with \( Pe^{0.06} \) dollars in one year. So since you get \( Pe^{0.06} \) dollars following this option, and it should be no more attractive than waiting one year for your $100, then

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

But maybe you have other choices with your \( P \) in hand now, and can do better than \( Pe^{0.06} \) dollars in one year if you pursue the right one. It turns out that under mild assumptions, any different result will involve some risk, where you might end up with more than $100 but you might end up with less. So in fact

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

To see that you are just as flexible waiting for the money, you know that you can borrow \( $100e^{-0.06} \) today and pay back the loan with with the $100 you
receive in the future. Now you can do whatever you like with the $100e^{-0.06}$ today (just as you can with the $P$). This shows that

$$94.18 = 100e^{-0.06} = P$$

which is another derivation of the equation we justified before.

The idea that $100$ in one year is equivalent to $94.18$ today (under the assumption that the risk-free rate is 6% per annum with continuous compounding) is called the time-value of money, sometimes abbreviated as TVM.

**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 1.8.** 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 $5000 + 3000e^{(-0.045)(0.5)} = 7,933.25$. The second opportunity has a present value of $8100e^{(-0.045)(0.5)} = 7,919.79$. Thus the first opportunity is worth more today and should be chosen.

## 1.8 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 1.9.** 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 invest money in 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 was offering a 1-year CD with an effective annual rate of 3.3%. The 6-month rate was 2.6% and the 2-year rate was 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 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 1.10.** 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

---

2You 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.
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 $100e^{0.0353(2)} = 107.32$.

1.9 Force of interest

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

$$Pe^{r(t_1)\Delta t},$$

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 $T$, 

$$Pe^{\int_0^T r(t)dt}.$$
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
\[ Pe^{\int_0^T r(t) dt} \]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
\[ r(t) = \frac{Q'(t)}{Q(t)}. \] (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 \( Q(0) = P \).

**Example 1.11.** 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 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 1.12.** 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 1.13.** 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%. 
1.11 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? Do the problem assuming the interest rate is the effective rate with continuous compounding and assuming it is the effective rate with annual compounding.

6. What annual rate of interest will cause an investment to double in 10 years? Do the problem assuming the interest rate is the effective rate with continuous compounding and assuming it is the effective rate with annual compounding.

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?
1.11. PROBLEMS

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 nominal 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

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

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 $i > 0$ leads to a higher effective annual rate versus compounding twice per year at the same rate.

(b) Show if $n > m$, then compounding $n$ times per year at the nominal rate $i > 0$ always leads to a higher effective annual rate versus compounding $m$ times per year.

Hint: we need to show $(1 + \frac{i}{n})^n > (1 + \frac{i}{m})^m$. This is equivalent to showing $\ln((1 + \frac{i}{n})^n) > \ln((1 + \frac{i}{m})^m)$ since $\ln(x)$ is an increasing function. Now consider the function

\[ F(i) = \ln((1 + \frac{i}{n})^n) - \ln((1 + \frac{i}{m})^m). \]

Show that $F'(i) > 0$ when $i > 0$ and that $F(0) = 0$. Then use your calculus skills to conclude that $F(i)$ is an increasing function and therefore positive for $i > 0$. 

15. A $10,000 business loan that begins today will need to be repaid in 9 months, and no sooner. The balance due on the loan at that time will be $10,260. What is the implied 9-month zero rate? In other words, what value of $r_{0.75}$ would imply that the loan balance grows to this amount. Express your answer in per annum continuous compounding. (If you haven’t read about zero rates for this class, you can assume the question asks for $r$, the risk-free rate with continuous compounding.)

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

A survey of financial instruments and markets

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.

2.1 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. If you sell short the stock of a company, then you are speculating.

2.1.1 Bonds

Governments, corporations, and non-profit institutions use bonds as a way to raise money. The bond is essentially a loan that these entities take out. But instead one person, or counterparty, making the loan to the entity, 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 in the fall of 2008 where bondholders were reported to receive about 9 cents for each dollar of bonds that they held 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-
2.1. 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.

2.1.2 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.

2.1.3 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.\footnote{Dividend payouts are set by companies when they report their quarterly earnings. In the fall of 2008 many financial companies slashed their dividends in order to conserve cash. As another example, in the summer of 2010 BP canceled its generous dividend following the outcry over its oil spill in the Gulf of Mexico.}

\footnote{As of this writing in October 2008.}
Stock can be issued as common shares or preferred shares. Preferred 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.

2.2 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, 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. Although it now appears that even these derivatives will now be regulated to some extent following the passage of the Financial Reform Bill in the summer of 2010.

---

2 Depending on the type of swap, it may not be considered a derivative.
2.2.1 Swaps

A swap is a generic name for any agreement to exchange payments, usually at regular intervals. The payments that are exchanged are determined by the initial agreement. For example, a bank and a company may enter a swap. The bank agrees to pay the company a fixed 5% annual interest rate on a fixed amount of money each year, while the company agrees to pay the bank a floating interest rate (determined by some index) on the same amount of money. The agreement might last several years. Companies enter into such agreements to manage their interest rate exposure.

A **floating interest rate** or **variable interest rate** is one that changes over time. An example of such a rate is the interest on a checking account, which could possibly change every day. Another example is the prime rate, which is used to compute the interest on many loans.

2.2.2 Credit default swaps

Despite its name, a credit default swap (CDS) is not really a swap, but more like a form of insurance. A CDS is a contract between two parties, A and B, where A agrees to pay B a fixed periodic payment. These payments last either a fixed pre-determined number of times, or until a pre-specified third party C defaults on its obligations (usually on a bond it wrote). When such a default occurs, A ceases its payments to B, and B pays A a lump sum.

Note that a CDS is not really insurance, since A does not need to suffer any lose in order to receive money from B. (Imagine you, party A, bought car insurance from party B, but on a car belonging to someone else unrelated to you, party C.)

There are other significant differences between insurance and CDSs. Party B does not have to maintain collateral in the event C defaults. Party A does not need to disclose to party B the risks involved. CDSs were a root problem of the economic collapse in 2008 because Party B was often unaware that Party A would try to manipulate markets to force party C to default. (Imagine you secretly but legally trying to cause party C to get into a car accident.)
2.2.3 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 exchange something in the future at a pre-determined price. See Chapter 5 for more details on forwards.

2.2.4 Futures

A futures contract (or future) on an asset is a more advanced way to handle the idea behind a forward. A bit of terminology: a holder of a long position in a futures contract on, say, gold with a delivery date in June is agreeing to buy gold in June at whatever the price of gold is at the time the futures contract expires. The holder of a short position is agreeing to sell gold in June at whatever the price of gold is at that time. Notice that the price is not fixed ahead of time; instead the two parties are essentially agreeing to exchange the gold at the later date at whatever the price is at that time.

Now the futures contract has a price, called the futures price, and it can be considered the price to buy (or sell) the asset at the later time. But this futures price will change and as the time for delivery of the asset draws near, the futures price will become the actual price, called the spot price, of the asset at that moment.

Futures are traded on an exchange, just like stocks are. Unlike a stock, however, the investor does not actually buy or sell anything when entering into a position in a futures contract. Instead the investor leaves money in a margin account with a broker. The price of the futures contract varies throughout the day and if the investor does not close out the position, then at the end of the day, the futures contract is marked-to-market; that is, money is taken from the margin account by the broker to cover the loss on the contract for that day. Or conversely, money is put into the margin account by the broker to represent the gain on the contract for the day. In a sense, it is as if the investor sells the contract at the end of each day, and
then immediately buys it back for the next day.

The futures contract 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. Most futures contracts, however, are closed out before this happen, meaning that if an investor holds a long position in a contract, the investor will go to the futures exchange and take a short position in the same contract. Then the long position and short position cancel each other out and the investor is free of any obligations. Because the investor can avoid actually buying or selling the asset, it is possible for an investor/speculator to participate in the futures market without being a producer or user of the asset.

Most commodities are traded in the futures market, including gasoline, oil, heating oil, wheat, corn, and gold. In addition, financial instruments are also traded on the futures market including futures on stock indices such as the S&P 500 and Nasdaq 100 as well as bonds and interest rates. Currencies are also traded in the futures market, although they are more commonly traded by forward contracts with delivery dates in the very near future.

### 2.2.5 Options

Options are similar to forward contracts except there is an option involved; one side of the transaction can choose whether to complete the transaction or not. There are two situations depending upon whether the buyer of the asset has the option or the seller does.

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 follow through with the option to buy the asset, but may if she chooses. If she chooses to buy the asset, this is called exercising the option. The investor on the opposite side of the transaction, the one 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 not known to each other but 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
2.3. TRADING ON A MARKET VERSUS OVER-THE-COUNTER (OTC)

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.

2.3 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). Stocks that
are listed on one exchange can trade on other exchanges. There are several
smaller exchanges where stocks also trade, such as BATS and DirectEdge.
High-frequency traders often exploit small differences in the price of a stock
on each exchange. When the price of a stock on one exchange differs greatly
from its price on another this can be problematic and this was certainly a
factor in the ‘flash crash’ on May 6, 2010. Surprisingly perhaps, many stock
trades do not even take place on a stock exchange: an article in the Wall
Street Journal reported (September 2010) that, in fact, 30% of all stock trades
take place in so-called ‘dark pools’ or over-the-counter in private transactions.

Futures in the U.S. trade on the CME (Chicago Mercantile Exchange),
NYMEX (New York Mercantile Exchange), or the ICE (Intercontinental Ex-
In 2007, CME acquired CBOT, its Chicago competitor since the 1800’s.

Options trade on several exchanges, with the largest being the CBOE
(Chicago Board of Options Exchange). NYSE and Nasdaq-OMX are also
big players in options trading.
2.4 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.

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.

---

3 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 little terminology: after you sell short an asset, you are said to have a **short position** in the asset. This will be listed in your brokerage account, so that while you have acquired the cash from selling the asset, you have also acquired this short position. The short position is a liability and will have a negative value.

**Example 2.1.** NVIDIA stock (NVDA) is trading at $10 today. Jen sells short 100 shares of NVDA today. This means she now has $1000 in her account, but she has the obligation to return 100 shares of NVDA. The obligation is a liability of exactly $1000, so her total account is worth nothing.

Next, suppose that in 1 week NVDA goes up to $11. Jen still have $1000 in cash, but her obligation to return the 100 shares is now a liability of $11 \times 100 = 1100$ dollars. Her account is now actually worth $-100$ dollars. Her broker will call her up and ask her to deposit 100 dollars (or perhaps more). Let’s suppose Jen deposits this 100 dollars in her brokerage account.

Finally, suppose that in 2 weeks NVDA is trading at $8. What is Jen’s account worth? First, she has $1100 in cash after her recent deposit. Her short position in the stock is now worth $-800$. Her account is therefore worth $300$. Suppose she decides to return the stock she borrowed, that is, she closes out her short position in the stock. To do this she spends $800 to buy 100 shares of NVDA, returns the borrowed shares, and her account now contains $300$ in cash and nothing else.

**Example 2.2.** 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 semiannual 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.

2.5 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). In other words, in 3 months, the owner of a share of stock will receive $2 dollars, and again in 6 months, 9 months, etc.

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. (This means you can assume the risk-free rate is 0%). Suppose the stock trades at $52 in three months time and $49 in six months time. Suppose that Sue buys the stock back in 7 months time 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 semiannually, 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.


   (a) Write down the date and time you are doing this.
2.5. PROBLEMS

(b) What is the price of Apple stock (AAPL)?

c) How many shares have exchanged hands so far today (this is called the **volume**)?

d) Look at the call options on Apple stock with expiration in February 2011 (this refers to the third Friday of February). Explain the meaning of the strike price of an option. How does it differ from the number in the third column (marked Last)?

e) Write down the prices for all the options with strike prices starting at 250 and going up by multiples of 10 to the price of the stock, using the Last column. Why should a call option with strike 250 be worth more than one with strike 260?

4. Search on the web to find:

(a) the U.S. stock in the S&P 500 that has the highest dividend yield

(b) the nickname for the major stock index in: India, Korea, France, UK, the country where you were born.

(c) the most actively traded stock in the U.S. on the day you are looking.

(d) what is SPY? what is QQQQ?

(e) the exchange rate between Japanese Yen (JPY) and U.S. dollars (USD)

5. Go to the CME website (www.cmegroup.com).

(a) Find the (most common) abbreviation for futures contracts on the following underlying assets: E-mini S&P 500 in dollars, Heating Oil, Light Sweet Crude Oil

(b) Look up the difference between cash settlement and physical settlement of a futures contract. Give an example of a futures contract that is cash settled and one that is physically settled.

(c) Look up the price for the front contract of NG and GC. The front contract refers to the one with the closest expiration date. This price is usually fairly close to the actual price that these assets trade for today.
Chapter 3

Bonds

In this section we introduce the risk-free rate 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.

3.1 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.

### 3.2 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
3.2. RISK-FREE RATE

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 3.1.** 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? How much would be owed if $100 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 of $100 would lead to an obligation to pay $100.85after six months 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.

### 3.3 Portfolio valuation

How do you 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 [4.1] 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

$$B = \sum_{i=1}^{n} P_i e^{-r_i t_i} \quad (3.1)$$

**Example 3.2.** 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:

$$200(e^{-0.03(25)} + e^{-0.03(5)} + e^{-0.03(75)} + e^{-0.03(1)}) = 785.17.$$
Example 3.3. 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:

\[ 200(e^{-0.03(25)} + e^{-0.03(5)} + e^{-0.0325(75)} + e^{-0.033(1)}) = 784.02. \]

3.4 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 3.4.** 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 3.5.** 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 4.4 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 3.6.** 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.

3.5 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 $r_{25} = .04$ 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
3.5. LIBOR AND TREASURIES AS MEASURES OF THE RISK-FREE RATE

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.

### 3.6 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:

$$98 = 100e^{-r(.5)} \Rightarrow r = \frac{\ln(100/98)}{.5} = .0404.$$

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 3.7.** 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 3.8.** 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.

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

Recall from Section 1.10 that forwards rates are interest rates promised today for investments that start at some time \( T_1 \) and last until a later time \( T_2 \). 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 \( r_{5,1.5} \), 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 \( Pe^{(0.05)(2)} \) 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:

\[ P e^{0.05(2)} = P e^{0.04+r_{1,2}}. \]

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,

\[ P e^{0.05(2)} > P e^{0.04+r_{1,2}}. \]

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 \( P e^{0.04} \) 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 \( P e^{0.05(2)} \), while our loan debt is now \((P e^{0.04}) e^{r_{1,2}}. \) Since the investment was assumed to be worth more than the amount owed, we have earned a guaranteed profit of

\[ P e^{0.05(2)} - P e^{0.04+r_{1,2}} > 0. \]

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
3.7. FORWARD RATE RELATIONSHIP

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 $P_{e^{.05(2)}} > P_{e^{.04+r_{1,2}}}$ cannot 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 $P_{e^{.05(2)}} < P_{e^{.04+r_{1,2}}}$ cannot hold either. Thus, we have given a complete argument to explain the equality

$$P_{e^{(.05)(2)}} = P_{e^{.04+r_{1,2}}},$$

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 giving 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 last 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
CHAPTER 3. BONDS

$T_2$, which last until time $T_3$ at the forward rate $c = r_{T_2, T_3}$. Then the investment grows to $P e^{a(T_3 - T_1)}$ and the loan grows to $P e^{b(T_2 - T_1) + c(T_3 - T_2)}$. Hence, the profit (or loss) for the combined portfolio is $P e^{a(T_3 - T_1)} - P e^{b(T_2 - T_1) + c(T_3 - T_2)}$. Since arbitrage opportunities do not exist, we cannot have $P e^{a(T_3 - T_1)} > P e^{b(T_2 - T_1) + c(T_3 - T_2)}$.

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 $P e^{a(T_3 - T_1)} < P e^{b(T_2 - T_1) + c(T_3 - T_2)}$ is not possible since arbitrage opportunities do not exist.

We conclude that

$$P e^{a(T_3 - T_1)} = P e^{b(T_2 - T_1) + c(T_3 - T_2)}.$$

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 3.9.** 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, $r_{1, 4.5} = \frac{0.021 \times 2 + 0.036 \times 1.5}{3.5} = 2.74\%$

**Example 3.10.** 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 3.1

$$B = 30e^{-0.04(0.25)} + 130e^{-0.05(0.75)} = 154.92,$$

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

\[ 0.05(0.75) - 0.04(0.25) = 0.05. \]

After 9 months, we have received a total of $30e^{0.05(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.

### 3.8 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 3.11.** 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

\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}
\]  

(3.2)

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

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

(3.3)

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

\[
\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} P_i e^{-r_i t_i} = B.
\]
3.9. PARALLEL SHIFT IN THE ZERO CURVE

Example 3.12. Calculate the duration of the bond in Example 3.6 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.

3.9 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

\[
B_{\text{new}} = \sum_{i=1}^{n} P_i e^{-(r_i + \Delta r) t_i} = \sum_{i=1}^{n} P_i e^{-r_i t_i} e^{-\Delta r t_i}.
\] (3.4)
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

$$B_{new} \approx \sum P_i e^{-r_i t_i} (1 - \Delta r t_i)$$

$$= \sum P_i e^{-r_i t_i} - (t_i P_i e^{-r_i t_i}) \Delta r$$

$$= B - BD \Delta r.$$ 

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

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

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 3.13.** 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 3.14. Calculate both the exact change in price and the approximate change in price of the bond in Example 3.12 if all zero rates move down by 12 bips.

Solution: The new relevant zero rates are 
\[0.018 - 0.0012 = 0.0168, 0.0148, \text{ and } 0.0198,\]
so the new price of the bond is
\[250e^{-1.0168(1/12)} + 250e^{-0.0148(7/12)} + 10250e^{-0.0198(13/12)} = \$10,529.98,\]
and thus the exact change in price is 
\[10,529.98 - 10,516.75 = 13.23 \text{ dollars.}\]

On the other hand, the approximate change is
\[-BD\Delta r = -(10,516.75)(1.0478)(-0.0012) = 13.223 \text{ dollars, which is essentially the same and easier to compute once you know the duration.}\]

3.10 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 3.4. If we approximate \(e^x\) to second order by
\[e^x \approx 1 + x + \frac{x^2}{2},\]
then we find that
\[B_{new} \approx \sum P_i e^{-r_i t_i} \left(1 - \Delta r \, t_i + \frac{1}{2} t_i^2 (\Delta r)^2 \right) = B - BD\Delta r + \frac{1}{2} BC(\Delta r)^2,\]
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 3.15. Continuing with the Example 3.14, 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

\[
\frac{(1/12)^2250e^{-0.018(1/12)} + (7/12)^2250e^{-0.016(7/12)} + (13/12)^210250e^{-0.021(13/12)}}{B} = 1.126.
\]

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.

3.11 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 3.16 (Yield). The yield \( y \) of a bond with a price of \( B \) is the value \( y \) which solves the equation:

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

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 3.1. 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 3.16 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 3.17.** 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 3.16 we can view the bond price as function of its yield $y$:

$$B(y) = \sum_{i=1}^{n} P_i e^{-yt_i}. \quad (3.7)$$

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:

$$D \approx \sum_{i=1}^{n} \frac{t_i P_i e^{-y t_i}}{B}.$$ 

With this approximation in mind, we see that

$$\frac{1}{B} \frac{dB}{dy} = -\frac{\sum_{i=1}^{n} t_i P_i e^{-y t_i}}{B} \approx -D,$$  \hspace{1cm} (3.8)

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:

$$\frac{\Delta B}{B} \approx -D \Delta y.$$  \hspace{1cm} (3.9)

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 3.18. Estimate the duration of the bond in Example 3.17 and use the yield-duration approximation to estimate the change in the bond price when the yield moves to 1%.
3.12. PORTFOLIO OF BONDS

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

\[
D = \frac{0.3333 \cdot 10 e^{-0.00381(0.3333)} + (1.3333) \cdot 10 e^{-0.00381(1.3333)} + (2.3333) \cdot 10 e^{-0.00381(2.3333)}}{1021},
\]

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

Since \( \Delta y = 0.01 - 0.038 = 0.0062 \), we get

\[
\Delta B = -BD\Delta y = -(1021)(2.304)(0.0062) = -14.58
\]
dollars.

We conclude the section by showing why the yield is well-defined (that is, why Equation 3.16 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 3.16 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 \( B \).

3.12 Portfolio of bonds

As discussed in Section 3.3, 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 \( B = \sum P_i e^{-r_i t_i} \) and duration \( D = \sum P_i t_i e^{-r_i t_i} \), while the second bond has a price of \( B' = \sum P'_i e^{-r'_i t'_i} \) and duration \( D' = \sum P'_i t'_i e^{-r'_i t'_i} \), 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:

\[ \frac{\sum P_i t_i e^{-r_i t_i} + \sum P'_i t'_i e^{-r'_i t'_i}}{B + B'} \]

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 3.19.** 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 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 \( \Delta r = -.0015 \), then \( \Delta B = -(3028)(2.34)(-.0015) = 10.63 \) dollars.
3.13. PROBLEMS

3.13 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 6 set $T_1 = 0$ (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 3.6).

