

An Undergraduate Introduction to Financial Mathematics



J Robert Buchanan

An Undergraduate
Introduction to
Financial Mathematics



This page is intentionally left blank

AN UNDERSTANDING OF
THE INSTRUCTIONS TO
Financial Mathematics

J Robert Buchanan

Millersville University, USA



NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

AN UNDERGRADUATE INTRODUCTION TO FINANCIAL MATHEMATICS

Copyright © 2006 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-256-637-6

Dedication

For my wife, Monika.

This page is intentionally left blank

Preface

This book is intended for an audience with an undergraduate level of exposure to calculus through elementary multivariable calculus. The book assumes no background on the part of the reader in probability or statistics. One of my objectives in writing this book was to create a readable, reasonably self-contained introduction to financial mathematics for people wanting to learn some of the basics of option pricing and hedging. My desire to write such a book grew out of the need to find an accessible book for undergraduate mathematics majors on the topic of financial mathematics. I have taught such a course now three times and this book grew out of my lecture notes and reading for the course. New titles in financial mathematics appear constantly, so in the time it took me to compose this book there may have appeared several superior works on the subject. Knowing the amount of work required to produce this book, I stand in awe of authors such as those.

This book consists of ten chapters which are intended to be read in order, though the well-prepared reader may be able to skip the first several with no loss of understanding in what comes later. The first chapter is on interest and its role in finance. Both discretely compounded and continuously compounded interest are treated there. The book begins with the theory of interest because this topic is unlikely to scare off any reader no matter how long it has been since they have done any formal mathematics.

The second and third chapters provide an introduction to the concepts of probability and statistics which will be used throughout the remainder of the book. Chapter Two deals with discrete random variables and emphasizes the use of the binomial random variable. Chapter Three introduces continuous random variables and emphasizes the similarities and differences between discrete and continuous random variables. The nor-

mal random variable and the close related lognormal random variable are introduced and explored in the latter chapter.

In the fourth chapter the concept of arbitrage is introduced. For readers already well versed in calculus, probability, and statistics, this is the first material which may be unfamiliar to them. The assumption that financial calculations are carried out in an “arbitrage free” setting pervades the remainder of the book. The lack of arbitrage opportunities in financial transactions ensures that it is not possible to make a risk free profit. This chapter includes a discussion of the result from linear algebra and operations research known as the Duality Theorem of Linear Programming.

The fifth chapter introduces the reader to the concepts of random walks and Brownian motion. The random walk underlies the mathematical model of the value of securities such as stocks and other financial instruments whose values are derived from securities. The choice of material to present and the method of presentation is difficult in this chapter due to the complexities and subtleties of stochastic processes. I have attempted to introduce stochastic processes in an intuitive manner and by connecting elementary stochastic models of some processes to their corresponding deterministic counterparts. Itô’s Lemma is introduced and an elementary proof of this result is given based on the multivariable form of Taylor’s Theorem. Readers whose interest is piqued by material in Chapter Five should consult the bibliography for references to more comprehensive and detailed discussions of stochastic calculus.

Chapter Six introduces the topic of options. Both European and American style options are discussed though the emphasis is on European options. Properties of options such as the Put/Call Parity formula are presented and justified. In this chapter we also derive the partial differential equation and boundary conditions used to price European call and put options. This derivation makes use of the earlier material on arbitrage, stochastic processes and the Put/Call Parity formula.

The seventh chapter develops the solution to the Black-Scholes PDE. There are several different methods commonly used to derive the solution to the PDE and students benefit from different aspects of each derivation. The method I choose to solve the PDE involves the use of the Fourier Transform. Thus this chapter begins with a brief discussion of the Fourier and Inverse Fourier Transforms and their properties. Most three- or four-semester elementary calculus courses include at least an optional section on the Fourier Transform, thus students will have the calculus background necessary to follow this discussion. It also provides exposure to the Fourier

Transform for students who will be later taking a course in PDEs and more importantly exposure for students who will not take such a course. After completing this derivation of the Black-Scholes option pricing formula students should also seek out other derivations in the literature for the purposes of comparison.

Chapter Eight introduces some of the commonly discussed partial derivatives of the Black-Scholes option pricing formula. These partial derivatives help the reader to understand the sensitivity of option prices to movements in the underlying security's value, the risk-free interest rate, and the volatility of the underlying security's value. The collection of partial derivatives introduced in this chapter is commonly referred to as "the Greeks" by many financial practitioners. The Greeks are used in the ninth chapter on hedging strategies for portfolios. Hedging strategies are used to protect the value of a portfolio against movements in the underlying security's value, the risk-free interest rate, and the volatility of the underlying security's value. Mathematically the hedging strategies remove some of the low order terms from the Black-Scholes option pricing formula making it less sensitive to changes in the variables upon which it depends. Chapter Nine will discuss and illustrate several examples of hedging strategies.

Chapter Ten extends the ideas introduced in Chapter Nine by modeling the effects of correlated movements in the values of investments. The tenth chapter discusses several different notions of optimality in selecting portfolios of investments. Some of the classical models of portfolio selection are introduced in this chapter including the Capital Assets Pricing Model (CAPM) and the Minimum Variance Portfolio.

It is the author's hope that students will find this book a useful introduction to financial mathematics and a springboard to further study in this area. Writing this book has been hard, but intellectually rewarding work.

During the summer of 2005 a draft version of this manuscript was used by the author to teach a course in financial mathematics. The author is indebted to the students of that class for finding numerous typographical errors in that earlier version which were corrected before the camera ready copy was sent to the publisher. The author wishes to thank Jill Bachstadt, Jason Buck, Mark Elicker, Kelly Flynn, Jennifer Gomulka, Nicole Hundley, Alicia Kasif, Stephen Kluth, Patrick McDevitt, Jessica Paxton, Christopher Rachor, Timothy Refi, Pamela Wentz, Joshua Wise, and Michael Zrncic.

A list of errata and other information related to this book can be found at a web site I created:

<http://banach.millersville.edu/~bob/book/>

Please feel free to share your comments, criticism, and (I hope) praise for this work through the email address that can be found at that site.

J. Robert Buchanan
Lancaster, PA, USA
October 31, 2005

Contents

<i>Preface</i>	vii
1. The Theory of Interest	1
1.1 Simple Interest	1
1.2 Compound Interest	3
1.3 Continuously Compounded Interest	4
1.4 Present Value	5
1.5 Rate of Return	11
1.6 Exercises	12
2. Discrete Probability	15
2.1 Events and Probabilities	15
2.2 Addition Rule	17
2.3 Conditional Probability and Multiplication Rule	18
2.4 Random Variables and Probability Distributions	21
2.5 Binomial Random Variables	23
2.6 Expected Value	24
2.7 Variance and Standard Deviation	29
2.8 Exercises	32
3. Normal Random Variables and Probability	35
3.1 Continuous Random Variables	35
3.2 Expected Value of Continuous Random Variables	38
3.3 Variance and Standard Deviation	40
3.4 Normal Random Variables	42
3.5 Central Limit Theorem	49

3.6	Lognormal Random Variables	51
3.7	Properties of Expected Value	55
3.8	Properties of Variance	58
3.9	Exercises	61
4.	The Arbitrage Theorem	63
4.1	The Concept of Arbitrage	63
4.2	Duality Theorem of Linear Programming	64
4.2.1	Dual Problems	66
4.3	The Fundamental Theorem of Finance	72
4.4	Exercises	74
5.	Random Walks and Brownian Motion	77
5.1	Intuitive Idea of a Random Walk	77
5.2	First Step Analysis	78
5.3	Intuitive Idea of a Stochastic Process	91
5.4	Stock Market Example	95
5.5	More About Stochastic Processes	97
5.6	Itô's Lemma	98
5.7	Exercises	101
6.	Options	103
6.1	Properties of Options	104
6.2	Pricing an Option Using a Binary Model	107
6.3	Black-Scholes Partial Differential Equation	110
6.4	Boundary and Initial Conditions	112
6.5	Exercises	114
7.	Solution of the Black-Scholes Equation	115
7.1	Fourier Transforms	115
7.2	Inverse Fourier Transforms	118
7.3	Changing Variables in the Black-Scholes PDE	119
7.4	Solving the Black-Scholes Equation	122
7.5	Exercises	127
8.	Derivatives of Black-Scholes Option Prices	131
8.1	Theta	131
8.2	Delta	133

8.3	Gamma	135
8.4	Vega	136
8.5	Rho	138
8.6	Relationships Between Δ , Θ , and Γ	139
8.7	Exercises	141
9.	Hedging	143
9.1	General Principles	143
9.2	Delta Hedging	145
9.3	Delta Neutral Portfolios	149
9.4	Gamma Neutral Portfolios	151
9.5	Exercises	153
10.	Optimizing Portfolios	155
10.1	Covariance and Correlation	155
10.2	Optimal Portfolios	164
10.3	Utility Functions	165
10.4	Expected Utility	171
10.5	Portfolio Selection	173
10.6	Minimum Variance Analysis	177
10.7	Mean Variance Analysis	186
10.8	Exercises	191
	Appendix A Sample Stock Market Data	195
	Appendix B Solutions to Chapter Exercises	203
B.1	The Theory of Interest	203
B.2	Discrete Probability	206
B.3	Normal Random Variables and Probability	212
B.4	The Arbitrage Theorem	225
B.5	Random Walks and Brownian Motion	231
B.6	Options	235
B.7	Solution of the Black-Scholes Equation	239
B.8	Derivatives of Black-Scholes Option Prices	245
B.9	Hedging	249
B.10	Optimizing Portfolios	255
	Bibliography	265
	Index	267

Chapter 1

The Theory of Interest

One of the first types of investments that people learn about is some variation on the savings account. In exchange for the temporary use of an investor's money, a bank or other financial institution agrees to pay **interest**, a percentage of the amount invested, to the investor. There are many different schemes for paying interest. In this chapter we will describe some of the most common types of interest and contrast their differences. Along the way the reader will have the opportunity to renew their acquaintanceship with exponential functions and the geometric series. Since an amount of capital can be invested and earn interest and thus numerically increase in value in the future, the concept of **present value** will be introduced. Present value provides a way of comparing values of investments made at different times in the past, present, and future. As an application of present value, several examples of saving for retirement and calculation of mortgages will be presented. Sometimes investments pay the investor varying amounts of money which change over time. The concept of **rate of return** can be used to convert these payments in effective interest rates, making comparison of investments easier.

1.1 Simple Interest

In exchange for the use of a depositor's money, banks pay a fraction of the account balance back to the depositor. This fractional payment is known as **interest**. The money a bank uses to pay interest is generated by investments and loans that the bank makes with the depositor's money. Interest is paid in many cases at specified times of the year, but nearly always the fraction of the deposited amount used to calculate the interest is called the **interest rate** and is expressed as a percentage paid per year.

For example a credit union may pay 6% annually on savings accounts. This means that if a savings account contains \$100 now, then exactly one year from now the bank will pay the depositor \$6 (which is 6% of \$100) provided the depositor maintains an account balance of \$100 for the entire year.

In this chapter and those that follow, interest rates will be denoted symbolically by r . To simplify the formulas and mathematical calculations, when r is used it will be converted to decimal form even though it may still be referred to as a percentage. The 6% annual interest rate mentioned above would be treated mathematically as $r = 0.06$ per year. The initially deposited amount which earns the interest will be called the **principal amount** and will be denoted P . The sum of the principal amount and any earned interest will be called the **compound amount** and A will represent it symbolically. Therefore the relationship between P , r , and A for a single year period is

$$A = P + Pr = P(1 + r).$$

The interest, once paid to the depositor, is the depositor's to keep. Banks and other financial institutions "pay" the depositor by adding the interest to the depositor's account. Unless the depositor withdraws the interest or some part of the principal, the process begins again for another interest period. Thus two interest periods (think of them as years) after the initial deposit the compound amount would be

$$A = P(1 + r) + P(1 + r)r = P(1 + r)^2.$$

Continuing in this way we can see that t years after the initial deposit of an amount P , the compound amount A will grow to

$$A = P(1 + r)^t \tag{1.1}$$

This is known as the **simple interest formula**. A mathematical "purist" may wish to establish Eq. (1.1) using the principle of induction.

Banks and other interest-paying financial institutions often pay interest more than a single time per year. The simple interest formula must be modified to track the compound amount for interest periods of other than one year.

1.2 Compound Interest

The typical interest bearing savings or checking account will be described an investor as earning a specific annual interest rate *compounded monthly*. In this section will be compare and contrast compound interest to the simple interest case of the previous section. Whenever interest is allowed to earn interest itself, an investment is said to earn **compound interest**. In this situation, part of the interest is paid to the depositor more than once per year. Once paid, the interest begins earning interest. We will let the number of compounding periods per year be n . For example for interest “compounded monthly” $n = 12$. Only two small modifications to the simple interest formula (1.1) are needed to calculate the compound interest. First, it is now necessary to think of the interest rate per compounding period. If the annual interest rate is r , then the interest rate per compounding period is r/n . Second, the elapsed time should be thought of as some number of compounding periods rather than years. Thus with n compounding periods per year, the number of compounding periods in t years is nt . Therefore the formula for compound interest is

$$A = P \left(1 + \frac{r}{n}\right)^{nt}. \quad (1.2)$$

Example 1.1 Suppose an account earns 5.75% annually compounded monthly. If the principal amount is \$3104 then after three and one half years the compound amount will be

$$A = 3104 \left(1 + \frac{0.0575}{12}\right)^{(12)(3.5)} = 3794.15.$$

The reader should verify that if the principal in the previous example earned a *simple* interest rate of 5.75% then the compound amount after 3.5 years would be only \$3774.88. Thus happily for the depositor, compound interest builds wealth faster than simple interest. Frequently it is useful to compare an annual interest rate with compounding to an equivalent simple interest, *i.e.* to the simple annual interest rate which would generate the same amount of interest as the annual compound rate. This equivalent interest rate is called the **effective interest rate**. For the amounts and rates mentioned in the previous example we can find the effective interest

rate by solving the equation

$$\begin{aligned} 3104 \left(1 + \frac{0.0575}{12}\right)^{12} &= 3104(1 + r_e) \\ 1.05904 &= 1 + r_e \\ 0.05904 &= r_e \end{aligned}$$

Thus the annual interest rate of 5.75% compounded monthly is equivalent to an effective annual simple rate of 5.904%.

Intuitively it seems that more compounding periods per year implies a higher effective annual interest rate. In the next section we will explore the limiting case of frequent compounding going beyond semiannually, quarterly, monthly, weekly, daily, *etc.* to continuously.

1.3 Continuously Compounded Interest

Mathematically when considering the effect on the compound amount of more frequent compounding, we are contemplating a limiting process. In symbolic form we would like to find the compound amount A which satisfies the equation

$$A = \lim_{n \rightarrow \infty} P \left(1 + \frac{r}{n}\right)^{nt}. \quad (1.3)$$

Fortunately there is a simple expression for the value of the limit on the right-hand side of Eq. (1.3). We will find it by working on the limit

$$\lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^n.$$

This limit is indeterminate of the form 1^∞ . We will evaluate it through a standard approach using the natural logarithm and l'Hôpital's Rule. The reader should consult an elementary calculus book such as [Smith and Minton (2002)] for more details. We see that if $y = (1 + r/n)^n$, then

$$\begin{aligned} \ln y &= \ln \left(1 + \frac{r}{n}\right)^n \\ &= n \ln(1 + r/n) \\ &= \frac{\ln(1 + r/n)}{1/n} \end{aligned}$$

which is indeterminate of the form $0/0$ as $n \rightarrow \infty$. To apply l'Hôpital's Rule we take the limit of the derivative of the numerator over the derivative of

the denominator. Thus

$$\begin{aligned}\lim_{n \rightarrow \infty} \ln y &= \lim_{n \rightarrow \infty} \frac{\frac{d}{dn} (\ln(1 + r/n))}{\frac{d}{dn} (1/n)} \\ &= \lim_{n \rightarrow \infty} \frac{r}{1 + r/n} \\ &= r\end{aligned}$$

Thus $\lim_{n \rightarrow \infty} y = e^r$. Finally we arrive at the formula for **continuously compounded interest**,

$$A = Pe^{rt}. \quad (1.4)$$

This formula may seem familiar since it is often presented as the exponential growth formula in elementary algebra, precalculus, or calculus. The quantity A has the property that A changes with time t at a rate proportional to A itself.

Example 1.2 Suppose \$3585 is deposited in an account which pays interest at an annual rate of 6.15% compounded continuously. After two and one half years the principal plus earned interest will have grown to

$$A = 3585e^{(0.0615)(2.5)} = 4180.82.$$

The effective simple interest rate is the solution to the equation

$$e^{0.0615} = 1 + r_e$$

which implies $r_e = 6.34305\%$.

1.4 Present Value

One of the themes we will see many times in the study of financial mathematics is the comparison of the value of a particular investment at the present time with the value of the investment at some point in the future. This is the comparison between the **present value** of an investment versus its **future value**. We will see in this section that present and future value play central roles in planning for retirement and determining loan payments. Later in this book present and future values will help us determine a fair price for stock market derivatives.

The future value t years from now of an invested amount P subject to an annual interest rate r compounded continuously is

$$A = Pe^{rt}.$$

Thus by comparison with Eq. (1.4), the future value of P is just the compound amount of P monetary units invested in a savings account earning interest r compounded continuously for t years. By contrast the present value of A in an environment of interest rate r compounded continuously for t years is

$$P = Ae^{-rt}.$$

In other words if an investor wishes to have A monetary units in savings t years from now and they can place money in a savings account earning interest at an annual rate r compounded continuously, the investor should deposit P monetary units now. There are also formulas for future and present value when interest is compounded at discrete intervals, not continuously. If the interest rate is r annually with n compounding periods per year then the future value of P is

$$A = P \left(1 + \frac{r}{n}\right)^{nt}.$$

Compare this equation with Eq. (1.2). Simple algebra shows then the present value of P earning interest at rate r compounded n times per year for t years is

$$P = A \left(1 + \frac{r}{n}\right)^{-nt}.$$

By choosing $n = 1$, the case of simple interest can be handled.

Example 1.3 Suppose an investor will receive payments at the end of the next six years in the amounts shown in the table.

Year	1	2	3	4	5	6
Payment	465	233	632	365	334	248

If the interest rate is 3.99% compounded monthly, what is the present value of the investments? Assuming the first payment will arrive one year from

now, the present value is the sum

$$\begin{aligned} & 465 \left(1 + \frac{0.0399}{12}\right)^{-12} + 233 \left(1 + \frac{0.0399}{12}\right)^{-24} + 632 \left(1 + \frac{0.0399}{12}\right)^{-36} \\ & + 365 \left(1 + \frac{0.0399}{12}\right)^{-48} + 334 \left(1 + \frac{0.0399}{12}\right)^{-60} + 248 \left(1 + \frac{0.0399}{12}\right)^{-72} \\ = & 2003.01 \end{aligned}$$

Notice that the present value of the payments from the investment is different from the sum of the payments themselves (which is 2277).

Unless the reader is among the very fortunate few who can always pay cash for all purchases, you may some day apply for a loan from a bank or other financial institution. Loans are always made under the assumptions of a prevailing interest rate (with compounding), an amount to be borrowed, and the lifespan of the loan, *i.e.* the time the borrower has to repay the loan. Usually portions of the loan must be repaid at regular intervals (for example, monthly). Now we turn our attention to the question of using the amount borrowed, the length of the loan, and the interest rate to calculate the loan payment.

A very helpful mathematical tool for answering questions regarding present and future values is the **geometric series**. Suppose we wish to find the sum

$$S = 1 + a + a^2 + \cdots + a^n \quad (1.5)$$

where n is a positive whole number. If both sides of Eq. (1.5) are multiplied by a and then subtracted from Eq. (1.5) we have

$$\begin{aligned} S - aS &= 1 + a + a^2 + \cdots + a^n - (a + a^2 + a^3 + \cdots + a^{n+1}) \\ S(1 - a) &= 1 - a^{n+1} \\ S &= \frac{1 - a^{n+1}}{1 - a} \end{aligned}$$

provided $a \neq 1$.

Now we will apply this tool to the task of finding out the monthly amount of a loan payment. Suppose someone borrows P to purchase a new car. The bank issuing the automobile loan charges interest at the annual rate of r compounded n times per year. The length of the loan will be t years. The monthly payment can be calculated if we apply the principle that the present value of all the payments made must equal the amount borrowed. Suppose the payment amount is the constant x . If the first

payment must be made at the end of the first compounding period, then the present value of all the payments is

$$\begin{aligned} & x(1 + \frac{r}{n})^{-1} + x(1 + \frac{r}{n})^{-2} + \cdots + x(1 + \frac{r}{n})^{-nt} \\ &= x(1 + \frac{r}{n})^{-1} \frac{1 - (1 + \frac{r}{n})^{-nt}}{1 - (1 + \frac{r}{n})^{-1}} \\ &= x \frac{1 - (1 + \frac{r}{n})^{-nt}}{\frac{r}{n}} \end{aligned}$$

Therefore the relationship between the interest rate, the compounding frequency, the length of the loan, the principal amount borrowed, and the payment amount is expressed in the following equation.

$$P = x \frac{n}{r} \left(1 - \left[1 + \frac{r}{n} \right]^{-nt} \right) \quad (1.6)$$

Example 1.4 If a person borrows \$25000 for five years at an interest rate of 4.99% compounded monthly and makes equal monthly payments, the payment amount will be

$$x = 25000(0.0499/12) \left(1 - [1 + (0.0499/12)]^{-(12)(5)} \right)^{-1} = 471.67.$$

Similar reasoning can be used when determining how much to save for retirement. Suppose a person is 25 years of age now and plans to retire at age 65. For the next 40 years they plan to invest a portion of their monthly income in securities which earn interest at the rate of 10% compounded monthly. After retirement the person plans on receiving a monthly payment (an annuity) in the absolute amount of \$1500. The amount of money the person should invest monthly while working can be determined by equating the present value of all their deposits with the present value of all their withdrawals. The first deposit will be made one month from now and the first withdrawal will be made 481 months from now. The monthly deposit amount will be denoted by the symbol x . The present value of all the deposits made into the retirement account is

$$\begin{aligned} \sum_{i=1}^{480} \left(1 + \frac{0.10}{12} \right)^{-i} &= x \left(1 + \frac{0.10}{12} \right)^{-1} \frac{1 - (1 + \frac{0.10}{12})^{-480}}{1 - (1 + \frac{0.10}{12})^{-1}} \\ &\approx 117.765x. \end{aligned}$$

Meanwhile the present value of all the annuity payments is

$$1500 \sum_{i=481}^{840} \left(1 + \frac{0.10}{12}\right)^{-i} = 1500 \left(1 + \frac{0.10}{12}\right)^{-481} \frac{1 - (1 + \frac{0.10}{12})^{-360}}{1 - (1 + \frac{0.10}{12})^{-1}} \\ \approx 3182.94.$$

Thus $x \approx 27.03$ dollars per month. This seems like a small amount to invest, but such is the power of compound interest and starting a savings plan for retirement early. If the person waits ten years to begin saving for retirement, but all other factors remain the same then

$$x \sum_{i=1}^{360} \left(1 + \frac{0.10}{12}\right)^{-i} \approx 113.951x \\ 1500 \sum_{i=361}^{720} \left(1 + \frac{0.10}{12}\right)^{-i} \approx 8616.36$$

which implies the person must invest $x \approx 75.61$ monthly. Waiting ten years to begin saving for retirement nearly triples the amount which the future retiree must set aside for retirement.

The initial amounts invested are of course invested for a longer period of time and thus contribute a proportionately greater amount to the future value of the retirement account.

Example 1.5 Suppose two people will retire in twenty years. One begins saving immediately for retirement but due to unforeseen circumstances must abandon their savings plan after four years. The amount they put aside during those first four years remains invested, but no additional amounts are invested during the last sixteen years of their working life. The other person waits four years before putting any money into a retirement savings account. They save for retirement only during the last sixteen years of their working life. Let us explore the difference in the final amount of retirement savings that each person will possess. For the purpose of this example we will assume that the interest rate is $r = 0.05$ compounded monthly and that both workers will invest the same amount x , monthly. The first worker has upon retirement an account whose present value is

$$x \sum_{i=1}^{48} \left(1 + \frac{0.05}{12}\right)^{-i} \approx 43.423x.$$

The present value of the second worker's total investment is

$$x \sum_{i=49}^{240} \left(1 + \frac{0.05}{12}\right)^{-i} \approx 108.102x.$$

Thus the second worker retires with a larger amount of retirement savings; however, the ratio of their retirement balances is only $43.423/108.102 \approx 0.40$. The first worker saves approximately 40% of what the second worker saves in only one fifth of the time.

The discussion of retirement savings makes no provision for rising prices. The economic concept of **inflation** is the phenomenon of the decrease in the purchasing power of a unit of money relative to a unit amount of goods or services. The rate of inflation (usually expressed as an annual percentage rate, similar to an interest rate) varies with time and is a function of many factors including political, economic, and international factors. While the causes of inflation can be many and complex, inflation is generally described as a condition which results from an increase in the amount of money in circulation without a commensurate increase in the amount of available goods. Thus relative to the supply of goods, the value of the currency is decreased. This can happen when wages are arbitrarily increased without an equal increase in worker productivity.

We now focus on the effect that inflation may have on the worker planning to save for retirement. If the interest rate on savings is r_s and the inflation rate is r_i , the monthly amount put into savings by the worker must be discounted by an effective rate $r_s + r_i$. Then once the worker is retired the monthly annuity must be discounted by the rate $r_s - r_i$. Returning to the earlier example consider the case in which $r_s = 0.10$, the worker will save for 40 years and live on a monthly annuity whose inflation adjusted value will be \$1500 for 30 years, and the rate of inflation will be $r_i = 0.03$ for the entire lifespan of the worker/retiree. Assuming the worker will make the first deposit in one month the present value of all deposits to be made is

$$\begin{aligned} \sum_{i=1}^{480} \left(1 + \frac{0.13}{12}\right)^{-i} &= x \left(1 + \frac{0.13}{12}\right)^{-1} \frac{1 - (1 + \frac{0.13}{12})^{-480}}{1 - (1 + \frac{0.13}{12})^{-1}} \\ &\approx 91.784x. \end{aligned}$$

The present value of all the annuity payments is given by

$$1500 \sum_{i=481}^{840} \left(1 + \frac{0.07}{12}\right)^{-i} = 1500 \left(1 + \frac{0.07}{12}\right)^{-481} \frac{1 - (1 + \frac{0.07}{12})^{-360}}{1 - (1 + \frac{0.07}{12})^{-1}} \\ \approx 13822.30.$$

Thus the monthly deposit amount is \$150.60. This is roughly six times the monthly investment amount when inflation is ignored. However, since inflation does tend to take place over the long run, ignoring a 3% inflation rate over the lifetime of the individual would mean that the present value purchasing power of the last annuity payment would be

$$1500 \left(1 + \frac{0.03}{12}\right)^{840} \approx 184.17.$$

This is not much money to live on for an entire month.

1.5 Rate of Return

The present value of an item is one way to determine the absolute worth of the item and to compare its worth to that of other items. Another way to judge the value of an item which an investor may own or consider purchasing is known as the **rate of return**. If a person invests an amount P now and receives an amount A one time unit from now, the rate of return can be thought of as the interest rate per time unit that the invested amount would have to earn so that the present value of the payoff amount is equal to the invested amount. Since the rate of return is going to be thought of as an equivalent interest rate, it will be denoted by the symbol r . Then by definition

$$P = A(1 + r)^{-1} \quad \text{or equivalently} \quad r = \frac{A}{P} - 1.$$

Example 1.6 If you loan a friend \$100 today with the understanding that they will pay you back \$110 in one year's time, then the rate of return is $r = 0.10$ or 10%.

In a more general setting a person may invest an amount P now and receive a sequence of positive payoffs $\{A_1, A_2, \dots, A_n\}$ at regular intervals. In this case the rate of return per period is the interest rate such that the present value of the sequence of payoffs is equal to the amount invested. In

this case

$$P = \sum_{i=1}^n A_i(1+r)^{-i}.$$

It is not clear from this definition that r has a unique value for all choices of P and payoff sequences. Defining the function $f(r)$ to be

$$f(r) = -P + \sum_{i=1}^n A_i(1+r)^{-i} \quad (1.7)$$

we can see that $f(r)$ is continuous on the open interval $(-1, \infty)$. In the limit as r approaches -1 from the right, the function values approach positive infinity. On the other hand as r approaches positive infinity, the function values approach $-P < 0$ asymptotically. Thus by the Intermediate Value Theorem (p. 108 of [Smith and Minton (2002)]) there exists r^* with $-1 < r^* < \infty$ such that $f(r^*) = 0$. The reader is encouraged to show that r^* is unique in the exercises.

Rates of return can be either positive or negative. If $f(0) > 0$, i.e., the sum of the payoffs is greater than the amount invested then $r^* > 0$ since $f(r)$ changes sign on the interval $[0, \infty)$. If the sum of the payoffs is less than the amount invested then $f(0) < 0$ and the rate of return is negative. In this case the function $f(r)$ changes sign on the interval $(-1, 0]$.

Example 1.7 Suppose you loan a friend \$100 with the agreement that they will pay you at the end of the next five years amounts $\{21, 22, 23, 24, 25\}$. The rate of return per year is the solution to the equation,

$$-100 + \frac{21}{1+r} + \frac{22}{(1+r)^2} + \frac{23}{(1+r)^3} + \frac{24}{(1+r)^4} + \frac{25}{(1+r)^5} = 0.$$

Newton's Method (Sec. 3.2 of [Smith and Minton (2002)]) can be used to approximate the solution $r^* \approx 0.047$.

1.6 Exercises

- (1) Suppose that \$3659 is deposited in a savings account which earns 6.5% simple interest. What is the compound amount after five years?
- (2) Suppose that 3993 is deposited in an account which earns 4.3% interest. What is the compound amount after two years if the interest is compounded

- (a) monthly?
 - (b) weekly?
 - (c) daily?
 - (d) continuously?
- (3) Find the effective annual simple interest rate which is equivalent to 8% interest compounded quarterly.
- (4) You are preparing to open a bank which will accept deposits into savings accounts and which will pay interest compounded monthly. In order to be competitive you must meet or exceed the interest paid by another bank which pays 5.25% compounded daily. What is the minimum interest rate you can pay and remain competitive?
- (5) Suppose you have \$1000 to deposit in one of two types of savings accounts. One account pays interest at an annual rate of 4.75% compounded daily, while the other pays interest at an annual rate of 4.75% compounded continuously. How long would it take for the compound amounts to differ by \$1?
- (6) Many textbooks determine the formula for continuously compounded interest through an argument which avoids the use of l'Hôpital's Rule (for example [Goldstein *et al.* (1999)]). Beginning with Eq. (1.3) let $h = r/n$. Then

$$P \left(1 + \frac{r}{n}\right)^{nt} = P(1+h)^{(1/h)rt}$$

and we can focus on finding the $\lim_{h \rightarrow 0} (1+h)^{1/h}$. Show that

$$(1+h)^{1/h} = e^{(1/h)\ln(1+h)}$$

and take the limit of both sides as $h \rightarrow 0$. Hint: you can use the definition of the derivative in the exponent on the right-hand side.

- (7) Which of the two investments described below is preferable? Assume the first payment will take place exactly one year from now and further payments are spaced one year apart. Assume the continually compounded annual interest rate is 2.75%.

Year	1	2	3	4
Investment A	200	211	198	205
Investment B	198	205	211	200

- (8) Suppose you wish to buy a house costing \$200000. You will put a down payment of 20% of the purchase price and borrow the rest from a bank for 30 years at a fixed interest rate r compounded monthly. If you

wish your monthly mortgage payment to be \$1500 or less, what is the maximum annual interest rate for the mortgage loan?

- (9) Use the Mean Value Theorem (p. 235 of [Stewart (1999)]) to show the rate of return defined by the root of the function in Eq. (1.7) is unique.
- (10) Suppose for an investment of \$10000 you will receive payments at the end of each of the next four years in the amounts {2000, 3000, 4000, 3000}. What is the rate of return per year?
- (11) Suppose you have the choice of investing \$1000 in just one of two ways. Each investment will pay you an amount listed in the table below at the end of the next five years.

Year	1	2	3	4	5
Investment A	225	215	250	225	205
Investment B	220	225	250	250	210

- (a) Using the present value of the investment to make the decision, which investment would you choose? Assume the simple annual interest rate is 4.33%.
- (b) Using the rate of return per year of the investment to make the decision, which investment would you choose?

Chapter 2

Discrete Probability

Since the number and interactions of forces driving the values of investments are so large and complex, development of a deterministic mathematical model of a market is likely to be impossible. In this book a probabilistic or stochastic model of a market will be developed instead. This chapter presents some elementary concepts of probability and statistics. Here the reader will find explanations of discrete events and their outcomes. A discrete outcome can take on only one of a finite number of values. For example the outcome of a roll of a fair die can be only one of the six values in the set $\{1, 2, 3, 4, 5, 6\}$. No one ever rolls a die and discovers the outcome to be π for example. Basic methods for determining the probabilities of outcomes will be presented. The concept of the **random variable**, a numerical quantity whose value is not known until an experiment is conducted, will be explained. There are many different kinds of discrete random variables, but one that frequently arises in financial mathematics is the **binomial random variable**. While statistics is a field of study unto itself, two important descriptive statistics will be introduced in this chapter, **expected value** (or mean) and **standard deviation**. The expected value provides a number which is representative of typical values of a random variable, while the standard deviation provides a measure of the “spread” in the values the random variable can take on.

2.1 Events and Probabilities

To the layman an event is something that happens. To the statistician, an **event** is an outcome or set of outcomes of an experiment. This brings up the question of what is an experiment? For our purposes an **experiment** will be any activity that generates an observable outcome. Some simple

examples of experiments include flipping a coin, rolling a pair of dice, and drawing cards from a deck. An outcome of each of these experiments could be “heads”, 7, or the ace of hearts respectively. For the example of the coin flip or drawing a card from the deck, the example outcomes given can be thought of as “atomic” in the sense that they cannot be further broken down into simpler events. The outcome of achieving a 7 on a roll of a pair of dice could be thought of as consisting of a pair of outcomes, one for each die. For example the 7 could be the result of 2 on the first die and 5 on the second. Having a flipped coin land heads up cannot be similarly decomposed. An event can also be thought of as a collection of outcomes rather than just a single outcome. For example the experiment of drawing a card from a standard deck, the events could be segregated into hearts, diamonds, spades, or clubs depending on the suit of the card drawn. Then any of the atomic events 2, 3, . . . , 10, jack, queen, king, or ace of hearts would be a “heart” event for the experiment of drawing a card and observing its suit.

In this chapter the outcomes of experiments will be thought of as **discrete** in the sense that the outcomes will be from a set whose members are isolated from each other by gaps. The discreteness of a coin flip, a roll of a pair of dice, and card draw are apparent due to the condition that there is no outcome between “heads” and “tails”, or between 6 and 7, or between the two of clubs and the three of clubs respectively. Also in this chapter the number of different outcomes of an experiment will be either finite or countable (meaning that the outcomes can be put into one-to-one correspondence with a subset of the natural numbers).

The **probability** of an event is a real number measuring the likelihood of that event occurring as the outcome of an experiment. To begin the more formal study of events and probabilities, let the symbol A represent an event. The probability of event A will be denoted $P(A)$. By convention, probabilities are always real numbers in the interval $[0, 1]$, that is, $0 \leq P(A) \leq 1$. If A is an event for which $P(A) = 0$, then A is said to be an **impossible** event. If $P(A) = 1$, then A is said to be a **certain** event. Impossible events never occur, while certain events always occur. Events with probabilities closer to 1 are more likely to occur than events whose probabilities are closer to 0.

There are two approaches to assigning a probability to an event, the **classical** approach and the **empirical** approach. Adopting the empirical approach requires an investigator to conduct (or at least simulate) the experiment N times (where N is usually taken to be as large as practical).

During the N repetitions of the experiment the investigator counts the number of times that event A occurred. Suppose this number is x . Then the probability of event A is estimated to be $P(A) = x/N$. The classical approach is a more theoretical exercise. The investigator must consider the experiment carefully and determine the total number of different outcomes of the experiment (call this number M), assume that each outcome is equally likely, and then determine the number of outcomes among the total in which event A occurs (suppose this number is y). The probability of event A is then assigned the value $P(A) = y/M$. In practice the two methods closely agree, especially when N is very large.

Some experiments involve events which can be thought of as the result of two or more outcomes occurring simultaneously. For example, suppose a red coin and a green coin will be flipped. One compound outcome of the experiment is the red coin lands on “heads” and the green coin lands on “heads” also. The next section contains some simple rules for handling the probabilities of these compound events.

2.2 Addition Rule

Suppose A and B are two events which may occur as a result of conducting an experiment. An investigator may wish to know the probability that A or B occurs. Symbolically this would be represented as $P(A \vee B)$. If an investigator rolls a pair of fair dice they may want to know the probability that a total of 2 or 12 results. Let event A be the outcome of 2 and B be the outcome of 12. Since $P(A) = 1/36$ and $P(B) = 1/36$ and the two events are **mutually exclusive**, that is, they cannot both simultaneously occur, the $P(A \vee B) = P(A) + P(B) = 1/18$. Suppose instead that the investigator wants to know the probability that a total of less than 6 or an odd total results. We can let event A be the outcome of a total less than 6 (that is, a total of 2, 3, 4, or 5) and let event B be the outcome of an odd total (specifically 3, 5, 7, 9, or 11). We see this time that events A and B are not mutually exclusive, there are outcomes which overlap both events, namely the odd numbers 3 and 5 are less than 6. To adjust the calculation the probabilities of the non-exclusive events should be counted only once. Thus the $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$ where $P(A \wedge B)$ is the probability that one of the events in the overlapping non-exclusive set of

outcomes occurs. Hence

$$\begin{aligned} P(A \vee B) &= \left(\frac{1}{36} + \frac{1}{18} + \frac{1}{12} + \frac{1}{9} \right) + \left(\frac{1}{18} + \frac{1}{9} + \frac{1}{6} + \frac{1}{9} + \frac{1}{18} \right) - \left(\frac{1}{18} + \frac{1}{9} \right) \\ &= \frac{11}{18}. \end{aligned}$$

Thus the calculation of the probability of event A or event B occurring is different depending on whether A and B are mutually exclusive.

The concept outlined above is known as the **Addition Rule** for Probabilities and can be stated in the form of a theorem.

Theorem 2.1 (Addition Rule) *For events A and B , the probability of A or B occurring is*

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B). \quad (2.1)$$

If A and B are mutually exclusive events then $P(A \wedge B) = 0$ and the Addition Rule simplifies to

$$P(A \vee B) = P(A) + P(B).$$

Determining the probability of the occurrence of A or B rests on determining the probability that *both* A and B occur. This topic is explored in the next section.

2.3 Conditional Probability and Multiplication Rule

During the past decade a very famous puzzle involving probability has come to be known as the “Monty Hall Problem”. This paradox of probability was published in a different but equivalent form in Martin Gardner’s “Mathematical Games” feature of *Scientific American* [Gardner (1959)] in 1959 and in the *American Statistician* [Selvin (1975)] in 1975. In 1990 it appeared in its present form in the “Ask Marilyn” column of *Parade Magazine* [vos Savant (1990)].

A game show host hides a prize behind one of three doors. A contestant must guess which door hides the prize. First, the contestant announces the door they have chosen. The host will then open one of the two doors not chosen in order to reveal the prize is not behind it. The host then tells the contestant they may keep their original choice or switch to the other unopened door. Should the contestant switch doors?

At first glance when faced with two identical unopened doors, it may seem that there is no advantage to switching doors; however, if the contestant switches they will win with probability $2/3$. When the contestant makes the first choice they have a $1/3$ chance of being correct and a $2/3$ chance of being incorrect. When the host reveals the non-winning, unchosen door, the contestants first choice still has a $1/3$ chance of being correct, but now the unchosen unopened door has a $2/3$ probability of being correct, so the contestant should switch.

This example illustrates the concept known as **conditional probability**. Essentially the decision the contestant faces is “given that I have seen that one of the doors I did not choose is not the winning door, should I alter my choice?” The probability that one event occurs given that another event has occurred is called conditional probability. The probability that event A occurs given that event B occurs is denoted $P(A|B)$. One of the classical thought experiments of discrete probability involves selecting balls from an urn. Suppose an urn contains 20 balls, 6 of which are blue and the remaining 14 are green. Two balls will be drawn, the second will be drawn without replacing the first. The question “what is the probability that the second ball is green, given that the first ball was green?” could be asked. The answer to this question will motivate the statement of the multiplication rule of probability. One approach to the answer involves determining the probability that when two balls are drawn without replacement they are both green. The probability that both selections are green would be the number of two green ball outcomes divided by the total number of outcomes. There are 20 candidates for the first ball selected and there are 19 candidates for the second ball selected. Thus the total number of outcomes is 380. Of those outcomes $(14)(13) = 182$ are both green balls. Thus the probability that both balls are green is $182/380 = 91/190$. The reader may be asking what this situation has to do with the question originally posed. The outcome in which both balls are green is a subset of all the outcomes in which the first ball is green. Consider the diagram in Fig. 2.1. Thus the probability that both balls are green is the product of the probability that the first ball is green multiplied the probability the second ball is green. Let event A be the set of outcomes in which the first ball is green and event B be the set of outcomes in which the second ball is green. Numerically $P(A) = 7/10$ and symbolically

$$P(A \wedge B) = P(A) P(B|A)$$

Thus $P(B|A) = P(A \wedge B)/P(A) = (91/190)/(7/10) = 13/19$.

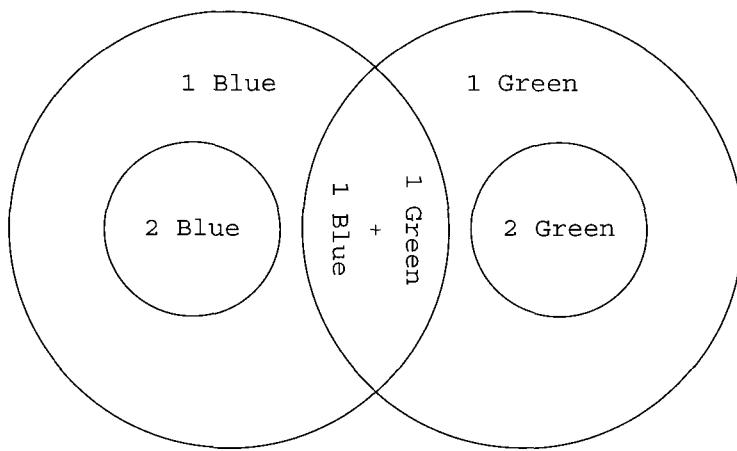


Fig. 2.1 The sets of outcomes of drawing one or two balls from an urn containing blue and green balls.

The concept illustrated above is known as the **Multiplication Rule** for Probabilities and can be stated in the form of a theorem.

Theorem 2.2 (Multiplication Rule) *For events A and B, the probability of A and B occurring is*

$$P(A \wedge B) = P(A)P(B|A). \quad (2.2)$$

Equation 2.2 can be used to find $P(B|A)$ directly

$$P(B|A) = \frac{P(A \wedge B)}{P(A)}.$$

This expression is meaningful only when $P(A) > 0$.

Example 2.1 One type of roulette wheel, known as the American type, has 38 potential outcomes represented by the integers 1 through 36 and two special outcomes 0 and 00. The positive integers are placed on alternating red and black backgrounds while 0 and 00 are on green backgrounds. What is the probability that the outcome is less than 10 and more than 3 given that the outcome is an even number?

Let event A be the set of outcomes in which the number is even. $P(A) = 10/19$ if 0 and 00 are treated as even numbers. Let B be the set of outcomes in which the number is greater than 3 and less than 10. Then $P(A \wedge B) =$

3/38 and

$$P(B|A) = \frac{3/38}{10/19} = 3/20.$$

To expand on the previous example, suppose the roulette wheel will be spun twice. One could ask what is the probability that both spins have a red outcome. If event A is the outcome of red on the first spin and event B is the outcome of red on the second spin, then we as before $P(A \wedge B) = P(A)P(B|A)$. However there is no reason to believe that the wheel somehow “remembers” the outcome of the first spin while it is being spun the second time. The first outcome has no effect on the second outcome. In any experiment, if event A has no effect on event B then A and B are said to be **independent**. In this situation $P(B|A) = P(B)$. Thus for independent events the Multiplication Rule can be modified to

$$P(A \wedge B) = P(A)P(B).$$

Therefore the probability that both spins will have red outcomes is $P(A \wedge B) = (9/19)(9/19) = 81/361$.

2.4 Random Variables and Probability Distributions

The outcome of an experiment is not known until after the experiment is performed. For example the number of people who vote in an election is not known until the election is concluded. In a more formal sense we can describe a **random variable** as a function which maps the set of outcomes of an experiment to some subset of the real numbers. In the election example (assuming the number of registered voters is N and that there will be no fraudulent voting) the sample space of outcomes of voter turnout is the set $S = \{0, 1, 2, \dots, N\}$. Symbolically a random variable for the voter turnout is the function $X : S \rightarrow \mathbb{R}$. Often X is thought of as the eventual numeric result of the experiment.

A **probability distribution** is a function which assigns a probability to each element in the sample space of outcomes of an experiment. If S is the set outcomes of an experiment and f is the associated probability distribution function, then f maps each element in S to a unique real number in the interval $[0, 1]$. If x is a potential outcome of an experiment with sample space S then $f(x) = P(X = x)$, in other words $f(x)$ is the probability that x occurs as the outcome of the experiment. Since a probability

Table 2.1 The genders of four children born by the same set of parents.

Child			
1	2	3	4
B	B	B	B
G	B	B	B
B	G	B	B
B	B	G	B
B	B	B	G
G	G	B	B
G	B	G	B
G	B	B	G
B	G	G	B
B	G	B	G
B	B	G	G
B	G	G	G
G	B	G	G
G	G	B	G
G	G	G	B
G	G	G	G

distribution function maps an outcome to a probability then the following two characteristics are true of the function.

- (1) If x_i is one of the N outcomes of an experiment then $0 \leq f(x_i) \leq 1$.
- (2) The sum of the values of the probability distribution function is unity, *i.e.*

$$1 = \sum_{i=1}^N f(x_i).$$

Example 2.2 Consider a family with four children. The random variable X will represent the number of children who are male. The sample space for this experiment (having four children and counting the number of boys) is the set $S = \{0, 1, 2, 3, 4\}$. The gender of each child is independent from the genders of their siblings, so assuming that the probability of a child being male is $1/2$ then the 16 events shown in table 2.1 are equally likely. There is one outcome in which there are no male children, thus $f(0) = P(X = 0) = 1/16$. There are four cases in which there is a single male child and hence $f(1) = 1/4$. The reader can readily determine from the table that $f(2) = 3/8$, $f(3) = 1/4$, and $f(4) = 1/16$.

Several common types of random variables and their associated probability distributions will be important to the study of financial mathematics. The binomial random variable will be discussed in the next section. It will be seen to be related to the **Bernoulli random variable** which takes on only one of two possible values, often thought of as true or false (or sometimes as success and failure). It is mathematically convenient to designate the outcomes as 0 and 1. The probability distribution function of a Bernoulli random is particularly simple, $f(1) = P(X = 1) = p$ where $0 \leq p \leq 1$ and $f(0) = 1 - p$.

2.5 Binomial Random Variables

Returning to the last example of the previous section, we can think of the births of four children as four independent events. The gender of one child in no way influences the genders of children born before or after. Since the gender of a child can take on only one of two values it is possible to think of the birth of each child as a Bernoulli event. The probability of having a male or female child does not change between births. Thus the experiment of producing four children in a family is the same as repeating the experiment of having a single child four times or, repeating a Bernoulli experiment four times. This is the idea of the **binomial random variable** defined next. A binomial random variable X is the number of successful outcomes out of n independent Bernoulli random variables.

A binomial random variable is parameterized by the number of repetitions of the Bernoulli experiment (referred to as trials from here on) and by the probability of success on a single trial. If the number of trials is n and the probability of success on a single trial is p , then the set of possible outcomes of the binomial experiment is the set $\{0, 1, \dots, n\}$. The number of combinations of x successes out of n trials is

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

The probability of x successes out of n independent trials in a specified combination is, according to the Multiplication Rule, $p^x(1-p)^{n-x}$. Since the various combinations are mutually exclusive, by the Addition Rule the probability of x successes out of n trials is given by the function

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad (2.3)$$

Thus if the probability of an individual child being born male or female is $1/2$, the probability that a family with four children will have two female children is

$$P(X = 2) = \frac{4!}{2!(4-2)!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{4-2} = \frac{3}{8}.$$

This result agrees with the result of the more cumbersome method used to determine this probability in the previous section.

Example 2.3 The probability that a computer memory chip is defective is 0.02 . A SIMM (single in-line memory module) contains 16 chips for data storage and a 17th chip for error correction. The SIMM can operate correctly if one chip is defective, but not if two or more are defective. The probability that a SIMM will not function is

$$P(X \geq 2) = \sum_{x=2}^{17} \binom{17}{x} (0.02)^x (0.98)^{17-x} \approx 0.044578.$$

2.6 Expected Value

When faced with experimental data, summary statistics are often useful for making sense of the data. In this context “statistics” refers to numbers which can be calculated from the data rather than the means and algorithms by which these numbers are calculated. In financial mathematics the statistical needs are somewhat more specialized than in a general purpose course in statistics. Here we wish to answer the hypothetical question, “if an experiment was to be performed an infinite number of times, what would be the typical outcome?” Thus we will introduce only the statistical concepts to be used later in this text. A reader interested in a broader, deeper, and more rigorous background in statistics should consult one of the many textbooks devoted to the subject for example [Ross (2006)].

To take an example, if a fair die was rolled an infinite number of times, what would be the typical result? The notion to be explored in this section is that of **expected value**. In some ways expected value is synonymous with the mean or average of a list of numerical values; however, it can differ in at least two important ways. First, the expected value usually refers to the typical value of a random variable whose outcomes are not necessarily equally likely whereas the mean of a list of data treats each observation as equally likely. Second, the expected value of a random variable is the

typical outcome of an experiment performed an infinite number of times whereas the statistical mean is calculated based on a finite collection of observations of the outcome of an experiment.

If X is a discrete random variable with probability distribution $P(X)$ then the expected value of X is denoted $E[X]$ and defined as

$$E[X] = \sum_X (X \cdot P(X)). \quad (2.4)$$

It is understood that the summation is taken over all values that X may assume. In the case that X takes on only a finite number of values with nonzero probability, then this sum is well-defined. If X may assume an infinite number of values with probabilities greater than zero, we will assume that the sum converges. Since each value of X is multiplied by its corresponding probability, the expected value of X is a weighted average of the variable X . Returning to the question posed in the previous paragraph as to the typical outcome achieved when rolling a fair die an infinite number of times, we may determine this number from the formula for expected value. Since $X \in \{1, 2, 3, 4, 5, 6\}$ and $P(X) = 1/6$ for all possible values of X , then the expected value of X is

$$E[X] = \sum_{X=1}^6 \frac{X}{6} = \frac{1}{6} \sum_{X=1}^6 X = \frac{1}{6} \cdot \frac{(6)(7)}{2} = \frac{7}{2}.$$

Thus the average outcome of rolling a fair die is 3.5.

Example 2.4 Let random variable X represent the number of female children in a family of four children. Assuming that births of males and females are equally likely and that all births are independent events, what is the $E[X]$? The sample space of X is the set $\{0, 1, 2, 3, 4\}$. Using the binomial probability formula $P(0) = 1/16$, $P(1) = 1/4$, $P(2) = 3/8$, $P(3) = 1/4$, and $P(4) = 1/16$, we have

$$E[X] = 0 \cdot \frac{1}{16} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16} = 2.$$

In families having four children, typically there are two female children and consequently two male children.

The notion of the expected value of a random variable X can be extended to the expected value of a function of X . Thus we say that if F is

a function applied to X , then

$$\mathbb{E}[F(X)] = \sum_X F(X)P(X).$$

When the function F is merely multiplication by a constant then the expected value takes on a simple form.

Theorem 2.3 *If X is a random variable and a is a constant, then $\mathbb{E}[aX] = a\mathbb{E}[X]$.*

Proof. By the definition of expected value

$$\mathbb{E}[aX] = \sum_X ((aX) \cdot P(X)) = a \sum_X (X \cdot P(X)) = a\mathbb{E}[X].$$

□

Later in this work sums of random variables will become important. Thus some attention must be given to the expected value of the sum of random variables. However, this requires the probability of two or more random variables be considered simultaneously. The reader should already be familiar with one example of this situation, namely the rolling of a pair of dice. Suppose that the two dice can be distinguished from one another (imagine that one of them is red while the other is green). Let X be the random variable denoting the outcome of the green die while Y is a random variable denoting the outcome of the red die. If the experiment to be performed is rolling the pair of dice and considering the total of the upward faces then the random variable denoting the outcome of this experiment is $X + Y$. This naturally leads us to the issue of describing the probabilities associated with various values of the random variable $X + Y$. The **joint probability function** is denoted $P(X, Y)$ and we will understand it to mean $P(X, Y) = P(X \wedge Y)$. Thus $P(1, 3)$ symbolizes the probability that the outcome of the red die is 1 while the outcome of the green die is 3. If the individual dice are independent then $P(1, 3) = P(1)P(3) = 1/36$ according to the multiplication rule. A couple of additional comments are in order. First, joint probabilities exist even for random events which are not independent. Second, realize that in general $P(X + Y) \neq P(X) + P(Y)$. This is an abuse of notation, but is not likely to cause confusion in what follows. $P(X + Y)$ refers to the probability of the sum $X + Y$ which depends on the joint probabilities of X and Y . $P(X)$ and $P(Y)$ refer respectively to the individual probabilities of random variable X and Y . The following is true, if we wish to know the probability that the sum of the discrete random

variables X and Y is m then by using the addition rule for probabilities

$$P(X + Y = m) = \sum_{X+Y=m} P(X, Y).$$

The summation is taken over all combinations of X and Y such that $X + Y = m$. Returning to the dice example introduced earlier in the paragraph we see that the probability that the sum of the dice is 4 is

$$\begin{aligned} P(X + Y = 4) &= \sum_{X+Y=4} P(X, Y) \\ &= P(1, 3) + P(2, 2) + P(3, 1) \\ &= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{12}. \end{aligned}$$

The joint probability distribution of a pair of random variables possesses many of the same properties that the probability distribution of a single random variable possesses. For example $0 \leq P(X, Y) \leq 1$ for all X and Y in the discrete sample space. It is also true that

$$\sum_X \sum_Y P(X, Y) = \sum_Y \sum_X P(X, Y) = 1.$$

An important property will be used in the proof of the next theorem. The sum of the joint probability of X and Y where Y is allowed to take on each of its possible values is called the **marginal probability of X** . Without confusion we will denote the marginal probability of X as $P(X)$ and realize that

$$P(X) = \sum_Y P(X, Y).$$

Similarly the marginal probability of Y is denoted $P(Y)$ and defined as

$$P(Y) = \sum_X P(X, Y).$$

Conveniently the expected value of a sum of random variables is the sum of the expected values of the random variables. This notion is made more precise in the following theorem.

Theorem 2.4 *If X_1, X_2, \dots, X_k are random variables then*

$$E[X_1 + X_2 + \dots + X_k] = E[X_1] + E[X_2] + \dots + E[X_k].$$

Proof. If $k = 1$ then the proposition is certainly true. If $k = 2$ then

$$\begin{aligned} E[X_1 + X_2] &= \sum_{X_1, X_2} ((X_1 + X_2)P(X_1, X_2)) \\ &= \sum_{X_1} \sum_{X_2} ((X_1 + X_2)P(X_1, X_2)) \\ &= \sum_{X_1} \sum_{X_2} X_1 P(X_1, X_2) + \sum_{X_2} \sum_{X_1} X_2 P(X_1, X_2) \\ &= \sum_{X_1} X_1 \sum_{X_2} P(X_1, X_2) + \sum_{X_2} X_2 \sum_{X_1} P(X_1, X_2) \\ &= \sum_{X_1} X_1 P(X_1) + \sum_{X_2} X_2 P(X_2) \\ &= E[X_1] + E[X_2]. \end{aligned}$$

For a finite value of $k > 2$ the result is true by induction. Suppose the result is true for $n < k$ where $k > 2$, then

$$\begin{aligned} E[X_1 + \cdots + X_{k-1} + X_k] &= E[X_1 + \cdots + X_{k-1}] + E[X_k] \\ &= E[X_1] + \cdots + E[X_{k-1}] + E[X_k] \end{aligned}$$

The last step is true by the induction hypothesis. \square

We can use Theorem 2.4 to determine the expected value of a binomial random variable. Along the way we will also find the expected value of a Bernoulli random variable. Suppose n trials of a Bernoulli experiment will be conducted for which the probability of success on a single trial is $0 \leq p \leq 1$. Random variable X represents the number of successes out of n trials. By assumption the trials are independent of one another and the outcomes are mutually exclusive. The result of the binomial experiment can be thought of as the sum of the results of n Bernoulli experiments. Let the random variable X_i be the number of successes of the i^{th} Bernoulli trial, then

$$E[X] = E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] = p + \cdots + p = np.$$

Later we will have need to calculate the expected value of a product of random variables. This situation is not as straightforward as the case of a sum of random variables.

Theorem 2.5 *Let X_1, X_2, \dots, X_k be pairwise independent random vari-*

ables, then

$$\mathbb{E}[X_1 X_2 \cdots X_k] = \mathbb{E}[X_1] \mathbb{E}[X_2] \cdots \mathbb{E}[X_k].$$

Proof. Naturally we see that when $k = 1$ the theorem is true. Next we will consider the case when $k = 2$. Let X_1 and X_2 be independent random variables with joint probability distribution $P(X_1, X_2)$. Since the random variables are assumed to be independent then $P(X_1, X_2) = P(X_1)P(X_2)$. Once again we are lax in our use of notation, since in the previous equation the symbol P is used in three senses ((1) the joint probability distribution of X_1 and X_2 , (2) the probability distribution of X_1 , and (3) the probability distribution of X_2); however, there is little chance of confusion in this elementary proof.

$$\begin{aligned}\mathbb{E}[X_1 X_2] &= \sum_{X_1, X_2} X_1 X_2 P(X_1, X_2) \\ &= \sum_{X_1} \sum_{X_2} X_1 X_2 P(X_1) P(X_2) \\ &= \sum_{X_1} X_1 P(X_1) \sum_{X_2} X_2 P(X_2) \\ &= \mathbb{E}[X_1] \mathbb{E}[X_2]\end{aligned}$$

For a finite value of $k > 2$ the result is true by induction. Suppose the result is true for $n < k$ where $k > 2$, then

$$\begin{aligned}\mathbb{E}[X_1 \cdots X_{k-1} X_k] &= \mathbb{E}[X_1 \cdots X_{k-1}] \mathbb{E}[X_k] \\ &= \mathbb{E}[X_1] \cdots \mathbb{E}[X_{k-1}] \mathbb{E}[X_k]\end{aligned}$$

The last step is true by the induction hypothesis. \square

The expected value of a random variable specifies the average outcome of an infinite number of repetitions of an experiment. In the next section the notions of variance and standard deviation are introduced. They specify measures of the spread of the outcomes from the expected value.

2.7 Variance and Standard Deviation

The **variance** of a random variable is a measure of the spread of values of the random variable about the expected value of the random variable. The

variance is defined as

$$\text{Var}(X) = E[(X - E[X])^2]. \quad (2.5)$$

As the reader can see from Eq. (2.5), the variance is always non-negative. The expression $X - E[X]$ is the signed deviation of X from its expected value. The variance may be interpreted as the average of the squared deviation of a random variable from its expected value. An alternative formula for the variance is sometimes more convenient in calculations.

Theorem 2.6 *Let X be a random variable, then the variance of X is $E[X^2] - E[X]^2$.*

Proof. By definition,

$$\begin{aligned} \text{Var}(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2] \\ &= E[X^2] - E[2XE[X]] + E[E[X]^2] \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2. \end{aligned}$$

The third and fourth steps of this derivation made use of theorems 2.4 and 2.3 respectively. \square

Returning to the previous example of the hypothetical family with four children, we can now investigate the variance in the number of female children. We already know that $E[X] = 2$. If we make use of the result of Theorem 2.6 then

$$\begin{aligned} \text{Var}(X) &= E[X^2] - E[X]^2 \\ &= (0^2)(1/16) + (1^2)(1/4) + (2^2)(3/8) + (3^2)(1/4) + (4^2)(1/16) - 2^2 \\ &= 1. \end{aligned}$$

Before investigating the variance of a binomial random variable, we should determine the variance of a Bernoulli random variable. If the prob-

ability of success is $0 \leq p \leq 1$ then according to Eq. (2.5),

$$\begin{aligned}\text{Var}(X) &= E[(X - E[X])^2] \\ &= E[(X - p)^2] \\ &= (1-p)^2 p + (0-p)^2(1-p) \\ &= (1-p)^2 p + p^2(1-p) \\ &= p(1-p)(1-p+p) \\ &= p(1-p).\end{aligned}$$

The following theorem provides an easy formula for calculating the variance of *independent* random variables.

Theorem 2.7 *Let X_1, X_2, \dots, X_k be pairwise independent random variables, then*

$$\text{Var}(X_1 + X_2 + \dots + X_k) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_k).$$

Proof. If $k = 1$ then the result is trivially true. Take the case when $k = 2$. By the definition of variance,

$$\begin{aligned}\text{Var}(X_1 + X_2) &= E[((X_1 + X_2) - E[X_1 + X_2])^2] \\ &= E[((X_1 + X_2) - E[X_1] - E[X_2])^2] \\ &= E[((X_1 - E[X_1]) + (X_2 - E[X_2]))^2] \\ &= E[(X_1 - E[X_1])^2 + (X_2 - E[X_2])^2 + \\ &\quad 2(X_1 - E[X_1])(X_2 - E[X_2])] \\ &= E[(X_1 - E[X_1])^2] + E[(X_2 - E[X_2])^2] + \\ &\quad 2E[(X_1 - E[X_1])(X_2 - E[X_2])] \\ &= \text{Var}(X_1) + \text{Var}(X_2) + 2E[(X_1 - E[X_1])(X_2 - E[X_2])]\end{aligned}$$

Since we are assuming that random variables X_1 and X_2 are independent, then by Theorem 2.5

$$\begin{aligned}E[(X_1 - E[X_1])(X_2 - E[X_2])] &= E[X_1 - E[X_1]] E[X_2 - E[X_2]] \\ &= (E[X_1] - E[X_1])(E[X_2] - E[X_2]) \\ &= 0,\end{aligned}$$

and thus

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2).$$

The result can be extended to any finite value of k by induction. Suppose the result has been shown true for $n < k$ with $k > 2$. Then

$$\begin{aligned}\text{Var}(X_1 + \cdots + X_{k-1} + X_k) &= \text{Var}(X_1 + \cdots + X_{k-1}) + \text{Var}(X_k) \\ &= \text{Var}(X_1) + \cdots + \text{Var}(X_{k-1}) + \text{Var}(X_k)\end{aligned}$$

where the last equality is justified by the induction hypothesis. \square

Readers should think carefully about the validity of the claim that $X_1 - E[X_1]$ and $X_2 - E[X_2]$ are independent in light of the assumption that X_1 and X_2 are independent.

Example 2.5 Suppose a binomial experiment is characterized by n independent repetitions of a Bernoulli trial for which the probability of success on a single trial is $0 \leq p \leq 1$. Random variable X denotes the total number of successes accrued over the n trials.

$$\begin{aligned}\text{Var}(X) &= \text{Var}\left(\sum_{j=1}^n X_j\right) \\ &= \sum_{j=1}^n \text{Var}(X_j) \quad (\text{since trials are independent}) \\ &= \sum_{j=1}^n p(1-p) \\ &= np(1-p)\end{aligned}$$

So far we have made no mention of the other topic in the heading for this section, namely **standard deviation**. There is little more that must be said since by definition the standard deviation is the square root of the variance. Standard deviation of a random variable X is denoted by $\sigma(X)$ and thus

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

2.8 Exercises

- (1) Suppose the four sides of a regular tetrahedron are labeled 1 through 4. If the tetrahedron is rolled like a die, what is the probability of it landing on 3?

- (2) Use the classical approach and the assumption of fair dice to find the probabilities of the outcomes obtained by rolling a pair of dice and summing the dots shown on the upward faces.
- (3) If the probability that a batter strikes out in the first inning of a baseball game is $1/3$ and the probability that the batter strikes out in the fifth inning is $1/4$, and the probability that the batter strikes out in both innings is $1/10$, then what is the probability that the batter strikes out in either inning?
- (4) Part of a well-known puzzle involves three people entering a room. As each person enters, at random either a red or a blue hat is placed on the person's head. The probability that an individual receives a red hat is $1/2$. No person can see the color of their own hat, but they can see the color of the other two persons' hats. The three will split a prize if at least one person guesses the color of their own hat correctly and no one guesses incorrectly. A person may decide to pass rather than to guess. The three people are not allowed to confer with one another once the hats have been placed on their heads, but they are allowed to agree on a strategy prior to entering the room. At the risk of spoiling the puzzle, one strategy the players may follow instructs a player to pass if they see the other two persons wearing mis-matched hats and to guess the opposite color if their friends are wearing matching hats. Why is this a good strategy and what is the probability of winning the game?
- (5) Suppose cards will be drawn without replacement from a standard 52-card deck. What is the probability that the first two cards will be aces?
- (6) Suppose cards will be drawn without replacement from a standard 52-card deck. What is the probability that the second card drawn will be an ace and that the first card was not an ace?
- (7) Suppose cards will be drawn without replacement from a standard 52-card deck. What is the probability that the fourth card drawn will be an ace given that the first three cards drawn were all aces?
- (8) Suppose cards will be drawn without replacement from a standard 52-card deck. On which draw is the card mostly likely to be the first ace drawn?
- (9) On the last 100 spins of an American style roulette wheel, the outcome has been black. What is the probability of the outcome being black on the 101st spin?

