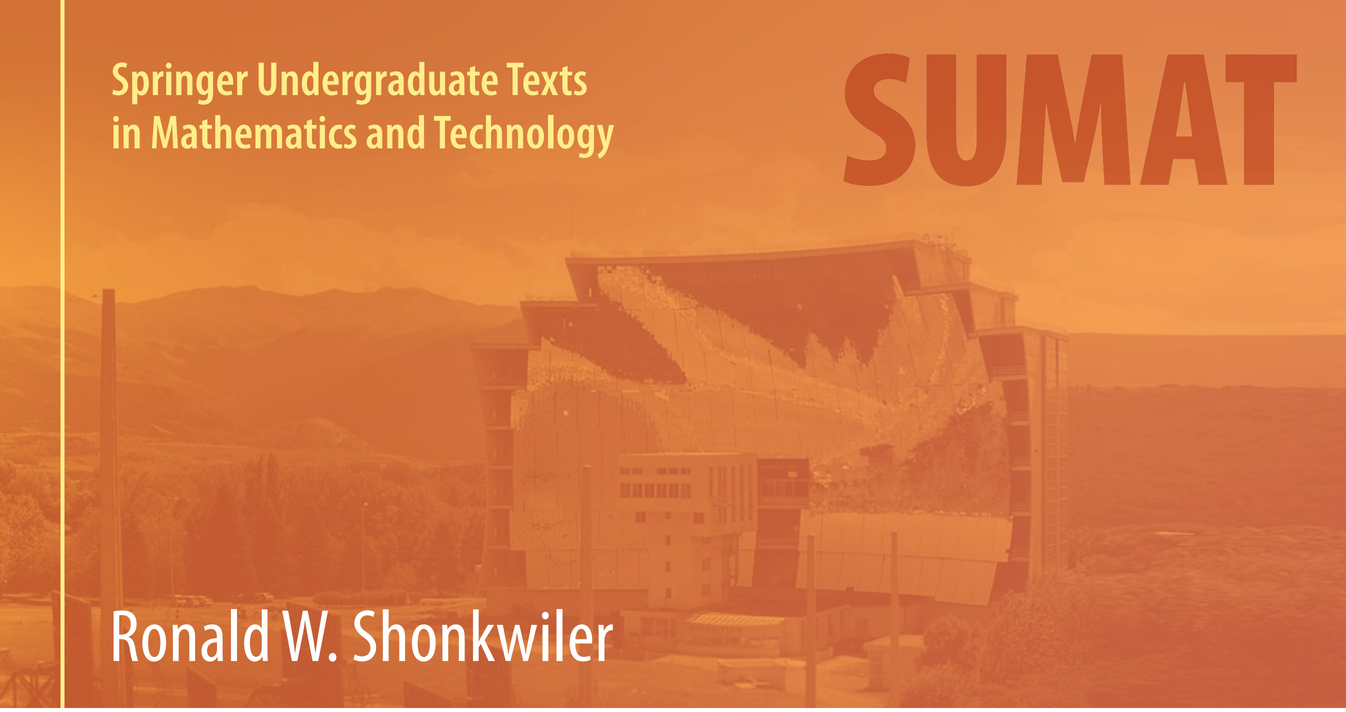
Springer Undergraduate Texts in Mathematics and Technology

SUMAT



Ronald W. Shonkwiler

Finance with Monte Carlo

Springer

Springer Undergraduate Texts in Mathematics and Technology

For further volumes:
http://www.springer.com/series/7438

Ronald W. Shonkwiler

Finance with Monte Carlo

Springer

Ronald W. Shonkwiler
School of Mathematics
Georgia Institute of Technology
Atlanta, GA, USA

ISSN 1867-5506 ISSN 1867-5514 (electronic)
ISBN 978-1-4614-8510-0 ISBN 978-1-4614-8511-7 (eBook)
DOI 10.1007/978-1-4614-8511-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013945890

Mathematics Subject Classification (2010): 65C05, 91Bxx

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The aim of this text is to introduce the reader to the core topics that constitute an introductory course in finance and financial engineering. Our particular emphasis is on illustrating principles and modeling through the Monte Carlo method. Monte Carlo is the uniquely appropriate tool for this purpose because the driving factors underlying the market are primarily random in nature. And just as the random dynamics works its way through the system and into financial observables, we may track the chain of influence every step of the way computationally.

The intended audience for the book is upper division undergraduates or beginning graduate students in mathematics, finance or economics. The reader is assumed to have knowledge of calculus through partial derivatives, Taylor series and LaGrange optimization, probability through an understanding of random variables, expectation, distribution and density functions including the normal distrubution, and basic matrix algebra through the solution of linear systems. A refresher for these topics is presented in the Appendices.

For additional background on probability with Monte Carlo methods, it may be useful to read through *Explorations in Monte Carlo Methods*, a textbook that I co-authored with Franklin Mendivil, published in Springer's *Undergraduate Texts in Mathematics* series ©2009.

In keeping with our presentation of the material in parallel with the Monte Carlo method, a majority of the exercises are primarily programming in nature. Hopefully this is where the real understanding takes place. A great enjoyment can derive from experimenting with parameters and seeing the results unfold, sometimes surprisingly, always in an interesting way.

Regarding programming, I prefer allowing students to use whatever language with which they are familiar; many use Matlab™, Maple®, R, C, and Java. However, some of the results can only be appreciated if presented graphically, for example as histograms or x-y plots. In the case of the first three, graphics is built-in; otherwise there is ample public domain software for rendering numerical output.

The programming background needed is quite modest, basically, branching, loops, and subroutines. In many cases, other than the boiler plate, programs span fewer than a dozen lines. Furthermore, programming code, given in a mathematical format, is presented in line with the text as encountered. I treat these in exactly the same way as displayed equations. Like equations, they are condensed and must be read with care.

Organization of the Book

Two fundamental systems for analyzing market prices are presented in Chapter 1, the geometric Brownian motion (GBM) model and the binomial lattice model. GBM is preceded by a first principles introduction to Wiener processes. A more in-depth treatment is given in the Appendices if needed. Thus a rationale is provided for the Monte Carlo method and price simulation by the geometric random walk. In a kind of turn-about, we use the numerical method of the geometric random walk to derive the theoretical distribution for maturity stock prices, the lognormal distribution.