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 3.14 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 3.6 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 3.10 called “Convexity?”
Chapter 4

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.

4.1 Bootstrapping

As we saw in Section 3.6 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 4.1. 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 is $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:

\[ 992 = 1000e^{-r_4 \frac{4}{12}} \Rightarrow r_4 = 0.02410 = 2.41\% . \]

The equation for Bond B’s price is:

\[ 1062 = 40e^{-r_4 \frac{4}{12}} + 1040e^{-r_{10} \frac{10}{12}}, \]

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 4.2. 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

\[ 1001 = 27.5e^{-0.01204(1)} + 27.5.5e^{-0.06525(1)} + 1027.5e^{r_{1.5}(1.5)}, \]

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

### 4.2 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

\[ 1000 \times \frac{.037}{2} = 18.50 \]
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 4.1.

<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 4.1: A 2-year floating-rate bond, with semiannual coupons

In the example from Table 4.1 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

\[ Pe^{rT} - P = P(e^{rT} - 1), \]

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
compounding, then the next coupon will be $1000(e^{r(5)} - 1)$, 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

\[ 1000 + 1000(e^{r'_6(6/12)} - 1) = 1000e^{r'_6(6/12)}, \]

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 \( 1000(e^{r'_6(6/12)} - 1) \) and then we will be left with a bond worth 1000. The total sum is \( 1000e^{r'_6(6/12)} \). We can repeat the above argument and deduce that the price of the bond is

\[ [1000e^{r'_6(6/12)}] \times e^{-r'_6(6/12)} = 1000 \]

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 4.3.** 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 4.4. 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 4.5. 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

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

4.3 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\textsuperscript{1}. 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

\[ 1000000 \times \frac{.04}{4} = 10000 \]

dollars, while Jefferson County pays out

\[ 1000000 \times \frac{.028}{4} = 7000 \]

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

\[ 1000000 \times \frac{.056}{4} = $14000. \]

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

\textsuperscript{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 2.4 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 4.6.** 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 4.2 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.
Example 4.7. 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 $B_{\text{fixed}} - B_{\text{float}}$, where $B_{\text{fixed}}$ is a fixed-rate bond with 22 months until maturity and $B_{\text{float}}$ 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

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

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

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

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.

4.4 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
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 4.8.** 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:

$$P = \frac{Pc}{2}(e^{-0.04(5)} + e^{-0.045(1)}) + Pe^{-0.045(1)}.$$
78  

CHAPTER 4. SWAPS

The \( P \) cancels out and we get

\[
c = 4.55\%.
\]

4.5 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

\[
V_{\text{swap}} = B_{US} - S_0 B_J.
\]
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 4.9.** 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 4.10.** 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

\[
V_{\text{swap}} = B_{US} - S_{BE} = -0.67.
\]

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

\[
V_{\text{swap}} = B_{E} - S' B_{US} = 0.52.
\]

### 4.6 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.

### 4.7 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 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 \( S_t \) denote its spot price at time \( t \). 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 positions: 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 \( T = .5 \) years later, the portfolio is worth

\[-80e^{0.04(5)} + F_0,\]

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

\[ F_0 \leq 80e^{0.04(5)} = 81.62. \]

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 \( 75 + 5 = 80 \) dollars. Carrying through the analysis for this new portfolio, we deduce the opposite inequality and hence that

\[ F_0 = 80e^{0.04(5)} = 81.62. \]

This example again leads to the following generalization.

**Formula 13. (Forward price with costs and income)** Let \( I \) be the present value of all income that the asset pays between the time the forward contract is negotiated and the delivery time \( T \). Let \( C \) be the present value of all costs to hold the asset in the same time period. Then,

\[ F_0 = (S_0 - I + C)e^{rT}. \]

One way to encompass all of these variations of the forward price formula is to define the **effective spot price**. The effective spot price is the value of the asset at time \( t \) given that you will not actually possess the asset between times \( t \) 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 \( t \) and \( T \); thus, the effective spot price is the price of the stock minus the present value of the dividends to be received. Naturally, the effective spot price is lower 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.
CHAPTER 5. FORWARD CONTRACTS

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

$$76.52e^{-0.05(\frac{5}{12})} = 74.94,$$

and so we get the equality:

$$90 + f_s = 76.52e^{-0.05(\frac{5}{12})} = 74.94,$$

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

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 costumer 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

$$F \leq 60e^{0.05 \times \left(\frac{18}{12} - 1\right)}.$$ 

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. 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?

8. 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?

9. 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 2.2.5 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 of an option is said to have a short position. The seller is also called the **writer** of the option, because in order for an option contract to be created, there needs to be someone 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 15,233rd contract. Options are also commonly written on equity indices and on futures contracts.

Let $S_t$ denote the price of a stock a time $t$. Suppose that an option expires at time $T$. At the current time $t$, where $t < T$, 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 refers to the fact that when an American option is out-the-money, there is no reason to exercise the option since it will entail a guaranteed loss. Conversely, if the option is
in-the-money, exercising the option will definitely yield some positive cash flow. This is not to suggest that exercising the option is the best course of action when the option is in-the-money. 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.

It is important to keep in mind that for the owner of an American option, there are always three possible courses of action at any any time. The owner can do nothing (keep holding the option). The owner can sell the option and receive the premium that the option is currently trading for. Finally, the owner can exercise the option.

On the other hand, for the writer of an American option, there are only two courses of action. Either the writer can keep holding the option or the writer can buy the same type of option contract. In the latter case, the investor no longer owns any options; we say the investor has closed out her short position in the option. Of course, if the option is assigned before she buys an offsetting contract, then the writer will be forced to sell the stock at $K$ (for a call option) or to buy the stock at $K$ (for a put option).

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 and the value of money at different times, 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 again 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 put 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 put 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. Consider a long position in a call option with expiration time $T$. A person with this portfolio has the option to buy the stock for $K$ at time $T$ (European) or up until time $T$ (American). Let $S_T$ be the price of the stock at the expiration time $T$.

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 of the long position in the call option is

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

In other words, this formula gives the value of the call option at expiration.
Note that the writer of the call option has the exact opposite payoff

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

**Example 6.2.** 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 6.3.** 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 \( S_T - K \).

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. To be totally accurate, when we compute the profit, we should take the future value of the initial cost, but often we will omit this step to simplify the discussion.

### 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. SOME COMMON PORTFOLIOS WITH OPTIONS

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 10.

**Example 6.4.** Consider a straddle with strike price 50 expiring in $T = 0.75$ years that was constructed from a call option that cost 2 dollars and a put option that cost 3 dollars. Assume the risk-free rate is zero. What values of $S_T$ will lead to a profit on the straddle?

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

$$|S_T - 50| - 5.$$

Therefore a profit is earned whenever

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

For example, if the current stock price were 49 (so the put was in-the-money and the 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.

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, $S_0 < K_1$, where $S_0$ denotes the spot price of the stock (i.e., the current stock price). 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 6.5.** 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 contain some examples involving butterfly spreads.

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 6.6. 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 cannot happen.

---

1Technically, we 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 \( T \).
Applying this principle twice provides a special case that we will often use: if the value of portfolio A exactly equals that of portfolio B at time $T$, then their current values are equal.

**Example 6.7.** Consider two call options on the same stock with the same expiration $T$. Call option A has strike price $50$ and call option B has strike price $60$. If $S_T \geq 60$, then the payoff for A, which is $S_T - 50$, will definitely be more than the payoff for B, which is $S_T - 60$. If $50 < S_T \leq 60$, then the payoff for A is positive, but the payoff for B is zero. Finally, if $S_T \leq 50$, both payoffs are zero. In all cases, the payoff for option A is greater than or equal to the payoff for option B. It follows that price of option A today must be at least as much as the price of option B.

### 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 6.8.** Show that for a put option with strike $K$ and expiration at time $T$ that

$$p \geq Ke^{-rT} - S_0$$

where $p$ is price of the put option at the present time and $r$ is the $T$-year zero rate.

**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 $K$ at time $T$ 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 $S_T$. Hence this portfolio will be guaranteed to be worth the bigger of the two numbers $K$ or $S_T$.

Next consider a second portfolio with $Ke^{-rT}$ dollars in cash. Then this portfolio will be worth $(Ke^{-rT})e^{rT} = K$ 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 long the option in the first portfolio and the holder of the long position gets to decide whether to exercise the option or not if the put is American. 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, we could make a guaranteed by buying the option and immediately exercising it). And this is a better inequality assuming that \( r > 0 \), which is typical. In other words, for an American put with value \( P \), we have
\[ P \geq K - S_0 > Ke^{-rT} - S_0. \]

We can improve the inequalities. Neither the American or the European put is ever negatively priced, so \( p, P \geq 0 \). Also neither put can be worth more than \( K \) when exercised; such a maximum payoff occurs if the stock price is zero. Combining these observations with the previous inequalities, we get
\[ Ke^{-rT} \geq p \geq \max(Ke^{-rT} - S_0, 0) \]
\[ K \geq P \geq \max(K - S_0, 0). \]

To prove the upper bounds rigorously, we would construct a portfolio consisting of a short position in the put and \( Ke^{-rT} \) dollars to prove the first upper bound; and with \( K \) dollars to prove the second.

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 \( S_0 - K e^{-rT} \) 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 on the stock and this may be advantageous.

**Example 6.9.** IBM stock trades at $100. The 1-year risk-free zero rate is 3%. A European call option on IBM with strike $90 expiring in one year, costs costs $10. Construct an arbitrage opportunity.

**Solution:** The lower bound on the call option that we just obtained is violated:

\[
100 - 90e^{-0.03 \times 1} = 12.66 > 10.
\]

To take advantage of this, we sell short the stock, buy the call option, and invest the remaining \( 100 - 10 = 90 \) dollars for one year. In one year, we withdraw the cash (which is now \( 90e^{0.03 \times 1} \)) and buy the stock to close out the short position by 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 \( S_1 \) is the unknown price of IBM in one year.

We can also use the Fundamental Principle with various option strategies. Let’s compare European options with different strike prices as we did in the first example above, but this time using a spread.

**Example 6.10.** Compare the price \( c_1 \) of a European call with strike price \( K_1 \) and expiration \( T \) with the price \( c_2 \) of a European call with strike price \( K_2 > K_1 \) and the same expiration \( T \).
6.7 PUT-CALL PARITY

Solution: Consider the portfolio which is long the first call and short the second 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 zero; thus, the present value of the portfolio, \(c_1 - c_2\), must be also greater than or equal to zero:

\[c_1 \geq c_2.\]

A similar argument using either a bull spread or bear spread with European puts shows that

\[p_2 \geq p_1\]

by considering puts instead of calls in the example. 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:

- Portfolio \(A\) consists of 1 European put option on a stock with strike \(K\) and expiration \(T\) and 1 share of the stock

- Portfolio \(B\) consists of 1 European call option on a stock with strike \(K\) and expiration \(T\) and \(Ke^{-rT}\) 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:

\[p + S_0 = c + Ke^{-rT},\]
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:

\[
K \geq P - C + S_0 \geq Ke^{-rT}
\]

(see the exercises).

**Example 6.11.** 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.

**Solution:** Let Portfolio A be the call and \( Ke^{-rT} \) dollars. 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:

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

Portfolio B is overpriced compared to Portfolio A. So buy Portfolio A and sell Portfolio B. Then right now we receive $33.25 - 32.26 = $0.96. In 3 months, no matter what happens, our cash-flow is \( \max(S_T, K) - \max(S_T, K) = 0 \), so we are guaranteed a risk-free profit.

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

**Example 6.12.** 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

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

\footnote{We do not know if \( B \) is overpriced or \( A \) underpriced. We only know that \( B \) is overpriced compared to \( A \).}
Apply put-call parity to both sides of the first inequality

\[ p_1 + S_0 - K_1 e^{-rT} \geq p_2 + S_0 - K_2 e^{-rT}. \]

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. 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. Because of this drop, it might be better to exercise the option 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 the dividends (or perhaps other reasons). 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

\[ Ke^{-rT} \geq p \geq \max(K e^{-rT} - S_0^{\text{eff}}, 0). \]
6.9 Problems

For these problems, all interest rates are the same (which in the language of Chapter 3 means a flat yield curve) and are expressed in terms the effective annual rate compounded continuously, which we refer to this as the risk-free rate.

1. Pfizer stock (PFE) is trading at $17.40 today. Consider a European call option on 1 share of PFE with strike $K = 18$ and expiration in $T = .25$ years. The call is selling today for $0.44$. Assume $r = 0.01$.

   (a) You buy one of these calls. In 3 months at expiration, PFE is trading at $19$. What is the call worth at expiration? What is the only course of action at expiration (assuming you want to maximize your profit)? What is your profit or loss?

   (b) Repeat the problem if instead PFE trades at $17$ at expiration. What about if it trades at $8$ at expiration?

   (c) Now instead suppose that today you write one of the calls. What will happen if PFE is trading at $19$ at expiration and what course of action will you have to take? What is your profit or loss? Repeat the problem if PFE trades at $17$ at expiration.

2. Delta stock (DAL) is trading at $11.70$ today. Consider an American put option on 1 share of DAL with strike $K = 12$ and expiration in $T = .25$ years. The put is selling today for $1.46$. Assume the risk-free rate is 1%. Recall that American options can be exercised at any time by the holder of the option.

   (a) You buy one of these puts. You get confused because you did not study your 537 notes closely enough and you decide to exercise the put right away. What is the payoff from exercising the option (that is, what amount of money do you receive from exercising the option)? What is your profit or loss? Was the exercise a mistake?

   (b) Actually you did study your 537 notes and you did not exercise the option. Instead 1 month later, the put option is trading at $0.85$. However, you have had enough of 537 and finance in general and so you decide to sell your put option. What is your profit or loss?
(c) Actually you really enjoy 537 and finance in general. You also
decide to hold the option until expiration. At that time DAL is
trading at $10. What is the payoff of the put option? What is
your profit?

3. 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 $S_T$.

4. 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, and C, respectively, today (time 0). Suppose the price on the
stock 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 at time $T$ taking into consider-
ation the time value of the initial cost. By how much would the
asset have to change over the year for your butterfly portfolio to
be a loss?

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

(a) Use put-call parity to compute the premiums of the corresponding
European puts, $A', B', C'$?

(b) Graph the pay-off and profit functions for the straddle using call
$B$ and put $B'$.

6. Assume the risk-free rate is 2%. 
(a) A portfolio is guaranteed to be worth exactly $5 in 1 month. What is the portfolio worth today? What name have we given to a financial instrument that behaves the same as such a portfolio?

(b) A portfolio is guaranteed to be worth at least $5 in 1 month. What can you say about the price of the portfolio today?

(c) In the previous part, describe an arbitrage if the portfolio is trading at $4 today. What is your guaranteed profit from the arbitrage?

7. 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, and $1.00 for the strikes at 50, 60, 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? Assume the risk-free rate is 0%.

(c) Explain, using the fundamental principle, why the spread must have a positive value now. (Your argument should hold for a general butterfly spread, not just for the special case of premiums given in part (b).)

(d) Deduce that the current price of the call with \( K = 60 \) is at most the average of the current prices of the other two calls.

8. 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

\[
P - C + S_0 \geq Ke^{-rT}.
\]
9. 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:

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

10. 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. Show that

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

using the fundamental principal.

11. 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.
12. 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?

13. 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.

14. 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.

(a) Show that if $r = 0$, then the portfolio is guaranteed to be worth at least $P - p$ at time $T$. Deduce that $P = p$. Hint: Draw a timeline from 0 to $T$. Plot the different cash flows under different scenarios.

(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.
15. 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 \( \frac{1}{2} \) 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 \(i = 1, 2, \ldots, 5, 6\) is number on the blue die and \(j = 1, 2, \ldots, 5, 6\) 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 \( \frac{1}{36} \) 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 $\Omega$ (the Greek letter omega). In the coin toss example from the previous section,

$$\Omega = \{\text{Heads, Tails}\},$$

and in the roll of a pair of dice,

$$\Omega = \{(i, j) \mid i \in \{1, 2, \ldots, 6\} \text{ and } j \in \{1, 2, \ldots, 6\}\}.$$ 

For the experiment recording the price of a stock at 4 p.m. on November 14, $\Omega$ 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, $\Omega$ 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 $\Omega$ 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 $A$.

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 $A \cap B = \emptyset$.

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.

In words, the first condition of the definition 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 ($\Omega$). 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,

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = \frac{6}{36} + \frac{5}{36} - \frac{1}{36} = \frac{10}{36}.$$ 

Alternatively, we could have listed the number of elements in $A \cup B$ and seen that there 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

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

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

$$X(\{(H,H)\}) = 2, \ X(\{(H,T)\}) = 1, \ X(\{(T,H)\}) = 1, \ X(\{(T,T)\}) = 0.$$ 

**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 the random variable $X$, 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?

Solution: The expected return is \( 0.4(10) + 0.6(-8) = -8 \). 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 call options with strike prices 90 and 105. That is, you are long the call with strike 90 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

\[
E[\text{payoff}] = E[\max(S_1 - 90, 0) - \max(S_1 - 105, 0)] \\
= E[\max(S_1 - 90, 0)] - E[\max(S_1 - 105, 0)].
\]

Notice that \(\max(S_1 - 105, 0)\) and \(\max(S_1 - 90, 0)\) are also random variables.

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

\[
E[\max(S_1 - 90, 0)] = 0.4 \times \max(110 - 90, 0) + 0.6 \times \max(92 - 90, 0) \\
= 0.4 \times 20 + 0.6 \times 2 = 9.2
\]

Similarly,

\[
E[\max(S_1 - 105, 0)] = 0.4 \times 5 + 0.6 \times 0 = 2.
\]
Thus the expected payoff is $9.2 - 2 = 7.2$.

Later, we will define the expectation for an infinite sample space. The sum in Definition 7.7 will be replaced by an integral.

### 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 $Var(X)$, is

\[
Var(X) = E[(X - \bar{X})^2].
\]

Using the properties of expectation, it is possible to show that

\[
Var(X) = E[X^2] - \bar{X}^2.
\]

We leave this as an exercise.

Since the random variable $(X - \bar{X})^2$ is always nonnegative, its expectation is also nonnegative. That is, $Var(X)$ is a nonnegative number. This allows us to define:

**Definition 7.12.** The **standard deviation** of $X$, denoted $SD(X)$, is

\[
SD(X) = \sqrt{Var(X)}.
\]

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 $SD(X)$ 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.
7.6. COVARIANCE AND CORRELATION

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 \( SD(X) = \frac{1}{\sqrt{2}} \).

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

\[
\frac{1}{3}(1) + \frac{1}{3}(1000) + \frac{1}{3}(10000) = \$3667.
\]

The standard deviation from this scenario is \$4496.6. On the other hand, the banker’s offers have zero standard deviation, i.e., no risk. If you are offered \$5,000, you should 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 a later 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 **covariance** of $X$ and $Y$, denoted $Cov(X,Y)$, is

$$Cov(X,Y) = E[(X - \bar{X})(Y - \bar{Y})].$$

Notice that $Cov(X,X) = Var(X)$. 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

$$\rho_{X,Y} = \frac{Cov(X,Y)}{SD(X)SD(Y)}.$$

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. As the correlation gets closer to 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 Gathering statistics on data sets

Suppose we have an experiment with possible outcomes $\Omega$ and a probability $p$ on $\Omega$. Suppose that we can run the experiment repeatedly. Let $X$ be a random variable on $\Omega$. 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 first formula since 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

\[
\frac{\frac{1}{n} \left( \sum x_i y_i \right) - \bar{x} \bar{y}}{s_x s_y}.
\]

**Example 7.17.** 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

\[
\frac{\frac{2}{5}(4) + (3)(1) + (-5)(0) + (-4)(-3) + (3)(-1)}{3.554 \cdot 2.6077} - (-.2)(.4) = .519.
\]
7.8 Problems

1. Consider a game where a fair coin is flipped. If the coin comes up heads, you win your bet; if it comes up tails, you lose your bet. At each new flip, however, you must double your bet from the previous flip. You begin with $31 in your bank account. Your initial bet is $1. The game continues until you see a head or when you go negative for the first time.

(a) What is the set of outcomes of this experiment? What is the probability of each outcome?

(b) Let $X$ be the random variable which represents the gain or loss from each outcome. What is the expectation of $X$?

(c) Would you agree to play this game? If you had more money, would you play this game?

2. In a distant land, parents continue to have children until they have a girl and then they stop having kids. Assume there is no limit to the number of births possible to each couple.

(a) What is the expected number of girls in each family?

(b) What is the expected number of boys in each family? (Hint: you might need to remember the formula for summing a geometric series).

3. 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 using the axioms of a probability in Definition 7.1:

\[ p(A \cup B) = p(A) + p(B) - p(A \cap B) \]

for any two events \( A \) and \( B \). (Hint: consider the set \( C = A \cap B \) and also the two sets \( A - C \) and \( B - C \).)