- (10) On the last 5000 spins of an American style roulette wheel, the outcome has been 00. What is the probability of the outcome being 00 on the 5001st spin?
- (11) Suppose that a random variable X has a probability distribution function f so that $f(x) = c/x$ for $x = 1, 2, \dots, 10$ and is zero otherwise. Find the appropriate value of the constant c .
- (12) Suppose that a box contains 15 black balls and 5 white balls. Three balls will be selected without replacement from the box. Determine the probability distribution function for the number of black balls selected.
- (13) Quality control for a manufacturer of integrated circuits is done by randomly selecting 25 chips from the previous days manufacturing run. Each of the 25 chips is tested. If two or more chips are faulty, then the entire run is discarded. Previously gathered evidence indicates that the defect rate for chips is 0.0016. What is the probability that a manufacturing run of chips will be discarded?
- (14) The probabilities of a child being born male or female are not exactly equal to 1/2. Typically there are nearly 105 live male births per 100 live female births. Determine the expected number of female children in a family of 6 total children using these birth ratios and ignoring infant mortality.
- (15) Suppose a standard deck of 52 cards is well shuffled and one card at a time will be drawn without replacement from the deck. What is the expected value of the first ace drawn (in other words, of the first, second, third, *etc* cards drawn, on average which will be the first ace drawn)?
- (16) Show that for constants a and b and discrete random variable X that $E[aX + b] = aE[X] + b$.
- (17) For the situation described in exercise 15 determine the variance in the occurrence of the first ace drawn.
- (18) Show that for constants a and b and discrete random variable X that $\text{Var}(aX + b) = a^2\text{Var}(X)$.

Chapter 3

Normal Random Variables and Probability

Whereas in Chapter 2 random variables could take on only a finite number of values taken from a set with gaps between the values, in the present chapter, **continuous** random variables will be described. A continuous random variable can take on an infinite number of different values from a range without gaps. Calculus-based methods for determining the expected value and standard deviation of a continuous random variable will be described. A very useful type of continuous random variable for the study of financial mathematics is the **normal** random variable. We will see that when many random factors and influences come together (as in the complex situation of a financial market), the sum of the influences can be modeled by a normal random variable. Finally we will examine some stock market data and detect the presence of normal randomness in the fluctuations of stock prices.

3.1 Continuous Random Variables

Our understanding of probability must change if we are to understand the differences between discrete and continuous random variables. Suppose we consider the interval $[0, 1]$. If we think of only the integers contained in this interval, then only the discrete set $\{0, 1\}$ need be considered. If the probability of selecting either integer from this set is $1/2$ then we remain in the realm of discrete probability, random variables, and distributions. If we think of selecting, with equal likelihood, a number from the interval $[0, 1]$ of the form $k/10$ where $k \in \{0, 1, \dots, 10\}$, then once again our approach to probability remains discrete. The $P(X = k/10) = \frac{1}{11}$. Continuing in this way we see that if consider only the real numbers in $[0, 1]$ of the form k/n where $k \in \{0, 1, \dots, n\}$, then $P(X = k/n) = \frac{1}{n+1}$. So long as n is finite we

are dealing with the familiar concept of discrete probability. What happens as $n \rightarrow \infty$? In one sense we can think of this limiting case as the case of **continuous probability**. We say continuous because in the limit, the gaps between the outcomes in the sample space disappear. However, if we retain the old notion of probability, the likelihood of choosing a particular number from the interval $[0, 1]$ becomes

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} = 0.$$

This is correct but brings to mind a paradox. Does this imply that the probability of choosing any number from $[0, 1]$ is 0? Surely we must choose some number. What is needed is a shift in our notion of probability when dealing with continuous random variables. Instead of determining the probability that a continuous random variable equals some real number, the meaningful determination is the likelihood that the random variable lies in a set, usually an interval or finite union of intervals on the real line. These notions are defined next.

A random variable X has a **continuous distribution** if there exists a non-negative function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that for an interval $[a, b]$ the

$$P(a \leq X \leq b) = \int_a^b f(x) dx. \quad (3.1)$$

The function f which is known as the **probability distribution function** must, in addition to satisfying $f(x) \geq 0$, have the following property,

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (3.2)$$

These properties are the analogues of properties of discrete random variables and their distributions. The probabilities of individual values of the random variable are non-negative. In fact, as explained above $P(X = x) = 0$ for a continuous random variable provided $f(t)$ is continuous at $t = x$. By contrast $P(X = x)$ can be greater than zero for a discrete random variable. The sum, expressed now as an integral over the real number line, of the values of the probability distribution function must be one. Interpreted graphically, if X represents a continuous random variable then $P(a \leq X \leq b)$ represents the area of the region bounded by the graph of the probability distribution function, the x -axis, and the lines $x = a$ and $x = b$. See Fig. 3.1. The total area under the graph of the probability density function is unity.

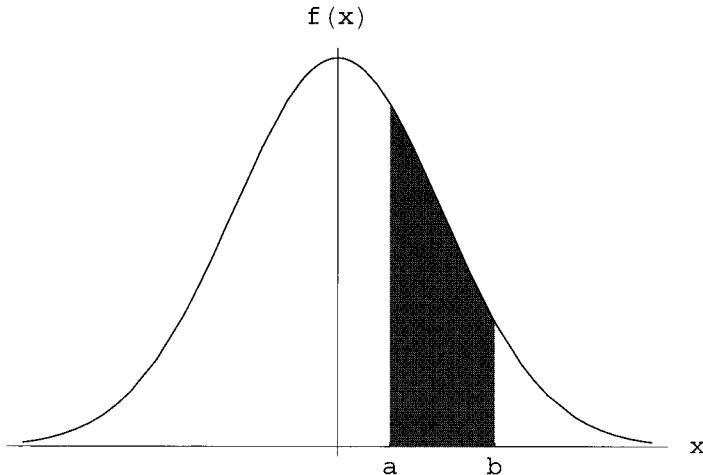


Fig. 3.1 An example of a probability distribution function. The shaded area represents the $P(a \leq X \leq b)$.

Perhaps the most elementary of the continuous random variables is the **uniformly distributed random variable**. A random variable X is uniformly distributed in the interval $[a, b]$ if the probability that X belongs to any subinterval of $[a, b]$ is proportional to the length of the subinterval. The definition of the continuous uniform random variable allows us to determine the probability density function for X . Since the length of the subinterval is proportional to the probability that X lies in the subinterval then the probability density function $f(x)$ must be constant. Suppose this constant is k . From the property expressed in Eq. (3.2) we know that

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_a^b k dx = \int_a^b \frac{1}{b-a} dx.$$

We are assuming here that the probability distribution function vanishes outside of the interval $[a, b]$. This simplifies the evaluation of the improper integral. Thus the probability density function for a continuous random variable on interval $[a, b]$ is the piecewise-defined function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

Example 3.1 Random variable X is continuously randomly distributed in the interval $[-1, 12]$. Find the probability that $2 \leq X \leq 7$.

Since according to the previous discussion of uniformly distributed random variables, the probability density function of X is

$$f(x) = \begin{cases} \frac{1}{13} & \text{if } -1 \leq x \leq 12 \\ 0 & \text{otherwise,} \end{cases}$$

we have

$$\mathrm{P}(2 \leq X \leq 7) = \int_2^7 f(x) dx = \int_2^7 \frac{1}{13} dx = \frac{5}{13}.$$

Now that continuous random variables and their probability distributions have been introduced we can begin to discuss formulae for determining their means and variances.

3.2 Expected Value of Continuous Random Variables

By definition the **expected value** or **mean** of a continuous random variable with probability density function $f(x)$ is

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} xf(x) dx. \quad (3.3)$$

This equation is analogous to Eq. (2.4) for the case of a discrete random variable. In the continuous case the product of the value of the random variable and its probability density function is summed (this time through the use of a definite integral) over all values of the random variable. The expected value is only meaningful in cases in which the improper integral in (3.3) converges.

Example 3.2 Find the expected value of X if X is a continuously uniformly distributed random variable on the interval $[-50, 75]$.

Using Eq. (3.3) we have

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} \frac{x}{75 - (-50)} dx = \frac{1}{125} \int_{-50}^{75} x dx = \frac{1}{250} (75^2 - (-50)^2) = \frac{25}{2}.$$

The results which follow in this section will be similar to the results that were presented for discrete random variables. In most cases the notion of a discrete sum need only be replaced by a definite integral to justify these new results. To keep the exposition as brief as possible, many of the proofs of results in this section will be left to the reader. The expected value of

a function g of a continuously distributed random variable X which has probability distribution function f is defined as

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) dx, \quad (3.4)$$

provided the improper integral converges absolutely, *i.e.*, $\mathbb{E}[g(X)]$ is defined if and only if

$$\int_{-\infty}^{\infty} |g(x)|f(x) dx < \infty.$$

Joint probability distributions for continuous random variables are defined similarly to those for discrete random variables. A joint probability distribution for a pair of random variables is a non-negative function $f(x, y)$ for which

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy.$$

Consequently the expected value of a function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ of the two random variables X and Y is defined as

$$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y) dx dy,$$

again, provided the integral is absolutely convergent.

Example 3.3 Consider the jointly distributed random variables $(X, Y) \in [1, \infty] \times [-1, 2]$ whose distribution is the function $f(x, y) = \frac{2}{3x^3}$. We can find the mean of $X + Y$ as follows.

$$\begin{aligned} \mathbb{E}[X + Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) \frac{2}{3x^3} dx dy \\ &= \int_{-1}^2 \int_1^{\infty} (x + y) \frac{2}{3x^3} dx dy \\ &= \int_{-1}^2 \int_1^{\infty} \frac{2x}{3x^3} dx dy + \int_{-1}^2 \int_1^{\infty} \frac{2y}{3x^3} dx dy \\ &= \int_1^{\infty} \int_{-1}^2 \frac{2x}{3x^3} dy dx + \int_{-1}^2 \int_1^{\infty} \frac{2y}{3x^3} dx dy \\ &= 3 \int_1^{\infty} \frac{2}{3x^2} dx + \int_{-1}^2 \frac{y}{3} dy \\ &= 2 + \frac{1}{2} = \frac{5}{2} \end{aligned}$$

If we examine the example above we recognize that $E[X + Y] = E[X] + E[Y]$. This is true in general, not just for the previous example. The additivity property of the expected value of continuous random variables is stated in the following theorem.

Theorem 3.1 *If X_1, X_2, \dots, X_k are continuous random variables with joint probability distribution $f(x_1, x_2, \dots, x_k)$ then $E[X_1 + X_2 + \dots + X_k] = E[X_1] + E[X_2] + \dots + E[X_k]$.*

The reader should again note that, just as was the case in the previous chapter, the previous theorem is true for random variables which are dependent or independent. Before turning our attention to products of continuous random variables we will define the notion of a marginal distribution. If X and Y are continuous random variables with joint distribution $f(x, y)$ then the **marginal distribution** for X is defined as the function

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy.$$

The marginal distribution for Y is defined similarly. By definition two continuous random variables are **independent** if the joint distribution factors as the product of the marginal distributions of X and Y .

$$f(x, y) = g(x)h(y)$$

This property is used in the proof of the following theorem.

Theorem 3.2 *Let X_1, X_2, \dots, X_k be pairwise independent random variables with joint distribution $f(x_1, x_2, \dots, x_k)$, then*

$$E[X_1 X_2 \cdots X_k] = E[X_1] E[X_2] \cdots E[X_k].$$

Just as was the case for discrete random variables in the previous chapter, the expected value of a continuous random variable can be thought of as the average value of the outcome of an infinite number of experiments. The variance and standard deviation of a continuous random variable again mirror the concept earlier defined for discrete random variables.

3.3 Variance and Standard Deviation

The variance of a continuous random variable is a measure of the spread of values of the random variable about the expected value of the random

variable. The variance is defined as

$$\text{Var}(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx, \quad (3.5)$$

where $\mu = E[X]$ and $f(x)$ is the probability distribution function of X . The variance may be interpreted as the squared deviation of a random variable from its expected value. An alternative formula for the variance is given in the following theorem.

Theorem 3.3 *Let X be a random variable with probability distribution f and mean μ , then the variance of X is $E[X^2] - \mu^2$.*

Example 3.4 Find the variance of the continuous random variable whose probability distribution is given by

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

The reader may readily check that $\mu = E[X] = 2/3$. Now by use of Theorem 3.3 we have

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} x^2(2x) dx - \left(\frac{2}{3}\right)^2 \\ &= \int_0^1 2x^3 dx - \frac{4}{9} \\ &= \int_0^1 2x^3 dx - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}. \end{aligned}$$

In the exercises the reader will be asked to prove the following two theorems which extend to continuous random variables results we have already seen for discrete random variables.

Theorem 3.4 *Let X be a continuous random variable with probability distribution $f(x)$ and let $a, b \in \mathbb{R}$, then*

$$\text{Var}(aX + b) = a^2 \text{Var}(X).$$

Theorem 3.5 *Let X_1, X_2, \dots, X_k be pairwise independent continuous random variables with joint probability distribution $f(x_1, x_2, \dots, x_k)$, then*

$$\text{Var}(X_1 + X_2 + \dots + X_k) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_k).$$

By definition the standard deviation of a continuous random variable is the square root of its variance. The standard deviation is sometimes

denoted σ . On occasion we will denote the variance of a random variable as σ^2 .

3.4 Normal Random Variables

For our purposes one of the most important and useful continuous random variables will be the **normally distributed random variable**. A continuous random variable obeying a normal distribution is frequently said to follow the “bell curve”. Many measurable quantities found in nature seem to have normal distributions, for example adult heights and weights. Statisticians, mathematicians, and physical scientists frequently assume any quantity subject to a large number of small independently acting forces (regardless of their distributions) is normally distributed. The normal distribution even finds its way into the financial arena via the assumption that movements in the price of an asset are subject to a large number of incompletely understood political, economic, and social forces, thus justifying the assumption that these changes in value are related to a normal random variable (this assumption will be explored further later).

The normal random variable can also be thought of as the limiting behavior of the binomial random variable introduced in Chapter 2. We will explore the derivation of the probability distribution for the normal random variable in this section. This derivation is outlined in [Bleecker and Csordas (1996), pg. 139]. We begin by supposing that a particle sits at the origin of the x -axis and may move the left or the right a distance Δx in a time interval of length Δt . The (discrete) probability of moving left is $1/2$ which also happens to be the probability of moving right. We will show that if $n \in \mathbb{N}$ and $m \in \mathbb{Z}$ with $-n \leq m \leq n$ and if $n - m$ is even then the probability that the particle is at location $X = m\Delta x$ at time $t = n\Delta t$ is

$$\frac{n! \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)! \left(\frac{1}{2}(n-m)\right)!}. \quad (3.6)$$

We will assume that all the steps taken by the particle are independent and identically distributed. Let the total number of steps taken be n and the number of steps taken to the right be r . Consequently the number of steps taken to the left will be $n - r$. The position of the particle after n steps will be

$$(r - (n - r))\Delta x = (2r - n)\Delta x = m\Delta x,$$

where we have assigned $m = 2r - n$. Using the binomial probability distribution given in Eq. (2.3) we can state that the probability that the particle is at position $m\Delta x$ after n steps is

$$\begin{aligned} P(X = m\Delta x) &= P(X = (2r - n)\Delta x) \\ &= \binom{n}{r} \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{n-r} \\ &= \frac{n!}{r!(n-r)!} \left(\frac{1}{2}\right)^n \\ &= \frac{n! \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)! \left(\frac{1}{2}(n-m)\right)!} \end{aligned}$$

since $n + m = 2r$. This is the same probability as in Eq. (3.6).

This claim can also be proved by induction on n . If $n = 1$ then $m = \pm 1$ (the particle is not allowed to remain in place).

$$\begin{aligned} P(X = -\Delta x) &= \frac{1}{2} = \frac{1! \left(\frac{1}{2}\right)^1}{\left(\frac{1}{2}(1 + (-1))\right)! \left(\frac{1}{2}(1 - (-1))\right)!} \\ P(X = \Delta x) &= \frac{1}{2} = \frac{1! \left(\frac{1}{2}\right)^1}{\left(\frac{1}{2}(1 + 1)\right)! \left(\frac{1}{2}(1 - 1)\right)!} \end{aligned}$$

Thus the claim is true for $n = 1$. Now suppose the claim is true for $k \leq n-1$. If the particle will move to $m\Delta x$ at time $n\Delta t$, then at time $(n-1)\Delta t$ the particle must be at either $(m-1)\Delta x$ or $(m+1)\Delta x$. Therefore at time $t = n\Delta t$

$$\begin{aligned} P(X = m\Delta x) &= \frac{1}{2}P(X = (m-1)\Delta x) + \frac{1}{2}P(X = (m+1)\Delta x) \\ &= \frac{1}{2} \left(\frac{(n-1)! \left(\frac{1}{2}\right)^{n-1}}{\left(\frac{1}{2}(n-1+m-1)\right)! \left(\frac{1}{2}(n-1-(m-1))\right)!} \right) \\ &\quad + \frac{1}{2} \left(\frac{(n-1)! \left(\frac{1}{2}\right)^{n-1}}{\left(\frac{1}{2}(n-1+m+1)\right)! \left(\frac{1}{2}(n-1-(m+1))\right)!} \right) \\ &= \frac{(n-1)! \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)! \left(\frac{1}{2}(n-m-2)\right)!} \\ &\quad + \frac{(n-1)! \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m-2)\right)! \left(\frac{1}{2}(n-m)\right)!} \\ &= \frac{n! \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)! \left(\frac{1}{2}(n-m)\right)!}, \end{aligned}$$

which is the probability given in expression (3.6).

Next we will make use of Stirling's Formula which approximates $n!$ when n is large.

$$n! \approx \sqrt{2\pi} e^{-n} n^{n+1/2} \quad (3.7)$$

If we apply this formula to every factorial present in expression (3.6) we obtain the following sequence of equivalent expressions.

$$\begin{aligned} & \frac{\sqrt{2\pi} e^{-n} n^{n+1/2} \left(\frac{1}{2}\right)^n}{\sqrt{2\pi} e^{-(n+m)/2} \left(\frac{1}{2}(n+m)\right)^{(n+m+1)/2} \sqrt{2\pi} e^{-(n-m)/2} \left(\frac{1}{2}(n-m)\right)^{(n-m+1)/2}} \\ &= \frac{1}{\sqrt{2\pi}} \frac{n^{n+1/2} \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)^{(n+m+1)/2} \left(\frac{1}{2}(n-m)\right)^{(n-m+1)/2}} \\ &= \frac{2}{\sqrt{2\pi}} \frac{n^{n+1/2}}{(n+m)^{(n+m+1)/2} (n-m)^{(n-m+1)/2}} \\ &= \frac{2}{\sqrt{2n\pi}} \frac{n^{(n+m+1)/2} n^{(n-m+1)/2}}{(n+m)^{(n+m+1)/2} (n-m)^{(n-m+1)/2}} \\ &= \frac{2}{\sqrt{2n\pi}} \left(\frac{n}{n+m}\right)^{(n+m+1)/2} \left(\frac{n}{n-m}\right)^{(n-m+1)/2} \\ &= \frac{2}{\sqrt{2n\pi}} \left(1 + \frac{m}{n}\right)^{-(n+m+1)/2} \left(1 - \frac{m}{n}\right)^{-(n-m+1)/2} \\ &= \frac{2}{\sqrt{2n\pi}} \left(1 + \frac{m}{n}\right)^{-m/2} \left(1 - \frac{m}{n}\right)^{m/2} \left(1 - \frac{m^2}{n^2}\right)^{-(n+1)/2} \end{aligned}$$

Now we will make use of the fact that $m = x/\Delta x$ and $n = t/\Delta t$. The last expression above is then equal to

$$\frac{2\sqrt{\Delta t}}{\sqrt{2\pi t}} \left(1 + \frac{x\Delta t}{t\Delta x}\right)^{-\frac{x}{2\Delta x}} \left(1 - \frac{x\Delta t}{t\Delta x}\right)^{\frac{x}{2\Delta x}} \left(1 - \left[\frac{x\Delta t}{t\Delta x}\right]^2\right)^{-\frac{1+t/\Delta t}{2}}.$$

We will suppose the square of the step size along the x -axis is related to the step size in t via the formula $(\Delta x)^2 = 2k\Delta t$. Using this relationship to replace the Δt 's in the expression above yields

$$\frac{\Delta x}{\sqrt{k\pi t}} \left(1 + \frac{x}{2kt}\Delta x\right)^{-\frac{x}{2\Delta x}} \left(1 - \frac{x}{2kt}\Delta x\right)^{\frac{x}{2\Delta x}} \left(1 - \left[\frac{x}{2kt}\Delta x\right]^2\right)^{-\frac{kt}{(\Delta x)^2} - \frac{1}{2}}.$$

So far what we have derived is an approximation (which we will treat as an equality for large n or, what is equivalent, for small Δx) for the probability that the particle is located at position $x = m\Delta x$ at time $t = n\Delta t$. To derive

a probability distribution for a continuously distributed random variable we must divide the expression above by $2\Delta x$ and take the limit as $\Delta x \rightarrow 0$. Readers accustomed to numerical approximations for derivatives will recognize that we have used the centered difference formula for the first derivative in order to develop a difference quotient whose limit will be the probability distribution. Background information on finite difference approximations to derivatives can be found in [Burden and Faires (2005)]. We will make use of the result that $\lim_{\Delta x \rightarrow 0} (1 + a\Delta x)^{1/\Delta x} = e^a$. Let us define $f(x, t)$ as

$$\begin{aligned} f(x, t) &= \frac{1}{2\sqrt{k\pi t}} \lim_{\Delta x \rightarrow 0} \left(1 + \frac{x}{2kt}\Delta x\right)^{\frac{-x}{2\Delta x}} \left(1 - \frac{x}{2kt}\Delta x\right)^{\frac{x}{2\Delta x}} \left(1 - \left[\frac{x}{2kt}\Delta x\right]^2\right)^{-\frac{kt}{(\Delta x)^2} - \frac{1}{2}} \\ &= \frac{1}{2\sqrt{k\pi t}} \left(e^{\frac{x}{2kt}}\right)^{-\frac{x}{2}} \left(e^{-\frac{x}{2kt}}\right)^{\frac{x}{2}} \left(e^{-\frac{x^2}{2kt^2}}\right)^{-\frac{t}{2}} \\ &= \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} e^{-\frac{x^2}{4kt}} e^{\frac{x^2}{4kt}} \\ &= \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}. \end{aligned}$$

We must verify that this expression satisfies the properties of a probability distribution function. For $k, t > 0$ the expression is non-negative. Since for $|x| \geq 1$ it is true that $0 < e^{-x^2} \leq e^{-|x|}$, then the integral of $f(x, t)$ over the entire real number line converges, suppose we write:

$$0 < \int_{-\infty}^{\infty} \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx = S < \infty.$$

Therefore we have that

$$\begin{aligned} S^2 &= \int_{-\infty}^{\infty} \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx \int_{-\infty}^{\infty} \frac{1}{2\sqrt{k\pi t}} e^{-\frac{y^2}{4kt}} dy \\ &= \frac{1}{4k\pi t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/4kt} dx dy. \end{aligned}$$

Switching to polar coordinates by making the substitutions $x = r \cos \theta$,

$y = r \cos \theta$, and $dx dy = r dr d\theta$ produces

$$\begin{aligned} S^2 &= \frac{1}{4k\pi t} \int_0^{2\pi} \int_0^\infty r e^{-r^2/4kt} dr d\theta \\ &= \frac{1}{2kt} \int_0^\infty r e^{-r^2/4kt} dr \\ &= \int_0^\infty e^{-u} du \\ &= \lim_{M \rightarrow \infty} \int_0^M e^{-u} du \\ &= \lim_{M \rightarrow \infty} (-e^{-M} + 1) = 1. \end{aligned}$$

Therefore $S = 1$. Notice in the third step we made use of the substitution $u = r^2/4kt$. The reader may consult a book on complex analysis such as [Marsden and Hoffman (1987)] for a more formal proof of this condition. Hence we see that the function given above satisfies the non-negativity condition and unit area condition of a probability distribution for a continuous random variable. Thus we have derived the probability distribution function for the particle. For a fixed value of t the probability density function has the familiar bell shape as seen in Fig. 3.2. The definition of expected value given in Eq. (3.3) can be used to determine the average position of

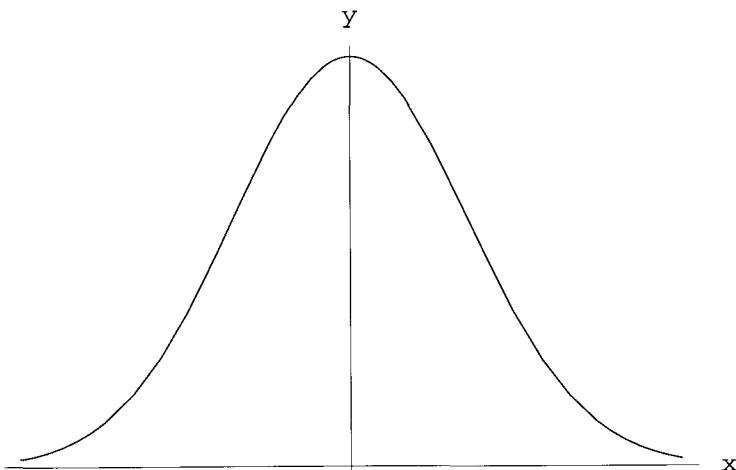


Fig. 3.2 For a fixed value of $t > 0$ the function $f(x, t) = \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}$ will have a graph resembling the often discussed “bell curve.”

the particle.

$$\begin{aligned}
 E[X] &= \int_{-\infty}^{\infty} \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx \\
 &= \lim_{M \rightarrow \infty} \int_0^M \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx + \lim_{N \rightarrow -\infty} \int_{-N}^0 \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx \\
 &= \lim_{M \rightarrow \infty} \sqrt{\frac{k\pi}{t}} \left(1 - e^{-M^2/4kt} \right) + \lim_{N \rightarrow -\infty} \sqrt{\frac{k\pi}{t}} \left(e^{-N^2/4kt} - 1 \right) \\
 &= 0
 \end{aligned}$$

Thus, as might be expected, the average location of the particle is at the origin independent of t . This is due to the assumption that the particle has no preference for movement to the left or right. In a similar manner we may determine the variance of the position of the particle.

$$\begin{aligned}
 \text{Var}(X) &= \int_{-\infty}^{\infty} \frac{x^2}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx - (E[X])^2 \\
 &= \frac{1}{\sqrt{k\pi t}} \lim_{M \rightarrow \infty} \int_0^M x^2 e^{-\frac{x^2}{4kt}} dx \\
 &= \frac{1}{\sqrt{k\pi t}} \lim_{M \rightarrow \infty} 2kt \left(Me^{-M^2/4kt} - \int_0^M e^{-x^2/4kt} dx \right) \\
 &= \frac{1}{\sqrt{k\pi t}} \lim_{M \rightarrow \infty} 2kt \left(Me^{-M^2/4kt} - \int_0^M e^{-x^2/4kt} dx \right) \\
 &= \frac{2kt}{\sqrt{k\pi t}} \int_0^{\infty} e^{-x^2/4kt} dx \\
 &= 2kt \cdot \frac{1}{2\sqrt{k\pi t}} \int_{-\infty}^{\infty} e^{-x^2/4kt} dx \\
 &= 2kt
 \end{aligned}$$

We were able to avoid evaluating the last improper integral by making use of the unit area property of the probability distribution. From the derivation that $\sigma^2 = 2kt$ we know that the “spread” in the location of the particle increases with time. This can be readily seen in the surface plot shown in Fig. 3.3.

If the particle were initially placed at location μ but all other assumptions remained the same then the probability distribution would simply be shifted in the x direction by μ . Thus we may define the **normal proba-**

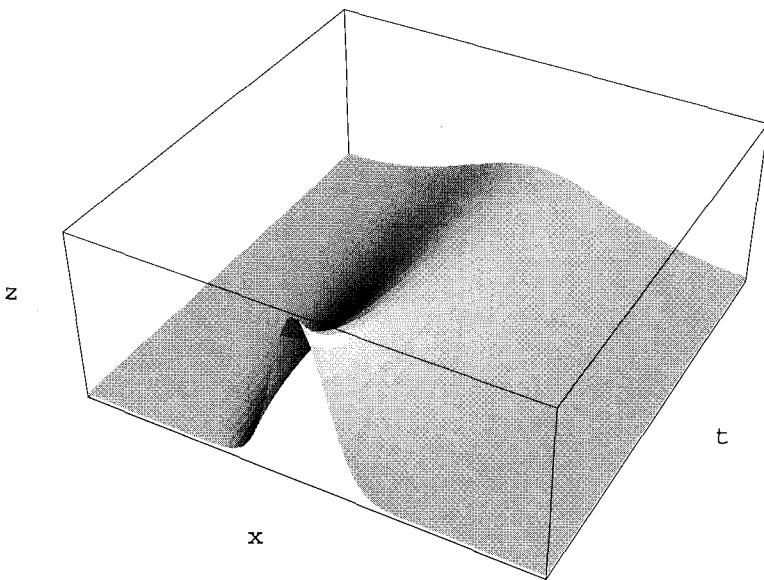


Fig. 3.3 For $t > 0$ the function $f(x, t) = \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}$ retains the bell-shaped cross section, but the profile flattens and broadens as t increases.

bility distribution with mean μ and variance σ^2 to be the function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (3.8)$$

When $\mu = 0$ and $\sigma = 1$, this is referred to as the **standard normal probability distribution**.

In the sequel the **cumulative distribution function** $\phi(x)$ where

$$\phi(x) = P(X < x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (3.9)$$

will be frequently used in discussions involving probabilities of normal random variables. For example if X is a standard normal random variable then $P(X < 0) = \phi(0) = 1/2$. It is frequently helpful when dealing with normally distributed random variables to perform a mathematical change of variable which produces a standard normal random variable.

Theorem 3.6 *If X is a normally distributed random variable with expected value μ and variance σ^2 , then $Z = (X - \mu)/\sigma$ is normally distributed*

with an expected value of zero and a variance of one.

The proof of this result is left to the reader in exercise 17. In the sequel the symbol Z will be used to denote a normally distributed random variable with standard normal distribution.

3.5 Central Limit Theorem

The name Central Limit Theorem is given to several results concerned with the distribution of sample means or the distribution of the sum of random variables. The proofs of these theorems are beyond the scope of this work. The interested reader should consult a book on the mathematical underpinnings of statistics for proofs of the various versions of the Central Limit Theorem (for example [DeGroot (1975)]). Nevertheless we can observe the consequences of the Central Limit Theorem on data from a random number simulation. The reader should keep the following points in mind. Given a random variable X which can be either discrete or continuous and which may have any probability distribution, we may collect a sample of size n and denote the mean of that sample \bar{X}_n . If the process of collecting multiple samples and calculating their means is repeated then we can treat the sample means as random variables in their own right. One version of the Central Limit Theorem due to Lindeberg and Lévy implies that the sample means become normally distributed as the sample size becomes large.

Theorem 3.7 *If random variables X_1, X_2, \dots, X_n form a random sample of size n from a probability distribution with mean μ and standard deviation σ then for all x*

$$\lim_{n \rightarrow \infty} P\left(\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq x\right) = \phi(x).$$

Example 3.5 The reader can replicate this example on most programmable computing devices. Collect 5000 samples of size n (where $n \in \{2, 5, 10, 20, 40\}$) of a uniformly distributed continuous random variable on the interval $[0, 1]$. Compute the means of each sample of size n and plot the frequency histogram of the means. As can be seen in Fig. 3.4 as n increases the histograms take on the appearance of a normal probability distribution function.

Another version of the Central Limit Theorem is due to Liapounov and concerns the asymptotic distribution of a sum of random variables.

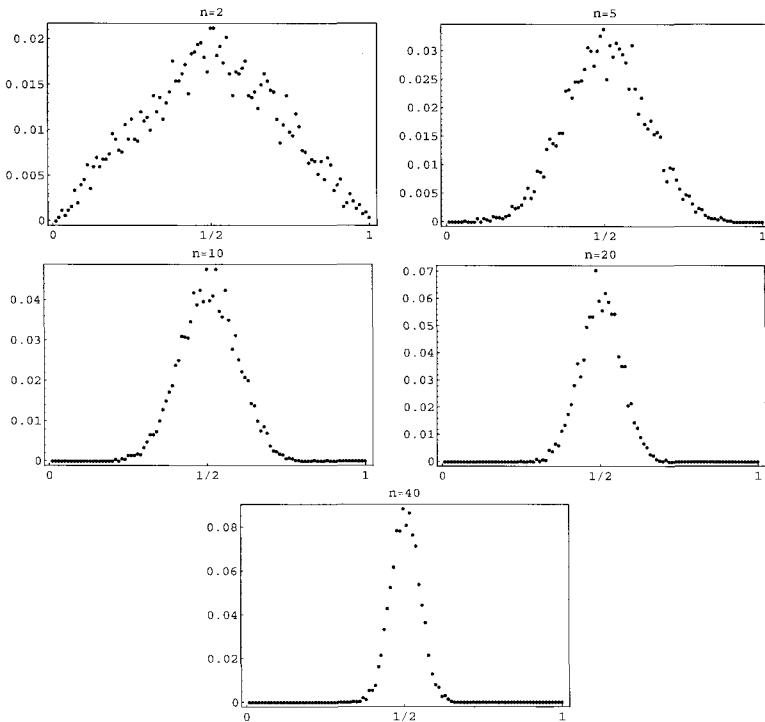


Fig. 3.4 An illustration of the Central Limit Theorem due to Lindeberg and Lévy. As the sample size increases the distribution of the sample means becomes more normal in appearance.

Suppose the random variables X_1, X_2, \dots, X_n are pairwise independent but not necessarily identically distributed. We will assume that for each $i \in \{1, 2, \dots, n\}$, $E[X_i] = \mu_i$ and that $\text{Var}(X_i) = \sigma_i^2$. Now define a new random variable Y_n as

$$Y_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{\sqrt{\sum_{i=1}^n \sigma_i^2}}.$$

Using the assumption that the random variables are pairwise independent and Theorems 3.1, 3.4, and 3.5 we can determine that $E[Y_n] = 0$ and $\text{Var}(Y_n) = 1$. The following theorem establishes that as n becomes large Y_n is approximately normally distributed.

Theorem 3.8 *Suppose that the infinite collection $\{X_i\}_{i=1}^{\infty}$ of random variables are pairwise independent and that for each $i \in \mathbb{N}$ we have*

$\mathbb{E} [|X_i - \mu_i|^3] < \infty$. If in addition,

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \mathbb{E} [|X_i - \mu_i|^3]}{\left(\sum_{i=1}^n \sigma_i^2\right)^{3/2}} = 0$$

then for any $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} \mathbb{P}(Y_n \leq x) = \phi(x)$$

where random variable Y_n is defined as above.

Example 3.6 Once again, in place of a proof of Theorem 3.8 we will present a numerical simulation illustrating this result. The reader is encouraged to pursue their own similar numerical exploration. The data used in the following plots were generated by defining n continuous uniform distributions on $[\alpha_i, \beta_i]$ for $i = 1, 2, \dots, n$ where each of α_i and β_i were randomly, uniformly distributed in $[-100, 100]$ and independently selected. The mean and variance of each uniform distribution were calculated. Then a random variable X_i was randomly chosen from each uniform distribution on $[\alpha_i, \beta_i]$ and Y_n was computed as above. A sample for Y_n of size 5000 was collected and frequency histograms were created. These are shown in Fig. 3.5. It is readily seen that as n increases the distribution of Y_n becomes more normal in appearance.

3.6 Lognormal Random Variables

While normal random variables are central to our discussion of probability and the upcoming Black-Scholes option pricing formula, of equal importance will be continuous random variables which are distributed in a log-normal fashion. A random variable X is a **lognormal random variable** with parameters μ and σ if $Y = \ln X$ is a normally distributed random variable with mean μ and variance σ^2 . When referring to a lognormal random variable, the parameters μ and σ are often called the **drift** and **volatility** respectively. A lognormal random variable is a continuous random variable which takes on values in the interval $(0, \infty)$. From Eq. (3.8) we see that for the normal random variable Y , the probability that $Y < y$ is expressed as the integral

$$\mathbb{P}(Y < y) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^y e^{-(t-\mu)^2/2\sigma^2} dt.$$

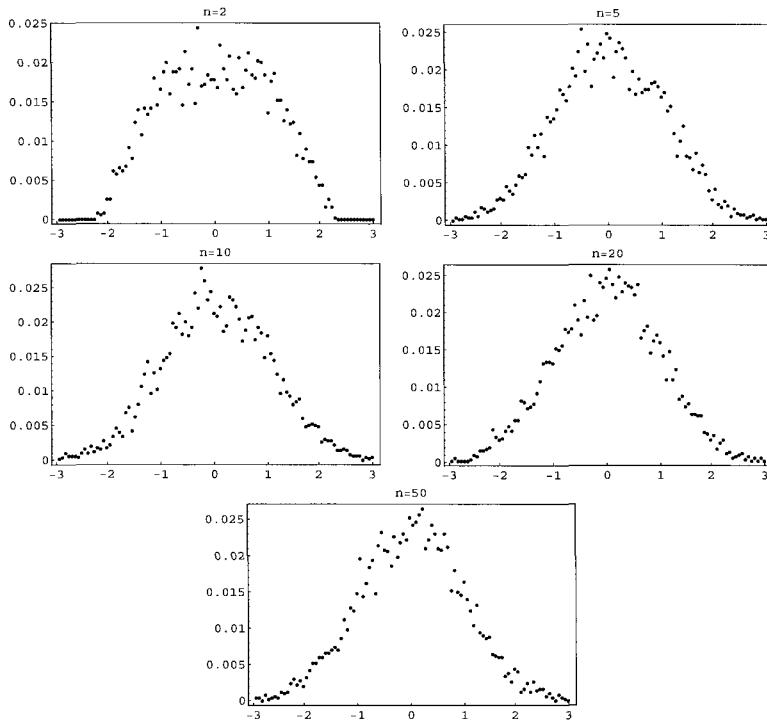


Fig. 3.5 An illustration of the Central Limit Theorem due to Liapounov. As the sample size increases the distribution of the sum of the random variables becomes more normal in appearance.

Making the change of variable $t = \ln u$ we see then that

$$\begin{aligned} P(\ln X < y) &= \frac{1}{\sigma\sqrt{2\pi}} \int_0^{e^y} \frac{1}{u} e^{-(\ln u - \mu)^2/2\sigma^2} du \\ &= P(X < e^y). \end{aligned}$$

Therefore a lognormally distributed random variable with parameters μ and σ^2 has a probability distribution function of the form

$$f(x) = \frac{1}{(\sigma\sqrt{2\pi})x} e^{-(\ln x - \mu)^2/2\sigma^2} \quad (3.10)$$

for $0 < x < \infty$. The probability that the lognormally distributed random variable X is less than $y > 0$ is then

$$P(X < y) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^y \frac{1}{x} e^{-(\ln x - \mu)^2/2\sigma^2} dx = \phi(\ln y), \quad (3.11)$$

where $\phi(x)$ is the cumulative probability distribution function for a normal random variable with mean zero and variance one. Similarly the $P(X > y) = 1 - \phi(\ln y)$ for a lognormally distributed random variable.

Using the definitions of expected value and variance we can prove the following lemma.

Lemma 3.1 *If X is a lognormal random variable with parameters μ and σ then*

$$E[X] = e^{\mu + \sigma^2/2} \quad (3.12)$$

$$\text{Var}(X) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) \quad (3.13)$$

Proof. Since X is assumed to be lognormal, then $Y = \ln X$ is normally distributed with $E[Y] = \mu$ and $\text{Var}(Y) = \sigma^2$. By Eq. (3.3)

$$\begin{aligned} E[X] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^y e^{-(y-\mu)^2/2\sigma^2} dy \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(y-(\mu+\sigma^2))^2/2\sigma^2} e^{\mu+\sigma^2/2} dy \\ &= e^{\mu+\sigma^2/2} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(y-(\mu+\sigma^2))^2/2\sigma^2} dy \\ &= e^{\mu+\sigma^2/2} \end{aligned}$$

The last equality is true since the integral represents the area under the probability distribution curve for a normal random variable with mean $\mu + \sigma^2$ and variance σ^2 .

Likewise by Eq. (3.5)

$$\begin{aligned}
 \text{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{2y} e^{-(y-\mu)^2/2\sigma^2} dy - \left(e^{\mu+\sigma^2/2}\right)^2 \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(y-(\mu+2\sigma))^2/2} e^{2(\mu+\sigma^2)} dy - e^{2\mu+\sigma^2} \\
 &= e^{2(\mu+\sigma^2)} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(y-(\mu+2\sigma))^2/2} dy - e^{2\mu+\sigma^2} \\
 &= e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2} \\
 &= e^{2\mu+\sigma^2} (e^{\sigma^2} - 1)
 \end{aligned}$$

□

Now we will apply the concept of the lognormally distributed random variable to the situation of the price of a stock or other security. Suppose that the selling price of a security will be measured daily and that the starting measurement will be denoted $S(0)$. For $n \geq 1$, we will let $S(n)$ denote the selling price on day n . To a good approximation the ratios of consecutive days' selling prices are lognormal random variables, *i.e.* the expressions $X(n) = S(n)/S(n-1)$ for $n \geq 1$ are lognormally distributed. It is also generally assumed that the ratios are independent. If a sufficient number of measurements have been made that a financial analyst estimates the parameters of the random variable X are $\mu = 0.0155$ and $\sigma = 0.0750$, then questions regarding the likelihood of future prices of the security can be answered.

Example 3.7 First, what is the probability that the selling price of the stock on the next day will be higher than the price on the present day?

$$\begin{aligned}
 \mathbb{P}(X(n) > 1) &= \mathbb{P}(\ln X(n) > 0) \\
 &= \mathbb{P}\left(z > \frac{0 - \mu}{\sigma}\right) \\
 &= \mathbb{P}(z > -0.206667) \\
 &= 1 - \phi(-0.206667) \\
 &\approx 0.581865
 \end{aligned}$$

Thus for a security with the parameters measured and reported above, there is a better than one half probability of the selling price increasing on any given day.

Second, we may ask what is the probability that the selling price two days hence will be higher than the present selling price? At first glance the reader may be tempted to use the fact that each observation of the random variable is independent and hence declare that the sought after probability is $0.562^2 \approx 0.316$. This implies the mistaken assumption that the price of the security increased on each of the two days. This ignores the possibility that the price of the security could decrease on either (but not both) of the two days and still make a net gain over the two-day period. To correctly approach the problem we must make use of the Central Limit Theorem.

$$\begin{aligned}
P(S(n+2)/S(n) > 1) &= P\left(\frac{S(n+2)}{S(n+1)} \frac{S(n+1)}{S(n)} > 1\right) \\
&= P((X(n))^2 > 1) \\
&= P(2 \ln X(n) > 0) \\
&= P\left(z > \frac{0 - 2(0.0155)}{0.0750\sqrt{2}}\right) \\
&= P(z > -0.292271) \\
&= 1 - \phi(-0.292271) \\
&\approx 0.61496
\end{aligned}$$

Thus the probability of a gain over a two-day period is nearly twice as large as may have been first suspected.

The assumption that ratios of selling prices sampled at regular intervals for securities are lognormally distributed will be explored further in Chapter 5. Appendix A contains a sample of closing prices for Sony Corporation stock. Readers can collect their own data set from many corporations' websites or from finance sites such as finance.yahoo.com. The eager reader may want to collect data for a particular stock to further test the lognormal hypothesis.

3.7 Properties of Expected Value

In later chapters the reader will encounter frequent references to the positive part of the difference of two quantities. For instance if an investor may purchase a security worth X for price K , then the excess profit (if any) of the transaction is the positive part of the difference $X - K$. This is typically denoted $(X - K)^+ = \max\{0, X - K\}$. In this section a theorem and several corollaries are presented which will enable the reader to calculate

the expected value of $(X - K)^+$ when X is a continuous random variable and K is a constant.

Theorem 3.9 *Let X be a continuous random variable with probability distribution function $f(x)$. If K is a constant then*

$$\mathbb{E}[(X - K)^+] = \int_K^\infty \left(\int_x^\infty f(t) dt \right) dx. \quad (3.14)$$

Proof. For any $x \geq 0$, $(X - K)^+ \geq x$ if and only if $X - K \geq x$, which is true if and only if $X \geq K + x$.

$$\begin{aligned} \mathbb{P}((X - K)^+ \geq x) &= 1 - \mathbb{P}((X - K)^+ < x) \\ &= 1 - \mathbb{P}(X < K + x) \\ &= 1 - \int_{-\infty}^{K+x} f(t) dt \end{aligned}$$

Therefore we have

$$\begin{aligned} \mathbb{E}[(X - K)^+] &= \int_0^\infty \left(1 - \int_{-\infty}^{K+x} f(t) dt \right) dx \\ &= \int_K^\infty \left(1 - \int_{-\infty}^x f(u) du \right) dx \quad \text{where } u = t - K \\ &= \int_K^\infty \left(\int_x^\infty f(t) dt \right) dx. \end{aligned}$$

The last equality is true because $f(x)$ is a probability distribution for the continuous random variable X . \square

We can immediately put Theorem 3.9 to work finding the expected value of the positive part of a normal random variable. Corollary 3.1 is a special case of Theorem 3.9.

Corollary 3.1 *If X is normal random variable with mean μ and variance σ^2 and K is a constant, then*

$$\mathbb{E}[(X - K)^+] = \frac{\sigma}{\sqrt{2\pi}} e^{-(\mu-K)^2/2\sigma^2} + (\mu - K)\phi\left(\frac{\mu - K}{\sigma}\right). \quad (3.15)$$

Proof. By Theorem 3.9 the expected value of $(X - K)^+$ is

$$\mathbb{E}[(X - K)^+] = \int_K^\infty \frac{1}{\sqrt{2\pi}\sigma} \int_x^\infty e^{-(t-\mu)^2/2\sigma^2} dt dx$$

Employing integration by parts with

$$\begin{aligned} u &= \frac{1}{\sqrt{2\pi}\sigma} \int_x^\infty e^{-(t-\mu)^2/2\sigma^2} dt & v = x \\ du &= -\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} dx & dv = dx, \end{aligned}$$

the expected value becomes

$$\begin{aligned} \mathbb{E}[(X - K)^+] &= -\frac{K}{\sqrt{2\pi}\sigma} \int_K^\infty e^{-(t-\mu)^2/2\sigma^2} dt + \frac{1}{\sqrt{2\pi}\sigma} \int_K^\infty te^{-(t-\mu)^2/2\sigma^2} dt \\ &= -\frac{K}{\sqrt{2\pi}} \int_{-\infty}^{(\mu-K)/\sigma} e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_{(K-\mu)/\sigma}^\infty (\sigma z + \mu) e^{-z^2/2} dz \\ &= -K\phi\left(\frac{\mu-K}{\sigma}\right) + \frac{\mu}{\sqrt{2\pi}} \int_{(K-\mu)/\sigma}^\infty e^{-z^2/2} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{(K-\mu)/\sigma}^\infty ze^{-z^2/2} dz \\ &= -K\phi\left(\frac{\mu-K}{\sigma}\right) + \mu\left(1 - \phi\left(\frac{K-\mu}{\sigma}\right)\right) + \frac{\sigma}{\sqrt{2\pi}} \int_{(K-\mu)^2/2\sigma^2}^\infty e^{-u} du \\ &= (\mu - K)\phi\left(\frac{\mu-K}{\sigma}\right) + \frac{\sigma}{\sqrt{2\pi}} e^{-(\mu-K)^2/2\sigma^2}. \end{aligned}$$

□

For the purpose of completing the mean-variance analysis encountered later in Chapter 10 we must also be able to apply Theorem 3.9 when X is a lognormally distributed random variable.

Corollary 3.2 *If X is a lognormally distributed random variable with parameters μ and σ^2 and $K > 0$ is a constant then*

$$\mathbb{E}[(X - K)^+] = e^{\mu+\sigma^2/2} \phi\left(\frac{\mu - \ln K}{\sigma}\right) - K\phi\left(\frac{\mu - \ln K}{\sigma}\right) \quad (3.16)$$

Proof. According to Theorem 3.9

$$\begin{aligned} \mathbb{E}[(X - K)^+] &= \int_K^\infty \left(\int_x^\infty f(t) dt \right) dx \\ &= \int_K^\infty \left(\frac{1}{\sqrt{2\pi}\sigma} \int_x^\infty \frac{1}{t} e^{-(\ln t - \mu)^2/2\sigma^2} dt \right) dx \end{aligned}$$

where the last equation is obtained using Eq. (3.10). Making the substitution $z = (\ln t - \mu)/\sigma$ allows us to write

$$\mathbb{E}[(X - K)^+] = \int_K^\infty \left(\frac{1}{\sqrt{2\pi}} \int_{(\ln x - \mu)/\sigma}^\infty e^{-z^2/2} dz \right) dx$$

This last integral will be evaluated using integration by parts with

$$\begin{aligned} u &= \frac{1}{\sqrt{2\pi}} \int_{(\ln x - \mu)/\sigma}^{\infty} e^{-z^2/2} dz & v = x \\ du &= -\frac{1}{\sqrt{2\pi}\sigma} e^{-(\ln x - \mu)^2/2\sigma^2} dx & dv = dx. \end{aligned}$$

Therefore we have

$$\begin{aligned} \mathbb{E}[(X - K)^+] &= -\frac{K}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^{\infty} e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}\sigma} \int_K^{\infty} e^{-(\ln x - \mu)^2/2\sigma^2} dx \\ &= -K\phi\left(\frac{\mu - \ln K}{\sigma}\right) + \frac{1}{\sqrt{2\pi}\sigma} \int_{\ln K}^{\infty} e^z e^{-(z - \mu)^2/2\sigma^2} dz \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{\mu + \sigma^2/2} \int_{\ln K}^{\infty} e^{-(z - (\mu + \sigma^2))^2/2\sigma^2} dz - K\phi\left(\frac{\mu - \ln K}{\sigma}\right) \\ &= e^{\mu + \sigma^2/2} \phi\left(\frac{\mu - \ln K}{\sigma} + \sigma\right) - K\phi\left(\frac{\mu - \ln K}{\sigma}\right). \end{aligned}$$

□

3.8 Properties of Variance

In this section we will derive properties of the variance for the positive part of the difference of a continuous random variable and a constant. These results will be of use in later sections of this text.

Lemma 3.2 *Let X be a normally distributed random variable with mean μ and variance σ^2 . If K is a constant then*

$$\begin{aligned} \mathbb{E}[((X - K)^+)^2] &= ((\mu - 2K)^2 + \sigma^2) \phi\left(\frac{\mu - 2K}{\sigma}\right) + \frac{(\mu - 2K)\sigma}{\sqrt{2\pi}} e^{-\frac{(\mu - 2K)^2}{2\sigma^2}}. \quad (3.17) \end{aligned}$$

Proof. Since X is normally distributed with mean μ and variance σ^2 , then $X - K$ is normally distributed with mean $\mu - K$ and variance σ^2 .

$$\begin{aligned} \mathbb{E} [((X - K)^+)^2] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} ((x - K)^+)^2 e^{-(x-(\mu-K))^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_K^{\infty} (x - K)^2 e^{-(x-(\mu-K))^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{(2K-\mu)/\sigma}^{\infty} (\sigma t + \mu - 2K)^2 e^{-t^2/2} dt \end{aligned}$$

upon making the substitution $t = (x - (\mu - K))/\sigma$. Expanding the square in the last integrand and integrating produces

$$\begin{aligned} \mathbb{E} [((X - K)^+)^2] &= (\mu - 2K)^2 \phi\left(\frac{\mu - 2K}{\sigma}\right) + \frac{2(\mu - 2K)\sigma}{\sqrt{2\pi}} \int_{(2K-\mu)/\sigma}^{\infty} te^{-t^2/2} dt \\ &\quad + \frac{\sigma^2}{\sqrt{2\pi}} \int_{(2K-\mu)/\sigma}^{\infty} t^2 e^{-t^2/2} dt \\ &= (\mu - 2K)^2 \phi\left(\frac{\mu - 2K}{\sigma}\right) + \frac{2(\mu - 2K)\sigma}{\sqrt{2\pi}} e^{-(\mu-2K)^2/2\sigma^2} \\ &\quad - \frac{(\mu - 2K)\sigma}{\sqrt{2\pi}} e^{-(\mu-2K)^2/2\sigma^2} + \sigma^2 \phi\left(\frac{\mu - 2K}{\sigma}\right) \\ &= ((\mu - 2K)^2 + \sigma^2) \phi\left(\frac{\mu - 2K}{\sigma}\right) + \frac{(\mu - 2K)\sigma}{\sqrt{2\pi}} e^{-(\mu-2K)^2/2\sigma^2}. \end{aligned}$$

□

Now the following result is immediately established.

Corollary 3.3 *Let X be a normally distributed random variable with mean μ and variance σ^2 . If K is a constant then*

$$\begin{aligned} \text{Var}((X - K)^+) &= ((\mu - 2K)^2 + \sigma^2) \phi\left(\frac{\mu - 2K}{\sigma}\right) + \frac{(\mu - 2K)\sigma}{\sqrt{2\pi}} e^{-(\mu-2K)^2/2\sigma^2} \\ &\quad - \left(\frac{\sigma}{\sqrt{2\pi}} e^{-(\mu-K)^2/2\sigma^2} + (\mu - K) \phi\left(\frac{\mu - K}{\sigma}\right) \right)^2. \end{aligned} \tag{3.18}$$

Now we must establish a similar result for lognormally distributed random variables.

Lemma 3.3 Let X be a lognormally distributed random variable with parameters μ and σ^2 . If $K > 0$ is a constant then

$$\begin{aligned} \mathbb{E} [((X - K)^+)^2] \\ = e^{2(\mu+\sigma^2)}\phi(w + 2\sigma) - 2Ke^{\mu+\sigma^2/2}\phi(w + \sigma) + K^2\phi(w) \end{aligned} \quad (3.19)$$

where $w = (\mu - \ln K)/\sigma$.

Proof.

$$\begin{aligned} \mathbb{E} [((X - K)^+)^2] &= \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty ((x - K)^+)^2 \frac{1}{x} e^{-(\ln x - \mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_K^\infty (x - K)^2 \frac{1}{x} e^{-(\ln x - \mu)^2/2\sigma^2} dx \end{aligned}$$

The expression $\sigma z = \ln x - \mu$ will be substituted into the last integral to yield:

$$\begin{aligned} \mathbb{E} [((X - K)^+)^2] &= \frac{1}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^\infty (e^{\sigma z + \mu} - K)^2 e^{-z^2/2} dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^\infty (e^{\sigma z + \mu} - K)^2 e^{-z^2/2} dz \\ &= \frac{e^{2(\mu+\sigma^2)}}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^\infty e^{-(z-2\sigma)^2/2} dz \\ &\quad - \frac{2Ke^{\mu+\sigma^2/2}}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^\infty e^{-(z-\sigma)^2/2} dz \\ &\quad + \frac{K^2}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^\infty e^{-(z-\sigma)^2/2} dz \\ &= e^{2(\mu+\sigma^2)}\phi\left(\frac{\mu - \ln K}{\sigma} + 2\sigma\right) + K^2\phi\left(\frac{\mu - \ln K}{\sigma}\right) \\ &\quad + 2Ke^{\mu+\sigma^2/2}\phi\left(\frac{\mu - \ln K}{\sigma} + \sigma\right). \end{aligned}$$

□

At last we have an expression for the variance of the positive part of the difference of a lognormal random variable and a constant.

Corollary 3.4 Let X be a lognormally distributed random variable with parameters μ and σ^2 . If $K > 0$ is a constant then

$$\begin{aligned} \text{Var}((X - K)^+) &= e^{2(\mu+\sigma^2)}\phi(w + 2\sigma) - 2Ke^{\mu+\sigma^2/2}\phi(w + \sigma) + K^2\phi(w) \\ &\quad - \left(e^{\mu+\sigma^2/2}\phi(w + \sigma) - K\phi(w) \right)^2 \end{aligned} \quad (3.20)$$

where $w = (\mu - \ln K)/\sigma$.

3.9 Exercises

- (1) Random variable X is continuously uniformly distributed in the interval $[-4, 1]$. Find $P(X \geq 0)$.
- (2) Random variable X is continuously distributed on the interval $(1, \infty)$, with probability distribution function

$$f(x) = \begin{cases} \frac{c}{x^3} & \text{if } 1 \leq x \\ 0 & \text{otherwise.} \end{cases}$$

Determine the value of c .

- (3) Show using properties of the definite integral that for a continuous random variable X with probability distribution function $f(x)$,

$$P(X \geq a) = 1 - P(X < a).$$

- (4) A random variable X has a continuous Cauchy distribution with probability density function

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

Show that the mean of this random variable does not exist.

- (5) If X is a continuous random variable with probability density function $f(x)$, show that $E[aX + b] = aE[X] + b$ where $a, b \in \mathbb{R}$.
- (6) Prove Theorem 3.1.
- (7) Prove Theorem 3.2.
- (8) Prove Theorem 3.3.
- (9) Prove Theorem 3.4.
- (10) Prove Theorem 3.5.
- (11) Find the expected value and variance of the continuous random variable X whose probability distribution function is given by

$$f(x) = \begin{cases} \frac{2}{5}|x| & \text{if } -1 \leq x \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

- (12) Show that if m and n are integers then $n - m$ is even if and only if $n + m$ is even.

- (13) Evaluate the following indefinite integral.

$$\int \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx$$

- (14) Use the technique of integration by parts to evaluate

$$\int \frac{x^2}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx.$$

- (15) Evaluate the following limit with $k > 0$ and $t > 0$.

$$\lim_{M \rightarrow \infty} 2ktM e^{-M^2/4kt}$$

- (16) Using a computer algebra system, graphing calculator, or some numerical method evaluate the following probabilities for a standard normal random variable:
- (a) $P(-1 < X < 1)$
 - (b) $P(-2 < X < 2)$
 - (c) $P(-3 < X < 3)$
 - (d) $P(1 < X < 3)$
- (17) Prove Theorem 3.6.
- (18) The annual rainfall amount in a geographical area is normally distributed with a mean of 14 inches and a standard deviation of 3.2 inches. What is the probability that the sum of the annual rainfalls in two consecutive years will exceed 30 inches?
- (19) The ratio of selling prices of a security on consecutive days is a log-normally distributed random variable with parameters $\mu = 0.01$ and $\sigma = 0.05$. What is the probability of a one-day increase in the selling price? What is the probability of a one-day decrease in the selling price? What is the probability of a four-day decrease in the selling price of the security?
- (20) Let X be a uniformly distributed continuous random variable on the interval $[a, b]$. If K is a constant, find an expression for $E[(X - K)^+]$.
- (21) Show that for the standard normal random variable $1 - \phi(x) = \phi(-x)$.
- (22) Fill in the details of the proof of Corollary 3.3.
- (23) Fill in the details of the proof of Corollary 3.4.

Chapter 4

The Arbitrage Theorem

The concept known as **arbitrage** is subtle and can seem counter-intuitive. In its basic form arbitrage exists whenever two financial instruments are mis-priced relative to one another. Due to the mis-pricing, it becomes possible to make a financial gain. For example suppose bank A issues loans at a 5% interest rate and bank B offers a savings account which pays 6% interest. A person could take out a loan from bank A and place the loan into savings with bank B. When it becomes time to repay bank A for the loan, the person closes the savings account with bank B, repays the principal and interest and still has 1% of the loaned amount as profit. The assumption that financial markets are efficient prevents such obvious arbitrage opportunities from being commonplace. When arbitrage opportunities arise, investors wanting to make a profit flock to the mis-priced instruments and the financial market reacts by correcting the pricing of the instruments.

We will see that financial instruments such as options, bonds, and stocks must be priced so as to be “arbitrage free.” It is this absence of arbitrage which forms the basis of the derivation of the Black-Scholes equation found in Chapter 6.

4.1 The Concept of Arbitrage

As mentioned previously one way to think about arbitrage is as the situation which arises when two financial instruments are mis-priced relative to one another. In this section we will refine and make precise our definition. Suppose there is a set of possible outcomes for some experiment and that wagers can be placed on those outcomes. The **Arbitrage Theorem** states that either the probabilities of the outcomes are such that all bets are fair, or there is a betting scheme which produces a positive gain independent

of the outcome of the experiment. Since we are adopting the language of wagering, a simple example will illustrate the Arbitrage Theorem.

To simplify the process of placing and paying off bets, **odds** are used rather than probabilities, though the two concepts are related by a simple formula. Suppose the “odds” of a particular sporting events outcome are quoted as “2 : 1 against.” We can think of this as implying that there are three outcomes to the experiment and in two of them the desired outcome does not occur and in one it does. Thus the probability of the desired outcome arising is $1/3$. In general then if the “odds against” a particular outcome are $n : 1$ then the probability of the outcome is $1/(n + 1)$. Odds simplify the paying off of bets in the following way. If the odds against an outcome are $n : 1$ then a unit bet will pay n units on a win and otherwise the unit bet is lost. The payoffs scale multiplicatively for non-unit bets. Odds against the occurrence of an event can also be expressed in the form $n : m$ where n and m are natural numbers. In this case the probability of the event occurring is $m/(n + m)$. Odds of the form $n : m$ are equivalent to the odds $\frac{n}{m} : 1$.

Now consider this example. Suppose the odds against player A defeating player B in a tennis match are $\frac{3}{2} : 1$ and the odds against player B defeating player A are $\frac{3}{7} : 1$. Converting these to probabilities we see that player A defeats player B with probability 0.4 while player B defeats player A with probability 0.7. There is obviously something wrong with these probabilities since they should add to one, but do not. According to the Arbitrage Theorem there is a betting strategy which generates a positive gain regardless of the outcome of the match. Now suppose we wager -1 on player A and -2 on player B. If player A wins we will lose $3/2$ on the first bet and gain 2 on the second, producing a net gain of $1/2$. If player B wins we will gain 1 on the first bet and lose $6/7$ on the second, yielding a positive payoff of $1/7$. This example is hardly transparent so study of the Arbitrage Theorem would be beneficial in avoiding arbitrage opportunities in more complex financial situations. Before a statement and proof of the Arbitrage Theorem is presented a preliminary result from linear programming is needed.

4.2 Duality Theorem of Linear Programming

In this section we will derive and discuss a result from which the Arbitrage Theorem easily follows. The **Duality Theorem** is a familiar result to

people having studied linear programming and operations research. We give a brief introduction to it here following the style of [Strang (1986)].

A classic problem of linear programming is that of minimizing the quantity $\mathbf{c} \cdot \mathbf{x}$ subject to \mathbf{x} being feasible. The expression $\mathbf{x} \in \mathbb{R}^{n \times 1}$ is a column vector of n components. The vector \mathbf{x} is **feasible** by definition if $x_i \geq 0$ for $i = 1, 2, \dots, n$ and \mathbf{x} satisfies the set of constraint equations $A\mathbf{x} = \mathbf{b}$. A is an $m \times n$ matrix while $\mathbf{b} \in \mathbb{R}^{m \times 1}$ is another column vector. The expression $\mathbf{c} \in \mathbb{R}^{1 \times n}$ is a row vector of n components. The inner product

$$\mathbf{c} \cdot \mathbf{x} = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

is referred to as the **cost function**.

If m and n are not too large this optimization task is easily accomplished. For example suppose we try to minimize $5x_1 + 4x_2 + 8x_3$ subject to the constraints $x_1 + x_2 + x_3 = 1$ and \mathbf{x} is feasible. The constraints require the solution to the optimization problem to lie in a subset of a plane in the positive orthant of \mathbb{R}^3 . See Fig. 4.1. If the minimum of the cost function is the as yet unknown value k , then $5x_1 + 4x_2 + 8x_3 = k$ defines yet another plane. The smallest value of k for which the level set of the cost function intersects the constrained set of points will be the minimum of the cost function. Thus the minimum of the cost function is 4 as can be seen in Fig. 4.2. The cost function is minimized at the point $(x_1, x_2, x_3) = (0, 1, 0)$.

The constraint in the previous example specified strict equality. Some optimization problems may include **inequality constraints**. If the constraints of the previous example had been $x_1 + x_2 + x_3 \leq 1$ and \mathbf{x} feasible, then the set of points where the minimum must be found would resemble a tetrahedron with vertices at $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$. Feasible sets of points in these types of optimization problems are always **convex sets**. A set is convex if for every pair of points A and B contained in the set, the line segment connecting them also lies completely in the set. Inequality constraints can be converted to equality constraints by introducing **slack variables**. Thus

$$x_1 + x_2 + x_3 \leq 1 \quad \text{becomes} \quad x_1 + x_2 + x_3 + x_4 = 1,$$

where $x_4 \geq 0$ and “takes up the slack” to produce equality. No modification of the cost function is necessary, although it is always possible to think of the cost function in the form now of $5x_1 + 4x_2 + 8x_3 + 0x_4$. The cost function is minimized when the plane of the cost function first intersects the tetrahedron of feasible points, in other words at $(0, 0, 0)$.