Starting out in this way allows for the implementation of the Monte Carlo method immediately. We take advantage of that by investigating one of the tenets of modern finance, the efficient market hypothesis (EMH). More generally we show that Monte Carlo can be used to test the antithesis of EMH, namely technical analysis (TA).

But in order to do that, we must have access to a database of historical prices. One of the best free sources of historical price information can be found at http://finance.yahoo.com. However this text must be independent of that lest its access change in the future. Therefore a database of prices, the FIMCOM prices, is supplied at the following URL:

http://people.math.gatech.edu/~shenk.

The FIMCOM database is in exactly the same format as that of finance.yahoo. Either can be used to test programs and to answer TA queries. Additionally, I have placed utility programs for working with the FIMCOM database or the finance.yahoo prices, on the aforementioned website. Further, any eventual errata to this text will also be found there.

Chapter 2 is devoted to basic investment science and to the important mean-variance theory of portfolio management. The central tenet of this theory is that diversification ameliorates risk. But, equally important, it is a completely quantitative theory providing for an exact measurement of risk. Strikingly specific investment policy, widely implemented in practice, derives from the theory. With regard to risk, the GBM model for stock prices lends itself to a natural explanation of the value at risk (VAR) and its Monte Carlo calculation.

Chapter 3 introduces forward contracts and options as tools for the alleviation of risk. This leads to the important topic of option pricing and its solution by the fundamental principle of no-arbitrage and the risk-neutral probability. In the interest of pedagogy, our approach builds on the binomial lattice model. Using it we derive techniques for pricing both European and American options. Separately, we obtain the Black–Scholes formulas for European options by straightforward integration of the maturity distribution. As a Monte Carlo technique for American puts, we introduce the notion of the exercise boundary.

Chapter 4 exhibits the power of the Monte Carlo method for it is here that we introduce exotic options and use the method to price them. Often these options require knowledge of the path prices take to maturity and some even require future knowledge in order to price. For some of these options Monte Carlo is the only applicable method. Moreover, the techniques illustrated here are not restricted to the GBM model for prices. They work just as well, for example, with Lévy models, the main topic of Chapter 6.

Chapter 5 deals with financial engineering and some practical aspects of options, namely option trading. Many of the most popular option strategies are investigated, among them are: covered calls, spreads, butterflies, straddles and condors. Here we introduce a novel use of Monte Carlo as a tool for the prediction of expected outcomes of these strategies under differing market conditions. Here also the option "greeks" are defined and studied. Their practical use for insulating a portfolio against market fluctuations in price and volatility is demonstrated.

In Chapter 6 we delve into more advanced processes for market prices, exponential Lévy processes. These are recent developments that include models allowing for discontinuities in prices and models that support "heavytailed" phenomena. The latter refers to market collapses that occur too frequently as predicted by the Gaussian model upon which Black–Scholes is based. Although a rigorous treatment of this material is beyond the scope of this text, we nevertheless strive to convey the essentials of the relevant mathematics without getting too involved in technicalities. Instead we focus on the application of these processes to financial modeling and the implications for option pricing.

Chapter 7 addresses the problem of optimally allocating resources among risky ventures. The solution, as formulated in a 1950s paper out of Bell Laboratory, and its application to finance has been controversial. Still the method is sound, interesting and valid in its claims.

Atlanta, GA, USA Ronald W. Shonkwiler

Acknowledgements

The author would like to thank several of his colleagues for discussions on some of the material that appears in this text. Most helpful were John Elton, Steve Demko, Joe Murad, Gunter Meyer, and Carl Spruill. The author would also like to thank all those who reviewed the initial manuscript; they made many valuable suggestions to me. Finally, I thank Elizabeth Loew at Springer, who gave encouragement, advice, and kept the project on track.

Contents

List of Notation

Symbol	*Page*	*Meaning*
Δx	4	the change or increment in the variable x
$C(n,r)$	5	combinatorial function, n choose r
$\binom{n}{r}$	5	combinatorial function, n choose r
$\mathbb{E}(X)$	5	mathematical expectation of X
$\mathrm{var}(X)$	6	variance of the random variable X
μ	6	mean of a probabillity distribution; for a stock, its drift
σ^2	6	variance of a probabillity distribution
σ	6	standard deviation of a probability distribution, for a stock, its volatility
$\phi(x)$	6	normal probability density function
$N(\mu,\sigma^2)$	6	normal probability distribution with mean μ and variance σ^2
Z	8	a $N(0,1)$ sample
$\sim$	8	is a sample from
ARW	10	arithmetic random walk
$\triangleright$	10	comment in program code
GBM	11	geometric Brownian motion (continuous motion)
GRW	12	geometric random walk (discrete motion)
S_0	12	initial stock price
S_T	12	final stock price
$N_k(i)$	18	the ith node at the kth step of binomal tree
EMH	21	efficient market hypothesis
TA	22	technical analysis
ma_t	23	simple moving average
ema_t	24	exponential moving average
FIMCOM	26	financial database of prices for this text
DMI	27	direction movement indicator
$\mathbb{1}_A(x)$	28	indicator function of A, 1 if $x \in A$ 0 otherwise