4. 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 \( A^c \)?

5. 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 \( E[Y] \)?

6. 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 7.3. Either of these formulas can be taken as the definition for the variance of \( X \).

7. Show that \( Var(aX + b) = a^2 Var(X) \), using the properties of expectation.

8. 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 \( a = 0 \)?

9. Consider the model of a stock where it either goes up by 7% or down by 4% each year. Let \( p_{up} \) be the probability that stock goes up and \( p_{down} \) the probability that stock goes down.
(a) Suppose that \( p_{\text{up}} = 52\% \) and \( p_{\text{down}} = 48\% \). The **expected rate of return** \( \mu \) is defined as the number \( \mu \) which satisfies the equation \( E[S_1] = S_0 e^{\mu \times 1} \) where \( S_0 \) is the known asset spot price and \( S_1 \) is the unknown asset price in one year from now. This will be revisited later. What is the expected rate of return of the stock?

(b) If the risk-free rate is 5\%, what probabilities \( p_{\text{up}} \) and \( p_{\text{down}} \) lead to an expected return equal to the risk-free rate?

10. 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

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 8.1 with the most basic of trees, known as the one-step tree. We generalize this to the n-step tree in Section 8.3. In Sections 8.4 and 8.5, we describe how to change the models when the options are American, or when the stocks have dividends or carrying costs.

In Section 8.2, 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 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.
8.1 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 8.3 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.

8.1.1 A call option on a one-step tree

To illustrate how to value a call option on a one-step tree, we begin with the following example.

Example 8.1. 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 $Pe^{0.12 \times 3/12}$ 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

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.

8.1.2 An example of an exotic derivative on a stock

Let us extend the discussion of the single call option in the above example to more general derivatives.

Example 8.2. 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

$$\max(S_T - 425, 0) + \max(425 - S_T, 0) = |S_T - 425|.$$
8.1. **ONE-STEP MODEL**

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

$$h_d = \max(420 - 410, 0) + 2 \times 410 + |410 - 425| = 845.$$  

### 8.1.3 Pricing formula for any derivative on a 1-step tree

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 8.1.

Chose Portfolio $A$ to be long $N$ shares and short one such derivative.

- Value of Portfolio $A$ if stock goes up = $S_0uN - h_u$
- Value of Portfolio $A$ 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:

$$S_0uN - h_u = S_0dN - h_d.$$  

Solving for $N$, we get

$$N = \frac{h_u - h_d}{S_0u - S_0d}. \quad (8.1)$$  

With this value of $N$, Portfolio $A$ is guaranteed to be worth

$$S_0uN - h_u = S_0u\left(\frac{h_u - h_d}{S_0u - S_0d}\right) - h_u,$$

at time $T$, using the first expression for the value in Equation 8.1. 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

$$\text{current value of Portfolio } A = e^{-rT}\left(S_0u\left(\frac{h_u - h_d}{S_0u - S_0d}\right) - h_u\right).$$
Figure 8.1: 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.
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,

**Formula 15.** 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 \( S_0u \) and a known value of \( h_d \) when the stock price moves to \( S_0d \). 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 8.3.** Find the current value of the derivative in Example 8.2 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
\[ p = \frac{e^{.07 \times .5} - 1.025}{1.125 - 1.025} = 0.106. \]
Thus
\[ h = e^{-0.07 \times 0.5} (.106 \times 925 + .894 \times 845) = 824.12 \]
dollars.
8.2 Risk-neutral world

We now discuss one of the most important features in mathematical finance: risk-neutral derivative pricing.

Let us revisit Example 8.1. Translate the numbers into the notation for the general case: $u = 22/20 = 1.1$, $d = 18/20 = 0.9$, $r = 0.12$, $S_0 = 20$, $T = 0.25$, $h_u = 1$, $h_d = 0$

\[
p = \frac{e^{0.12 	imes 0.25} - 0.9}{1.1 \times 0.9} = 0.65, \quad 1 - p = 0.35.
\]

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 the expected value of the stock price equals

\[
E[S_T] = pS_0u + (1 - p)S_0d = 0.65 \times 22 + 0.35 \times 18 = 20.60.
\]

But notice all that the future value of $S_0$ using the risk-free rate $r$ is

\[
S_0e^{rT} = 20e^{0.12 \times 0.25} = 20.60,
\]

the same number. Therefore if $p$ represented the probability that the stock will go up and $1 - p$ represented the probability that stock will go down, then the expected (not guaranteed) growth rate of the stock would be the risk-free rate.

In a risk-neutral world, investors are indifferent to risk. That is, an investor seeks the highest return regardless of the risk involved. Investors would borrow money at the risk-free rate and invest in investments with high expected returns. The result would be that these investments with high expected returns would become more costly as investors move money into them and this would have the effect of lowering the potential returns. Eventually, an equilibrium would be reached where all investments have the exact same expected return, namely the risk-free rate.

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 rate. We denote the expectation of an asset 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 events.
Let $P$ denote some cash sitting in a portfolio and earning the risk-free rate $r$. Note that the value of the portfolio at time $T$ is deterministic: we know exactly what the value $P_T$ will be. Thus, in both the risk-neutral world and the real-world

$$E_{RW}[P_T] = P_T = Pe^{rT}, \quad E_{RNW}[P_T] = P_T = Pe^{rT}.$$ 

In the risk-neutral world any stock is expected to grow at this rate as well

$$E_{RNW}[S_T] = S_0e^{rT}.$$ 

Whereas in the real world, the (annual) expected rate of return of the stock $\mu$ is defined by

$$E_{RW}[S_T] = S_0e^{\mu T}.$$ 

Most people believe that $\mu > r$ since there is a risk involved in owning the stock. In order to invest in a risky asset, investors must be compensated for the higher risk. This means that stocks are expected (but not guaranteed) to grow at a rate greater than the risk-free rate.

Returning to our discussion of computing the price of a derivative, let $h_T$ denote the payoff at time $T$ of a derivative on a stock obeying the binomial model. Then $h_T$ is a random variable and we introduced the notation that $h_T = h_u$ if the stock price goes up and $h_T = h_d$ if the stock price goes down. Then translating Formula 15 into the language of this section, we see that the current price $h$ of the derivative is the present value of the 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

$$p = \frac{e^{rT} - d}{u - d},$$

as a quick calculation will show.
Applying this formula to our example
\[ c = e^{-0.12 \times 0.25} E_{RNW} \left[ \max \left( S_{0.25} - 21, 0 \right) \right] = e^{-0.12 \times 0.25} \left( 0.65 \times 1 + 0.35 \times 0 \right) = 0.63 \]
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 8.4.** *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 payoff of the derivative 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 hypothesis, 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 8.1.

### 8.3 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 8.5.** Consider a European put option on Amazon.com Inc stock (symbol: AMZN) which expires in one year, with strike $52. Amazon’s spot
8.3. MORE THAN ONE TIME STEP

Figure 8.2: 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 stock 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 8.2):

$$S_0 ud = S_0 du$$

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 8.2 above, $h_u$ and $h_d$ denote the two (not yet computed) values of the put at $t = 0.5$ and $h_{uu}, h_{ud}, h_{dd}$ 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 8.1. Let

$$p = \frac{e^{r\Delta t} - d}{u - d} = \frac{e^{0.05 \times 0.5} - 0.8}{1.2 - 0.8} = 0.56. \quad 1 - p = 0.44.$$ 

Then our formula from Section 8.1 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 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
8.3. MORE THAN ONE TIME STEP

each of two time periods of length $\Delta t$. Let $r$ be the risk-free rate and assume it holds steady over the entire time period in question. Let $h$ be the current value of a derivative paying $h_{uu}$, $h_{ud}$, and $h_{dd}$ 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

$$h = e^{-r\Delta t}(ph_u + (1-p)h_d)$$

$$= e^{-r\Delta t}(p(e^{-r\Delta t}(ph_{uu} + (1-p)h_{ud})) + (1-p)(e^{-r\Delta t}(ph_{ud} + (1-p)h_{dd})))$$

$$= e^{-r2\Delta t}(p^2h_{uu} + 2p(1-p)h_{ud} + (1-p)^2h_{dd})$$

where

$$p = \frac{e^{r\Delta t} - d}{u - d}.$$ giving a single formula for the two-step model.

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 $S_0u^k d^{n-k}$. 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

$$h = e^{-rT} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} h_{u^k d^{n-k}}$$

(8.3)

The proof is similar to the two-step model using a backwards-in-time induction step to reduce to the $(n-1)$-step model. Again we can interpret this as risk-neutral valuation, since

$$e^{-rT} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} h_{u^k d^{n-k}} = e^{-rT} \mathbb{E}_{RNW}[h_T].$$

In Section 8.4, 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.
8.4 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 8.3.

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 8.3.

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 \( S_0 u = 60 \). 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 8.3 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 \( S_0 u = 60 \),

\[
H_u = \max(52 - 60, 1.72) = 1.72.
\]

Clearly it is not advantageous to exercise early at this node.
Figure 8.3: This tree includes the computed values of the European put from Figure 8.2. Note that at the expiration date, $t = 1$, the American put and European put have the same values.
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 value from the two possible cases at the first node: immediate exercise versus no exercise. If there is immediate exercise, the put is worth \(52 - 50 = 2\) dollars. If we do not exercise, then we proceed as in the one-step tree and value the option. Note that we use \(H_d = 12, H_u = 1.72\) in the formula, 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 \(t = 0.5\).

### 8.5 Options with dividends

In the one-step tree model without dividends, recall the risk-neutral probability \((p, 1 - p)\) was characterized by
\[
p S_0u + (1 - p) S_0d = S_0 e^{rT} = E_{RNW}[S_T].
\]

If the stock had dividends, this would affect the risk-free growth in exactly the same way as it affects the formula for put-call parity: the stock price in the formula needs to be replaced by the effective stock price. For example, if the stock pays a dividend before time \(T\) whose present value is \(D\), then the expected return of the stock at time \(T\) must reflect the pay out of the dividend; thus
\[
p S_0u + (1 - p) S_0d = E_{RNW}[S_T] = (S_0 - D)e^{rT},
\]
Solving for \(p\) we get
\[
p = \frac{S_0 - D}{S_0} \frac{e^{rT} - d}{u - d}.
\]

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.
8.6 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 8.4 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 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 8.1 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.

8.7 Problems

1. A stock is currently priced at $30. In 3 months its price 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 $20. In 4 months its price will either go down 10% or up 20%. The risk-free rate is 5% per annum with continuous compounding.

(a) Compute the premium of a 4-month European put with strike price $19.
(b) Use put-call parity to price a 4-month European call with strike price $19.

(c) Compute the premium of the call in part (b) directly and verify that you get the same answer as in part (b).

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 $25$. In 6 months it will be either $26$ or $30$. The risk-free rate is 12\% per annum with continuous compounding.

(a) Verify the assumption on $u$ and $d$ given following Theorem 8.4.

(b) What is the price of 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 your portfolio have to be risk-free?

6. Let $S_0$ denote the spot price of a stock, which after one-time step $T$ will be either $S_0u$ or $S_0d$ 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?
7. A stock is currently priced at $50. In 4 months its price will be either $48 or $56. The risk-free rate is 5% per annum with continuous compounding.

   (a) Compute the price of a derivative consisting of one European call with strike price $50 and one European put with strike $54. Both expire in 4 months.

   (b) Now assume that the above options are American and price the derivative.

8. For a one-step binomial tree on a stock paying no dividend, show that an American call option is always worth the same as a European call option.

9. A stock price is currently $30. Every 6 months the price will either go 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.

10. 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; and a straddle with strike $35 which is comprised of American options. Compute the price of the Meirkiec.

11. A stock price is currently priced at $25. In 1 year its price 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.
12. Consider a one-step binomial model for a stock. Now, the stock is worth $S_0$, but will be worth either $S_0 u$ or $S_0 d$ 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 $K > S_0 u$. Show that it is always better to exercise the put early when $r > 0$.

(b) Consider an American put with strike price $K$ where $K < S_0 d$. 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.

13. (Hard!) Find the range of strike prices for which the two-step American put from Section 8.4 has an early exercise at the intermediate time $t = 0.5$ but not at the initial time $t = 0$. (The risk-free rate, the up and down factors of the underlying asset price, etc., are all the same.)

14. A stock price is currently priced at $25. Every week it will either go up or down 5%. The risk-free rate is 4% per annum with continuous compounding.

(a) Write a sum representing the price of an at-the-money European put expiring in one year. If programming is not a prerequisite or covered in class, you can leave your answer as a sum.

(b) Write a program to compute the price of an at-the-money American put expiring in one year.

15. 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 $q > p$.

16. Consider a one-step model (from now to time $T > 0$) with up and down factors $u$ and $d$, risk free rate is $r$, and spot price $S_0$. Let $(q, 1 - q)$ represent the up and down probability in the real world.
(a) Compute the price of an at-the-money call option.

(b) Compute the expected (in the real world) profit today of this option.

(c) If your answer in (b) is positive, what can you say about the real world? When is (b) zero?
CHAPTER 8. BINOMIAL TREES
Chapter 9

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.

9.1 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 9.1. 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 9.2.** The cumulative distribution function or c.d.f of the random variable \( Z \) is the function

\[ F_Z : R \to R \]

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 9.3.** 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 \( Z \).\(^1\)

\(^1\)Strictly speaking certain hypotheses on \( Z \) must be satisfied.
Definition 9.4. Let $Z$ be a random variable on $\Omega$. The **probability density function** of $Z$, denoted by $f_Z(x)$, has the defining property that

$$F_Z(x) = \int_{-\infty}^{x} f_Z(y)dy.$$ (9.1)

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,

$$\int_{a}^{b} f_Z(y)dy = p(a \leq Z \leq b).$$ (9.2)

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 $F_Z(x)$

$$\frac{dF_Z(x)}{dx} = f_Z(x).$$

When the sample space $\Omega$ is finite or discrete or the image of $Z$ is finite, the p.d.f. will not be a function in the usual sense. In that case, the proper mathematical idea to introduce is the **probability mass function**. It is defined by

$$f_Z(x) = p(Z = x).$$

Then it is true that

$$F_Z(x) = \sum_{y}^{\infty} f_Z(y),$$

where the sum replaces the integral in the general definition 9.1.

**Example 9.5.** 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 a constant when it is nonzero 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 9.6.** Let \( Y \) be the price of Pepsi stock in 17 days from today. Suppose that 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 9.2

\[
p(48 \leq Y \leq 52) = \int_{48}^{52} f_Y(x) dx = .916 = 91.6\%
\]
\[
p(58 \leq Y \leq 62) = \int_{58}^{62} f_Y(x) dx = 0.0004 = 0.04\%
\]

So Pepsi will probably hover around \$50.
9.2 Expectation

We can now extend the definition of expectation to handle all possible sample spaces \( \Omega \) whether finite or infinite.

**Definition 9.7.** 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 \( X \).

When the sample space is finite or discrete or the image of \( Z \) is finite, it is preferable to use the probability mass function definition of \( f_X(y) \) in the previous definition and take a sum over all \( y \) rather than an integral. This recovers the definition of expectation from Chapter 7, which remains valid in this new context.

**Example 9.8.** 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 9.9. Let $X$ and $Y$ be random variables. Let $g : \mathbb{R} \to \mathbb{R}$. Let $a$ be any constant. Then

- $E[X + Y] = E[X] + E[Y]$
- $E[aX] = aE[X]$
- More generally, the last two properties can be written as:
  \[E[aX + Y] = aE[X] + E[Y]\]

- $E[a] = a$
- $E[g(X)] = \int_{-\infty}^{\infty} g(z)f_X(z)dz$

Example 9.10. 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 price of Coca-Cola today 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 the time-value of money. Hence the expected return is

\[E[30(Y - 49) + 60(W - 80)]\]
\[= 30(E[Y] - 49) + 60(E[W] - 80)\]
\[= 30(50 - 49) + 60 \int_{72}^{92} 0.05xdx - 4800\]
\[= 30 + 4920 - 4800 = 150\]

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
9.3. THE NORMAL VARIABLE

Definition 9.11. The variance of $X$ is equal to

\[ \text{Var}(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2 = \int_{-\infty}^{\infty} y^2 f_X(y) \, dy - \left( \int_{-\infty}^{\infty} y f_X(y) \, dy \right)^2. \]

9.3 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
\[ u = \frac{1}{\sqrt{2\pi}} y, \quad dv = ye^{-y^2/2} dy \]
to get from the second to third line below.

\[
\begin{align*}
\text{Var}(Z) &= E[Z^2] - E[Z]^2 \\
&= E[Z^2] - 0^2 \\
&= \frac{1}{\sqrt{2\pi}} \left. y \cdot \left( -e^{-y^2/2} \right) \right|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \left( -e^{-y^2/2} \right) dy \\
&= 0 - 0 - (-1) = 1.
\end{align*}
\]

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

\[ F_Z(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \]

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
9.3. THE NORMAL VARIABLE

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 9.12.** Find a range of values about 0 in which \( Z \) has a 90% chance of being in that range. This is called a 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 9.2 and Chapter 7,

\[
E[X] = E[\sigma Z + \mu] = \sigma E[Z] + \mu = \sigma \cdot 0 + \mu = \mu
\]

\[
SD(X) = SD(\sigma Z + \mu) = \sigma SD(Z) = \sigma \cdot 1 = \sigma.
\]
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 9.13.** 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 9.14.** 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 9.9. In the third step we complete the squares. In the fourth step we substitute $w$ for $y - t$.

$$
\Phi_Z(t) = \int_{-\infty}^{\infty} e^{ty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy
$$

$$
= \int_{-\infty}^{\infty} \sqrt{\frac{t^2}{2\pi}} e^{-\frac{(y - ty)^2}{2}} dy
$$

$$
= \sqrt{\frac{t^2}{2\pi}} \cdot 1 = e^{\frac{t^2}{2}}.
$$

We could repeat the above integral calculus when computing $\Phi_X(t)$, or we could take a short-cut and use Theorem 9.9 to reduce the work to the previous calculation:

$$
\Phi_X(t) = E[e^{tX}]
$$

$$
= E[e^{t(\sigma Z + \mu)}]
$$

$$
= e^{t\mu} E[e^{t\sigma Z}]
$$

$$
= e^{t\mu} \Phi_Z(t\sigma)
$$

$$
= e^{t\mu} e^{-\frac{(t\sigma)^2}{2}} = e^{\frac{(\sigma t)^2}{2} + t\mu}.
$$

Example 9.15. 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

$$
E[10000e^{tX}] = 10000e^{\frac{(\sigma t)^2}{2} + t\mu}
$$

$$
= 10000 \times 2, \text{ that is}
$$

$$
\frac{(\sigma t)^2}{2} + t\mu = \ln 2
$$
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

\[
p(10000e^{t_*X} \geq 20000) = p(t_*X \geq \ln 2)
= p(t_* (\sigma Z + \mu) \geq \ln 2)
= p \left( Z \geq \frac{\ln 2 - \mu}{\sigma} \right)
= p \left( Z \leq -\frac{\ln 2 - \mu}{\sigma} \right)
= F_Z \left( \frac{\mu - \ln 2}{\sigma t_*} \right)
\]

where \( t_* \) is expressed in terms of \( \mu \) and \( \sigma \) in the previous paragraph.

We return to the moment generating function in Chapter 10.

### 9.4 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 9.16.** 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 \textbf{independent} if

\[ p(A \cap B) = p(A)p(B). \quad (9.3) \]

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 \textbf{independent} if for any numbers \( a, b \)

\[ p(X \leq a \text{ and } Y \leq b) = p(X \leq a)p(Y \leq b). \quad (9.4) \]

\textbf{Example 9.17.} 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

\[ p(A \cap C) \neq p(A)p(C). \]
CHAPTER 9. PROBABILITY II

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.

9.5 Central Limit Theorem

The reason why we focused on the normal variable out of the many possible random variables is due to one of the most important theorems in math finance, and all of probability for that matter: the Central Limit Theorem. We revisit its role in finance in Section 10.2.

We state the simplest version of the Central Limit Theorem.

**Theorem 9.18.** Suppose you have a sequence of random variable \( X_1, X_2, X_3, \ldots \) with the following properties.

- All the random variables are independent from each other.

- All the random variables have the same distribution, that is, probability density function (if it exists). So for example, \( P(X_{17} \leq -23) = P(X_{101} \leq -23) \).

Let \( \mu \) and \( \sigma \) denote the (common) mean and standard deviation of any one of these random variables: for any \( i \), \( E[X_i] = \mu, SD(X_i) = \sigma \). Then as \( n \) approaches \( \infty \), the “average” random variable \( (X_1 + X_2 + \cdots + X_n)/n \) is approximately a normal random variable with distribution \( N(\mu, \sigma/\sqrt{n}) \). This is true regardless of the actual probability density function of the random variables \( X_1, X_2, \ldots \)

An equivalent way to state this theorem is in terms of the sum instead of the average: the random variable \( X_1 + X_2 + \cdots + X_n \) is approximately a normal random variable with distribution \( N(\mu n, \sigma \sqrt{n}) \).