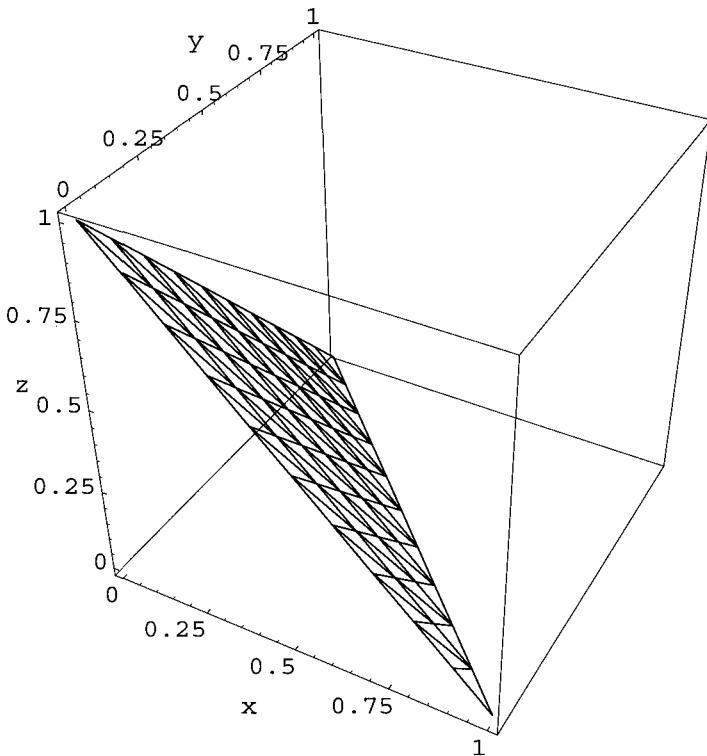


Fig. 4.1 The set of points in \mathbf{R}^3 where the cost function's minimum will occur.

4.2.1 Dual Problems

In the remainder of this chapter we will adopt a convenient extension of the notion of inequality. We will say that vector \mathbf{u} is less than (less than or equal to) vector \mathbf{v} if the vectors are both elements of $\mathbb{R}^{1 \times k}$ and $u_i < v_i$ ($u_i \leq v_i$) for $i = 1, 2, \dots, k$. Similarly we will say that vector \mathbf{u} is greater than (greater than or equal to) vector \mathbf{v} if $u_i > v_i$ ($u_i \geq v_i$) for $i = 1, 2, \dots, k$. These inequalities will be denoted as appropriate $\mathbf{u} < \mathbf{v}$, $\mathbf{u} \leq \mathbf{v}$, $\mathbf{u} > \mathbf{v}$, or $\mathbf{u} \geq \mathbf{v}$. This new notation will simplify the statements of the theorems in this section.

In mathematics we frequently benefit from the ability to solve one problem by means of finding the solution to a related, but simpler, problem. This is certainly true of the linear programming problems which are the subject of this chapter. For every linear programming problem of the type discussed

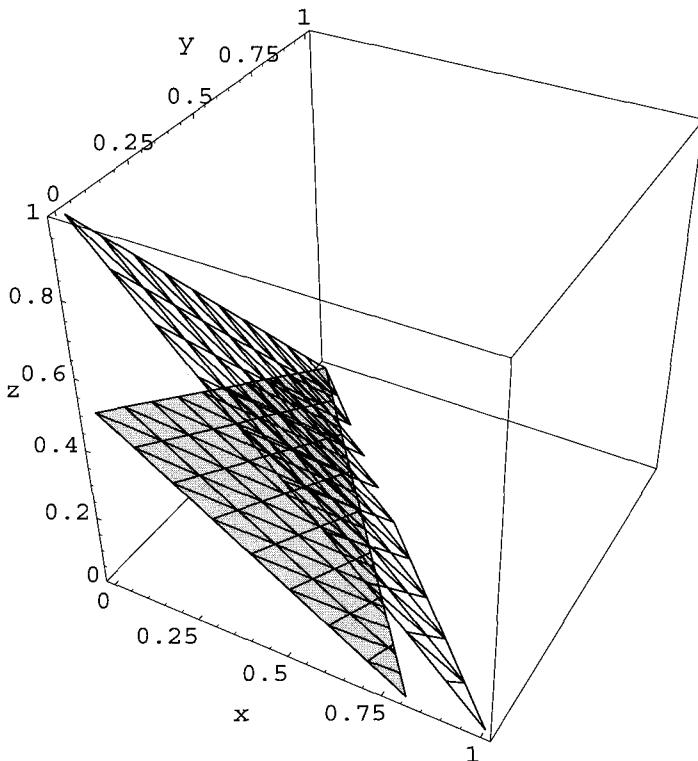


Fig. 4.2 The minimum of the cost function is the first level set of the cost function to intersect the simplex of constrained solution points.

above, there is an associated problem known as its **dual**. Henceforth the original problem will be known as the **primal**. These paired optimization problems are related in the following ways.

Primal: Minimize $\mathbf{c} \cdot \mathbf{x}$ subject to $A\mathbf{x} = \mathbf{b}$ and \mathbf{x} is feasible.

Dual: Maximize $\mathbf{y} \cdot \mathbf{b}$ subject to $\mathbf{y}A \leq \mathbf{c}$.

We should note that:

- (1) the unknown of the dual is the row vector $\mathbf{y} \in \mathbb{R}^{1 \times n}$ with n components,
- (2) the vector \mathbf{b} moves from the constraint of the primal to the cost function of the dual,

- (3) the vector \mathbf{c} moves from the cost of the primal to the constraint of the dual,
- (4) the constraints of the dual are inequalities and there are n of them,
- (5) there is no feasibility constraint (*i.e.* no $\mathbf{y} \geq 0$) in the dual.

Given a primal or a dual problem it is a routine matter to construct its partner. Perhaps it is no surprise then that the solutions to the primal and dual are related.

Theorem 4.1 (Weak Duality Theorem) *If \mathbf{x} and \mathbf{y} are the unknowns of the primal and dual problems respectively, then $\mathbf{y} \cdot \mathbf{b} \leq \mathbf{c} \cdot \mathbf{x}$. If $\mathbf{y} \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{x}$ then these vectors are optimal for their respective problems.*

Proof. Solutions to the primal and the dual problems must satisfy the constraints $A\mathbf{x} = \mathbf{b}$, $x_i \geq 0$ for $i = 1, \dots, n$ and $\mathbf{y}A \leq \mathbf{c}$. We multiply the constraint in the primal by \mathbf{y} and the constraint in the dual by \mathbf{x} to obtain

$$\mathbf{y}A\mathbf{x} = \mathbf{y} \cdot \mathbf{b} \quad \text{and} \quad \mathbf{y}A\mathbf{x} \leq \mathbf{c} \cdot \mathbf{x}.$$

The non-negativity of the components of \mathbf{x} implied by the feasibility assumption on \mathbf{x} preserve the inequality in the latter case. Therefore we have $\mathbf{y} \cdot \mathbf{b} = \mathbf{y}A\mathbf{x} \leq \mathbf{c} \cdot \mathbf{x}$ or naturally $\mathbf{y} \cdot \mathbf{b} \leq \mathbf{c} \cdot \mathbf{x}$. If $\mathbf{y} \cdot \mathbf{b} = \mathbf{c} \cdot \mathbf{x}$ then \mathbf{y} and \mathbf{x} must be optimal since no \mathbf{y} can make $\mathbf{y} \cdot \mathbf{b}$ larger than $\mathbf{c} \cdot \mathbf{x}$ and no \mathbf{x} can make $\mathbf{c} \cdot \mathbf{x}$ smaller than $\mathbf{y} \cdot \mathbf{b}$. \square

Example 4.1 Consider the paired primal and dual problems:

Primal: Minimize $4x_1 + 3x_2$ subject to $x_1 + x_2 = 2$ and $x_1, x_2 \geq 0$.

Dual: Maximize y_2 subject to $y_2 \leq 3$ and $y_2 \leq 4$.

While both problems may seem simple, the dual is trivial. The maximum value of y_2 subject to the constraints must be $y_2 = 3$. According to the Weak Duality Theorem then the minimum of the cost function of the primal must be at least 3. Applying the level set argument as before, the smallest value of k for which $4x_1 + 3x_2 = k$ intersects the set of constrained, feasible points for the primal is $k = 6$. See Fig. 4.3.

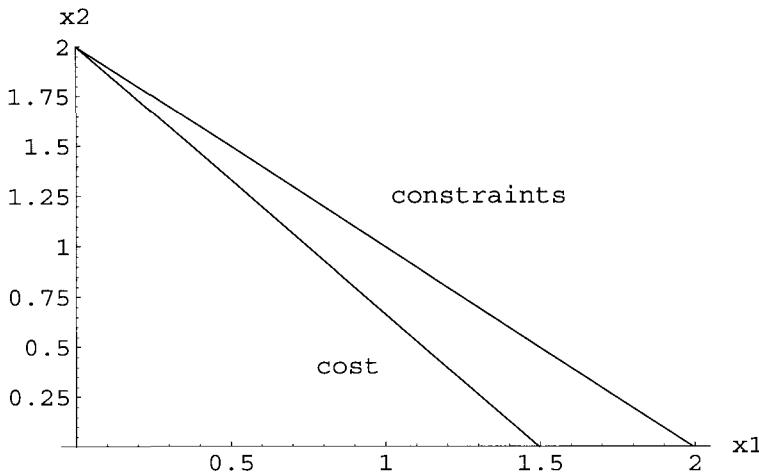


Fig. 4.3 The minimum of the cost function $4x_1 + 3x_2$ subject to the constraints $x_1 + x_2 = 2$ and $x_1, x_2 \geq 0$, occurs at the point with coordinates $(x_1, x_2) = (0, 2)$.

As a consequence of the Weak Duality Theorem we can determine that when \mathbf{x} and \mathbf{y} are optimal for their respective problems then

$$\begin{aligned} \mathbf{y} \cdot \mathbf{b} &= \mathbf{y} \cdot \mathbf{Ax} = \mathbf{c} \cdot \mathbf{x} \\ \mathbf{c} \cdot \mathbf{x} - \mathbf{y} \cdot \mathbf{Ax} &= 0 \\ \mathbf{c} \cdot \mathbf{x} - (\mathbf{yA}) \cdot \mathbf{x} &= 0 \\ (\mathbf{c} - \mathbf{yA}) \cdot \mathbf{x} &= 0. \end{aligned} \tag{4.1}$$

Since \mathbf{x} is feasible then each component of that vector is non-negative. Likewise by assumption in the dual $\mathbf{yA} \leq \mathbf{c}$ which implies that each component of the vector $\mathbf{c} - \mathbf{yA}$ is non-negative. Using Eq. (4.1) we can conclude that vector \mathbf{x} must be zero in every component for which vector $\mathbf{c} - \mathbf{yA}$ is positive and vice versa. Thus we have proved the following theorem.

Theorem 4.2 *Optimality in the primal and dual problems requires either $x_j = 0$ or $(\mathbf{yA})_j = c_j$ for $j = 1, \dots, n$.*

Example 4.2 We will use the previous theorem to find the optimal value of the following linear programming problem.

Primal: Minimize $\mathbf{c} \cdot \mathbf{x} = x_1 + 2x_2 + 7x_3 + 3x_4$ subject to $x_i \geq 0$ for

$i = 1, 2, 3, 4$ and

$$\begin{bmatrix} 1 & 1 & -1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \quad (4.2)$$

Readers unfamiliar with matrix notation may wish to re-write Eq. (4.2) as a system of two linear equations with four unknowns.

Since the set of points described by the constraints exists in four-dimensional space, it will be more difficult to use geometrical and graphical thinking to analyze this problem. Fortunately we can also analyze this problem's associated dual problem.

Dual: Maximize $\mathbf{y} \cdot \mathbf{b} = 5y_1 + 3y_2$ subject to

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix} \leq \begin{bmatrix} 1 & 2 & 7 & 3 \end{bmatrix} \quad (4.3)$$

The space of points on which the dual is to be optimized exists in two-dimensional space and thus is easily pictured. The constraints of the dual can be thought of as a system of inequalities.

$$y_1 - 2y_2 \leq 1$$

$$y_1 \leq 2$$

$$-y_1 + y_2 \leq 7$$

$$y_2 \leq 3$$

This solution to this set of inequalities can be pictured as the shaded region shown in Fig. 4.4. The maximum of the cost function for the dual occurs at the point with coordinates $(y_1, y_2) = (2, 3)$. Thus the maximum value is 19 which will also be the minimum value of the primal problem's cost function. At the optimal point for the dual problem, strict inequality is present in the first and third constraints since

$$2 - 2(3) = -4 < 3$$

$$-2 + 3 = 1 < 7.$$

By the previous theorem then the first and third components of \mathbf{x} in the primal problem must be zero. Therefore the primal can be recast as

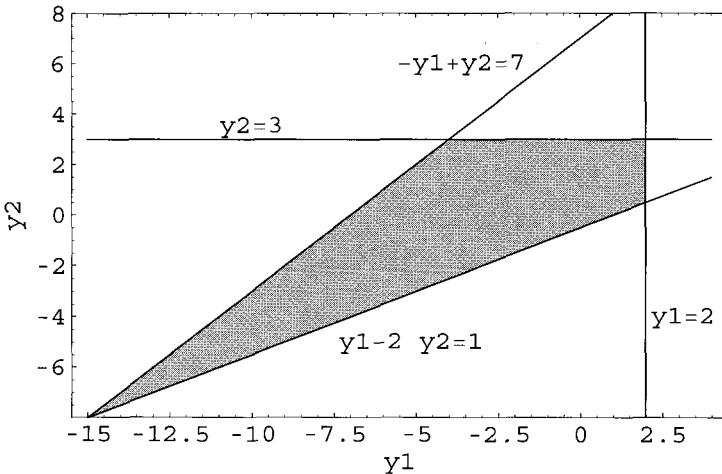


Fig. 4.4 The shaded region in the plot denotes the constrained set of points from Eq. (4.3) for the dual problem.

Primal: Minimize $2x_2 + 3x_4$ subject to $x_i \geq 0$ for $i = 2, 4$ and

$$\begin{bmatrix} 1 & 1 & -1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \quad (4.4)$$

By inspection we must have $x_2 = 5$ and thus $x_4 = 3$. Consequently the minimum of the cost function for the primal is seen to be 19 and it occurs at the point with coordinates $(x_1, x_2, x_3, x_4) = (0, 5, 0, 3)$.

We have assumed that both the primal and dual have optimal solutions. This is in fact the conclusion of the Duality Theorem, which is stated below. The proof of this theorem may be found in [Gale *et al.* (1951)].

Theorem 4.3 (Duality Theorem) *If there is an optimal \mathbf{x} in the primal, then there is an optimal \mathbf{y} in the dual and the minimum of $\mathbf{c} \cdot \mathbf{x}$ equals the maximum of $\mathbf{y} \cdot \mathbf{b}$.*

Before stating and proving the Arbitrage Theorem in the next section, a few remarks concerning the Duality Theorem are in order. First, the dual problem has a dual – namely the primal. Thus the Duality Theorem could be restated in the symmetric form: if either the dual or the primal has an optimal solution vector then so does the other and $\mathbf{c} \cdot \mathbf{x} = \mathbf{y} \cdot \mathbf{b}$. Second,

if the primal has an inequality constraint $A\mathbf{x} \geq \mathbf{b}$ in place of the equality constraint $A\mathbf{x} = \mathbf{b}$, then the dual possesses the extra constraint $\mathbf{y} \geq 0$. To see this imagine introducing the slack variables \mathbf{x}_s in the primal. The primal in matrix form now resembles

$$[A - I] \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_s \end{bmatrix} = \mathbf{b}$$

while the dual becomes

$$\mathbf{y} [A - I] \leq [\mathbf{c} \ 0].$$

This implies the two vector inequalities $\mathbf{y}A \leq \mathbf{c}$ and $-\mathbf{y} \leq 0$. This last inequality is equivalent to $\mathbf{y} \geq 0$. In conclusion the symmetric form the primal/dual pair can be stated as follows.

Primal: Minimize $\mathbf{c} \cdot \mathbf{x}$ subject to $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq 0$.

Dual: Maximize $\mathbf{y} \cdot \mathbf{b}$ subject to $\mathbf{y}A \leq \mathbf{c}$ and $\mathbf{y} \geq 0$.

4.3 The Fundamental Theorem of Finance

Consider an experiment with m possible outcomes numbered 1 through m . Suppose we can place n wagers (numbered 1 through n) on the outcomes of the experiment. Let r_{ij} be the return for a unit bet on wager i when the outcome of the experiment is j . The vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is called a **betting strategy**. Component x_i is the amount placed on wager i . The composite return from a betting strategy is then $\sum_{i=1}^n x_i r_{ij}$. The Arbitrage Theorem states that the probabilities of the m outcomes of the experiment are such that the expected value of the payoff is zero, or there exists a betting strategy for which the payoff is positive regardless of the outcome of the experiment.

Theorem 4.4 (Arbitrage Theorem) *Exactly one of the following is true: either*

- (1) *there is a vector of probabilities $\mathbf{y} = (y_1, y_2, \dots, y_m)$ for which*

$$\sum_{j=1}^m y_j r_{ij} = 0, \quad \text{for } i = 1, 2, \dots, n, \text{ or}$$

(2) there is a betting strategy $\mathbf{x} = (x_1, x_2, \dots, x_n)$ for which

$$\sum_{i=1}^n x_i r_{ij} > 0, \quad \text{for } j = 1, 2, \dots, m.$$

Proof. Let vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ be the betting strategy. We will append another element x_{n+1} to this vector. This last element represents the amount of the positive payoff. We would like to maximize x_{n+1} under the constraint that $\sum_{i=1}^n x_i r_{ij} \geq x_{n+1}$ for $j = 1, 2, \dots, m$. This is equivalent to $\sum_{i=1}^n x_i (-r_{ij}) + x_{n+1} \leq 0$ for $j = 1, 2, \dots, m$. Let the row vector \mathbf{c} be the zero vector of m components and let the column vector $\mathbf{b} = (0, \dots, 0, 1)$ have $n + 1$ components. Then the maximization problem outlined above can be stated in matrix form as “Maximize $\mathbf{x} \cdot \mathbf{b}$ subject to $\mathbf{x}A \leq \mathbf{c}$,” where matrix A has the form

$$\begin{bmatrix} -r_{11} & -r_{12} & \cdots & -r_{1m} \\ -r_{21} & -r_{22} & \cdots & -r_{2m} \\ \vdots & \vdots & & \vdots \\ -r_{n1} & -r_{n2} & \cdots & -r_{nm} \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

This is the dual of the primal problem “Minimize $\mathbf{c} \cdot \mathbf{y}$ subject to $A\mathbf{y} = \mathbf{b}$ and $\mathbf{y} \geq 0$.” The unknown \mathbf{y} is a column vector of m components. Since \mathbf{c} is the zero vector the cost function to be minimized is really just the constant 0. Therefore in matrix form the primal can be thought of as minimizing 0 subject to $\mathbf{y} \geq 0$ and

$$\begin{bmatrix} -r_{11} & -r_{12} & \cdots & -r_{1m} \\ -r_{21} & -r_{22} & \cdots & -r_{2m} \\ \vdots & \vdots & & \vdots \\ -r_{n1} & -r_{n2} & \cdots & -r_{nm} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

This matrix equation is equivalent to the system of equations, $\sum_{j=1}^m y_j r_{ij} = 0$ for $i = 1, 2, \dots, n$; $\sum_{j=1}^m y_j = 1$, and $y_j \geq 0$ for $j = 1, 2, \dots, m$. The primal problem is feasible with a minimum value of zero if and only if \mathbf{y} is a probability vector for which all bets have an expected return of zero. Suppose the primal problem is feasible, the dual problem is also feasible since $x_i = 0$ for $i = 1, 2, \dots, n + 1$ satisfies the inequality constraints of the dual. According to the Duality Theorem then the maximum of the dual problem is zero which means no guaranteed profit is possible. In other

words if (1) is true then (2) is false. Now suppose the primal problem is not feasible. According to the Duality Theorem the dual problem has no optimal solution. Therefore zero is not the maximum payoff, thus there is a betting strategy which produces a positive payoff. Thus we see that if (1) is false then (2) is true. \square

4.4 Exercises

- (1) Suppose the odds against the three possible outcomes of an experiment are as given in the table below.

Outcome	Odds
A	2:1
B	3:1
C	1:1

Find a betting strategy which produces a positive net profit regardless of the outcome of the experiment.

- (2) Sketch the region in the plane which satisfies the following inequalities:

$$x_1 + 3x_2 \geq 6$$

$$3x_1 + x_2 \geq 6$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

Is the region convex? What are the coordinates of the points at the corners of the region?

- (3) Minimize the cost function $x_1 + x_2$ on the region described in exercise 2.
- (4) Find inequality constraints which will describe the rectangular region with corners at $(0, 0)$, $(2, 0)$, $(2, 3)$, and $(0, 3)$.
- (5) Introduce slack variables in the inequality constraints found in exercise 4 to produce equality constraints $A\mathbf{x} = \mathbf{b}$. What are A and \mathbf{b} ?
- (6) Minimize the cost function $x_1 + x_2 + 2x_3$ on the set where $x_1 + 2x_2 + 3x_3 = 15$ and $x_i \geq 0$ for $i = 1, 2, 3$.
- (7) Minimize the cost function $7x_2 + 9x_3$ subject to the constraints $x_1 + x_2 + x_3 = 5$ and $x_2 + x_3 + 2x_4 = 1$ with $x_i \geq 0$.
- (8) Write down the dual problem of the following primal problem: Minimize $x_1 + x_2 + x_3$ subject to $2x_1 + x_2 = 4$ and $x_3 \leq 6$ with $x_i \geq 0$.

- (9) Find the solutions to the primal and dual problems of exercise 8.
- (10) Write down the primal companion problem to the dual problem: Maximize $2y_1 + 4y_3$ subject to $y_1 - y_2 \leq 1$ and $y_2 + 2y_3 \leq -1$.
- (11) Find the solutions to the primal and dual problems of exercise 10.

This page is intentionally left blank

Chapter 5

Random Walks and Brownian Motion

In this chapter we will introduce and explain some of the concepts surrounding the probabilistic models used to capture the behavior of stock, security, option, and index prices. Topics covered here could be expanded into an entire book of their own. For our purposes we will explore just enough of the stochastic calculus to provide some justification for Itô’s Lemma, the main result of this chapter.

5.1 Intuitive Idea of a Random Walk

Earlier in Chapter 3 we analyzed the motion of a particle taking a discrete step along the x -axis during every “tick” of the clock. At first it was revealed that the location of the particle relative to the origin followed a binomial distribution. As the step size was decreased, the probability of the particle lying in a particular interval was seen to obey the normal distribution. In this chapter we will extend this discussion to a more general setting. Imagine a person standing at the origin of the real number line. They will flip a coin. For every time the coin lands on heads, they will take a unit step to the right (the positive direction). For every time the coin lands on tails, they will move a unit step to the left (the negative direction). The evolution of this stochastic discrete dynamical system is called a **random walk**. A plot of the person’s location versus the number of times the coin has been flipped might resemble the graph in Fig. 5.1. Random walks do not require that the probabilities of moving left or right be equal (*i.e.* the coin does not have to be fair). The magnitude of the movement at each stage does not have to a single unit. The person could move left or right depending on the value of a random number chosen from a normal distribution for example. The bias in choosing to move left or right and the possibility of moving

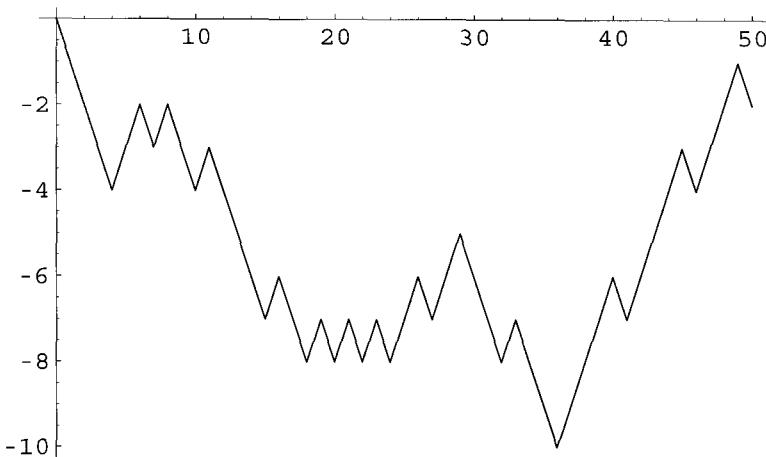


Fig. 5.1 A realization of a random walk on the real number line.

different distances during different steps are generalizations not present in the discussion presented in Chapter 3. In the remainder of this chapter we will apply the concept of the random walk to modeling the movement of the value of a stock or other security.

5.2 First Step Analysis

A simple first application of a random walk is the situation of a person owning a share of stock whose current value is denoted S . At discrete intervals (for example, once per day) the stock's value can increase or decrease by one unit, in other words if its current value on day n is $S(n)$, then tomorrow the value will either be $S(n+1) = S(n) + 1$ or $S(n+1) = S(n) - 1$. To keep the following discussion simple, the probability of increase in value will be $p = 1/2$ and consequently the probability of decrease is the same. This assumption implies the random walk is **symmetric**. If we think of X_n as a random variable which takes on value 1 with probability $p = 1/2$ and value -1 with the probability, we can think of the N^{th} state of the random walk as a partial sum where for $N > 0$,

$$S(N) = S(0) + X_1 + X_2 + \cdots + X_N. \quad (5.1)$$

The random variable X_n is a generalization of the Bernoulli random variable discussed earlier. $S(N)$ can be defined inductively by the formula $S(N) = S(N-1) + X_N$ for $N > 0$ where $S(0)$ is a specified initial state of the walk. If we assume that the random variables X_i and X_j are identically distributed and independent for $i \neq j$ then we see that the transitions between states in the random walk, $S(N) - S(N-1) = X_N$ are independent random variables. If the initial state of the random walk is $S(0)$ and $n \geq 0$ steps are taken, where $k \geq 0$ of the steps are in the positive direction and $n - k \geq 0$ are in the negative direction, then

$$\begin{aligned} P(S(n) = S(0) + k - (n - k)) &= P(S(n) - S(0) = 2k - n) \\ &= \binom{n}{k} \left(\frac{1}{2}\right)^n. \end{aligned}$$

The last expression is the familiar probability formula for a binomial random variable with $p = 1/2$.

If $S(0) \neq 0$ then we can perform a change of variable via $T(n) = S(n) - S(0)$ for $n = 0, 1, \dots$ and define a related random walk where $T(0) = 0$. The transitions in the states of the new random walk are independent, identically distributed random variables just as before. Likewise we see that

$$P(T(n) = 2k - n) = \binom{n}{k} \left(\frac{1}{2}\right)^n. \quad (5.2)$$

Thus without loss of generality we will assume that $S(0) = 0$. Now that we have settled on a starting state for the random walk, what states can be visited in n steps? The next lemma provides the answer and the respective probabilities of reaching these states.

Lemma 5.1 *For the random walk defined in Eq. (5.1) with initial state $S(0) = 0$,*

- (1) $P(S(n) = m) = 0$ if $|m| > n$,
- (2) $P(S(n) = m) = 0$ if $n + m$ is odd,
- (3) $P(S(n) = m) = \binom{n}{(n+m)/2} \left(\frac{1}{2}\right)^n$, otherwise .

Proof.

- (1) According to Eq. (5.1), $-n \leq S(n) \leq n$, thus if $|m| > n$ the partial sum cannot attain this value.
- (2) If we let $m = 2k - n$ in Eq. (5.2) then $n + m = 2k$ which implies $n + m$ is even, contradicting the assumption that $n + m$ is odd.

- (3) We may assume that $n + m$ is even and thus the result is shown by Eq. (5.2). □

In the previous discussion the random walk was free to move in either the positive or negative direction any integer amount as the number of steps increased. Suppose that bounds are placed in the path the particle may follow. Once again, for the sake of simplicity, we will assume the initial state of the random walk is positive, *i.e.* $S(0) > 0$ and we will assume the random walk has a lower boundary of 0. If the state of the random walk reaches the lower boundary in a finite number of steps then the state remains at that boundary value. In the study of random walks, this is known as an **absorbing boundary condition**. In a financial setting we may think of the value of a security as following a random walk (though in reality a much more complicated one than the type we are exploring here). If the value of the security drops to 0, the security becomes worthless to the owner. We must further generalize our discussion of random walks to include the absorbing boundary. There are two questions related to this situation that we wish to explore.

- (1) What is the probability that the state of the random walk crosses a threshold value of $A > 0$ before it hits the boundary at 0 (and hence remains there)?
- (2) What is the expected value of the number of steps which will elapse before the state of the random variable first crosses the $A > 0$ threshold?

Figure 5.2 shows a random walk of a stock whose value was initially 10. This situation has been so simplified and abstracted from the reality of the stock market that it is of little use in making investment decisions, but serves as an instructive example of the use of the concept of a random walk. The answers to the two primary questions will be found following the lines of reasoning laid out in [Kijima (2003)] and [Steele (2001)].

Since the boundary at 0 is absorbing we must keep track of the smallest value which $S(n)$ takes on. Thus we define $m_n = \min\{S(k) : 0 \leq k \leq n\}$. If i is the smallest non-negative integer such that $m_i = 0$ (and hence that $S(i) = 0$) then by the absorbing boundary condition $S(k) = 0$ for all $k \geq i$. Our attention is now focused on developing an understanding of the following conditional probability:

$$\mathrm{P}(S(n) = A \wedge m_n > 0 \mid S(0) = i).$$

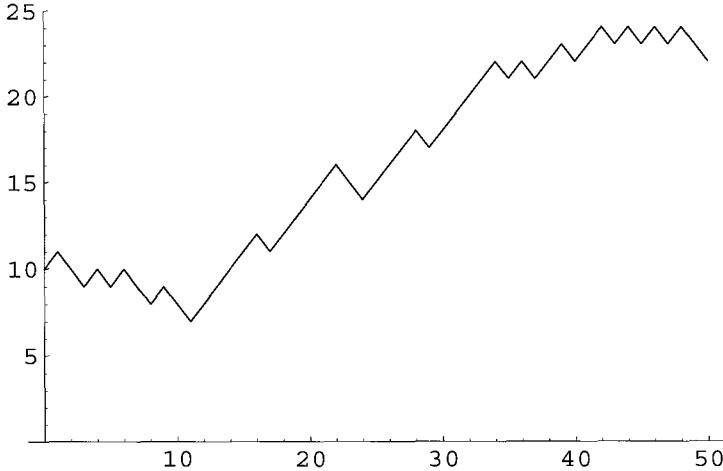


Fig. 5.2 A realization of a simple random walk attempting to capture the changes in the price of a stock.

We will assume that $A > 0$ and $i > 0$. This probability should depend on the three parameters A , i , and n . A formula for this conditional probability is derived in the proof of the following lemma.

Lemma 5.2 *Suppose a random walk has the form described in Eq. (5.1) in which the X_i for $i = 1, 2, \dots$ are independent, identically distributed random variables taking on the values ± 1 with probability $p = 1/2$. Suppose further that the boundary at 0 is absorbing, then if $A, i > 0$,*

$$\begin{aligned} P(S(n) = A \wedge m_n > 0 | S(0) = i) \\ = & \left[\binom{n}{(n+A-i)/2} - \binom{n}{(n-A-i)/2} \right] \left(\frac{1}{2}\right)^n. \end{aligned} \quad (5.3)$$

Proof. In order to prove the lemma we will start by considering a random walk with no boundary, that is, the random variable $S(n)$ has an initial state of $S(0) = i > 0$ and $S(k)$ is allowed to wander into negative territory (and back) arbitrarily. In this situation

$$\begin{aligned} P(S(n) = A | S(0) = i) \\ = P(S(n) = A \wedge m_n > 0 | S(0) = i) + P(S(n) = A \wedge m_n \leq 0 | S(0) = i) \end{aligned}$$

by the Addition Rule (Theorem 2.1). By the spatial homogeneity of the

random walk

$$\begin{aligned} P(S(n) = A \mid S(0) = i) &= P(S(n) = A - i \mid S(0) = 0) \\ &= \binom{n}{(n + A - i)/2} \left(\frac{1}{2}\right)^n \end{aligned}$$

so long as $|A - i| \leq n$ and $n + A - i$ is even (Lemma 5.1). Now we must make use of the symmetry property of the un-restricted random walk.

$$P(S(n) = j \mid S(0) = 0) = P(S(n) = -j \mid S(0) = 0)$$

As seen in Fig. 5.3 any realization of a random walk which intersects the n -axis at $n = i$ can be reflected about the axis. The probabilities associated

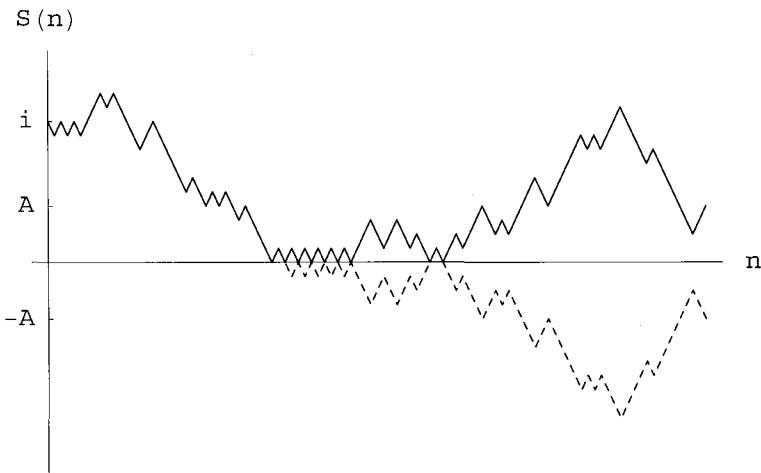


Fig. 5.3 The probabilities that a symmetric random walk will follow either the solid or dashed (reflected) path are equal.

with each path are the same. Therefore we have

$$\begin{aligned} P(S(n) = A \wedge m_n \leq 0 \mid S(0) = i) &= P(S(n) = -A \mid S(0) = i) \\ &= P(S(n) = -A - i \mid S(0) = 0) \\ &= \binom{n}{(n - A - i)/2} \left(\frac{1}{2}\right)^n \end{aligned}$$

provided $|A + i| \leq n$ and $n - A - i$ is even. Finally we are able to determine

that

$$\begin{aligned} \mathbb{P}(S(n) = A \wedge m_n > 0 \mid S(0) = i) \\ &= \binom{n}{(n+A-i)/2} \left(\frac{1}{2}\right)^n - \binom{n}{(n-A-i)/2} \left(\frac{1}{2}\right)^n. \end{aligned}$$

which is equivalent to the expression in Eq. (5.3). \square

The previous theorem now provides an answer to the first of the two questions posed.

Example 5.1 For a symmetric random walk with initial state $S(0) = 10$, what is the probability that $S(50) = 16$ and $S(n) > 0$ for $n = 0, 1, \dots, 50$?

Making use of the formula found in Eq. (5.3) we have

$$\begin{aligned} \mathbb{P}(S(50) = 16 \wedge m_{50} > 0 \mid S(0) = 10) &= \left[\binom{50}{28} - \binom{50}{12} \right] 2^{-50} \\ &\approx 0.0787178. \end{aligned}$$

Since the current notation is so bulky, we will define the function

$$f_{A,i}(n) = \left[\binom{n}{(n+A-i)/2} - \binom{n}{(n-A-i)/2} \right] \left(\frac{1}{2}\right)^n$$

and use it from now on.

We will call the first time that the random walk $S(n)$ equals A the **stopping time**, denoted Ω_A . The stopping time is a random variable. In the remainder of this section we will explore the probability that Ω_A takes on the various natural number values given the initial state of the random walk is $S(0) = i > 0$, parameter $A > 0$, and the boundary at 0 is absorbing. Suppose the initial state of the random walk $S(0) = i > A > 0$, then due to the spatial homogeneity of the random walk

$$\mathbb{P}(S(n) = A \wedge m_n \geq A \mid S(0) = i) = \mathbb{P}(S(n) = 0 \wedge m_n \geq 0 \mid S(0) = i - A).$$

$\Omega_0 = n$ if and only if $S(n-1) = 1$, $m_{n-1} > 0$ and $X_n = -1$. Therefore

$$\begin{aligned} \mathbb{P}(\Omega_0 = n \mid S(0) = i - A) \\ &= \mathbb{P}(X_n = -1 \wedge S(n-1) = 1 \wedge m_{n-1} > 0 \mid S(0) = i - A) \\ &= \frac{1}{2} \mathbb{P}(S(n-1) = 1 \wedge m_{n-1} > 0 \mid S(0) = i - A) \\ &= \frac{1}{2} f_{1,(i-A)}(n-1). \end{aligned}$$

Consequently $\mathbb{P}(\Omega_A = n \mid S(0) = i) = \frac{1}{2} f_{1,(i-A)}(n-1)$ as well.

Next we take up the case in which the initial state of the random walk $0 < S(0) = i < A$. Since the stopping time is the *first* time that $S(n) = A$ then we can think of the spatial domain of the random walk as having two absorbing boundaries: the usual one at $S = 0$ and a new one at $S = A$. The analysis of this situation is modeled after the discussion in the first two chapters of [Redner (2001)].

The symbol $p_{i \rightarrow A}$ will denote any path in the discrete interval $[0, A]$ from $i \neq 0$ to A which avoids 0. The symbol $P_{p_{i \rightarrow A}}$ will denote that probability that the random walk starting at $S(0) = i$ follows path $p_{i \rightarrow A}$. Finally the symbol $\mathcal{P}_A(i)$ will denote the probability that a random walk which starts at $S(0) = i$ will achieve state $S = A$ while avoiding the state $S = 0$. By the Addition Rule for probability

$$\mathcal{P}_A(i) = \sum_{p_{i \rightarrow A}} P_{p_{i \rightarrow A}}.$$

Since at each step of the random walk, the increment to the left or right is independent of the previous increments, we can decompose the state $S(0) = i$ into the superposition of the two states $S(-1) = i - 1$ and $S(-1) = i + 1$. Therefore the probability above can be written as

$$\begin{aligned} \mathcal{P}_A(i) &= \sum_{p_{i \rightarrow A}} \left(\frac{1}{2} P_{p_{(i-1) \rightarrow A}} + \frac{1}{2} P_{p_{(i+1) \rightarrow A}} \right) \\ &= \frac{1}{2} \mathcal{P}_A(i-1) + \frac{1}{2} \mathcal{P}_A(i+1) \\ 0 &= \mathcal{P}_A(i-1) - 2\mathcal{P}_A(i) + \mathcal{P}_A(i+1). \end{aligned} \quad (5.4)$$

Equation (5.4) is the discrete difference equation approximation to the second derivative [Burden and Faires (2005)]. It is also sometimes referred to as the discrete **Laplacian equation**. Since we know *a priori* that $\mathcal{P}_A(0) = 0$ and $\mathcal{P}_A(A) = 1$ then we may derive a system of linear equations from Eq. (5.4) whose solution will give us $\mathcal{P}_A(i)$ for $i = 1, 2, \dots, A-1$. In matrix form the system of equations resembles:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathcal{P}_A(0) \\ \mathcal{P}_A(1) \\ \mathcal{P}_A(2) \\ \vdots \\ \mathcal{P}_A(A-1) \\ \mathcal{P}_A(A) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

The reader can verify that the solution to this matrix equation is

$$\begin{bmatrix} \mathcal{P}_A(0) \\ \mathcal{P}_A(1) \\ \vdots \\ \mathcal{P}_A(i) \\ \vdots \\ \mathcal{P}_A(A-1) \\ \mathcal{P}_A(A) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{A} \\ \vdots \\ \frac{i}{A} \\ \vdots \\ \frac{A-1}{A} \\ 1 \end{bmatrix}.$$

By the symmetry of the random walk we also have

$$\begin{bmatrix} \mathcal{P}_0(0) \\ \mathcal{P}_0(1) \\ \vdots \\ \mathcal{P}_0(i) \\ \vdots \\ \mathcal{P}_0(A-1) \\ \mathcal{P}_0(A) \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{A-1}{A} \\ \vdots \\ \frac{A-i}{A} \\ \vdots \\ \frac{1}{A} \\ 0 \end{bmatrix}.$$

The expression $\mathcal{P}_0(i)$ is the probability that a symmetric random walk initially in state $S(0) = i$ achieves the state $S = 0$ while avoiding the state $S = A$. We can summarize these results in the following theorem.

Theorem 5.1 *Suppose a random walk has the form described in Eq. (5.1) in which the X_i for $i = 1, 2, \dots$ are independent, identically distributed random variables taking on the values ± 1 with probability $p = 1/2$. Suppose further that the boundaries at 0 and A are absorbing, then if $0 \leq S(0) = i \leq A$*

- (1) *the probability that the random walk achieves state A without achieving state 0 is $\mathcal{P}_A(i) = i/A$,*
- (2) *the probability that the random walk achieves state 0 without achieving state A is $\mathcal{P}_0(i) = 1 - i/A$.*

Before answering the question as to the expected stopping time for the random walk attaining state A , we will explore the simpler issue of the stopping time of reaching either of the two absorbing boundaries. Let $B = \{0, A\}$ represent the discrete boundary set of the one-dimensional, discrete random walk. To help in the derivation we will define the following quantities:

$\omega_{p_{i \rightarrow B}}$: the exit time of the random walk which starts at $S(0) = i$, where $0 \leq i \leq A$ and which follows path $p_{i \rightarrow B}$.

$\Omega_B(i)$: the expected value of the exit time for a random walk which starts at $S(0) = i$, where $0 \leq i \leq A$.

By the definition of expected value,

$$\Omega_B(i) = \sum_{p_{i \rightarrow B}} P_{p_{i \rightarrow B}} \omega_{p_{i \rightarrow B}}.$$

Certainly it is the case that if the random walk starts on the boundary the expected stopping time is 0. Thus it is true that $\Omega_B(0) = \Omega_B(A) = 0$. There is also a recursive relationship between the stopping times at neighboring starting points in space. The path from $i \rightarrow B$ can be decomposed into paths from $(i-1) \rightarrow B$ or $(i+1) \rightarrow B$ with the addition of a single step. Therefore

$$\begin{aligned} \Omega_B(i) &= \sum_{p_{i \rightarrow B}} \left(\frac{1}{2} P_{p_{(i-1) \rightarrow B}} [\omega_{p_{(i-1) \rightarrow B}} + 1] + \frac{1}{2} P_{p_{(i+1) \rightarrow B}} [\omega_{p_{(i+1) \rightarrow B}} + 1] \right) \\ &= \frac{1}{2} \sum_{p_{i \rightarrow B}} P_{p_{(i-1) \rightarrow B}} \omega_{p_{(i-1) \rightarrow B}} + \frac{1}{2} \sum_{p_{i \rightarrow B}} P_{p_{(i+1) \rightarrow B}} \omega_{p_{(i+1) \rightarrow B}} \\ &\quad + \sum_{p_{i \rightarrow B}} \left(\frac{1}{2} P_{p_{(i-1) \rightarrow B}} + \frac{1}{2} P_{p_{(i+1) \rightarrow B}} \right) \\ &= \frac{1}{2} \Omega_B(i-1) + \frac{1}{2} \Omega_B(i+1) + 1 \\ &\quad - 2 = \Omega_B(i-1) - 2\Omega_B(i) + \Omega_B(i+1). \end{aligned} \tag{5.5}$$

Equation (5.5) is often called the **Poisson equation** for the stopping time. From the Poisson equation we can derive a system of linear equations for the stopping times which are parameterized by the various initial conditions of the random walk. In matrix form the system of equations has the form:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Omega_B(0) \\ \Omega_B(1) \\ \Omega_B(2) \\ \vdots \\ \Omega_B(A-1) \\ \Omega_B(A) \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ -2 \\ \vdots \\ -2 \\ 0 \end{bmatrix}.$$

The matrix generated by the Poisson equation (as well as the earlier matrix generated by the Laplacian equation) are known as **tridiagonal matrices**

due to the fact that all entries in the matrices are zero except for possibly the entries on the diagonal, the first sub-diagonal, and the first super-diagonal. There are many methods for solving tridiagonal linear systems, but we will employ one of the simplest methods, **Gaussian elimination** followed by **backwards substitution**. This technique as well as many others is explained in detail in [Golub and Van Loan (1989)]. All that will be important during this process are the coefficients present in the matrix and on the right-hand side of the equation. Thus will be work with the **augmented matrix** formed by appending the column vector on the right-hand side of the equation to the matrix as an addition column. Thus for the Poisson system we obtain the following.

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array} \right] \quad (5.6)$$

A vertical bar is used to visually set off the augmented column from the original matrix. Gaussian elimination works by adding scalar multiples of rows of the augmented matrix to rows below in such a way as to produce an **upper triangular matrix**, that is, a matrix for which all entries below the main diagonal are zero. If -1 is multiplied by the first row and then that scalar product is added to the second row, then the system in Eq. (5.6) becomes

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array} \right]. \quad (5.7)$$

Notice that the first column of the matrix has now achieved the upper triangular form. Now we proceed to the second column. Multiplying the second row of Eq. (5.7) by $1/2$ and adding to the third row yields an augmented

matrix of the form:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array} \right] . \quad (5.8)$$

The conversion to upper triangular form will be completed by induction (see [Courant and Robbins (1969)]) on the rows. Suppose that Gaussian elimination has been performed on the first i rows of the augmented matrix and the i^{th} and $(i+1)^{th}$ rows have the following form:

$$\left[\begin{array}{cccccc|c} 0 & \cdots & 0 & \frac{-i}{i-1} & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 & 0 & \cdots & 0 \end{array} \right] \begin{array}{l} -i \\ -2 \end{array}$$

The i^{th} row is multiplied by $1 - 1/i$ and added to the $(i+1)^{th}$ row to produce the following.

$$\left[\begin{array}{cccccc|c} 0 & \cdots & 0 & \frac{-i}{i-1} & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \frac{-(i+1)}{i} & 1 & 0 & \cdots & 0 \end{array} \right] \begin{array}{l} -i \\ -(i+1) \end{array}$$

Thus once the Gaussian elimination procedure is completed the augmented matrix has the upper triangular form:

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\frac{A}{A-1} & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array} \right] . \quad (5.9)$$

The Gaussian elimination process is the equivalent of adding multiples of linear equations to one another. If we re-write the upper triangular aug-

mented matrix in the form of an equation we obtain the following.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \cdots 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \cdots 0 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 \cdots 0 & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 \cdots 0 & -\frac{A}{A-1} & 1 \\ 0 & 0 & 0 & 0 \cdots 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Omega_B(0) \\ \Omega_B(1) \\ \Omega_B(2) \\ \vdots \\ \Omega_B(A-1) \\ \Omega_B(A) \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ -3 \\ \vdots \\ -A \\ 0 \end{bmatrix}$$

Now the system is readily solved by the process of backwards substitution. If the value of $\Omega_B(j)$ is known then the value of $\Omega_B(j-1)$ can be determined from the equation

$$\begin{aligned} -\frac{j}{j-1}\Omega_B(j-1) + \Omega_B(j) &= -j \\ \Omega_B(j-1) &= j-1 + \frac{j-1}{j}\Omega_B(j). \end{aligned} \quad (5.10)$$

The reader can understand why this method is called backwards substitution — the components of the solution are found in reverse index order. The boundary condition $\Omega_B(A) = 0$ is used to seed the substitution. Setting $j = A$ in Eq. (5.10) produces $\Omega_B(A-1) = A-1$. The remainder of the unknowns can be found by induction. Suppose $\Omega_B(A-k) = k(A-k)$ is known, then by setting $j = A-k$ in Eq. (5.10) we obtain

$$\Omega_B(A-k-1) = (k+1)(A-k-1). \quad (5.11)$$

Thus in general $\Omega_B(A-j) = j(A-j)$. A more convenient way of expressing the solution is found upon making the substitution $i = A-j$ in Eq. (5.11). The results are summarized in the following theorem.

Theorem 5.2 *Suppose a random walk has the form described in Eq. (5.1) in which the X_i for $i = 1, 2, \dots$ are independent, identically distributed random variables taking on the values ± 1 with probability $p = 1/2$. Suppose further that the boundaries at 0 and A are absorbing, then if $0 \leq S(0) = i \leq A$ the random walk intersects the boundary ($S = 0$ or $S = A$) after a mean number of steps given by the formula*

$$\Omega_B(i) = i(A-i). \quad (5.12)$$

Example 5.2 Suppose a symmetric random walk takes place on the discrete interval $\{0, 1, 2, 3, 4\}$ for which the boundaries at 0 and 4 are absor-

ing. If $S(0) = i$ then the expected value of the stopping time of the random walk is

i	0	1	2	3	4
$\Omega_B(i)$	0	3	4	3	0

Theorem 5.2 yields a formula for the mean number of steps required for a symmetric random walk to reach either boundary point of the interval $[0, A]$ based on the starting point. Now we want to determine a related quantity, the conditional exit time. This is the average number of steps required for a symmetric random walk to reach the boundary at A while avoiding the other boundary at 0. The symbol $\Omega_A(i)$ will denote the conditional exit time of the random walk in the initial state $S(0) = i$ which exits at boundary A while avoiding the boundary at 0. By definition this conditional exit time is

$$\Omega_A(i) = \frac{\sum_{p_i \rightarrow A} P_{p_i \rightarrow A} \omega_{p_i \rightarrow A}}{\sum_{p_i \rightarrow A} P_{p_i \rightarrow A}} = \frac{\sum_{p_i \rightarrow A} P_{p_i \rightarrow A} \omega_{p_i \rightarrow A}}{\mathcal{P}_A(i)}.$$

Note that the sums above are taken over all paths which start at $S(0) = i$ and exit at A and avoid 0. This implies the conditional exit times and path probabilities are related by the equation:

$$\Omega_A(i)\mathcal{P}_A(i) = \sum_{p_i \rightarrow A} P_{p_i \rightarrow A} \omega_{p_i \rightarrow A}. \quad (5.13)$$

Once again the idea of decomposing a path into a first step and the remainder of the step will be used on the right-hand side of Eq. (5.13).

$$\begin{aligned} & \Omega_A(i)\mathcal{P}_A(i) \\ &= \sum_{p_i \rightarrow A} \left(\frac{1}{2} P_{p_{(i-1)} \rightarrow A} [\omega_{p_{(i-1)} \rightarrow A} + 1] + \frac{1}{2} P_{p_{(i+1)} \rightarrow A} [\omega_{p_{(i+1)} \rightarrow A} + 1] \right) \\ &= \frac{1}{2}\mathcal{P}_A(i-1)\Omega_A(i-1) + \frac{1}{2}\mathcal{P}_A(i+1)\Omega_A(i+1) + \mathcal{P}_A(i) \\ -2\mathcal{P}_A(i) &= \mathcal{P}_A(i-1)\Omega_A(i-1) - 2\mathcal{P}_A(i)\Omega_A(i) + \mathcal{P}_A(i+1)\Omega_A(i+1) \end{aligned} \quad (5.14)$$

Equation (5.14) generates a system of linear equations which can be solved by elementary matrix algebra. The solution to this linear system will be the vector of products $\mathcal{P}_A(i)\Omega_A(i)$ for $i = 0, 1, \dots, A$. To save space and simplify the notation we will define $z_i = \mathcal{P}_A(i)\Omega_A(i)$. At the boundary where $i = 0$, $z_0 = 0$ since $\mathcal{P}_A(0) = 0$. At the other boundary, $z_A = 0$ since

$\Omega_A(A) = 0$. If we make use of Theorem 5.1 in order to provide expressions for $\mathcal{P}_A(i)$ we obtain the following tridiagonal linear system.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_{A-1} \\ z_A \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{2}{A} \\ -\frac{4}{A} \\ \vdots \\ -\frac{2(A-1)}{A} \\ 0 \end{bmatrix}. \quad (5.15)$$

Finally we have established the following theorem.

Theorem 5.3 Suppose a random walk has the form described in Eq. (5.1) in which the X_i for $i = 1, 2, \dots$ are independent, identically distributed random variables taking on the values ± 1 with probability $p = 1/2$. Suppose further that the boundary at 0 is absorbing. The random walk will stop the first time that $S(n) = A$. Let $\mathbf{z} = \langle z_0, z_1, \dots, z_A \rangle$ be the solution to the matrix equation in (5.15). Then the expected value of the stopping time is

$$\Omega_A(i) = \frac{i \cdot z_i}{A}, \quad \text{for } i = 1, 2, \dots, A. \quad (5.16)$$

Example 5.3 Suppose a symmetric random walk takes place on the integer-valued number line. Suppose the boundary at 0 is absorbing. If $0 < S(0) = i \leq 5$ then the first passage or stopping time of the state in which the random walk attains a value of 5 is

i	1	2	3	4	5
z_i	$\frac{8}{5}$	$\frac{14}{5}$	$\frac{16}{5}$	$\frac{12}{5}$	0
$\Omega_5(i)$	8	7	$\frac{16}{3}$	3	0

5.3 Intuitive Idea of a Stochastic Process

In this section we will bridge between the discrete random walk and the continuous random walk. We now have the need for a more traditional mathematical description of these random walks. An appropriate mathematical model must include some random component like the flipping coin of the previous sections. However a more useful random variable may be one which has a normal distribution with a prescribed mean and standard deviation. A familiar place to start this investigation is the deterministic process of exponential growth and decay. Most introductions to calculus

(see for example [Stewart (1999)]) contain the mathematical model which states that the rate of change of a non-negative quantity P , is proportional to P . Expressed as an equation this statement becomes

$$\frac{dP}{dt} = \mu P \quad (5.17)$$

where μ is the proportionality constant and t usually represents time. The proportionality constant is called a “growth rate” if $\mu > 0$ and a “decay rate” if $\mu < 0$. If the value of P is known at a specified value of t (usually at $t = 0$) then this differential equation has solution $P(t) = P(0)e^{\mu t}$, see exercise 6. This mathematical model is described as **deterministic** since there is no place for random events to express themselves in the model. Once the initial value of P and the proportionality constant are set, the future evolution of $P(t)$ is completely determined.

Equation (5.17) can be rewritten as

$$\frac{dP}{P} = \mu dt, \quad (5.18)$$

and if we make the change of variable $X = \ln P$ then this equation becomes

$$dX = \mu dt. \quad (5.19)$$

The following discussion may be easier to understand if we consider the discrete analogue of Eq. (5.19), namely

$$X(t + \Delta t) - X(t) = \Delta X = \mu \Delta t. \quad (5.20)$$

This equation is still deterministic but becomes **stochastic** if we introduce a random element. Suppose a normal random variable, $dz(t)$ with a mean of zero and a standard deviation of one is introduced as in the following equation.

$$\Delta X = \mu \Delta t + \sigma dz(t)\sqrt{\Delta t}. \quad (5.21)$$

We write the random variable as dz rather than z since we want to think of it as the change or “delta” in a random walk. The t dependency of dz comes from the need to think of z as taking on a random value at each time t . Furthermore we will assume that the $dz(t)$ is independent of $dz(s)$ when $t \neq s$. The constant σ is frequently referred to as **volatility** in financial applications. The product $\sigma dz(t)\sqrt{\Delta t}$ could naturally be thought of as a normal random variable with a mean of zero and a standard deviation of $\sigma\sqrt{\Delta t}$. The new random part of Eq. (5.21) represents “noise” whose

amplitude is a function of the volatility and the value of a normal random variable, but what is the purpose of the expression $\sqrt{\Delta t}$? It will turn out that the inclusion of this expression yields a couple of simple and yet elegant results regarding continuous stochastic processes. In the remainder of this section we will explore this question briefly.

Consider the cumulative change in X over several time steps each of size Δt . For this discrete model let $t_i < t_j = t_i + (j - i)\Delta t$ and consider the expression

$$\begin{aligned}\Delta X &= X(t_j) - X(t_i) \\ &= \sum_{k=i}^{j-1} (X(t_k + \Delta t) - X(t_k)) \\ &= \sum_{k=i}^{j-1} (\mu\Delta t + \sigma dz(t_k)\sqrt{\Delta t}).\end{aligned}\tag{5.22}$$

Since all of the random variables $\{dz(t_k)\}_{k=i}^{j-1}$ in Eq. (5.22) are independent for $t_k \neq t_{k'}$ and identically distributed with a mean of zero, then the average value of ΔX is $\mu(t_j - t_i)$. This is justified in the following set of equations.

$$\begin{aligned}\mathbb{E}[\Delta X] &= \mathbb{E}[X(t_j) - X(t_i)] \\ &= \mathbb{E}\left[\sum_{k=i}^{j-1} (\mu\Delta t + \sigma dz(t_k)\sqrt{\Delta t})\right] \\ &= \sum_{k=i}^{j-1} \mu\Delta t + \sigma\sqrt{\Delta t} \sum_{k=i}^{j-1} \mathbb{E}[dz(t_k)] \\ &= \mu(j - i)\Delta t \\ &= \mu(t_j - t_i)\end{aligned}\tag{5.23}$$

Notice that the expected value of $\Delta X = X(t_j) - X(t_i)$ over the time interval $[t_j, t_i]$ is the deterministic quantity $\mu(t_j - t_i)$. This might not seem so surprising if we keep in mind that the expected value of a random variable is usually a deterministic quantity. So far in this discussion the expression $\sqrt{\Delta t}$ has not played a significant role.

Now consider the variance of $\Delta X = X(t_j) - X(t_i)$.

$$\begin{aligned}\text{Var}(\Delta X) &= \text{Var}\left(\sum_{k=i}^{j-1} (\mu\Delta t + \sigma dz(t_k)\sqrt{\Delta t})\right) \\ &= \sigma^2 \Delta t \text{Var}\left(\sum_{k=i}^{j-1} dz(t_k)\right) \\ &= \sigma^2 \Delta t(j-i) = \sigma^2(t_j - t_i)\end{aligned}\tag{5.24}$$

We have assumed once again that the random variables $dz(t_k)$ for $k = i, \dots, j-1$ are pairwise independent. Now the purpose of the expression $\sqrt{\Delta t}$ in Eq. (5.21) becomes apparent. The random portion of the stochastic model in Eq. (5.21) contains $\sqrt{\Delta t}$ so that the variance of ΔX will be proportional to Δt . Without this assumption, the variance of ΔX would contain a nonlinear dependence on Δt . Since we will generally be working with a normal random variable with mean zero and variance $\sqrt{\Delta t}$ we introduce the symbol $dW(t)$ to represent it. This symbol will be used from now on.

The difference Eq. (5.21) can be used to generate a random walk if $X(t_i)$ is known for some value of t_i . In this case

$$\begin{aligned}X(t_{i+1}) - X(t_i) &= \mu\Delta t + \sigma dW(t_i) \\ X(t_{i+1}) &= X(t_i) + \mu\Delta t + \sigma dW(t_i) \\ \ln P(t_{i+1}) &= \ln P(t_i) + \mu\Delta t + \sigma dW(t_i) \\ P(t_{i+1}) &= P(t_i)e^{\mu\Delta t + \sigma dW(t_i)}\end{aligned}\tag{5.25}$$

The third line in the derivation above makes use of the assignment $\ln P = X$. For example if $\mu = 0.01$, $\sigma = 0.05$, $\Delta t = 0.004$, and $P(0) = 10$, then the random walk generated by the stochastic equation may resemble that shown in Fig. 5.4. Only a little imagination is needed to believe that the random walk pictured in Fig. 5.4 is the profile of end-of-day closing stock prices for a publicly traded company.

As $\Delta t \rightarrow 0$ this discrete model assumes the behavior of a continuous stochastic process. In the next section we will investigate the use of this model for describing the behavior of a corporation's stock price.

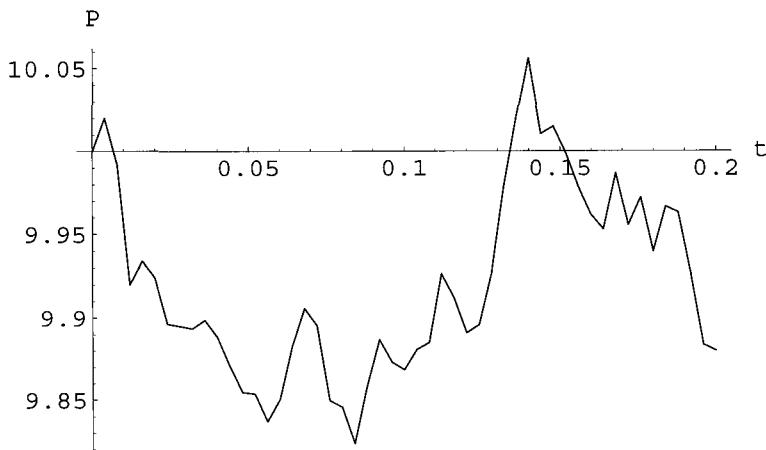


Fig. 5.4 A realization of a random walk generated by a stochastic process of the form in Eq. (5.21).

5.4 Stock Market Example

In the previous section we explored a discrete approximation to a continuous model for generating a random walk which seemed to mimic the fluctuations of the price of a corporation's stock. The model depended on several parameters which were called μ , σ , and Δt . In this section we intend to analyze the data associated with the stock price of an actual corporation and determine how to assign values to these parameters which are appropriate to that corporation. The end-of-day closing price of a corporation's stock is generally available on the World Wide Web. One convenient source of such information is the website, finance.yahoo.com. In this section we will use data for Sony Corporation stock. Appendix A contains the raw data analyzed here. The data is graphed in Fig. 5.5. The closing prices of the stock for 248 consecutive trading days are represented in this figure.

Since the data is sampled once every trading day (and days on which no trading took place, such as weekends and holidays, are irrelevant) then the value of Δt appropriate for the model to be developed is $\Delta t = 1$. Simple descriptive statistics can be used to determine μ and σ . The raw stock prices are playing the role of the variable P in our model. To fit the discrete form of the model used in Eq. (5.21) we must take the natural logarithm of the stock prices, since $X = \ln P$. Then we must form ΔX

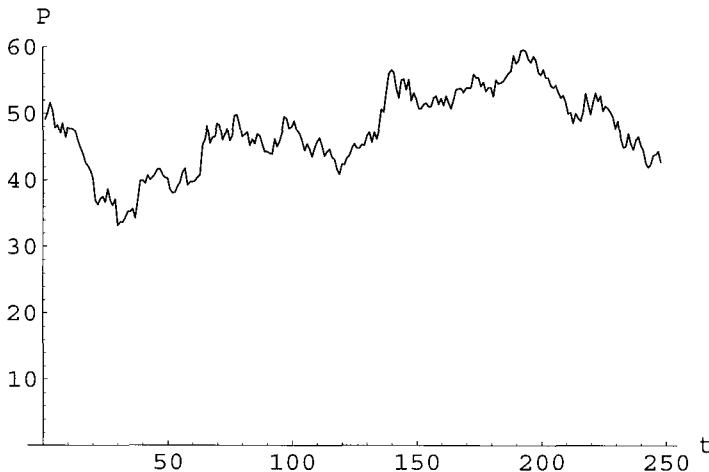


Fig. 5.5 The closing prices of Sony Corporation stock between August 13, 2001 (day 1) and August 12, 2002 (day 248).

by subtracting consecutive days logarithmic prices. If the values of ΔX just calculated fit the model then they will appear to be normally distributed. The mean of this data will be our estimate of μ and the standard deviation will be an estimate of σ which is also known as the volatility. A histogram of the values of ΔX calculated from the data in Appendix A appears in Fig. 5.6. It is certainly plausible that the values of ΔX summarized there are normally distributed. The mean of data is $\mu \approx -0.000555 \text{ day}^{-1}$ and the standard deviation is $\sigma \approx 0.028139 \text{ day}^{-1}$. Thus, despite the apparent randomness of the closing prices there is some structure and organization hidden in them.

The discrete approximation of the stochastic Eq. (5.21) evidently matches well the behavior of the stock price of an actual company. In the next section we will give some more background on stochastic processes. A reader interested in a more rigorous introduction to the mathematics and statistics of stochastic processes and stochastic differential equations should consult one or more of the references such as [Durrett (1996)] or [Mikosch (1998)]. In this book we are merely taking an intuitive, non-rigorous approach to their use as tools for modeling prices of stocks and their derivatives.

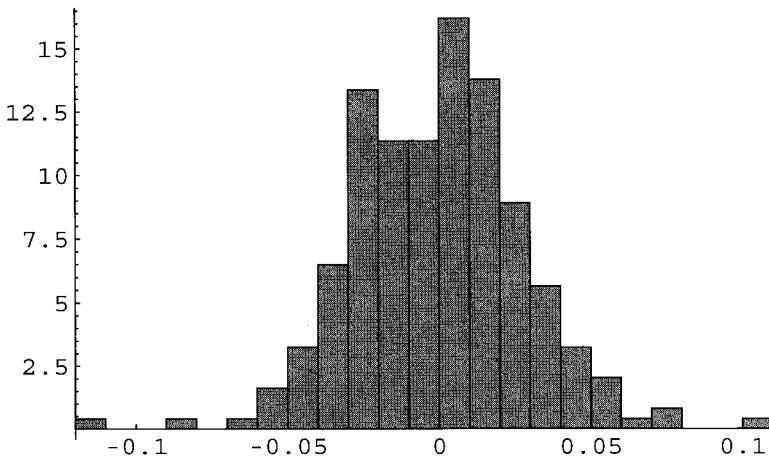


Fig. 5.6 A histogram of the values of the $\ln P(t_{i+1}) - \ln P(t_i)$ for $i = 1, 2, \dots, 247$.