Symbol	*Page*	*Meaning*
NYSE	29	New York stock exchange
NASDAQ	29	National Association of Securities Dealers Automated Quotations stock exchange
$U(0,1)$	30	uniform probability distribution on $[0,1)$
$\mathrm{covar}(X,Y)$	50	co-varance between X and Y
ρ_{XY}	50	correlation coefficient betwee X and Y
VaR	55	value at risk
OTM	78	out-of-the-money
ATM	78	at-the-money
ITM	78	in-the-money
$(S-X)^+$	78	same as $\max(S-X,0)$
$\Phi(x)$	101	standared normal cumulative distribution function
d_1	101	Black-Scholes pricing parameter
d_2	101	Black-Scholes pricing parameter
$G(S_T)$	101	option payoff for ending price S_T
IV	104	implied volatility
Δ	138	the delta of an option portfolio
Γ	140	the gamma of an option portfolio
Θ	142	the theta of an option portfolio
ν	144	the vega of an option portfolio
ρ	146	the rho of an option portfolio
$a:b::c$	160	allocation of strikes of a butterfly trade
$\mathbb{E}_Q(X)$	167	expectation with respect to the probability Q
$(L_t)_{t\geq 0}$	169	a Lévy process
$\mathrm{Po}(\lambda)$	170	Poisson process
N_t	170	Poisson random variable
$E(\lambda)$	170	exponential distribution with parameter λ
IG	172	inverse Gaussian distribution
$\nu(dx)$	173	Lévy measure
e^{L_t}	177	exponential-Lévy process
t_ν	184	Student t-distribution with ν degrees of freedom
dof	184	degrees of freedom of a t distribution
$\Gamma(z)$	184	the gamma function
$(a:b)$	192	odds paying a upon a win for a bet of b
$\tanh(x)$	208	hyperbolic tangent function
cdf	214	cumulative distribution function
pdf	215	probability density function

List of Algorithms

1

Geometric Brownian Motion and the Efficient Market Hypothesis

1.1 Stock Prices as a Random Walk

This book is about the nature of stock prices and its attendant consequences in terms of risk and money management. What we know today is the culmination of many years of observation and profound insight by many influential thinkers. It is rightfully so that there should be such devotion to the subject for understanding the mysteries of the ups and downs of stock prices has far reaching consequences. We will encounter many such examples in this text.

It is also a book about how to use the nature of stock prices to deliver verifiable answers to financial questions. As we will see, stock prices appear to follow a geometric random walk (GRW) and it follows that Monte Carlo methods are the appropriate scientific tool for the job. So in tandem with the goal of studying the workings of financial systems, we also aim to demonstrate the use of the computer to derive and develop the consequences of the financial models.

We start with an examination of our chief object of study, a typical chart of stock prices. In Fig. 1.1 we show prices for the Southern Copper Corporation (compensated for dividend disbursements).[1] It is for the period from Jan. 1st, 2007 to Feb. 26th, 2010.

Probably the most striking feature of the chart is that stock prices are unpredictable. On a short term basis, for example day-to-day, they seem to go up and down in a random, jittery fashion. Besides this fine scale structure, there appears to be large scale structure as well. The prices experience big up and down moves over periods of time on the order of months or even years. This chart is very typical of stock price charts generally.

Although the random walk model for prices can generate charts similar to Fig. 1.1 all by itself, without the need for multiple levels of randomness, nevertheless the market for a particular stock is affected both by its own contemporary news and by the economic environment which tends to operate on a longer time scale.

[1] The amount of a dividend is subtracted from the close price. See the discussion on page 25.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_1,

Fig. 1.1. Adjusted closing prices for SCCO from 2007 to 2010. Its features are typical for stock prices

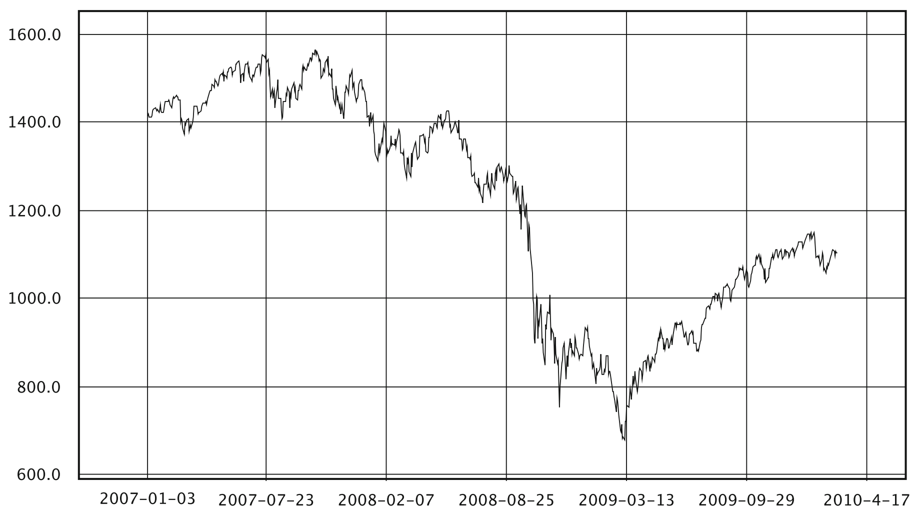


Fig. 1.2. Prices for the S&P-500 from 2007 to 2010. It shows how the market in general behaved over that period

A chart of the S&P-500 over the same time period is shown in Fig. 1.2. The S&P-500 data consists of a weighted average of a basket of 500 stocks and so represents the market as a whole more accurately than any single stock. (For an up-to-date list of the S&P stocks, visit http://www.indexarb.com/indexComponentWtsSP500.html.) It shows similar large scale trends as in the SCCO figure above indicating, at least, that individual stocks are correlated to a degree with the market as a whole. We discuss this point further in Chapter 2.

In this chapter we want to understand the fine scale structure. As mentioned, the day-to-day price seems to jitter about with no particular direction. It is prudent to ask what good can come of such a study if the price of a stock in the future, say 1 month ahead, is random anyway? The answer is that the future stock price, while random, nevertheless obeys constraints; it follows a probability distribution, there is more chance of ending at some values and less at other values. For better or worse, financial decisions must be worked out on the basis of the future price probability distribution if it can be determined.

1.2 Brownian Motion

In 1900 the jittery motion of stock prices reminded mathematics student Louis Bachelier of a phenomenon reported by a botanist three quarters of a century earlier. In 1827 Robert Brown described observing the jittery motion of pollen grains in water as viewed in a microscope. The pollen grains seemingly moved by themselves, but how?