If \( Z \) is a standard normal variable \( Z \sim N(0, 1) \), then we can restate the Central Limit Theorem as an approximation

\[
\frac{X_1 + X_2 + \cdots + X_n}{n} \approx \mu + \frac{\sigma}{\sqrt{n}}Z \quad \text{or} \quad X_1 + X_2 + \cdots + X_n \approx \mu n + \sigma \sqrt{n}Z.
\]
We illustrate the Central Limit Theorem with a specific discrete example.

**Example 9.19.** For \( i = 1, 2, \ldots \), let \( X_i \) be a random variable which has a 1/3 chance of being 1, 1/3 chance of being 2 and 1/3 chance of being 3. Write down the probability density functions for \( \frac{X_1 + X_2}{2} \) and \( \frac{X_1 + X_2 + X_3}{3} \).

**Solution:** For the first problem, \( \frac{X_1 + X_2}{2} \), we consider a sample space with nine outcomes where the pair of random variables \((X_1, X_2)\) can be the pairs

\[(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3).\]

Each outcome has a probability of 1/9 of occurring. If you look at the random variable \( \frac{X_1 + X_2}{2} \) at each outcome, it takes values 2/2=1 and 6/2=3 at one outcome each, 3/2 and 5/2 at two outcomes, and 4/2=2 at three outcomes. To simplify notation, denote \( Y = \frac{X_1 + X_2}{2} \). Thus the p.d.f. of \( Y \) is

\[
f_Y(1) = 1/9, f_Y(3/2) = 2/9, f_Y(2) = 3/9, f_Y(5/2) = 2/9, f_Y(3) = 1/9.
\]

If we repeat the computation for \( Y = \frac{X_1 + X_2 + X_3}{3} \), we see that

\[
f_Y(1) = 1/27, f_Y(4/3) = 3/27, f_Y(5/3) = 6/27, f_Y(2) = 7/27, \\
f_Y(7/3) = 6/27, f_Y(8/3) = 3/27, f_Y(3) = 1/27.
\]

If you plot these p.d.f.'s, and maybe one more for \( \frac{X_1 + X_2 + X_3 + X_4}{4} \), you can see that what started out as flat (the p.d.f. for \( X_1/1 = X_1 \)) tends towards the p.d.f. of some normal variable. We will see in the exercises how to deduce the mean and standard deviations, and use it to approximate distributions of long-run averages.
9.6 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 9.4. 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. (The “return” means the gain, or increase,
if positive, in your portfolio value. If the return is negative it means your portfolio value decreased.) Suppose the p.d.f. of $X$ is given by

$$f_X(x) = \begin{cases} 80000 & \text{if } -50 \leq x \leq 350, \\ 0 & \text{otherwise}. \end{cases}$$

(a) Compute the expected rate of 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 random variable with mean $\mu = -2$ and standard deviation $\sigma = 3.2$. You will need to use a Z-score table or your calculator or MATLAB for this problem.

(a) What is the probability that $X \geq 1$?

(b) What is the probability that $0 \leq X \leq 1$?

(c) Find a symmetric interval about the mean so that you can be 97% sure that $X$ lies in that interval. In other words, find $a$ so that

$$p(-2 - a \leq X \leq -2 + a) = .97.$$

7. 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.”

8. Let $X$ be a normally distributed variable with mean $\mu$ and standard deviation $\sigma$.

(a) Compute $E[e^{2X}]$.

(b) Compute $E[X^2 + e^{2X} + 17]$.

---

$^2N$ 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.
(c) Compute $Var(e^{2X} + 17)$.

9. 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 at least triples after two year?

10. Recall from the example in Section [9.5], the sequence of random variables $X_1, X_2, \ldots$ where $X_i$ has a $1/3$ probability of being 1, 2 and 3, and where all the $X_i$ are independent.

(a) Compute $E[X_i]$ and $SD(X_i)$.

(b) Let $Y = \frac{X_1 + X_2 + \ldots + X_{100}}{100}$. Rather then explicitly computing the p.d.f. of $Y$ by hand (as done when averaging 2 and 3 of the $X_i$ in the example), the Central Limit Theorem says you can approximate $Y$ by what random variable? Write down the cumulative distribution function of this (approximating) random variable.

(c) Use part (b) to approximate $P(Y \leq 2.013)$.

(d) Approximate $P(X_1 + X_2 + \ldots + X_{100} \leq 201.3)$.

11. Let $X_1, X_2, \ldots$ be a sequence of random variables, all independent and with identical distributions. Assume that $X_i$ has a $1/9$ chance of being 0, $2/3$ chance of being 2, and $2/9$ chance of being 3. Approximate $P(X_1 + X_2 + \ldots + X_{200} \leq 441)$. (Hint: repeat steps in preceding problem.)
Chapter 10

The Black-Scholes formula

In Section 8.3, Equation 8.3 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 10.2. This model can also be derived from some basic economic assumptions, including the efficient market hypothesis of Section 10.1.

As in Chapter 8, we use the risk-neutral pricing principle to value European-style derivatives. Again we see that the (real world) expected rate of return of the stock is not relevant in the final answer; however, the stock volatility (see Section 10.3) 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 10.4 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 per annum continuous compounding by $r$. 
10.1 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. 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 9.4, one could ask if these two random variables are independent. If they are, then

$$p(S_4 \leq 29|S_3 \leq 30) = p(S_4 \leq 29)$$

Could this be possible? If today’s price is $400 then $p(S_4 \leq 29)$ should be tiny. It is highly unlikely that the stock will lose more than 90% of 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 not so unlikely. There is a much greater 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

$$X = \frac{S_4 - S_3}{S_3}, \quad Y = \frac{S_7 - S_6}{S_6}$$

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.

10.2 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. As described in Section 9.5, 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).

Example 10.1. For $i = 1, 2, \ldots, 100$ let $X_i$ denote the return of the stock over the $i$-th (trading) day. Suppose the returns are independent and identically distributed, with mean 0.001 and standard deviation 0.01. (We will see later that these values are reasonable.) The spot price of the stock is $S_0 = 60$. Approximate the stock price $S_{100}$ in terms of a standard normal variable $Z \sim N(0, 1)$.

Solution: Rewrite

$$S_{100} = S_0 \frac{S_1}{S_0} \frac{S_2}{S_1} \cdots \frac{S_{100}}{S_{99}}.$$ 

Dividing both sides by $S_0 = 60$ and taking ln

$$\ln(S_{100}/60) = \ln \left( \frac{S_1}{S_0} \frac{S_2}{S_1} \cdots \frac{S_{100}}{S_{99}} \right) = \ln \left( \frac{S_1}{S_0} \right) + \ln \left( \frac{S_2}{S_1} \right) + \cdots + \ln \left( \frac{S_{100}}{S_{99}} \right).$$

The default unit of time is a year; thus, we think of one day as being a small unit of time. This allows us to use the first Taylor polynomial approximation: for $x \approx 0$, $\ln(1 + x) \approx x$; or equivalently, for $y \approx 1$, $\ln(y) = y - 1$. Since $S_{i+1}/S_i - 1 = \frac{S_{i+1} - S_i}{S_i}$, the latter version of the Taylor polynomial implies

$$\ln(S_{100}/60) \approx \frac{S_1 - S_0}{S_0} + \frac{S_2 - S_1}{S_1} + \cdots + \frac{S_{100} - S_{99}}{S_{99}}.$$ 

We now use the Central Limit Theorem from Section 9.5. On the right hand side, we are summing a bunch of independent and identically distributed
random variables. Thus, the random variable \( \ln(S_{100}/60) \) can be approximated by a normal variable with distribution \( N(0.001 \times 100, 0.01 \times \sqrt{100}) \). That is,
\[
\ln(S_{100}/60) \approx 0.1 + 0.1Z
\]
for some standard normal variable \( Z \sim N(0, 1) \). Solving for \( S_{100} \) we get
\[
S_{100} \approx 60e^{0.1Z + 0.1}.
\]

Since we could look at returns not just every day, but every hour, minute, second, ...., it turns out that the approximation in the above example is in fact an equality. A fancier version of the Central Limit Theorem proves that the three consequences of the efficient market hypothesis, mentioned at the beginning of this subsection, 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 \( \sigma \).

We will derive this log-normal model again in Section 12 when we discuss how 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.

Log-normality can be restated in 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 8.2, 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 rate of 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

\[
S_t e^{\mu(T-t)} = E_{RW}[S_T] \\
= E_{RW}[S_t e^{\sigma\sqrt{T-t}Z + a(T-t)}] \\
= S_t e^{a(T-t)} E_{RW}[e^{\sigma\sqrt{T-t}Z}] \\
= S_t e^{a(T-t)} \Phi_Z(\sigma\sqrt{T-t}) \\
= S_t e^{a(T-t)} e^{\frac{\sigma^2(T-t)}{2}} \\
= S_t e^{\left(\frac{\sigma^2}{2} + a\right)(T-t)},
\]

where we have used the moment generating formula in Equation [9.3]. 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 16 (Log-normal model).** The stock price \( S_t \) follows the log-normal model. That is,

\[
S_T = S_t e^{\sigma\sqrt{T-t}Z + (\mu - \frac{\sigma^2}{2})(T-t)}.
\]
where $\mu$ is the expected rate of 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.

When $t = 0$, the formula simplifies to:

$$S_T = S_0 e^{\sigma \sqrt{T}Z + (\mu - \frac{\sigma^2}{2})T}.$$  

10.3 Volatility and expected rate of 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 Using Taylor’s series for the exponential function $e^x$

$$\frac{S_{t+\Delta t}}{S_t} = e^{\sigma \sqrt{\Delta t}Z + (\mu - \frac{\sigma^2}{2})(\Delta t)}$$

$$= 1 + \sigma \sqrt{\Delta t}Z + \left( \mu - \frac{\sigma^2}{2} \right) (\Delta t) + \frac{1}{2} \left( \sigma \sqrt{\Delta t}Z + (\mu - \frac{\sigma^2}{2})(\Delta t) \right)^2 + \ldots$$

$$= 1 + \sigma \sqrt{\Delta t}Z + O(\Delta t)$$

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 $\ln \frac{S_{t+\Delta t}}{S_t}$.

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 10.2.** Suppose the current stock price of a stock is $20, the expected rate of 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

$$E_{RW}[S_T] = 20e^{0.12 \times \frac{4}{256}} = 20.04.$$ 

Therefore the expected percent change is

$$E\left[\frac{S_T - S_0}{S_0}\right] = \frac{E[S_T]}{S_0} - \frac{S_0}{S_0} = \frac{20.04}{20} - 1 = 0.002 = 0.2\%.$$
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

\[ SD(S_T) = S_0 e^{\mu T} \sqrt{e^{\sigma^2 T} - 1}. \]

In that case, we find that \( SD(S_T) = 0.501 \), not much different from our approximation.

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

\[
\alpha = (\mu - \sigma^2/2)T = \left(0.12 - \frac{0.2^2}{2}\right) \times 0.0016 = 0.0016 \\
\beta = \sigma \sqrt{T} = 0.20 \times \sqrt{\frac{4}{256}} = 0.025
\]

then

\[
S_T = 20e^{0.0016 + 0.025Z} = 20e^X
\]
To get a 90% confidence interval for $S_T$ from the confidence interval for $X$,

\[
0.9 = p(-1.645 \times 0.025 + 0.0016 \leq X \leq 1.645 \times 0.025 + 0.0016) = p(20e^{-1.645 \times 0.025 + 0.0016} \leq S_T \leq 20e^{1.645 \times 0.025 + 0.0016})
\]

So a 90%-confidence interval for the future stock price $S_T$ is

\[
20e^{-1.645 \times 0.025 + 0.0016} = 19.225 \leq S_T \leq 20e^{1.645 \times 0.025 + 0.0016} = 20.873.
\]

### 10.4 Risk neutral pricing

Recall the risk-neutral pricing principle in the binomial tree model from Section 8.2. 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 $E_{RNW}[:]$ denotes the expectation in the risk-neutral world. This formula implies that under the binomial tree model of a stock, a derivative can be priced in the real world by using the expectation of the derivative in the risk neutral world.

To prove this formula, 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 given by

\[
h = e^{-rT}E_{RNW}[\text{payoff of derivative}].
\]

We continue to call this risk neutral valuation or risk neutral pricing for a derivative.

**Example 10.3.** Consider the derivative that pays you $100 in $T$ years. This derivative is just a $T$-year zero-coupon bond. Suppose today is time $t = 0$. Let $r$ be the $T$-year risk-free zero rate. 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. This is simply the present value of $100, to be paid out in \( T \) years.

**Example 10.4.** 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 \( t = 0 \) and the spot price is \( S_0 \). Then the price of this derivative is

\[
h = e^{-rT} E_{RNW} [S_T - K] = e^{-rT} (S_0 e^{rT} - K) = S_0 - e^{-rT} K.
\]

If you call \( K \) the “delivery price” then this derivative is a forward contract with delivery price \( \$K \) and \( h \) is the current price of the forward contract. We discuss forwards in Chapter 5.

**Example 10.5.** 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 stock price is \( \$20 \). The expected rate of return of the stock is 10\%, the risk-free rate is 6\%. The price of the derivative \( h \) is then

\[
h = e^{-0.06 \times 9/12} E_{RNW} [3S_{9/12} + 75] = 3 \times 20 + 75 e^{-0.06 \times 9/12}.
\]

Note that the expected rate of return is irrelevant to this problem; the relevant interest rate is the risk-free rate of 6\%. In the calculation, we have used the fact that

\[
E_{RNW} [3S_{9/12}] = 3E_{RNW} [S_{9/12}] = 3S_0 e^{0.06 \times 9/12}.
\]

**Example 10.6.** A “power derivative” has a payoff 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.

\[
h = e^{-rT} E_{RNW} [S_T^2]
\]

\[
= e^{-rT} E_{RNW} [S_0^2 e^{2(r-\sigma^2/2)T + 2\sigma\sqrt{T}Z}]
\]

\[
= e^{-0.06 \times 0.75} E_{RNW} [20^2 e^{2(0.06-(0.2^2/2))0.75+2\times0.2\sqrt{0.75}Z}]
\]

\[
= e^{-0.06 \times 0.75} 400 e^{2(0.06-(0.2^2/2))0.75} e^{(2\times0.2\sqrt{0.75})^2/2} = 431.15,
\]
10.4. RISK NEUTRAL PRICING

using the moment generating formula in Equation 9.3.

**Example 10.7.** Consider the stock from the previous example. Consider a European call expiring in 9 months with strike price $19. The premium of the call using risk neutral pricing is given by

\[ h = e^{-0.06 \times 9/12} E_{RNW}[\max(S_{9/12} - 19, 0)]. \]

This expectation is harder to compute. It computation leads to the Black-Scholes formula for European options.

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 17 (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

\[ d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma \sqrt{T}} \quad \text{and} \quad d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma \sqrt{T}}. \]

Notice that \( d_1 = d_2 + \sigma \sqrt{T} \).

The case of the European put with the same specifications follows from put-call parity.

**Formula 18 (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 \( Z \). These values can be looked up in a Z-score table or from a calculator.

Although these formula make seem a bit mysterious, they make sense in “extreme conditions.” For example, suppose that the call is extremely in-the-money, that is \( S_0 \) is much much bigger than \( K \). Then there is a large (like
99.99\% chance that the call will be exercised, in which case the call holder’s payoff is approximately

\[ e^{-rT} E_{RNW}[S_T - K] = S_0 - Ke^{-rT}. \]

On the other hand, as \( S_0 \) gets large much larger than \( K \), \( \ln(S_0/K) \) approaches infinity, and so \( d_1, d_2 \) approach infinity as well (holding \( \sigma, r, T \) constant). Thus, the limit of the Black-Scholes price of the call as \( S_0/K \to \infty \) is

\[
\lim_{S_0/K \to \infty} h = \lim_{S_0/K \to \infty} e^{-rT}(S_0e^{rT}F_Z(d_1) - KF_Z(d_2)) = e^{-rT}(S_0e^{rT}\lim_{S_0/K \to \infty} d_1) - KF_Z(\lim_{S_0/K \to \infty} d_2) = e^{-rT}(S_0e^{rT} \times 1 - K \times 1) = S_0 - Ke^{-rT}.
\]

So we see that the Black-Scholes price agrees with “common sense” in this extreme case. (Note that we are not assuming \( S_0 \to \infty \) necessarily, just \( S_0/K \to \infty \).

Similarly, one can argue (and check for the Black-Scholes formula) that way out-of-the-money calls price at zero:

\[
\lim_{S_0/K \to 0} h = \lim_{S_0/K \to 0} e^{-rT}(S_0e^{rT}F_Z(d_1) - KF_Z(d_2)) = 0.
\]

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 \( S_0^{\text{eff}} \) as in Section 8.5. For example, if the present value of the dividends that a stock pays between \( t = 0 \) and \( t = T \) is \( D \), then \( S_0^{\text{eff}} = S_0 - D \) and the Black-Scholes formula for a European call with the same specifications as above becomes

\[ h = (S_0 - D)F_Z(d_1) - KF_Z(d_2)e^{-rT} \]

where

\[
d_1 = \frac{\ln \frac{S_0 - D}{K} + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\ln \frac{S_0 - D}{K} + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}.
\]

In other words one must replace \( S_0 \) with \( S_0^{\text{eff}} = S_0 - D \) wherever \( S_0 \) appears in the formula.
10.5 Proof of the Black-Scholes formula

In this section, we prove the Black-Scholes formula\(^\text{[17]}\) for the premium \(c\) of a European call with strike \(K\) and expiration \(T\). We assume that the risk-free rate is \(r\) for all zero rates and this remains true between \(t = 0\) and \(t = T\).

By risk neutral pricing we know that \(c\) is given by

\[
c = e^{-rT} E_{RNW}[\max(S_T - K, 0)]
\]

since

\[
\max(S_T - K, 0)
\]

is the payoff of a European call at expiration time \(T\). So our goal is to compute the expectation using the fact that \(S_T\) obeys the log-normal model. The payoff is zero for \(S_T < K\) and equals \(S_T - K\) for \(S_T \geq K\), so the definition of expectation and Theorem 9.9 imply that

\[
E_{RNW}[\max(S_T - K, 0)] = \int_{K}^{\infty} (S_T - K) f_{S_T}(x) dx,
\]

where \(f_{S_T}(x)\) is the p.d.f. for \(S_T\).

The next step is to write \(S_T\) in terms of the standard normal variable \(Z\). Then \(S_T = g(Z)\), where

\[
g(z) = S_0 e^{a + bZ},
\]

where

\[
a = (r - \sigma^2/2) T
\]

and

\[
b = \sigma \sqrt{T}.
\]

Using Theorem 9.9 again to avoid explicitly writing out \(f_{S_T}(x)\), we have

\[
E_{RNW}[\max(S_T - K, 0)] = \int_{m}^{\infty} (S_0 e^{a + bZ} - K) \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz.
\]

The value of \(m\) is determined by setting \(S_T = S_0 e^{a + bm}\) equal to \(K\) and solving for \(m\). Thus

\[
m = \ln(K/S_0) - a/b.
\]

Essentially all that is going on here is that we are doing a change of variables in the integral and adjusting the limits of the integral. Breaking the integral into two parts gives

\[
\int_{m}^{\infty} S_0 e^{a + bZ} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz - K \int_{m}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz.
\]
CHAPTER 10. THE BLACK-SCHOLES FORMULA

Now consider the second term above. It equals $KF_Z(-m)$ from the even symmetry of $Z$, where recall that

$$F_Z(-m) = \int_{-\infty}^{-m} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$ 

The first term is harder:

$$\frac{1}{\sqrt{2\pi}} \int_m^{\infty} S_0 e^{a+bz} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_m^{\infty} S_0 e^{a+bz-z^2/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_m^{\infty} S_0 e^{-\frac{(z-b)^2}{2}} e^{a+b^2/2} dz$$

after completing the square. Now we can pull out the $S_0 e^{a+b^2/2}$ factor to get

$$S_0 e^{a+b^2/2} \int_m^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-b)^2}{2}} dz.$$ 

What is left in the integral is coming from a normal variable with standard deviation 1 and mean $b$. Hence the integral reduces to $F_Z(b-m)$.

We can verify that $d_2 = -m$ in the Black-Scholes formula using the fact that $-\ln(K/S_0) = \ln(S_0/K)$. Also $b-m = \sigma \sqrt{T} + d_2 = d_1$. Finally note that $a + \frac{b^2}{2} = rT$, so that we have proven that

$$E_{RNW}[\max(S_T-K,0)] = S_0 e^{rT} F_Z(d_1) - K F_Z(d_2)$$

and the formula 17 follows for the price $c$ of the call by taking the present value.

10.6 Digital options

In this section we derive the pricing formula for another kind of derivative.