5.5 More About Stochastic Processes

The equation shown in (5.21) is an example of a **generalized Wiener process**. In particular a generalized Wiener process can be described by a straight forward extension of an ordinary differential equation. This stochastic differential equation is obtained by adding a normally distributed random variable to the ODE.

$$dX = a dt + b dW(t) \quad (5.26)$$

The expressions a and b are constants and $dW(t)$ is a normal random variable. The random variable X can be found by integrating both sides of Eq. (5.26) to obtain

$$X(t) = X(0) + at + \int_0^t b dW(\tau). \quad (5.27)$$

Other types of stochastic processes are more general still. An **Itô process** has the form

$$dX = a(X, t) dt + b(X, t) dW(t) \quad (5.28)$$

where the expressions a and b are each now function of time t and the random variable X . In general the stochastic differential equation given in Eq. (5.28) cannot be solved explicitly through integration of both sides.

Nevertheless it is an important type of equation for us to use because it will enable us to work with stochastic processes of quantities which depend in turn on other random variables having their own individual stochastic process descriptions. In this way we will develop the fundamentally important Black-Scholes partial differential equation. Before we can do this we must develop an analogue of the chain rule for differentiation found in elementary calculus.

In standard differential calculus if one makes the assignment $X = \ln P$ then it is well understood that $dX = dP/P$ according to the chain rule. Using this on a generalized Wiener process may lead us to believe that the following equations are equivalent.

$$\begin{aligned} dX &= a dt + b dW(t) \\ dP &= aP dt + bP dW(t) \end{aligned}$$

First we note that while the first equation is a generalized Wiener process the second is an Itô process. Secondly, if a and b represent the mean and standard deviation respectively of random variable dX (as they did for the stock price data analyzed earlier), does the second equation imply that a and b are the mean and standard deviation of the underlying variable P ? The answer is no. What is needed is a valid procedure for changing variables in a stochastic process, in other words, a stochastic calculus version of the chain rule for differentiation. This change of variable differentiation result is known as Itô's Lemma and is covered in the next section.

5.6 Itô's Lemma

The proper procedure for changing variables in a stochastic differential equation can be derived using the multivariable version of Taylor's Theorem. A standard reference containing this result is [Taylor and Mann (1983)]. Here we will give a brief overview of a two-variable version of Taylor's formula with remainder. We start with the single-variable Taylor's formula. If $f(x)$ is an $(n+1)$ -times differentiable function on an open interval containing x_0 then the function may be written as

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 \\ &\quad + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\theta)}{(n+1)!}(x - x_0)^{n+1} \end{aligned} \tag{5.29}$$

The last term above is usually called the Taylor remainder formula and denoted R_{n+1} . The quantity θ lies between x and x_0 . The other terms form a polynomial in x of degree at most n and can be used as an approximation for $f(x)$ in a neighborhood of x_0 . This form of Taylor's will be used to derive the two-variable form.

Suppose the function $F(y, z)$ has partial derivatives up to order three on an open disk containing the point with coordinates (y_0, z_0) . Define the function $f(x) = F(y_0 + xh, z_0 + xk)$ where h and k are chosen small enough that $(y_0 + h, z_0 + k)$ lie within the disk surrounding (y_0, z_0) . Since f is a function of a single variable then we can use the single-variable form of Taylor's formula in (5.29) with $x_0 = 0$ and $x = 1$ to write

$$f(1) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + R_3. \quad (5.30)$$

Using the multi-variable chain rule for derivatives we have, upon differentiating $f(x)$ and setting $x = 0$,

$$f'(0) = hF_y(y_0, z_0) + kF_z(y_0, z_0) \quad (5.31)$$

$$f''(0) = h^2F_{yy}(y_0, z_0) + 2hkF_{yz}(y_0, z_0) + k^2F_{zz}(y_0, z_0). \quad (5.32)$$

We have made use of the fact that $F_{yz} = F_{zy}$ for this function under the smoothness assumptions. The remainder term R_3 contains only third order partial derivatives of F evaluated somewhere on the line connecting the points (y_0, z_0) and $(y_0 + h, z_0 + k)$. Thus if we substitute Eqs. (5.31) and (5.32) into (5.30) we obtain

$$\begin{aligned} \Delta F &= f(1) - f(0) \\ &= F(y_0 + h, z_0 + k) - F(y_0, z_0) \\ &= hF_y(y_0, z_0) + kF_z(y_0, z_0) \\ &\quad + \frac{1}{2}(h^2F_{yy}(y_0, z_0) + 2hkF_{yz}(y_0, z_0) + k^2F_{zz}(y_0, z_0)) + R_3. \end{aligned} \quad (5.33)$$

This last equation can be used to derive Itô's Lemma.

Let X be a random variable described by an Itô process of the form

$$dX = a(X, t)dt + b(x, t)dW(t) \quad (5.34)$$

where $dW(t)$ is a normal random variable and a and b are functions of X and t . Let $Y = F(X, t)$ be another random variable defined as a function of X and t . Given the Itô process which describes X we will now determine the Itô process which describes Y .

Using a Taylor series expansion for Y detailed in (5.33) we find

$$\begin{aligned}\Delta Y &= F_X \Delta X + F_t \Delta t + \frac{1}{2} F_{XX} (\Delta X)^2 + F_{Xt} \Delta X \Delta t + \frac{1}{2} F_{tt} (\Delta t)^2 + R_3 \\ &= F_X (a \Delta t + b dW(t)) + F_t \Delta t + \frac{1}{2} F_{XX} (a \Delta t + b dW(t))^2 \\ &\quad + F_{Xt} (a \Delta t + b dW(t)) \Delta t + \frac{1}{2} F_{tt} (\Delta t)^2 + R_3.\end{aligned}$$

Upon simplifying, the expression ΔX has been replaced by the discrete version of the Itô process. The dependence of a and b on X and t has been suppressed to simplify the notation. The reader is reminded that the Taylor remainder term R_3 contains terms of order $(\Delta t)^k$ where $k \geq 2$. Thus as Δt becomes small

$$\Delta Y \approx F_X (a \Delta t + b dW(t)) + F_t \Delta t + \frac{1}{2!} F_{XX} b^2 (dW(t))^2.$$

Recall that $dW(t)$ is a normal random variable with expected value zero and variance Δt . By definition

$$\text{Var}(dW(t)) = E[(dW(t))^2] - E[dW(t)]^2$$

which implies that $E[(dW(t))^2] = \Delta t$ which we will use as an approximation to $(dW(t))^2$. Thus we obtain the approximation

$$\Delta Y \approx F_X (a \Delta t + b dW(t)) + F_t \Delta t + \frac{1}{2!} F_{XX} b^2 \Delta t. \quad (5.35)$$

The reader should be aware that this is merely an outline of a proof of Itô's Lemma. In particular the replacement of $(dW(t))^2$ by Δt , while true, requires a more rigorous justification. The interested reader should consult [Neftci (2000), Chap. 10]. To summarize what we have done, the statement of the lemma is given below.

Lemma 5.3 (Itô's Lemma) *Suppose that the random variable X is described by the Itô process*

$$dX = a(X, t) dt + b(X, t) dW(t) \quad (5.36)$$

where $dW(t)$ is a normal random variable. Suppose the random variable $Y = F(X, t)$. Then Y is described by the following Itô process.

$$dY = \left(a(X, t) F_X + F_t + \frac{1}{2} (b(X, t))^2 F_{XX} \right) dt + b(X, t) F_X dW(t) \quad (5.37)$$

Now we can return to the question raised earlier regarding the stochastic process followed by a stock price P . If $X = \ln P$ and

$$dX = \mu dt + \sigma dW(t), \quad (5.38)$$

then we can let $F(X, t) = e^X = P$. According to Itô's Lemma then

$$\begin{aligned} dP &= \left(\mu e^X + 0 + \frac{1}{2} \sigma^2 e^X \right) dt + \sigma e^X dW(t) \\ &= \left(\mu + \frac{1}{2} \sigma^2 \right) P dt + \sigma P dW(t) \end{aligned}$$

Note that the coefficient of the deterministic portion of the equation contains the expression $\frac{1}{2}\sigma^2$ which would have been absent had the chain rule of deterministic calculus been used instead of Itô's Lemma. This lemma plays a role of central importance in the next chapter in which the Black-Scholes partial differential equations is derived.

5.7 Exercises

- (1) Another, less formal, method for obtaining the probability that a symmetric random walk which begins in state $S(0) = i > 0$ and finishes in state $S = A$ while avoiding an absorbing boundary at 0 is to treat the Laplacian Eq. (5.4) as a continuous boundary value problem. Find all the functions $f(x)$ which satisfy $f''(x) = 0$ subject to the conditions that $f(0) = 0$ and $f(A) = 1$.
- (2) Another, less formal, method for obtaining the stopping time for a symmetric random walk which begins in state $0 \leq S(0) = i \leq A$ and finishes in either state $S = 0$ or $S = A$ is to treat the Poisson Eq. (5.5) as a continuous boundary value problem. Find all the functions $g(x)$ which satisfy $g''(x) = -2$ subject to the conditions that $g(0) = 0$ and $g(A) = 0$.
- (3) If a stock price can change (up or down) by only one dollar per day and the probability of a unit increase is $1/2$ what is the probability that the stock price will increase by \$25 before decreasing by \$50?
- (4) What is the expected number of days that will elapse before the stock described in exercise 3 increases in price by \$25 or decreases by \$50?
- (5) What is the expected number of days that will elapse before the stock described in exercise 3 increases in price by \$25 while avoiding a decrease of \$50?

- (6) Assuming that μ and $P(0)$ are constants, verify by differentiation and substitution that $P(0)e^{\mu t}$ solves Eq. (5.17).
- (7) Using the growth constant and volatility of the Sony Corporation stock found earlier, use Eq. (5.25) to generate a realization of the random walk of stock prices for another year's duration.
- (8) Investigate the random walk of another company's stock. Collect if possible a year's worth of closing prices $(P(t_1), P(t_2), \dots, P(t_n))$ for the stock of your favorite company. Plot a histogram of $\Delta X = \ln P(t_{i+1}) - \ln P(t_i)$ and also calculate the mean and standard deviation of this random variable.
- (9) Using the growth constant and volatility of the stock data you found in exercise 8, use Eq. (5.25) to generate a realization of the random walk of stock prices for another year's duration.
- (10) Using the chain rule for derivatives from ordinary calculus (not the stochastic version governed by Itô's Lemma) verify the expressions found for $f'(0)$ and $f''(0)$ in equations (5.31) and (5.32). Find an expression for the Taylor remainder R_3 .
- (11) Suppose that P is governed by the stochastic process

$$dP = \mu P dt + \sigma P dW(t)$$

and that $Y = P^n$. Find the process which governs Y .

- (12) Suppose that P is governed by the stochastic process

$$dP = \mu P dt + \sigma P dW(t)$$

and that $Y = \ln P$. Find the process which governs Y .

Chapter 6

Options

In the present world of finance there are many types of financial instruments which go by the name of **options**. At its simplest, an option is the right, but not the obligation, to buy or sell a security such as a stock for an agreed upon price at some time in the future. The agreed upon price for buying or selling the security is known as the **strike price**. Options come with a time limit at which (or prior to) they must be exercised or else they expire and become worthless. The deadline by which they must be exercised is known as the **exercise time, strike time, or expiry date**. In the remainder of the text the three terms will be used interchangeably. An option to buy a security in the future is called a **call option**. An option to sell is known as a **put option**. Types of options can also be distinguished by their handling of the expiry date. The **European option** can only be exercised at maturity, while an **American option** can be exercised at or before the strike time. Of the two types, the European option is simpler to treat mathematically, and will be the focus of much of the rest of this book. However, in practice, American options are more commonly traded. There is a mathematical price to be paid in terms of complexity for the added flexibility of the American-style option.

Suppose stock in a certain company is selling today for \$100 per share. An investor may not want to buy this stock today, but may want to own it in the future. To reduce the risk of financial loss due to a potential large increase in the price of the stock during the next three-month period, they buy a European call option on the stock with a three-month strike time and a strike price of \$110. At the expiry date, if the price of the stock is above \$110 and the investor now wishes to buy the stock, they have the right as holder of the call option to purchase it for \$110, even if the market value of the stock is higher. Otherwise if the value of the stock is below

\$110 and the investor still wishes to buy the stock, they will let the call option expire without exercising it, and purchase the stock at its market price. Call options allow investors to protect themselves against paying an unexpectedly high price in the future for a stock which they are considering purchasing.

We have used the language of “buying” an option. Thus options themselves have some value, just as the securities which underlie the options have a value. An important issue to consider is how are values assigned to options? In light of the Arbitrage Theorem of Chapter 4, if the option is mis-priced relative to the security, arbitrage opportunities may be created. In this chapter we will explore some of the relationships between option and security prices. We will derive the **Black-Scholes** partial differential equation, together with boundary and final conditions which govern the prices of European call options.

During the explanations of many of the concepts in this chapter, we will refer to buying or selling a financial instrument. In the financial markets it is often possible to sell something which we do not yet own by borrowing the object from a true owner. For example investor A may borrow 100 shares of stock from investor B and sell them with the understanding that by some time in the future investor A will purchase 100 shares of the same stock and return them to investor B. Normally one buys an object first and sells it later, but in the financial arena one can often sell first and then buy later. Borrowing and selling an object with the agreement to purchase it later is called adopting a **short position** in the object, or sometimes **shorting** the object. Adopting a short position can be a profitable transaction if the investor believes the price of the object is going to decrease. Purchasing an object first and then selling it in the future is called adopting a **long position**.

6.1 Properties of Options

There are many relationships between the values of options and their underlying securities. The maintenance of most of these relationships is necessary to eliminate the possibility of arbitrage. We will cover some these relationships here and give the reader a taste of the way that one can prove these properties. Throughout this section we will use the following definitions.

C_a : value of an American-style call option

C_e : value of a European-style call option

K : strike price of an option

P_a : value of an American-style put option

P_e : value of a European-style put option

r : continuously compounded, risk-free interest rate

S : price of a share of a security

T : exercise time of an option

We have used the term “risk-free” to describe the interest rate r . It is assumed that this is the rate of return of an investment which carries no risk. While the reader may wish to debate whether a complete absence of risk can ever be achieved, there are some investments which carry very little risk, for example, U.S. Treasury Bonds.

Consider the cost of an American option compared with a European option. In the absence of arbitrage an American-style option must be worth at least as much as its European counterpart, *i.e.* $C_a \geq C_e$ and $P_a \geq P_e$. Naturally we are assuming that the strike prices, strike times, and underlying securities are the same. To the contrary, suppose that $C_a < C_e$. Taking the position of an informed consumer, knowing that the American-style option has all the characteristics of the European option and in addition has the increased flexibility that it may be exercised early, no investor would purchase a European-style call option if it cost more than the American-style option. To formulate a more mathematical proof, suppose an investor sells the European-style call option and purchases the American-style option. Since $C_e - C_a > 0$, the investor may purchase a risk-free bond paying interest at rate r compounded continuously. At the expiry date the bond will have value $(C_e - C_a)e^{rT}$. If the owner of the European option wishes to exercise the option, the investor insures this is possible by exercising their own American option. If the owner of the European option lets it expire unused, then the investor can do the same with the American option. Thus in both cases the investor makes a risk-free profit of $(C_e - C_a)e^{rT}$ which is an example of arbitrage.

It is also the case that $C_e \geq S - Ke^{-rT}$ or arbitrage is present. Assuming an absence of arbitrage and that $C_e < S - Ke^{-rT}$, an investor can purchase the call option and sell the security at time $t = 0$. Since $0 < Ke^{-rT} < S - C_e$, the investor can invest $S - C_e$ in a risk-free bond paying interest rate r . At the strike time the bond is worth $(S - C_e)e^{rT}$ which is greater than K . Thus if the investor chooses to exercise the option (because S at time T is greater than K), the bond can be cashed out, the security purchased for K , and there is still capital left over. In other words there is risk-free

profit. On the other hand if S at time T is less than K , the short position on the security is eliminated using the bond and purchasing the stock for S (which is less than K). Again there will still be capital left over. Once again there is a risk-free profit.

There is strict arbitrage-free relationship between the value of a European call and put with the same strike price and expiry date on the same underlying security. This is the important **Put-Call Parity Formula** of Eq. (6.1).

$$P_e + S = C_e + Ke^{-rT} \quad (6.1)$$

To see why this relationship must be true, imagine portfolio A represents the left-hand side of (6.1) while the right-hand side is represented by portfolio B. The Put-Call Parity Formula implies that in an arbitrage-free setting these portfolios must have the same value.

Suppose portfolio A is worth less than portfolio B, *i.e.*

$$P_e + S < C_e + Ke^{-rT}. \quad (6.2)$$

An investor can borrow at interest rate r an amount equal to $P_e + S - C_e$. This would allow the investor to purchase the security, the European put option, and to sell the European call option. At the strike time of the two options, the investor must repay principal and interest in the amount of $(P_e + S - C_e)e^{rT}$. If the security is worth more than K at time T , the put expires worthless and the call will be exercised by its owner. The investor must sell the security for K . Thus the net proceeds of this transaction are

$$K - (P_e + S - C_e)e^{rT} > 0 \quad (6.3)$$

since inequality (6.3) is equivalent to the one in (6.2). If the security is worth less than K at time T , the call expires worthless and the put will be exercised by the investor. Again the investor will sell the security for K . The net proceeds of this transaction are the same as in the previous case. Thus there is a risk-free profit to be realized if portfolio A is worth less than portfolio B.

Now suppose portfolio A is worth more than portfolio B, *i.e.*

$$P_e + S > C_e + Ke^{-rT}. \quad (6.4)$$

An investor can sell the security and the European put option and buy the call option. This generates an initial positive flow of capital in the amount of $S + P_e - C_e$. This amount will be invested in a risk-free bond earning

interest at rate r until the expiry date arrives. At that time the investor will have $(S + P_e - C_e)e^{rT}$. If the security is worth more than K at time T , then the put option is worthless and investor will exercise the call option. The investor purchases the security for K (thus canceling their short position). This leaves the investor with a net gain of

$$(P_e + S - C_e)e^{rT} - K > 0 \quad (6.5)$$

since inequality (6.5) is equivalent to the one in (6.4). If the security is worth less than K at time T , the call option expires unused and the owner of the put option will exercise it. Thus the investor will clear their short position by buying the security for K and their net gain is as before. Consequently there exists an arbitrage opportunity if portfolio B is worth less than portfolio A. Therefore the two portfolios must have the same value, *i.e.* the Put-Call Parity Formula is true.

Assuming that the prices of European call and put options are non-negative (a mild assumption), we can derive two corollary inequalities to the Put-Call Parity Formula.

$$C_e \geq S - Ke^{-rT} \quad (6.6)$$

$$P_e \geq -S + Ke^{-rT} \quad (6.7)$$

While none of the formulas or inequalities developed so far enable us to price options directly, they are useful checks for arbitrage opportunities.

6.2 Pricing an Option Using a Binary Model

In this section we will examine a very simple option/security situation and determine the arbitrage-free price for the security. The analysis done here is too simple to be applied directly to any real-world example of a security; however, it provides an example application of the Arbitrage Theorem (Theorem 4.4). This type of analysis can be generalized and extended to more relevant real-world examples and thus provides a good starting point for the goal of developing option pricing formulae. The analysis performed here is based on a binary model, so named since the security is assumed to take on one of two values in the future.

Suppose the price of a share of stock is \$100 currently. After a single, indivisible unit of time T , the price of the share will either be \$50 or \$200. The stock will be worth \$50 after time T with probability p or will be worth \$200 with probability $1 - p$. Currently an investor can purchase a

European call option whose value is C . The exercise time and strike price of the option are respectively T and \$150. In the absence of arbitrage what is the value of C ?

The investor could buy either the option or the stock at the beginning of the time interval. In the absence of arbitrage, the expected value of the investor's profit from either course of action should be zero regardless of which direction the price of the stock moves. Since the option or stock may be purchased at the beginning of the time interval but the profit (if any) arrives after time T has elapsed, we must calculate the present value of any potential profit. We will assume the interest rate per T unit of time is r . Suppose the investor purchases the stock initially. At time T their net gain is $-100 + 200(1+r)^{-1}$ with probability p or $-100 + 50(1+r)^{-1}$ with probability $1-p$. In an arbitrage-free setting the expected value of this gain is zero.

$$\begin{aligned} 0 &= (-100 + 200(1+r)^{-1}) p + (-100 + 50(1+r)^{-1})(1-p) \\ 0 &= -100(1+r) + 150p + 50 \\ p &= \frac{1+2r}{3} \end{aligned} \tag{6.8}$$

On the other hand, if the investor purchases the option initially, at time T their net gain is $-C + (200 - 150)(1+r)^{-1}$ with probability p or $-C$ with probability $1-p$. Again the expected value of this gain will be zero in the absence of arbitrage.

$$\begin{aligned} 0 &= (-C + 50(1+r)^{-1}) p + (-C)(1-p) \\ 0 &= 50p(1+r)^{-1} - C \\ C &= \frac{50 + 100r}{3(1+r)} \end{aligned} \tag{6.9}$$

Equation (6.9) was found using Eq. (6.8). The arbitrage-free probability vector and the option cost both depend on the interest rate. Figure 6.1 contains a parametric plot of the points $(p(r), C(r))$. If the probability p and the call option price deviate from the curve shown, then a risk-free profit can be made.

Example 6.1 Suppose the current value of the stock is $S(0) = \$100$ and a time $T = 1$

$$S(1) = \begin{cases} \$200 & \text{with probability } p = 11/30, \\ \$50 & \text{with probability } 1-p = 19/30. \end{cases}$$

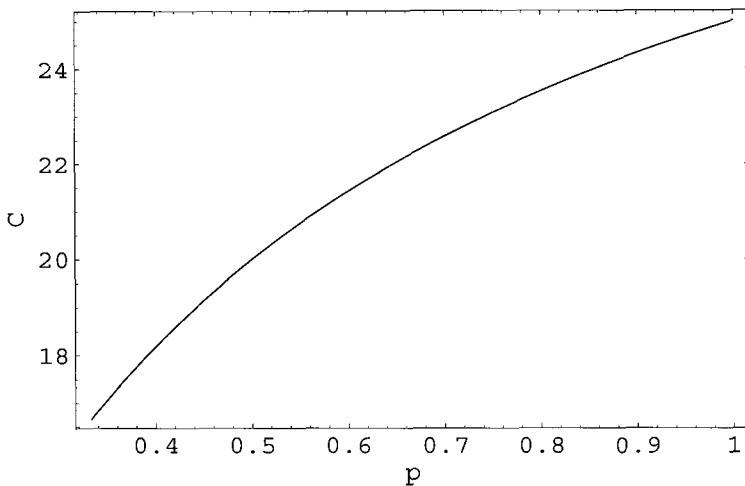


Fig. 6.1 A parametric plot of probability and call option cost as a function of interest rate.

Suppose further that the risk-free interest rate is $r = 5\%$ and that a European call option for the stock can be purchased for $C = \$17.50$. The strike price of the option is $K = \$150$. A risk-free profit can be generated in this situation since the option has been mis-priced according to Eqs. (6.9) and (6.8).

Suppose an investor borrows sufficient cash to purchase x shares of stock and y call options. The portfolio is originally worth $100x + 17.50y$ and they must pay back $e^{0.05}(100x + 17.50y)$ at time $T = 1$. If $S(1) = \$200$ then the call option will be exercised. Liquidating the portfolio yields $200x + 50y$ in revenue. If $S(1) = \$50$ the call option is not be exercised. Cashing out the portfolio produces $50x$ in revenue. Thus if x and y can be found so that the following set of linear inequalities are simultaneously satisfied then a risk free profit can be realized.

$$\begin{aligned} 200x + 50y &> e^{0.05}(100x + 17.50y) \\ 50x &> e^{0.05}(100x + 17.50y) \end{aligned}$$

One such solution is the point with coordinates $(x, y) = (1000, -3000)$. Therefore if the investor takes a long position in 1000 shares of the stock and shorts 3000 call options on the stock, their risk-free profit will be \$64.6229.

6.3 Black-Scholes Partial Differential Equation

In this section we will derive the fundamental equation governing the pricing of options, the famous **Black-Scholes partial differential equation**. The ideas of arbitrage, stochastic processes, and present value converge at this point in the study of financial mathematics.

We begin by supposing that S is the current value of a security and that it obeys a stochastic process of the form

$$dS = \left(\mu + \frac{\sigma^2}{2} \right) S dt + \sigma S dW(t) \quad (6.10)$$

If $F(S, t)$ is the value of any type of option (more generally called a financial derivative), then according to Itô's Lemma (Lemma 5.3), F obeys the following stochastic process.

$$dF = \left(\left[\mu + \frac{\sigma^2}{2} \right] SF_S + \frac{1}{2}\sigma^2 S^2 F_{SS} + F_t \right) dt + \sigma SF_S dW(t) \quad (6.11)$$

Suppose a portfolio of value P is created by selling the option and buying Δ units of the security. Thus the value of the portfolio is $P = F - \Delta S$. Since the portfolio is a linear combination of the the option and the security then the stochastic process governing the portfolio is

$$\begin{aligned} dP &= d(F - \Delta S) \\ &= dF - \Delta dS \\ &= \left(\left[\mu + \frac{\sigma^2}{2} \right] SF_S + \frac{1}{2}\sigma^2 S^2 F_{SS} + F_t \right) dt + \sigma SF_S dW(t) \\ &\quad + \Delta \left(\left(\mu + \frac{\sigma^2}{2} \right) S dt + \sigma S dW(t) \right) \\ &= \left(\left[\mu + \frac{\sigma^2}{2} \right] S [F_S - \Delta] + \frac{1}{2}\sigma^2 S^2 F_{SS} + F_t \right) dt + \\ &\quad + \sigma S (F_S - \Delta) dW(t) \end{aligned} \quad (6.12)$$

Notice the coefficient of the normal random variable, $dW(t)$, contains the factor $F_S - \Delta$. Equation (6.12) can be simplified if we assume $\Delta = F_S$. This has the beneficial effect of reducing the number of stochastic terms in Eq. (6.12). Randomness is not completely eliminated since the value of the security S remains and is stochastic. Under the assumption that $\Delta = F_S$,

Eq. (6.12) becomes

$$dP = \left(\frac{1}{2} \sigma^2 S^2 F_{SS} + F_t \right) dt. \quad (6.13)$$

In an arbitrage-free setting the difference in the returns from investing in the portfolio described above or investing an equal amount of capital in a risk-free bond paying interest at rate r should be zero. Thus the following equations are true.

$$\begin{aligned} 0 &= rP dt - dP \\ 0 &= rP dt - \left(\frac{1}{2} \sigma^2 S^2 F_{SS} + F_t \right) dt \\ rP &= \frac{1}{2} \sigma^2 S^2 F_{SS} + F_t \\ r(F - \Delta S) &= \frac{1}{2} \sigma^2 S^2 F_{SS} + F_t \\ rF &= \frac{1}{2} \sigma^2 S^2 F_{SS} + F_t + r\Delta S \\ rF &= F_t + rSF_S + \frac{1}{2} \sigma^2 S^2 F_{SS} \end{aligned} \quad (6.14)$$

Equation (6.14) is an example of a **partial differential equation** or **PDE**, for short. While the general theory of solving PDEs is beyond the scope of this book, we will briefly mention a few concepts related to the Black-Scholes PDE. PDEs are often described by their order, type, and linearity properties. The Black-Scholes PDE is a **second order** equation since the highest order derivative of the unknown function F present in the the equation is the second derivative. This PDE is of **parabolic type**. The best known example of the parabolic PDE is the **heat equation** which is used to describe the distribution of temperature along an object, for example a metal rod. For an introductory treatment of the heat equation see [Boyce and DiPrima (2001)]. In [Wilmott *et al.* (1995)] the Black-Scholes PDE is solved by appropriate changes of variables until it becomes the heat equation, which is easily solved via a technique known as **separation of variables**. Lastly, the Black-Scholes PDE is an example of a **linear** partial differential equation since if F_1 and F_2 are two solutions to the equation then $c_1 F_1 + c_2 F_2$ is also a solution where c_1 and c_2 are any constants (see exercise 9).

Financial derivative products of many types obey the Black-Scholes equation. Different solutions correspond to different initial/final and

boundary side conditions imposed while solving the equation. In the next section we will describe in more detail the initial and boundary conditions relevant to the Black-Scholes PDE. In Sec. 7.4 we will derive a solution to the Black-Scholes equation for the price of a European style call option.

6.4 Boundary and Initial Conditions

Without the imposition of side conditions in the form of initial or final conditions and boundary conditions, a partial differential equation can have many different solutions. Mathematicians refer to a differential equation with many possible solutions as “ill-posed.” In this section we will discuss the appropriate conditions to impose in order to obtain a solution to Eq. (6.14) which describes the value of a European call option.

The domain of the unknown function F of Eq. (6.14) is a region in (S, t) -space. We will refer to this region as $\Omega \subset (S, t)$ -space. For a European call option we are only interested in the value of the option during the time interval $[0, T]$, where T is the exercise time of the option. During this time interval the price of the security underlying the option may be any non-negative value. Thus the domain of the solution to Eq. (6.14) is

$$\Omega = \{(S, t) \mid 0 \leq S < \infty \text{ and } 0 \leq t \leq T\}.$$

At time $t = T$ the value of the security will either exceed the strike price (in which case the call option will be exercised generating an income flow of $S(T) - K > 0$) or the security will have a value less than or equal to the strike price (in which case the call option expires unused and has value 0). When $S(T) > K$ the call option is said to be **in the money**. When $S(T) \leq K$ the call option is said to be **out of the money**. Thus the terminal value of a European call option is

$$(S(T) - K)^+ = \max\{S(T) - K, 0\},$$

where $S(T)$ is the value of the underlying security at the exercise time and K is the strike price of the option. Graphically the payoff of the portfolio is said to resemble a “hockey stick”. See Fig. 6.2. Hence we can use the following equation as the **final condition** of the Black-Scholes PDE.

$$F(S, T) = \max\{S(T) - K, 0\} \tag{6.15}$$

Since there is a final condition and no initial condition, the Black-Scholes PDE is sometimes referred to as a **backwards** parabolic equation.

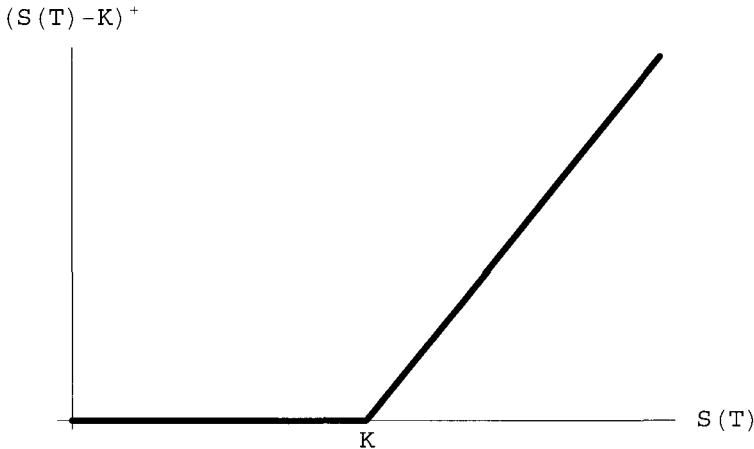


Fig. 6.2 The piecewise linear curve representing the payoff of a European call option.

We see from Eq. (6.10) that if $S = 0$, then $dS = 0$, i.e. S never changes and hence remains zero. The state $S = 0$ is an absorbing boundary condition for the random walk of S . Thus on the portion of the boundary of Ω where $S = 0$, the call option would never be exercised and hence must have zero value. Thus we have derived one boundary condition, namely

$$F(0, t) = 0. \quad (6.16)$$

Now if we suppose that S is approaching infinity, it becomes increasingly likely that the call option will be exercised, since as $S \rightarrow \infty$, S will exceed any finite value of K . Also as $S \rightarrow \infty$, the difference $S - K \approx S$ and hence we have the remaining boundary condition

$$F(S, t) = S, \quad \text{as } S \rightarrow \infty. \quad (6.17)$$

Thus to summarize the Black-Scholes equation and its final and boundary conditions for a European call option we have the following set of equations.

$$\begin{aligned} rF &= F_t + rSF_S + \frac{1}{2}\sigma^2 S^2 F_{SS} \quad \text{for } (S, t) \text{ in } \Omega, \\ F(S, T) &= (S(T) - K)^+ \quad \text{for } S > 0, \\ F(0, t) &= 0 \quad \text{for } 0 \leq t < T, \\ F(S, t) &= S \quad \text{as } S \rightarrow \infty. \end{aligned}$$

6.5 Exercises

- (1) Show that in the absence of arbitrage $P_a \geq P_e$ where the underlying security, exercise time, and strike price for both options are the same.
- (2) Consider a European call option with a strike price of \$60 which costs \$10. Draw a graph illustrating the net payoff of the option for stock prices in the interval $[0, 100]$.
- (3) Show that in the absence of arbitrage the value of a call option never exceeds the value of the underlying security.
- (4) What is the minimum price of a European style call option with an exercise time of three months and a strike price of \$26 for a security whose current value is \$29 while the continuously compounded interest rate is 6%?
- (5) If a share of a security is currently selling for \$31, a three-month European call option is \$3 with a strike price of \$31, and the risk-free interest rate is 10%, what is the arbitrage-free European put option price for the security?
- (6) If a share of a security is currently selling for \$31, a three-month European call option is \$3, a three-month European put option is \$2.25, and the risk-free interest rate is 10%, what is the arbitrage-free strike price for the security?
- (7) What is the minimum price of a European put option with an exercise time of two months and a strike price of \$14 for a stock whose value is \$11 while the continuously compounded interest rate is 7%?
- (8) If a share of a security is currently selling for \$31, a three-month European call option is \$3, a three-month European put option is \$1, the strike price for both options is \$30, and the risk-free interest rate is 10%, what arbitrage opportunities are open to investors?
- (9) Suppose the functions $f_1(S, t)$ and $f_2(S, t)$ each solve the Black-Scholes Eq. (7.5). Show that if c_1 and c_2 are constants then the function $f(S, t) = c_1 f_1(S, t) + c_2 f_2(S, t)$ also solves the Black-Scholes equation.
- (10) What final and boundary conditions are appropriate for the Black-Scholes PDE (6.14) when F represents a European-style put option?

Chapter 7

Solution of the Black-Scholes Equation

In Chapter 6 the Black-Scholes partial differential equation was derived and summarized in Eq. (6.14). Every European style option satisfies this PDE. The differences in the options are due to different boundary and final conditions. In the present chapter we will solve the Black-Scholes equation with boundary and final conditions appropriate for a European style call option. Several methods are available to solve this equation. Some take the limit of a discrete time binomial model of security prices (see [Ross (1999)]). In [Wilmott *et al.* (1995)] the Black-Scholes PDE is transformed through a sequence of changes of variable to an ordinary differential equation which is solved by elementary means. Their change of variable approach will be mimicked here, but the **Fourier Transform** will be used to solve the equation. For readers unfamiliar with the Fourier Transform, a brief introduction to the relevant operations are included in this chapter. The Fourier Transform is a standard mathematical device used to convert certain types of partial differential equations into ordinary differential equations. A more complete introduction to the Fourier Transform can be found in either [Greenberg (1998)] or [Jeffrey (2002)].

7.1 Fourier Transforms

Suppose the function $f(x)$ is defined for all real numbers. and that The Fourier Transform of f is denoted $\mathcal{F}\{f(x)\}$ or often as $\hat{f}(w)$. By definition the Fourier Transform of f is

$$\mathcal{F}\{f(x)\} = \hat{f}(w) = \int_{-\infty}^{\infty} f(x)e^{-iwx} dx, \quad (7.1)$$

provided the improper integral converges. There are many theorems in the literature on the Fourier Transform which list conditions under which the Fourier Transform will exist. It is possible to state very general but technical conditions in such an existence theorem. For our purposes we will adopt a less technical theorem which states that if the domain of f is all real numbers and if

- f and f' are piecewise continuous on every interval of the form $[-M, M]$ for arbitrary $M > 0$, and
- $\int_{-\infty}^{\infty} |f(x)| dx$ converges,

then the Fourier Transform of f exists according to Theorem 17.9.1 of [Greenberg (1998)]. The symbol $i = \sqrt{-1}$. The Fourier Transform maps the original function f which is a function of x to a new function \hat{f} which depends on w which is often thought of as a frequency.

Example 7.1 Consider the piecewise-defined function

$$f(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Its Fourier Transform is readily found through the following calculation.

$$\begin{aligned}\hat{f}(w) &= \int_{-\infty}^{\infty} f(x)e^{-iwx} dx \\ &= \int_{-1}^1 e^{-iwx} dx \\ &= -\frac{1}{iw} (e^{-iw} - e^{iw}) \\ &= \frac{2}{w} \sin w\end{aligned}$$

In the previous example we carried out integration of a complex exponential as if it followed the same rules of integration as real-values functions. This is in fact true. The **Euler Identity**, $e^{i\theta} = \cos \theta + i \sin \theta$ was also used to simplify the result [Churchill *et. al.* (1976)].

The Fourier Transform also has an interesting effect on derivatives of functions. Suppose the Fourier Transform of f exists and that f is differ-

entiable for all real numbers.

$$\begin{aligned}
 \mathcal{F}\{f'(x)\} &= \int_{-\infty}^{\infty} f'(x)e^{-iwx} dx \\
 &= f(x)e^{-iwx} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)(-iw)e^{-iwx} dx \\
 &= iw \int_{-\infty}^{\infty} f(x)e^{-iwx} dx \\
 &= iw\hat{f}(w)
 \end{aligned} \tag{7.2}$$

The technique of integration by parts was used to shift the derivative off of f and onto the exponential function. Since f vanishes as $|x| \rightarrow \infty$ the leading integral is zero. This property of the Fourier Transform can be extended to higher order derivatives.

Theorem 7.1 *If $f(x)$, $f'(x)$, \dots , $f^{(n-1)}(x)$ are all Fourier transformable and if $f^{(n)}(x)$ exists (where $n \in \mathbb{N}$) then $\mathcal{F}\{f^{(n)}(x)\} = (iw)^n \hat{f}(w)$.*

Proof. The result has already been shown in Eq. (7.2) for the case where $n = 1$. The technique of proof by induction is used to establish this result for all higher order derivatives. Suppose the result has been proved for $n = k$.

$$\begin{aligned}
 \mathcal{F}\{f^{(k+1)}(x)\} &= \int_{-\infty}^{\infty} f^{(k+1)}(x)e^{-iwx} dx \\
 &= f^{(k)}(x)e^{-iwx} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f^{(k)}(x)(-iw)e^{-iwx} dx \\
 &= iw \int_{-\infty}^{\infty} f^{(k)}(x)e^{-iwx} dx \\
 &= iw(iw)^k \hat{f}(w) = (iw)^{k+1} \hat{f}(w)
 \end{aligned}$$

The induction assumption was used in the second to the last line of this equation. \square

Another Fourier Transform result which we will have need for later concerns the transform of the convolution of two functions. By definition the **Fourier Convolution** of two functions f and g is

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-z)g(z) dz. \tag{7.3}$$

The Fourier Convolution is not the same as the Fourier Transform. The convolution involves no complex integration, merely the integration of the

product of two real-valued functions. It is possible to calculate the Fourier Transform of this convolution.

Theorem 7.2 $\mathcal{F}\{(f * g)(x)\} = \hat{f}(w)\hat{g}(w)$, in other words the Fourier Transform of the Fourier Convolution of f and g is the product of the Fourier Transforms of f and g .

Proof. We begin by applying the definitions of Fourier Transform and Convolution.

$$\begin{aligned}\mathcal{F}\{(f * g)(x)\} &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x-z)g(z) dz \right] e^{-iwx} dx \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x-z)g(z)e^{-iwx} dz \right] dx \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x-z)g(z)e^{-iwx} dx \right] dz \\ &= \int_{-\infty}^{\infty} g(z) \left[\int_{-\infty}^{\infty} f(x-z)e^{-iwx} dx \right] dz\end{aligned}$$

Since the order of integration in a multiple integral is irrelevant (only the region over which the integration takes place is meaningful), equality is preserved above. Now the change of variable $x = u + z$ is used to rewrite the interior integral above.

$$\begin{aligned}\mathcal{F}\{(f * g)(x)\} &= \int_{-\infty}^{\infty} g(z) \left[\int_{-\infty}^{\infty} f(u)e^{-iw(u+z)} du \right] dz \\ &= \int_{-\infty}^{\infty} g(z)e^{-iwz} \left[\int_{-\infty}^{\infty} f(u)e^{-iwu} du \right] dz \\ &= \hat{f}(w) \int_{-\infty}^{\infty} g(z)e^{-iwz} dz \\ &= \hat{f}(w)\hat{g}(w)\end{aligned}$$

□

7.2 Inverse Fourier Transforms

So far we have seen that the Fourier Transform can be used to map functions of x , their derivatives, and their convolutions to other functions which depend on a frequency-like variable w . We intend to apply this transformation to the Black-Scholes partial differential equation and its associated side conditions. However, any useful solution to the Black-Scholes equation

should depend on the variables and constants in the original equation and not on the frequency variable w . It is necessary to have a method for transforming a function of w back into a function of x . This is the purpose of the **Inverse Fourier Transform**. By definition the inverse Fourier Transform of $\hat{f}(w)$ given by

$$\mathcal{F}^{-1}\{\hat{f}(w)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(w) e^{iwx} dw. \quad (7.4)$$

Example 7.2 Suppose a is a positive constant, then the inverse Fourier Transform of $e^{-a|w|}$ is

$$\begin{aligned} \mathcal{F}^{-1}\{e^{-a|w|}\} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-a|w|} e^{iwx} dw \\ &= \frac{1}{2\pi} \int_{-\infty}^0 e^{(a+ix)w} dw + \frac{1}{2\pi} \int_0^{\infty} e^{(-a+ix)w} dw \\ &= \frac{1}{2\pi(a+ix)} + \frac{1}{2\pi(a-ix)} \\ &= \frac{a}{\pi(a^2+x^2)} \end{aligned}$$

Readers with some prior exposure to the Fourier Transform may prefer more symmetric definitions of the transform and the inverse transform. An infinite number of variations on the formulas can be stated, though they are all equivalent and compatible with the versions given above in Eqs. (7.1) and (7.4). Some people prefer each of the forward and inverse transforms to be scaled by a factor involving $\sqrt{2\pi}$ as in the following.

$$\begin{aligned} \hat{f}(w) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-iwx} dx \\ f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w) e^{iwx} dw \end{aligned}$$

7.3 Changing Variables in the Black-Scholes PDE

To recap the partial differential equation and its side conditions for a European style call option, we have the following three equations

$$rF = F_t + \frac{1}{2}\sigma^2 S^2 F_{SS} + rSF_S \quad \text{for } t \in [0, T), S \in (0, \infty) \quad (7.5)$$

$$F(S, T) = (K - S)^+ \quad \text{for } S \in (0, \infty) \quad (7.6)$$

$$F(0, t) = 0, \quad F(S, t) \rightarrow S \quad \text{as } S \rightarrow \infty \text{ for } t \in [0, T]. \quad (7.7)$$

The final condition given in (7.6) states that when the expiry date for the option arrives, the call option will be worth nothing if the value of the stock is less than the strike price. Otherwise the call option is worth the excess of the stock's value over the strike price. The boundary conditions are specified in Eq. (7.7). The first condition implies that if the stock itself becomes worthless before maturity, the call option is also worthless. An investor could buy the stock for nothing and then let the call option expire unused. The second part of this condition follows from the European call option property discussed in Chapter 6, namely $C_e \geq S - Ke^{-rT}$. Thus as $S \rightarrow \infty$ the value of the call option becomes S asymptotically.

A judicious change of variables for the Black-Scholes equation can simplify it. In fact, in only a few steps we can convert Eq. (7.5) into a more well known partial differential equation, namely the **Heat Equation**. Suppose F , S , and t are defined in terms of the new variables v , x , and τ as in the following equations.

$$S = Ke^x \Leftrightarrow x = \ln \frac{S}{K} \quad (7.8)$$

$$t = T - \frac{2\tau}{\sigma^2} \Leftrightarrow \tau = \frac{\sigma^2}{2}(T - t) \quad (7.9)$$

$$F(S, t) = Kv(x, \tau) \quad (7.10)$$

How are the Black-Scholes equation and its side conditions altered by this change of variables? The multivariable form of the chain rule can be used to determine new expressions for the derivatives present in Eq. (7.5).

$$F_S = (Kv(x, \tau))_S = K(v_x x_S + v_t t_S) = K\left(v_x \frac{1}{S} + 0\right) = e^{-x} v_x \quad (7.11)$$

The reader will be asked to verify that

$$F_t = -\frac{K\sigma^2}{2}v_\tau \quad \text{and} \quad (7.12)$$

$$F_{SS} = \frac{e^{-2x}}{K}(v_{xx} - v_x) \quad (7.13)$$

in exercises 4 and 5. Substituting the results in Eqs. (7.10)–(7.13) in the Black-Scholes Eq. (7.5) and simplifying produces the equation

$$v_\tau = v_{xx} + (k - 1)v_x - kv \quad (7.14)$$

where $k = 2r/\sigma^2$ (see exercise 6).

The final condition for F is converted by this change of variables into an initial condition since when $t = T$, $\tau = 0$. The initial condition $v(x, 0)$ is then found to be

$$\begin{aligned} Kv(x, 0) &= F(S, T) \\ &= (S - K)^+ \\ &= K(e^x - 1)^+ \\ v(x, 0) &= (e^x - 1)^+ \end{aligned} \quad (7.15)$$

Since $\lim_{S \rightarrow 0^+} x = -\infty$, then

$$0 = \lim_{S \rightarrow 0^+} F(S, t) = \lim_{x \rightarrow -\infty} Kv(x, \tau) \implies \lim_{x \rightarrow -\infty} v(x, \tau) = 0.$$

Likewise since as $\lim_{S \rightarrow \infty} x = \infty$,

$$S = \lim_{S \rightarrow \infty} F(S, t) = \lim_{x \rightarrow \infty} Kv(x, \tau) = Ke^x \implies \lim_{x \rightarrow \infty} v(x, \tau) = e^x.$$

Thus we have derived a pair of boundary conditions for the partial differential equation in (7.14). Thus the original Black-Scholes initial, boundary value problem has been recast in the form of the following.

$$v_\tau = v_{xx} + (k - 1)v_x - kv \quad \text{for } x \in (-\infty, \infty), \tau \in (0, \frac{T\sigma^2}{2}) \quad (7.16)$$

$$v(x, 0) = (e^x - 1)^+ \quad \text{for } x \in (-\infty, \infty) \quad (7.17)$$

$$v(x, \tau) \rightarrow 0 \quad \text{as } x \rightarrow -\infty \text{ and} \quad (7.18)$$

$$v(x, \tau) \rightarrow e^x \quad \text{as } x \rightarrow \infty, \tau \in (0, \frac{T\sigma^2}{2}) \quad (7.19)$$

The independent variable x can be thought of as corresponding to a spatial variable. The reader should note that where previously the price of the security S was assumed to take on only non-negative values, now x can be any real number.

Now another change of variables is needed. If α and β are constants then we can introduce the new dependent variable u .

$$v(x, \tau) = e^{\alpha x + \beta t} u(x, \tau) \quad (7.20)$$

$$v_x = e^{\alpha x + \beta t} (\alpha u(x, \tau) + u_x) \quad (7.21)$$

$$v_{xx} = e^{\alpha x + \beta t} (\alpha^2 u(x, \tau) + 2\alpha u_x + u_{xx}) \quad (7.22)$$

$$v_\tau = e^{\alpha x + \beta t} (\beta u(x, \tau) + u_\tau) \quad (7.23)$$

Substituting the expressions found in (7.20)–(7.23) into Eq. (7.16) produces

$$u_\tau = (\alpha^2 + (k - 1)\alpha - k - \beta)u + (2\alpha + k - 1)u_x + u_{xx} \quad (7.24)$$

Since α and β are arbitrary constants they can now be chosen appropriately to simplify Eq. (7.24). Ideally the coefficients of u_x and u would be zero. Solving the two equations:

$$\begin{aligned} 0 &= \alpha^2 + (k - 1)\alpha - k - \beta \\ 0 &= 2\alpha + k - 1 \end{aligned}$$

yields $\alpha = (1 - k)/2$ and $\beta = -(k + 1)^2/4$. The initial condition for u can be derived from the initial condition given in (7.17).

$$\begin{aligned} v(x, 0) &= (e^x - 1)^+ \\ e^{\alpha x} u(x, 0) &= \\ u(x, 0) &= e^{(k-1)x/2}(e^x - 1)^+ \\ &= e^{(k-1)x/2} \begin{cases} e^x - 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases} \\ &= e^{(k-1)x/2} \begin{cases} e^x e^{(k-1)x/2} - e^{(k-1)x/2} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases} \\ &= e^{(k-1)x/2} \begin{cases} e^{(k+1)x/2} - e^{(k-1)x/2} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases} \\ &= (e^{(k+1)x/2} - e^{(k-1)x/2})^+ \end{aligned}$$

Likewise we can derive boundary conditions at $x = \pm\infty$ for u from Eqs. (7.18) and (7.19). To summarize then the original Black-Scholes partial differential equation, initial, and boundary conditions have been converted to the following system of equations.

$$u_\tau = u_{xx} \quad \text{for } x \in (-\infty, \infty) \text{ and } \tau \in (0, T\sigma^2/2) \quad (7.25)$$

$$u(x, 0) = (e^{(k+1)x/2} - e^{(k-1)x/2})^+ \quad \text{for } x \in (-\infty, \infty) \quad (7.26)$$

$$u(x, \tau) \rightarrow 0 \quad \text{as } x \rightarrow \pm\infty \text{ for } \tau \in (0, T\sigma^2/2) \quad (7.27)$$

The PDE of Eq. (7.25) is the well-known **heat equation** of mathematical physics. The Fourier Transform technique introduced earlier will now be used to solve this initial boundary value problem.

7.4 Solving the Black-Scholes Equation

We begin by taking the Fourier transform of both sides of the heat equation (since the Black-Scholes equation has been through two changes of variables

and has been converted into the heat equation on the real number line) in Eq. (7.25).

$$\begin{aligned}\mathcal{F}\{u_\tau\} &= \tau \mathcal{F}\{u_{xx}\} \\ \int_{-\infty}^{\infty} u_\tau e^{-iwx} dx &= \int_{-\infty}^{\infty} u_{xx} e^{-iwx} dx \\ \frac{d}{d\tau} \int_{-\infty}^{\infty} u(x, \tau) e^{-iwx} dx &= (iw)^2 \int_{-\infty}^{\infty} u(x, \tau) e^{-iwx} dx \\ \frac{d\hat{u}}{d\tau} &= -w^2 \hat{u} \end{aligned} \quad (7.28)$$

The reader should note that in deriving Eq. (7.28) Theorem 7.1 is used. This last equation is an ordinary differential equation of the type used to model exponential decay. Separating variables (see Sec. 5.3, especially Eq. (5.17) and the paragraphs which follows it) and solving this equation produces a solution of the form

$$\hat{u}(w, \tau) = D e^{-w^2 \tau}.$$

The expression represented by D is any quantity which is constant with respect to τ . Even though this equation was solved using ordinary differential equations techniques, the quantity \hat{u} is a function of both w and τ . To evaluate D we can set $\tau = 0$ and determine that $\hat{u}(w, 0) = D$. Thus D is merely the Fourier transform of the initial condition in Eq. (7.26). For simplicity of notation we will write $D = \hat{f}(w)$. Thus the Fourier transformed solution to the heat equation is

$$\hat{u}(w, \tau) = \hat{f}(w) e^{-w^2 \tau}. \quad (7.29)$$

Now this solution must be inverse Fourier transformed and then have its variables changed back to the original variables of the Black-Scholes equation.

$$\begin{aligned}\mathcal{F}^{-1}\{\hat{u}(w, \tau)\} &= \mathcal{F}^{-1}\{\hat{f}(w) e^{-w^2 \tau}\} \\ u(x, \tau) &= (e^{(k+1)x/2} - e^{(k-1)x/2})^+ * \frac{1}{2\sqrt{\pi\tau}} e^{-x^2/(4\tau)} \\ &= \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} (e^{(k+1)\frac{z}{2}} - e^{(k-1)\frac{z}{2}})^+ e^{-\frac{(x-z)^2}{4\tau}} dz \end{aligned} \quad (7.30)$$

The reader should note that the Fourier convolution theorem was used and also the result of exercise 3. If the substitution $z = x + \sqrt{2\tau}y$ is made then

Eq. (7.30) becomes

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(e^{(k+1)(x+\sqrt{2\tau}y)/2} - e^{(k-1)(x+\sqrt{2\tau}y)/2} \right)^+ e^{-y^2/2} dy \quad (7.31)$$

$$= \frac{e^{(k+1)x/2} e^{(k+1)^2 \tau/4}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{1}{2}(k+1)\sqrt{2\tau})^2/2} dy \quad (7.32)$$

$$- \frac{e^{(k-1)x/2} e^{(k-1)^2 \tau/4}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{1}{2}(k-1)\sqrt{2\tau})^2/2} dy$$

The reader will be asked to fill in the details of this derivation in exercises 7 and 8. The two improper integrals of Eq. (7.32) can be expressed in terms of the cumulative distribution function for a normal random variable. We will make the substitution $w = y - \frac{1}{2}(k+1)\sqrt{2\tau}$ in the first integral and $w' = y - \frac{1}{2}(k-1)\sqrt{2\tau}$ in the second. The first integral of Eq. (7.32) equals the following expression

$$\int_{-\infty}^{x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau}} e^{-w^2/2} dw = \sqrt{2\pi} \phi \left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau} \right)$$

where ϕ is the cumulative distribution function for a normal random variable with mean zero and standard deviation one. The reader should verify this calculation (exercise 9) and a similar calculation for the second improper integral of Eq. (7.32).

Thus far the solution to the initial boundary value problem in Eqs. (7.25)–(7.27) is given by

$$u(x, \tau) = e^{(k+1)x/2 + (k+1)^2 \tau/4} \phi \left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau} \right) \quad (7.33)$$

$$- e^{(k-1)x/2 + (k-1)^2 \tau/4} \phi \left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k-1)\sqrt{2\tau} \right).$$

Now we begin the task of reconverting variables to those of the Black-Scholes initial boundary value problem as stated in Eqs. (7.5)–(7.7). Using the change of variables in Eq. (7.20) this solution can be re-written in terms

of the function $v(x, \tau)$.

$$\begin{aligned} v(x, \tau) &= e^{-(k-1)x/2 - (k+1)^2\tau/4} u(x, \tau) \\ &= e^x \phi \left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau} \right) \\ &\quad - e^{-k\tau} \phi \left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k-1)\sqrt{2\tau} \right) \end{aligned} \quad (7.34)$$

Now using the change of variables described in Eqs. (7.8) and (7.9) the reader can show in exercise 11 that

$$w = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau} = \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \quad (7.35)$$

$$w' = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k-1)\sqrt{2\tau} = w - \sigma\sqrt{T-t}. \quad (7.36)$$

Using these in Eq. (7.34) yields

$$v(x, \tau) = \frac{S}{K} \phi(w) - e^{-r(T-t)} \phi(w - \sigma\sqrt{T-t})$$

which upon using Eq. (7.10) produces the sought after **Black-Scholes European call option pricing formula**. For the sake of simplicity and meaningful notation the value of the European call will be denoted C .

$$C(S, t) = S\phi(w) - Ke^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \quad (7.37)$$

A surface plot of the value of the European call option as a function of S and t is shown in Fig. 7.1.

The value of a European put option could be found via a similar set of calculations or the already determined value of the European call could be used along with the Put-Call Parity formula (6.1) to derive

$$P(S, t) = Ke^{-r(T-t)}\phi(\sigma\sqrt{T-t} - w) - S\phi(-w). \quad (7.38)$$

A surface plot of the value of the European put option as a function of S and t is shown in Fig. 7.2.

Example 7.3 Suppose the current price of a security is \$62 per share. The continuously compounded interest rate is 10% per year. The volatility of the price of the security is $\sigma = 20\%$ per year. The cost of a five-month European call option with a strike price of \$60 per share can be found using Eqs. (7.35) and (7.37). If we summarize the quantities we know, in

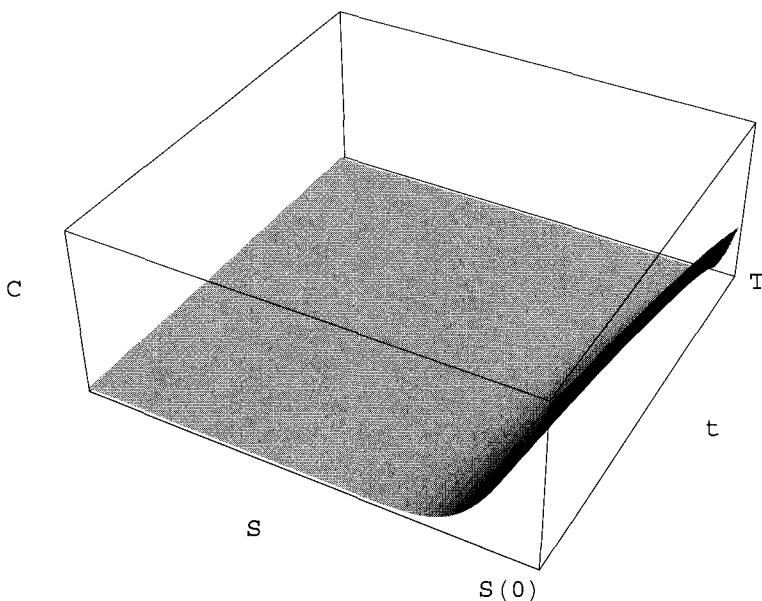


Fig. 7.1 A surface plot of C_e as a function of S and t .

the notation of this section, we have:

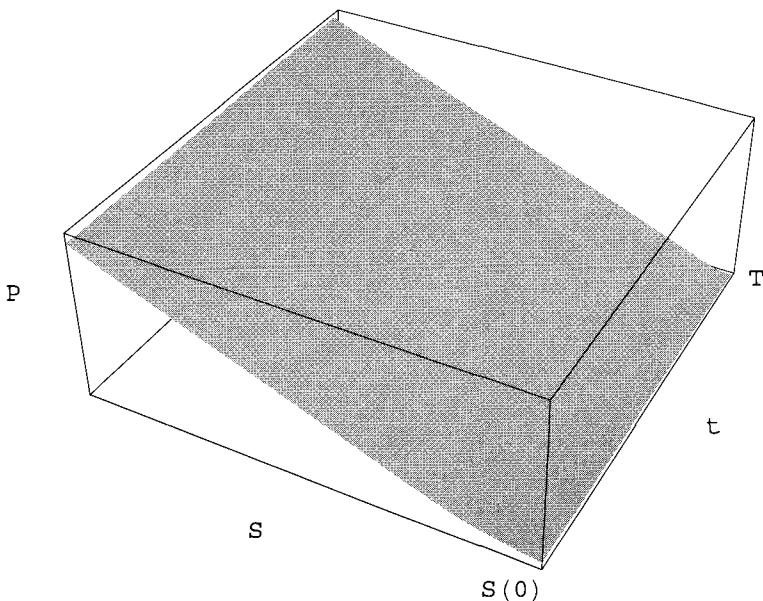
$$T = 5/12, \quad t = 0, \quad r = 0.10, \quad \sigma = 0.20, \quad S = 62, \quad \text{and} \quad K = 60.$$

Thus we have $w \approx 0.641287$ and therefore the price of the European call option is $C = \$5.789$.

Example 7.4 Suppose the current price of a security is \$97 per share. The continuously compounded interest rate is 8% per year. The volatility of the price of the security is $\sigma = 45\%$ per year. The cost of a three-month European put option with a strike price of \$95 per share can be found using Eqs. (7.35), and (7.38). If we summarize the quantities we know, in the notation of this section, we have:

$$T = 1/4, \quad t = 0, \quad r = 0.08, \quad \sigma = 0.45, \quad S = 97, \quad \text{and} \quad K = 95.$$

Thus we have $w \approx 0.293985$ and consequently the put option price is $P = \$6.71343$.

Fig. 7.2 A surface plot of P_e as a function of S and t .

7.5 Exercises

- (1) Let the function $f(x)$ be defined as

$$f(x) = \begin{cases} e^{-ax} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where a is a positive constant. Show that the Fourier Transform of f is $\hat{f}(w) = \frac{1}{a+iw}$.

- (2) Let the function $\hat{f}(w)$ be defined as

$$\hat{f}(w) = \begin{cases} e^{-aw} & \text{if } w \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where a is a positive constant. Show that the inverse Fourier Transform of \hat{f} is $f(x) = \frac{1}{2\pi(a-ix)}$.

- (3) If a is a positive constant verify that the inverse Fourier Transform of e^{-aw^2} is $\frac{1}{2\sqrt{\pi a}}e^{-x^2/4a}$.
- (4) Using the new variables described in Eqs. (7.8), (7.9), (7.10), and the multivariable chain rule, verify the identity shown in Eq. (7.12).

- (5) Using the new variables described in Eqs. (7.8), (7.9), (7.10), the expression for F_S shown in (7.11), and the multivariable chain rule, verify the identity shown in Eq. (7.13).
- (6) Verify that the Black-Scholes partial differential Eq. (7.5) can be simplified to the form shown in Eq. (7.14) by using the variables described in (7.8), (7.9), and (7.10) and the partial derivatives found in (7.11), (7.12), and (7.13).
- (7) Verify that the expression

$$\left(e^{(k+1)(x+\sqrt{2\tau}y)/2} - e^{(k-1)(x+\sqrt{2\tau}y)/2} e^{-y^2/2} \sqrt{2\tau} \right)^+$$

is non-zero whenever $y > -x/\sqrt{2\tau}$.

- (8) Use completion of the square in the exponents of the integrand in Eq. (7.31) to derive Eq. (7.32).
- (9) Using the change of variable $w = -y + \frac{1}{2}(k+1)\sqrt{2\tau}$ in the first improper integral of Eq. (7.32) verify that

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y-\frac{1}{2}(k+1)\sqrt{2\tau})^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau}} e^{-\frac{w^2}{2}} dw \\ &= \phi\left(x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau}\right) \end{aligned}$$

where ϕ is the cumulative distribution function for a normal random variable with mean zero and standard deviation one.

- (10) Using the change of variable $w' = -y + \frac{1}{2}(k-1)\sqrt{2\tau}$ in the second improper integral of Eq. (7.32) verify that

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y-\frac{1}{2}(k-1)\sqrt{2\tau})^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/\sqrt{2\tau} + \frac{1}{2}(k-1)\sqrt{2\tau}} e^{-\frac{w'^2}{2}} dw' \\ &= \phi\left(x/\sqrt{2\tau} + \frac{1}{2}(k-1)\sqrt{2\tau}\right) \end{aligned}$$

where ϕ is the cumulative distribution function for a normal random variable with mean zero and standard deviation one.

- (11) Verify using the change of variable formulas in Eqs. (7.8) and (7.9) and the fact that $k = 2r/\sigma^2$ that

$$x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau} = \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

$$x/\sqrt{2\tau} + \frac{1}{2}(k-1)\sqrt{2\tau} = \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} - \sigma\sqrt{T-t}.$$

- (12) What is the price of a European call option on a stock when the stock price is \$52, the strike price is \$50, the interest rate is 12%, the stock's volatility is 30%, and the exercise time is three months?
- (13) What is the price of a European put option on a stock when the stock price is \$69, the strike price is \$70, the interest rate is 5%, the stock's volatility is 35%, and the exercise time is six months?
- (14) A European call option on a stock has a market value of \$2.50. The stock price is \$15, the strike price is \$13, the interest rate is 5%, and the exercise time is three months. What is the volatility of the stock?
- (15) Show by differentiation and substitution in Eq. (7.5) that the price of the security itself, *i.e.* $f(S, t) = S$ solves the Black-Scholes equation.

This page is intentionally left blank

Chapter 8

Derivatives of Black-Scholes Option Prices

Now that the solution to the Black-Scholes equation is known we can investigate the sensitivity of the solution to changes in the independent variables. Knowledge of these sensitivities allows a portfolio manager to minimize changes in the value of a portfolio when underlying variables such as the risk-free interest rate change. The material of the present chapter will be used extensively in Chapter 9. An understanding of the sensitivity of an option's value to changes in its independent variables will also provide a new way of interpreting the Black-Scholes equation itself.

To mathematicians the word *derivative* means an instantaneous rate of change in a quantity. To a quantitative analyst a *derivative* is a financial entity whose value is derived from the value of some underlying asset. Hence a European call option is a derivative whose value is a function of (among other things) the value of the security underlying the option. In this chapter we explore derivatives from the mathematician's perspective. In the Chapter 9 we will use these derivatives in the manner of a quantitative analyst. Members of the quantitative financial profession refer to the subject matter of this chapter as the "Greeks" since a Greek letter is used to name each derivative (except for one which we will meet in due time).