Brownian motion, as it is called, is now known to be the result of random impacts on the pollen grains by water molecules. The water molecules are themselves invisible. This explanation was worked out rigorously by Albert Einstein in 1905 who showed that, statistically, Brownian motion particles must satisfy the partial differential equation

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2} \tag{1.1}$$

where p is the distribution of the particles over space and time and D is a physical constant. Equation (1.1) is called the *Diffusion equation*.[2]

Bachelier went on to pioneer several fundamental advancements in finance in his Ph.D. thesis, *The Theory of Speculation*, based on the analogy between Brownian motion and stock prices.

In Fig. 1.3 we show the graph of an approximation to a one dimensional Brownian motion. (This is an approximation because the "events" in our figure occur at regular time intervals while those of true Brownian motion are not regular.) The exact sequence of movements a particle experiences over a period of time is called a *world state* or *scenario*. A photographic record can show the scenario that actually took place. But looking to the future, literally infinitely many scenarios are possible. A simulation such as that of the figure depicts one of the possibilities.

At this point Brownian motion looks like a good candidate for the movement of market prices. But before making that commitment, we continue to analyze the motion of Brownian particles

A crude approximation to one-dimensional Brownian motion along the real line (x-axis) may be made by means of a simple coin toss experiment. Remarkably, by varying the conditions of the approximation carefully, accurate Brownian motion emerges in the limit.

Assume time is divided into discrete periods Δt and in each such period the Brownian particle moves a step right or left by one unit Δx, the choice being

[2] On the basis of his derivation, Einstein was able to predict the size of water molecules. At the time the existence of atoms and molecules was still in doubt.

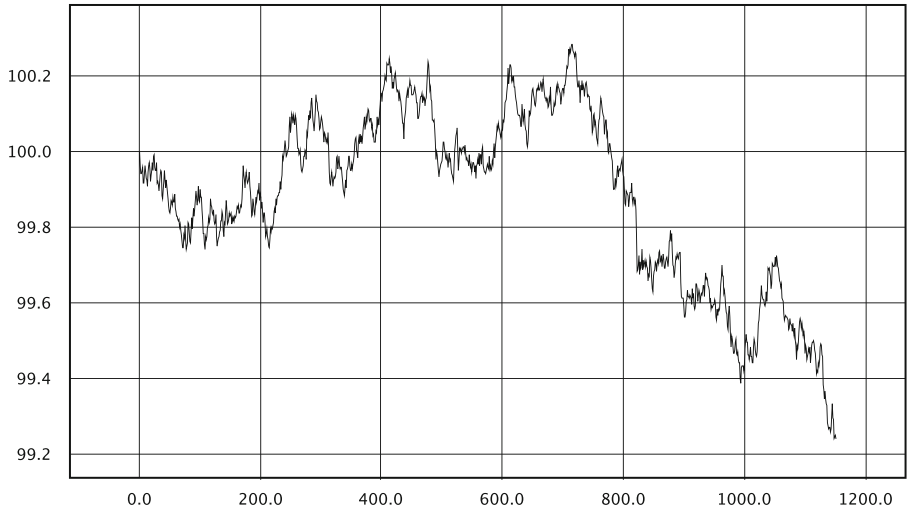


Fig. 1.3. An instance of a simulated arithmetic random walk

random. We refer to this as a *random walk*. Assume the particle's initial position along the line is $x_0 = 0$. After n time periods the walk has taken n steps. The particle's new position along the x-axis at that time will be between $-n(\Delta x)$ and $n(\Delta x)$.

For example, suppose $n = 4$. If all 4 choices are to the left, the particle will be at $-4(\Delta x)$; if 3 are left and 1 right, it will be at $-2(\Delta x)$. The other possible outcomes are, omitting the Δx, 0, 2, and 4. Notice the outcomes are separated by 2 steps. Also notice there are several ways most of the outcomes can arise, the outcome 2 for instance. We can see this as follows. Let R denote a step right and L a step left. Then a path of 4 steps can be coded as a string of 4 choices of the symbols R or L. For example, $LRRR$ means the first step is to the left and the next three are to the right. In order that the outcome of 4 steps be a net 2 to the right, 3 steps must be taken to the right and one to the left but the order doesn't matter. There are four possibilities that do it, they are $LRRR$, $RLRR$, $RRLR$, or $RRRL$.

In general, let $p(X, t)$ denote the probability that the particle is at position $X = m(\Delta x)$, m steps to the right of the origin, after n time periods, $t = n(\Delta t)$.[3] We wish to calculate $p(X, t)$. It will help to recognize that our random walk with n steps is like tossing n coins. For every coin that lands heads we step right and for tails we step left. Let r be the number of steps taken to the right and l the number left, then to be at position $m(\Delta x)$ it must be that their difference is m,

$$m = r - l \qquad \text{where} \qquad n = r + l.$$

So m and n are given in terms of r and l. But these two equations can be inverted to find r and l in terms of m and n. To find r, add the two and for l, subtract.

[3] It is customary to use upper case letters to denote random variables and lower case letters to denote an instance of the random variable.

$$r = \frac{1}{2}(n+m) \qquad \text{and} \qquad l = \frac{1}{2}(n-m). \tag{1.2}$$

Now the number of ways of selecting r moves to the right out of n possibilities is the problem of counting combinations and is given by

$$C(n,r) = \frac{n!}{r!\,(n-r)!}. \tag{1.3}$$

$C(n,r)$ is referred to as "n choose r" and is often denoted as well by $\binom{n}{r}$. For example, 3 moves right out of 4 possible moves can happen in $4!/(3!\,1!) = 4$ ways in agreement with the explicitly written L R possibilities noted above. Therefore, if the probabilities of going left or right are equally likely, then

$$p(X,t) = (\text{probability of ending at } X = m\Delta x) = \frac{C(n,r)}{2^n}, \tag{1.4}$$

where

$$r = \frac{1}{2}(n+m).$$

The solid curve in Fig. 1.4 is a graph of $p(x,t)$ for $n = 24$ and $\Delta x = 1$. If the random walk experiment with $n = 24$ steps were conducted 2^{24} times, then a frequency chart of the end points of the walk will closely approximate this curve. The result of such an experiment is in fact shown as an overlay in the figure.