In general if we take any derivative whose payoff at time $T$ is given by $g(S_T)$, where $g(x)$ is some function, then as we saw in the last section, the price $h$ today of this derivative is given by

$$h = e^{-rT} E_{RNW}[g(S_T)]$$

$$= e^{-rT} E_{RNW}[g(S_0 e^{(r-a^2/2)T+\sigma\sqrt{T}Z})]$$

$$= e^{-rT} \int_{-\infty}^{\infty} g(S_0 e^{(r-a^2/2)T+\sigma\sqrt{T}z}) \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz$$

We will use this formula to compute the value of the digital option.
Example 10.8. 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 a simplified 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?

$$S_T > K \leftrightarrow S_0 e^{(r-\frac{\sigma^2}{2})T+\sigma \sqrt{T}Z} > K \leftrightarrow Z > \frac{\ln \frac{K}{S_0} - (r - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}} = -d_2$$

So our payoff function is

$$g(S_0 e^{(r-\frac{\sigma^2}{2})T+\sigma \sqrt{T}Z}) = 0 \text{ if } Z \leq -d_2 \text{ and }$$
$$g(S_0 e^{(r-\frac{\sigma^2}{2})T+\sigma \sqrt{T}Z}) = 1 \text{ if } Z > -d_2.$$ 

Plugging this into the integral for the price of a general derivative,

$$h = e^{-rT} \int_{\infty}^{-d_2} g(S_0 e^{(r-\frac{\sigma^2}{2})T+\sigma \sqrt{T}Z}) e^{-z^2/2} \frac{dz}{\sqrt{2\pi}}$$

$$= e^{-rT} \left( \int_{-\infty}^{-d_2} 0 \times \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz + \int_{-d_2}^{\infty} 1 \times \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \right)$$

$$= e^{-rT} \int_{-\infty}^{d_2} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz$$

$$= e^{-rT} F_Z(d_2).$$

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 $K$ at time $T$.

10.7 Problems

Unless otherwise stated, assume the log normal model for all stock prices.
1. The current spot price of a stock is $50, the expected rate of 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. Also compute the expected percent change in the stock between months 0 and 6.

2. The current spot price of a stock is $34, the expected rate of return of the stock is 8%, and the volatility of the stock is 20%. The risk-free rate is 3% for all times. Compute the price of a European call option on the stock with strike price $35 expiring in 4 months. Do the same for a European put with the same strike and the same expiration date.

3. Repeat the previous problem if the stock pays a dividend of 25 cents in 2 months.

4. The current spot price of a stock is $68 and the volatility of the stock is 14%. The risk-free rate is 1%. Compute the price of a straddle on the stock with strike $70 and expiration in 3 weeks built from a European call and put.

5. Verify that the Black-Scholes formulas for the call and put satisfy put-call parity.

6. Approximate the price of a European put with strike \( K \) and expiration \( T \) assume that the spot price \( S_0 \) is a lot less than \( K \) or a lot more than \( K \). This was partially done for the European call in the text.

7. Consider a stock which is currently trading at $16. It has a volatility of 18%. The risk-free rate is 2%. Calculate the premiums of the three European call options on the stock with strikes 10, 15, 20 and expiration in 2 months. Also verify that the butterfly spread inequality holds in this case.

8. The current spot price of a stock is $21, the expected rate of 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.
9. The current spot price of a stock is $20, the expected rate of 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.

10. If the example in Section 10.6 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.

11. The current spot price of a stock is $34, the expected rate of 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.

12. Consider a derivative whose payoff at expiration is $S_T^2 + 1$ if $S_T < 100$ and $S_T^2 + S_T - 100$ if $S_T \geq 100$. Assuming the volatility is $\sigma$, the risk-free rate $r$, and the spot price $S_0$, compute the price of this derivative today.

Hint: Write the derivative as 3 derivatives (a power option, an option and a digital option) and price each separately.

13. Assuming the log-normal price for a stock

$$S_T = S_t e^{\sigma \sqrt{T-t} Z + (\mu - \frac{\sigma^2}{2})(T-t)}$$

verify the formula

$$SD(S_T) = S_0 e^{\mu T} \sqrt{e^{\sigma^2 T} - 1}$$
14. 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 implied volatility, \( \sigma_{BS} \). Some calculus, shows that \( \sigma_{BS} \) exists and is unique.

The current spot price is $40, the expected rate of 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.

15. Assume in the real world \( S_T = S_0e^{(\mu-\sigma^2/2)T-\sigma\sqrt{T}Z} \) and in the risk-neutral world \( S_T = S_0e^{(r-\sigma^2/2)T-\sigma\sqrt{T}Z} \). Suppose the values of \( r, \sigma, \mu, T, K, S_0 \) (and hence \( d_1, d_2 \) as well) are unknown. Assume the real world is risk averse. Consider a European call on the asset with expiration \( T \) and strike \( K \).

(a) Suppose \( P(Z > d_2) = 0.7 \) is known. What if anything is known about the probability of the call being exercised in the real world or risk neutral world?

(b) Suppose \( P(Z > d_1) = 0.7 \) is known. What if anything is known about the probability of the call being exercised in the real world or risk neutral world?

Note that the probability of the call being exercised in the risk-neutral world is not useful for math finance.
Chapter 11

The Greek letters

When holding a portfolio, it is important to understand how the portfolio will react to changes in the underlying asset prices or other parameters which affect the portfolio’s value. For example, if market volatility increases, how will the portfolio value change? If the portfolio value changes drastically with respect to a parameter, it is important to understand this and adjust the portfolio accordingly.

The way a portfolio reacts to changes in the price of an asset, asset volatility, the risk-free rate, and other variables is referred to by specific letters from the Greek alphabet. Collectively, we call this group of letters “the Greeks”.

11.1 Delta

Perhaps the most important of the Greeks is $\Delta$ (Delta). Delta measures the change of the portfolio’s value with respect to a change in the price of an underlying asset. If $\Pi$ represents the value of a portfolio and $S$ the price of the asset, then the delta of the portfolio is given by

$$\Delta = \frac{\partial \Pi}{\partial S}.$$  

\[ ^1 \text{Actually one of the symbols, vega, is not a Greek letter.} \]
11.1.1 Delta for a call option and for a put option

For a portfolio consisting of a single European call option with strike $K$ and expiration $T$ written on one share of stock, the delta can be computed by taking the partial derivative of the Black-Scholes formula with respect to the spot price $S_0$.

**Formula 19.** For a European call option

$$\Delta = F_Z(d_1),$$

where

$$d_1 = \frac{\ln \frac{S_0}{K} + (r + \frac{\sigma^2}{2})T}{\sigma \sqrt{T}}$$

as in Formula 17.

To establish the formula, we take the partial derivative of

$$c = S_0 F_Z(d_1) - Ke^{-rT} F_Z(d_2)$$

to get

$$\frac{\partial c}{\partial S_0} = F_Z(d_1) + S_0 (F_Z(d_1))' - Ke^{-rT} (F_Z(d_2))'$$

(11.1)

where we have used the product rule. The notation $(F_Z(d_i))'$ refers to the partial derivative of the expression with respect to $S_0$. The important thing to keep in mind is that $d_i$ depends on $S_0$, so we have to use the chain rule next.

Recall that

$$F_Z(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

and so we use the fundamental theorem of calculus to conclude that

$$F'_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Then by the chain rule

$$(F_Z(d_1))' = \frac{e^{-(d_1)^2/2}}{\sqrt{2\pi}} \cdot \frac{1}{S_0 \sigma \sqrt{T}}$$
since
\[ \frac{\partial d_1}{\partial S_0} = \frac{1}{S_0 \sigma \sqrt{T}}. \]

A similar calculation gives
\[ (F_Z(d_2))' = \frac{e^{-(d_2)^2/2}}{\sqrt{2\pi}} \cdot \frac{1}{S_0 \sigma \sqrt{T}}. \]

Next using the fact that \( d_1 = d_2 + \sigma \sqrt{T} \) and
\[ d_2 = \frac{\ln S_0 - (r - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}}, \]
we get
\[
e^{-\frac{(d_1)^2}{2}} = e^{-\frac{(d_2)^2}{2}} \cdot e^{\ln(S_0/K)} e^{r T} = e^{-\frac{(d_2)^2}{2}} (S_0/K) e^{r T}.
\]

and plugging this back into 11.1 gives
\[ \frac{\partial c}{\partial S_0} = F_Z(d_1) \]

since the last two terms cancel out.

**Formula 20.** For a European put option
\[ \Delta = F_Z(d_1) - 1 \]

or equivalently,
\[ \Delta = -F_Z(-d_1). \]

This follows by taking the derivative with respect to \( S_0 \) in the put-call parity formula and then using Formula [19]
11.1.2 Properties of Delta

For a call option, $\Delta = F_Z(d_1)$. The latter is the c.d.f. for the standard normal variable evaluated at $d_1$ and is therefore always positive, so we see that the $\Delta$ of a call option is always positive. Since $\Delta$ is the derivative of the call price with respect to the stock price, it follows that the call price is an increasing function of $S_0$, and thus increasing the value of $S_0$ will always lead to an increase in the value of $c$, something which is intuitively clear.

Moreover, $0 \leq F_Z(d_1) \leq 1$ and so the $\Delta$ for a call option is always between 0 and 1. A value of $\Delta$ closer to 0 indicates that the call option is mostly unaffected by changes in the stock price and a value closer to 1 indicates that the call option will change approximately by the same amount that the stock price changes.

The two extremes for $\Delta$ occur when the call becomes very much in-the-money or very much out-of-the-money. When the stock is very much in-the-money, $S_0$ is large relative to $K$ and this makes $d_1$ tends to positive infinity. Thus $F_Z(d_1) \approx 1$ and hence $\Delta \approx 1$. This is consistent with the fact that a call very much in-the-money has a value approximately equal to $S_0 - Ke^{-rT}$ (see Section 10.4) and so if the stock price moves up (or down) by an amount, the call price will move up (or down) by approximately the same amount.

When the stock is very much out-of-the-money, on the other hand, $S_0$ is small relative to $K$ and this makes $\ln(S_0/K)$ and hence $d_1$ tend to negative infinity. Thus $F_Z(d_1) \approx 0$ and hence $\Delta \approx 0$. This is consistent with the fact that the call option will be worth approximately 0 and even if the stock price moves a little, the call will still be worth close to nothing.

For a put option, $\Delta = -F_Z(-d_1)$, which is always negative. This means that when the stock price increases the put price will decrease, which is intuitively clear. Moreover, as the stock price increases relative to $K$, the put option is very far out-of-the-money and the $\Delta$ tends to 0. On the other hand, if the stock price decreases relative to $K$, the put option becomes very much in-the-money and the $\Delta$ tends toward $-1$. This means that when the stock price moves up by an amount, the put option moves down by the corresponding amount.
11.1. DELTA

11.1.3 Delta hedging

The discussion in the previous suggestion can be extended to the concept of delta hedging. Since \( \Delta = \frac{\partial \Pi}{\partial S} \), we see that for a small change \( \Delta S \) in the stock price, \( \Delta \approx \frac{\Delta \Pi}{\Delta S} \). Or in other words,

\[ \Delta \cdot \Delta S \approx \Delta \Pi. \]

This means that when the stock price changes by \( \Delta S \), the portfolio value changes by \( \Delta \) times \( \Delta S \).

Another way to use this equation is to construct a portfolio that is unaffected by a small change in the stock price at the given moment in time. To do this we would subtract from our portfolio \( \Delta \) shares of stock; in other words, if \( \Delta \) is positive, we would go short \( \Delta \) shares of stock. This leads to a portfolio with a \( \Delta \) of zero, which is called a \textit{delta-neutral} portfolio.

More precisely, if a portfolio \( \Pi \) has a delta equal to \( \Delta \), then we could create a new portfolio \( \Pi' \) which consists of the original portfolio \( \Pi \) together with \(-\Delta\) shares of stock and then \( \Pi' \) will be a delta-neutral portfolio.

\textbf{Example 11.1.} A stock is trading at $10. Its annual volatility is 25%. The risk-free rate is 2%.

1. Calculate \( \Delta \) for call options which expire in 6 months with strikes 5, 10, and 15.

2. If you own one of the call options with strike 10, how many shares would you need to short to create a delta-neutral portfolio \( \Pi' \)?

3. If the stock price moves up by 10 cents, calculate the change in the value of the portfolio from the previous part directly using the Black-Scholes formula.

\textit{Solution:}

1. We calculate \( d_1 \) in each case and get 4.06599, 0.14496, and \(-2.14870\), respectively, which leads to a \( \Delta \) of 0.999976, 0.557628, and 0.015829 for the three different strike prices.

2. For the \( K = 10 \) option, \( \Delta = 0.5576 \) and so we would short 0.5576 shares to create a delta-neutral portfolio. In other words, our delta-neutral portfolio consists of the original portfolio of 1 call option together with a short position of 0.5576 shares of stock.
3. With the stock price at 10, the call with strike $K = 10$ is worth $c = 0.752$ from Formula 17 and so the value of portfolio $\Pi'$ is

$$0.752 - (0.5576)10 = -4.824$$

If the stock moves up to 10.10, then we calculate $c = 0.8086$ from Formula 17 and now the value of $\Pi'$ is

$$0.809 - (0.5576)(10.10) = -4.823$$

So we have successfully created a portfolio that is resistant to small moves in the stock prices at the particular moment in time.

11.2 The other Greeks

The term Vega, written with the Greek letter $\nu$ (pronounced 'nu'), expresses the change in the portfolio value $\Pi$ with respect to volatility $\sigma$:

$$\nu = \frac{\partial \Pi}{\partial \sigma}.$$  

**Formula 21.** For a European call or put option

$$\nu = \frac{S_0 \sqrt{T}}{\sqrt{2\pi}} e^{-\frac{d_2}{2}}.$$  

This formula can be established in the same way that we established the formula for $\Delta$. Notice that $\nu$ is always positive for both a call and put option. This means that increasing volatility will always increase the price of a call option and the price of a put option.

The Greek letter Gamma, $\Gamma$, calculates the change in $\Delta$ with respect to asset price. In other words, $\Gamma$ represents the second partial derivative of the portfolio value with respect to the asset price $S$:

$$\Gamma = \frac{\partial^2 \Pi}{\partial S^2}.$$  

The formula for $\Gamma$ for both a call and a put is

$$\Gamma = \frac{e^{-\frac{d_2}{2}}}{\sqrt{2\pi}S_0\sigma \sqrt{T}}.$$
11.2. THE OTHER GREEKS

The Greek letter rho, ρ, calculates the change in Π with respect to the risk-free rate r. The Greek letter Theta, Θ calculates the change in Π with respect to remaining time until expiration T (or T − t if today is time t instead of 0)

\[ \rho = \frac{\partial \Pi}{\partial r}, \quad \Theta = \frac{\partial \Pi}{\partial T}. \]

We omit the formulas for these Greeks.

**Example 11.2.** Let \( S_0 \) be the asset spot price. The asset has expected rate of return 0.1, and volatility 0.2. Let \( B_0 \) be the price of a (zero-coupon) bond which pays $1 in one year. Assume the risk-free rate is \( r = 0.03 \) per annum, with continuous compounding. Calculate the Greeks for the portfolio Π which is made of 1 asset and 3 bonds.

**Solution:** First we calculate the Greeks of the asset and the bond individually, denoting them with subscripts ‘S’ or ‘B’ to distinguish them. Recall that \( B_0 = 1e^{-rT} \).

\[
\begin{align*}
\Delta_S &= \frac{\partial S_0}{\partial S_0} = 1, \quad \nu_S = \frac{\partial S_0}{\partial \sigma} = 0, \quad \Gamma_S = \frac{\partial^2 S_0}{\partial S_0^2} = 0 \\
\rho_S &= \frac{\partial S_0}{\partial r} = 0, \quad \Theta_S = \frac{\partial S_0}{\partial T} = 0, \\
\Delta_B &= \frac{\partial B_0}{\partial S_0} = 0, \quad \nu_B = \frac{\partial B_0}{\partial \sigma} = 0, \quad \Gamma_B = \frac{\partial^2 B_0}{\partial S_0^2} = 0 \\
\rho_B &= \frac{\partial B_0}{\partial r} = -Te^{-rT} = -e^{-0.03}, \quad \Theta_B = \frac{\partial B_0}{\partial T} = -re^{-rT} = -0.03e^{-0.03},
\end{align*}
\]

The Greeks for the portfolio are then

\[
\Delta = \Delta_S + 3\Delta_B = 1, \quad \nu = 0, \quad \Gamma = 0, \quad \rho = -3e^{-0.03}, \quad \Theta = -0.09e^{-0.03}
\]

If a bank writes a derivative, it often likes to hedge its position against fluctuations in the underlying asset price. Section 11.1.3 describes how to achieve delta-neutrality. But this only hedges against small changes in stock price. For example, if interest rates or volatility changes, the position could increase or decrease in value. Sometimes a bank wants to be delta-, gamma- and vega-neutral all at once. Each Greek that it wants to set equal to zero requires trading in a different derivative. 

\footnote{This is not entirely correct, since as we see in Section 13.2 not all the Greeks are independent.}
Example 11.3. You have written a derivative $Z$ which has the following Greeks: $\Delta_Z = 2, \nu_Z = 1$. Suppose derivative $X$ has the following Greeks: $\Delta_X = 3, \nu_X = 3$. Suppose derivative $Y$ has the following Greeks: $\Delta_Y = -3, \nu_Y = 2$. What position should you take in $X$ and $Y$ to make yourself both delta- and vega-neutral?

Solution: Suppose the portfolio $\Pi$ is made of $-1$ of derivative $Z$ (since you wrote $Z$), $a$ derivatives $X$ and $b$ derivatives $Y$. Then we want to solve the linear system of equations

$$
0 = \Delta_{\Pi} = -\Delta_Z + a\Delta_X + b\Delta_Y = -2 + 3a - 3b \\
0 = \nu_{\Pi} = -\nu_Z + a\nu_X + b\nu_Y = -1 + 3a + 2b
$$

The positions are $a = -\frac{1}{3}, b = 1$.

11.3 Greeks for binomial trees

When we studied the binomial tree model for a stock, we actually already encountered a discrete version of $\Delta$; the role of $\Delta$ was played by the number $N$ in Equation 8.2 in 8.1.3. At that time, we introduced $N$ as the number of shares of stock needed to combine with a short position in a derivative to make a portfolio whose value is unchanged by a move in the stock price (at that moment in time). In other words, we asked for the number of shares to add to a portfolio with one short derivative in order to make a new portfolio which was delta-neutral. We found that $N$ was equal to

$$
\frac{h_u - h_d}{S_0u - S_0d},
$$

where $h_u$ was the value of the derivative at the up node and $h_d$ was its value at the down node. This formula for $N$ looks just like a change in the value of $h$ divided by the change in the value of the stock at the next node in the tree. In fact, this formula can be shown to be the correct formula for the discrete version of $\frac{\partial \Pi}{\partial S}$. In conclusion, $N$ is nothing more than the discrete value of $\Delta$ for the binomial tree model.

Can we recover any of the other Greeks from a binomial tree? The answer is some, but not all.
The risk-free rate \( r \) is fixed, so the tree says nothing about \( \rho \). Similarly, the underlying asset’s volatility is also held fixed: although not explicitly stated, the volatility can be recovered from the up- and down-factors \( u \) and \( d \), the up- and down-probabilities, \( p_d \) and \( p_u \), as well as the time-step-size \( \Delta t \). Thus, \( \nu \) also cannot be inferred from, the tree.

Since \( \Delta = \frac{\partial h}{\partial S} \), we estimated it by computing the ratio of the change of the derivative price at two adjacent nodes (at the same time), \( h_u - h_d \), divided by the change of the underlying asset at those two nodes, \( S_{0u} - S_{0d} \). To approximate \( \Gamma = \frac{\partial \Delta}{\partial S} \), we need to compute the ratio of the change of the delta at two adjacent nodes (at the same time), divided by the change of the underlying asset \( S_{0u} - S_{0d} \). For this we need at least two time steps.

**Example 11.4.** Estimate \( \Gamma \) of the European put option described in Figures 8.2 and 8.3.

**Solution:** First we approximate the \( \Delta \) at two nodes, \((t, s) = (6/12, S_{0u})\) and \((t, s) = (6/12, S_{0d})\). We label these \( \Delta_u \) and \( \Delta_d \)

\[
\Delta_u = \frac{h_{uu} - h_{ud}}{S_{uu} - S_{ud}} = -\frac{4}{24} \\
\Delta_d = \frac{h_{ud} - h_{dd}}{S_{ud} - S_{dd}} = \frac{16}{16}
\]

\( \Gamma \) at the initial node \((t, s) = (0, 50)\) is then approximately

\[
\Gamma = \frac{\Delta_u - \Delta_d}{S_{0u} - S_{0d}} = \frac{-\frac{1}{6} - \frac{-1}{20}}{20} = \frac{1}{24}.
\]

This example is a pretty crude estimate. But if we consider an \( n \)-step tree with \( n \) large, the approximation of \( \Gamma \) can become quite good.