8.1 Theta

The quantity Θ is defined to be the rate of change of the value of an option with respect to t . We will calculate Θ for the European call and put options derived in Chapter 7. Since the valuation of European-style options depends heavily on the cumulative distribution function of the standard normal random variable and a composite variable w , we will begin by investigating their partial derivatives. The derivative with respect to t of quantity w

defined in Eq. (7.35) is given by

$$\frac{\partial w}{\partial t} = \frac{1}{2\sigma\sqrt{T-t}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right). \quad (8.1)$$

By the Fundamental Theorem of Calculus [Smith and Minton (2002), Chap. 4] the cumulative distribution function for the standard normal distribution has derivative

$$\phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (8.2)$$

Hence using the Chain Rule [Stewart (1999), Chap. 3] for derivatives we can see that the time rate of change in the value of a European call option (defined in Eq. (7.37)) is

$$\begin{aligned} \frac{\partial C}{\partial t} &= S\phi'(w)\frac{\partial w}{\partial t} - Ke^{-r(T-t)} \left(r\phi(w - \sigma\sqrt{T-t}) \right. \\ &\quad \left. + \phi'(w - \sigma\sqrt{T-t}) \left[\frac{\partial w}{\partial t} + \frac{\sigma}{2\sqrt{T-t}} \right] \right) \\ &= \left(S\phi'(w) - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \right) \frac{\partial w}{\partial t} \\ &\quad - Ke^{-r(T-t)} \left(r\phi(w - \sigma\sqrt{T-t}) + \frac{\sigma}{2\sqrt{T-t}}\phi'(w - \sigma\sqrt{T-t}) \right) \\ &= \frac{S\phi'(w) - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t})}{2\sigma\sqrt{T-t}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\ &\quad - Ke^{-r(T-t)} \left(r\phi(w - \sigma\sqrt{T-t}) + \frac{\sigma}{2\sqrt{T-t}}\phi'(w - \sigma\sqrt{T-t}) \right) \\ \Theta_C &= \frac{Se^{-w^2/2} - Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\ &\quad - Ke^{-r(T-t)} \left(r\phi(w - \sigma\sqrt{T-t}) + \frac{\sigma e^{-(w-\sigma\sqrt{T-t})^2/2}}{2\sqrt{2\pi(T-t)}} \right) \end{aligned} \quad (8.3)$$

To take an example, consider a European call option on a stock whose current price is \$250. The strike price is \$245, the annual risk-free interest rate is 2.5%, the volatility of the stock price is 20%, and the strike time is four months. Thus substituting $S = 250$, $K = 245$, $r = 0.025$, $\sigma = 0.20$, $T = 1/3$, and $t = 0$ into Eq. (8.3) produces

$$\Theta = \frac{\partial C}{\partial t} = -19.9836.$$

The case that Θ is negative for an option is typical since options lose value as the expiry date approaches.

The calculation of Θ for a European put option is carried out in a similar fashion to that for the call option. The reader will be asked to provide the details of the following derivation.

$$\Theta_P = \frac{\partial P}{\partial t} \quad (8.4)$$

$$\begin{aligned} &= \frac{Se^{-w^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\ &\quad + K e^{-r(T-t)} \phi(\sigma\sqrt{T-t} - w) \\ &\quad - \frac{Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r + \sigma^2/2 \right) \end{aligned} \quad (8.5)$$

Equation (8.5) will be put to use in the following example. Consider a European put option on a stock whose current price is \$325. The strike price is \$330, the annual risk-free interest rate is 3.5%, the volatility of the stock price is 27%, and the strike time is three months. Thus substituting $S = 325$, $K = 330$, $r = 0.035$, $\sigma = 0.27$, $T = 1/4$, and $t = 0$ into Eq. (8.5) produces

$$\Theta = \frac{\partial P}{\partial t} = -28.7484.$$

The derivative Θ , is unique among the derivatives covered in this chapter. The passage of time t , is the only non-stochastic variable among the list of independent variables upon which C or P depend.

8.2 Delta

The value of an option is also sensitive to changes in the price of the underlying stock. This rate of change is called Delta, Δ . The astute reader may recall that Δ was introduced in Sec. 6.3 during the derivation of the Black-Scholes partial differential equation. Delta was used to eliminate a stochastic term from Eq. (6.12). The clever choice of Δ left us with a deterministic partial differential equation to solve rather than a stochastic differential equation. The starting point for deriving Δ is again Eq. (7.37).

$$\frac{\partial C}{\partial S} = \phi(w) + \left(S\phi'(w) - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \right) \frac{\partial w}{\partial S} \quad (8.6)$$

We can make use of Eq. (8.2) and the following result.

$$\frac{\partial w}{\partial S} = \frac{1}{\sigma S \sqrt{T-t}} \quad (8.7)$$

Substituting Eqs. (8.2) and (8.7) into Eq. (8.6) produces

$$\begin{aligned} \frac{\partial C}{\partial S} &= \phi(w) + \left(\frac{Se^{-w^2/2}}{\sqrt{2\pi}} - \frac{Ke^{-r(T-t)}e^{-(w-\sigma\sqrt{T-t})^2/2}}{\sqrt{2\pi}} \right) \frac{1}{\sigma S \sqrt{T-t}} \\ &= \phi(w) + \frac{e^{-w^2/2}}{\sigma \sqrt{2\pi(T-t)}} \left(1 - \frac{Ke^{-r(T-t)}}{S} e^{-(\sigma^2(T-t)-2\sigma w \sqrt{T-t})/2} \right) \\ &= \phi(w) + \frac{e^{-w^2/2}}{\sigma \sqrt{2\pi(T-t)}} \times \\ &\quad \left(1 - \frac{Ke^{-r(T-t)}}{S} e^{-(\sigma^2(T-t)-2(r+\sigma^2/2)(T-t)+2\ln(K/S))/2} \right) \\ &= \phi(w) + \frac{e^{-w^2/2}}{\sigma \sqrt{2\pi(T-t)}} \left(1 - \frac{Ke^{-r(T-t)}}{S} e^{r(T-t)+\ln(S/K)} \right) \\ &= \phi(w) + \frac{e^{-w^2/2}}{\sigma \sqrt{2\pi(T-t)}} \left(1 - \frac{K}{S} e^{\ln(S/K)} \right) \\ &= \phi(w) + \frac{e^{-w^2/2}}{\sigma \sqrt{2\pi(T-t)}} (1 - 1) \\ &= \phi(w) \end{aligned} \quad (8.8)$$

We can put this formula to use in the following example. Suppose the price of a security is \$100, the risk-free interest rate is 4% per annum, the annual volatility of the stock price is 23%, the strike price is \$110, and the strike time is 6 months. Under these conditions $w = -0.381747$ and

$$\Delta_C = \frac{\partial C}{\partial S} = 0.351325.$$

Delta for a European put option could be calculated directly from the definition for P ; however, the Put-Call parity formula (Eq. (6.1)) provides a convenient shortcut.

$$\begin{aligned} \frac{\partial}{\partial S}(P+S) &= \frac{\partial}{\partial S}(C+Ke^{-r(T-t)}) \\ \frac{\partial P}{\partial S} + 1 &= \frac{\partial C}{\partial S} \\ \Delta_P = \frac{\partial P}{\partial S} &= \phi(w) - 1 \end{aligned} \quad (8.9)$$

Thus using the same parameter values as in the previous example,

$$\frac{\partial P}{\partial S} = 0.351325 - 1 = -0.648675.$$

So as the European call option increases in value with an increase in the value of the underlying security, the put option decreases in value.

8.3 Gamma

The **gamma** Γ , of an option or collection of options is defined to be the second derivative of the option with respect to S , the price of the underlying security. Hence Γ is the partial derivative of Δ with respect to S . Thus for a European call option Γ is

$$\begin{aligned}\frac{\partial^2 C}{\partial S^2} &= \frac{\partial}{\partial S} \left(\frac{\partial C}{\partial S} \right) \\ &= \frac{\partial}{\partial S} (\phi(w)) \quad (\text{using Eq. (8.8)}) \\ &= \phi'(w) \frac{\partial w}{\partial S} \\ &= \frac{e^{-w^2/2}}{\sqrt{2\pi}} \left(\frac{1}{\sigma S \sqrt{T-t}} \right) \quad (\text{using exercise 4}) \\ &= \frac{e^{-w^2/2}}{\sigma S \sqrt{2\pi(T-t)}}\end{aligned}\tag{8.10}$$

We can readily see from Eq. (8.9) that

$$\Gamma = \frac{\partial^2 P}{\partial S^2} = \frac{\partial^2 C}{\partial S^2} = \frac{e^{-w^2/2}}{\sigma S \sqrt{2\pi(T-t)}}.\tag{8.11}$$

If all independent variables are held fixed except for S , then Γ measures the concavity of the value of the option as a function of S . Suppose we consider a European call option on a security whose value is \$295. The strike price is \$290, the annual risk-free interest rate is 4.25%, the volatility of the security is 25% per annum, and the expiry date is two months. We can calculate the price of the option, its Δ , and its Γ as follows.

$$C = 15.7173, \quad \Delta_C = 0.613297, \quad \Gamma = 0.0127122$$

Thus the value of the option will increase with the price of the underlying security at an increasing rate. As the price of the security increases, the

price of the option will increase at an increasing rate. This will have important implications for the practice known as hedging which will be explored in Chapter 9. In terms of elementary calculus, the graph of C as a function of S (holding all other variables fixed) is increasing and concave upward on an interval containing $S = 295$. See Fig. 8.1.

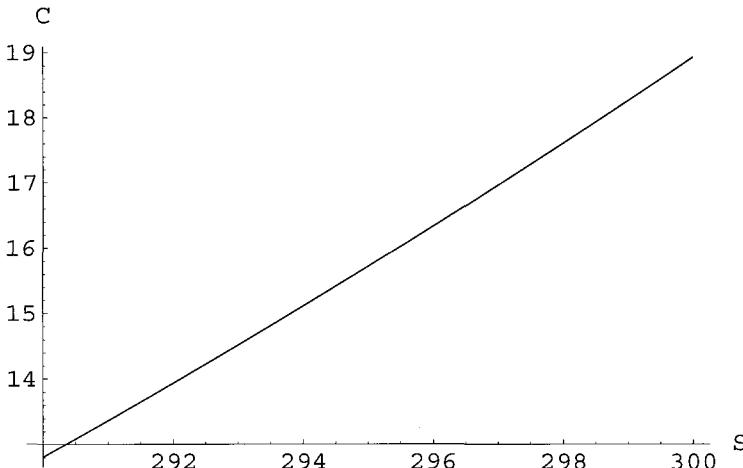


Fig. 8.1 Since the function $C(S)$ is increasing and concave upward on an interval containing $S = 295$ the price of the European call option increases at an increasing rate when the price of the underlying security is near \$295.

8.4 Vega

The **vega** of an option is the change in the value of the option as a function of the volatility of the underlying security. Among the partial derivatives of the value of an option vega distinguishes itself as the only partial derivative given a non-Greek letter name. Vega comes to us through medieval Latin from Arabic.

For a European call option defined as in Eq. (7.37),

$$\frac{\partial C}{\partial \sigma} = S\phi'(w)\frac{\partial w}{\partial \sigma} - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t})\left(\frac{\partial w}{\partial \sigma} - \sqrt{T-t}\right), \quad (8.12)$$

where, according to exercise 8,

$$\frac{\partial w}{\partial \sigma} = \sqrt{T-t} - w/\sigma. \quad (8.13)$$

Substituting Eq. (8.13) into Eq. (8.12) produces

$$\begin{aligned}
 \frac{\partial C}{\partial \sigma} &= S\phi'(w) \left(\sqrt{T-t} - \frac{w}{\sigma} \right) \\
 &\quad - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \left[\left(\sqrt{T-t} - \frac{w}{\sigma} \right) - \sqrt{T-t} \right] \\
 &= S\phi'(w) \left(\sqrt{T-t} - \frac{w}{\sigma} \right) + Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \left(\frac{w}{\sigma} \right) \\
 &= S\sqrt{T-t}\phi'(w) - \frac{w}{\sigma} \left(S\phi'(w) - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \right) \\
 &= S\sqrt{T-t}\phi'(w) - \frac{w}{\sigma\sqrt{2\pi}} \left(Se^{-w^2/2} - Ke^{-r(T-t)}e^{-(w-\sigma\sqrt{T-t})^2/2} \right) \\
 &= S\sqrt{T-t}\phi'(w) \\
 &\quad - \frac{w}{\sigma\sqrt{2\pi}} e^{-w^2/2} \left(S - Ke^{-r(T-t)+w\sigma\sqrt{T-t}-\sigma^2(T-t)/2} \right) \\
 &= S\sqrt{T-t}\phi'(w) - \frac{w}{\sigma\sqrt{2\pi}} e^{-w^2/2} \\
 &\quad \times \left(S - Ke^{-r(T-t)+\ln(S/K)+(r+\sigma^2/2)(T-t)-\sigma^2(T-t)/2} \right) \\
 &= S\sqrt{T-t}\phi'(w) - \frac{w}{\sigma\sqrt{2\pi}} e^{-w^2/2} \left(S - Ke^{\ln(S/K)} \right) \\
 &= S\sqrt{T-t}\phi'(w) - \frac{w}{\sigma\sqrt{2\pi}} e^{-w^2/2}(S - S) \\
 &= \frac{S\sqrt{T-t}}{\sqrt{2\pi}} e^{-w^2/2} \tag{8.14}
 \end{aligned}$$

Note that the value of the European call option increases with increasing volatility in the underlying security. It is easily seen from the Put-Call Parity Formula (Eq. (6.1)) that

$$\frac{\partial P}{\partial \sigma} = \frac{S\sqrt{T-t}}{\sqrt{2\pi}} e^{-w^2/2}, \tag{8.15}$$

in other words the rate of change in the value of a European put option with respect to volatility is the same as the rate of change in the value of the corresponding European call option with respect to volatility. Thus we may unambiguously define vega for a European style option to be

$$\mathcal{V} = \frac{S\sqrt{T-t}}{\sqrt{2\pi}} e^{-w^2/2}. \tag{8.16}$$

Thus the vega for a European style option (either call or put) on a security who current value is \$160, whose strike price is \$150, and whose expiry

date is 5 months can be calculated using Eq. (8.16). Suppose the volatility of the security is 20% per year and the risk-free interest rate is 2.75% per year. Then $w = 0.653219$ and $\mathcal{V} = 33.2866$.

8.5 Rho

The **rho** of an option is the rate of change in the value of the option as a function of the risk-free interest rate r . Once again the European call option will be used to demonstrate the calculation of rho. Using Eq. (7.37) it is seen that

$$\frac{\partial C}{\partial r} = \left(S\phi'(w) - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \right) \frac{\partial w}{\partial r} + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \quad (8.17)$$

Calling on the definition of w given in Eq. (7.35) we see that

$$\frac{\partial w}{\partial r} = \frac{\sqrt{T-t}}{\sigma}. \quad (8.18)$$

Substituting Eq. (8.18) into Eq. (8.17) yields

$$\begin{aligned} \frac{\partial C}{\partial r} &= \left(S\phi'(w) - Ke^{-r(T-t)}\phi'(w - \sigma\sqrt{T-t}) \right) \frac{\sqrt{T-t}}{\sigma} \\ &\quad + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \\ &= \left(Se^{-w^2/2} - Ke^{-r(T-t)}e^{-(w-\sigma\sqrt{T-t})^2/2} \right) \frac{\sqrt{T-t}}{\sigma\sqrt{2\pi}} \\ &\quad + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \\ &= \left(S - Ke^{-r(T-t)+w\sigma\sqrt{T-t}-(T-t)\sigma^2/2} \right) \frac{e^{-w^2/2}\sqrt{T-t}}{\sigma\sqrt{2\pi}} \\ &\quad + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \\ &= \left(S - Ke^{-r(T-t)+\ln(S/K)+(r+\sigma^2/2)(T-t)-(T-t)\sigma^2/2} \right) \frac{e^{-w^2/2}\sqrt{T-t}}{\sigma\sqrt{2\pi}} \\ &\quad + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \\ &= \left(S - Ke^{\ln(S/K)} \right) \frac{e^{-w^2/2}\sqrt{T-t}}{\sigma\sqrt{2\pi}} + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \\ &= (S - S)\frac{e^{-w^2/2}\sqrt{T-t}}{\sigma\sqrt{2\pi}} + K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \\ \rho_C &= K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) \end{aligned} \quad (8.19)$$

Consider the situation of a European call option on a security whose current value is \$225. The strike price is \$230, while the strike time is 4 months. The volatility of the price of the security is 27% annually and the risk-free interest rate is 3.25%. Under these conditions, $w = 0.0064434$ and

$$\frac{\partial C}{\partial r} = 33.4156.$$

The rho for a European put option can be determined from the Put-Call Parity Formula (Eq. (6.1)) and Eq. (8.19).

$$\begin{aligned}\frac{\partial P}{\partial r} &= \frac{\partial C}{\partial r} - K(T-t)e^{-r(T-t)} \\ &= K(T-t)e^{-r(T-t)}\phi(w - \sigma\sqrt{T-t}) - K(T-t)e^{-r(T-t)} \\ &= -K(T-t)e^{-r(T-t)}(1 - \phi(w - \sigma\sqrt{T-t})) \\ \rho_P &= -K(T-t)e^{-r(T-t)}\phi(\sigma\sqrt{T-t} - w)\end{aligned}\tag{8.20}$$

In the last equation above we made use of the fact that $\phi(x) = 1 - \phi(-x)$ for the cumulative distribution function of a normally distributed random variable.

To put Eq. (8.20) to work in an example, consider the situation of a European put option on a security whose underlying value is \$150. The strike price is \$152, the volatility is 15% per annum, the expiry date is three months, and the risk-free interest rate is 2.5%. Then $w = -0.0557697$ and

$$\frac{\partial P}{\partial r} = -20.8461.$$

8.6 Relationships Between Δ , Θ , and Γ

The value F of an option on a security S must satisfy the Black-Scholes partial differential Eq. (6.14). This equation is repeated below for convenience.

$$rF = F_t + rSF_S + \frac{1}{2}\sigma^2 S^2 F_{SS}$$

This equation assumes the risk-free interest rate is r and the volatility of the security is σ . In this chapter we have given names to the partial derivatives F_t , F_S , and F_{SS} , namely

$$\Theta = F_t, \quad \Delta = F_S, \quad \Gamma = F_{SS}.$$

Table 8.1 Listing of the derivatives of European option values and some of their properties.

Name	Symbol	Definition	Property
Theta	Θ_C	Eq. (8.3)	
	Θ_P	Eq. (8.5)	
Delta	Δ_C	$\phi(w)$	$\Delta_C > 0$
	Δ_P	$\phi(w) - 1$	$\Delta_P < 0$
Gamma	Γ	Eq. (8.10)	$\Gamma > 0$
Vega	\mathcal{V}	Eq. (8.16)	$\mathcal{V} > 0$
Rho	ρ_C	Eq. (8.19)	$\rho_C > 0$
	ρ_P	Eq. (8.20)	$\rho_P < 0$

Thus the Black-Scholes partial differential equation can be re-written as

$$rF = \Theta + rS\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma. \quad (8.21)$$

The value of the option is a linear combination of Δ , Θ , and Γ .

When the value of the option is insensitive to the passage of time $\Theta = 0$ and Eq. (8.21) becomes

$$F = S \left(\Delta + \frac{\sigma^2 S}{2r} \Gamma \right).$$

If Δ should be of large magnitude then Γ must be large and of opposite sign. Returning to Eq. (8.21), if the rate of change in the value of the option with respect to the value underlying security is zero (or at least very small) then the Black-Scholes equation becomes

$$rF = \Theta + \frac{1}{2}\sigma^2 S^2 \Gamma.$$

Thus Θ and Γ tend to be of opposite algebraic sign. In practice, calculation of Θ can be used to approximate Γ .

Table 8.1 summarizes the option value derivatives discussed in this chapter and lists some of their relevant properties. This table is provided as a convenient place to find the definitions of these quantities in the future.

In the next chapter the practice of hedging will be described. Hedging is performed on a **portfolio** of securities, a collection which may include stocks, put and call options, cash in savings accounts, etc. The mathematical aspect of hedging is made possible by the linearity property of the Black-Scholes partial differential equation. An operator $L[\cdot]$ is linear if $L[aX + Y] = aL[X] + L[Y]$ for all scalars a and vectors X, Y in the domain of L . In the case of the Black-Scholes PDE the vectors are solution

functions. If we define the operator $L[\cdot]$ to be

$$L[X] = \frac{\partial X}{\partial t} + rS \frac{\partial X}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 X}{\partial S^2} - rX,$$

then any solution X , to the Black-Scholes partial differential Eq. (6.14) will satisfy $L[X] = 0$.

8.7 Exercises

- (1) Find Θ for a European call option on a stock whose current value is \$300. The annual risk-free interest rate is 3%. The strike price is \$310 while the strike time is three months. The volatility of the stock is 25% annually.
- (2) Carefully work through the details of the derivation of Θ for a European put option. Confirm that you obtain the result in Eq. (8.5).
- (3) Find Θ for a European put option on a stock whose current value is \$275. The annual risk-free interest rate is 2%. The strike price is \$265 while the expiry date is four months. The volatility of the stock is 20% annually.
- (4) Find the partial derivative of w (as defined in Eq. 7.35) with respect to S .
- (5) Find Δ for a European call option on a stock whose current value is \$150. The annual risk-free interest rate is 2.5%. The strike price is \$165 while the strike time is five months. The volatility of the stock is 22% annually.
- (6) Find $\partial P/\partial S$ for a European put option on a stock whose current value is \$125. The annual risk-free interest rate is 5.5%. The strike price is \$140 while the expiry date is eight months. The volatility of the stock is 15% annually.
- (7) Calculate Γ for a European put option for a security whose current price is \$180. The strike price is \$175, the annual risk-free interest rate is 3.75%, the volatility in the price of the security is 30% annually, and the strike time is four months.
- (8) Using partial differentiation derive Eq. (8.13) from Eq. (7.35).
- (9) Calculate the \mathcal{V} of a European style option on a security who current value is \$300, whose strike price is \$305, and whose expiry date is 6 months. Suppose the volatility of the security is 25% per year and the risk-free interest rate is 4.75% per year.

- (10) Calculate the rho of a European style put option on a security who current value is \$270, whose strike price is \$272, and whose strike time is 2 months. Suppose the volatility of the security is 15% per year and the risk-free interest rate is 3.75% per year.
- (11) Calculate the rho of a European style call option on a security who current value is \$305, whose strike price is \$325, and whose expiry date is 4 months. Suppose the volatility of the security is 35% per year and the risk-free interest rate is 2.55% per year.

Chapter 9

Hedging

Hedging is the practice of making a portfolio of investments less sensitive to changes in market variables such as the prices of securities and interest rates. If $F(S, t)$ is a solution to the Black-Scholes partial differential Eq. (6.14), the quantity $\Delta = F_S$ is central to hedging a portfolio. In fact this same quantity Δ , was used to derive the Black-Scholes PDE. The Black-Scholes equation can be thought of as the statement that when Δ is zero, the rate of return from a portfolio consisting of ownership of the underlying security and sale of the option (or the opposite positions) should be the same as the rate of return from the risk-free interest rate on the same net amount of cash.

In Chapter 8 the partial derivatives of the values of European put and call options were calculated. In this chapter they will be used to hedge portfolios of investments.

9.1 General Principles

Before launching into a discussion of hedging, let us pause and recall what happens when a one entity (say, a bank) sells a call option on a stock to another entity (say, an investor). The bank has promised the investor that they will be able to purchase the stock at a predetermined price (the strike price), at a predetermined time (the expiry date) in the future. The stock must be available to the investor at the strike price even if the market value of the stock exceeds the strike price. The bank is in a favorable position if, at the strike time, the market price of the stock is below the strike price set when the call option was sold to the investor. In this case the call option will not be exercised and the bank keeps the money it received when it sold the option. Conversely, if the market price of the stock exceeds the strike

price of the option, the bank must ensure the investor can find a seller of the stock who will agree to accept the strike price. One way to accomplish this would be for the bank itself to be the seller to the investor.

Consider this strategy: the bank creates a European call option on a stock whose current price is \$50 while the strike price is \$52, the risk-free interest rate is 2.5% per annum, the expiry date is four months, and the volatility of the stock price is 22.5% per annum. According to the Black-Scholes option pricing formula, the value of this European call option is \$1.91965. So that we deal with whole numbers, suppose an investor purchases one million of these call options. The bank accepts revenue equal to \$1,919,650. The bank could wait until the expiry date to purchase the one million shares of the stock, or they could purchase one million shares at the time the investor bought the million call options. Suppose the bank takes the latter course, they will expend \$50M, considerably more than the revenue generated by the sale of the call options. If S represents the price per share of the stock at the strike time then the net revenue generated by the transaction obeys the formula

$$1919650 + (\min\{52, S\}e^{-0.025/3} - 50) \cdot 10^6.$$

Notice that the present value of the stock price must be incorporated rather than just S . The reader can check that when $S \approx 48.4827$ the net expenditure is zero. Figure 9.1 shows if the stock is below that price the bank loses money (at $S = 46$ the loss comes to more than \$2.4M), and if the stock price is higher the bank profits (at $S = 52$ the profit nearly \$3.5M). Since the strike price is \$52, the maximum net revenue for the bank occurs at that price.

Another option for the bank is to purchase the stock at the strike time and instantaneously sell it to the investor at the strike price. In this scenario the net revenue generated by this sequence of transactions is

$$1919650 + \min\{0, 52 - S\}e^{-0.025/3} \cdot 10^6.$$

The net revenue is zero when S is approximately \$53.9357. As long as the stock price remains below the strike price the bank keeps its revenue from the sale of the options. However the bank's losses could be dramatic when the price of the stock rises before the expiry date. At a price of \$56 per share the bank's net loss would be approximately \$2.0M. Figure 9.2 illustrates the profit or loss to the seller of the European call option.

Neither of these schemes are practical for hedging portfolios in the fi-

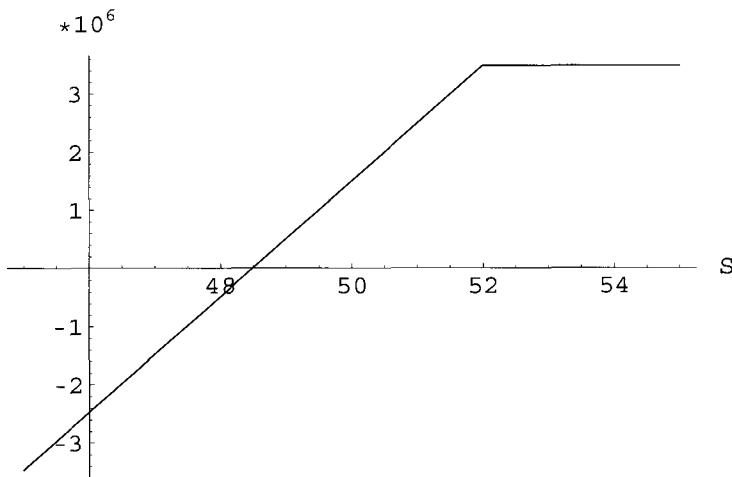


Fig. 9.1 The line shows the profit or loss to the seller of a European call option who adopts a covered position by purchasing the underlying security at the time they sell the option.

nancial world due to their potentially large costs. In the next section a simple, practical method for hedging is explored.

9.2 Delta Hedging

The purpose of hedging is to eliminate or at least minimize the change in the value of an investor's or an institution's portfolio of investments as conditions change in the financial markets. One variable subject to change is the value of the stock or security underlying an option. The reader will recall from the previous chapter that Delta is the partial derivative of the value of an option with respect to the value of the underlying security. In Sec. 8.2 expressions for Delta for European call and put options were derived. One can think of Delta in the following way: for every unit change in the value of the underlying security, the value of the option changes by Δ . A portfolio consisting of securities and options is called **Delta-neutral** if for every option sold, Δ units of the security are bought.

Example 9.1 Consider the case of a security whose price is \$100, while the risk-free interest rate is 4% per annum, the annual volatility of the stock price is 23%, the strike price is set at \$105, and the expiry date is 3 months. Under these conditions $w = -0.279806$, the value of a European call option

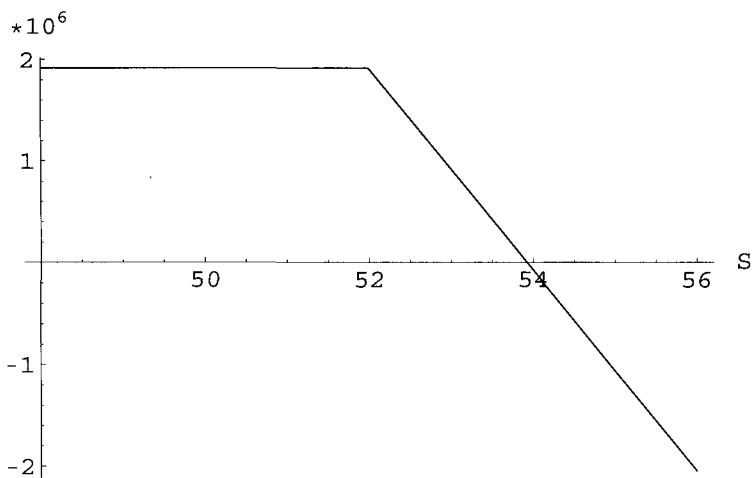


Fig. 9.2 The line shows the profit or loss to the seller of a European call option who adopts a naked position by purchasing the underlying security at the expiry date of the option.

is $C = 2.96155$ and Delta for the option is

$$\Delta = \frac{\partial C}{\partial S} = 0.389813.$$

Thus if a firm sold an investor European call options on ten thousand shares of the security¹, the firm would receive \$29615.50 and purchase $\$389813 = (10000)(0.389813)(100)$ worth of the security, most likely with borrowed money.

The firm may choose to do nothing further until the strike time of the option. In that case, this is referred to as a “hedge and forget” scheme. On the other hand, since the price of the security is dynamic, the firm may choose to make periodic adjustments to the number shares of the security it holds. This strategy is known as **rebalancing** the portfolio. The example begun above can be extended to include weekly rebalancing.

Assume that the value of the security follows the random walk illustrated in Fig. 9.3. In this case the European call option will be exercised since the price of the security at the expiry date exceeds the strike price of \$105. At the end of the first week the value of the security has declined to \$98.79.

¹Typically an option is keyed to 100 shares of the underlying security, so in practice options on 10000 shares would amount to 100 options. To avoid confusion we will assume a one-to-one ratio of options to shares.

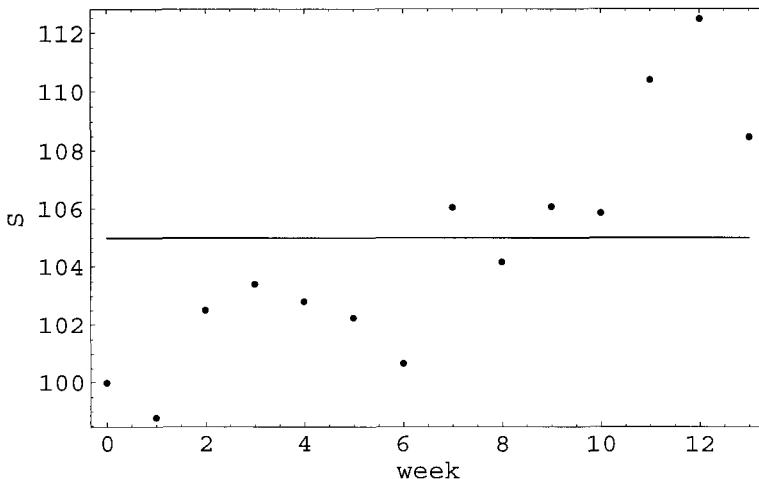


Fig. 9.3 A realization of a random walk taken by the value of a security. The horizontal line indicates the strike price of a European call on the security.

Upon re-computing, $\Delta = 0.339811$. Thus the investment firm would adjust its security holdings so that it now owned 3398 shares of the security with a total value of \$335699. Also at the end of the first week the firm will have incurred costs in the form of interest on the money initially borrowed to purchase shares of the security. This cost amounts to

$$(389813)(e^{0.04/52} - 1) = \$299.97.$$

The interest for the current is added to the cumulative cost entry of the next week. Table 9.1 summarizes the weekly rebalancing up to the expiry date. Notice at the expiry date the investment firm has in its portfolio 10000 shares of the security for which the investor will pay a total of \$1,050,000. Thus the net proceeds to the investment firm of selling the call option and hedging this position are

$$1050000 + 29615.50 - 1069460 = \$10155.50.$$

In other words, the firm profits from issuing the call option. The reader should note that this assumes the investor does not pay for the call option until the strike time.

Alternatively the value of the security may evolve in such a fashion that the call option is not exercised. Such a scenario is depicted in Fig. 9.4. Since the value of the security finishes below the strike price, the call option

Table 9.1 Delta hedging using portfolio rebalancing at weekly intervals.

Week	S	Δ	Shares Held	Interest Cost	Cumulative Cost
0	100.00	0.389813	3898	300	389800
1	98.79	0.339811	3398	262	340705
2	102.52	0.462922	4629	359	467169
3	103.41	0.490192	4902	382	495760
4	102.82	0.460541	4605	358	465604
5	102.25	0.428236	4282	333	432935
6	100.67	0.347145	3471	271	351625
7	106.05	0.589204	5892	468	608643
8	104.17	0.491348	4913	390	507129
9	106.08	0.595047	5950	475	617524
10	105.86	0.585915	5859	468	608366
11	110.40	0.878690	8787	717	932085
12	112.46	0.985811	9858	811	1053247
13	108.47	1.0	10000	0	1069460

expires unused. Table 9.2 summarizes the weekly rebalancing activity of the investment firm as it practices Delta hedging. Notice that at the expiry date the investment firm owns no shares of the security (none are needed since the call option will not be exercised). The net proceeds to the investment firm of selling the call option and hedging its position are

$$29615.50 - 29669 = -\$53.50.$$

In this case the firm loses a small amount of money on the collection of all these transactions.

Rebalancing of the portfolio can take place either more or less frequently than weekly as was done in the two previous extended examples. The quantity discussed in Sec. 8.3 known as Gamma (Γ) is the second partial derivative of the portfolio with respect to the value of the underlying security. In other words Gamma is the rate of change of Delta with respect to S . If $|\Gamma|$ is large then Δ changes rapidly with relatively small changes in S . In this case frequent rebalancing of the investment firm's position may be necessary. If $|\Gamma|$ is small then Δ is relatively insensitive to changes in S and hence rebalancing may be needed infrequently. Thus an investment firm can monitor Gamma in order to determine the frequency at which the firm's position should be rebalanced.

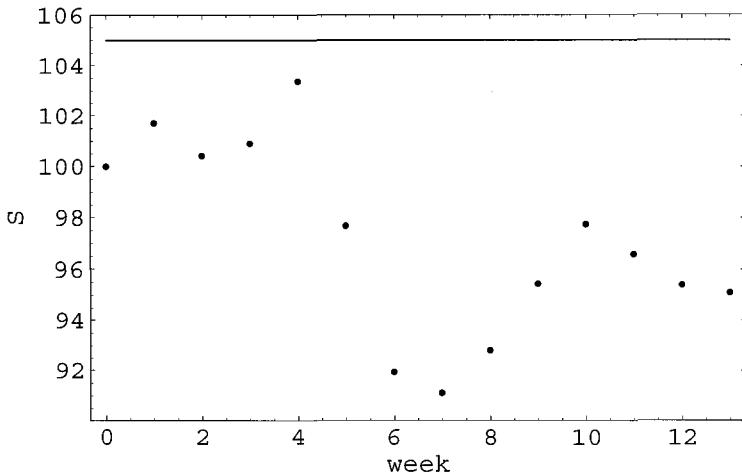


Fig. 9.4 Another realization of a random walk taken by the value of a security. The horizontal line indicates the strike price of a European call on the security.

9.3 Delta Neutral Portfolios

In exercise 15 of Chapter 7 the reader was asked to verify that a stock or security itself satisfies the Black-Scholes PDE (6.14). Now suppose that a portfolio consists of a linear combination of options and shares of the underlying security. The portfolio contains a short position in a European call option and a long position in the security (hedged appropriately as described in the previous section). Thus the net value \mathcal{P} of the portfolio is

$$\mathcal{P} = C - \Delta S = C - \frac{\partial C}{\partial S} \Big|_{S_0} S, \quad (9.1)$$

where S_0 is the price of the security at the time the hedge is created. The quantity \mathcal{P} satisfies the Black-Scholes equation since C and S separately do, and the equation is linear. Thus it reasonable to contemplate Delta for the portfolio. The partial derivative of the entire portfolio with respect to S represents the sensitivity of the value of the portfolio to changes in S . Differentiating both sides of Eq. (9.1) with respect to S produces

$$\frac{\partial \mathcal{P}}{\partial S} = \frac{\partial C}{\partial S} - \frac{\partial C}{\partial S} \Big|_{S_0}$$

This quantity equals zero whenever $S = S_0$ (for example at the moment the hedge is set up) and remains very close to zero so long as S is near S_0 . For

Table 9.2 Delta hedging using portfolio rebalancing at weekly intervals for an option which will expire unused.

Week	S	Δ	Shares Held	Interest Cost	Cumulative Cost
0	100.00	0.389813	3898	300	389800
1	101.71	0.440643	4406	340	441769
2	100.43	0.386757	3868	299	388077
3	100.91	0.394649	3946	305	396247
4	103.37	0.482725	4827	375	487621
5	97.69	0.246176	2462	198	256959
6	91.95	0.071229	712	74	96244
7	91.12	0.043022	430	54	70623
8	92.81	0.050427	504	60	77545
9	95.45	0.078574	786	80	104521
10	97.75	0.110154	1102	104	135491
11	96.58	0.036209	362	49	64126
12	95.40	0.001508	15	24	31072
13	95.10	0.0	0	0	29669

this reason a portfolio which is hedged using Delta hedging is sometimes called **Delta neutral**.

Assuming that the risk-free interest rate remains constant and the volatility of the security does not change then a Taylor series expansion of the value of the portfolio in terms of t and S yields

$$\begin{aligned}\mathcal{P} &= \mathcal{P}_0 + \frac{\partial \mathcal{P}}{\partial t}(t - t_0) + \frac{\partial \mathcal{P}}{\partial S}(S - S_0) + \frac{\partial^2 \mathcal{P}}{\partial S^2} \frac{(S - S_0)^2}{2} + \dots \\ \delta \mathcal{P} &= \Theta \delta t + \Delta \delta S + \frac{1}{2} \Gamma(\delta S)^2 + \dots\end{aligned}$$

All omitted terms in the Taylor series involve powers of δt greater than 1. The Gamma term is retained since the stochastic random variable S follows a stochastic process dependent on $\sqrt{\delta t}$, see Eq. (5.38). If the portfolio is hedged using Delta hedging then Δ for the portfolio is zero and thus

$$\delta \mathcal{P} \approx \Theta \delta t + \frac{1}{2} \Gamma(\delta S)^2, \quad (9.2)$$

where the approximation omits terms involving powers of δt greater than 1. The term involving Θ is not stochastic and thus must be retained; however, the approximation can be further refined if the portfolio can be made **Gamma neutral**, *i.e.* if the makeup of the portfolio can be adjusted so that $\Gamma = 0$.

9.4 Gamma Neutral Portfolios

Gamma for a portfolio is the second partial derivative of the portfolio with respect to S . The portfolio cannot be made gamma neutral using a linear combination of just an option and its underlying security, since the second derivative of S with respect to itself is zero. As explained in the previous section, the portfolio can be made Delta neutral with the appropriate combination of the option and the underlying security. The portfolio can be made Gamma neutral by manipulating a component of the portfolio which depends non-linearly on S . One such component is the option. However, as mentioned earlier, the portfolio cannot contain only the option and its underlying security. One way to achieve the goal of a Gamma neutral portfolio is to include in the portfolio two (or perhaps more) different types of option dependent on the same underlying security. For example suppose the portfolio contains options with two different strike times written on the same stock. An investment firm may sell a number of European call options with the expiry date three months hence and buy some other number of the European call options on the same stock with an expiry date arriving in six months. Let the number of the early option sold be w_e and the number of the later option be w_l . The Gamma of the portfolio would be

$$\Gamma_P = w_e \Gamma_e - w_l \Gamma_l,$$

where Γ_e and Γ_l denote the Gammas of the earlier and later options respectively. The relative weights of the earlier and later options can be chosen so that $\Gamma_P = 0$. Once the portfolio is made Gamma neutral, the underlying security can be added to the portfolio in such a way as to make the portfolio Delta neutral. The reader should keep in mind that the underlying security will have a $\Gamma = 0$ and thus Gamma neutrality for the portfolio will be maintained while Delta neutrality is achieved. Thus Eq. (9.2) for a Gamma neutral portfolio reduces to

$$\delta\mathcal{P} \approx \Theta\delta t. \quad (9.3)$$

Example 9.2 Suppose a stock's current value is \$100 while its volatility is $\sigma = 0.22$ and the risk-free interest rate is 2.5% per year. An investment firm sells a European call option on this stock with a strike time of three months and a strike price of \$102. The firm buys European call options on the same stock with the same strike price but with a strike time of six months. According to Eq. (8.10) the Gamma of the three-month option is $\Gamma_3 = 0.03618$, while that of the six-month options is $\Gamma_6 = 0.02563$. The

portfolio becomes Gamma neutral at any point in the first quadrant of $w_3 w_6$ -space where the equation

$$0.03618w_3 - 0.02563w_6 = 0$$

is satisfied. Suppose $w_3 = 100000$ of the three-month option were sold, then the portfolio is Gamma neutral if $w_6 = 141163$ six-month options are purchased. Thus prior to the inclusion of the underlying stock in the portfolio, the Delta of the portfolio is

$$w_3 \Delta_3 - w_6 \Delta_6 = (100000)(0.4728) - (141163)(0.5123) = -25038.$$

Therefore the portfolio can be made Delta neutral if 25038 shares of the underlying stock are sold short. Figure 9.5 shows that over a fairly wide range of values for the underlying stock the value of the portfolio remains nearly the same.

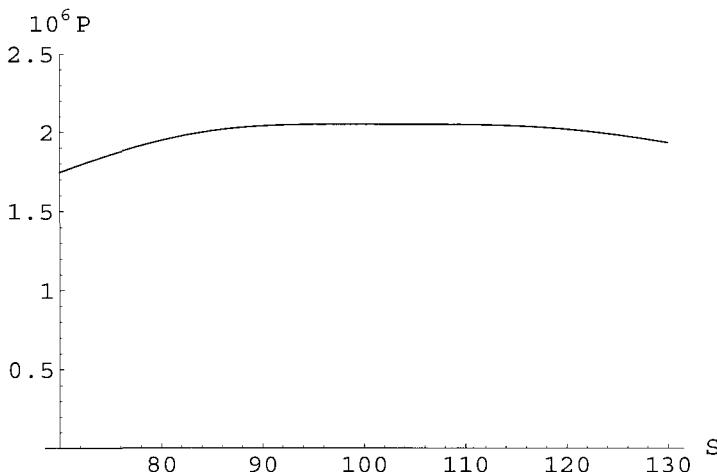


Fig. 9.5 The aggregate value of a Gamma neutral portfolio is insensitive to a range of changes in the value of the underlying securities.

This discussion of hedging is far from complete. The present chapter has focused mainly on making a portfolio's value resistant to changes in the value of a security. In reality the risk-free interest rate and the volatility of the stock also affect the value of a portfolio. The quantities Rho and Vega discussed in Chapter 8 can be used to set up portfolios hedged against changes in the interest rate and volatility. In the present chapter it was also

assumed that the necessary options and securities could be bought so as to form the desired hedge. In practice this may not always be possible. For example an investment firm may not be able to purchase sufficient quantities of a stock to make their portfolio Delta or Gamma neutral. In this case they may have to substitute a different, but related security or other financial instrument in order to set up the hedge. This strategy will be discussed in the next chapter after the prerequisite statistical concepts are introduced.

9.5 Exercises

- (1) A call option closes “in the money” if the market price of the underlying security exceeds the strike price for the call option at the time the option matures. Use Eq. (7.35) to show that $\Delta \rightarrow 1$ as $t \rightarrow T^-$ for an “in the money” European call option.
- (2) A call option closes “out of the money” if the strike price for the call option exceeds the market price of the underlying security at the time the option matures. Use Eq. (7.35) to show that $\Delta \rightarrow 0$ as $t \rightarrow T^-$ for an “out of the money” European call option.
- (3) Suppose the price of a stock is \$45 per share with volatility $\sigma = 0.20$ per annum and the risk free interest rate is 4.5% per year. Create a table similar to Table 9.1 for European call options on 5000 shares of the stock with a strike price of \$47 per share and a strike time of 15 weeks. Rebalance the portfolio weekly assuming the weekly prices of the stock are as follows.

Week	0	1	2	3	4	5	6	7
Week	8	9	10	11	12	13	14	15
S	47.79	48.33	48.81	51.36	52.06	51.98	54.22	54.31

- (4) Repeat exercise 3 for the following scenario in which the call option finishes out of the money.

Week	0	1	2	3	4	5	6	7
Week	8	9	10	11	12	13	14	15
S	41.94	42.72	44.83	44.93	44.19	41.77	39.56	40.62

- (5) Repeat exercise 3 assuming this time that the financial institution has issued a European put option for the stock.
- (6) Repeat exercise 4 assuming this time that the financial institution has issued a European put option for the stock.
- (7) Set up a Gamma neutral portfolio consisting of European call options on a security whose current value is \$85. Suppose the risk-free interest rate is 5.5% and the volatility of the security is 17% per year. The portfolio should contain a short position in four-month options with a strike price of \$88 and a long position in six-month options with the same strike price.
- (8) For the portfolio described in exercise 7 add a position in the underlying security so that the portfolio is also Delta neutral.
- (9) Set up a Gamma neutral portfolio consisting of European call options on a security whose current value is \$95. Suppose the risk-free interest rate is 4.5% and the volatility of the security is 23% per year. The portfolio should contain a short position in three-month options with a strike price of \$97 and a long position in five-month options with a strike price of \$98.
- (10) For the portfolio described in exercise 9 add a position in the underlying security so that the portfolio is also Delta neutral.

Chapter 10

Optimizing Portfolios

There are several notions of the idea of optimizing a portfolio of securities, options, bonds, cash, *etc.* In this chapter we will explore optimality in the sense of maximizing the rate of return, minimizing the variance in the rate of return, and minimizing risk to the investor. We will also introduce the Capital Assets Pricing Model which attempts to relate the rate of return of a specific investment to the rate of return for the entire market of investments. We will see in this chapter that the difference in the expected rates of return for a specific security and the risk-free interest rate is proportional to the difference in the expected rates of return for the market and the risk-free interest rate. In order to explain this proportionality relationship we will introduce two additional statistical measures, covariance and correlation. We will also make use of these concepts in developing additional hedging strategies.

10.1 Covariance and Correlation

The concept of **covariance** is related to the degree to which two random variables tend to change in the same or opposite direction relative to one another. If X and Y are the random variables then mathematically the covariance, denoted $\text{Cov}(X, Y)$, is defined as

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]. \quad (10.1)$$

Making use of the definition and properties of expected value one can see that

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}[XY - Y\mathbb{E}[X] - X\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]] \\ &= \mathbb{E}[XY] - \mathbb{E}[Y]\mathbb{E}[X] - \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y],\end{aligned}\tag{10.2}$$

where Eq. (10.2) is generally more convenient to use than the expression in the right-hand side of Eq. (10.1).

Example 10.1 Table 10.1 lists the names, heights, and weights of a sample of *Playboy* centerfolds [Dean *et al.* (2001)]. This data will be used to illustrate the concept of covariance. Let X represent height and Y represent weight for each woman. The reader can readily calculate that $\mathbb{E}[X] = 67.05$ inches and the $\mathbb{E}[Y] = 117.6$ pounds. The expected value of the pairwise product of height and weight is $\mathbb{E}[XY] = 7887.85$. Thus the covariance of height and weight for this sample is $\text{Cov}(X, Y) = 2.77$. Note that the covariance is positive, indicating that, in general, as height increases so does weight.

Table 10.1 A sample of names, heights, and weights for women appearing as *Playboy* centerfolds.

Name	Height (in.)	Weight (lb.)
Allias, Henriette	68.5	125
Anderson, Pamela	67	105
Broady, Eloise	68	125
Butler, Cher	67	123
Clark, Julie	65	110
Sloan, Tiffany	66	120
Stewart, Liz	67	116
Taylor, Tiffany	67	115
Witter, Cherie	69	117
York, Brittany	66	120

From the definition of covariance several properties of this concept follow almost immediately. If X and Y are independent random variables then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ by Theorem 2.5 and thus the covariance of independent random variables is zero. The following set of relationships can also be established.

Theorem 10.1 Suppose X , Y , and Z are random variables, then the following statements are true:

- (1) $\text{Cov}(X, X) = \text{Var}(X)$,
- (2) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$,
- (3) $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$.

The proofs of the first two statements are left for the reader as exercises. Justification for statement 3 is given below.

Proof. Let X , Y , and Z be random variables, then

$$\begin{aligned}\text{Cov}(X + Y, Z) &= E[(X + Y)Z] - E[X + Y]E[Z] \\ &= E[XZ] + E[YZ] - E[X]E[Z] - E[Y]E[Z] \\ &= E[XZ] - E[X]E[Z] + E[YZ] - E[Y]E[Z] \\ &= \text{Cov}(X, Z) + \text{Cov}(Y, Z)\end{aligned}$$

□

The third statement of Theorem 10.1 can be generalized as in the following corollary.

Corollary 10.1 Suppose $\{X_1, X_2, \dots, X_n\}$ and $\{Y_1, Y_2, \dots, Y_m\}$ are sets of random variables where $n, m \geq 1$.

$$\text{Cov}\left(\sum_{i=1}^n X_i, \sum_{i=1}^m Y_i\right) = \sum_{i=1}^n \sum_{j=1}^m \text{Cov}(X_i, Y_j) \quad (10.3)$$

Proof. We begin by demonstrating the result when $m = 1$. The corollary holds trivially when $n = 1$ and follows from the third statement of Theorem 10.1 when $n = 2$. Suppose the claim holds when $n \leq k$ where $k \in \mathbb{N}$. Let $\{X_1, X_2, \dots, X_k, X_{k+1}\}$ be random variables.

$$\begin{aligned}\text{Cov}\left(\sum_{i=1}^{k+1} X_i, Y_1\right) &= \text{Cov}\left(\sum_{i=1}^k X_i, Y_1\right) + \text{Cov}(X_{k+1}, Y_1) \\ &= \sum_{i=1}^k \text{Cov}(X_i, Y_1) + \text{Cov}(X_{k+1}, Y_1) \\ &= \sum_{i=1}^{k+1} \text{Cov}(X_i, Y_1)\end{aligned}$$

Therefore by induction we may show that the result is true for any finite, integer value of n (at least when $m = 1$). Now we can relax the assumption on m . Suppose the claim has been proved for $m \leq k$ where again $k \in \mathbb{N}$. Let $\{Y_1, Y_2, \dots, Y_k, Y_{k+1}\}$ be random variables, then

$$\begin{aligned} \text{Cov} \left(\sum_{i=1}^n X_i, \sum_{j=1}^{k+1} Y_j \right) &= \text{Cov} \left(\sum_{j=1}^{k+1} Y_j, \sum_{i=1}^n X_i \right) \\ &= \sum_{j=1}^k \text{Cov} \left(Y_j, \sum_{i=1}^n X_i \right) + \text{Cov} \left(Y_{k+1}, \sum_{i=1}^n X_i \right) \\ &= \sum_{j=1}^k \text{Cov} \left(\sum_{i=1}^n X_i, Y_j \right) + \text{Cov} \left(\sum_{i=1}^n X_i, Y_{k+1} \right) \\ &= \sum_{j=1}^k \sum_{i=1}^n \text{Cov}(X_i, Y_j) + \sum_{i=1}^n \text{Cov}(X_i, Y_{k+1}) \\ &= \sum_{i=1}^n \sum_{j=1}^k \text{Cov}(X_i, Y_j) + \sum_{i=1}^n \text{Cov}(X_i, Y_{k+1}) \\ &= \sum_{i=1}^n \sum_{j=1}^{k+1} \text{Cov}(X_i, Y_j) \end{aligned}$$

Therefore Eq. (10.3) holds all finite, positive integer values of m , and n . \square

Yet another corollary follows from Corollary 10.1. This one generalizes statement 1 of Theorem 10.1.

Corollary 10.2 *If $\{X_1, X_2, \dots, X_n\}$ are random variables then*

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i}^n \text{Cov}(X_i, X_j). \quad (10.4)$$

Proof. Let $Y = \sum_{i=1}^n X_i$, then according to the first statement of Theorem 10.1,

$$\begin{aligned}\text{Var}(Y) &= \text{Cov}(Y, Y) \\ \text{Var}\left(\sum_{i=1}^n X_i\right) &= \text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) \quad (\text{by Corollary 10.1}) \\ &= \sum_{i=1}^n \text{Cov}(X_i, X_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(X_i, X_j)\end{aligned}$$

Once more the first statement of Theorem 10.1 was used to reintroduce the variance in the last line of the derivation. \square

Often a quantity related to covariance is used as a measure of the degree to which large values of a random variable X are associated with large values of another random variable Y . This quantity is known as **correlation** and is denoted $\rho(X, Y)$. The correlation of two random variables is defined as

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}. \quad (10.5)$$

The correlation of two random variables can be interpreted as a measure of the degree to which one of the variables can be written as a linear function of the other. The correlation is more than just a simple re-scaling of the covariance. While the covariance of X and Y may numerically be positive, negative, or zero, the correlation always lies in the interval $[-1, 1]$. Once again we see that independent random variables have a correlation of zero and hence are described as **uncorrelated**.

Theorem 10.2 Suppose X and Y are random variables such that $Y = aX + b$ where $a, b \in \mathbb{R}$ with $a \neq 0$. If $a > 0$ then $\rho(X, Y) = 1$, while if $a < 0$ then $\rho(X, Y) = -1$.

Proof. We start by calculating the covariance of X and Y .

$$\begin{aligned}\text{Cov}(X, Y) &= \text{Cov}(X, aX + b) \\ &= E[X(aX + b)] - E[X]E[aX + b] \\ &= E[aX^2 + bX] - E[X](aE[X] + b) \\ &= aE[X^2] + bE[X] - aE[X]E[X] - bE[X] \\ &= a(E[X^2] - E[X]^2) \\ &= a\text{Var}(X)\end{aligned}$$

The reader should make note of the use of Theorem 2.3. Therefore using the result of exercise 18 from Chapter 2

$$\rho(X, Y) = \frac{a\text{Var}(X)}{\sqrt{\text{Var}(X) \cdot a^2\text{Var}(X)}} = \frac{a}{|a|}$$

which is -1 when $a < 0$ and 1 when $a > 0$. \square

Before we can bound the correlation in the interval $[-1, 1]$ we will need to prove the following lemma.

Lemma 10.1 (Schwarz Inequality) *If X and Y are random variables then $(E[XY])^2 \leq E[X^2]E[Y^2]$.*

Proof. The cases in which $E[X^2] = 0$, $E[X^2] = \infty$, $E[Y^2] = 0$, or $E[Y^2] = \infty$ are left as exercises. Suppose for the purposes of this proof that $0 < E[X^2] < \infty$ and $0 < E[Y^2] < \infty$. If a and b are real numbers then the following two inequalities hold:

$$\begin{aligned}0 &\leq E[(aX + bY)^2] = a^2E[X^2] + 2abE[XY] + b^2E[Y^2] \\ 0 &\leq E[(aX - bY)^2] = a^2E[X^2] - 2abE[XY] + b^2E[Y^2]\end{aligned}$$

If we let $a^2 = E[Y^2]$ and $b^2 = E[X^2]$ then the first inequality above yields

$$\begin{aligned}0 &\leq E[Y^2]E[X^2] + 2\sqrt{E[Y^2]E[X^2]}E[XY] + E[X^2]E[Y^2] \\ -2E[X^2]E[Y^2] &\leq 2\sqrt{E[Y^2]E[X^2]}E[XY] \\ -\sqrt{E[X^2]E[Y^2]} &\leq E[XY].\end{aligned}$$

By a similar set of steps the second inequality produces

$$E[XY] \leq \sqrt{E[X^2]E[Y^2]}.$$

Therefore, since

$$\begin{aligned}-\sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]} &\leq \mathbb{E}[XY] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]} \\ (\mathbb{E}[XY])^2 &\leq \mathbb{E}[X^2]\mathbb{E}[Y^2].\end{aligned}$$

Occasionally $(\mathbb{E}[XY])^2$ is written as $\mathbb{E}[XY]^2$ or even as $\mathbb{E}^2[XY]$. A careful reading of the expression will avoid any possible confusion. \square

Now we are in a position to prove the following theorem.

Theorem 10.3 *If X and Y are random variables then $-1 \leq \rho(X, Y) \leq 1$.*

Proof. Consider the covariance of X and Y .

$$\begin{aligned}(\text{Cov}(X, Y))^2 &= (\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])])^2 \\ &\leq \mathbb{E}[(X - \mathbb{E}[X])^2]\mathbb{E}[(Y - \mathbb{E}[Y])^2] \quad (\text{by Lemma 10.1}) \\ &= \text{Var}(X)\text{Var}(Y)\end{aligned}$$

Thus $|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$, which is equivalent to the inequality,

$$\begin{aligned}-\sqrt{\text{Var}(X)\text{Var}(Y)} &\leq \text{Cov}(X, Y) < \sqrt{\text{Var}(X)\text{Var}(Y)} \\ -1 &\leq \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} < 1 \\ -1 &\leq \rho(X, Y) < 1,\end{aligned}$$

which follows from the definition of correlation. \square

Example 10.2 Referring to the data in Table 10.1 the correlation between height and weight is calculated as

$$\frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{2.77}{\sqrt{(1.46944)(41.8222)}} \approx 0.3533.$$

For this data set there does not seem to be an especially strong linear relationship between height and weight. Figure 10.1 illustrates this as well.

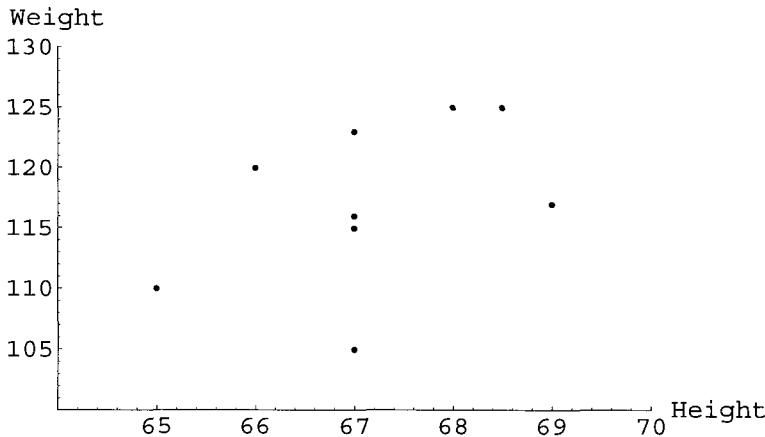


Fig. 10.1 A scatter plot of height *versus* weight for a sample of ten women who appeared as *Playboy* centerfolds.

We end this section with a more financial application of the concepts of covariance and correlation. A portfolio may consist of n different investments. Suppose that an investor may invest an amount w_i for one time period and at the end of the time period, the investment will return wealth in the amount of $w_i X_i$. In Sec. 1.5 the concept of rate of return was defined to be the interest rate equivalent to this growth in value. For the i^{th} investment the rate of return will be determined by

$$w_i(1 + R_i) = w_i X_i \quad \text{which implies} \quad R_i = X_i - 1.$$

Thus the total accumulated wealth after one time period from all n investments is $W = \sum_{i=1}^n w_i(1 + R_i)$. If the rate of return for each of the n investments is treated as a random variable, then

$$\mathbb{E}[W] = \mathbb{E}\left[\sum_{i=1}^n w_i(1 + R_i)\right] = \sum_{i=1}^n w_i + \sum_{i=1}^n w_i \mathbb{E}[R_i].$$

The variance in the accumulated wealth is, according to Eq. (10.4),

$$\begin{aligned}\text{Var}(W) &= \text{Var} \left(\sum_{i=1}^n w_i (1 + R_i) \right) \\ &= \text{Var} \left(\sum_{i=1}^n w_i R_i \right) \\ &= \sum_{i=1}^n \text{Var}(w_i R_i) + \sum_{i=1}^n \sum_{j \neq i} \text{Cov}(w_i R_i, w_j R_j) \\ &= \sum_{i=1}^n w_i^2 \text{Var}(R_i) + \sum_{i=1}^n \sum_{j \neq i} w_i w_j \text{Cov}(R_i, R_j).\end{aligned}$$

Later sections will expand on this idea and develop a method of selecting a portfolio of investments which minimizes the variance in the wealth generated. Minimizing the variance in the rate of return on a portfolio of investments makes the return more “predictable”. Remember that the smaller the variance of a random variable, the smaller the spread about the expected value of that variable.

Before concluding this section we will exercise our understanding of covariance and lognormal random variables to derive a technical result which will be needed in Sec. 10.6.

Lemma 10.2 *If X is a lognormal random variable with drift parameter μ and volatility σ^2 and $K > 0$ is a constant then*

$$\text{Cov}(X, (X - K)^+) = \mathbb{E}[X(X - K)^+] - \mathbb{E}[X]\mathbb{E}[(X - K)^+], \quad (10.6)$$

where

$$\mathbb{E}[X(X - K)^+] = e^{2(\mu+\sigma^2)}\phi(w + 2\sigma) - Ke^{\mu+\sigma^2/2}\phi(w + \sigma), \quad (10.7)$$

with $w = (\mu - \ln K)/\sigma$.

Proof. Equation 10.6 follows from the definition of covariance. The technical portion of the lemma is the derivation of Eq. 10.7.

$$\begin{aligned}\mathbb{E}[X(X - K)^+] &= \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty x(x - K)^+ \frac{1}{x} e^{-(\ln x - \mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_K^\infty (x - K) e^{-(\ln x - \mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^\infty (e^{\sigma z + \mu} - K) e^{\sigma z + \mu} e^{-z^2/2} dz\end{aligned}$$

The last equation is derived by making the substitution $\sigma z = \ln x - \mu$. Therefore

$$\begin{aligned} E[X(X - K)^+] &= \frac{e^{2(\mu+\sigma^2)}}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^{\infty} e^{-(z-2\sigma)^2/2} dz \\ &\quad - \frac{e^{\mu+\sigma^2/2}}{\sqrt{2\pi}} \int_{(\ln K - \mu)/\sigma}^{\infty} e^{-(z-\sigma)^2/2} dz \\ &= e^{2(\mu+\sigma^2)} \phi\left(\frac{\mu - \ln K}{\sigma} + 2\sigma\right) \\ &\quad - Ke^{\mu+\sigma^2/2} \phi\left(\frac{\mu - \ln K}{\sigma} + \sigma\right). \end{aligned}$$

The proof is complete if we let $w = (\mu - \ln K)/\sigma$. \square

10.2 Optimal Portfolios

In the previous chapter on hedging, all of the discussion and examples assumed that a hedged position with respect to a particular option could be set up by taking a position in the stock or security underlying the option. However, it may not always be possible to purchase (or sell) sufficient shares of the underlying security (or the option itself) to create the hedged position. In these cases the manager of a portfolio may have to find a surrogate for the security. Upon reflection, a particular financial instrument will be a better or worse surrogate for a security depending on how the values of the security and the other instrument change. If the changes in the two values are highly correlated then the portfolio manager will have more confidence in using the surrogate.

Suppose an investment firm sells a European call option for stock A and wishes to create a Delta-neutral portfolio. If insufficient shares of stock A can be purchased to create the hedge then the firm can investigate hedging their short position in the option for stock A with a long position in n shares of stock B. If the hedge could be created with stock A in place of stock B then the Delta-neutral hedge is achieved when $n = \Delta$ as we saw in Sec. 9.2. The issue confronting the investment firm now is, what should n be when the hedge must be created with stock B? An **optimal portfolio** is one for which the variance in the value of the portfolio is minimized. The value of the portfolio consisting of a short position on a call option for

stock A and a long position of n shares of stock B will be denoted

$$\mathcal{P} = C_A - nB.$$

The variance in the value of the portfolio is

$$\begin{aligned}\text{Var}(\mathcal{P}) &= \text{E}[(C_A - nB)^2] - \text{E}[C_A - nB]^2 \\ &= \text{Var}(C_A) + n^2\text{Var}(B) - 2n\text{Cov}(C_A, B) \\ &= n^2\text{Var}(B) - 2n\text{Cov}(C_A, B) + \text{Var}(C_A).\end{aligned}$$

Thus it is seen that the variance in the value of the portfolio is a quadratic function of the hedging parameter n . Thus the minimum variance is achieved when

$$n = \frac{\text{Cov}(C_A, B)}{\text{Var}(B)} = \rho(C_A, B) \sqrt{\frac{\text{Var}(C_A)}{\text{Var}(B)}}.$$

Thus if the correlation between the values of the option on stock A and stock B were unity then the hedging parameter would be the ratio of the standard deviations in the value of the option on stock A and the value of stock B. A small correlation between C_A and B would indicate that stock B is a poor surrogate for stock A.

The idea of selecting a portfolio by minimizing the variance of its value will be extended in later sections of this chapter. Before returning to minimum variance analysis, we must introduce the concept known as utility and describe its properties.

10.3 Utility Functions

This section will focus on defining and understanding a class of functions which can be used as the basis of rational decision making. Suppose that an investor is faced with the choice of two different investment products. Suppose further that the set of outcomes resulting from investing in either of the products is $\{C_1, C_2, \dots, C_n\}$. The reader should think of this set of outcomes as the union of the possible outcomes of investing in the first product and the possible outcomes resulting from investing in the second product. The probability of outcome C_i coming about as the result of investing in the first product will be denoted p_i for $i = 1, 2, \dots, n$. If outcome C_i cannot result from investing in the first product then, of course,

$p_i = 0$. For the second product the values of the probabilities will be denoted q_i .