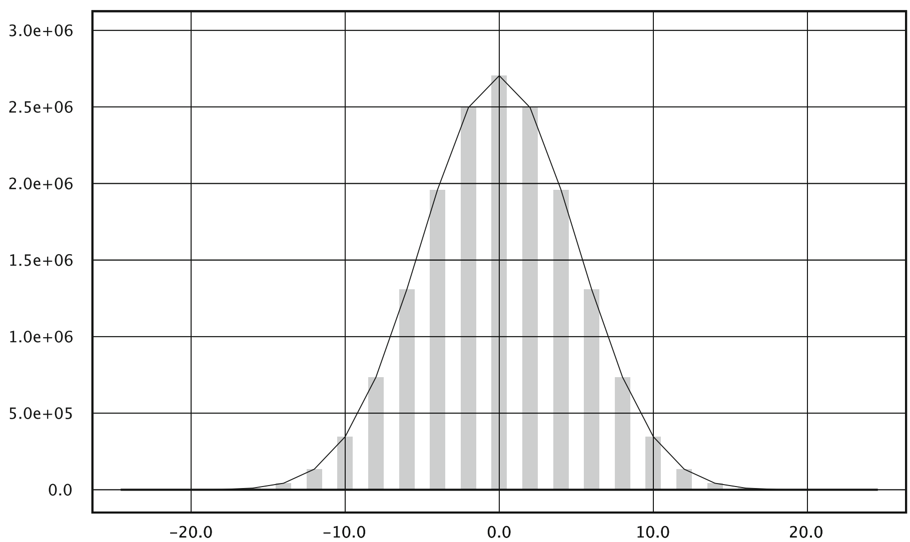


Fig. 1.4. A graph of $p(m, 24)$, showing the end point distribution of a 24 step coin toss random walk

Next we calculate the *mean* and *variance* of the ending position. The mean or average position, $\mathbb{E}(X)$, after a random walk of n steps is 0 when the probabilities of stepping left or right are equal. In this notation, $\mathbb{E}$ is a mnemonic for "expected" or "expectation." To show that the expectation is 0, start with a walk of just 1 step; the expectation of X is

$$(-\Delta x)\frac{1}{2} + (+\Delta x)\frac{1}{2} = 0,$$

because with probability 1/2 the step is to $-\Delta x$ and with probability 1/2 the step is to $+\Delta x$. For a walk of n independent steps the expectation is just n times this or 0.

But knowing that the average end point is 0 is not the whole story. It does not tell us how far from 0 the ending position is likely to be. Walks that go to the left cancel their distance with those that go to the right.

We can avoid the left versus right cancellation by using the squares of the positions. The mean square position, $\mathbb{E}(X^2)$, for a single step is

$$(-\Delta x)^2\frac{1}{2} + (+\Delta x)^2\frac{1}{2} = \Delta x^2.$$

Again because the steps are independent, the mean square position for n such steps is the sum, $n\Delta x^2$. Therefore the square root of this is a measure of how far the price averagely moves from the start point, it is referred to as the *root mean square position*. For a walk of n steps

$$\text{root mean square position} = \sqrt{\mathbb{E}(X^2)} = \sqrt{n}\Delta x. \tag{1.5}$$

In summary, a random walk moves away from its start point in proportion to the square root of time, at least statistically it does.[4]

In fact, $n(\Delta x)^2$ is the variance of the position and $\sqrt{n}\Delta x$ the standard deviation since $\mathbb{E}(X) = 0$.[5]

The exact equation for $p(X, t)$, equation (1.4), has a simple approximation. There is a genuine need for such an approximation because it is difficult to compute the combinatorial factor $C(n, r)$ for large values of n. Moreover, the approximation improves with an error that tends to 0 as $n \to \infty$. One may notice in Fig. 1.4 that the graph of $p(X, t)$ looks very much like that of a *normal distribution*. The probability density of a normal distribution with *mean* μ and *variance* σ^2 is given by

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{1.6}$$

see Section A.3. It is customary to refer to this distribution notationally as $N(\mu, \sigma^2)$. The square root of variance, σ, is the *standard deviation*. By the *Central Limit Theorem* (CLT) of probability, the distribution of the random walk particles will in fact tend in the limit to a normal distribution, see Section A.6.

To obtain the approximation, we match up the means and variances of the two distributions. Recall $t = n\Delta t$, $x = m\Delta x$, $\mu = 0$, and $\sigma^2 = n\Delta x^2$. Therefore

[4] Note that $\sqrt{\mathbb{E}(X^2)}$ is not necessarily equal to $\mathbb{E}(|X|)$ and usually they are not equal. More generally, for a given function f, $f(\mathbb{E}(X))$ is not necessarily equal to $\mathbb{E}(f(X))$. A simple demonstration is provided by the function $f(x) = x^2$ in the present case since $(\mathbb{E}(X))^2 = 0$ but $\mathbb{E}(X^2) = n(\Delta x)^2$.

[5] By definition the variance is $\mathrm{var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$; see Section A.4.

$$p(x,t) \approx \frac{1}{\sqrt{2\pi n \Delta x^2}} e^{-\frac{x^2}{2n\Delta x^2}}$$
$$\approx \frac{1}{\sqrt{2\pi t \frac{\Delta x^2}{\Delta t}}} e^{-\frac{x^2}{2t\frac{\Delta x^2}{\Delta t}}}. \tag{1.7}$$

By putting $D = \Delta x^2/(2\Delta t)$ we get

$$p(x,t) \approx \frac{1}{\sqrt{4\pi t D}} e^{-\frac{x^2}{4tD}}. \tag{1.8}$$

Now let n and m tend to ∞ while at the same time letting Δx and Δt tend to 0, so that t and x remain fixed and so that the ratio $D = \Delta x^2/(2\Delta t)$ remains constant. (This is the same D of the Diffusion equation (1.1).) The result is (1.8). We leave it to the reader to show that $p(x,t)$ given by this equation indeed satisfies (1.1).