In Section 13.2, we will see how to estimate the last Greek \( \Theta \) in terms of \( \Delta \) and \( \Gamma \), so we defer computing \( \Theta \) for binomial trees until then.

### 11.4 Problems

1. A stock trades at $20. Its annual volatility is 18%. The risk-free rate is 3%. Calculate the price of a European call option and put option with strike \( K = 20 \) and \( T \) equal to 4 months.
2. (Continued) Calculate the ∆ of a portfolio consisting of 2 long calls and 1 short put from the previous problem. How many shares of the stock would you short in order to build a new portfolio that is delta-neutral?

3. (Continued) If the stock moves down 15 cents, what is the exact change in price of the delta-neutral portfolio from the previous problem?

4. (Continued) What is the ν of one of these calls? of one of the puts?

5. (Continued) Calculate the ∆ and the ν of a straddle built from one of the calls and one of the puts. This example illustrates why a straddle built from at-the-money options is close to being delta-neutral and is a bet on volatility.

6. Assume the log-normal model. Compute the delta of the digital call option. (As usual, σ is volatility, r the risk-free rate, T the expiration, S₀ the underlying spot price, and K the strike.)

7. Apply put-call parity to compare the ν, Γ, ρ and Θ of a European put and a European call on the same asset with the same strike and expiration. Note that this holds without assuming the log-normal model for the asset, that is, the Black-Scholes formula for the options.

8. Recall the implied volatility σ_{BS} defined in an exercise in Section 10.7. Prove that it is well-defined. That is, prove that there cannot be two distinct values σ_{BS} and σ'_{BS} that plug into the Black-Scholes formula (along with the same values S₀, r, K, T) that produce the same call option price S₀F_{Z}(d₁) − KF_{Z}(d₂)e^{−rT}. (Hint: there is a useful property of ν of a call.)

9. Assume the log-normal model. Compute the limit of the ∆ of a put as the spot price S₀ goes to zero. This should make intuitive sense even without assuming the log-normal model.

10. Assume the log-normal model. The spot price is $100. The expected rate of return is 10%. The volatility is 20%. The risk-free rate is 3%.

A “power derivative” pays you the square of the underlying asset price in 4 months. Calculate its ∆ and Γ today.
11. (Continued). Suppose the asset price jump up by $0.05 today. Use your computation of the $\Delta$ to estimate the new price. Without computing the new price exactly, argue whether this is an underestimate or and overestimate, based on your computation for $\Gamma$. (Possibly Useless Hint: consider the parabola $y = ax^2$; for what values of the constant $a$ is the tangent-line approximation of the parabola at the origin an overestimate and for what values is it an underestimate?)

12. Derivative $A$ has the following Greeks today: $\Delta_A = 0$, $\nu_A = 0.3$ and $\Gamma_A = 1.1$. Derivative $B$ has the following Greeks today: $\Delta_B = 0$, $\nu_B = 0.4$ and $\Gamma_B = 2$. You write a derivative $X$ with the following Greeks today: $\Delta_X = 300$, $\nu_X = 120$ and $\Gamma_X = 48$. All three derivatives are for the same underlying asset. What position should you take in the underlying asset, derivative $A$, and derivative $B$ to be delta- gamma-, and vega-neutral today?

13. What can say about the signs of the $\nu$ and $\rho$ of a European call and of a European put? Give a qualitative (verbal) justification. For inspiration, you can look at the formulas in the case when the Black-Scholes formula holds; however, your answer should be whether or not the Black-Scholes formula is valid.

14. Consider the 2-step binomial tree from Problem 9 in Chapter 8. Compute the value of $\Delta$ of the European put option at the initial node and each of the two intermediary nodes.

15. (Continued) Build a portfolio consisting of cash (borrowed from or deposited in a bank at the risk-free rate), a certain amount of stocks, and 1 put option so that its value at the initial node is zero and so that its value will be zero regardless of where the stock goes at the intermediate time step.

16. (Continued) Suppose the stock goes up at the intermediate time step. What should your new portfolio holdings (1 put, some cash, and some stock) be so that its value will still be zero and so that its value at the final time step is also guaranteed to be zero. This is an example of dynamic hedging.

17. Estimate the $\Delta$ and the $\Gamma$ of the American put from Figure 8.3.
18. (Continued) Apply your results from the previous problem to a Taylor polynomial to estimate the change in price of the American put if the asset price immediately (at time 0) increases by 25 cents. (Assume all other factors, like the volatility and risk-free rate, are unchanged.)
Chapter 12

Stochastic calculus

Stochastic calculus combines two major branches of mathematics: probability, which we have seen, and partial differential equations, which we will touch on later. Stochastic calculus tries to frame the dynamics of something that moves with time, given certain known external forces as well as unknown “noise.” If you throw a paper airplane, you could try to model its dynamics, that is, how does it move with time given the effects of wind and gravity. Your answer will probably involve the calculus of differential equations. But if you repeated the experiment, would the plane take the same trajectory? No, because there is some uncertainty in the dynamics. However, the trajectory is not completely random: you know the plane won’t fly to the moon! Balancing the known and unknown forces is what stochastic calculus tries to do.

Stochastic calculus underlies much of the research in modern-day mathematical finance. Because so many people use stochastic calculus, from Wall Street traders to mathematics professors, the theory can be explained with different degrees of rigor and difficulty. Some traders spend their whole careers using stochastic calculus, without knowing the precise definition of Brownian motion, the basic building block of the theory which we cover in the first couple of pages of this chapter. Some professors spend their whole careers proving results about Brownian motion, without ever thinking about the concepts beyond the first couple of pages of this chapter. Since we are only devoting one chapter to this broad topic, we take a lighter approach, which while less thorough, should hopefully instill some intuition on the subject.
12.1 Brownian motion

Suppose you foolishly (but hopefully legally and temporarily!) alter your mind to such an extent that as you walk down the sidewalk of Wall Street, at any point in time, you are just as likely to take a step forward as you are backwards, regardless of where you are, what time it is, or what you have done so far. Further assume that you take a step every second, and that the step size is always one foot (forwards or backwards). This is called a random walk.

Mathematically, the $t$ denote time (in seconds) and $X_t$ your position on the sidewalk. For notational convenience, you start at 0, $X_0 = 0$. Where will you go at time $t = 13$ given where you are at time $t = 12$?

$$P(X_{13} = 9 \mid X_{12} = 8) = \frac{1}{2} = P(X_{13} = 7 \mid X_{12} = 8).$$

Recall from Chapter 9 $X_{12} = 8$ is the event that at that time 12 you are position 8. And $P(X_{13} = 9 \mid X_{12} = 8)$ is the probability that the event $X_{13} = 9$ occurs conditioned on the event $X_{12} = 8$ occurring. So the first equality is just rephrasing that you have a 50% chance of taking a step forward. The statement that you have 50% chance of taking a step forward “regardless of where you are, what time it is, or what you have done so far,” is illustrated, respectively, by the following equations:

$$\frac{1}{2} = P(X_{t+1} = k + 1 \mid X_t = k) \quad \text{for any } k$$
$$\frac{1}{2} = P(X_{t+1} = k + 1 \mid X_t = k) \quad \text{for any } t$$
$$\frac{1}{2} = P(X_{t+1} = k + 1 \mid X_t = k \text{ and } X_{t-1} = k - 1)$$
$$= P(X_{t+1} = k + 1 \mid X_t = k \text{ and } X_{t-1} = k + 1)$$

In that last equation, we only emphasize that where you go, $X_{t+1}$, depends on where are now, $X_t$, but not where you were one second ago, $X_{t-1}$. To be more thorough (which this chapter is not), we would also have to say it does not depend on where you are two seconds ago, $X_{t-2}$, three seconds ago, $X_{t-3}$, etc.

**Example 12.1.** Estimate the probability that after a minute you will have moved at least 8 feet forward.
Solution: We apply the Central Limit Theorem from Section 9.5. Let $Y_t$ be the distance traveled at time $t$. So $P(Y_t = +1) = P(Y_t = -1) = 0.5$ and all the $Y_t$ are independent and identically distributed. Let

$$
\mu = E[Y_t] = \frac{1}{2} \times (-1) + \frac{1}{2} \times 1 = 0
$$

$$
\sigma = SD(Y_t) = \sqrt{E[Y_t^2] - E[Y_t]^2} = \sqrt{\frac{1}{2} \times (-1)^2 + \frac{1}{2} \times 1^2 - 0^2} = 1.
$$

Since $X_{60} = Y_1 + Y_2 + \ldots + Y_{60}$, the Central Limit Theorem implies

$$
X_{60} \approx \mu \times 60 + \sigma \sqrt{60} Z = \sqrt{60} Z
$$

where $Z \sim N(0, 1)$. So

$$
P(X_{60} \geq 8) \approx P(\sqrt{60} Z \geq 8) = 1 - P(\sqrt{60} Z \leq -8) = 1 - F_Z \left( \frac{-8}{\sqrt{60}} \right) = .
$$

Now suppose you take smaller steps more frequently (again with uniform frequency and size). More specifically, suppose you take a step every $\Delta t$ seconds, and the step size is $\sqrt{\Delta t}$. Imagine $\Delta t$ going to 0.

If you are a physics/calculus fan, you may be thinking, “Wait a second! My instantaneous speed is

$$
\text{Speed} = \frac{\text{Displacement}}{\text{Time}} = \left| \frac{X_{t+\Delta} - X_t}{(t+\Delta t) - t} \right| = \frac{\sqrt{\Delta t}}{\Delta t} \rightarrow \infty. \quad (12.1)
$$

What kind of mind-altering substance is this anyway??”

But actually, because you are frequently reversing your direction, in the limit, you are walking (shuffling, really) a continuous path on the Wall Street sidewalk, with no probability of “reaching $\infty$.” With the Central Limit Theorem one can show the following.

Let $Z_t$ denote the limiting process of the random walk, as the time step increment $\Delta t$ and spacial step increment $\sqrt{\Delta t}$ converge to 0. Physically, $Z_t$ denotes your (uncertain) position on the sidewalk at time $t$ as you shuffle along. Then $Z_t$ has the following properties

- $Z_t$ as a limit exists and is continuous with respect to time (not obvious!).
- $Z_0 = 0$. 

The increments between non-overlapping time periods are independent. For example, how you move from $Z_{15} - Z_{12,1}$ is a random variable independent of, for example, $Z_{\pi + 3} - Z_2$ or $Z_{1,1} = Z_{1,1} - Z_0$.

The increment $Z_{t+s} - Z_t$ has a normal distribution $N(0, \sqrt{s})$. Note that this distribution does not depend on $t$.

The process $Z_t$ is called Brownian motion. In fact Brownian motion can be characterized by the above properties, without thinking of it as some limiting random walk process. It is an example of a stochastic process, which is a family of random variables, one for each time $t$. Henceforth, we reserve the symbol $Z_t$ for Brownian motion.

We make a quick notational aside remark. In previous math classes, one uses "$f(x)$" to denote both the function $f(\cdot)$ and the particular real number which this function achieves at the real number $x$. The context clarifies any ambiguity. Similarly, we use "$Y_t$" for both the stochastic process $Y_\cdot$ and the particular random variable which this process achieves at the time $t$. Hopefully, the context clarifies any ambiguity.

Consider the stochastic process $Y_t = 3 + 2Z_t$. This can be viewed as a limiting process (as $\Delta t \to 0$) where the step sizes are twice as big and the initial position is $Y_0 = 3$. In general, for constants $a, b$, the process $Y_t = at + bZ_t$ is called a Weiner process or generalized (versus standard) Brownian motion. It too has independent increments for non-overlapping time periods, and the increments $(Y_{t+s} - Y_t) \sim N(as, b\sqrt{s})$.

**Example 12.2.** Confirm some of these properties of the Weiner process $Y_t = at + bZ_t$ by looking at some specific computations.

**Solution:** We compute the first two moments of $Y_{3+4} - Y_3$.

$$E[Y_{3+4} - Y_3] = E[a((3 + 4) - 3) + b(Z_{3+4} - Z_3)]$$
$$= 4a + b(E[Z_{3+4}] - E[Z_3])$$
$$= 4a + b \cdot (0 - 0).$$

$$SD(Y_{3+4} - Y_3) = SD(a(3 + 4 - 3) + b(Z_{3+4} - Z_t))$$
$$= bSD(Z_{3+4} - Z_3)$$
$$= b\sqrt{4}. $$

Rather than showing $Y_3 - Y_1$ and $Y_1 - Y_0 = Y_1$ are independent (since they have non-overlapping time-periods), we confirm an important property of
12.1. **BROWNIAN MOTION**

independent random variables:

\[
E[(Y_3 - Y_1)(Y_1)] = E[Y_3 - Y_1]E[Y_1].
\]

To get (part of) the third equality below, we use that non-overlapping Brownian steps are independent. To get (part of) the fifth equality below, \(E[Z_1] = 0\) so \(a = a + bE[Z_1]\).

\[
E[(Y_3 - Y_1)(Y_1)] = E[(3a + bZ_3 - 1a - bZ_1)(1a + bZ_1)]
\]
\[
= (2a)(a) + 2abE[Z_1] + abE[Z_3 - Z_1] + b^2E[(Z_3 - Z_1)(Z_1)]
\]
\[
= (2a)(a) + (2ab)(0) + (ab)(0) + b^2E[Z_3 - Z_1]E[Z_1]
\]
\[
= (2a)(a) + (b^2)(0)(0)
\]
\[
= (2a + bE[Z_3 - Z_1])(a + bE[Z_1])
\]
\[
= E[3a + bZ_3 - (a + bZ_1)]E[a + bZ_1]
\]
\[
= E[Y_3 - Y_1]E[Y_1].
\]

**Example 12.3.** Suppose that at time \(t\) the difference in stock price of LED Corp and ZEP Inc is modeled by \(Y_t = 2t + 4Z_t\) where \(Z_t\) is Brownian motion. What is the probability that LED Corp will price at or more then ZEP Inc at time \(t = 3\)?

**Solution:** LED Corp prices at or above ZEP Inc if and only if at \(t = 3\)

\[
P(Y_3 \geq 0) = P((2)(3) + 4Z_3 \geq 0) = P(Z_3 \geq -6/4)
\]

Both \(Z_3\) and \(3Z\) have the same distribution \(N(0,3)\) (where \(Z\) is the standard normal variable), so

\[
P(Z_3 \geq -6/4) = P(3Z \geq -6/4) = P(Z \geq -1/2)
\]
\[
= 1 - P(Z \leq -1/2) = 1 - F_Z(-1/2) = 0.69.
\]

where the numerical result follows from the computer or table.

The example above models the difference of two stock prices as a Wiener process, as oppose to modeling a single stock. This is because while a stock is always non-negative, any Weiner process has a positive probability of becoming negative at any time \(t\) :

\[
P(at + bZ_t < 0) = P(at + btZ < 0) = P\left(Z \leq \frac{-at}{bt}\right) = F_Z(-a/b) > 0.
\]
12.2 A review of the chain rule

At this point, we take a slight deviation from probability theory, and review some concepts from single- and multi-variable calculus.

Consider a single-variable function \( y = f(t) \). The derivative

\[
f'(t) = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}
\]

can be used to approximate \( \Delta y = f(t + \Delta t) - f(t) \), with

\[
\Delta y \approx f'(t) \Delta t
\]

This is the first order approximation of \( \Delta y \) in the Taylor series

\[
\Delta y = f'(t) \Delta t + \frac{1}{2} f''(t) \Delta t^2 + \cdots
\]

Some calculus classes introduce (without formal definitions) the notion of a differential, replacing the approximation \( \Delta y \approx f'(t) \Delta t \) with the equality \( dy = f'(t) dt \). For our purposes, we think of the differential \( dy \) conceptually as a tiny increment, \( \Delta y \), which is small enough to neglect the discrepancy between equalities "=" and approximations "\( \approx \)."

Using the language of the Taylor expansion

\[
dy = f'(t) dt + \frac{1}{2} f''(t) dt^2 + \cdots
\]

we are essentially saying that \( dt^2 = 0, dt^3 = 0, \ldots \). Why does \( dt \neq 0 \) but \( dt^2 = 0? \) Because \( dy \) and \( dt \) are of the same order while \( dt^2, dt^3, \ldots \) are of smaller orders (as \( dt \to 0 \)).

Now consider a multi-variable function \( y = f(t, z) \). We repeat the discussion to conclude that the differential \( dy \) is given by

\[
dy = \frac{\partial f}{\partial t}(t, z) dt + \frac{\partial f}{\partial z}(t, z) dz + \frac{1}{2} \frac{\partial^2 f}{\partial t^2}(t, z) dt^2 + \frac{1}{2} \frac{\partial^2 f}{\partial z^2}(t, z) dz^2 + \frac{\partial^2 f}{\partial t \partial z}(t, z) dtdz + \cdots
\]

\[
= \frac{\partial f}{\partial t}(t, z) dt + \frac{\partial f}{\partial z}(t, z) dz
\]

**Example 12.4.** Suppose \( y = t^3 z \). Choose the small values for the differentials \( dt \) and \( dz \) to both be \( 10^{-6} \). Compute the differential \( dy \) at \((t, z) = (2, -1)\).
Solution: Since $\frac{\partial f}{\partial t} = 3t^2z$ and $\frac{\partial f}{\partial z} = t^3$, we compute

$$\frac{\partial f}{\partial t}(2, -1) = 3 \times 2^2 \times (-1) = -12, \quad \frac{\partial f}{\partial z}(2, -1) = 2^3 = 8.$$ 

Then the differential $dy$ at $(2, -1)$ is

$$dy(2, -1) = \frac{\partial f}{\partial t}(2, -1)dt + \frac{\partial f}{\partial z}(2, -1)dz = (-12) \times 10^{-6} + 8 \times 10^{-6} = 4 \times 10^{-6}.$$ 

Note here that the terms $dt^2 = 10^{-12}, dtdz = 10^{-12}, dz^2 = 10^{-12}, dt^3 = 10^{-18}, dt^2dz = 10^{-18}, \ldots$ in the Taylor expansion are set equal to zero.

After reviewing the derivative (and partial derivatives), the next important concept is the chain rule.

Consider the single-variable functions $y = f(x)$ and $x = g(t)$. The chain rule says how varying $t$ influences $y$, via the effects of the intermediate variable $x$. In class, this often leads to the formula

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}, \quad \text{or} \quad \frac{d(f \circ g)}{dt} = \frac{df}{dx} \frac{dg}{dt}.$$ 

We interpret this in the language of differentials. We have $dy = \frac{dy}{dx}dx$ and $dx = \frac{dx}{dt}dt$. Combining these together we get a reformulation of the chain rule

$$dy = \frac{df}{dx}dx = \frac{df}{dx} \frac{dg}{dt}dt.$$ 

The multi-dimensional versions of the chain rule is similar. Rather than trying to write out the general form that captures all such multi-dimensional chain rules, we focus on one, albeit unconventional example. This example will prove useful later.

Suppose $x = g(t, z)$ and $y = f(t, x)$. Here $x$ and $y$ depend on the same $t$. How does changing $t$ affect $y$? In two different ways: directly, since $t$ appears as the first argument of $f$; and indirectly, via the intermediate variable $x$. We will sort this out more precisely with a multi-dimensional chain rule, in the language of differentials. We have

$$dx = \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial z}dz, \quad dy = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx.$$
Substituting the first into the second gives us the chain rule

\[
dy = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} \left( \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial z} dz \right) = \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \frac{\partial g}{\partial t} \right) dt + \frac{\partial f}{\partial x} \frac{\partial g}{\partial z} dz. \tag{12.2}
\]

**Example 12.5.** Suppose \( x = g(t, z) = t^3z \) and \( y = f(t, x) = 7t + x^2 \). Estimate the change \( \Delta y \) at \((t, z) = (2, -1)\) if \( \Delta t = 10^{-6} = \Delta z \).

**Solution:** The value of \( x \) at the point of interest is \( x = g(2, -1) = -8 \). We compute the partial derivatives at \((t, z) = (2, -1)\) and \((t, x) = (2, -8)\)

\[
\begin{align*}
\frac{\partial f}{\partial t}(2, -8) &= 7, \quad \frac{\partial f}{\partial x}(2, -8) = (2)(-8) = -16 \\
\frac{\partial g}{\partial t}(2, -1) &= (3)(2^2)(-1) = -12, \quad \frac{\partial g}{\partial z}(2, -1) = 2^3 = 8.
\end{align*}
\]

So \( \Delta y \) at \((t, z) = (2, -1)\) is

\[
\Delta y \approx (7 + (-16)(-12))(10^{-6}) + (-12)(8)(10^{-6}) = (199 - 96) \times 10^{-6} = 103 \times 10^{-6}.
\]

**Example 12.6.** Consider the same example. Estimate \( \Delta y \) again but this time when \( \Delta t = 10^{-6} \) and \( \Delta z = 10^{-3} \). What is wrong with this estimate?