The investor can rank the outcomes in order of desirability, with the most desirable outcome being C_α and the least desirable being C_ω . To each of the possible outcomes a **utility** can be assigned. The utility function $u(C_i)$ will have a range of $[0, 1]$. To start, $u(C_\alpha) = 1$ and $u(C_\omega) = 0$. The values of $u(C_i)$ for $i \neq \alpha, \omega$ will lie in the open interval $(0, 1)$. They are assigned as follows. The investor imagines that they have the choice of receiving outcome C_i with certainty or participating in a random experiment in which they will receive the most desirable outcome C_α with probability u and the least desirable outcome C_ω with probability $1 - u$. If $u = 1$ then the investor will surely participate in the random experiment because they will receive C_α with certainty, and they rank this as a more desirable outcome than C_i . If $u = 0$ then the investor will definitely not participate in the random experiment since that would result in them receiving the least desirable outcome. They are better off taking the sure outcome C_i . As u decreases from 1 to 0 there will exist a value of $0 < u < 1$ at which the investor will switch from the behavior of participating in the random experiment to taking outcome C_i with certainty. This value of u will be the value of $u(C_i)$. This probability is called the utility of outcome C_i . It is assumed that the probability at which the behavior changes is unique. The rational investor once having decided to accept the sure outcome C_i will not at a lower probability of receiving C_α decide to once again participate in the random experiment. Thus the utility function is well-defined.

The utility function provides a means by which two outcomes of unequal desirability may be compared. Outcome C_i is equivalent to receiving C_α with probability $u(C_i)$ or receiving C_ω with probability $1 - u(C_i)$. Returning to the original task of deciding between two different investment products, the investor can then understand that investing in the first product is equivalent to receiving C_α with probability

$$\sum_{i=1}^n p_i u(C_i),$$

while the second investment is equivalent to C_α with probability

$$\sum_{i=1}^n q_i u(C_i).$$

Therefore the investor will choose the first product whenever

$$\sum_{i=1}^n p_i u(C_i) > \sum_{i=1}^n q_i u(C_i), \quad (10.8)$$

and otherwise will choose the second product.

In this section we have spoken of outcomes or consequences C_i of making an investment decision. More concretely these outcomes can be thought of as receiving differing amounts of money for an investment (some of which may be negative). Thus in general the **utility function** denoted $u(x)$ is the investor's utility of receiving an amount x . Utility functions are specific (like personality traits) to each investor. In the remainder of this section categories and properties of the utility function will be explored.

A general property of utility functions is that the amount of extra utility that an investor experiences when x is increased to $x + \Delta x$ is non-increasing. In other words $u(x + \Delta x) - u(x)$ is a non-increasing function of x . This property is illustrated in Fig. 10.2. We will describe a function $f(t)$ as

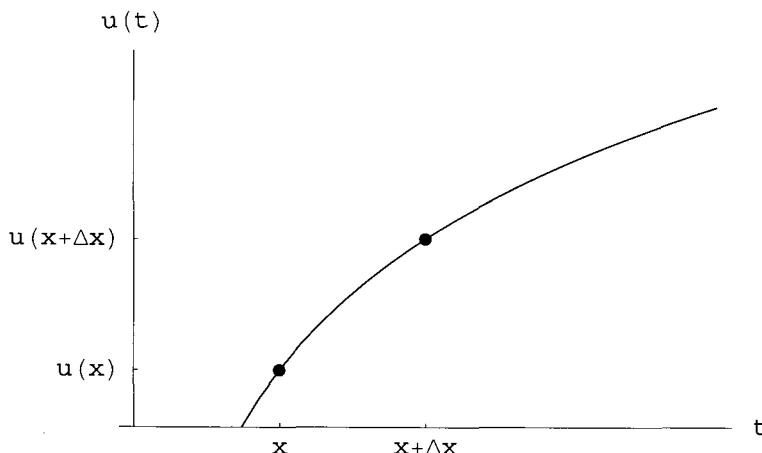


Fig. 10.2 The extra utility received, $u(x + \Delta x) - u(x)$, is a decreasing function of x .

concave on an open interval (a, b) if for every $x, y \in (a, b)$ and every $\lambda \in [0, 1]$ we have

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y). \quad (10.9)$$

Graphically this may be interpreted as meaning that all the secant lines

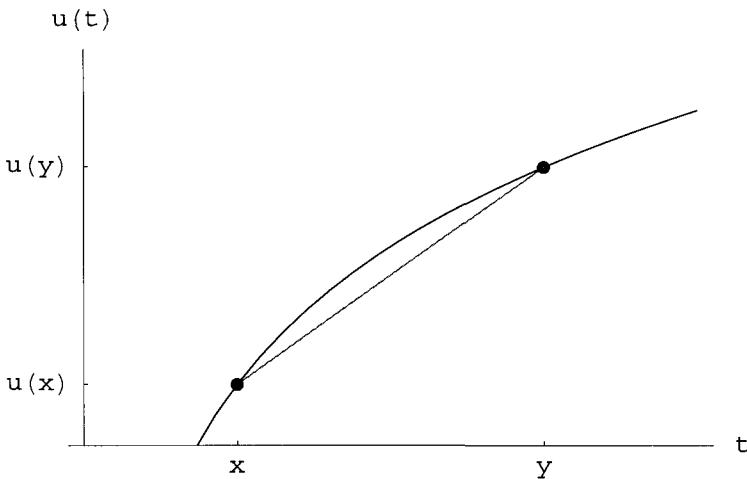


Fig. 10.3 For concave functions the secant line parameterized by $\lambda \in [0, 1]$ will lie below the corresponding point on the graph of the function.

lie below the graph of $f(t)$. See Fig. 10.3. A utility function $u(t)$, obeying inequality (10.9) will also be called concave. An investor whose utility function is concave is said to be **risk-averse**. An investor with a linear utility function of the form $u(x) = ax + b$ with $a > 0$ is said to be **risk-neutral**.

The following result can be demonstrated regarding concave functions.

Theorem 10.4 *If $f \in C^2(a, b)$ then f is concave on (a, b) if and only if $f''(t) \leq 0$ for $a < t < b$.*

Proof. If f is concave on (a, b) then by definition f satisfies inequality (10.9). Let $y \in (a, b)$. If $w = \lambda x + (1 - \lambda)y$ and if $0 < \lambda < 1$ then $a < x < w < y < b$. By definition of w ,

$$w - x = (1 - \lambda)(y - x) > 0 \quad (10.10)$$

$$y - w = \lambda(y - x) > 0 \quad (10.11)$$

Making use of inequality (10.9) produces the following pair of inequalities:

$$f(w) - f(x) \geq (1 - \lambda)(f(y) - f(x)) \quad (10.12)$$

$$f(y) - f(w) \leq \lambda(f(y) - f(x)) \quad (10.13)$$

Through the use of inequalities (10.12) and (10.13) and Eq. (10.10) and

(10.11) we obtain

$$\begin{aligned}\frac{f(y) - f(w)}{y - w} &\leq \frac{\lambda(f(y) - f(w))}{y - w} = \frac{f(y) - f(x)}{y - x} \\ &= \frac{(1 - \lambda)(f(y) - f(x))}{w - x} \\ &\leq \frac{f(w) - f(x)}{w - x}.\end{aligned}$$

By the Mean Value Theorem [Stewart (1999)] there exist α and β with $x < \beta < w < \alpha < y$ such that

$$f'(\alpha) = \frac{f(y) - f(w)}{y - w} \leq \frac{f(w) - f(x)}{w - x} = f'(\beta).$$

Consequently

$$\frac{f'(\alpha) - f'(\beta)}{\alpha - \beta} \leq 0,$$

and in the limit we see that

$$\lim_{x \rightarrow y} \frac{f'(\alpha) - f'(\beta)}{\alpha - \beta} = f''(y) \leq 0.$$

To prove the converse suppose $f''(x) \leq 0$ on (a, b) . Then for any x, y such that $a < x < y < b$, we can for any $\lambda \in (0, 1)$ assign $w = \lambda x + (1 - \lambda)y$ and note that $x < w < y$.

$$\frac{f(y) - f(w)}{y - w} \leq \frac{f(w) - f(x)}{w - x}.$$

From the inequality immediately above it is seen that

$$\begin{aligned}(w - x) \frac{f(y) - f(w)}{y - w} &\leq f(w) - f(x) \\ f(x) + (\lambda x + (1 - \lambda)y - x) \frac{f(y) - f(\lambda x + (1 - \lambda)y)}{y - \lambda x - (1 - \lambda)y} &\leq f(\lambda x + (1 - \lambda)y) \\ f(x) + (1 - \lambda)(y - x) \frac{f(y) - f(\lambda x + (1 - \lambda)y)}{\lambda(y - x)} &\leq f(\lambda x + (1 - \lambda)y) \\ \lambda f(x) + (1 - \lambda)(f(y) - f(\lambda x + (1 - \lambda)y)) &\leq \lambda f(\lambda x + (1 - \lambda)y) \\ \lambda f(x) + (1 - \lambda)f(y) &\leq f(\lambda x + (1 - \lambda)y)\end{aligned}$$

which implies that f is concave on (a, b) . □

In the next section we will make use of the following result known as Jensen's Inequality.

Theorem 10.5 (Jensen's Inequality (Discrete Version)) *Let f be a concave function on the interval (a, b) , suppose $x_i \in (a, b)$ for $i = 1, 2, \dots, n$, and suppose $\lambda_i \in [0, 1]$ for $i = 1, 2, \dots, n$ with $\sum_{i=1}^n \lambda_i = 1$, then*

$$\sum_{i=1}^n \lambda_i f(x_i) \leq f\left(\sum_{i=1}^n \lambda_i x_i\right) \quad (10.14)$$

Proof. Let $\mu = \sum_{i=1}^n \lambda_i x_i$ and note that since $\lambda_i \in [0, 1]$ for $i = 1, 2, \dots, n$ and $\sum_{i=1}^n \lambda_i = 1$, then $a < \mu < b$. The equation of the line tangent to the graph of f at the point $(\mu, f(\mu))$ is $y = f'(\mu)(x - \mu) + f(\mu)$. Since f is concave on (a, b) then

$$f(x_i) \leq f'(\mu)(x_i - \mu) + f(\mu) \quad \text{for } i = 1, 2, \dots, n.$$

Therefore

$$\begin{aligned} \sum_{i=1}^n \lambda_i f(x_i) &\leq \sum_{i=1}^n (\lambda_i [f'(\mu)(x_i - \mu) + f(\mu)]) \\ &= f'(\mu) \sum_{i=1}^n (\lambda_i x_i - \lambda_i \mu) + f(\mu) \sum_{i=1}^n \lambda_i \\ &= f'(\mu) \left(\mu - \mu \sum_{i=1}^n \lambda_i \right) + f(\mu) \\ &= f(\mu) \\ &= f\left(\sum_{i=1}^n \lambda_i x_i\right). \end{aligned}$$

□

Jensen's inequality may be interpreted as saying that for a concave function, the mean of the values of the function applied to a set of values of a random variable is no greater than the function applied to the mean of the values of the random variable. If u is a concave utility function, then $E[u(X)] \leq u(E[X])$. Thus a risk-averse investor prefers the certain return of $C = E[X]$ to receiving a random return with this expected value.

Example 10.3 Suppose $f(x) = \tanh x$ and let

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5\} = \{0.30, 0.07, 0.23, 0.20, 0.20\}.$$

The reader can verify that $f(x)$ is concave on $(0, \infty)$. If $X_i = i^2$ for $i = 1, 2, \dots, 5$ then

$$\sum_{i=1}^5 \lambda_i f(X_i) = 0.928431 \leq 1 = f\left(\sum_{i=1}^5 \lambda_i X_i\right)$$

There is also a continuous version of Jensen's Inequality.

Theorem 10.6 (Jensen's Inequality (Continuous Version)) *Let $\phi(t)$ be an integrable function on $[0, 1]$ and let f be a concave function, then*

$$\int_0^1 f(\phi(t)) dt \leq f\left(\int_0^1 \phi(t) dt\right). \quad (10.15)$$

Proof. For the sake of compactness of notation let $\alpha = \int_0^1 \phi(t) dt$, and let $y = f'(\alpha)(x - \alpha) + f(\alpha)$, the equation of the tangent line passing through the point with coordinates $(\alpha, f(\alpha))$. Since f is concave then

$$f(\phi(t)) \leq f'(\alpha)(\phi(t) - \alpha) + f(\alpha),$$

which implies that

$$\begin{aligned} \int_0^1 f(\phi(t)) dt &\leq \int_0^1 f'(\alpha)(\phi(t) - \alpha) + f(\alpha) dt \\ &= f'(\alpha) \int_0^1 \phi(t) dt - \int_0^1 f'(\alpha)\alpha dt + \int_0^1 f(\alpha) dt \\ &= f'(\alpha)\alpha - f'(\alpha)\alpha + f(\alpha) \\ &= f(\alpha) \\ &= f\left(\int_0^1 \phi(t) dt\right). \end{aligned}$$

□

10.4 Expected Utility

As the name suggests, **expected utility** is the expected value of a utility function. Returning to the previous discussion of choosing between two possible investment instruments, we see that the investment with the greater expected utility is preferable. The inequality in (10.8) indicated that if the first investment choice results in outcomes $\{C_1, C_2, \dots, C_n\}$ with respective probabilities p_i for $i = 1, 2, \dots, n$ and the second investment choice

produces the same outcomes but with probabilities $\{q_1, q_2, \dots, q_n\}$ then a rational investor would select the first investment whenever

$$\sum_{i=1}^n p_i u(C_i) > \sum_{i=1}^n q_i u(C_i).$$

Suppose the random variable X is thought of as the set of outcomes with probabilities $\{p_1, p_2, \dots, p_n\}$ while the random variable Y represents the outcomes with probabilities $\{q_1, q_2, \dots, q_n\}$. Then the first investment is preferable whenever

$$\mathbb{E}[u(X)] > \mathbb{E}[u(Y)].$$

Example 10.4 An investor must choose between the following two “investments”:

- A:** Flip a fair coin, if the coin lands heads up the investor receives \$10, otherwise they receive nothing.
- B:** Receive an amount $\$M$ with certainty.

The investor is risk-averse with a utility function $u(x) = x - x^2/25$. The rational investor will select the investment with the greater expected utility. The expected utility for investment A is

$$\frac{1}{2}u(10) + \frac{1}{2}u(0) = \frac{1}{2}\left(10 - \frac{10^2}{25}\right) = 3.$$

The expected utility for B is $u(M) = M - M^2/25$. Thus the investor will choose the coin flip whenever

$$\begin{aligned} 3 &> M - \frac{M^2}{25} \\ M^2 - 25M + 75 &> 0 \\ \frac{25 - 5\sqrt{13}}{2} &> M. \end{aligned}$$

Thus investment A is preferable to B whenever $M < \$3.49$.

If $M \geq \$3.49$ in the previous example, then the investor will choose the second investment. This illustrates the concept known as the **certainty equivalent** which is defined as the value C of a random variable X at which $u(C) = \mathbb{E}[u(X)]$.

Example 10.5 An investor wishes to find the certainty equivalent C for the following investment choice:

A: Flip a fair coin, if the coin lands heads up the investor receives $0 < X \leq 10$, otherwise they receive $0 < Y < X$.

B: Receive an amount C with certainty.

Assuming the investor's utility function is the same as in the previous example, the certainty equivalent and payoffs of investment A must satisfy the equation

$$\begin{aligned} \frac{1}{2} \left(X - \frac{X^2}{25} + Y - \frac{Y^2}{25} \right) &= C - \frac{C^2}{25} \\ C^2 - 25C + 12(X^2 + Y^2) &= 0 \\ C &= \frac{25 - \sqrt{625 - 48(X^2 + Y^2)}}{2}. \end{aligned}$$

It is clear from the expression under the radical that the certainty equivalent is real only when $X^2 + Y^2 \leq 625/48$. The design of investment A specifies that $0 < Y < X < 10$. Thus the certainty equivalent can be thought of as a surface plotted in cylindrical polar coordinates over the region where $0 < r < \frac{25}{4\sqrt{3}}$ and $0 < \theta < \pi/4$. Figure 10.4 illustrates the certainty equivalent as a function of X and Y .

10.5 Portfolio Selection

In the previous sections we have treated the investor's choice between two different investments as a binary, all-or-nothing decision. In reality an investor may choose to split investment capital between two (or more) investments. In this section we will explore the **portfolio selection problem**, the task of allocating funds in an optimal fashion among several investment options.

We begin with a simple example. Suppose an investor has a total of x amount of capital to invest. Assuming that the investor may use any proportion of this capital, let $\alpha \in [0, 1]$ be the proportional they invest. The investment is structured such that an allocation of αx will earn αx with probability p and lose αx with probability $1 - p$. Thus for an investment of αx the random variable representing the investor's financial position after the conclusion of the investment is

$$X = \begin{cases} x(1 + \alpha) & \text{with probability } p, \\ x(1 - \alpha) & \text{with probability } 1 - p. \end{cases}$$

The allocation proportion α will be optimal when the expected value of the

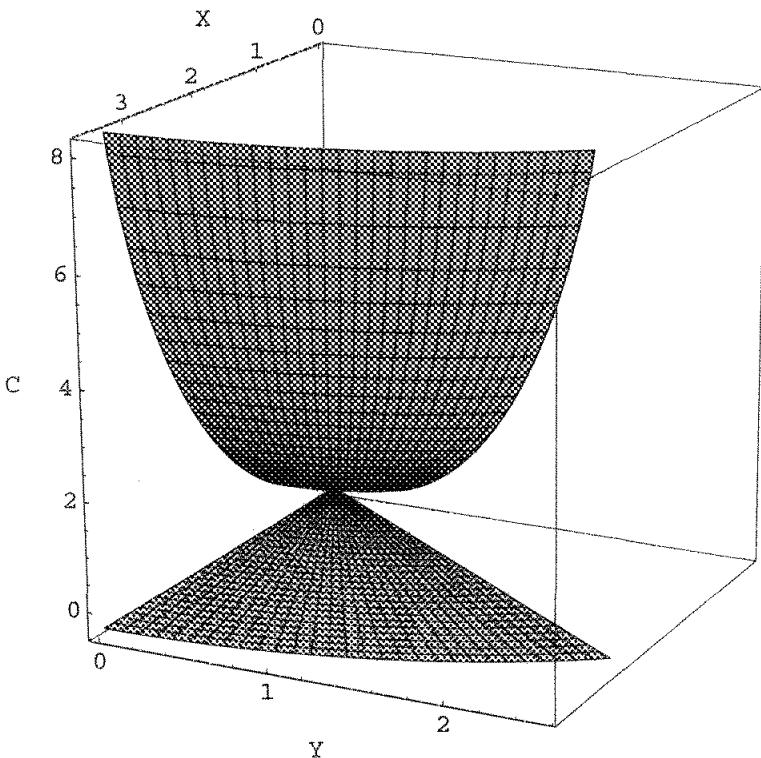


Fig. 10.4 The certainty equivalent in this example is a section of an ellipsoid. The domain of interest in the XY -plane is drawn as a shadow below the surface.

utility is maximized.

$$\begin{aligned} E[u(X)] &= pu(x(1 + \alpha)) + (1 - p)u(x(1 - \alpha)) \\ \frac{d}{d\alpha} E[u(X)] &= pxu'(x(1 + \alpha)) - x(1 - p)u'(x(1 - \alpha)) \\ 0 &= pu'(x(1 + \alpha)) - (1 - p)u'(x(1 - \alpha)) \end{aligned}$$

The critical value of α which solves the last equation above will correspond to a maximum for the expected value of the utility as long as the utility function is concave.

Example 10.6 If $u(x) = \ln x$, then the expected value of the investor's utility is maximized when

$$\frac{p}{1 + \alpha} - \frac{1 - p}{1 - \alpha} = 0, \quad \text{which implies } \alpha = 2p - 1.$$

Thus if $p > 1/2$ the investor should allocate $100(2p - 1)\%$ of their capital. If $0 \leq p \leq 1/2$, then $E[u(X)]$ is maximized when $\alpha = 0$, i.e. when no investment is made. A curve depicting the expected value of the investor's utility as a function of α and p is presented in Fig. 10.5.

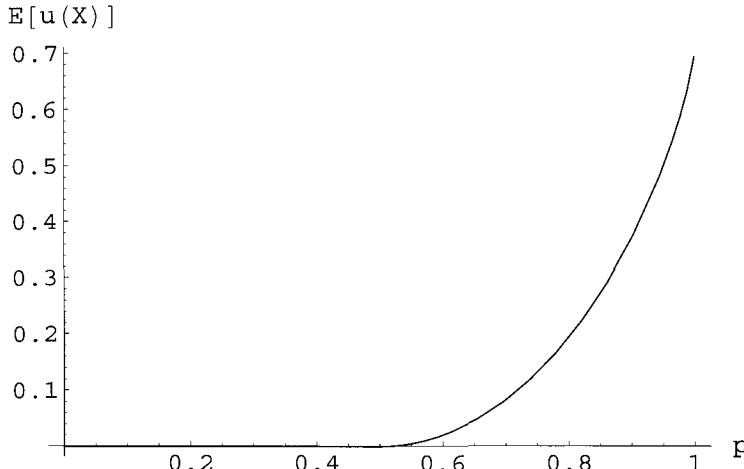


Fig. 10.5 The expected value of utility for a risk-averse investor allocating $100(2p - 1)\%$ of their capital.

With this simple example understood, we are now ready to consider a more general situation. Suppose that an amount of capital x may be allocated among n different investments. A proportion x_i will be allocated to security i for $i = 1, 2, \dots, n$. The vector notation $\langle x_1, x_2, \dots, x_n \rangle$ will be used to represent the portfolio of investments. The return from investment i will be denoted W_i for $i = 1, 2, \dots, n$. The portfolio selection problem is then defined to be that of determining x_i for $i = 1, 2, \dots, n$ such that

- (1) $0 \leq x_i \leq 1$ for $i = 1, 2, \dots, n$, and
- (2) $\sum_{i=1}^n x_i = 1$,

and which maximizes the investor's total expected utility $E[u(W)]$, where

$$W = x \sum_{i=1}^n x_i W_i.$$

Throughout this section it will be assumed that n is large and that the W_i 's are not too highly correlated. Under these assumptions the Central

Limit Theorem of Sec. 3.5 implies that W is a normally distributed random variable.

The allocation of investment funds can be driven by the objective of maximizing the expected value of the investor's utility function. Suppose an investor's utility function is given by $u(x) = 1 - e^{-bx}$ where $b > 0$. The reader is asked in exercise 12 to show that this utility function is concave and monotonically increasing. Assuming that W is a normally distributed random variable then $-bW$ is also normally distributed with

$$\mathbb{E}[-bW] = -b\mathbb{E}[W] \quad \text{and} \quad \text{Var}(-bW) = b^2\text{Var}(W).$$

Consequently,

$$\mathbb{E}[u(W)] = \mathbb{E}[1 - e^{-bW}] = 1 - \mathbb{E}[e^{-bW}].$$

Making the assignment $Y = e^{-bW}$, then Y is a lognormal random variable (see Sec. 3.6). According to Lemma 3.1,

$$\mathbb{E}[Y] = \mathbb{E}[e^{-bW}] = e^{-b\mathbb{E}[W]+b^2\text{Var}(W)/2}$$

which implies that

$$\mathbb{E}[u(W)] = 1 - e^{-b(\mathbb{E}[W]-b\text{Var}(W)/2)}.$$

Thus the expected utility is maximized when $\mathbb{E}[W] - b\text{Var}(W)/2$ is maximized. If two different portfolios, $\langle x_1, x_2, \dots, x_n \rangle$ and $\langle y_1, y_2, \dots, y_n \rangle$, give rise respectively to returns X and Y then the portfolio represented by $\langle x_1, x_2, \dots, x_n \rangle$ will be preferable if

$$\mathbb{E}[X] \geq \mathbb{E}[Y] \quad \text{and} \quad \text{Var}(X) \leq \text{Var}(Y).$$

To extend this discussion and make it more concrete numerically, suppose $b = 0.005$ and that an investor wishes to choose among an infinite number of different portfolios which can be created by investing in two securities denoted A and B. The investor has \$100 to invest and will allocate y dollars to security A and $100 - y$ dollars to B. The following table summarizes what is known about the rates of return on the hypothetical securities A and B.

Security	A	B
$\mathbb{E}[\text{rate of return}]$	0.16	0.18
$\text{Var}(\text{rate of return})$	0.20	0.24

Assume that the correlation between the rates of return is $\rho = -0.35$. Once again W will represent the wealth returned.

$$\begin{aligned} E[W] &= 100 + 0.16y + 0.18(100 - y) \\ &= 118 - 0.02y \end{aligned}$$

$$\begin{aligned} \text{Var}(W) &= y^2(0.20)^2 + (100 - y)^2(0.24)^2 + 2y(100 - y)(0.20)(0.24)(-0.35) \\ &= 0.04y^2 + 0.0576(100 - y)^2 - 0.0336y(100 - y) \end{aligned}$$

Since the optimal portfolio is the one which maximizes $E[W] - b\text{Var}(W)/2$ then we must maximize

$$\begin{aligned} 118 - 0.02y - 0.0025(0.04y^2 + 0.0576(100 - y)^2 - 0.0336y(100 - y)) \\ = -0.000328y^2 + 0.0172y + 116.56. \end{aligned}$$

This occurs when $y \approx 26.2195$. For this choice of y

$$\begin{aligned} E[W] &\approx 117.476 \\ \text{Var}(W) &\approx 276.049 \\ E[u(W)] &\approx 0.442296 \end{aligned}$$

In the next section we will generalize and extend these ideas to situations in which investment capital is partitioned n ways.

10.6 Minimum Variance Analysis

Building on the ideas contained in the previous section we consider the case of an investor allocating a portion of their capital between two potential investments. Without loss of generality it is assumed that the investor has an infinitely divisible unit of capital to invest and will invest a fraction $\alpha \in [0, 1]$ in one security and $1 - \alpha$ in a second security. The rates of return of the two securities are respectively R_1 and R_2 . The variances of these rates of return will be denoted σ_1^2 and σ_2^2 . The covariance in the rates of return is assumed to be c . Thus the variance in the wealth returned from the investments is

$$\text{Var}(W) = \alpha^2\sigma_1^2 + (1 - \alpha)^2\sigma_2^2 + 2c\alpha(1 - \alpha). \quad (10.16)$$

One measure of an optimal portfolio of investments would be the one with the least variance in the returned wealth. Differentiating $\text{Var}(W)$ with

respect to α and solving for the critical value of α produces

$$0 = 2(\alpha\sigma_1^2 - (1-\alpha)\sigma_2^2 + c(1-2\alpha))$$

$$\alpha^* = \frac{\sigma_2^2 - c}{\sigma_1^2 + \sigma_2^2 - 2c}$$

The denominator of α^* is the covariance of $R_1 - R_2$ and thus is non-negative. In order for the critical value of α to be defined, it is assumed that $\sigma_1^2 + \sigma_2^2 > 2c$. This condition is also sufficient to guarantee that the variance of the wealth in Eq. (10.16) will have a global minimum. The parameter α^* must also fall within the closed interval $[0, 1]$. This occurs whenever

$$c \leq \min\{\sigma_1^2, \sigma_2^2\}.$$

If $c \leq 0$ then this condition is always satisfied. According to the Extreme Value Theorem [Stewart (1999)], the minimum variance will occur when $\alpha = 0$, $\alpha = 1$, or $\alpha = \alpha^*$.

$$\begin{aligned}\text{Var}(W)|_{\alpha=0} &= \sigma_2^2 \\ \text{Var}(W)|_{\alpha=1} &= \sigma_1^2 \\ \text{Var}(W)|_{\alpha=\alpha^*} &= \frac{\sigma_1^2\sigma_2^2 - c^2}{\sigma_1^2 + \sigma_2^2 - 2c}\end{aligned}$$

Consider the special case when the rates of return of the two investments are uncorrelated. Under this assumption $c = 0$ and the critical value of α at which the variance in the returned wealth is minimized reduces to

$$\alpha^* = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \frac{\frac{1}{\sigma_1^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}.$$

There is an appealing simplicity and symmetry to this critical value. This is not due to coincidence, it is in fact a pattern seen in the more general case of allocating a unit of investment among n different potential investments.

Theorem 10.7 *Suppose that $0 \leq \alpha_i \leq 1$ will be invested in security i for $i = 1, 2, \dots, n$ subject to the constraint that $\alpha_1 + \alpha_2 + \dots + \alpha_n = 1$. Suppose the rate of return of security i is a random variable R_i and that all the rates of return are mutually uncorrelated. The optimal, minimum variance portfolio described by the allocation vector $\langle \alpha_1, \alpha_2, \dots, \alpha_n \rangle$, is the*

one for which

$$\alpha_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^n \frac{1}{\sigma_j^2}} \quad \text{for } i = 1, 2, \dots, n,$$

where $\sigma_i^2 = \text{Var}(R_i)$.

Proof. Since the rates of return are uncorrelated then the variance in the returned wealth W is

$$\text{Var}(W) = \sum_{i=1}^n \alpha_i^2 \sigma_i^2,$$

and is subject to the constraint that $1 = \sum_{i=1}^n \alpha_i$. Minimizing the variance subject to the constraint is accomplished by use of Lagrange Multipliers [Stewart (1999)]. This technique states that $\text{Var}(W)$ will be optimized at one of the solutions $(\alpha_1, \alpha_2, \dots, \alpha_n, \lambda)$ of the set of simultaneous equations:

$$\begin{aligned} \nabla \left(\sum_{i=1}^n \alpha_i^2 \sigma_i^2 \right) &= \lambda \nabla \left(\sum_{i=1}^n \alpha_i \right) \\ \sum_{i=1}^n \alpha_i &= 1 \end{aligned}$$

The symbol ∇ denotes the gradient operator. These equations are equivalent to respectively:

$$\begin{aligned} 2\alpha_i \sigma_i^2 &= \lambda \quad \text{for } i = 1, 2, \dots, n, \text{ and} \\ \sum_{i=1}^n \alpha_i &= 1 \end{aligned}$$

Solving for α_i in the first equation and substituting into the second equation determines that

$$\lambda = \frac{2}{\sum_{j=1}^n \frac{1}{\sigma_j^2}}.$$

Substituting this expression for λ into the first equation yields

$$\alpha_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^n \frac{1}{\sigma_j^2}} \quad \text{for } i = 1, 2, \dots, n.$$

The reader can confirm that for each i , $\alpha_i \in [0, 1]$. □

The previous discussion can be generalized yet again to the situation in which the portfolio of securities is financed with borrowed capital. Let $\mathbf{w} = \langle w_1, w_2, \dots, w_n \rangle$ represent the portfolio of investments. As before, R_i will represent the return on investing w_i in the i^{th} security. For simplicity's sake we will assume this is a one-period model in which investments are purchased by borrowing money which must be paid back with simple interest of rate r . The return on the portfolio after one time period is

$$R(\mathbf{w}) = \sum_{i=1}^n w_i(1 + R_i) - (1 + r) \sum_{i=1}^n w_i = \sum_{i=1}^n w_i(R_i - r). \quad (10.17)$$

The expected rate of return on the portfolio and the variance in the rate of return on the portfolio are functions of the vector \mathbf{w} and are defined to be respectively,

$$r(\mathbf{w}) = E[R(\mathbf{w})] \quad (10.18)$$

$$\sigma^2(\mathbf{w}) = \text{Var}(R(\mathbf{w})). \quad (10.19)$$

In this situation, the borrowed amounts w_1, w_2, \dots, w_n do not have to sum to unity, or to any other prescribed value.

Lemma 10.3 *Assuming the rates of return on the securities are uncorrelated, the optimal portfolio with an expected return of unity which minimizes the variance of the portfolio is*

$$\mathbf{w}^* = \left\langle \frac{\frac{r_1 - r}{\sigma_1^2}}{\sum_{j=1}^n \frac{(r_j - r)^2}{\sigma_j^2}}, \frac{\frac{r_2 - r}{\sigma_2^2}}{\sum_{j=1}^n \frac{(r_j - r)^2}{\sigma_j^2}}, \dots, \frac{\frac{r_n - r}{\sigma_n^2}}{\sum_{j=1}^n \frac{(r_j - r)^2}{\sigma_j^2}} \right\rangle \quad (10.20)$$

where $r_i = E[R_i]$ and $\sigma_i^2 = \text{Var}(R_i)$ for $i = 1, 2, \dots, n$.

The proof of this lemma is similar to the proof of Theorem 10.7 and is left as an exercise.

The existence of the optimal portfolio \mathbf{w}^* provided by Lemma 10.3 will be used in the following result known as the **Portfolio Separation Theorem**.

Theorem 10.8 (Portfolio Separation Theorem) *If b is any positive scalar, the variance of all portfolios with expected rate of return equal to b is minimized by portfolio $b\mathbf{w}^*$ where \mathbf{w}^* is described in Eq. (10.20).*

Proof. Suppose \mathbf{x} is a portfolio for which $r(\mathbf{x}) = b$, then

$$\begin{aligned}\frac{1}{b}r(\mathbf{x}) &= r\left(\frac{1}{b}\mathbf{x}\right) \quad (\text{by exercise 15}) \\ &= 1.\end{aligned}$$

Thus we see that $\frac{1}{b}\mathbf{x}$ is a portfolio with unit expected rate of return. For the portfolio \mathbf{w}^* ,

$$\begin{aligned}\sigma^2(b\mathbf{w}^*) &= b^2\sigma^2(\mathbf{w}^*) \\ &\leq b^2\sigma^2\left(\frac{1}{b}\mathbf{x}\right) \quad (\text{by Lemma 10.3}) \\ &= \sigma^2(\mathbf{x}) \quad (\text{by exercise 15}).\end{aligned}$$

□

As a consequence of the Portfolio Separation Theorem all expected rates of return on portfolios can be normalized to unity when determining the optimal portfolio.

As a special case of the preceding discussion the investor may wish to divide their portfolio between a risk-free investment such as a savings account and a “risky” investment such as a security. The risk-free interest rate will be symbolized by r_f . By assumption the rate of return on the risk-free investment is guaranteed, *i.e.* has a variance of zero. If the rate of return of the security is R_S and $x \in [0, 1]$ is invested in the security while $1 - x$ is placed in the risk-free investment, the rate of return on the portfolio is

$$R = (1 - x)r_f + xR_S,$$

which implies the expected rate of return for this portfolio is

$$\begin{aligned}\mathbb{E}[R] &= (1 - x)\mathbb{E}[r_f] + x\mathbb{E}[R_S] \\ &= (1 - x)r_f + xR_S \\ r &= r_f + (R_S - r_f)x\end{aligned}\tag{10.21}$$

To achieve a more compact notation we have set $r = \mathbb{E}[R]$ and $r_S = \mathbb{E}[R_S]$. Since the risk-free return is a constant, its expected value is this same constant value. The expected rate of return on the portfolio is a linear function of x . The line has its r -intercept at r_f and slope $r_S - r_f$. Thus if the expected rate of return of the security exceeds the risk-free rate of return, the slope of the line is positive and the expected rate of return of

the portfolio increases with x , otherwise it decreases. The variance in the rate of return on the portfolio is

$$\begin{aligned}\text{Var}(R) &= (1-x)^2\text{Var}(r_f) + x^2\text{Var}(R_S) + 2x(1-x)\text{Cov}(r_f, R_S) \\ \sigma^2 &= x^2\sigma_S^2\end{aligned}\tag{10.22}$$

where $\text{Var}(R) = \sigma^2$ and $\text{Var}(R_S) = \sigma_S^2$. The reader will note that since the rate of return on the risk-free investment is constant, its variance is zero and the covariance with the rate of return on the security also vanishes. The standard deviation of the return on the portfolio is also an increasing linear function of x . Equation (10.22) can be used to eliminate the parameter x from Eq. (10.21) and to derive a formula relating the expected rate of return and the standard deviation of the rate of return.

$$r = r_f + \frac{r_S - r_f}{\sigma_S} \sigma$$

Once again, if the expected rate of return on the security exceeds the rate of return of the risk-free investment, the expected rate of return of the portfolio increases with the standard deviation in the return on the portfolio. Thus an investor willing to accept a wider variation in the return on funds invested can expect a higher rate of return.

If the idealized single security of the previous discussion is replaced with an investment spread evenly throughout the securities market then the same form of linear relationship is determined, but the symbols r_M and σ_M will be used in place of r_S and σ_S respectively. This is known as the equation of the **capital market line**.

$$r = r_f + \frac{r_M - r_f}{\sigma_M} \sigma\tag{10.23}$$

The capital market line is illustrated in Fig. 10.6.

Before moving on, the reader should reflect on the σr -plane. For a fixed level of expected return, risk-adverse investors will want to minimize the variance (or standard deviation) in the rate of return. A fixed level of return corresponds to a horizontal line in the σr -plane. Thus a risk-adverse investor would prefer a portfolio for which the standard deviation in the rate of return is as far to the left on the horizontal line as is feasible. The phrase “as is feasible” is used because it may not be possible to achieve a given rate of return with zero standard deviation. Referring to the illustration of the capital market line in Fig. 10.6, for a rate of return above r_f it is not feasible to have zero variance in the rate of return. Similarly, for a

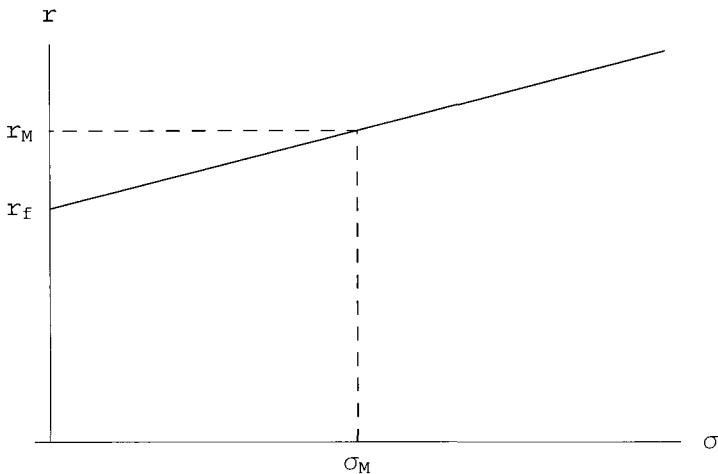


Fig. 10.6 The capital market line illustrates the linear relationship between the standard deviation in the rate of return on a portfolio and the return on the portfolio.

fixed standard deviation in the rate of return, an investor should prefer the highest expected return. Thus along a vertical line in the σr -plane an investor should prefer a portfolio as high on the line as is feasible. Again, for a fixed σ not all levels of r are achievable. These ideas could be extended a great deal and the interested reader is encouraged to consult [Luenberger (1998)].

Finally we arrive at a discussion of the main topic of this chapter, the **Capital Asset Pricing Model** (CAPM). This mathematical model will relate the return on the investment in an individual security to the return on the entire market. In this way an investor can weigh the return on the investment against the risk involved in the investment. The expression R_i will be the return on investment i and R_M will denote the return on the entire market. An investor will invest a portion x of his wealth in investment i and the remainder $1-x$ in the market. The following calculations will seem familiar to the reader. They are similar to those done when determining the optimal allocation of a portfolio between two securities. This time the second security is the entire market. The rate of return on the portfolio is given by

$$R = xR_i + (1-x)R_M,$$

which implies the expected rate of return on the portfolio and the variance

in the expected rate of return are respectively,

$$\mathbb{E}[R] = x\mathbb{E}[R_i] + (1-x)\mathbb{E}[R_M]$$

$$\text{Var}(R) = x^2\text{Var}(R_i) + (1-x)^2\text{Var}(R_M) + 2x(1-x)\text{Cov}(R_i, R_M).$$

Without further information we cannot assume that the rate of return for investment i is uncorrelated from the rate of return for the market. To simplify the notation we will make the following assignments: $r = \mathbb{E}[R]$, $r_i = \mathbb{E}[R_i]$, $r_M = \mathbb{E}[R_M]$, $\sigma^2 = \text{Var}(R)$, $\sigma_i^2 = \text{Var}(R_i)$, and $\sigma_M^2 = \text{Var}(R_M)$. Furthermore we will replace $\text{Cov}(R_i, R_M)$ by its slightly more compact equivalent $\rho_{i,M}\sigma_i\sigma_M$.¹ Thus the equations above become

$$r = xr_i + (1-x)r_M \quad (10.24)$$

$$\sigma^2 = x^2\sigma_i^2 + (1-x)^2\sigma_M^2 + 2x(1-x)\rho_{i,M}\sigma_i\sigma_M. \quad (10.25)$$

Equations (10.24) and (10.25) can be thought of as the parametric form of a curve in the σr -plane. When $x = 0$, corresponding to the situation in which the entire portfolio is invested in the market, the parametric curve intersects the capital market line. In fact the parametric curve must be tangent to the capital market line at $x = 0$. If the intersection were transverse rather than tangential, then for some value of x near zero the parametric curve would lie above the capital market line and hence represent an infeasible expected return. Thus we can obtain a relationship between r_i , r_M , and r_f . Using Eq. (10.23) to obtain the slope of the capital market line we have

$$\begin{aligned} \frac{r_M - r_f}{\sigma_M} &= \left. \frac{dr}{d\sigma} \right|_{x=0} \\ &= \left. \frac{\frac{dr}{dx}}{\frac{d\sigma}{dx}} \right|_{x=0} \\ &= \frac{r_i - r_M}{\frac{\rho_{i,M}\sigma_i\sigma_M}{\sigma_M} - \sigma_M}. \end{aligned}$$

Solving the last equation for $r_i - r_f$ yields

$$r_i - r_f = \frac{\rho_{i,M}\sigma_i}{\sigma_M}(r_M - r_f). \quad (10.26)$$

The expression $\rho_{i,M}\sigma_i/\sigma_M$ is commonly denoted β_i and is referred to as the **beta** of security i . Equation (10.26) can be interpreted as stating that the excess rate of return of security i above the risk-free interest rate is

¹This is a small abuse of notation since we normally use $\rho(R_i, R_M)$ to denote the correlation of R_i and R_M . Here there is little chance of confusion.

proportional to the excess rate of return of the market above the risk-free interest rate. The proportionality constant is β_i . For the case in which the rate of return on security i is uncorrelated with the rate of return on the market, the excess expected return will be zero. When the correlation is positive, excess positive expected return for the market implies excess positive expected return for security i . When the correlation is negative, excess positive expected return for the market will imply excess negative return for security i .

Example 10.7 Suppose the risk-free interest rate is 5% per year, the expected rate of return on the market is 9% per year, and the standard deviation in the return on the market is 15% per year. If the covariance in the expected returns on a particular stock and a the market is 10% then

$$\beta = \frac{0.10}{(0.15)^2} = 4.44444.$$

Therefore the expected rate of return on the stock will be

$$r = 0.05 + 4.44444(0.09 - 0.05) = 0.22777,$$

or expressed as a percentage return, 22.78%.

The variance in the rate of return of a portfolio can be thought of as a measure of the risk in the portfolio. For a given level of variance in the rate of return, the investor will expect the highest expected value of the rate of return. Thus once again we turn to Lagrange Multipliers to solve this constrained optimization problem. The maximum and minimum value of expected rate of return will occur when

$$\frac{dr}{dx} = \lambda \frac{d(\sigma^2)}{dx} \quad (10.27)$$

subject to the constraint expressed in Eq. (10.25). The reader will be asked to verify in the exercises that the solution to this set of equations is

$$\lambda = \pm \frac{r_i - r_M}{2\sqrt{(\rho_{i,M}^2 - 1)\sigma_i^2\sigma_M^2 + (\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M)\sigma^2}} \quad (10.28)$$

$$x = \frac{\sigma_M^2 - \rho_{i,M}\sigma_i\sigma_M}{\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M} \pm \frac{\sqrt{(\rho_{i,M}^2 - 1)\sigma_i^2\sigma_M^2 + (\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M)\sigma^2}}{\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M}. \quad (10.29)$$

Under these conditions the maximum expected rate of return is

$$r = r_M - \frac{(r_i - r_M)(\rho_{i,M}\sigma_i\sigma_M - \sigma_M^2)}{\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M} \quad (10.30)$$

$$\mp \frac{(r_i - r_M)\sqrt{(\rho_{i,M}^2 - 1)\sigma_i^2\sigma_M^2 + (\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M)\sigma^2}}{\sigma_i^2 + \sigma_M^2 - 2\rho_{i,M}\sigma_i\sigma_M}.$$

In most scenarios the variance in the rate of return for investment i will exceed the variance for the market. Thus we will assume that $\sigma_i > \sigma_M$. As two special cases, we will explore the return on the portfolio when the rates of return on the market and investment i are perfectly correlated, *i.e.* when $\rho_{i,m} = 1$. In this case

$$r = r_M - \frac{(r_i - r_M)(\sigma_M \pm \sigma)}{\sigma_i - \sigma_M}.$$

If the rate of return on investment i is anti-correlated with the rate of return on the market, then

$$r = r_M - \frac{(r_i - r_M)(\sigma_M \pm \sigma)}{\sigma_i + \sigma_M}.$$

Example 10.8 Suppose the expected rate of return for investment i is $r_i = 0.07$ with standard deviation $\sigma_i = 0.25$ and the expected rate of return for the market is $r_M = 0.05$ with standard deviation $\sigma_M = 0.20$. Suppose the rates of return have a correlation of $\rho_{i,M} = 0.57$. If an investor is willing to hold a portfolio with a standard deviation in its rate of return at $\sigma = 0.22$ then the maximum rate of return can be found via the method outlined above. According to Eq. (10.30), the maximum rate of return on the portfolio is 0.0650248. This occurs when $x = 0.751242$ or when slightly more than 75% of investment capital is placed in investment i . The curve depicting r and x as parametric functions of σ is shown in Fig. 10.7.

10.7 Mean Variance Analysis

This chapter concludes with a discussion of the stochastic process approach to determining the expected rate of return of a portfolio. If necessary, the reader may wish to review Chapter 5 before proceeding. In this section the portfolio will consist of positions in a security and a European call option on the security. The reader should not be confused by the portfolio. The investor will purchase a units of the security and b units of the call option,

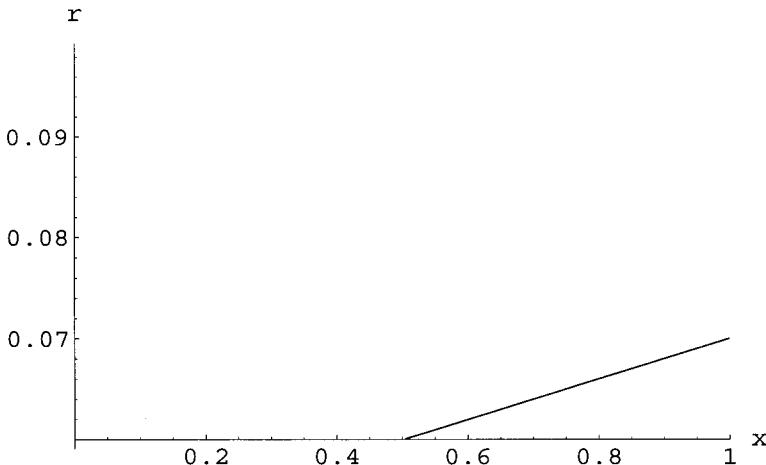


Fig. 10.7 This curve shows the relationship between the rate of return r , and the proportion of capital invested x .

where either a or b could be positive or negative. For example, the investor might take a short position in the security and hedge this position with a long position in the call option. Likewise, a short position in the call option can be offset with a long position in the security.

We will assume the present value of the security is $S(0)$, the value of the option is C , the strike price is K , and the expiry date is T . The risk-free interest rate is r . Suppose the value of the security obeys a Brownian motion-type, stochastic process with drift parameter μ and volatility σ . The reader will recall that this means that the logarithm of the value of the security follows a Wiener process of the form:

$$d(\ln S(t)) = \mu dt + \sigma dW(t).$$

By Itô's Lemma (Lemma 5.3) then

$$dS(t) = \left(\mu + \frac{\sigma^2}{2} \right) S(t) dt + \sigma S(t) dW(t).$$

At time $t = 0$ the investor devotes resources in the amount of $aS(0) + bC$ to create the portfolio. At the expiry date T the positions in the portfolio are zeroed out. The present value of any gain or loss is then

$$R = e^{-rT} (aS(T) + b(S(T) - K)^+) - (aS(0) + bC).$$

Using the linearity property of the expected value, the expected return is

$$\begin{aligned}\mathbb{E}[R] &= \mathbb{E}[e^{-rT}(aS(T) + b(S(T) - K)^+) - aS(0) - bC] \\ &= e^{-rT}\mathbb{E}[aS(T) + b(S(T) - K)^+] - aS(0) - bC \\ &= e^{-rT}(a\mathbb{E}[S(T)] + b\mathbb{E}[(S(T) - K)^+]) - aS(0) - bC.\end{aligned}$$

According to Eqs. (5.23) and (5.24) the expected value and variance of $\ln S(T)$ are respectively

$$\mathbb{E}[\ln S(T)] = \mu T \quad \text{and} \quad \text{Var}(\ln S(T)) = \sigma^2 T.$$

Thus by Lemma 3.1, $\mathbb{E}[S(T)] = S(0)e^{(\mu+\sigma^2/2)T}$. Provided that $r < \mu+\sigma^2/2$ the expression

$$e^{-rT}\mathbb{E}[S(T)] = S(0)e^{(-r+\mu+\sigma^2/2)T} > S(0).$$

In order to determine a closed form expression for $\mathbb{E}[(S(T) - K)^+]$ we will make use of Corollary 3.2. $S(T)$ is a lognormal random variable with parameters $\ln S(0) + \mu T$ (the drift parameter) and $\sigma\sqrt{T}$ (the volatility parameter). Thus according to Corollary 3.2,

$$\begin{aligned}\mathbb{E}[(S(T) - K)^+] &= S(0)e^{(\mu+\sigma^2/2)T}\phi\left(\frac{\ln S(0) + \mu T - \ln K}{\sigma\sqrt{T}} + \sigma\sqrt{T}\right) \\ &\quad - K\phi\left(\frac{\ln S(0) + \mu T - \ln K}{\sigma\sqrt{T}}\right) \\ &= S(0)e^{(\mu+\sigma^2/2)T}\phi\left(\frac{\ln(S(0)/K) + \mu T}{\sigma\sqrt{T}} + \sigma\sqrt{T}\right) \\ &\quad - K\phi\left(\frac{\ln(S(0)/K) + \mu T}{\sigma\sqrt{T}}\right) \\ &= S(0)e^{(\mu+\sigma^2/2)T}\phi(w + \sigma\sqrt{T}) - K\phi(w),\end{aligned}$$

where $w = (\ln(S(0)/K) + \mu T)/\sigma\sqrt{T}$. Thus the expected value of the rate of return is

$$\begin{aligned}r &= \mathbb{E}[R] \\ &= e^{-rT}\left(S(0)e^{(\mu+\frac{\sigma^2}{2})T}\left(a + b\phi(w + \sigma\sqrt{T})\right) - bK\phi(w)\right) - aS(0) - bC.\end{aligned}\tag{10.31}$$

The level sets of the expected return are lines in the ab -plane.

The variance in the rate of return can be calculated as

$$\begin{aligned}\text{Var}(R) &= \text{Var}(e^{-rT} (aS(T) + b(S(T) - K)^+) - aS(0) - bC) \\ &= \text{Var}(e^{-rT} (aS(T) + b(S(T) - K)^+)) \\ &= e^{-2rT} \text{Var}(aS(T) + b(S(T) - K)^+) \\ &= e^{-2rT} (a^2 \text{Var}(S(T)) + b^2 \text{Var}((S(T) - K)^+) \\ &\quad + 2ab \text{Cov}(S(T), (S(T) - K)^+)).\end{aligned}$$

Referring to the second part of Lemma 3.1 we have

$$\text{Var}(S(T)) = S^2(0)e^{(2\mu+\sigma^2)T} (e^{\sigma^2 T} - 1).$$

According to Corollary 3.4,

$$\begin{aligned}\text{Var}((S(T) - K)^+) &= S^2(0)e^{2(\mu+\sigma^2)T} \phi(w + 2\sigma\sqrt{T}) \\ &\quad - 2KS(0)e^{(\mu+\sigma^2/2)T} \phi(w + \sigma\sqrt{T}) + K^2\phi(w) \\ &\quad - (S(0)e^{(\mu+\sigma^2/2)T} \phi(w + \sigma\sqrt{T}) - K\phi(w))^2.\end{aligned}$$

Lemma 10.2 implies that

$$\begin{aligned}\text{Cov}(S(T), (S(T) - K)^+) &= KS(0)e^{(\mu+\sigma^2/2)T} (\phi(w) - \phi(w + \sigma\sqrt{T})) \\ &\quad - S^2(0)e^{2(\mu+\sigma^2)T} (\phi(w + \sigma\sqrt{T}) - \phi(w + 2\sigma\sqrt{T}))\end{aligned}$$

Putting these last three results together yields, at long last, the complicated, but complete, expression for $\text{Var}(R)$.

$$\begin{aligned}\text{Var}(R) &= e^{-2rT} (a^2 S^2(0)e^{(2\mu+\sigma^2)T} (e^{\sigma^2 T} - 1) \\ &\quad + b^2 (S^2(0)e^{2(\mu+\sigma^2)T} \phi(w + 2\sigma\sqrt{T}) - 2KS(0)e^{(\mu+\sigma^2/2)T} \phi(w + \sigma\sqrt{T}) \\ &\quad + K^2\phi(w) - (S(0)e^{(\mu+\sigma^2/2)T} \phi(w + \sigma\sqrt{T}) - K\phi(w))^2) \\ &\quad + 2ab (KS(0)e^{(\mu+\sigma^2/2)T} (\phi(w) - \phi(w + \sigma\sqrt{T})) \\ &\quad - S^2(0)e^{2(\mu+\sigma^2)T} (\phi(w + \sigma\sqrt{T}) - \phi(w + 2\sigma\sqrt{T})))\end{aligned}\tag{10.32}$$

The level sets of $\text{Var}(R)$ are parabolas in the ab -plane.

Example 10.9 Suppose a share of a security has a current price of \$100 while the drift parameter and volatility of the price of the security are respectively 0.08 and 0.21 per annum. The risk free interest rate is 3.5%

annually. The value of a European call option for a strike price of \$102 with an exercise time of 12 months on the security is \$9.075. An investor can use the preceding analysis to determine that the expected rate of return on a portfolio consisting of a position x in the security and a position y in the call option is given by Eq. (10.31) as

$$\text{E}[R] = 6.93489x + 4.3834y.$$

The line of zero return is given by the equation $y = -1.58208x$. Thus at a point in the second or fourth quadrant above this line the expected return will be positive. We look for a point in the second or fourth quadrant since typically x and y must be of opposite sign due to the fact that the investor will take opposite positions in the security and the option. A density plot in Fig. 10.8 shows the return in various subsets of the plane. If the investor

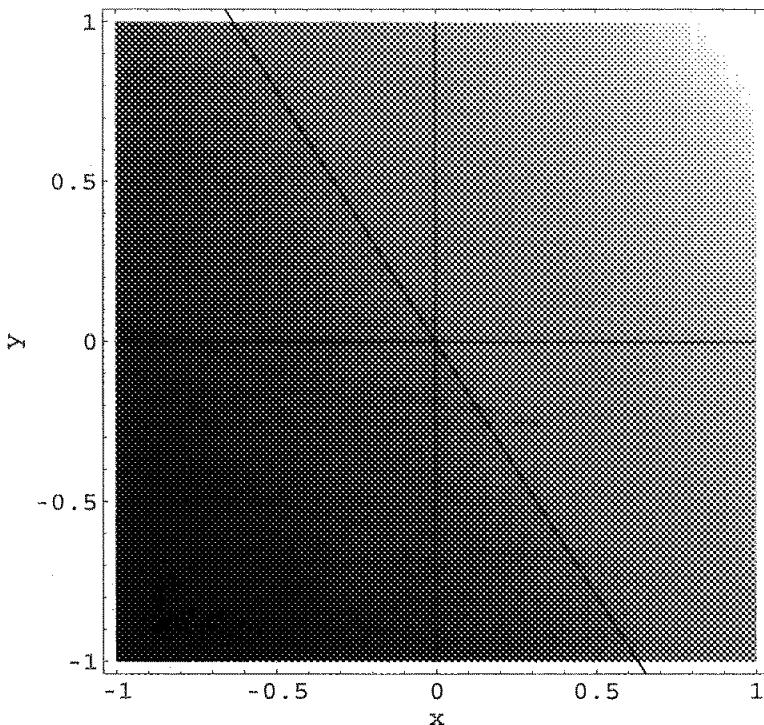


Fig. 10.8 The expected return of the portfolio is positive above the diagonal line shown in the density plot.

decides that a return of $E[R] = 2$ is desirable they will then select x and y so that the variance in the return is minimized over all portfolios for which $E[R] = 2$. This is yet another constrained optimization exercise. According to Eq. (10.32) the variance in the return is

$$\text{Var}(R) = 515.571x^2 + 748.802xy + 306.902y^2$$

Rather than use the method of Lagrange multipliers to find the minimum of $\text{Var}(R)$ we will take the more elementary approach of solving the equation $2 = 6.93489x + 4.3834y$ for y and substituting into the expression for variance. In this case we get the quadratic expression

$$\text{Var}(R) = 63.891 - 101.421x + 99.0737x^2$$

which is minimized when $x = 0.511844$. Under these conditions $y = -0.353511$ and the minimum variance is $\text{Var}(R) = 37.9349$.

10.8 Exercises

- (1) Using the definition of covariance and variance prove the first two statements of Theorem 10.1.
- (2) Show that for any two random variables X and Y

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

If X and Y are independent random variables, show this formula reduces to the result mentioned in Theorem 2.7.

- (3) Fill in the remaining details of the proof of Lemma 10.1 for the cases in which $E[X^2] = 0$, $E[X^2] = \infty$, $E[Y^2] = 0$, or $E[Y^2] = \infty$
- (4) The data shown in the table below was originally published in the *New York Times*, Section 3, pg. 1, 05/31/1998. It lists the name of a corporate CEO, the CEO's corporation, the CEO's golf handicap, and a rating of the stock of the CEO's corporation. Determine the covariance and correlation between the CEO's golf handicap, and the corporation's stock rating.

Name	Corp.	Handicap	Rating
Terrence Murray	Fleet Financial	10.1	67
William T. Esrey	Sprint	10.1	66
Hugh L. McColl Jr.	Nationsbank	11	64
James E. Cayne	Bear Stearns	12.6	64
John R. Stafford	Amer. Home Products	10.9	58
John B. McCoy	Banc One	7.6	58
Frank C. Herring	Transamerica	10.6	55
Ralph S. Larsen	Johnson&Johnson	16.1	54
Paul Hazen	Wells Fargo	10.9	54
Lawrence A. Bossidy	Allied Signal	12.6	51
Charles R. Shoemate	Bestfoods	17.6	49
James E. Perrella	Ingersoll-Rand	12.8	49
William P. Stiritz	Ralston Purina	13	48

- (5) Suppose that a company sells a European call option on a security. The standard deviation on the price of the call option is $\sigma_O = 0.045$. The company hedges its position by buying a different security for which the standard deviation in value is $\sigma_D = 0.037$. The correlation between the values of the two instruments is $\rho(O, D) = 0.86$. What is the optimal ratio of the number of shares of security purchased to the number of options sold so that the variance in the value of the portfolio is minimized?
- (6) Which of the following utility functions are concave on their domain?
- $u(x) = \ln x$
 - $u(x) = \ln(x^2)$
 - $u(x) = (\ln x)^2$
 - $u(x) = \tan^{-1} x$
- (7) The mean of a discrete random variable as defined in Eq. (2.4) is sometimes called the **arithmetic mean**. The **harmonic mean** is defined as

$$\mathcal{H} = \frac{1}{\sum_X \frac{P(X)}{X}}. \quad (10.33)$$

The **geometric mean** is defined as

$$\mathcal{G} = \prod_X X^{P(X)}. \quad (10.34)$$

Use Jensen's inequality (10.14) to show that $\mathcal{H} \leq \mathcal{G} \leq E[X]$ and that

equality holds only when $X_1 = X_2 = \dots = X_n$. You may assume that all the random variables are positive.

- (8) Verify inequality (10.15) for $f(x) = \tanh x$ and $\phi(t) = t$.
- (9) Find the certainty equivalent for the choice between (a) flipping a fair coin and either winning \$10 or losing \$2, or (b) receiving an amount C with certainty. Assume that your utility function is $f(x) = x - x^2/50$.
- (10) Suppose you must choose between receiving an amount C with certainty or playing a game in a fair die is rolled. If a prime number results you win \$15, but if a non-prime number results you lose that amount of money. Assume that your utility function is $f(x) = x - x^2/2$.
- (11) Suppose a risk-averse investor with utility function $u(x) = \ln x$ will invest a proportion α of their total capital x in an investment which will pay them either $2\alpha x$ with probability p or nothing with probability $1-p$. The amount of capital not invested will earn interest for one time period at the simple rate r . What proportion of their capital should the investor allocate if they wish to achieve the maximum expected utility?
- (12) Consider the function $u(x) = 1 - e^{-bx}$ where $b > 0$.
 - (a) Show that $u(x)$ is concave.
 - (b) Show that $0 \leq x_1 < x_2$ implies $u(x_1) < u(x_2)$.
- (13) Suppose an investor whose utility function is $u(x) = 1 - e^{-x/100}$ has a total of \$1000 to invest in two securities. The expected rate of return for the first investment is $r_1 = 0.08$ with a standard deviation in the rate of return of $\sigma_1 = 0.03$. The expected rate of return for the second investment is $r_2 = 0.13$ with a standard deviation in the rate of return of $\sigma_2 = 0.09$. The correlation in the rates of return is $\rho = -0.26$. What is the optimal amount of money to place into each investment?
- (14) Suppose that an investor will split a unit of wealth between the following securities possessing variances in their rates of return as listed in the table below. Assuming that the rates of return are uncorrelated, determine the proportion of the portfolio which will be allocated to each security so that the variance in the returned wealth is minimized.

Security	σ^2
A	0.24
B	0.41
C	0.27
D	0.16
E	0.33

- (15) For any real scalar c show that for $r(\mathbf{w})$ and $\sigma^2(\mathbf{w})$ as defined in Eqs. (10.18) and (10.19) the following conditions hold.

$$\begin{aligned} r(c\mathbf{w}) &= cr(\mathbf{w}) \\ \sigma^2(c\mathbf{w}) &= c^2\sigma^2(\mathbf{w}) \end{aligned}$$

- (16) Prove Lemma 10.3.
 (17) In this exercise we will extend the work that was done in exercise 14 by determining the optimal portfolio which minimizes the variance in the return under the condition that the portfolio is financed with money borrowed at the interest rate of 11%.

Security	r	σ^2
A	0.13	0.24
B	0.12	0.41
C	0.15	0.27
D	0.11	0.16
E	0.17	0.33

- (18) Verify that the solutions to the set of Eqs. (10.25) and (10.27) are given by the values of λ and x in (10.28) and (10.29).
 (19) Suppose the risk-free interest rate is 4.75% per year, the expected rate of return on the stock market is 7.65% per year, and the standard deviation in the return on the market is 22% per year. If the covariance in the expected returns on a particular stock and the market is 15%, determine the expected rate of return on the stock.
 (20) Suppose a share of a security has a current price of \$83 while the drift parameter and volatility of the price of the security are respectively 0.13 and 0.25 per annum. The risk free interest rate is 5.5% annually. Find the value of 3-month European call option on the security with a strike price of \$86. Use the mean-variance analysis of Sec. 10.7 to determine the positions of the optimal portfolio with an expected return of \$10.

Appendix A

Sample Stock Market Data

In this appendix are end-of-day closing stock market prices for Sony Corporation stock. The data were collected from the website www.siliconinvestor.com and cover the one year period of August 13, 2001 until August 12, 2002.