By comparing (1.4) and (1.8) it is possible to work out an approximation for $C(n,r)$,

$$C(n,r) \approx 2^n \sqrt{\frac{2}{\pi n}} e^{-(2r-n)^2/(2n)}.$$

However *Stirling's formula*[6] leads to a better approximation, especially for r near 0 or n,

$$C(n,r) \approx \sqrt{\frac{n}{2\pi r(n-r)}} \frac{n^n}{r^r(n-r)^{n-r}} \tag{1.9}$$

The meaning of 1.8 is this: the probability that the Brownian particle will lie between two values, say $X = a$ and $X = b$ after time t has passed is equal to

$$\Pr(a < X < b) = \int_a^b \frac{1}{\sqrt{4\pi t D}} e^{-\frac{x^2}{4tD}}\, dx.$$

1.3 Wiener Processes

The mathematical theory of Brownian Motion was developed by Norbert Wiener and is often referred to as a *Wiener process*. Let W_t, $t \geq 0$, denote the position of a Brownian particle at time t with $W_0 = 0$. The axioms of a Wiener process are:

1. Every increment $W_{t+h} - W_t$ is normally distributed with mean 0 and variance $\sigma^2 h$ where σ is a fixed parameter.
2. For every pair of disjoint time intervals $[t_1, t_2]$ and $[t_3, t_4]$, the increments $W_{t_4} - W_{t_3}$ and $W_{t_2} - W_{t_1}$ are independent random variables with distributions as in Axiom (1).
3. W_t is continuous at $t = 0$.

[6] Stirling's formula is $n! \approx \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$.

By *standard Brownian motion* we mean a Wiener process with parameter $\sigma = 1$.

An immediate consequence of Axiom (2) is that a Wiener process satisfies the *Markov property*. It means that for any time t, the future realization of the process, W_{t+h}, only depends on W_t, the present state, and not on the path the process took to W_t. This is because $W_{t+h} - W_t$ is independent of $W_t - W_0$.

Likewise it follows from Axiom (1) that, since the expectation of the increment $\Delta W_t = W_{t+h} - W_t$ is zero, the future expectation of the process equals the present value W_t. Given the process is at W_t at time t, then

$$\mathbb{E}(W_{t+h} \,|\, W_t) = W_t + \mathbb{E}(\Delta W_t) = W_t. \tag{1.10}$$

This is the *martingale* property.

It is perhaps remarkable that the Axioms of a Wiener process hold for finite increments h as well as infinitesimal ones dt. With respect to an infinitesimal increment, we write

$$dW_t = W_{t+dt} - W_t. \tag{1.11}$$

An entire differential and integral calculus can be built upon such increments including the ability to integrate functions against the Wiener process. This matter is discussed further in appendix Section B.

According to Axiom (1) we may write

$$W_t = \sigma\sqrt{t}Z \tag{1.12}$$

where $Z \sim N(0,1)$.[7] Among other uses, this equation shows that as the increments of time between jumps tend to zero, so also does the size of the jumps and they do so according to the square root of the time increments. For this reason, a Brownian motion path is *continuous*. Within every visible jump there are a large number of jumps of much smaller size including infinitesimal ones. In fact, although it is beyond our scope to prove it here, while continuous, Brownian paths are nowhere differentiable. Moreover, the total length of a Brownian path W_t, $0 \le t \le T$, adding both the increments to the left and to the right, is infinite for all T. Such a path is said to have *infinite variation*.

1.3.1 Simulating Brownian Motion End Points

Simulating the end point W_T of a Wiener process is just a matter of generating standard normal samples. Computer installations usually provide a means for this.[8] Such samples are readily converted to a normal sample of any desired variance by multiplying by the standard deviation. That is, if $Z \sim N(0,1)$ is a sample from the standard normal, then $\sigma\sqrt{T}Z$ is a sample from $N(0, \sigma^2 T)$. If in addition, one adds the constant μT, then $X = \mu T + \sigma\sqrt{T}Z$ is a sample from $N(\mu T, \sigma^2 T)$.

[7] The notation $Z \sim N(0,1)$ means that the random variable Z is a sample from the density indicated, here, the normal density with mean 0 and variance 1.

[8] If only uniform samples $U \sim U(0,1)$ are available, normal samples can be generated from them, see Section A.9.

1.3.2 Simulating Brownian Motion Paths

We will frequently need to simulate a Brownian motion path leading to the end point W_T. Of course only a crude approximation is possible since, as a Wiener process, such a path has infinite variation as noted above. Instead we simulate the value of the process at a sequence of discrete times, for example at Δt, $2\Delta t$, ..., $n\Delta t = T$. We have

$$\begin{aligned} W_0 = 0, \qquad W_{i\Delta t} &= W_{(i-1)\Delta t} + \Delta W_i \\ &= W_{(i-1)\Delta t} + \sigma\sqrt{\Delta t} Z_i, \quad i = 1, \ldots, n \end{aligned}$$

where $Z_i \sim N(0, 1)$ for each i.

1.3.3 Wiener Processes with Drift

The Wiener process described above has mean displacement 0; this is because for every path allowed by the process, its negative is also an admissible path with the same chance of occurring. Or more directly, mean displacement 0 is explicitly stated in Axiom (1). But a directional bias can be introduced to the process. Letting W_t denote a Wiener process with parameter σ, define X_t by

$$X_t = \mu t + W_t. \tag{1.13}$$

The constant parameter μ is called the *drift*. Its effect is to shift the mean position of the Brownian particle from 0 to μt at time t,

$$\mathbb{E}(X_t) = \mathbb{E}(\mu t + W_t) = \mu t + \mathbb{E}(W_t) = \mu t.$$

The variance of X_t however is the same as that of W_t,

$$\operatorname{var}(X_t) = \operatorname{var}(\mu t + W_t) = \operatorname{var}(\mu t) + \operatorname{var}(W_t) = 0 + \sigma^2 t.$$

In this way the probability density of the walk at time t becomes $N(\mu t, \sigma^2 t)$.