**Solution:** The computations of the partial derivatives is unchanged, so

\[
\Delta y \approx 199 \times 10^{-6} - 96 \times 10^{-3} \approx -96 \times 10^{-3}.
\]

There is nothing wrong with this approximation. The “problem” is that the chain rule does not give the right terms. The \( dt \)-term is on the order of \( 10^{-6} \) and the \( dz \)-term is on the order of \( 10^{-3} \). But the \( dz^2 \)-term of the Taylor expansion, which was dropped earlier, is also of the order \((10^{-3})^2 = 10^{-6}\). So either only the \( dz \)-term should be included in the chain rule, or else, the \( dz \), \( dt \)- and \( dz^2 \)-terms should all be included. (The other terms are of the order \( 10^{-9} \) and smaller.)

This example motivates an important result in stochastic calculus: Itô’s Chain Rule.
12.3 Itô’s Chain Rule: a first look

The stochastic version of Section 12.2 is known as “Itô’s Chain Rule” or “Itô’s Lemma.” The formula is very similar to the multi-variable chain rule of equation (12.2), except that in the stochastic case, some of the higher order terms from the Taylor expansion no longer vanish.

Recall how in equation (12.1), the “speed” of the random walk converged to infinity. Another way to state this with Brownian motion is as follows. Let \( \Delta Z = Z_{t+\Delta t} - Z_t \) denote a small “Brownian step” occurring over a small time interval \( \Delta t \).

\[
E[(\Delta Z)^2] = \text{Var}(\Delta Z) - E[\Delta Z]^2 = \Delta t - 0 = \Delta t.
\]

If we go to the limit, where this equation becomes \( dZ^2 = dt \). As with differentials, we will not make the statement “\( dZ^2 = dt \)” precise. Instead, think of it as saying the following: although the product of a small Brownian spatial step with itself is still a random variable, to a very good approximation, it equals the small known time step \( dt \).

We now try to replicate the construction of the chain rule (12.2) in the case of stochastic processes. This time, we will keep track of some of the lower order terms in the Taylor series.

Suppose \( X_t \) is a Weiner process given by \( X_t = at + bZ_t \). Then, letting \( \Delta X = X_{t+\Delta t} - X_t \), we get

\[
\Delta X = X_{t+\Delta t} - X_t \\
= a(t + \Delta t) + bZ_{t+\Delta t} - at - bZ_t \\
= a\Delta t + b\Delta Z.
\]

As \( \Delta t \to 0 \), we restate this using differentials as

\[
dX = adt + bdZ. \quad (12.3)
\]

Recall in Section 9.2 how if \( f : \mathbb{R} \to \mathbb{R} \) is a single-variable function and \( X \) a random variable, then \( Y = f(X) \) is a random variable. We generalize this idea. Let \( f(t, x) \) be a two-variable function and \( X_t \) the stochastic process above (or any stochastic process). Then \( f(t, X_t) \) is another stochastic process: at any time \( t \), we stick the random variable \( X_t \) into \( f(t, \cdot) \) to produce another random variable.
Let $Y_t = f(t, X_t)$ denote this new stochastic process. We want to determine the differential $dY = Y_{t+dt} - Y_t$. Using Taylor Series and without dropping higher terms, we know that

$$dY = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} dt^2 + f_{tx} dtdX + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dX^2 + \ldots \quad (12.4)$$

As when deriving the chain rules in Section 12.2, we input equation (12.3) into equation (12.4) to get

$$dY = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} (adt + bdZ) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (adt + bdZ)^2 + \ldots$$

We might be tempted to just keep the first two terms and set all others equal to zero, as we did in Section 12.2. But recall the comparison of differentials made earlier: $dZ^2 = dt$. This means that the third term $\frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} dZ^2$ cannot be set equal to 0 because it is the same order as the first term $(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} ) dt$. On the other hand, if we reinterpret $dZ^2 = dt$ to mean $dZ$ is the same order as $\sqrt{dt}$, the fourth term is order $dt dZ = (dt)^{\frac{3}{2}}$ which is smaller than the first two terms; thus we can set this equal to 0. The fifth term is of order $dt^2$ which is even smaller; so we set this also equal to zero. We leave it to you to confirm that all the higher terms “…” unmentioned in the Taylor expansion are smaller still. Replace $dZ^2$ with $dt$ and eliminating the fourth and later terms, we have now derived a preliminary version of Itô’s Chain Rule:

$$dY = \left( \frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \left( b \frac{\partial f}{\partial x} \right) dZ. \quad (12.5)$$

**Example 12.7.** Suppose $dX = 4 dt + 5 dZ$ and $X_0 = 0$. Let $f(t, x) = x^2$ and $Y_t = f(t, X_t)$.

1. Compute the differential $dY$.

2. Express the random variable $Y_{t+\Delta t}$ in terms of $Y_t$. 


Solution:

1. First we compute the partial derivatives

$$\frac{\partial f}{\partial t} = 0, \quad \frac{\partial f}{\partial x} = 2x, \quad \frac{\partial^2 f}{\partial x^2} = 2.$$  

Note that the term $\frac{\partial f}{\partial x}$ in equation (12.5) is a function $\frac{\partial f}{\partial x}(t, x)$ evaluated at $(t, x) = (t, X_t)$. That is, the first input is the number $t$ and the second input is the random variable $X_t$. So, for this example

$$\frac{\partial f}{\partial x}(t, X_t) = 2X_t.$$  

Plugging in the other partial derivatives into equation (12.5) we get

$$dY = \left(0 + (4)(2X_t) + \frac{5^2}{2}(2)\right) dt + (4)(2X_t)dZ$$

Recall that for us, $dt = \Delta t$, $dY = Y_{t+\Delta t} - Y_t$, and $dZ = Z_{t+\Delta t} - Z_t$.

2. Note $X_t$ is just a Weiner process given by

$$X_t = X_0 + 4t + 5Z_t = 4t + 5Z_t.$$  

We plug this into $X_t$ in the expression for $dY = \Delta Y = Y_{t+\Delta t} - Y_t$ and isolate the $Y_{t+\Delta t}$ to get

$$Y_{t+\Delta t} = Y_t + (8X_t + 25) \Delta t + (8X_t)(Z_{t+\Delta t} - Z_t)$$

$$= Y_t + (32t + 40Z_t + 25)dt + (32t + 40Z_t)(Z_{t+\Delta t} - Z_t)$$

3. Setting $t = 3$ and $\Delta t = 0.01$, bringing the $Y_t$ back over to the left side, and taking expectations, we get

$$E[Y_{3.01} - Y_3] = E[((32)(3) + 40Z_3 + 25)(0.01) + ((32)(3) + 40Z_3)(Z_{3.01} - Z_3)]$$


$$= 1.21.$$
Here we use some properties of Brownian motion:

\[ E[Z_3] = 0 = E[Z_{3.01} - Z_3]; \]

and,

\[ E[(Z_3)(Z_{3.01} - Z_3)] = E[Z_3]E[Z_{3.01} - Z_3] \]

since \( Z_3 \) is independent of \( Z_{3.01} - Z_3 \). ("What step, \( Z_{3.01} - Z_3 \), we take at time \( t = 3 \) has nothing to do with the location of where we are standing, \( Z_3 \), when we take it.")

12.4 Stochastic differential equations and revisiting Itô’s rule

Just as we reinterpreted regular calculus’ chain rule with differentials in Section 12.2, we can also interpret differential equations in the language of differentials. We first recall/introduce what differential equations are.

An algebraic equation is an equation where the solution is a number. For example, the algebraic equation \( x^2 - 4x + 3 = 0 \) has as solutions the real numbers \( x = 1 \) and \( x = 3 \). In the equation, \( x \) represents a variable which, a priori, can be any real number.

A **differential equation** is an equation where the solution is a function. In addition to squaring the function, adding numbers, and other operations as done in the case of an algebraic equation, one can also differentiate the function, possible multiple times. The function might depend on one or more variables. In this case, if partial derivatives are involved, the equation is called a **partial differential equation** and abbreviated PDE.

**Example 12.8.** Verify that the function \( f(t) = 2e^{3t} \) is a solution to the differential equation \( \frac{df}{dt} - 3f = 0 \).

**Solution:** We plug in \( \frac{df}{dt} = 6e^{3t} \) and \(-3f = -6e^{3t} \) to confirm \( \frac{df}{dt} - 3f = 6e^{3t} - 6e^{3t} = 0 \).

Note that \( f(t) = ke^{3t} \) for any value of \( k \) is a solution. For the differential equation to be unique, we need more information, namely a **boundary condition**.
Example 12.9. Find the (unique) solution to the differential equation \( \frac{df}{dt} - 3f = 0 \) with boundary condition \( f(1) = 7 \).

Solution: We already know that \( f(t) = ke^{3t} \) is a solution to the differential equation. Applying the boundary condition we get

\[
7 = f(1) = ke^{3(1)} \quad \longrightarrow \quad k = 7e^{-3}
\]

So the solution is \( f(t) = 7e^{3(t-1)} \). We do not discuss why it is unique.

We can rewrite this differential equation as an equation of differentials.

\[
\frac{df}{dt} - 3f = 0 \iff \frac{df}{dt} = 3f \iff df = 3f dt.
\]

This equation we can re-interpret as the approximation

\[
f(t + \Delta t) - f(t) = 3f(t)\Delta t, \quad \text{or},
\]

\[
f(t + \Delta t) = f(t) + 3f(t)\Delta t.
\]

The equation for the differential \( df \) could be even more involved. For example, the differential equation

\[
\frac{df}{dt} - 3f + t^4 = 0
\]

has an associated equation of differentials given by

\[
df = (3f - t^4)dt.
\]

Stochastic differentials are similar. Recall the Weiner process \( X_t \) whose differential is given by

\[
dX = adt + bdZ.
\]

Up to now, \( a \) and \( b \) were constants. However, one could define stochastic processes \( X_t \) whose differentials had non-constant coefficients. Not only could \( a \) and \( b \) depend on time, but they can depend on the random variable \( X_t \) itself\(^1\)

\[
dX = a(t, X_t)dt + b(t, X_t)dZ.
\]

\(^1\)For a number of reasons, including the “memory-less” Markov property, people rarely consider the case when \( a \) or \( b \) depend on the stochastic process at other times, such as \( X_{t \pm 1} \).
CHAPTER 12. STOCHASTIC CALCULUS

We call this formula, for given functions \( a(t, x) \) and \( b(t, x) \), a **stochastic differential equation** (SDE) because the stochastic differential \( dX \) is expressed in terms of the stochastic process \( X_t \). (This is similar to the differential equation re-expressed as \( df = 3f dt \).

It turns out that Itô’s Chain Rule still holds in this more general case.

**Formal 22.** Let \( X_t \) be a stochastic process whose differential is given as in equation (12.6). Let \( f(t, x) \) be a function. Define the process \( Y_t = f(t, X_t) \). The stochastic differential is given by

\[
dY = \left( \frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{b^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \left( b \frac{\partial f}{\partial x} \right) dZ.
\]

Note that \( a \) and \( b \) are no longer necessarily constant. They, as well as \( \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x} \) and \( \frac{\partial^2 f}{\partial x^2} \) are each functions of two variables, where the input for the first is time \( t \) and the input for the second is the random variable \( X_t \).

\[
dY = \left( \frac{\partial f}{\partial t}(t, X_t) + a(t, X_t) \frac{\partial f}{\partial x}(t, X_t) + \frac{b^2(t, X_t)}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t) \right) dt
+ \left( b(t, X_t) \frac{\partial f}{\partial x}(t, X_t) \right) dZ.
\]

**Example 12.10.** Fix constants \( \mu \) and \( \sigma \). Consider the Weiner process \( X_t = (\mu - \frac{\sigma^2}{2})t + \sigma Z_t \). Let \( f(t, x) = e^x \). Define the stochastic process

\[ Y_t = f(t, X_t) = e^{(\mu - \frac{\sigma^2}{2})t + \sigma Z_t} \]

Compute the stochastic differential \( dY \).

**Solution:** First we compute the stochastic differential for the process \( X_t \). This follows immediately from thinking of \( dX \) as \( X_{t+dt} - X_t \):

\[
dx = X_{t+dt} - X_t
= (\mu - \frac{\sigma^2}{2})(t + dt - t) + \sigma(Z_{t+dt} - Z_t)
= (\mu - \frac{\sigma^2}{2})dt + \sigma dZ.
\]

Then we compute some partial derivatives

\[
\frac{\partial f}{\partial t}(t, x) = 0, \quad \frac{\partial f}{\partial x}(t, x) = e^x, \quad \frac{\partial^2 f}{\partial x^2} = e^x.
\]
We apply Itô’s Chain Rule to get
\[
dY = \left( \frac{\partial f}{\partial t} + (\mu - \frac{\sigma^2}{2}) \frac{\partial f}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + (\sigma \frac{\partial f}{\partial x}) dZ
\]
\[
= \left( 0 + (\mu - \frac{\sigma^2}{2}) e^{X_t} + \frac{\sigma^2}{2} e^{X_t} \right) dt + \sigma e^{X_t} dZ
\]
\[
= \mu Y_t dt + \sigma Y_t dZ.
\]

We can rephrase this example as solving a stochastic differential equation: the equation
\[
dY = \mu Y_t dt + \sigma Y_t dZ
\]
has as solution the stochastic process
\[
Y_t = e^{\left( \mu - \sigma^2 t \right) t + \sigma Z_t}.
\]

In fact, if we scale this \( Y_t = 17 e^{\left( \mu - \frac{\sigma^2}{2} \right) t + \sigma Z_t} \), we still get a solution. Note that
\[
Y_0 = 17 e^{\left( \mu - \frac{\sigma^2}{2} \right) 0 + \sigma Z_0} = 17 e^{0+0} = 17.
\]
Thus we adopt the convention of writing this scaled solution by
\[
Y_t = Y_0 e^{\left( \mu - \frac{\sigma^2}{2} \right) t + \sigma Z_t}.
\]

Solving a general differential equation is hard. Solving a general SDE is even harder and very few tools are known to do this. We will content ourselves with knowing just a couple of examples of SDEs and their solutions. We summarize the one we just did, known as geometric Brownian motion, below.

**Formula 23 (Geometric Brownian motion).** Geometric Brownian motion is a stochastic process \( Y_t \) which can be characterized two ways. Explicitly, the formula for \( Y_t \) is
\[
Y_t = Y_0 e^{\left( \mu - \frac{\sigma^2}{2} \right) t + \sigma Z_t},
\]
where \( Y_0 \) is given. Alternatively, geometric Brownian motion is the solution to the SDE
\[
dY = \mu Y_t dt + \sigma Y_t dZ.
\]
CHAPTER 12. STOCHASTIC CALCULUS

Recall from Section 12.1 that \( Z_t \sim N(0, \sqrt{t}) \). Thus, for a fixed \( t \), we could replace \( \sigma Z_t \) by \( \sigma \sqrt{t} Z \) where \( Z \sim N(0, 1) \). Plugging that into the equation for \( Y_t \) above, we get the log-normal model for the stock introduced in Section 10.2:

\[
Y_t = Y_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma \sqrt{t} Z}.
\]

Now that \( t \) is a varying time, as oppose to some fixed (expiration) time as in Section 10.2, we need to be careful. If we replace \( \sigma Z_t \) by \( \sigma \sqrt{t} Z \) then we would need to replace \( \sigma Z_t' \) by \( \sigma \sqrt{t} Z' \) where \( Z' \) is a different standard normal variable, \( Z' \sim N(0, 1) \).

**Example 12.11.** Consider the geometric Brownian motion given by the SDE

\[
dY_t = 0.2Y_t dt + 0.4Y_t dZ
\]

\( Y_0 = 17 \).

The time units are in years. Compute the probabilities of the geometric Brownian motion at least doubling from year 1 to year 2, and from year 4 to year 5.

**Solution:** We know from Formula 23 the geometric Brownian motion \( Y_t \) can be written as

\[
Y_t = 17e^{0.2t+0.4Z_t}.
\]

We want to compute the probability that \( Y_2/Y_1 \geq 2 \).

\[
Y_2/Y_1 = \frac{17e^{(0.2)(2)+0.4Z_2}}{17e^{(0.2)(1)+0.4Z_1}} = e^{0.2+0.4(Z_2-Z_1)}
\]

Now \( Z_2 - Z_1 = \sqrt{2-1}Z = Z \) for some standard normal variable \( Z \).

\[
P(Y_2/Y_1 \geq 2) = P(e^{0.2+0.4(Z_2-Z_1)} \geq 2)
\]

\[
= P(Z \geq \frac{\ln(2) - 0.2}{0.4})
\]

\[
= 1 - F_Z \left( \frac{\ln(2) - 0.2}{0.4} \right)
\]

\[
= 0.11
\]

Similarly

\[
Y_5/Y_4 = \frac{17e^{(0.2)(5)+0.4Z_5}}{17e^{(0.2)(4)+0.4Z_4}} = e^{0.2+0.4(Z_5-Z_4)},
\]
and $Z_5 - Z_4$ has the same distribution as $Z_2 - Z_1$ (although it is not the same random variable!), we get

$$P(Y_5/Y_4 \geq 2) = P(Y_2/Y_1 \geq 2) = 0.11.$$  

Geometric Brownian motion is the most important example for introductory math finance. It underlies the log-normal model for a stock price, and is the framework for the Black-Scholes pricing model in Chapters 10 and 13.

Let $B_t$ be the price of a (zero-coupon) bond which pays $1$ at some fixed maturity time $T \geq t$. Then

$$B_t = e^{-r(T-t)} = e^{-rT}e^{rt}$$
and so

$$dB = re^{-rT}e^{rt}dt = rB_t dt$$

This means that, “technically,” we could count the bond price as an example of geometric Brownian motion with $\mu = r$ and $\sigma = 0$.

There are other SDEs that become important when studying volatility models, currency models, or stochastic interest rate models. Since these SDEs are beyond the scope of this text, we only briefly mention one that is commonly used by practitioners.

An Ornstein-Uhlenbeck process $Y_t$ solves the SDE

$$dY = \alpha(\bar{Y} - Y_t)dt + \sigma dZ$$

where $\alpha > 0$, $\sigma$ and $\bar{Y}$ are constants. A central feature that such a process possesses, which geometric Brownian motion does not, is called mean-reversion. Qualitatively, when $Y_t > \bar{Y}$, $dY$ has a negative “drift” of $\alpha(\bar{Y} - Y_t) < 0$, and when $Y_t < \bar{Y}$, $dY$ has a positive drift. This does not mean that $dY$ will be positive (respectively, negative) when its drift is positive (respectively, negative), since there is also the random component $\sigma dZ$. However, one “expects” $Y_t$ to want to drift towards the value $\bar{Y}$, which one can show is its long-term mean.

12.5 Problems

1. Let $Y_t = 3t + 2Z_t$ where $Z_t$ is Brownian motion. Compute the standard deviation $SD(Y_t)$.  


2. Consider a random walk where the increment size in space is $\sqrt{0.01}$ feet and the increment size in time is 0.01 seconds. Estimate the probability that after 5 seconds the random walker will have moved at least 1 foot in the positive direction.

3. “A random walk in the wind.” Consider a “biased” random walk where the increment size in space is $\sqrt{0.01}$ feet and the increment size in time is 0.01 seconds, but this time you have a 70% chance of taking a step forward (with the wind) and only 30% chance of taking a step back (against the wind).

   (a) Let $X_t$ be your position after $t$ seconds where $t$ is some integer. Write $X_t = Y_1 + \ldots + Y_N$ where the $Y_i$ are i.i.d. What is $N$? What is $\mu(Y_i)$? What is $\sigma(Y_i)$?

   (b) According to the CLT, approximate $X_t$ as some normally distributed random variable. What mean and standard deviation should this approximating random variable have?

   (c) What is the Weiner process that approximates this windy random walk? (The answer should be in the form $at + bZ_t$ where $a, b$ are to-be-determined constants and $t$ is measured in seconds.)

4. Suppose that at time $t$ the difference in stock price of LED Corp and ZEP Inc is modeled by $Y_t = 2t - 3Z_t + 1$ where $Z_t$ is Brownian motion. What is the probability that LED Corp is at least 3 more than ZEP Inc at $t = 2$?

5. Let $x = f(t, z) = e^t z$ and $y = g(t, x) = t^2 x^3$. Estimate

   $$g(2.01, f(2.01, 2.97)) - g(2, f(2, 3))$$

   using the techniques and notation of Section 12.2.

6. Assume the risk-free rate is 5% per annum, continuously compounded. Let $B_t$ be the price at time $t \leq 5$ of a (zero-coupon) bond paying $\$1$ at time $T = 5$. Let $f(t, x) = t^2 x^4 + x^2$. Let $Y_t = f(t, B_t)$. Compute the differential $dY$. (Note: you can do this without stochastic calculus although you might as well practice using the techniques of this section.)
7. Assume the risk-free rate is 5% per annum, continuously compounded. Let $S_t$ be the price of an asset given by $S_0 = e^{-0.05 \times 5}$ and

$$dS = 0.05S_t dt + 0.16S_t dZ.$$ 

Let $B_t$ be the price at time $t \leq 5$ of a (zero-coupon) bond paying $1$ at time $T = 5$. Note that $S_0 = B_0$. Compute the probability that the asset will outperform the bond after the first three years.