Date	Close
Aug-12-2002	42.86
Aug-9-2002	44.44
Aug-8-2002	44.0
Aug-7-2002	43.85
Aug-6-2002	42.5
Aug-5-2002	42.01
Aug-2-2002	42.71
Aug-1-2002	44.55
Jul-31-2002	45.33
Jul-30-2002	46.65
Jul-29-2002	46.14
Jul-26-2002	44.7
Jul-25-2002	45.49
Jul-24-2002	47.1
Jul-23-2002	45.21
Jul-22-2002	44.97
Jul-19-2002	46.3
Jul-18-2002	48.95
Jul-17-2002	47.82
Jul-16-2002	49.59

Date	Close
Jul-15-2002	50.3
Jul-12-2002	50.9
Jul-11-2002	51.2
Jul-10-2002	50.47
Jul-9-2002	52.75
Jul-8-2002	51.98
Jul-5-2002	53.17
Jul-3-2002	51.75
Jul-2-2002	50.0
Jul-1-2002	51.55
Jun-28-2002	53.1
Jun-27-2002	50.3
Jun-26-2002	48.95
Jun-25-2002	49.41
Jun-24-2002	50.1
Jun-21-2002	48.63
Jun-20-2002	50.23
Jun-19-2002	50.06
Jun-18-2002	51.8
Jun-17-2002	52.78
Jun-14-2002	52.45
Jun-13-2002	53.27
Jun-12-2002	54.31
Jun-11-2002	54.01
Jun-10-2002	54.36
Jun-7-2002	55.5
Jun-6-2002	55.5
Jun-5-2002	56.7
Jun-4-2002	55.91
Jun-3-2002	56.3
May-31-2002	58.11
May-30-2002	58.69
May-29-2002	57.8
May-28-2002	58.23

Date	Close
May-24-2002	59.4
May-23-2002	59.7
May-22-2002	59.56
May-21-2002	58.05
May-20-2002	57.65
May-17-2002	58.81
May-16-2002	56.48
May-15-2002	56.15
May-14-2002	55.6
May-13-2002	55.0
May-10-2002	54.75
May-9-2002	54.6
May-8-2002	55.14
May-7-2002	52.65
May-6-2002	54.0
May-3-2002	54.02
May-2-2002	53.45
May-1-2002	54.8
Apr-30-2002	54.2
Apr-29-2002	55.48
Apr-26-2002	55.5
Apr-25-2002	56.0
Apr-24-2002	53.98
Apr-23-2002	53.92
Apr-22-2002	53.93
Apr-19-2002	53.25
Apr-18-2002	53.87
Apr-17-2002	53.8
Apr-16-2002	53.7
Apr-15-2002	52.17
Apr-12-2002	50.86
Apr-11-2002	51.7
Apr-10-2002	52.72
Apr-9-2002	51.31

Date	Close
Apr-8-2002	52.35
Apr-5-2002	51.5
Apr-4-2002	52.71
Apr-3-2002	52.46
Apr-2-2002	51.15
Apr-1-2002	51.1
Mar-28-2002	51.7
Mar-27-2002	51.45
Mar-26-2002	50.85
Mar-25-2002	50.85
Mar-22-2002	52.3
Mar-21-2002	53.16
Mar-20-2002	52.01
Mar-19-2002	55.21
Mar-18-2002	53.66
Mar-15-2002	55.25
Mar-14-2002	55.09
Mar-13-2002	52.4
Mar-12-2002	53.9
Mar-11-2002	56.2
Mar-8-2002	56.59
Mar-7-2002	56.05
Mar-6-2002	53.4
Mar-5-2002	50.3
Mar-4-2002	50.7
Mar-1-2002	47.8
Feb-28-2002	46.2
Feb-27-2002	47.22
Feb-26-2002	45.75
Feb-25-2002	47.25
Feb-22-2002	46.71
Feb-21-2002	45.3
Feb-20-2002	45.4
Feb-19-2002	44.87

Date	Close
Feb-15-2002	44.85
Feb-14-2002	45.55
Feb-13-2002	44.85
Feb-12-2002	43.77
Feb-11-2002	43.45
Feb-8-2002	42.4
Feb-7-2002	42.5
Feb-6-2002	41.0
Feb-5-2002	41.75
Feb-4-2002	43.25
Feb-1-2002	43.5
Jan-31-2002	44.76
Jan-30-2002	44.35
Jan-29-2002	43.7
Jan-28-2002	45.2
Jan-25-2002	46.4
Jan-24-2002	45.86
Jan-23-2002	44.91
Jan-22-2002	43.56
Jan-18-2002	44.7
Jan-17-2002	45.51
Jan-16-2002	44.5
Jan-15-2002	45.91
Jan-14-2002	47.02
Jan-11-2002	47.55
Jan-10-2002	48.92
Jan-9-2002	48.0
Jan-8-2002	47.82
Jan-7-2002	49.31
Jan-4-2002	49.57
Jan-3-2002	47.25
Jan-2-2002	45.95
Dec-31-2001	45.1
Dec-28-2001	46.25

Date	Close
Dec-27-2001	44.0
Dec-26-2001	44.11
Dec-24-2001	44.36
Dec-21-2001	44.36
Dec-20-2001	45.5
Dec-19-2001	46.75
Dec-18-2001	47.0
Dec-17-2001	45.55
Dec-14-2001	46.2
Dec-13-2001	45.28
Dec-12-2001	47.3
Dec-11-2001	46.95
Dec-10-2001	46.61
Dec-7-2001	48.5
Dec-6-2001	49.86
Dec-5-2001	49.7
Dec-4-2001	46.8
Dec-3-2001	45.99
Nov-30-2001	47.7
Nov-29-2001	47.0
Nov-28-2001	46.11
Nov-27-2001	48.25
Nov-26-2001	48.55
Nov-23-2001	46.55
Nov-21-2001	46.49
Nov-20-2001	45.6
Nov-19-2001	48.14
Nov-16-2001	46.2
Nov-15-2001	45.3
Nov-14-2001	40.83
Nov-13-2001	40.5
Nov-12-2001	39.95
Nov-9-2001	39.8
Nov-8-2001	39.79

Date	Close
Nov-7-2001	39.3
Nov-6-2001	41.8
Nov-5-2001	41.2
Nov-2-2001	39.72
Nov-1-2001	39.14
Oct-31-2001	38.2
Oct-30-2001	38.12
Oct-29-2001	38.6
Oct-26-2001	40.33
Oct-25-2001	40.43
Oct-24-2001	40.84
Oct-23-2001	41.7
Oct-22-2001	41.74
Oct-19-2001	41.0
Oct-18-2001	40.5
Oct-17-2001	40.1
Oct-16-2001	40.75
Oct-15-2001	39.58
Oct-12-2001	40.01
Oct-11-2001	39.96
Oct-10-2001	37.0
Oct-9-2001	34.31
Oct-8-2001	35.7
Oct-5-2001	35.25
Oct-4-2001	35.3
Oct-3-2001	34.4
Oct-2-2001	33.72
Oct-1-2001	33.74
Sep-28-2001	33.2
Sep-27-2001	37.19
Sep-26-2001	36.25
Sep-25-2001	37.05
Sep-24-2001	38.7
Sep-21-2001	36.75

Date	Close
Sep-20-2001	37.56
Sep-19-2001	37.25
Sep-18-2001	36.35
Sep-17-2001	37.0
Sep-10-2001	40.4
Sep-7-2001	41.48
Sep-6-2001	42.2
Sep-5-2001	42.7
Sep-4-2001	44.0
Aug-31-2001	44.9
Aug-30-2001	45.92
Aug-29-2001	47.33
Aug-28-2001	47.75
Aug-27-2001	47.72
Aug-24-2001	47.91
Aug-23-2001	46.49
Aug-22-2001	48.55
Aug-21-2001	47.1
Aug-20-2001	48.25
Aug-17-2001	47.9
Aug-16-2001	50.48
Aug-15-2001	51.65
Aug-14-2001	50.1
Aug-13-2001	49.16

Appendix B

Solutions to Chapter Exercises

B.1 The Theory of Interest

- (1) The principal amount is $P = \$3659$, the annual interest rate is $r = 0.065$, the period is $t = 5$ years. Thus the final compound balance is

$$P(1 + r)^t = 3659(1 + 0.065)^5 = \$5013.15,$$

where the final balance has been rounded to the nearest cent.

- (2) The principal amount is $P = \$3993$, the annual interest rate is $r = 0.043$, the period is $t = 2$ years. If there are n compounding periods per year, the final compound balance is

$$A = P\left(1 + \frac{r}{n}\right)^{nt}.$$

- (a) If the interest is compounded monthly $n = 12$ and

$$A = 3993\left(1 + \frac{0.043}{12}\right)^{24} = \$4350.93$$

- (b) If the interest is compounded weekly $n = 52$ and

$$A = 3993\left(1 + \frac{0.043}{52}\right)^{104} = \$4351.44$$

- (c) If the interest is compounded daily $n = 365$ and

$$A = 3993\left(1 + \frac{0.043}{365}\right)^{730} = \$4351.57$$

- (d) If the interest is compounded continuously

$$A = Pe^{rt} = 3993e^{(0.043)(2)} = \$4351.60$$

- (3) The period is $t = 1$, the interest rate is $r = 0.08$ compounded $n = 4$ times per period. The effective simple annual interest rate R is

$$R = -1 + \left(1 + \frac{0.08}{4}\right)^4 = 0.08243216.$$

- (4) Since the competitor pays interest at rate $r = 0.0525$ compounded daily we will compute his effective simple annual interest rate and then determine the monthly compounded rate which matches it. The effective simple annual interest rate R is

$$R = -1 + \left(1 + \frac{0.0525}{365}\right)^{365} = 0.0538985832.$$

Thus we wish to solve the following equation for i .

$$\begin{aligned} 0.058985832 &= -1 + \left(1 + \frac{i}{12}\right)^{12} \\ 1.058985832 &= \left(1 + \frac{i}{12}\right)^{12} \\ \sqrt[12]{1.058985832} &= 1 + \frac{i}{12} \\ 1.004384268 &= 1 + \frac{i}{12} \\ 0.004384268 &= \frac{i}{12} \\ 0.052611220 &= i \end{aligned}$$

Thus if you pay an interest rate compounded monthly of at least 5.261122% you will compete favorably.

- (5) The continuously compounded interest-bearing account will increase in value more rapidly, thus we wish to solve the following equation.

$$1000e^{0.0475t} - 1000\left(1 + \frac{0.0475}{365}\right)^{365t} = 1$$

This equation cannot be solved algebraically for t , so we use Newton's method to approximate the solution, $t \approx 42.6583$ years.

- (6) We focus on finding the limit $\lim_{h \rightarrow 0} (1+h)^{1/h}$. If we let $y = (1+h)^{1/h}$, then

$$\begin{aligned}\ln y &= \ln((1+h)^{1/h}) \\ &= \frac{1}{h} \ln(1+h) \\ e^{\ln y} &= e^{\frac{1}{h} \ln(1+h)} \\ y &= e^{\frac{\ln(1+h)}{h}}.\end{aligned}$$

Therefore

$$\begin{aligned}\lim_{h \rightarrow 0} (1+h)^{1/h} &= \lim_{h \rightarrow 0} e^{\frac{\ln(1+h)}{h}} \\ &= e^{\lim_{h \rightarrow 0} \frac{\ln(1+h)}{h}}\end{aligned}$$

since the exponential function is continuous everywhere. The limit in the exponent is the definition of the derivative of $\ln x$ at $x = 1$. Thus the exponent limit is 1 and $\lim_{h \rightarrow 0} (1+h)^{1/h} = e^1 = e$.

- (7) To determine the preferable investment we must determine the present value of the payouts. Since the interest rate is $r = 0.0275$ compounded continuously we have

$$\begin{aligned}P_A &= 200e^{-0.0275} + 211e^{-0.055} + 198e^{-0.0825} + 205e^{-0.11} = 760.25 \\ P_B &= 198e^{-0.0275} + 205e^{-0.055} + 211e^{-0.0825} + 200e^{-0.11} = 760.12\end{aligned}$$

Thus investment A is preferable.

- (8) If the price of the house is \$200000 and your down payment is 20% then you will borrow $P = \$160000$. If the monthly payment on a $t = 30$ year, fixed rate mortgage should not exceed $x = \$1500$ then we can substitute these quantities into Eq. (1.6) and use Newton's Method to approximate r .

$$\begin{aligned}P &= x \frac{n}{r} \left(1 - \left[1 + \frac{r}{n} \right]^{-nt} \right) \\ 160000 &= 1500 \frac{12}{r} \left(1 - \left[1 + \frac{r}{12} \right]^{-(12)(30)} \right) \\ r &\approx 0.108034\end{aligned}$$

Thus the interest rate must not exceed 10.8034% annually.

- (9) Differentiating the function $f(r)$ given in Eq. (1.7) yields

$$f'(r) = \sum_{i=1}^n A_i(-i)(1+r)^{-i-1} = -\sum_{i=1}^n A_i(i)(1+r)^{-(i+1)} < 0$$

for $-1 < r < \infty$. Thus if there exist two distinct rates of return r_1 and r_2 then by the Mean Value Theorem $f'(r) = 0$ for some r between r_1 and r_2 , contradicting the inequality above.

- (10) Using Eq. (1.7) we setup the equation

$$10000 = \frac{2000}{1+r} + \frac{3000}{(1+r)^2} + \frac{4000}{(1+r)^3} + \frac{3000}{(1+r)^4}$$

whose solution is approximated using Newton's Method. The rate of return $r \approx 0.0717859$ or equivalently 7.17859% per year.

- (11) (a) The present value of initial investment and payout amounts for investment A is (assuming a simple annual interest rate of $r = 0.0433$), $-\$10.91$. The present value of the initial investment and payout amounts for investment B is $\$18.63$. Thus investment B is preferable if present value is the decision criterion.
 (b) The rate of return of investment A is $r_A = 0.039354$ while the rate of return of investment B is $r_B = 0.049921$. Thus investment B would be the preferable investment if rate of return were the deciding criterion.

B.2 Discrete Probability

- (1) Assuming the regular tetrahedron is fair so that it equally likely to land on any of its four faces, the probability of it landing on 3 is $p = 1/4$.
 (2) Let the ordered pair (m, n) represent the numeric outcomes of the rolling of the dice. The outcome of the first die is m and the outcome of the second die is n . If you are concerned about the problem of keeping track of which is the first die and which is the second, assume that the first die is painted green while the second is painted red. According to the Multiplication Rule the probability of the event (m, n) is

$$P(m, n) = P(\text{first die shows } m)P(\text{second die shows } n)$$

since the dice are independent of each other. Since each of the die has six equally likely simple outcomes, then $P(m, n) = 1/36$. The table of 36 outcomes is shown below.

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Since we are interested in the sums of the numbers shown on the upward faces of the dice we will tabulate them below.

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

Thus we can see the sample space of sums for the fair dice is the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Each of the entries in the previous table occurs with probability $1/36$ and therefore the probabilities of the outcomes of rolling a pair of fair dice are

Outcome	Prob.	Outcome	Prob.
2	$\frac{1}{36}$	7	$\frac{1}{6}$
3	$\frac{1}{18}$	8	$\frac{5}{36}$
4	$\frac{1}{12}$	9	$\frac{1}{9}$
5	$\frac{1}{9}$	10	$\frac{1}{12}$
6	$\frac{5}{36}$	11	$\frac{1}{18}$
		12	$\frac{1}{36}$

- (3) Let the probability that the batter strikes out in the first inning be denoted $P(1) = 1/3$. Let the probability that the batter strikes out

in the fifth inning be denoted $P(5) = 1/4$. Let the probability that the batter strikes out in both innings be denoted $P(1 \wedge 5) = 1/10$. According to the Addition Rule the probability that the batter strikes out in either the first or the fifth inning is

$$P(1 \vee 5) = P(1) + P(5) - P(1 \wedge 5) = \frac{1}{3} + \frac{1}{4} - \frac{1}{10} = \frac{29}{60}.$$

- (4) The eight possible outcomes of the experiment are listed in the table below.

		Person	1	2	3
		Hat	Red	Red	Red
		Red	Red	Blue	
		Red	Blue	Red	
		Blue	Red	Red	
		Red	Blue	Blue	
		Blue	Blue	Red	
		Blue	Red	Blue	
		Blue	Blue	Blue	

In 6 of the outcomes two of the three people will see their companions wearing mis-matched hats and pass while the third will see their two companions wearing matching hats opposite in color to their own hat's color. Thus one of the three will guess the correct color in these six cases. In the remaining two cases all three hats are of the same color, so the strategy of guessing the opposite color would fail. Thus the probability of winning under the strategy is $6/8 = 3/4 = 0.75$.

- (5) Since the cards are drawn without replacement the outcome of the second draw is dependent on the outcome of the first draw. Using the Multiplication Rule and conditional probability we have

$$\begin{aligned} P(2 \text{ aces}) &= P(\text{2nd ace} | \text{1st ace}) P(\text{1st ace}) \\ &= \left(\frac{3}{51}\right) \left(\frac{4}{52}\right) \\ &= \frac{1}{221}. \end{aligned}$$

- (6) Since the cards are drawn without replacement the outcome of the second draw is dependent on the outcome of the first draw. Using the Multiplication Rule and conditional probability we have

$$\begin{aligned} P(2\text{nd ace} \wedge 1\text{st not ace}) &= P(2\text{nd ace}|1\text{st not ace})P(1\text{st not ace}) \\ &= \left(\frac{4}{51}\right)\left(\frac{48}{52}\right) \\ &= \frac{16}{221}. \end{aligned}$$

- (7) Since the cards are drawn without replacement the outcome of the fourth draw is dependent on the outcome of the third draw which is dependent on the outcome of the second draw which is dependent on the outcome of the first draw. Using the Multiplication Rule and conditional probability we have

$$\begin{aligned} P(4 \text{ aces}) &= P(4\text{th ace}|3\text{rd ace}) \cdot P(3\text{rd ace}|2\text{nd ace}) \cdot \\ &\quad P(2\text{nd ace}|1\text{st ace}) \cdot P(1\text{st ace}) \\ &= \left(\frac{1}{49}\right)\left(\frac{2}{50}\right)\left(\frac{3}{51}\right)\left(\frac{4}{52}\right) \\ &= \frac{1}{270725}. \end{aligned}$$

- (8) If $n \in \{2, \dots, 49\}$ is the position of the first ace drawn then

$$P(n) = \frac{48 \cdot 47 \cdot 46 \cdots (50-n) \cdot 4}{52 \cdot 51 \cdot 50 \cdots (53-n)}.$$

When $n = 1$, $P(1) = 1/13$. By calculating $P(2), P(3), \dots, P(49)$ we obtain the following probabilities for the first ace appearing on the n^{th} draw.

n	$P(n)$	n	$P(n)$
2	$\frac{16}{221}$	26	$\frac{8}{833}$
3	$\frac{376}{5525}$	27	$\frac{92}{10829}$
4	$\frac{17296}{270725}$	28	$\frac{2024}{270725}$
5	$\frac{3243}{54145}$	29	$\frac{253}{38675}$
6	$\frac{3036}{54145}$	30	$\frac{44}{7735}$
7	$\frac{2838}{54145}$	31	$\frac{38}{7735}$
8	$\frac{1892}{38675}$	32	$\frac{228}{54145}$
9	$\frac{1763}{38675}$	33	$\frac{57}{15925}$
10	$\frac{328}{7735}$	34	$\frac{48}{15925}$
11	$\frac{164}{4165}$	35	$\frac{8}{3185}$
12	$\frac{152}{4165}$	36	$\frac{16}{7735}$
13	$\frac{703}{20825}$	37	$\frac{1}{595}$
14	$\frac{8436}{270725}$	38	$\frac{4}{2975}$
15	$\frac{222}{7735}$	39	$\frac{22}{20825}$
16	$\frac{12}{455}$	40	$\frac{44}{54145}$
17	$\frac{11}{455}$	41	$\frac{33}{54145}$
18	$\frac{352}{15925}$	42	$\frac{24}{54145}$
19	$\frac{5486}{270725}$	43	$\frac{12}{38675}$
20	$\frac{992}{54145}$	44	$\frac{8}{38675}$
21	$\frac{899}{54145}$	45	$\frac{1}{7735}$
22	$\frac{116}{7735}$	46	$\frac{4}{54145}$
23	$\frac{522}{38675}$	47	$\frac{2}{54145}$
24	$\frac{36}{2975}$	48	$\frac{4}{270725}$
25	$\frac{9}{833}$	49	$\frac{1}{270725}$

The draw with the highest probability of being the first ace is the first draw.

- (9) The outcomes of different spins of a roulette wheel are assumed to be independent, thus the probability that any spin of the wheel has an outcome of black is $P(\text{black}) = 9/19$.
- (10) The outcomes of different spins of a roulette wheel are assumed to be independent, thus the probability that any spin of the wheel has an outcome of 00 is $P(00) = 1/38$.
- (11) To be a probability distribution function

$$\sum_{x=1}^{10} f(x) = \sum_{x=1}^{10} \frac{c}{x} = c \sum_{x=1}^{10} \frac{1}{x} = \frac{7381c}{2520} = 1,$$

thus $c = 2520/7381$.

(12)

$$\begin{aligned} P(0 \text{ black}) &= \frac{5}{20} \cdot \frac{4}{19} \cdot \frac{3}{18} = \frac{1}{118} \\ P(1 \text{ black}) &= \frac{15}{20} \cdot \frac{5}{19} \cdot \frac{4}{18} + \frac{5}{20} \cdot \frac{15}{19} \cdot \frac{4}{18} + \frac{5}{20} \cdot \frac{4}{19} \cdot \frac{15}{18} = \frac{5}{38} \\ P(2 \text{ black}) &= \frac{15}{20} \cdot \frac{14}{19} \cdot \frac{5}{18} + \frac{15}{20} \cdot \frac{5}{19} \cdot \frac{14}{18} + \frac{5}{20} \cdot \frac{15}{19} \cdot \frac{14}{18} = \frac{35}{76} \\ P(3 \text{ black}) &= \frac{15}{20} \cdot \frac{14}{19} \cdot \frac{13}{18} = \frac{91}{228} \end{aligned}$$

- (13) This situation can be thought of as a binomial experiment with 25 trials and the probability of success on a single trial of 0.0016. Thus the probability that a manufacturing run will be rejected is

$$\begin{aligned} P(X \geq 2) &= \sum_{n=2}^{25} P(X = n) \\ &= \sum_{n=2}^{25} \binom{25}{n} (0.0016)^n (1 - 0.0016)^{25-n} \\ &\approx 0.000749405. \end{aligned}$$

- (14) The probability of a child being born female is $p = 100/205 = 20/41$. Thus the expected number of female children in a family of 6 total children is

$$\begin{aligned} \sum_{x=0}^6 (x \cdot P(x)) &= \sum_{x=0}^6 \left(x \cdot \binom{6}{x} \left(\frac{20}{41}\right)^x \left(1 - \frac{20}{41}\right)^{6-x} \right) \\ &= \frac{120}{41} \approx 2.92683. \end{aligned}$$

- (15) Using the probabilities calculated in exercise 8 and the definition of expected value we have the expected position of the first ace,

$$\mu = \sum_{n=1}^{49} (n \cdot P(n)) = \frac{53}{5} = 10.6.$$

- (16) By the definition of expected value

$$\begin{aligned}
 E[aX + b] &= \sum_X ((aX + b)P(X)) \\
 &= \sum_X (aXP(X) + bP(X)) \\
 &= \sum_X (aXP(X)) + \sum_X (bP(X)) \\
 &= a \sum_X (XP(X)) + b \sum_X P(X) \\
 &= aE[X] + b
 \end{aligned}$$

- (17) Using the probabilities calculated in exercise 8 and the expected value calculated in exercise 15, the variance in the position of the first ace drawn is

$$\sum_{n=1}^{49} (n^2 \cdot P(n)) - \mu^2 = \frac{901}{5} - \frac{2809}{25} = \frac{1696}{25} = 67.84.$$

This implies the standard deviation in the appearance of the first ace is approximately 8.2365.

- (18) By the definition of variance

$$\begin{aligned}
 \text{Var}(aX + b) &= E[(aX + b)^2] - (E[aX + b])^2 \\
 &= E[a^2X^2 + 2abX + b^2] - (aE[X] + b)^2 \\
 &= a^2E[X^2] + 2abE[X] + b^2 - a^2(E[X])^2 - 2abE[X] - b^2 \\
 &= a^2E[X^2] - a^2(E[X])^2 \\
 &= a^2\text{Var}(X)
 \end{aligned}$$

B.3 Normal Random Variables and Probability

- (1) The probability distribution function for X is given by the piecewise defined function

$$f(x) = \begin{cases} \frac{1}{5} & \text{if } -4 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$\mathrm{P}(X \geq 0) = \int_0^{\infty} f(x) dx = \int_0^1 \frac{1}{5} dx = \frac{1}{5}.$$

- (2) We know that for any random variable X with probability distribution function $f(x)$ we must have $\int_{-\infty}^{\infty} f(x) dx = 1$, thus

$$\begin{aligned}\int_{-\infty}^{\infty} f(x) dx &= \int_1^{\infty} \frac{c}{x^3} dx \\&= c \lim_{M \rightarrow \infty} \int_1^M \frac{1}{x^3} dx \\&= c \lim_{M \rightarrow \infty} \left(\frac{-1}{2x^2} \Big|_1^M \right) \\&= c \lim_{M \rightarrow \infty} \left(\frac{-1}{2M^2} + \frac{1}{2} \right) \\1 &= \frac{c}{2} \\c &= 2.\end{aligned}$$

(3)

$$\begin{aligned}1 &= \int_{-\infty}^{\infty} f(x) dx \\&= \int_{-\infty}^a f(x) dx + \int_a^{\infty} f(x) dx \\&= \mathrm{P}(X < a) + \mathrm{P}(X \geq a) \\1 - \mathrm{P}(X < a) &= \mathrm{P}(X \geq a)\end{aligned}$$

- (4) A random variable X has a continuous Cauchy distribution with probability density function

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

Show that the mean of this random variable does not exist.

By definition, the mean of a continuously distributed random variable is

$$\mu = \int_{-\infty}^{\infty} xf(x) dx.$$

Thus for the given probability distribution function

$$\begin{aligned}\mu &= \int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)} dx \\ &= \int_{-\infty}^0 \frac{x}{\pi(1+x^2)} dx + \int_0^{\infty} \frac{x}{\pi(1+x^2)} dx \\ &= \lim_{M \rightarrow -\infty} \int_M^0 \frac{x}{\pi(1+x^2)} dx + \lim_{N \rightarrow \infty} \int_0^N \frac{x}{\pi(1+x^2)} dx \\ &= \lim_{M \rightarrow -\infty} \left(\frac{1}{2\pi} \ln(1+x^2) \Big|_M^0 \right) + \lim_{N \rightarrow \infty} \left(\frac{1}{2\pi} \ln(1+x^2) \Big|_0^N \right) \\ &= \lim_{M \rightarrow -\infty} \left(\frac{-1}{2\pi} \ln(1+M^2) \right) + \lim_{N \rightarrow \infty} \left(\frac{1}{2\pi} \ln(1+N^2) \right).\end{aligned}$$

The improper integral above does not converge and hence the mean of the random variable does not exist.

(5)

$$\begin{aligned}\mathbb{E}[aX + b] &= \int_{-\infty}^{\infty} (ax + b)f(x) dx \\ &= \int_{-\infty}^{\infty} (axf(x) + bf(x)) dx \\ &= \int_{-\infty}^{\infty} (axf(x)) dx + \int_{-\infty}^{\infty} (bf(x)) dx \\ &= a \int_{-\infty}^{\infty} (xf(x)) dx + b \int_{-\infty}^{\infty} f(x) dx \\ &= a\mathbb{E}[X] + b\end{aligned}$$

- (6) **Proof.** Let X_1, X_2, \dots, X_k be continuous random variables with probability distribution function $f(x_1, x_2, \dots, x_k)$. By the definition of expected value

$$\begin{aligned} & \mathbb{E}[X_1 + X_2 + \cdots + X_k] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 + x_2 + \cdots + x_k) f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \\ &\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_2 f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \\ &\quad + \cdots + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_k f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \\ &= \int_{-\infty}^{\infty} x_1 f_1(x_1) dx_1 + \int_{-\infty}^{\infty} x_2 f_2(x_2) dx_2 + \cdots + \int_{-\infty}^{\infty} x_k f_k(x_k) dx_k \\ &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_k]. \end{aligned}$$

To keep the notation compact we have written the marginal probability density of X_i as $f_i(x_i)$ for $i = 1, 2, \dots, k$, where

$$f_i(x_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_k) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_k.$$

□

- (7) **Proof.** Let X_1, X_2, \dots, X_k be pairwise independent, continuous random variables with probability distribution function $f(x_1, x_2, \dots, x_k)$. Since the component random variables are assumed to be pairwise independent the joint probability distribution can be rewritten as

$$f(x_1, x_2, \dots, x_k) = f_1(x_1) f_2(x_2) \cdots f_k(x_k).$$

Thus by the definition of expected value

$$\begin{aligned}
 & \mathbb{E}[X_1 X_2 \cdots X_k] \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 x_2 \cdots x_k f(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 x_2 \cdots x_k f_1(x_1) f_2(x_2) \cdots f_k(x_k) dx_1 dx_2 \dots dx_k \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 f_1(x_1) x_2 f_2(x_2) \cdots x_k f_k(x_k) dx_1 dx_2 \dots dx_k \\
 &= \left(\int_{-\infty}^{\infty} x_1 f_1(x_1) dx_1 \right) \left(\int_{-\infty}^{\infty} x_2 f_2(x_2) dx_2 \right) \cdots \left(\int_{-\infty}^{\infty} x_k f_k(x_k) dx_k \right) \\
 &= \mathbb{E}[X_1] \mathbb{E}[X_2] \cdots \mathbb{E}[X_k].
 \end{aligned}$$

□

- (8) **Proof.** Let X be a continuous random variable with probability distribution function $f(x)$. Suppose the $\mathbb{E}[X] = \mu$. By the definition of variance

$$\begin{aligned}
 \text{Var}(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\
 &= \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x) dx \\
 &= \int_{-\infty}^{\infty} x^2 f(x) dx - 2\mu \int_{-\infty}^{\infty} x f(x) dx + \mu^2 \int_{-\infty}^{\infty} f(x) dx \\
 &= \mathbb{E}[X^2] - 2\mu^2 + \mu^2 \\
 &= \mathbb{E}[X^2] - \mu^2
 \end{aligned}$$

□

- (9) **Proof.** Let X be a continuous random variable with probability distribution function $f(x)$. Suppose $a, b \in \mathbb{R}$.

$$\begin{aligned}
 \text{Var}(aX + b) &= E[(aX + b)^2] - (E[aX + b])^2 \\
 &= \int_{-\infty}^{\infty} (ax + b)^2 f(x) dx - (aE[X] + b)^2 \\
 &= \int_{-\infty}^{\infty} (a^2x^2 + 2abx + b^2)f(x) dx \\
 &\quad - (a^2(E[X])^2 + 2abE[X] + b^2) \\
 &= a^2 \int_{-\infty}^{\infty} x^2 f(x) dx + 2ab \int_{-\infty}^{\infty} xf(x) dx + b^2 \int_{-\infty}^{\infty} f(x) dx \\
 &\quad - a^2(E[X])^2 - 2abE[X] - b^2 \\
 &= a^2E[X^2] + 2abE[X] + b^2 - a^2(E[X])^2 - 2abE[X] - b^2 \\
 &= a^2(E[X^2] - (E[X])^2) \\
 &= a^2\text{Var}(X)
 \end{aligned}$$

□

- (10) **Proof.** Let X_1, X_2, \dots, X_k be pairwise independent, continuous random variables with probability distribution function $f(x_1, x_2, \dots, x_k)$. Let the mean of X_i be μ_i for $i = 1, 2, \dots, k$. By

the definition of variance

$$\begin{aligned}
 & \text{Var}(X_1 + X_2 + \cdots + X_k) \\
 &= E[(X_1 + X_2 + \cdots + X_k)^2] - (E[X_1 + X_2 + \cdots + X_k])^2 \\
 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 + \cdots + x_k)^2 f(x_1, \dots, x_k) dx_1 \dots dx_k \\
 &\quad - (\mu_1 + \cdots + \mu_k)^2 \\
 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 + \cdots + x_k)^2 f_1(x_1) \cdots f_k(x_k) dx_1 \dots dx_k \\
 &\quad - (\mu_1 + \cdots + \mu_k)^2 \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1^2 + x_2^2 + \cdots + x_k^2 + 2x_1 x_2 \\
 &\quad + \cdots + 2x_{k-1} x_k) f_1(x_1) f_2(x_2) \cdots f_k(x_k) dx_1 dx_2 \dots dx_k \\
 &\quad - (\mu_1^2 + \mu_2^2 + \cdots + \mu_k^2 + 2\mu_1\mu_2 + \cdots + 2\mu_{k-1}\mu_k) \\
 &= E[X_1^2] + E[X_2^2] + \cdots + E[X_k^2] \\
 &\quad + 2E[X_1 X_2] + \cdots + 2E[X_{k-1} X_k] \\
 &\quad - (\mu_1^2 + \mu_2^2 + \cdots + \mu_k^2 + 2\mu_1\mu_2 + \cdots + 2\mu_{k-1}\mu_k) \\
 &= E[X_1^2] + E[X_2^2] + \cdots + E[X_k^2] \\
 &\quad + 2E[X_1] E[X_2] + \cdots + 2E[X_{k-1}] E[X_k] \\
 &\quad - (\mu_1^2 + \mu_2^2 + \cdots + \mu_k^2 + 2\mu_1\mu_2 + \cdots + 2\mu_{k-1}\mu_k) \\
 &= E[X_1^2] + E[X_2^2] + \cdots + E[X_k^2] + 2\mu_1\mu_2 + \cdots + 2\mu_{k-1}\mu_k \\
 &\quad - (\mu_1^2 + \mu_2^2 + \cdots + \mu_k^2 + 2\mu_1\mu_2 + \cdots + 2\mu_{k-1}\mu_k) \\
 &= E[X_1^2] - \mu_1^2 + E[X_2^2] - \mu_2^2 + \cdots + E[X_k^2] - \mu_k^2 \\
 &= \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_k)
 \end{aligned}$$

□

(11) The expected value of X is calculated as

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} xf(x) dx \\&= \int_{-1}^2 \frac{2}{5}x|x| dx \\&= \frac{2}{5} \left(\int_{-1}^0 (-x^2) dx + \int_0^2 x^2 dx \right) \\&= \frac{2}{5} \left(-\frac{x^3}{3} \Big|_{-1}^0 + \frac{x^3}{3} \Big|_0^2 \right) \\&= \frac{2}{15}(-1 + 8) \\&= \frac{14}{15}.\end{aligned}$$

The variance is calculated as

$$\begin{aligned}\text{Var}(X) &= \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 \\&= \int_{-1}^2 \frac{2}{5}x^2|x| dx - \frac{196}{225} \\&= \frac{2}{5} \left(\int_{-1}^0 (-x^3) dx + \int_0^2 x^3 dx \right) - \frac{196}{225} \\&= \frac{2}{5} \left(-\frac{x^4}{4} \Big|_{-1}^0 + \frac{x^4}{4} \Big|_0^2 \right) - \frac{196}{225} \\&= \frac{1}{10}(1 + 16) - \frac{196}{225} \\&= \frac{17}{10} - \frac{196}{225} \\&= \frac{373}{450}.\end{aligned}$$

(12) **Proof.** Suppose $m, n \in \mathbb{Z}$ and $n - m$ is even. This is equivalent to the supposition that $n - m = 2k$ for some $k \in \mathbb{Z}$.

$$\begin{aligned}n - m &= 2k \\n - m + 2m &= 2k + 2m \\n + m &= 2(k + m)\end{aligned}$$

Thus $n + m$ is even. Now suppose that $n + m$ is even, or equivalently, that $n + m = 2j$ for some $j \in \mathbb{Z}$,

$$\begin{aligned} n + m &= 2j \\ n + m - 2m &= 2j - 2m \\ n - m &= 2(j - m) \end{aligned}$$

Thus $n - m$ is even. \square

- (13) Using the method of integration by substitution, where $u = x^2/(4kt)$ we obtain

$$\begin{aligned} \int \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx &= \int \frac{kt}{\sqrt{k\pi t}} e^{-u} du \\ &= \sqrt{\frac{kt}{\pi}} \int e^{-u} du \\ &= -\sqrt{\frac{kt}{\pi}} e^{-u} + C \\ &= -\sqrt{\frac{kt}{\pi}} e^{-\frac{x^2}{4kt}} + C \end{aligned}$$

where C is a constant of integration.

- (14) Use the technique of integration by parts to evaluate

$$\int \frac{x^2}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx.$$

Using the technique of integration by parts with

$$\begin{aligned} u &= \frac{x}{\sqrt{k\pi t}} & v &= -kte^{-\frac{x^2}{4kt}} \\ du &= \frac{1}{\sqrt{k\pi t}} dx & dv &= \frac{x}{2} e^{-\frac{x^2}{4kt}} dx \end{aligned}$$

we obtain

$$\begin{aligned} \int \frac{x^2}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx &= uv - \int v du \\ &= -\frac{ktx}{\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} + \int \frac{kt}{\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}} dx \\ &= \sqrt{\frac{kt}{\pi}} \left(-xe^{-\frac{x^2}{4kt}} + \int e^{-\frac{x^2}{4kt}} dx \right) \end{aligned}$$

- (15) Assuming that $k > 0$ and $t > 0$ we have

$$\lim_{M \rightarrow \infty} 2ktMe^{-M^2/4kt} = \lim_{M \rightarrow \infty} \frac{2ktM}{e^{M^2/4kt}}.$$

The limit on the right-hand side of the equation is indeterminate of the form ∞/∞ and thus we apply L'Hôpital's Rule.

$$\begin{aligned}\lim_{M \rightarrow \infty} \frac{2ktM}{e^{M^2/4kt}} &= \lim_{M \rightarrow \infty} \frac{2kt}{\frac{M}{2kt} e^{M^2/4kt}} \\ &= \lim_{M \rightarrow \infty} \frac{(2kt)^2}{Me^{M^2/4kt}} \\ &= 0.\end{aligned}$$

- (16) The probability distribution function for the standard normal random variable is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

and thus

- (a) $P(-1 < X < 1) = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{-x^2/2} dx \approx 0.682689,$
- (b) $P(-2 < X < 2) = \frac{1}{\sqrt{2\pi}} \int_{-2}^2 e^{-x^2/2} dx \approx 0.9545,$
- (c) $P(-3 < X < 3) = \frac{1}{\sqrt{2\pi}} \int_{-3}^3 e^{-x^2/2} dx \approx 0.9973,$
- (d) $P(1 < X < 3) = \frac{1}{\sqrt{2\pi}} \int_1^3 e^{-x^2/2} dx \approx 0.157305.$

- (17) **Proof.** Suppose X is a normally distributed random variable with mean μ and variance σ^2 . We will let $z = g(x) = (x - \mu)/\sigma$.

$$\begin{aligned}E[Z] &= E[g(X)] \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{x - \mu}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx\end{aligned}$$

Making the substitution $u = (x - \mu)/\sigma$ yields

$$\begin{aligned}E[Z] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} ue^{-\frac{u^2}{2}} \sigma du \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ue^{-\frac{u^2}{2}} du \\ &= 0.\end{aligned}$$

Therefore

$$\begin{aligned}
 \text{Var}(Z) &= \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 \\
 &= \mathbb{E}[(g(X))^2] \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{x-\mu}{\sigma}\right)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}} \sigma du \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}} du \\
 &= 1.
 \end{aligned}$$

□

- (18) Think of the consecutive years as year 1 and 2. The rainfalls are normally distributed random variables, call them X_1 and X_2 in each year with means $\mu_1 = \mu_2 = 14$ and standard deviations $\sigma_1 = \sigma_2 = 3.2$. Let the random variable X be the sum of the rainfalls in two consecutive years, then $\mu = \mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = \mu_1 + \mu_2 = 28$. Likewise

$$\sigma^2 = \text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) = \sigma_1^2 + \sigma_2^2 = (3.2)^2 + (3.2)^2 = 20.48,$$

assuming that rainfalls in different years are independent. Now let $Z = (X - \mu)/\sigma$ and then

$$\begin{aligned}
 \mathbb{P}(X > 30) &= \mathbb{P}\left(Z > \frac{30 - 28}{\sqrt{20.48}}\right) \\
 &= \mathbb{P}(Z > 0.44192) \\
 &= 1 - \mathbb{P}(Z \leq 0.44192) \\
 &= 1 - \phi(0.44192) \\
 &\approx 0.329266
 \end{aligned}$$

- (19) Let the random variable X represent the ratio of consecutive days selling prices of the security.

$$\begin{aligned} P(X > 1) &= P(\ln X > 0) \\ &= P\left(Z > \frac{0 - \mu}{\sigma}\right) \\ &= P\left(Z > \frac{0 - 0.01}{0.05}\right) \\ &= P(Z > -0.20) \\ &\approx 0.57926 \end{aligned}$$

Thus the probability of a one-day increase in the price of the security is approximately 0.57926. Consequently the probability of a one-day decrease in the price of the security is approximately $1 - 0.57926 = 0.42074$. The probability of a four-day decrease in the selling price of the security is

$$\begin{aligned} P(X^4 < 1) &= P(4 \ln X < 0) \\ &= P\left(Z < \frac{0 - 4\mu}{2\sigma}\right) \\ &= P\left(Z < \frac{0 - 0.04}{0.10}\right) \\ &= P(Z < -0.40) \\ &\approx 0.344578. \end{aligned}$$

- (20) Since X is uniformly distributed on $[a, b]$ its probability distribution function is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

We will make use of Theorem 3.9.

$$\int_x^\infty f(t) dt = \begin{cases} 1 & \text{if } x < a \\ \frac{b-x}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{if } x > b \end{cases}$$

Thus

$$\begin{aligned} \mathbb{E}[(X - K)^+] &= \int_K^\infty \left(\int_x^\infty f(t) dt \right) dx \\ &= \begin{cases} \int_K^a 1 dx + \int_a^b \frac{b-x}{b-a} dx & \text{if } K < a \\ \int_K^b \frac{b-x}{b-a} dx & \text{if } a \leq K \leq b \\ 0 & \text{if } K > b \end{cases} \\ &= \begin{cases} \frac{a+b}{2} - K & \text{if } K < a \\ \frac{(b-K)^2}{2(b-a)} & \text{if } a \leq K \leq b \\ 0 & \text{if } K > b \end{cases} \end{aligned}$$

(21) By definition

$$\phi(-x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-t^2/2} dt.$$

Now we make the substitution $u = -t$.

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-t^2/2} dt &= -\frac{1}{\sqrt{2\pi}} \int_{\infty}^x e^{-u^2/2} du \\ &= \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} du \\ &= 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du \\ \phi(-x) &= 1 - \phi(x) \end{aligned}$$

(22) **Proof.** Using the definition of variance and Eqs. (3.15) and (3.17) we have

$$\begin{aligned} \text{Var}((X - K)^+) &= \mathbb{E}\left[\left((X - K)^+\right)^2\right] - (\mathbb{E}[(X - K)^+])^2 \\ &= ((\mu - 2K)^2 + \sigma^2) \phi\left(\frac{\mu - 2K}{\sigma}\right) + \frac{(\mu - 2K)\sigma}{\sqrt{2\pi}} e^{-(\mu - 2K)^2/2\sigma^2} \\ &\quad - \left(\frac{\sigma}{\sqrt{2\pi}} e^{-(\mu - K)^2/2\sigma^2} + (\mu - K)\phi\left(\frac{\mu - K}{\sigma}\right)\right)^2. \end{aligned}$$

□

- (23) ***Proof.*** Using the definition of variance and Eqs. (3.16) and (3.19) we have

$$\begin{aligned}\text{Var}((X - K)^+) &= \mathbb{E}[(X - K)^+)^2] - (\mathbb{E}[(X - K)^+])^2 \\ &= e^{2(\mu+\sigma^2)}\phi(w + 2\sigma) - 2Ke^{\mu+\sigma^2/2}\phi(w + \sigma) + K^2\phi(w) \\ &\quad - \left(e^{\mu+\sigma^2/2}\phi(w + \sigma) - K\phi(w)\right)^2\end{aligned}$$

where $w = (\mu - \ln K)/\sigma$. □

B.4 The Arbitrage Theorem

- (1) Converting from the odds against to probabilities that the outcomes may occur produces a table of probabilities like the following:

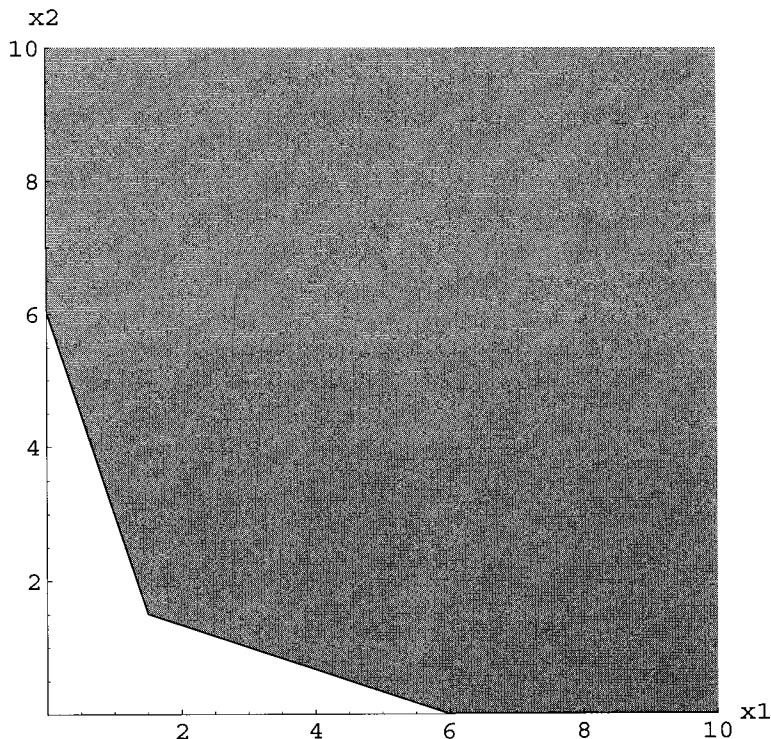
Outcome	Probability
A	$\frac{1}{3}$
B	$\frac{1}{4}$
C	$\frac{1}{2}$

Since the probabilities sum to $\frac{13}{12} > 1$ the Arbitrage Theorem guarantees the existence of a betting strategy which yields a positive payoff regardless of the outcome. Suppose we wager x , y , and z on outcomes A, B, and C respectively. The payoffs under different winning scenarios are listed in the next table.

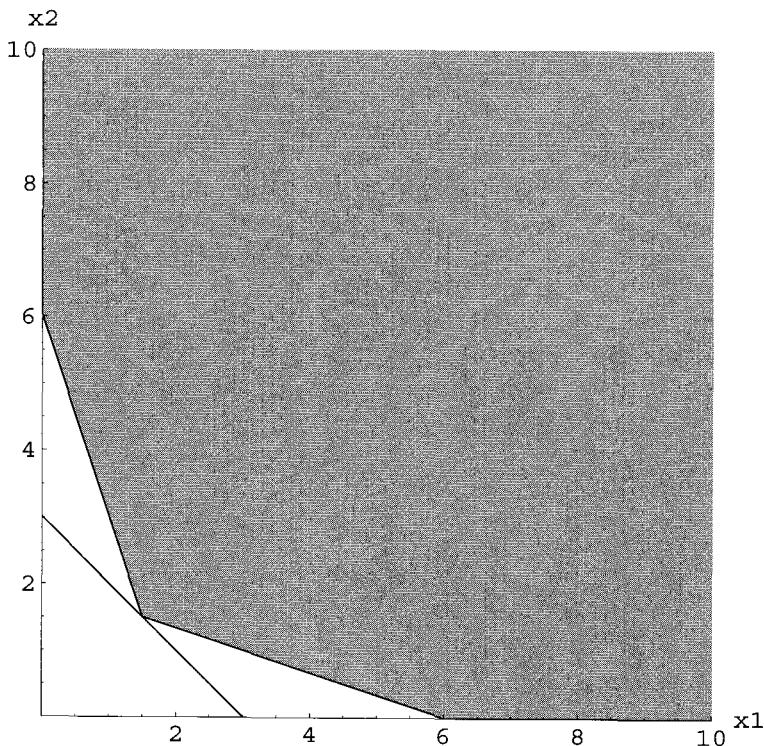
Winning Outcome	Payoff
A	$-2x + y + z$
B	$x - 3y + z$
C	$x + y - z$

We need only find values for x , y , and z which make the three expressions in the column on the right positive. By inspection we can see that $x = 1$, $y = 1$, and $z = 2$ is one solution to this set of inequalities.

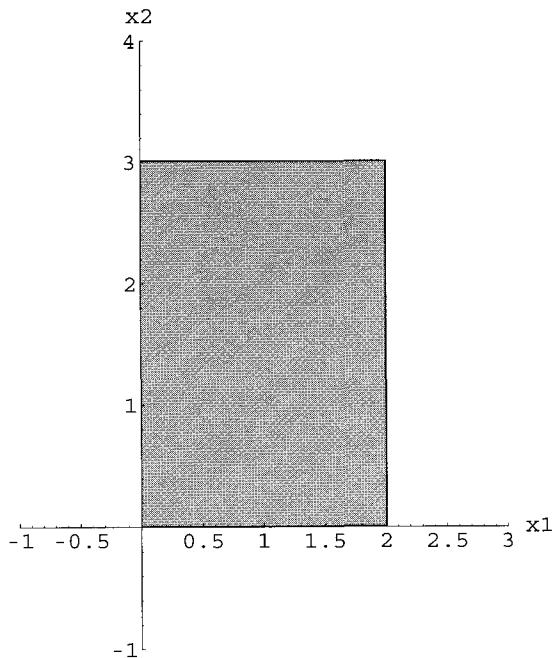
- (2) The region is convex and unbounded. There are three corners to the region located at $(0, 6)$, $(3/2, 3/2)$, and $(6, 0)$. The region is sketched below.



- (3) The cost function will be minimized at the smallest value of k for which the graph of $x_1 + x_2 = k$ intersects the feasible region sketched in exercise 2. This occurs when $(x_1, x_2) = (3/2, 3/2)$. Thus the minimum cost is $k = 3$.



- (4) $0 \leq x_1 \leq 2$ and $0 \leq x_2 \leq 3$



- (5) We can think of the solution as a system of inequalities:

$$\begin{aligned}x_1 &\geq 0 \\x_2 &\geq 0 \\x_3 &\geq 0 \\x_4 &\geq 0 \\x_1 + x_3 &= 2 \\x_2 + x_4 &= 3\end{aligned}$$

where x_3 and x_4 are the newly introduced slack variables. The last two equations can be re-written in matrix/vector form as

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \mathbf{b}$$

- (6) Let vector $\mathbf{c} = \langle 1, 1, 2 \rangle$ and $\mathbf{x} = \langle x_1, x_2, x_3 \rangle$. The problem as stated can be thought of as the primal problem:

$$\text{Minimize: } \mathbf{c} \cdot \mathbf{x} \quad \text{subject to } A\mathbf{x} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \mathbf{x} = 15 = b.$$

This is equivalent to the dual problem:

$$\text{Maximize: } by = 15y \quad \text{subject to } yA \leq \mathbf{c}.$$

From the inequality constraint in the dual problem we know that $y \leq 1/2$. Thus the maximum value of $15y$ is $15/2$. Thus the minimum value of $\mathbf{c} \cdot \mathbf{x} = 15/2$.

- (7) Let vector $\mathbf{c} = \langle 0, 7, 9, 0 \rangle$ and $\mathbf{x} = \langle x_1, x_2, x_3, x_4 \rangle$. The problem as stated can be thought of as the primal problem:

$$\text{Minimize: } \mathbf{c} \cdot \mathbf{x} \quad \text{subject to } A\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \mathbf{b}$$

This is equivalent to the dual problem:

$$\text{Maximize: } \mathbf{y} \cdot \mathbf{b} = [y_1 \ y_2] \mathbf{b} \quad \text{subject to } yA \leq \mathbf{c}.$$

The constraint inequality for the dual problem is equivalent to the following system of linear inequalities.

$$\begin{aligned} y_1 &\leq 0 \\ y_1 + y_2 &\leq 7 \\ y_1 + y_2 &\leq 9 \\ 2y_2 &\leq 0 \end{aligned}$$

The first and last of these inequalities imply strict inequality in the second and third. Thus by Theorem 4.2 $x_2 = x_3 = 0$. Therefore the primal problem can be restated as the following:

$$\text{Minimize: } \langle 0, 0 \rangle \cdot \langle x_1, x_4 \rangle \quad \text{subject to } \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

Thus the minimum value of the cost function is 0 and occurs where $\langle x_1, x_2, x_3, x_4 \rangle = \langle 5, 0, 0, 1/2 \rangle$.

- (8) In this case we must introduce a slack variable x_4 to convert the inequality constraint in the primal problem to an equality constraint. Thus we want to think of the primal problem as:

$$\begin{aligned} \text{Minimize } &x_1 + x_2 + x_3 \text{ subject to } 2x_1 + x_2 = 4 \text{ and } x_3 + x_4 = 6 \\ &\text{with } x_i \geq 0 \text{ for } i = 1, 2, 3, 4. \end{aligned}$$

If we define the following vectors and matrix

$$\mathbf{c} = \langle 1, 1, 1, 0 \rangle, \mathbf{x} = \langle x_1, x_2, x_3, x_4 \rangle, A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \text{ and } \mathbf{b} = \langle 1, 6 \rangle,$$

then the primal problem is the one of minimizing $\mathbf{c} \cdot \mathbf{x}$ subject to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. The dual to this problem is the one of maximizing $\mathbf{y} \cdot \mathbf{b}$ subject to $\mathbf{y}A \leq \mathbf{c}$ where $\mathbf{y} = \langle y_1, y_2 \rangle$. This is equivalent to maximizing $y_1 + 6y_2$ subject to the system of inequalities

$$\begin{aligned} 2y_1 &\leq 1 \\ y_1 &\leq 1 \\ y_2 &\leq 1 \\ y_2 &\leq 0. \end{aligned}$$

- (9) Paying attention to the inequality constraints of the dual problem indicates that $y_1 \leq 1/2$ and $y_2 \leq 0$. Thus by Theorem 4.2 $x_2 = x_3 = 0$. We can also see that the maximum of $\mathbf{y} \cdot \mathbf{b}$ will be $1/2$. This is also the minimum of the primal problem. However, we can also solve the primal problem directly after re-writing it as

$$\text{Minimize: } \langle 1, 0 \rangle \cdot \langle x_1, x_4 \rangle \text{ subject to } \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

From the constraint equation we see that now $\langle x_1, x_2, x_3, x_4 \rangle = \langle 2, 0, 0, 6 \rangle$ and the minimum value of the cost function is 2.

- (10) Written in matrix/vector form this dual problem becomes:

$$\text{Maximize } \mathbf{y} \cdot \mathbf{b} = \langle y_1, y_2, y_3 \rangle \cdot \langle 2, 0, 4 \rangle \text{ subject to}$$

$$\mathbf{y}A = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} \leq \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{c}.$$

Thus the primal problem is

$$\text{Minimize } \mathbf{c} \cdot \mathbf{x} = \langle 1, -1 \rangle \cdot \langle x_1, x_2 \rangle \text{ subject to}$$

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} = \mathbf{b}.$$

- (11) Find the solutions to the primal and dual problems of exercise 10.

We can see from the constraint equation for the primal problem that $\mathbf{x} = \langle x_1, x_2 \rangle = \langle 1, 1 \rangle$. Thus the minimum value of the cost function is

$\mathbf{c} \cdot \mathbf{x} = \langle 1, -1 \rangle \cdot \langle 1, 1 \rangle = 0$. Therefore the maximum of the dual problem is likewise 0.

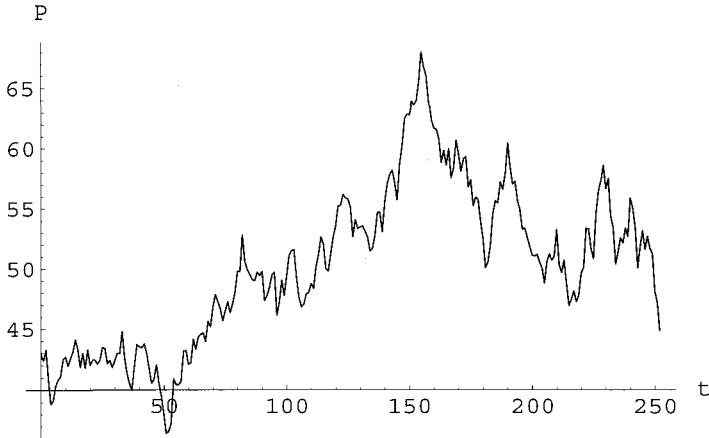
B.5 Random Walks and Brownian Motion

- (1) If f is a continuous, twice differentiable function whose second derivative is 0 on the interval $[0, A]$, then f must be a linear function of the form $f(x) = ax + b$ where a and b are constants. Since $f(0) = 0$ then $b = 0$. Since $f(A) = 1$ then $a = 1/A$. Thus we have $f(x) = x/A$. This agrees with our earlier derivation of exit probabilities for the discrete, symmetric random walk.
- (2) If g is a continuous, twice differentiable function whose second derivative is -2 on the interval $[0, A]$, then g must be a quadratic function of the form $g(x) = ax^2 + bx + c$ where a , b , and c are constants. Since $g(0) = 0$ then $c = 0$. Since $g''(x) = -2$ then $a = -1$ which implies $g(x) = x(b - x)$. Since $g(A) = 0$ then $b = A$. Thus we have $g(x) = x(A - x)$. This agrees with our earlier derivation of the conditional stopping times for the discrete, symmetric random walk.
- (3) This is an example of a symmetric random walk. Because of the spatial homogeneity property of the random walk, the desired probability is equivalent to determining the probability that a symmetric random walk initially in state $S(0) = 50$ will attain a value of 75 while avoiding the absorbing boundary at 0. Using Theorem 5.1, the probability is $P_{75}(50) = 50/75 = 2/3$.
- (4) As in exercise 3 the exit time is equivalent to the exit time of a symmetric random walk on the discrete interval $[0, 75]$ which begins in the state $S(0) = 50$. Using Theorem 5.2, the probability is $\Omega_{0,75}(50) = 50(75 - 50) = 1250$.
- (5) As in exercise 3 the conditional exit time is equivalent to the conditional exit time of a symmetric random walk on the discrete interval $[0, 75]$ which begins in the state $S(0) = 50$. We must set up and solve a tridiagonal linear system of the form shown in Eq. (5.15) where $A = 75$. This is best done with the aid of a computer or calculator since the linear system contains 76 unknowns (really only 74 since the boundary conditions are known). In this case $z_{50} = 6250/9 \approx 694.444$. Thus $\Omega_{75}(50) = 3125/3 \approx 1041.67$.

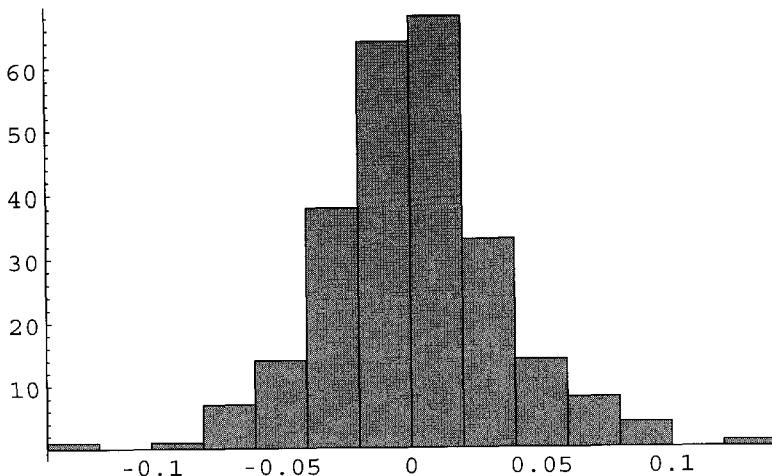
- (6) Let $P = P(0)e^{\mu t}$, then

$$\begin{aligned}\frac{dP}{dt} &= \frac{d}{dt}(P(0)e^{\mu t}) \\ &= P(0)\frac{d}{dt}(e^{\mu t}) \\ &= P(0)e^{\mu t}\frac{d}{dt}(\mu t) \\ &= P(0)e^{\mu t}\mu \\ &= \mu P.\end{aligned}$$

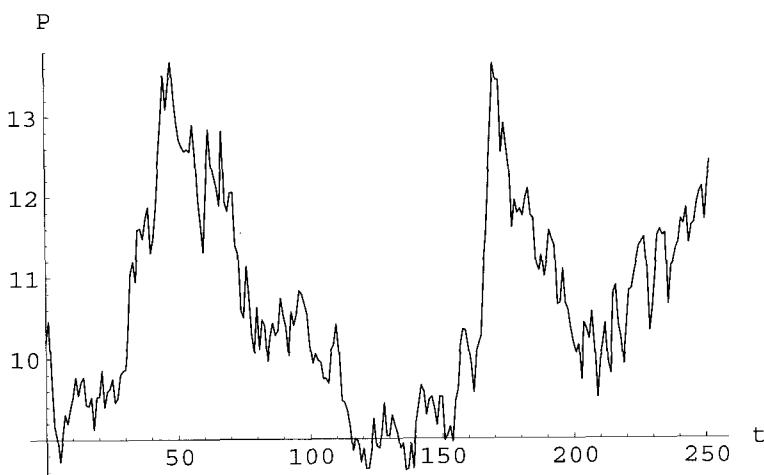
- (7) Various results are possible due to the fact that we are simulating a stochastic process. Using the data collected in Appendix A, the closing stock price on the last day for which data was collected was \$42.86. From the data the estimated drift parameter and volatility are $\mu = -0.000555 \text{ day}^{-1}$ and $\sigma = 0.028139 \text{ day}^{-1}$ respectively. If we substitute these values in Eq. (5.25) and set $P(t_0) = 42.86$ with $\Delta t = 1$ we can iteratively generate 252 (or any other amount for that matter) new days' closing prices. The random walk generated is pictured below.



- (8) From the website finance.yahoo.com the closing prices of stock for Continental Airlines, Inc. (symbol CAL) were captured for the time period of 06/15/2004 until 06/15/2005. A histogram of $\Delta X = \ln P(t_{i+1}) - \ln P(t_i)$ is shown below. The mean and standard deviation of ΔX are $\mu_{\Delta X} = 0.00121091 \text{ day}^{-1}$ and $\sigma_{\Delta X} = 0.0332183 \text{ day}^{-1}$ respectively.



- (9) Various results are possible due to the fact that we are simulating a stochastic process. The closing stock price on the last day for which data was collected was \$10.21. From the data the estimated drift parameter and volatility are $\mu = 0.00121091 \text{ day}^{-1}$ and $\sigma = 0.0332183 \text{ day}^{-1}$ respectively. If we substitute these values in Eq. (5.25) and set $P(t_0) = 10.21$ with $\Delta t = 1$ we can iteratively generate 252 (or any other amount for that matter) new days' closing prices. The random walk generated is pictured below.



(10) If $F \equiv F(y, z)$ and $f(x) = F(y_0 + xh, z_0 + xk)$ then

$$\begin{aligned} f'(x) &= F_y(y_0 + xh, z_0 + xk) \frac{dy}{dx} + F_z(y_0 + xh, z_0 + xk) \frac{dz}{dx} \\ &= F_y(y_0 + xh, z_0 + xk) \frac{d}{dx} [y_0 + xh] \\ &\quad + F_z(y_0 + xh, z_0 + xk) \frac{d}{dx} [z_0 + xk] \\ &= F_y(y_0 + xh, z_0 + xk)h + F_z(y_0 + xh, z_0 + xk)k \\ f'(0) &= F_y(y_0, z_0)h + F_z(y_0, z_0)k. \end{aligned}$$

Differentiating once again produces

$$\begin{aligned} f''(x) &= F_{yy}(y_0 + xh, z_0 + xk)h \frac{dy}{dx} + F_{yz}(y_0 + xh, z_0 + xk)h \frac{dz}{dx} \\ &\quad + F_{zy}(y_0 + xh, z_0 + xk)k \frac{dy}{dx} + F_{zz}(y_0 + xh, z_0 + xk)k \frac{dz}{dx} \\ &= hF_{yy}(y_0 + xh, z_0 + xk) \frac{d}{dx} [y_0 + xh] \\ &\quad + hF_{yz}(y_0 + xh, z_0 + xk) \frac{d}{dx} [z_0 + xk] \\ &\quad + kF_{zy}(y_0 + xh, z_0 + xk) \frac{d}{dx} [y_0 + xh] \\ &\quad + kF_{zz}(y_0 + xh, z_0 + xk) \frac{d}{dx} [z_0 + xk] \\ &= h^2 F_{yy}(y_0 + xh, z_0 + xk) + 2hkF_{yz}(y_0 + xh, z_0 + xk) \\ &\quad + k^2 F_{zz}(y_0 + xh, z_0 + xk) \\ f''(0) &= h^2 F_{yy}(y_0, z_0) + 2hkF_{yz}(y_0, z_0) + k^2 F_{zz}(y_0, z_0) \end{aligned}$$

By definition the Taylor remainder of order 3 for $f(x)$ is $R_3(x) = \frac{f'''(\alpha)}{3!}x^3$ for some α between 0 and x . If we notice the binomial coefficient pattern among the partial derivatives in the first and second derivatives of $f(x)$ then we can readily see that

$$\begin{aligned} f'''(x) &= h^3 F_{yyy}(y_0 + xh, z_0 + xk) + 3h^2 k F_{yyz}(y_0 + xh, z_0 + xk) \\ &\quad + 3hk^2 F_{yzz}(y_0 + xh, z_0 + xk) + k^3 F_{zzz}(y_0 + xh, z_0 + xk). \end{aligned}$$

From this expression we may directly determine the Taylor remainder of order 3.