1.4 Arithmetical Random Walk

Combining the foregoing, we define an *arithmetical random walk* (ARW) as the simulation $\{X_0, X_{\Delta t}, X_{2\Delta t}, \ldots, X_{n\Delta t}\}$ of Brownian motion with drift. We allow the walk to start at an arbitrary point, X_0, not necessarily at zero. The term arithmetical is used because the steps sizes are all the same in the sense of having the same mean and the same standard deviation. In the algorithm these are $\mu\Delta t$, and $\sigma\sqrt{\Delta t}$ respectively.

To implement an ARW, begin by dividing the interval 0 to T into a succession of subintervals of some desired length, Δt. Let n denote the number of subintervals required, $T = n\Delta t$. The inputs of the algorithm are the time periods T and Δt, the starting point X_0, and the walk parameters μ and σ.

Algorithm 1. Arithmetic random walk algorithm

```
inputs: X_0, T, Δt, μ, σ
  ▷signifies a comment
n = T/Δt    ▷number of Δt steps in time T
for t = 1,...,n
  Z_t ~ N(0,1)    ▷Z_t is a N(0,1) sample
  ΔX_t = μΔt + σ√(Δt) Z_t
  X_t = X_{t-1} + ΔX_t
endfor
  ▷the last X_t is X_T
```

The output is one possible path that could be taken by the particle over the given period. Starting from X_0 at $t = 0$, the simulation predicts X_1 at $t = \Delta t$, X_2 at $t = 2\Delta t$, ..., and X_n at $t = n\Delta t = T$. This is one realization or *instance* of an infinity of possible paths. Figure 1.3 was created using this algorithm.

1.5 Geometric Brownian Motion

There are two shortcomings with the use of an ARW for modeling stock prices. Even if started from a positive value, $X_0 > 0$, the walk can attain negative values; this is undesirable for stock prices. Secondly, stocks selling at small prices tend to have small increments in price while stocks selling at high prices tend to have much larger increments in price.

By happy circumstance, both problems are easily fixed by the same solution originally proposed by the MIT economist Paul Samuelson in 1965. The solution is that a stock's price increment should be *proportional* to the present price. In other words, if S_t is the current stock price, then an infinitesimal change in price, dS, will be given by

$$dS = S_t(\mu dt + dW_t) \tag{1.14}$$

where the dW_t are increments of a Wiener process. With this change, a stock's price can never go below 0 because when $S_t = 0$, the jump size is also 0 (recall, a Wiener process is continuous).

1.5.1 Price Volatility

By its definition a Wiener process has a parameter σ. Further the variance of W_t is $\sigma^2 t$ showing that σ controls the degree of dispersement of the process. It plays the same role in its application to stock prices where it is called *volatility*. It measures the tendency of a stock's price to oscillate or otherwise depart from a constant value. Each stock has its own volatility which can be estimated statistically using its recent price data. This is called the stock's *historical* or *statistical* volatility.

Henceforth we will use the notation W_t to refer to the standard Wiener process and the notation σW_t for the process with parameter σ. With this modification, (1.14) can be written as

$$
\begin{aligned}
\frac{dS}{S_t} &= \mu dt + \sigma dW_t \\
&= \mu dt + \sigma\sqrt{dt}Z_t, \quad Z_t \sim N(0,1).
\end{aligned}
\tag{1.15}
$$

This is called the *geometric Brownian motion* (GBM) model for stock prices. It is also referred to as a *drift-diffusion* model.

In the GBM model, two parameters characterize any given stock, its drift and its volatility. Both can be estimated from a sequence of recent prices, $\{S_0, S_{\Delta t}, S_{2\Delta t}, \ldots, S_{n\Delta t}\}$.[9]

Calculate the sequence of returns

$$
c_i = \frac{\Delta S_i}{S_i} = \frac{S_{i+1} - S_i}{S_i}, \quad i = 0, 1, \ldots, n-1. \tag{1.16}
$$

From (1.15) the mean of the c_i is $\mu\Delta t$ and the variance is $\sigma^2\Delta t$, so

$$
\begin{aligned}
\mu &\approx \frac{1}{n\Delta t}\sum_0^{n-1} c_i \\
\sigma^2 &\approx \frac{1}{(n-1)\Delta t}\sum_0^{n-1}(c_i - \mu\Delta t)^2.
\end{aligned}
\tag{1.17}
$$

Drifts and volatilities are reported on an annual basis. While an "annual basis" in finance often means 252 days/year because there are that many trading days,[10] we will use 365 days/year in order to be consistent across several applications. The conversion to trading day years is usually a straightforward matter of dividing by 365 and multiplying by 252. For example in the drift calculation above, if $\Delta t = 5$ days, the sum is divided by $n\Delta t = 5n$ as indicated and then multiplied by 365 days/year to get a calendar year drift or 252 days/year for a financial year drift.

The second calculation generates volatility squared, which has units per time same as the drift. A square root must be taken to get volatility. Thus its units are per square root time (and this must be taken into account when converting a volatility). Volatilities are normally quoted in percentages on a annual basis, for example one speaks of a 40 % volatility (annually).

Recall that (1.8) was derived as the end point probability density of the random walk on the line. By putting $D = \sigma^2/2$ this density is $N(0, \sigma^2 t)$. From (1.12) it is the same as that of W_t. This shows the connection between volatility and the diffusion process of a Brownian motion.

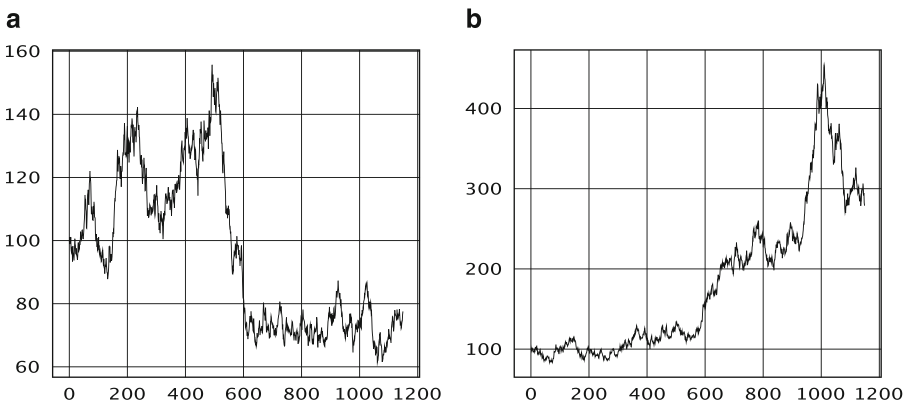
In the GBM model the drift term leads to exponential growth of the mean with growth rate μ. For, in the absence of the diffusion process, the differential equation is $dS/S = \mu dt$. And its solution is

[9] Theoretically this is so. But estimating the drift encounters a fundamental problem known as *statistical blur*. Error in the drift is given by the standard deviation. As Δt is reduced, the per period value of μ decreases by the same factor, but the per period value of standard deviation decreases by the square root of Δt. Hence for small periods, the error in drift exceeds the value of the drift itself.