8. Consider the process $X_t$ given by $X_0 = 3$ and

$$dX = t^2 dt + 3dZ.$$ 

Let $f(t, x) = e^{tx}$. Let $Y_t = f(t, X_t)$. Compute $Y_0$ and the differential $dY$.

9. Let $S_t$ be a geometric Brownian motion process given by $S_0 = 2$ and

$$dS = 0.04S_t dt + 0.16S_t dZ.$$ 

Let $f(t, x) = t^2x^4 + x^2$. Let $Y_t = f(t, S_t)$. Compute the differential $dY$.

10. Let $S_t$ be a geometric Brownian motion process given by $S_0 = 2$ and

$$dS = 0.04S_t dt + 0.16S_t dZ.$$ 

Let $f(t, x) = 4 \ln(x^3)$. Let $Y_t = f(t, S_t)$. Show that $Y_t$ is a Weiner process. That is, show that

$$dY = adt + bdZ$$ 

and compute the constants $a$ and $b$.

11. Let $S_t$ be a geometric Brownian motion process given by $S_0 = 2$ and

$$dS = 0.04S_t dt + 0.16S_t dZ.$$ 

Let $f(t, x) = x^2$. Let $Y_t = f(t, S_t)$. Show that $Y_t$ is also a geometric Brownian motion process. That is, show that

$$dY = \mu Y_t dt + \sigma Y_t dZ$$ 

and compute the constants $\mu$ and $\sigma$.

12. Let $X_t$ be an Ornstein-Uhlenbeck process given by $X_0 = 2$ and

$$dX = 0.4(3 - X_t) dt + 0.2dZ.$$ 

Let $f(t, x) = x^2$. Let $Y_t = f(t, X_t)$. Approximate $E[Y_{0.01}]$. 

12.5. PROBLEMS
Chapter 13

The Black-Scholes partial differential equation

In this chapter we use the stochastic calculus of Chapter 12 to derive the celebrated Black-Scholes partial differential equation (PDE). This is an equation relating several partial derivatives of the option (or any financial security) price $h(t, s)$, viewed as a function of time $t$ and the underlying asset price $s$. Although we derive the equation, we will not solve it.

Why go through the effort to derive the PDE and not try to solve it to get a formula for the option price $h$? Besides, didn’t we already derive an explicit formula for the option price $h$ via risk-neutral evaluation in Formula 17?

There are several answers that motivate this chapter.

First, the expectation computation underlying Formula 17 is only doable for certain derivative securities, notably European options and a few other derivatives with simply pay-offs. The Black-Scholes PDE applies not just to these, but to any derivative security price. The difference in the securities’ prices reflected in the different boundary conditions of the PDEs. (Boundary conditions were introduced in Section 12.4.)

Methods from an advanced undergraduate course in PDEs should be sufficient to solve the Black-Scholes PDE. However, many naturally occurring PDEs cannot be explicitly solved. This include PDEs from security pricing when we relax the model assumptions on our underlying asset price. For example, suppose the asset price could jump on days when certain information is released to the market. Depending what we say about these jumps can greatly change the option price PDE. For this reason and others, there
is a wealth of information on how to approximate the price \( h \) given its PDE, without actually knowing the explicit formula for \( h \). An approximation is often just as good for a practitioner.

There are other qualitative aspects one can determine about the price \( h \) just by knowing that it is a solution to a certain PDE and without having an explicit formula. We will discuss one aspect involving the Greeks.

As a purely theoretical motivation, note that in Chapter 10 we did not actually prove that for the log-normal asset price, the derivative price is the discounted risk-neutral expectation of the pay-off. We merely stated this fact, and referred to the justification from Chapter 8 that proved this when the asset price follows a binomial tree.

### 13.1 Deriving the Black-Scholes PDE

Let \( S_t \) denote the underlying asset price which we assume log-normal. Recall from Formula 23 and the subsequent discussion that this implies

\[
S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma Z_t} \quad \text{as well as} \quad dS = \mu S_t dt + \sigma S_t dZ.
\]

The constants \( \mu \) and \( \sigma \) denote the asset’s expected rate of return and volatility, per annum. Since today is time \( t = 0 \), the asset spot price \( S_0 \) is known.

Consider a financial derivative whose pay-off at time \( T > 0 \) is some function \( \Phi(S_T) \) of the underlying asset price \( S_T \) at expiration time \( T \). We assume that at time \( 0 \leq t \leq T \), the price of this derivative \( h \) is some function of the time \( t \) as well as the underlying asset price \( S_t \). This means that the derivative price \( h(t, S_t) \) is a stochastic process. As with all stochastic processes in this book, we denote it by some capital letter, say \( Y_t = h(t, S_t) \).

The goal is derive a PDE involving the function \( h(t, s) \), since ultimately, we want to know as much as possible about the derivative’s price today, which is \( h(0, S_0) \).

Suppose a trader writes one of these derivatives. The value of the trader’s position at time \( t \) is \(-h(t, S_t)\). Not wishing to be exposed to the risk of such a position, the trader delta hedges, acquiring \( \frac{\partial h}{\partial s} \) shares of the asset. See Section 11.1.3 to review why this is delta hedging. Section 11.3 points out how delta hedging is exactly what we did when we priced derivatives in the binomial tree framework.
13.1. DERIVING THE BLACK-SCHOLES PDE

The difference between the current hedging discussion and the ones in Sections 11.1.3 and 11.3 is that the trader must, at every increment of time \( \Delta t \), rehedge to ensure the portfolio is delta-neutral. (Delta-neutrality is defined in Section 11.1.3). This is called dynamic hedging, as oppose to static hedging discussed in the Chapter and Section 11.3. Because the position is always delta-neutral, the trader does not have to worry about the other Greeks, such as \( \Gamma \).

Let \( \Pi \) denote the value of this hedged position. So at time \( t \), the value is

\[
\Pi = -Y_t + \frac{\partial h}{\partial s} S_t.
\]

Note that \( \Pi \) changes with time. Indeed, the trader’s portfolio value is a stochastic process which is some (unknown so far to us) function of \( t \) and \( S_t \). To reflect this, we denote this portfolio value process by \( \Pi_t \).

Suppose the hedging does not change between time \( t \) and \( t + \Delta t \). Then the change in portfolio value \( \Pi_{t+\Delta t} - \Pi_t \) between time \( t \) and \( t + \Delta t \) equals the change in the value of the short derivative position plus the change in value of the \( \frac{\partial h}{\partial s} \) shares of the underlying asset.

\[
\Pi_{t+\Delta t} - \Pi_t = -(Y_{t+\Delta t} - Y_t) + \frac{\partial h}{\partial s} (S_{t+\Delta t} - S_t)
\]

In the language of differentials, this means

\[
d\Pi = -dY + \frac{\partial h}{\partial s} dS.
\]

Recall \( S_t \) is geometric Brownian motion

\[
dS = \mu S_t dt + \sigma S_t dZ.
\]

Even though we do not know the function \( h \), we can still apply Itô’s Chain Rule to \( Y_t = h(t, S_t) \) to determine the differential

\[
dY = \left( \frac{\partial h}{\partial t} + \mu S_t \frac{\partial h}{\partial s} + \frac{(\sigma S_t)^2 \partial^2 h}{2 \partial s^2} \right) dt + \left( \sigma S_t \frac{\partial h}{\partial s} \right) dZ.
\]

Plugging the equations for \( dS \) and \( dY \) into the formula for \( d\Pi \) we get

\[
d\Pi = - \left( \frac{\partial h}{\partial t} + \mu S_t \frac{\partial h}{\partial s} + \frac{(\sigma S_t)^2 \partial^2 h}{2 \partial s^2} \right) dt - \left( \sigma S_t \frac{\partial h}{\partial s} \right) dZ
\]

\[
+ \frac{\partial h}{\partial s} (\mu S_t dt + \sigma S_t dZ)
\]

\[
= - \left( \frac{\partial h}{\partial t} + \frac{(\sigma S_t)^2 \partial^2 h}{2 \partial s^2} \right) dt
\]
CHAPTER 13. THE BLACK-SCHOLES PARTIAL DIFFERENTIAL EQUATION

Now comes the crux of the argument. Note that
\[ d\Pi = \ldots dt + 0dZ. \]

Because there is no Brownian motion component, the portfolio value is not at all random. The short position in the derivative is perfectly hedged with \( \frac{\partial h}{\partial s} \) assets. This means the portfolio is risk-free. The no-arbitrage assumption implies that the portfolio must grow at the risk-free rate:
\[ \Pi_t = \Pi_0 e^{rt} \longrightarrow d\Pi = r\Pi_t dt. \]

Comparing the two expressions for the differentials \( d\Pi \) we get
\[
\begin{align*}
- \left( \frac{\partial h}{\partial t} + \frac{(\sigma S_t)^2}{2} \frac{\partial^2 h}{\partial s^2} \right) dt &= r\Pi_t dt \\
&= r \left( -h + \frac{\partial h}{\partial s} S_t \right) dt.
\end{align*}
\]

Bring all terms to one side and divide by \( dt \) to get
\[ \frac{\partial h}{\partial t} + \frac{(\sigma S_t)^2}{2} \frac{\partial^2 h}{\partial s^2} + rS_t \frac{\partial h}{\partial s} - rh = 0. \]

This is true for whatever the asset price \( S_t \) happens to be. Let us denote this value as the variable \( s \).

**Formula 24** (Black-Scholes PDE). Let \( h = h(t, s) \) denote the price of derivative security at time \( t \) when the underlying asset price is \( s \). Then \( h \) is a solution to the Black-Scholes partial differential equation
\[ \frac{\partial h}{\partial t} + \frac{(\sigma s)^2}{2} \frac{\partial^2 h}{\partial s^2} + rs \frac{\partial h}{\partial s} - rh = 0. \]

If the pay-off function at terminal time \( T \) is given by \( \Phi(S_T) \), then \( h \) also satisfies the boundary condition
\[ h(T, s) = \Phi(s). \]

**Example 13.1.** Suppose the risk-free rate is 5% per annum with continuous compounding. The annual volatility of an asset with price \( S_t \) is 20%. The expected rate of return is 10% per annum. Consider a derivative whose pay-off is \( S_t^{3/2} - S_T \). Write the SDE for \( S_t \) and the Black-Scholes PDE for the derivative price.
13.1. DERIVING THE BLACK-SCHOLES PDE

Solution: The asset price $S_t$ satisfies

$$dS_t = 0.1S_t dt + 0.2S_t dZ.$$ 

The pricing PDE is

$$\frac{\partial h}{\partial t} + \frac{(0.2s)^2}{2} \frac{\partial^2 h}{\partial s^2} + 0.05s \frac{\partial h}{\partial s} - 0.05h = 0.$$ 

with boundary condition

$$h(T, s) = s^{3.2} - s.$$ 

Note that the expected rate of return of 10% does not show up in the PDE, and hence does not show up in the formula for the derivative price $h(t, s)$. This may seem at first surprising, but recall that when we priced derivatives in Chapters 8 and 10, we used risk-neutral evaluation. That is, we left the real world where

$$dS_t = \mu S_t dt + \sigma S_t dZ.$$ 

and computed expectations in the risk-neutral world where

$$dS_t = rS_t dt + \sigma S_t dZ.$$ 

So the fact that $\mu = 0.1$ is no longer relevant for our derivative pricing.

Example 13.2. Let $h$ denote the price of a derivative that pays you 3 times the amount of the asset price in 9 months. The expected rate of return of the asset of 10%, the risk-free rate is 6% and the annual volatility is 16%. Verify that the price of the derivative satisfies the Black-Scholes PDE.

Solution: This derivative is similar to one of the examples from Section 10.4. Recall from risk-neutral valuation that the price is

$$h = e^{-0.06(9/12)} E_{RNW} [3S_{9/12}] = (e^{-0.06(9/12)}) (3e^{0.06(9/12)}S_0) = 3S_0.$$ 

At a general time $t$, is the asset price if $S_t = s$, then the derivative price is $h(t, s) = 3s$. We can see this either from generalizing the expectation computation above (replacing time 0 with time $t$) or from the following straightforward argument: if the derivative pays you three times the asset price at
expiration, no matter what, then it is equivalent to having 3 shares of the asset, whose price at time \( t \) is 3\( s \).

To check that \( h(t, s) \) satisfies the Black-Scholes PDE, we compute the partial derivatives
\[
\frac{\partial h}{\partial t} = 0, \quad \frac{\partial h}{\partial s} = 3, \quad \frac{\partial^2 h}{\partial s^2} = 0.
\]
Plugging these into the left-hand side of Black-Scholes equation in Formula 24 we get
\[
0 + (0.16s^2) \times 0 + (0.06)s \times 3 - (0.06)(3s) = 0
\]
confirming the equation holds.

13.2 More on the Black-Scholes PDE

13.2.1 Black-Scholes and the Greeks

Recall the Greeks from Chapter 11
\[
\Theta = \frac{\partial h}{\partial T}, \quad \Gamma = \frac{\partial^2 h}{\partial s^2}, \quad \Delta = \frac{\partial h}{\partial s}.
\]
\( \Theta \) measures how the derivative price changes as the time until expiration increases. Now a small decrease until expiration can be achieved two ways: either decrease \( T \) to \( T - \Delta t \); or increase the current time \( t \) (or 0) to \( t + \Delta t \) (or 0 = \( \Delta t \)). Thus
\[
\Theta = \frac{\partial h}{\partial T} = -\frac{\partial h}{\partial t}.
\]

With this observation in mind, we can translate the Black-Scholes PDE
\[
\frac{\partial h}{\partial t} + \frac{(\sigma s)^2}{2} \frac{\partial^2 h}{\partial s^2} + rs \frac{\partial h}{\partial s} - rh = 0.
\]
into a relationship involving three of the Greeks, the asset price \( s \), and the derivative price \( h \):
\[
-\Theta + \frac{(\sigma s)^2}{2} \Gamma + rs\Delta - rh = 0.
\]
This can be applied in a number of different ways, some of which we defer until the exercises.
Example 13.3. Assume the Black-Scholes PDE holds. The underlying asset is $50 with volatility 15%. A call option has a price of $5. It has a $\Theta$ of $2$ and a $\Gamma$ of $0.05$. The risk-free rate is 5%. Without any using the cumulative distribution function of $N(0,1)$, estimate the change in price of the option if the asset price immediately jumps by 50 cents. Assume all other factors remain constant.

Solution: We compute the $\Delta$ of the call to be

$$\Delta = -\Theta - \frac{(\sigma s)^2}{2} \times \Gamma + rh = 2 - \frac{(0.15 \times 50)^2}{2} \times 0.05 + 0.05 \times 5 = 0.89$$

The new price $h(50,50)$ compared to the original price $h(50)$ is then approximately

$$h(50,50) \approx h(50) + \Delta \times 0.50 + \frac{1}{2} \Gamma \times 0.50^2$$

$$= 5 + (0.89)(0.50) + \frac{0.05 \times 0.50^2}{2}$$

$$= 5.45$$

13.2.2 Dividends

Recall that the underlying asset may distribute dividends, in either lump sums if the asset represents an individual company, or as a continuous stream, if the asset represents a market index. The formula for the option price is adjusted accordingly. See Section 8.5 for the binomial tree model, and the end of Section 10.4 for the log-normal model.

Suppose the asset pays a constant dividend yield of $q$ per annum. Then one can show that the differential $dS$ in the risk-neutral world is

$$dS = (r - q)S_t dt + \sigma S_t dZ$$

where as usual, $r$ is the risk-free rate and $\sigma$ the asset volatility. The Black-Scholes PDE has to be slightly modified

$$\frac{\partial h}{\partial t} + \frac{(\sigma s)^2}{2} \frac{\partial^2 h}{\partial s^2} + (r - q)s \frac{\partial h}{\partial s} - rh = 0.$$
Example 13.4. The underlying (index) asset price $S_t$ has volatility 10%, expected rate of return 9%, and dividend yield $q$. The risk-free rate is 6%. There is a derivative whose price $h$ is given by $h(t,s) = e^{-0.06t}s^9$. Determine $q$.

Solution: First we compute the Greeks of $h$.

$$\frac{\partial h}{\partial t} = -0.06e^{-0.06t}s^9, \quad \frac{\partial h}{\partial s} = 9e^{-0.06t}s^8, \quad \frac{\partial^2 h}{\partial s^2} = 72e^{-0.06t}s^7.$$  

Plugging these into the modified Black-Scholes PDE, we get

$$0 = -0.06e^{-0.06t}s^9 - \frac{(0.1)^2}{2}(72e^{-0.06t}s^7) + (0.06 - q)s\left(9e^{-0.06t}s^8\right) - 0.06e^{-0.06t}s^9$$

We divide all terms by $e^{-0.06t}s^9$ to get

$$0 = -0.06 - \frac{(0.1)^2}{2}(72) + (0.06 - q)(9) - 0.06$$

We then solve $q = 0.06(-1 + 9 - 1)/9 - 0.36/9 = 0.67\%$.

13.2.3 Hedging generalizations

The idea of hedging a derivative with the underlying asset can be generalized. For example, sometimes there are derivatives whose underlying asset is not traded, such as “weather derivatives” whose pay-off is determined by some location’s temperature.

If a bank writes such a derivative, it could hedge its short position have with a pre-existing tradable derivative on the same underlying. The following example does not exactly fit into this picture, since both of the example’s derivatives are traded and so each has a price process; however, it illustrates part of the math involved.

Example 13.5. Consider a derivative $X$ and a derivative $Y$ whose price processes $X_t$ and $Y_t$ are geometric Brownian motions

$$dX = 0.1X_tdt + 0.2X_tdZ$$
$$dY = 0.08Y_tdt + 0.15Y_tdZ$$

The price today is $X_0 = 50$ and $Y_0 = 40$. Suppose a bank writes one of derivative $X$. What is the initial position the bank should take on the other derivative to make its portfolio risk-free?
Solution: As done earlier in this section, the portfolio is risk-free when the coefficient in front of \( dZ \) is 0. Suppose the portfolio has price \( \Pi_t \) and is made of \(-1\) shares of derivative \( X \) and \( a \) shares of derivative \( Y \). Then
\[
 d\Pi = -dX + adY = -0.1X_t dt - 0.2X_t dZ + a \times 0.08Y_t dt + a \times 0.15Y_t dZ
\]
At time 0, we need the coefficient for \( dZ \) to be 0
\[
0 = -0.2 \times X_0 + a \times 0.15Y_0 = -10 + 6a
\]
or \( a = 10/6 \).

13.3 Problems

For all of these problems, we assume the Black-Scholes PDE holds.

1. The underlying asset is $50 with volatility 15%. An at-the-money call option has a price of $5. The call has a \( \Theta \) of 2.65 and a \( \Gamma \) of 0.05. The risk-free rate is 5%. Compute the \( \Delta \) and \( \Gamma \) of the at-money-put and use them to estimate the change in price of the put if the asset price immediately jumps to 50.25, assuming all other factors remain constant.

2. Recall the definition of implied volatility introduced in the problems on Section 10.7. Assume the underlying asset price is 50, A call of unspecified strike has a price of $5. The call has a \( \Theta \) of 2, a \( \Delta \) of 0.5 and a \( \Gamma \) of 0.05. The risk-free rate is 5% Deduce the volatility implied by the Black-Scholes PDE.

3. The underlying asset spot price is $80. There are two call options on the same underlying asset. Call \( A \), pricing at $12, has the following Greeks: \( \Delta_A = 0.5, \Gamma_A = 0.0103, \Theta_A = 3 \). Call \( B \), pricing at $15, has the following Greeks: \( \Delta_B = 0.5, \Gamma_B = 0.0195, \Theta_B = 4 \). Determine the risk-free rate.

4. Estimate the \( \Theta \) of the European put from Figure 8.2 using the computations done in Section 11.3. Assume the volatility is 28%. \[\text{This is approximately correct since the annual vol can be approximated by } \sqrt{1/\Delta S/S} \text{ close to } 1/\sqrt{3} \times 0.56 \times 1.2^2 + 0.44 \times 0.8^2 - (0.56 \times 1.2 + 0.44 \times 0.8)^2 = 0.28.\]
5. Recall the “power derivative” example from Section 10.4. Compute the price of a power derivative whose payoff at expiration $T$ is $S_T^3$. Verify that your solution satisfies the Black-Scholes PDE.

Note: you will need a formula for $h(t, s)$, not just $h(0, s)$ because you will need to compute $\frac{\partial h}{\partial t}$. The method of computation is similar except that everywhere that $T$ appears in the formula $h(0, s)$, you replace it with a $(T - t)$ to get the formula for $h(t, s)$. 