- (11) If we apply Itô's Lemma with $a(P,t) = \mu P$ and $b(P,t) = \sigma P$ and $Y = P^n$, then

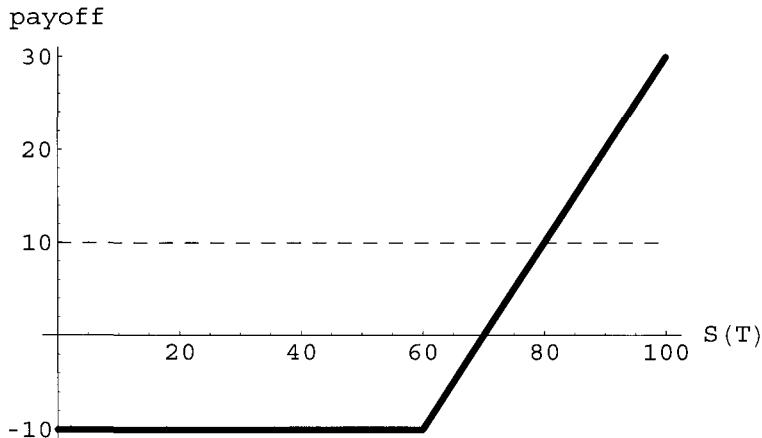
$$\begin{aligned} dY &= \left(\mu P n P^{n-1} + \frac{1}{2} (\sigma P)^2 n(n-1) P^{n-2} \right) dt + \sigma P n P^{n-1} dW(t) \\ &= \left(n \mu P^n + \frac{n(n-1)\sigma^2}{2} P^n \right) dt + n\sigma P^n dW(t) \end{aligned}$$

- (12) If we apply Itô's Lemma with $a(P,t) = \mu P$ and $b(P,t) = \sigma P$ and $Y = \ln P$, then

$$\begin{aligned} dY &= \left(\mu P \left[\frac{1}{P} \right] + \frac{1}{2} (\sigma P)^2 \left[\frac{-1}{P^2} \right] \right) dt + \sigma P \left[\frac{1}{P} \right] dW(t) \\ &= \left(\mu + \frac{\sigma^2}{2} \right) dt + \sigma dW(t) \end{aligned}$$

B.6 Options

- (1) Suppose the value of an American put option is P_a and that $P_a < P_e$, the value of a European put option on the same underlying security with identical strike price and exercise time. An investor could sell the European put option and buy the American put option and invest the net positive cash flow of $P_e - P_a$ in a risk-free bond at interest rate r . At the exercise time, in either the case where the owner of the European put exercises the option or not, the investor has a portfolio with value $(P_e - P_a)e^{rT} > 0$.
- (2) The net payoff will be the quantity $(S(T) - K)^+ - C$ where $K = 60$ and $C = 10$. The payoff is plotted below.



- (3) Suppose C is the value of a call option on a security whose value is S . Let the strike time of the option be T , the strike price be K , and the risk-free interest rate be r . Suppose at some time, say $t = 0$, it is the case that $C > S$, then an investor could take a long position in the security and a short position in the call option. This would generate cash for the investor in the amount of $C - S(0) > 0$ which could be invested in a risk-free bond at interest rate r . Thus at time t the investor owns a portfolio worth $S(t) + (C - S(0))e^{rt}$. If the owner of the call option chooses to exercise the option at time $0 \leq t^* \leq T$, then the investor sells the security for $\min\{S(t^*), K\}$. Hence at the time of exercise the investor's portfolio has a value of

$$S(t^*) + (C - S(0))e^{rt^*} - \min\{S(t^*), K\} > 0.$$

- (4) Since $C_e \geq S - Ke^{-rT}$ then when $S = 29$, $K = 26$, $r = 0.06$, and $T = 1/4$,

$$\begin{aligned} C_e &\geq 29 - 26e^{-(0.06)(0.25)} \\ &\geq 3.38709 \end{aligned}$$

- (5) Suppose $S = 31$, $T = 1/4$, $C_e = 3$, $K = 31$, and $r = 0.10$ then according to the Put-Call Parity Formula in Eq. (6.1):

$$\begin{aligned} P_e + S &= C_e + Ke^{-rT} \\ P_e + 31 &= 3 + 31e^{-(0.10)(0.25)} \\ P_e &= 3 + 31e^{-(0.10)(0.25)} - 31 \\ P_e &= 2.23461 \end{aligned}$$

- (6) Suppose $S = 31$, $T = 1/4$, $C_e = 3$, $P_e = 2.25$, and $r = 0.10$ then according to the Put-Call Parity Formula in Eq. (6.1):

$$\begin{aligned} P_e + S &= C_e + Ke^{-rT} \\ 2.25 + 31 &= 3 + Ke^{-(0.10)(0.25)} \\ 2.25 + 31 - 3 &= Ke^{-(0.10)(0.25)} \\ K &= (2.25 + 31 - 3)e^{(0.10)(0.25)} \\ K &= 31.0158 \end{aligned}$$

- (7) Suppose $T = 1/6$, $K = 14$, $S = 11$, and $r = 0.07$ then according to the Put-Call Parity Formula in Eq. (6.1):

$$\begin{aligned} P_e &= C_e - S + Ke^{-rT} \\ &\geq -S + Ke^{-rT} \\ &= -11 + 14e^{-(0.07)/6} \\ &= 2.83762 \end{aligned}$$

- (8) An investor could take a long position in the put option and the security and short position in the call option while borrowing, at the risk-free rate, the strike price of the options. This generates an initial cash flow of

$$K + C_e - P_e - S = 30 + 3 - 1 - 31 = \$1.$$

This amount could be invested at the risk-free rate until the strike time arrives. At $t = 3$ months if $S(3) > 30 = K$, then the call option will be exercised at the investor cancels all positions with a portfolio worth

$$1e^{(0.10)(0.25)} + 30 - 30e^{(0.10)(0.25)} = \$0.265862 > 0.$$

If the call option finishes out of the money, the investor exercises the put option and finishes with a portfolio worth again

$$1e^{(0.10)(0.25)} + 30 - 30e^{(0.10)(0.25)} = \$0.265862 > 0.$$

(9) If $f_1(S, t)$ and $f_2(S, t)$ both satisfy Eq. (7.5) then

$$\begin{aligned} rf_1 &= f_{1,t} + \frac{1}{2}\sigma^2 S^2 f_{1,ss} + rSf_{1,s} \quad \text{and} \\ rf_2 &= f_{2,t} + \frac{1}{2}\sigma^2 S^2 f_{2,ss} + rSf_{2,s}. \end{aligned}$$

Therefore if c_1 and c_2 are real numbers

$$\begin{aligned} rc_1 f_1 &= c_1 f_{1,t} + \frac{1}{2}\sigma^2 S^2 c_1 f_{1,ss} + rSc_1 f_{1,s} \quad \text{and} \\ rc_2 f_2 &= c_2 f_{2,t} + \frac{1}{2}\sigma^2 S^2 c_2 f_{2,ss} + rSc_2 f_{2,s}. \end{aligned}$$

Adding left-hand sides and right-hand sides of the previous two equations produces

$$\begin{aligned} rc_1 f_1 + rc_2 f_2 &= c_1 f_{1,t} + \frac{1}{2}\sigma^2 S^2 c_1 f_{1,ss} + rSc_1 f_{1,s} \\ &\quad + c_2 f_{2,t} + \frac{1}{2}\sigma^2 S^2 c_2 f_{2,ss} + rSc_2 f_{2,s} \\ r(c_1 f_1 + c_2 f_2) &= (c_1 f_{1,t} + c_2 f_{2,t}) + \frac{1}{2}\sigma^2 S^2 (c_1 f_{1,ss} + c_2 f_{2,ss}) \\ &\quad + rS(c_1 f_{1,s} + c_2 f_{2,s}) \\ &= (c_1 f_1 + c_2 f_2)_t + \frac{1}{2}\sigma^2 S^2 (c_1 f_1 + c_2 f_2)_{ss} \\ &\quad + rS(c_1 f_1 + c_2 f_2)_s \\ rf &= f_t + \frac{1}{2}\sigma^2 S^2 f_{ss} + rSf_s. \end{aligned}$$

(10) At time $t = T$ the payoff of the European put option will be

$$\max\{0, K - S(T)\} = (K - S(T))^+,$$

thus $F(S, T) = (K - S(T))^+$. If the stock is worthless (*i.e.*, $S = 0$) then the put option will be exercised and $F(0, t) = K$. As the price of the stock increases toward infinity the put option will not be exercised and thus $\lim_{S \rightarrow \infty} F(S, t) = 0$.

B.7 Solution of the Black-Scholes Equation

(1)

$$\begin{aligned}
 \hat{f}(w) &= \int_{-\infty}^{\infty} f(x)e^{-iwx} dx \\
 &= \int_0^{\infty} e^{-ax}e^{-iwx} dx \\
 &= \lim_{M \rightarrow \infty} \int_0^M e^{-(a+iw)x} dx \\
 &= \frac{-1}{a+iw} \lim_{M \rightarrow \infty} e^{-(a+iw)x} \Big|_0^M \\
 &= \frac{-1}{a+iw} \lim_{M \rightarrow \infty} (e^{-(a+iw)M} - 1)
 \end{aligned}$$

Since $0 \leq |e^{-iwM}|e^{-aM} \leq e^{-aM}$ then by the Squeeze Theorem (Theorem 2.7 of [Smith and Minton (2002)]),

$$\lim_{M \rightarrow \infty} (e^{-(a+iw)M} - 1) = -1 \implies \hat{f}(w) = \frac{1}{a+iw}.$$

(2)

$$\begin{aligned}
 f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(w)e^{iwx} dw \\
 &= \frac{1}{2\pi} \int_0^{\infty} e^{-aw}e^{iwx} dw \\
 &= \frac{1}{2\pi} \lim_{M \rightarrow \infty} \int_0^M e^{-(a-ix)w} dw \\
 &= \frac{1}{2\pi} \lim_{M \rightarrow \infty} \frac{-1}{a-ix} e^{-(a-ix)w} \Big|_0^M \\
 &= \frac{-1}{2\pi(a-ix)} \lim_{M \rightarrow \infty} (e^{-(a-ix)M} - 1)
 \end{aligned}$$

Since $0 \leq |e^{-ixM}|e^{-aM} \leq e^{-aM}$ then by the Squeeze Theorem (Theorem 2.7 of [Smith and Minton (2002)]),

$$\lim_{M \rightarrow \infty} (e^{-(a-ix)M} - 1) = -1 \implies f(x) = \frac{1}{2\pi(a-ix)}.$$

(3)

$$\begin{aligned}
f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(w) e^{iwx} dw \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-aw^2} e^{iwx} dw \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-a(w^2 - \frac{ix}{a}w)} dw \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-a(w^2 - \frac{ix}{a}w + (\frac{ix}{2a})^2)} e^{-x^2/4a} dw \\
&= \frac{1}{2\pi} e^{-x^2/4a} \int_{-\infty}^{\infty} e^{-a(w - \frac{ix}{2a})^2} dw \\
&= \frac{1}{2\pi} e^{-x^2/4a} \int_{-\infty}^{\infty} e^{-az^2} dz \\
&= \frac{1}{2\sqrt{\pi a}} e^{-x^2/4a}
\end{aligned}$$

(4)

$$\begin{aligned}
\frac{\partial}{\partial t} (F(S, t)) &= K \frac{\partial}{\partial t} (v(x, \tau)) \\
&= K \left(\frac{\partial v}{\partial x} \frac{dx}{dt} + \frac{\partial v}{\partial \tau} \frac{d\tau}{dt} \right) \\
&= K \left(\frac{\partial v}{\partial x} \cdot 0 + \frac{\partial v}{\partial \tau} \cdot \left[\frac{-\sigma^2}{2} \right] \right) \\
&= -\frac{K\sigma^2}{2} \frac{\partial v}{\partial \tau}
\end{aligned}$$

(5)

$$\begin{aligned}
\frac{\partial^2 F}{\partial S^2} &= \frac{\partial}{\partial S} \left(\frac{\partial F}{\partial S} \right) \\
&= \frac{\partial}{\partial S} (e^{-x} v_x) \quad \text{by Eq. (7.11)} \\
&= -e^{-x} v_x \frac{dx}{dS} + e^{-x} \left(v_{xx} \frac{dx}{dS} + v_{x\tau} \frac{d\tau}{dS} \right) \\
&= -e^{-x} v_x \cdot \frac{1}{S} + e^{-x} \left(v_{xx} \cdot \frac{1}{S} + v_{x\tau} \cdot 0 \right) \\
&= -\frac{e^{-x}}{S} (v_x - v_{xx}) \\
&= \frac{e^{-x}}{K e^x} (v_{xx} - v_x) \\
&= \frac{e^{-2x}}{K} (v_{xx} - v_x)
\end{aligned}$$

(6)

$$\begin{aligned}
rF &= F_t + \frac{1}{2} \sigma^2 S^2 F_{SS} + rSF_S \\
rKv &= -\frac{K\sigma^2}{2} v_\tau + \frac{1}{2} \sigma^2 S^2 \frac{e^{-2x}}{K} (v_{xx} - v_x) + rSe^{-x} v_x \\
&= -\frac{K\sigma^2}{2} v_\tau + \frac{1}{2} \sigma^2 K^2 e^{2x} \frac{e^{-2x}}{K} (v_{xx} - v_x) + rKe^x e^{-x} v_x \\
rv &= -\frac{\sigma^2}{2} v_\tau + \frac{1}{2} \sigma^2 (v_{xx} - v_x) + rv_x \\
\frac{\sigma^2}{2} v_\tau &= -rv + \frac{1}{2} \sigma^2 (v_{xx} - v_x) + rv_x \\
v_\tau &= -\frac{2r}{\sigma^2} v + v_{xx} - v_x + \frac{2r}{\sigma^2} v_x \\
&= -kv + v_{xx} - v_x + kv_x \\
&= v_{xx} + (k-1)v_x - kv
\end{aligned}$$

(7)

$$\left(e^{(k+1)(x+\sqrt{2\tau}y)/2} - e^{(k-1)(x+\sqrt{2\tau}y)/2} \right)^+ > 0$$

if and only if

$$\begin{aligned} e^{(k+1)(x+\sqrt{2\tau}y)/2} &> e^{(k-1)(x+\sqrt{2\tau}y)/2} \\ e^{x+\sqrt{2\tau}y} &> 1 \\ x + \sqrt{2\tau}y &> 0 \\ y &> -\frac{x}{\sqrt{2\tau}} \end{aligned}$$

(8) By the result of exercise 7 we know that

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(e^{(k+1)(x+\sqrt{2\tau}y)/2} - e^{(k-1)(x+\sqrt{2\tau}y)/2} \right)^+ e^{-y^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} \left(e^{(k+1)(x+\sqrt{2\tau}y)/2} - e^{(k-1)(x+\sqrt{2\tau}y)/2} \right) e^{-y^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{[(k+1)(x+\sqrt{2\tau}y)-y^2]/2} - e^{[(k-1)(x+\sqrt{2\tau}y)-y^2]/2} dy \end{aligned}$$

Completing the square in the exponents present in the integrand yields

$$\begin{aligned} &y^2 - \sqrt{2\tau}(k \pm 1)y - (k \pm 1)x \\ &= \left(y^2 - \sqrt{2\tau}(k \pm 1)y + \frac{(k \pm 1)^2\tau}{2} \right) - \frac{(k \pm 1)^2\tau}{2} - (k \pm 1)x \\ &= \left(y - \frac{(k \pm 1)\sqrt{2\tau}}{2} \right)^2 - \frac{(k \pm 1)^2\tau}{2} - (k \pm 1)x \end{aligned}$$

Thus the integral in Eq. (7.31) becomes

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \left[\int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{(k+1)\sqrt{2\tau}}{2})^2/2} e^{\frac{(k+1)^2\tau}{4} + (k+1)x/2} dy \right. \\ &\quad \left. - \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{(k-1)\sqrt{2\tau}}{2})^2/2} e^{\frac{(k-1)^2\tau}{4} + (k-1)x/2} dy \right] \\ &= \frac{e^{(k+1)^2\tau/4 + (k+1)x/2}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{(k+1)\sqrt{2\tau}}{2})^2/2} dy \\ &\quad - \frac{e^{(k-1)^2\tau/4 + (k-1)x/2}}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{(k-1)\sqrt{2\tau}}{2})^2/2} dy \end{aligned}$$

(9) If we let $w = -y + \frac{1}{2}(k+1)\sqrt{2\tau}$, then $dw = -dy$. When $y = -x/\sqrt{2\tau}$

then $w = x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau}$ and

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{1}{2}(k+1)\sqrt{2\tau})^2/2} dy \\ &= -\frac{1}{\sqrt{2\pi}} \int_{x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau}}^{-\infty} e^{-w^2/2} dw \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/\sqrt{2\tau} + \frac{1}{2}(k+1)\sqrt{2\tau}} e^{-w^2/2} dw \\ &= \phi\left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau}\right). \end{aligned}$$

- (10) If we let $w = -y + \frac{1}{2}(k-1)\sqrt{2\tau}$, then $dw = -dy$. When $y = -x/\sqrt{2\tau}$ then $w = x/\sqrt{2\tau} + \frac{1}{2}(k-1)\sqrt{2\tau}$ and

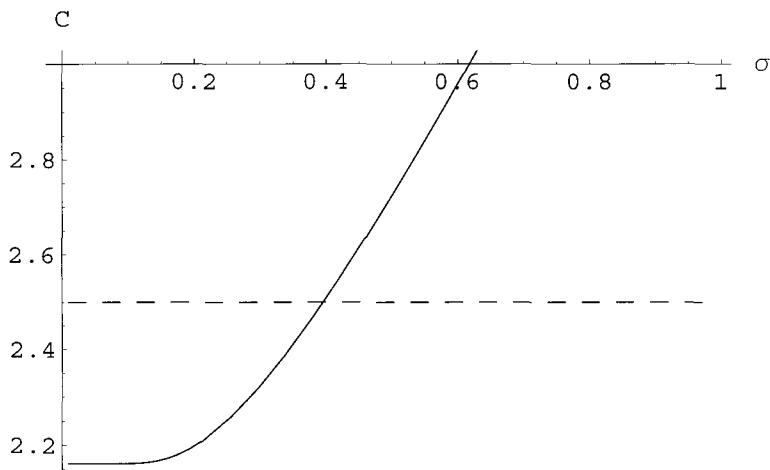
$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\tau}}^{\infty} e^{-(y - \frac{1}{2}(k-1)\sqrt{2\tau})^2/2} dy \\ &= -\frac{1}{\sqrt{2\pi}} \int_{x/\sqrt{2\tau} + \frac{1}{2}(k-1)\sqrt{2\tau}}^{-\infty} e^{-w^2/2} dw \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/\sqrt{2\tau} + \frac{1}{2}(k-1)\sqrt{2\tau}} e^{-w^2/2} dw \\ &= \phi\left(\frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k-1)\sqrt{2\tau}\right). \end{aligned}$$

- (11)

$$\begin{aligned} \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau} &= \frac{\ln(S/K)}{\sqrt{2 \cdot \frac{\sigma^2}{2}(T-t)}} + \frac{1}{2} \left(\frac{2r}{\sigma^2} + 1 \right) \sqrt{2 \cdot \frac{\sigma^2}{2}(T-t)} \\ &= \frac{\ln(S/K)}{\sqrt{\sigma^2(T-t)}} + \frac{1}{2} \left(\frac{2r}{\sigma^2} + 1 \right) \sqrt{\sigma^2(T-t)} \\ &= \frac{\ln(S/K)}{\sigma\sqrt{T-t}} + \left(\frac{r}{\sigma^2} + \frac{1}{2} \right) \sigma\sqrt{T-t} \\ &= \frac{\ln(S/K)}{\sigma\sqrt{T-t}} + \frac{\left(\frac{r}{\sigma^2} + \frac{1}{2} \right) \sigma^2(T-t)}{\sigma\sqrt{T-t}} \\ &= \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \end{aligned}$$

$$\begin{aligned}
 \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k-1)\sqrt{2\tau} &= \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k+1)\sqrt{2\tau} - \sqrt{2\tau} \\
 &= \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} - \sqrt{2 \cdot \frac{\sigma^2}{2}(T-t)} \\
 &= \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} - \sigma\sqrt{T-t}
 \end{aligned}$$

- (12) Using the parameter values specified $w \approx 0.536471$ and thus $C_e \approx \$5.05739$.
- (13) Using the parameter values specified $w \approx 0.16662$ and thus $P_e \approx \$6.40141$.
- (14) One way to find the volatility is to approximate it using Newton's Method (Sec. 3.2 of [Smith and Minton (2002)]). We wish to find a root of the equation $C_e(\sigma) = 2.50$. The graph below shows that the solution is near 0.80 which we will use as the initial approximation to the solution for Newton's Method.



Newton's Method approximates the solution as $\sigma \approx 0.827888$.

- (15) If $f(S, t) = S$, then $f_S = 1$, $f_{SS} = 0$, and $f_t = 0$. Thus

$$\begin{aligned} rf &= f_t + \frac{1}{2}\sigma^2 S^2 f_{SS} + rS f_S \\ rS &= 0 + \frac{1}{2}\sigma^2 S^2 \cdot 0 + rS \\ rS &= rS. \end{aligned}$$

Therefore the price of the security itself solve the Black-Scholes partial differential Eq. (7.5).

B.8 Derivatives of Black-Scholes Option Prices

- (1) Using Eq. (8.3) with $S = 300$, $K = 310$, $T = 1/4$, $t = 0$, $r = 0.03$, and $\sigma = 0.25$ we have

$$w = -0.139819$$

$$\Theta = -33.281$$

- (2) To find Θ for a European put option we will make use of the Put/Call parity formula (Eq. (6.1)) and Eq. (8.3). According to the Put/Call parity formula,

$$\frac{\partial P}{\partial t} = \frac{\partial C}{\partial t} + K r e^{-r(T-t)}.$$

Substituting the expression already found for $\frac{\partial C}{\partial t}$ in the right-hand

side of this equation yields

$$\begin{aligned}
 \frac{\partial P}{\partial t} &= \frac{Se^{-w^2/2} - Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\
 &\quad - Ke^{-r(T-t)} \left(r\phi(w - \sigma\sqrt{T-t}) + \frac{\sigma e^{-(w-\sigma\sqrt{T-t})^2/2}}{2\sqrt{2\pi(T-t)}} \right) \\
 &\quad + Kre^{-r(T-t)} \\
 &= \frac{Se^{-w^2/2} - Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\
 &\quad + \left(1 - \phi(w - \sigma\sqrt{T-t}) \right) - \frac{\sigma Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sqrt{2\pi(T-t)}} \\
 &= \frac{Se^{-w^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\
 &\quad + Kre^{-r(T-t)}\phi(\sigma\sqrt{T-t} - w) \\
 &\quad - \frac{Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{\sigma(T-t)} - r/\sigma - \sigma/2 + \sigma \right) \\
 &= \frac{Se^{-w^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r - \sigma^2/2 \right) \\
 &\quad + Kre^{-r(T-t)}\phi(\sigma\sqrt{T-t} - w) \\
 &\quad - \frac{Ke^{-r(T-t)-(w-\sigma\sqrt{T-t})^2/2}}{2\sigma\sqrt{2\pi(T-t)}} \left(\frac{\ln(S/K)}{T-t} - r + \sigma^2/2 \right)
 \end{aligned}$$

- (3) Using Eq. (8.5) with $S = 275$, $K = 265$, $T = 1/3$, $t = 0$, $r = 0.02$, and $\sigma = 0.20$ we have

$$w = 0.436257$$

$$\Theta = -15.3073$$

- (4) According to Eq. (7.35)

$$w = \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}},$$

which implies that

$$\begin{aligned}\frac{\partial w}{\partial S} &= \frac{\partial}{\partial S} \left[\frac{\ln(S/K) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T-t}} \right] \\&= \frac{\partial}{\partial S} \left[\frac{\ln(S/K)}{\sigma\sqrt{T-t}} \right] \\&= \frac{1}{\sigma\sqrt{T-t}} \frac{\partial}{\partial S} [\ln(S/K)] \\&= \frac{1}{\sigma\sqrt{T-t}} \frac{\partial}{\partial S} [\ln S - \ln K] \\&= \frac{1}{\sigma\sqrt{T-t}} \cdot \frac{1}{S} \\&= \frac{1}{\sigma S \sqrt{T-t}}.\end{aligned}$$

- (5) Using Eq. (8.8) with $S = 150$, $K = 165$, $T = 5/12$, $t = 0$, $r = 0.025$, and $\sigma = 0.22$ we have

$$\begin{aligned}w &= -0.526797 \\ \Delta &= 0.299167\end{aligned}$$

- (6) Using Eq. (8.9) with $S = 125$, $K = 140$, $T = 2/3$, $t = 0$, $r = 0.055$, and $\sigma = 0.15$ we have

$$\begin{aligned}w &= -0.564706 \\ \Delta &= -0.713863\end{aligned}$$

- (7) Using Eq. (8.11) with $S = 180$, $K = 175$, $T = 1/3$, $t = 0$, $r = 0.0375$, and $\sigma = 0.30$ we have

$$\begin{aligned}w &= 0.321416 \\ \Gamma &= 0.0121519\end{aligned}$$

- (8) According to Eq. (7.35)

$$w = \frac{\ln(S/K) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T-t}},$$

which implies that

$$\begin{aligned}
 \frac{\partial w}{\partial \sigma} &= \frac{\partial}{\partial \sigma} \left[\frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma \sqrt{T-t}} \right] \\
 &= \frac{\sigma(T-t)\sigma\sqrt{T-t} - (\ln(S/K) + (r + \sigma^2/2)(T-t))\sqrt{T-t}}{\sigma^2(T-t)} \\
 &= \frac{\sigma^2(T-t)^{3/2} - (\ln(S/K) + (r + \sigma^2/2)(T-t))\sqrt{T-t}}{\sigma^2(T-t)} \\
 &= \frac{\sigma^2(T-t) - (\ln(S/K) + (r + \sigma^2/2)(T-t))}{\sigma^2\sqrt{T-t}} \\
 &= \frac{\sigma^2(T-t)}{\sigma^2\sqrt{T-t}} - \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma^2\sqrt{T-t}} \\
 &= \sqrt{T-t} - \frac{1}{\sigma} \left(\frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \right) \\
 &= \sqrt{T-t} - \frac{w}{\sigma}
 \end{aligned}$$

- (9) Using Eq. (8.14) with $S = 300$, $K = 305$, $T = 1/2$, $t = 0$, $r = 0.0475$, and $\sigma = 0.25$ we have

$$w = 0.129235$$

$$\mathcal{V} = 83.9247$$

- (10) Using Eq. (8.19) with $S = 270$, $K = 272$, $T = 1/6$, $t = 0$, $r = 0.0375$, and $\sigma = 0.15$ we have

$$w = 0.012164$$

$$\rho = -23.4071$$

- (11) Using Eq. (8.19) with $S = 305$, $K = 325$, $T = 1/3$, $t = 0$, $r = 0.0255$, and $\sigma = 0.35$ we have

$$w = -0.171209$$

$$\rho = 38.0758$$

B.9 Hedging

- (1) If the call option is in the money then $S > K$. Start by considering the following limit.

$$\begin{aligned}\lim_{t \rightarrow T^-} w &= \lim_{t \rightarrow T^-} \left(\frac{\ln(S/K) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}} \right) \\ &= \lim_{t \rightarrow T^-} \left(\frac{\ln(S/K)}{\sigma\sqrt{T - t}} + \frac{(r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}} \right) \\ &= \lim_{t \rightarrow T^-} \left(\frac{\ln(S/K)}{\sigma\sqrt{T - t}} + \frac{(r + \sigma^2/2)}{\sigma} \sqrt{T - t} \right) \\ &= +\infty + 0 \\ &= +\infty\end{aligned}$$

We know the limit is $+\infty$ since $\ln(S/K) > 0$ for an in the money call option.

Now since $\Delta_C = \phi(w)$ and $\phi(w)$ is a continuous function then

$$\begin{aligned}\lim_{t \rightarrow T^-} \Delta &= \lim_{t \rightarrow T^-} \phi(w) \\ &= \phi \left(\lim_{t \rightarrow T^-} w \right) \\ &= \lim_{w \rightarrow \infty} \phi(w) \\ &= 1.\end{aligned}$$

- (2) If the call option is out of the money then $S < K$. Start by considering the following limit.

$$\begin{aligned}\lim_{t \rightarrow T^-} w &= \lim_{t \rightarrow T^-} \left(\frac{\ln(S/K) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}} \right) \\ &= \lim_{t \rightarrow T^-} \left(\frac{\ln(S/K)}{\sigma\sqrt{T - t}} + \frac{(r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}} \right) \\ &= \lim_{t \rightarrow T^-} \left(\frac{\ln(S/K)}{\sigma\sqrt{T - t}} + \frac{(r + \sigma^2/2)}{\sigma} \sqrt{T - t} \right) \\ &= -\infty + 0 \\ &= -\infty\end{aligned}$$

We know the limit is $-\infty$ since $\ln(S/K) < 0$ for an out of the money call option.

Now since $\Delta_C = \phi(w)$ and $\phi(w)$ is a continuous function then

$$\begin{aligned}\lim_{t \rightarrow T^-} \Delta &= \lim_{t \rightarrow T^-} \phi(w) \\ &= \phi\left(\lim_{t \rightarrow T^-} w\right) \\ &= \lim_{w \rightarrow -\infty} \phi(w) \\ &= 0.\end{aligned}$$

(3) These calculations are similar to those used to create Table 9.1.

Week	S	Δ	Shares Held	Interest Cost	Cumulative Cost
0	45.00	0.408940	2045	80	92025
1	44.58	0.366642	1833	72	82654
2	46.55	0.526428	2632	104	119919
3	47.23	0.581970	2910	115	133152
4	47.62	0.614837	3074	122	141077
5	47.28	0.583262	2916	116	133729
6	49.60	0.782988	3915	159	183396
7	50.07	0.824866	4124	168	194019
8	47.79	0.635479	3177	129	148930
9	48.33	0.698867	3494	142	164379
10	48.81	0.761174	3806	156	179750
11	51.36	0.954430	4772	199	229520
12	52.06	0.986327	4932	206	238048
13	51.98	0.995748	4979	208	240697
14	54.22	1.0	5000	210	242044
15	54.31	1.0	5000	0	242254

- (4) These calculations are similar to those used to create Table 9.2.

Week	S	Δ	Shares Held	Interest Cost	Cumulative Cost
0	45.00	0.408940	2045	80	92025
1	44.58	0.366642	1833	72	82654
2	45.64	0.447836	2239	88	101255
3	44.90	0.374621	1873	74	84909
4	43.42	0.238290	1191	48	55370
5	42.23	0.140594	703	30	34810
6	41.18	0.073032	365	18	20921
7	41.52	0.073130	366	18	20981
8	41.94	0.075921	380	19	21586
9	42.72	0.097654	488	23	26219
10	44.83	0.254169	1271	53	61343
11	44.93	0.235189	1176	49	57128
12	44.19	0.114050	570	26	30399
13	41.77	0.001621	8	6	6950
14	39.56	0.0	0	6	6640
15	40.62	0.0	0	0	6645

- (5) In this case the put option will not be exercised since the final price of the stock exceeds the strike price. In other words, the put option finishes out of the money. At the time the Put option is issued its price is \$2.73256. Delta for a European Put option is always negative, thus the financial institution will take a short position in the stock in creating the hedge. The revenue generated by the sale of the Put option and the short position in the stock will be invested in a bond at the risk-free interest rate $r = 0.045 \text{ yr}^{-1}$.

Week	S	Δ	Shares Shorted	Interest Earned	Cumulative Earnings
0	45.00	-0.591060	2955	0	132975
1	44.58	-0.633358	3167	115	142541
2	46.55	-0.473572	2368	123	105471
3	47.23	-0.418030	2090	91	92432
4	47.62	-0.385163	1926	80	84703
5	47.28	-0.416738	2084	73	92246
6	49.60	-0.217012	1085	80	42776
7	50.07	-0.175134	876	37	32348
8	47.79	-0.364521	1823	28	77633
9	48.33	-0.301133	1506	67	62380
10	48.81	-0.238826	1194	54	47205
11	51.36	-0.045570	228	41	-2368
12	52.06	-0.013673	68	-2	-10699
13	51.98	-0.004252	21	-9	-13152
14	54.22	0	0	-11	-14302
15	54.31	0	0	-12	-14314

Assuming the buyer of the option pays for the option at the strike time the net cost to the financial institution is

$$(5000)(2.73256) - 14314 = -651.20.$$

- (6) In this exercise the Put option will be exercised since the final stock price is smaller than the strike price.

Week	S	Δ	Shares Shorted	Interest Earned	Cumulative Earnings
0	45.00	-0.591060	2955	0	132975
1	44.58	-0.633358	3167	115	142541
2	45.64	-0.552164	2761	123	124135
3	44.90	-0.625379	3127	107	140676
4	43.42	-0.761710	3809	122	170410
5	42.23	-0.859406	4297	148	191166
6	41.18	-0.926968	4635	166	205250
7	41.52	-0.926870	4634	178	205386
8	41.94	-0.924079	4620	178	204977
9	42.72	-0.902346	4512	177	200540
10	44.83	-0.745831	3729	174	165612
11	44.93	-0.764811	3824	143	170024
12	44.19	-0.885950	4430	147	196950
13	41.77	-0.998379	4992	171	220595
14	39.56	-1.0	5000	191	221103
15	40.62	-1.0	5000	191	221294

At the strike time the financial institution must honor the Put option and cancel its short position in the stock. Thus the financial institution purchases 5000 shares of the stock at the strike price of \$47 per share. This final transaction changes the holdings of the financial institution by

$$221294 - (5000)(47) + (5000)(2.73256) = -43.20.$$

- (7) Let $S = 85$, $K = 88$, $r = 0.055$, and $\sigma = 0.17$. We will let the number of four-month options be w_4 and the number of six-month options be w_6 . If \mathcal{P} represents the value of the portfolio consisting of a short position in the four-month options and a long position in the six-month options, then the value of the portfolio is $\mathcal{P} = w_4 C(S, \frac{1}{3}) - w_6 C(S, \frac{1}{2})$. Note that $C(85, \frac{1}{3}) = 1.2972$ and $C(85, \frac{1}{2}) = 2.59313$. Gamma for the portfolio is

$$\begin{aligned}\Gamma &= \Gamma_4 w_4 - \Gamma_6 w_6 \\ &= 0.0474901 w_4 - 0.0390443 w_6.\end{aligned}$$

Thus the portfolio is Gamma-neutral whenever $w_4 = 0.822155 w_6$.

- (8) Using the results of exercise 7 and including a position of x shares of the security, the value of the portfolio is now $\mathcal{P} = Sx + w_4C(S, \frac{1}{3}) - w_6C(S, \frac{1}{2})$. Delta for the portfolio is

$$\begin{aligned}\Delta &= x + \Delta_4 w_4 - \Delta_6 w_6 \\ &= x + 0.45322 w_4 - 0.500131 w_6 \\ &= x + (0.45322)(0.822155)w_6 - 0.500131 w_6 \\ &= x - 0.127514 w_6.\end{aligned}$$

Thus the portfolio is both Gamma- and Delta-neutral whenever $x = 0.127514 w_6$.

- (9) Let $S = 95$, $r = 0.045$, and $\sigma = 0.23$. We will let the number of three-month options be w_3 and the number of five-month options be w_5 . If \mathcal{P} represents the value of the portfolio consisting of a short position in the three-month options and a long position in the five-month options, then the value of the portfolio is $\mathcal{P} = w_3C(S, \frac{1}{4}) - w_5C(S, \frac{5}{12})$. Note that $C(95, \frac{1}{4}) = 1.74863$ and $C(95, \frac{5}{12}) = 3.08418$. Gamma for the portfolio is

$$\begin{aligned}\Gamma &= \Gamma_3 w_3 - \Gamma_5 w_5 \\ &= 0.0365043 w_3 - 0.0282844 w_5.\end{aligned}$$

Thus the portfolio is Gamma-neutral whenever $w_3 = 0.774825 w_5$.

- (10) Using the results of exercise 9 and including a position of x shares of the security, the value of the portfolio is now $\mathcal{P} = Sx + w_3C(S, \frac{1}{4}) - w_5C(S, \frac{5}{12})$. Delta for the portfolio is

$$\begin{aligned}\Delta &= x + \Delta_3 w_3 - \Delta_5 w_5 \\ &= x + 0.489693 w_3 - 0.496454 w_5 \\ &= x + (0.489693)(0.774825) w_5 - 0.496454 w_5 \\ &= x - 0.117028 w_5.\end{aligned}$$

Thus the portfolio is both Gamma- and Delta-neutral whenever $x = 0.117028 w_5$.

B.10 Optimizing Portfolios

(1) Proof of statement 1:

$$\begin{aligned}\text{Cov}(X, X) &= \mathbb{E}[X \cdot X] - \mathbb{E}[X]\mathbb{E}[X] \quad (\text{by Eq. (10.2)}) \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= \text{Var}(X) \quad (\text{by Theorem (2.6)})\end{aligned}$$

Proof of statement 2:

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \quad (\text{by Eq. (10.2)}) \\ &= \mathbb{E}[YX] - \mathbb{E}[Y]\mathbb{E}[X] \\ &= \text{Cov}(Y, X) \quad (\text{by Eq. (10.2)})\end{aligned}$$

(2) Let X and Y be random variables, then by Theorem 10.1

$$\begin{aligned}\text{Var}(X + Y) &= \text{Cov}(X + Y, X + Y) \\ &= \mathbb{E}[(X + Y)(X + Y)] - \mathbb{E}[X + Y]\mathbb{E}[X + Y] \\ &= \mathbb{E}[X^2 + 2XY + Y^2] - (\mathbb{E}[X] + \mathbb{E}[Y])(\mathbb{E}[X] + \mathbb{E}[Y]) \\ &= \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] \\ &\quad - (\mathbb{E}[X])^2 - 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[Y])^2 \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 + \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\ &\quad + 2(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]) \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).\end{aligned}$$

Recall that if X and Y are independent random variables $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ and thus $\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$. Therefore, if X and Y are independent random variables

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

(3) Suppose that $\mathbb{E}[X^2] = 0$, then by definition

$$\mathbb{E}[X^2] = \sum_{i=1}^n x_i^2 \cdot P(X = x_i).$$

If $x_i \neq 0$ for some i , then $P(X = x_i) = 0$. Therefore we must have $P(X = 0) = 1$. In other words $X = 0$ which implied $XY = 0$ and thus

$$0 = (\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2] = 0.$$

A similar statement can be made when $E[Y^2] = 0$.

Suppose that $E[X^2] = \infty$, then

$$(E[XY])^2 < \infty = E[X^2] E[Y^2].$$

A similar statement can be made when $E[Y^2] = \infty$.

- (4) Let X be the golf handicap of the CEO and let Y be the stock rating of the CEO's corporation. Then we may calculate the following quantities:

$$E[X] = 11.9923$$

$$E[Y] = 56.6923$$

$$E[XY] = 671.069$$

$$\text{Var}(X) = 6.8741$$

$$\text{Var}(Y) = 45.5641$$

Thus we have

$$\text{Cov}(X, Y) = 671.069 - (11.9923)(56.6923) = -8.80207$$

$$\rho(X, Y) = \frac{-8.80207}{\sqrt{(6.8741)(45.5641)}} = -0.497354.$$

- (5) If n is the ratio of the number of shares of security D purchased compared to the number of call options C sold, then the value of the portfolio is $\mathcal{P} = C - nD$ which implies the variance in the value of the portfolio is

$$\begin{aligned}\text{Var}(\mathcal{P}) &= n^2 \text{Var}(D) - 2n\text{Cov}(C, D) + \text{Var}(C) \\ &= (0.037)^2 n^2 - 2n(0.86)(0.045)(0.037) + (0.045)^2 \\ &= 0.001369n^2 - 0.0028638n + 0.002025.\end{aligned}$$

This quadratic function is minimized when $n = 1.04595$.

- (6) According to Theorem 10.4, a twice differentiable function $u(x)$, is concave when $u''(x) \leq 0$.

- (a) $u(x) = \ln x$

$$u'(x) = \frac{1}{x}$$

$$u''(x) = -\frac{1}{x^2} < 0$$

for $0 < x < \infty$.

- (b) $u(x) = \ln(x^2)$ Since $\ln(x^2) = 2\ln x$, then by the previous result $\ln(x^2)$ is concave for $0 < x < \infty$.
 (c) $u(x) = (\ln x)^2$

$$u'(x) = \frac{2 \ln x}{x}$$

$$u''(x) = \frac{2(1 - \ln x)}{x^2}$$

This function is not concave on its whole domain. It is concave only for x such that $\ln x \geq 1$, in other words only for $x \geq e$.

- (d) $u(x) = \tan^{-1} x$

$$u'(x) = \frac{1}{1+x^2}$$

$$u''(x) = -\frac{2x}{(1+x^2)^2}$$

This function is not concave on its whole domain. It is concave only for $x \geq 0$.

- (7) According to Jensen's inequality (10.14), since $f(x) = \ln x$ is a concave function then

$$\begin{aligned} \ln \left(\sum_X X P(x) \right) &\geq \sum_X \ln X P(X) \\ &= \sum_X \ln \left(X^{P(X)} \right) \\ &= \ln \left(\prod_X X^{P(X)} \right). \end{aligned}$$

Since the exponential function is monotone then

$$E[X] = e^{\ln(\sum_X X P(x))} \geq e^{\ln(\prod_X X^{P(X)})} = G.$$

Now if $X > 0$ then $T = \frac{1}{X} > 0$. By the previous result

$$\begin{aligned}\prod_X T^{P(X)} &\leq E[T] \\ &= \sum_X T P(X) \\ \frac{1}{\prod_X X^{P(X)}} &\leq \sum_X \frac{P(X)}{X} \\ \prod_X X^{P(X)} &\geq \frac{1}{\sum_X \frac{P(X)}{X}} \\ \mathcal{G} &\geq \mathcal{H}.\end{aligned}$$

Thus we have shown that $\mathcal{H} \leq \mathcal{G} \leq E[X]$.

Now suppose that all the random variable X takes on only one value (with probability 1). Then we have

$$X = \mathcal{H} = \mathcal{G} = E[X].$$

Before finishing the exercise we must establish another inequality, an obvious one.

$$E[X] = \sum_X X P(X) \leq \max_X \{X\}$$

Equality holds if and only if X takes on only one value. From this inequality follows

$$\begin{aligned}\sum_X \frac{P(X)}{X} &\leq \max_X \left\{ \frac{1}{X} \right\} \\ \frac{1}{\sum_X \frac{P(X)}{X}} &\geq \min_X \{X\}.\end{aligned}$$

Again equality holds if and only if X takes on only one value. Thus the following string of inequalities is true.

$$\min_X \{X\} \leq \mathcal{H} \leq \mathcal{G} \leq E[X] \leq \max_X \{X\}$$

Now if $\min_X \{X\} = \mathcal{H}$ then X takes on only one value and we have $\min_X \{X\} = \max_X \{X\}$ and thus

$$\min_X \{X\} = \mathcal{H} = \mathcal{G} = E[X] = \max_X \{X\}$$

(8) $\phi(t) = t.$

$$\begin{aligned}
 \int_0^1 f(\phi(t)) dt &= \int_0^1 \tanh t dt \\
 &= \ln(\cosh t)|_0^1 \\
 &= \ln(\cosh 1) - \ln(\cosh 0) \\
 &= \ln(\cosh 1) \\
 &\approx 0.433781 \\
 &< 0.462117 \\
 &\approx \tanh\left(\frac{1}{2}\right) \\
 &= \tanh\left(\frac{t^2}{2}|_0^1\right) \\
 &= \tanh\left(\int_0^1 t dt\right) \\
 &= f\left(\int_0^1 \phi(t) dt\right)
 \end{aligned}$$

(9) The certainty equivalent is the solution C of the equation

$$\begin{aligned}
 \frac{1}{2}(f(10) + f(-2)) &= f(C) \\
 \frac{1}{2}\left(10 - \frac{10^2}{50} + \left[-2 - \frac{(-2)^2}{50}\right]\right) &= C - \frac{C^2}{50} \\
 \frac{1}{2}\left(8 - \frac{52}{25}\right) &= C - \frac{C^2}{50} \\
 \frac{74}{25} &= C - \frac{C^2}{50} \\
 C &= 25 \pm 3\sqrt{53}
 \end{aligned}$$

The certainty equivalent will be the smaller of the two roots of the quadratic equation, *i.e.*, $C = 25 - 3\sqrt{53} \approx 3.15967.$

- (10) The certainty equivalent is the solution C of the equation

$$\begin{aligned} \frac{1}{3}f(15) + \frac{2}{3}f(-15) &= f(C) \\ \frac{1}{3}\left(15 - \frac{15^2}{2}\right) + \frac{2}{3}\left(-15 - \frac{(-15)^2}{2}\right) &= C - \frac{C^2}{2} \\ \frac{1}{3}\left(\frac{-195}{2}\right) + \frac{2}{3}\left(\frac{-255}{2}\right) &= C - \frac{C^2}{2} \\ \frac{1}{3}\left(\frac{-195}{2}\right) + \frac{2}{3}\left(\frac{-255}{2}\right) &= C - \frac{C^2}{2} \\ -\frac{535}{4} &= C - \frac{C^2}{2} \\ C &= \frac{2 \pm \sqrt{1074}}{2} \end{aligned}$$

The certainty equivalent will be the smaller of the two roots of the quadratic equation, i.e., $C = \frac{2-\sqrt{1074}}{2} \approx -15.386$.

- (11) If we let X represent the random variable of the investor's return on the investment, then

$$X = \begin{cases} 2\alpha x + 1.11(1 - \alpha)x & \text{with probability } p, \\ 1.11(1 - \alpha)x & \text{with probability } 1 - p \end{cases}$$

If the investor's utility function is $u(x) = \ln x$, then the expected value of the utility function is

$$E[u(X)] = pu(2\alpha x + 1.11(1 - \alpha)x) + (1 - p)u(1.11(1 - \alpha)x).$$

This function is maximized when

$$\begin{aligned} 0 &= \frac{d}{d\alpha} E[u(X)] \\ &= \frac{p(2x - 1.11x)}{2\alpha x + 1.11(1 - \alpha)x} + \frac{(1 - p)(-1.11x)}{1.11(1 - \alpha)x} \\ \alpha &= 0.011236(200p - 111) \end{aligned}$$

- (12) Let $u(x) = 1 - e^{-bx}$ where $b > 0$, then $u'(x) = be^{-bx}$ and $u''(x) = -b^2e^{-bx}$. Since $b > 0$ then $-b^2 < 0$ and thus $u''(x) \leq 0$ for all x . Hence $u(x)$ is concave.

The function $u(x)$ is monotone increasing since $u'(x) > 0$ for all x . Therefore if $0 \leq x_1 < x_2$ then $u(x_1) < u(x_2)$.

- (13) Let W represent the wealth returned to the investor and let y represent the amount of money invested in the first investment ($0 \leq y \leq 1000$).

$$\begin{aligned} E[W] &= 1000 + 0.08y + 0.13(1000 - y) \\ &= 1130 - 0.05y \\ \text{Var}(W) &= (0.03)^2 y^2 + (0.09)^2 (1000 - y)^2 \\ &\quad + 2y(1000 - y)(0.03)(0.09)(-0.26) \\ &= 0.010404y^2 - 17.604y + 8100 \end{aligned}$$

The investor's utility function $u(x) = 1 - e^{-x/100}$, is maximized when $E[W] - \text{Var}(W)/200$ is maximized.

$$E[W] - \text{Var}(W)/200 = -0.00005202y^2 + 0.03802y + 1089.5$$

This quadratic expression is maximized when $y \approx 365.436$.

- (14) According to Theorem 10.7

$$\begin{aligned} \alpha_A &= \frac{\frac{1}{0.24}}{\frac{1}{0.24} + \frac{1}{0.41} + \frac{1}{0.27} + \frac{1}{0.16} + \frac{1}{0.33}} = 0.212697 \\ \alpha_B &= \frac{\frac{1}{0.41}}{\frac{1}{0.24} + \frac{1}{0.41} + \frac{1}{0.27} + \frac{1}{0.16} + \frac{1}{0.33}} = 0.124505 \\ \alpha_C &= \frac{\frac{1}{0.27}}{\frac{1}{0.24} + \frac{1}{0.41} + \frac{1}{0.27} + \frac{1}{0.16} + \frac{1}{0.33}} = 0.189064 \\ \alpha_D &= \frac{\frac{1}{0.16}}{\frac{1}{0.24} + \frac{1}{0.41} + \frac{1}{0.27} + \frac{1}{0.16} + \frac{1}{0.33}} = 0.319045 \\ \alpha_E &= \frac{\frac{1}{0.33}}{\frac{1}{0.24} + \frac{1}{0.41} + \frac{1}{0.27} + \frac{1}{0.16} + \frac{1}{0.33}} = 0.154689 \end{aligned}$$

- (15) If c is a real number then

$$\begin{aligned} r(c\mathbf{w}) &= E[R(c\mathbf{w})] \\ &= E\left[\sum_{i=1}^n cw_i(R_i - r)\right] \quad (\text{by Eq. (10.17)}) \\ &= E\left[c \sum_{i=1}^n w_i(R_i - r)\right] \\ &= E[cR(\mathbf{w})] \quad (\text{by Eq. (10.17)}) \\ &= cE[R(\mathbf{w})] \quad (\text{by Theorem 2.3}) \\ &= cr(\mathbf{w}). \end{aligned}$$

Likewise

$$\begin{aligned}
 \sigma^2(c\mathbf{w}) &= \text{Var}(R(c\mathbf{w})) \\
 &= \text{Var}\left(\sum_{i=1}^n cw_i(R_i - r)\right) \quad (\text{by Eq. (10.17)}) \\
 &= \text{Var}\left(c \sum_{i=1}^n w_i(R_i - r)\right) \\
 &= \text{Var}(cR(\mathbf{w})) \quad (\text{by Eq. (10.17)}) \\
 &= c^2 \text{Var}(R(\mathbf{w})) \quad (\text{by exercise 18 of Chap. 2}) \\
 &= c^2 \sigma^2(\mathbf{w})
 \end{aligned}$$

- (16) In this example, the variance in the rate of return is to be minimized such to the constraint that the expected value of the rate of return must be one. We will again use Lagrange multipliers to find the minimum variance. We must solve the following system of two simultaneous equations.

$$\begin{aligned}
 \nabla(\sigma^2(\mathbf{w})) &= \lambda \nabla(r(\mathbf{w})) \\
 r(\mathbf{w}) &= 1
 \end{aligned}$$

The set of equations above are equivalent to the following set of $n + 1$ equations.

$$\begin{aligned}
 2w_i \sigma_i^2 &= \lambda(r_i - r) \quad \text{for } i = 1, 2, \dots, n, \\
 \sum_{i=1}^n w_i(r_i - r) &= 1
 \end{aligned}$$

If we solve the i^{th} equation for w_i for $i = 1, 2, \dots, n$ and substitute into the last equation we have

$$\lambda \sum_{i=1}^n \frac{(r_i - r)^2}{2\sigma_i^2} = 1.$$

This implies that

$$\lambda = \frac{1}{\sum_{i=1}^n \frac{(r_i - r)^2}{2\sigma_i^2}}.$$

Therefore

$$w_i = \frac{r_i - r}{2\sigma_i^2 \sum_{i=1}^n \frac{(r_i - r)^2}{2\sigma_i^2}} = \frac{\frac{r_i - r}{\sigma_i^2}}{\sum_{i=1}^n \frac{(r_i - r)^2}{\sigma_i^2}}.$$

- (17) According to Lemma 10.3

$$\alpha_A = 4.44549$$

$$\alpha_B = 1.30112$$

$$\alpha_C = 7.9031$$

$$\alpha_D = 0.0$$

$$\alpha_E = 9.69925$$

- (18) We must show that the set of equations below is satisfied by the expressions in Eqs. (10.28) and (10.29). This can be accomplished by direct substitution into the equations.

$$\begin{aligned} r_i - r_M &= \lambda (2 [(\rho_{i,M}\sigma_i - \sigma_M)\sigma_M + x(\sigma_i^2 - 2\rho_{i,m}\sigma_i\sigma_M + \sigma_M^2)]) \\ \sigma^2 &= x^2\sigma_i^2 + (1-x)^2\sigma_M^2 + 2x(1-x)\rho_{i,M}\sigma_i\sigma_M \end{aligned}$$

- (19) Substituting $r_f = 0.0475$, $r_M = 0.0765$, $\sigma_M = 0.22$, and $\text{Cov}(i, M) = 0.15$ into Eq. (10.26) we obtain

$$\begin{aligned} r_i - r_f &= \frac{\text{Cov}(i, M)}{\sigma_M^2}(r_M - r_f) \\ r_i - 0.0475 &= \frac{0.15}{(0.22)^2}(0.0765 - 0.0475) \\ r_i &= 0.0475 + (3.09917)(0.029) \\ &= 0.137376 \end{aligned}$$

- (20) Using the values $S(0) = 83$, $\mu = 0.13$, $\sigma = 0.25$, $r = 0.055$, $T = 1/4$, $t = 0$, and $K = 86$ we can determine the value of the European call option as $C = 1.26152$.

Suppose a portfolio consists of a position x in the security and a position y in the option. According to Eq. (10.31) the expected return of the portfolio is a linear function of x and y given by

$$\mathbb{E}[R] = 2.23423x + 3.18342y.$$

The variance in the expected rate of return is a quadratic function of x and y described in detail by Eq. (10.32). In this example the

quadratic function takes the form

$$\text{Var}(R) = 114.405x^2 + 129.264xy + 46.6908y^2.$$

We would like to find the minimum variance portfolio having $E[R] = 10$. If we set the expected rate of return to \$10 and solve for y we obtain

$$y = 3.14128 - 0.701834x.$$

Substituting this expression into the variance of R produces the quadratic function of x given below.

$$\text{Var}(R) = 46.6819x^2 + 200.179x + 460.727$$

The minimum value occurs when $x = -2.14408$ and thus $y = 4.64607$.

Bibliography

- Bleecker, D. and Csordas, G. (1996). *Basic Partial Differential Equations*, International Press, Cambridge, Massachusetts, USA.
- Boyce, W. E. and DiPrima, R. C. (2001). *Elementary Differential Equations and Boundary Value Problems*, 7th edition, John Wiley and Sons, Inc., New York, USA.
- Burden, R. L. and Faires, J. D. (2005). *Numerical Analysis*, 8th edition, Thomson Brooks/Cole, Belmont, CA, USA.
- Churchill, R. V., Brown, J. W. and Verhey, R. F. (1976). *Complex Variables and Applications*, 3rd edition, McGraw-Hill Book Company, New York, USA.
- Courant, R. and Robbins, H. (1969). *What is Mathematics?*, Oxford University Press, Oxford, UK.
- Dean, J., Corvin, J. and Ewell, D. (2001). List of Playboy's playmates of the month with data sheet stats, Available at <http://www3.sympatico.ca/jimdean/pmstats.txt>.
- DeGroot, M. H. (1975). *Probability and Statistics*, Behavioral Science: Quantitative Methods. Addison-Wesley Publishing Company, Reading, Massachusetts, USA.
- Durrett, R. (1996). *Stochastic Calculus: A Practical Introduction*, CRC Press, Boca Raton, FL, USA.
- Gale, D., Kuhn, H. W. and Tucker, A. W. (1951). Linear programming and the theory of games, in T.C. Koopmans, editor, *Activity Analysis of Production and Allocation*, Wiley, New York, NY, USA, pp. 317–329.
- Gardner, M. (October 1959). “Mathematical Games” column, *Scientific American*, pp. 180–182.
- Goldstein, L. J., Lay, D. C. and Schneider, D. I. (1999). *Brief Calculus and Its Applications*, 8th edition, Prentice Hall, Upper Saddle River, NJ, USA.
- Golub, G. H. and Van Loan, C. F. (1989). *Matrix Computations*, 2nd edition, Johns Hopkins University Press, Baltimore, Maryland, USA.
- Greenberg, M. D. (1998). *Advanced Engineering Mathematics*, 2nd edition, Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA.
- Jeffrey, A. (2002). *Advanced Engineering Mathematics*, Harcourt Academic Press, Burlington, Massachusetts, USA.

- Kijima, M. (2003). *Stochastic Processes with Applications to Finance*, Chapman & Hall/CRC, Boca Raton, FL, USA.
- Luenberger, D. G. (1998). *Investment Science*, Oxford University Press, New York, USA.
- Marsden, J. E. and Hoffman, M. J. (1987). *Basic Complex Analysis*, W. H. Freeman and Company, New York, NY, USA.
- Mikosch, T. (1998). *Elementary Stochastic Calculus With Finance in View*, Vol. 6 of *Advanced Series on Statistical Science & Applied Probability*, World Scientific Publishing, River Edge, New Jersey, USA.
- Neftci, S. N. (2000). *An Introduction to the Mathematics of Financial Derivatives*, 2nd edition, Academic Press, San Diego, CA, USA.
- Redner, S. (2001). *A Guide to First Passage Processes*, Cambridge University Press, Cambridge, UK.
- Ross, S. M. (1999). *An Introduction to Mathematical Finance: Options and Other Topics*, Cambridge University Press, Cambridge, UK.
- Ross, S. M. (2006). *A First Course in Probability*, 7th edition, Prentice Hall, Inc., Upper Saddle River, NJ, USA.
- Selvin, S. (1975). A problem in probability, *American Statistician*, **29**, 1, p. 67.
- Smith, R. T. and Minton, R. B. (2002). *Calculus*, 2nd edition, McGraw-Hill, Boston, MA, USA.
- Steele, J. M. (2001). *Stochastic Calculus and Financial Applications*, Volume 45 in Applications of Mathematics, Springer-Verlag, New York, NY, USA.
- Stewart, J. (1999). *Calculus*, 4th edition, Brooks/Cole Publishing Company, Pacific Grove, CA, USA.
- Strang, G. (1986). *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, USA.
- Taylor, A. E. and Mann, W. R. (1983). *Advanced Calculus*, 3rd edition, John Wiley & Sons, Inc., New York, NY, USA.
- vos Savant, M. (February 1990). “Ask Marilyn” column, *Parade Magazine*, p. 12.
- Wilmott, P., Howison, S. and Dewynne, J. (1995). *The Mathematics of Financial Derivatives: A Student Introduction*, Cambridge University Press, Cambridge, UK.

Index

- absorbing boundary condition, 80, 113
- Addition rule, 18
- additivity, 40
- American option, 103
- annuity, 8
- arbitrage, 63, 110
- Arbitrage Theorem, 63, 72, 104
- area as probability, 36
- arithmetic mean, 192
- augmented matrix, 87
- average, *see* expected value
- backwards parabolic equation, 112
- backwards substitution, 87, 89
- bell curve, 42
- Bernoulli random variable, 23, 79
- beta, 184
- betting strategy, 72
- binary model, 107
- binomial model, 115
- binomial random variable, 23, 42
- Black-Scholes equation, 104, 110
- boundary condition, 89, 112
 - absorbing, 80
 - European call option, 113
- Brownian motion, 187
- call option, 103
- Capital Asset Pricing Model, 183
- capital market line, 182, 184
- CAPM, *see* Capital Asset Pricing Model
- centered difference formula, 45
- Central Limit Theorem, 49, 176
- certainty equivalent, 172
- chain rule, 98, 132
 - multivariable, 99
- closing price, 95
- compound amount, 2
- compound interest, 3
- compounding period, 3
- concavity, 135, 167
- conditional exit time, 90
- conditional probability, 19
- constrained optimization, 179
- constraints, 65
 - equality, 65
 - inequality, 65
- continuous distribution, 36
- continuous probability, 36
- continuous random variable, 35
- continuous random walk, 91
- continuous stochastic process, 93
- continuously compounded interest, 5
- convex set, 65
- correlation, 155, 159
- cost function, 65
- covariance, 155
- cumulative distribution function, 48
- Delta, 133, 145
- Delta neutrality, 145, 150

- dependent random variables, 40
- derivative, 131
 - financial, 110
- deterministic process, 92
- deviation, 41
- difference equation, 84
- differential equation
 - stochastic, 94, 96
- differentiation
 - chain rule, 98
- discrete outcome, 16
- distribution
 - continuous, 36
- distribution function, 36
- drift, 51, 163, 187
- dual, 67
- Duality Theorem, 65, 71
- effective interest rate, 3
- efficient markets, 63
- equality constraints, 65
- Euler Identity, 116
- European call option
 - boundary condition, 113
- European option, 103
 - call
 - final condition, 112
 - event, 15
 - certain, 16
 - discrete, 15
 - impossible, 16
 - events
 - compound, 17
 - independent, 21
 - mutually exclusive, 17
 - excess rate of return, 184
 - exercise time, 103
 - exit time, 90
 - expected rate of return, 180
 - expected utility, 171
 - expected value, 15, 24, 38
 - of normal random variable, 47
 - experiment, 15
 - expiry date, 103
 - exponential decay, 123
 - exponential growth, 5, 91
- Extreme Value Theorem, 178
- feasible vector, 65
- final condition, 111
 - European call option, 112
- financial derivative, 110
- finite differences, 45
- Fourier Convolution, 117
- Fourier Transform, 115
 - alternative definition, 119
 - definition, 115
 - existence, 116
 - of a convolution, 118
 - of a derivative, 117
- frequency, 116
- function
 - concave, 167
 - utility, 166
- Fundamental Theorem of Calculus, 132
- future value, 5
- Gamma, 135, 148, 151
- Gamma neutrality, 151
- Gaussian elimination, 87
- generalized Wiener process, 97
- geometric mean, 192
- geometric series, 7
- gradient operator, 179
- Greeks, 131
- harmonic mean, 192
- heat equation, 111, 120
- hedge and forget, 146
- hedging, 136, 143, 164
- hockey stick, 112
- in the money, 112, 153
- independent events, 21
- independent random variables, 40
- inequality constraints, 65
- inflation, 10
 - rate of, 10
- initial condition, 111
- interest, 1
 - compound, 3

- continuously compounded, 5
- effective rate, 3
- simple, 2
- interest rate, 1
- Inverse Fourier Transform, 119
- investor
 - risk-averse, 168
 - risk-neutral, 168
- Itô process, 97
- Itô's Lemma, 100, 110, 187
- Jensen's Inequality, 170, 171
 - Continuous Version, 171
 - Discrete Version, 170
- joint probability function, 26, 39
- Lagrange Multipliers, 179, 185
- Laplacian equation, 84
- linear programming, 65
- linear systems
 - tridiagonal, 87
- linearity, 140
- loans, 7
- lognormal random variable, 51, 163, 188
- long position, 104, 187
- marginal distribution, 40
- marginal probability, 27
- matrix
 - upper triangular, 87
- mean, *see* expected value, *see also* expected value
- Mean Value Theorem, 169
- minimum variance, 163, 165, 177
- Monty Hall problem, 18
- Multiplication rule, 20
- noise, 92
- normal probability distribution, 42, 48
- normal random variable, 35, 42
- odds, 64
- operations research, 65
- optimal portfolio, 164
- option, 103
 - American, 103
 - call, 103
 - European, 103
 - put, 103
- option pricing formula
 - European call, 125
 - European put, 125
- out of the money, 112, 153
- outcome, 15
 - discrete, 15, 16
- partial differential equation, 111
 - Black-Scholes, 104, 139
 - linear, 111
 - parabolic type, 111
 - second order, 111
- PDE, *see* partial differential equation
- Poisson equation, 86
- polar coordinates, 46
- portfolio, 140, 162
 - Delta-neutral, 145
 - Gamma neutral, 151
 - optimal, 164
 - rebalancing, 146
- portfolio selection problem, 173, 175
- Portfolio Separation Theorem, 180
- position
 - long, 104
 - short, 104
- positive part, 55
- present value, 1, 5, 110
- pricing formula
 - European call option, 125
 - European put option, 125
- primal, 67
- principal, *see* principal amount
- principal amount, 2
- probability, 16
 - classical, 16
 - conditional, 19
 - continuous, 36
 - empirical, 16
 - marginal, 27
- probability distribution, 21
 - normal, 48

- probability distribution function, 36
- put option, 103
- Put-Call Parity, 106
- Put-Call Parity Formula, 134
- random experiment, 166
- random variable, 15
 - Bernoulli, 79
 - binomial, 15, 42
 - continuous, 35
 - discrete, 15
 - lognormal, 51, 163
 - normal, 35, 42
 - uncorrelated, 159, 178
 - uniform, 37
- random walk, 77
 - continuous, 91
- rate of inflation, 10
- rate of return, 1, 11, 155, 162
 - excess, 184
 - expected value, 180
 - variance, 180
- rebalancing, 146
- retirement, 8
- return
 - rate of, *see* rate of return, 11
- Rho, 138
- risk, 185
- risk-averse investor, 168
- risk-neutral investor, 168
- samples means, 49
- Schwarz Inequality, 160
- secant lines, 168
- sensitivity, 131
- separation of variables, 111, 123
- short position, 104, 187
- simple interest, 2
- slack variables, 65
- spatial homogeneity, 82
- spread, *see* standard deviation
- standard deviation, 15, 32, 41
- standard normal random variable, 48
- Stirling's formula, 44
- stochastic calculus, 77
- stochastic differential equation, 94, 96
- stochastic process, 92, 96
 - continuous, 93
- stochastic processes, 110
- stopping time, 83
- strike price, 103
- strike time, 103
- surrogate, 164
- symmetric random walk, 78
- Taylor remainder, 99
- Taylor series, 150
- Taylor's Theorem, 98
- Theta, 131
- Treasury Bonds, 105
- tridiagonal linear systems, 87
- tridiagonal matrix, 87
- uncorrelated random variable, 159
- uniform random variable, 37
- upper triangular matrix, 87
- utility, 166
 - decreasing, 167
- utility function, 166, 167
- variance, 29, 41
 - alternative formula, 41
 - minimum, 163
 - of normal random variable, 47
- vector inequalities, 66
- Vega, 136
- volatility, 51, 92, 96, 137, 163, 187
- Weak Duality Theorem, 68
- Wiener process, 97, 187

An Undergraduate Introduction to Financial Mathematics

This textbook provides an introduction to financial mathematics and financial engineering for undergraduate students who have completed a three or four semester sequence of calculus courses. It introduces the theory of interest, random variables and probability, stochastic processes, arbitrage, option pricing, hedging, and portfolio optimization. The student progresses from knowing only elementary calculus to understanding the derivation and solution of the Black–Scholes partial differential equation and its solutions. This is one of the few books on the subject of financial mathematics which is accessible to undergraduates having only a thorough grounding in elementary calculus. It explains the subject matter without “hand waving” arguments and includes numerous examples. Every chapter concludes with a set of exercises which test the chapter’s concepts and fill in details of derivations.