[10] Approximately, it depends on the year.

$$S = S_0 e^{\mu t}$$

where S_0 is the initial value of S. See also (1.28).



An instance of a GRW. Another instance.

Fig. 1.5. Two instances of a geometric random walk (GRW) using all the same parameters showing the variety possible

In Fig. 1.5 we show the resulting path of two simulations of (1.15). The time scale is the same as in Fig. 1.1, 1,152 days. The starting value is $S_0 = 100$ and volatility is $\sigma = 40\,\%$. The figure shows that the geometric Brownian motion model is capable of a great deal of variety in terms of price paths.

The simulations depicted were generated by the following algorithm. Inputs are: starting price S_0 in units of currency, period of time of the study T in years (number of days divided by 365), volatility σ in per square root year, and drift μ in per year. We will refer to this as the *geometric random walk* (GRW) algorithm.

Algorithm 2. Simulating GBM

`inputs:` S_0, T, μ, σ
$\Delta t = 1/365.0$ $\quad\triangleright$1 day time increments in years
$n = T/\Delta t$ $\quad\triangleright$number of Δt steps in time T
`for` $t = 1, \ldots, n$
  $Z_t \sim N(0,1)$
  $\Delta S_t = S_{t-1}(\mu \Delta t + \sigma \sqrt{\Delta t} Z_t)$
  $S_t = S_{t-1} + \Delta S_t$;
`endfor`
  $\triangleright$the last S_t is S_T

1.5.2 Geometric Brownian Motion End Point Distribution

While individual realizations can give some idea of how stock prices behave, more important information for making inferences from the model derives from the distribution of the ending price, S_T, over all possible realizations. We call this the *maturity distribution*.

We can get a sense of that by running a large number of simulations and graphing the results. Figure 1.6 is a histogram of 100,000 trials of Algorithm 2. We see that the distribution is not a normal distribution; the tail on the upside is longer than on the downside. This is because downside prices are restricted to non-negative values.

To determine the maturity distribution, let the interval $[0, T]$ be divided into n subdivisions of equal length Δt, $T = n\Delta t$. From the algorithm above

$$S_i = S_{i-1}(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i)$$

for $i = 1, 2, \ldots, n$. To find the ending price S_n, just multiply the factors together[11]

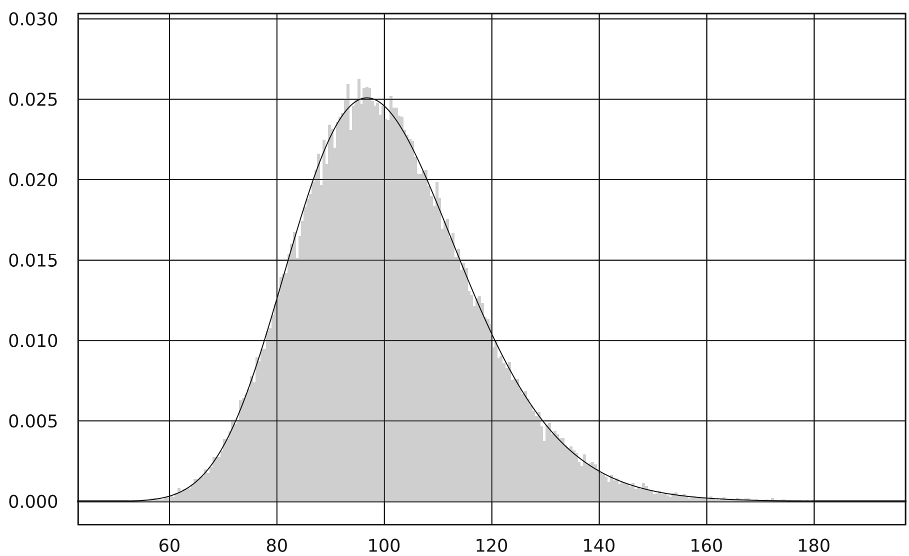


Fig. 1.6. Geometric random walk using the parameters $S_0 = 100$, $T = 60$ days, $\sigma = 0.4$. The *solid lines* gives the exact lognormal distribution

$$S_n = S_0 \prod_{i=1}^{n} \left(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i\right). \tag{1.18}$$

The Z_i are independent $N(0, 1)$ random variables. The starting price is S_0 at $i = 0$. The exact ending price S_T is the limit of this as $n \to \infty$ or as $\Delta t \to 0$ with $T = n\Delta t$ constant.

Start by dividing the equation by S_0 and take the logarithm of both sides. Since the logarithm of a product is the sum of the logarithms of its factors, we get

$$\log \frac{S_n}{S_0} = \sum_{i=1}^{n} \log \left(1 + \mu\Delta t + \sigma\sqrt{\Delta t}Z_i\right). \tag{1.19}$$

[11] The symbol $\prod_i^n a_i$ means the product of the a_i from $i = 1$ to $i = n$.

Recall the Taylor series expansion of the logarithm,[12]

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots.$$

Expanding logarithms, (1.19) becomes

$$\log \frac{S_n}{S_0} = \sum_{i=1}^{n} \Bigg((\mu \Delta t + \sigma\sqrt{\Delta t} Z_i) - \frac{1}{2}(\mu \Delta t + \sigma\sqrt{\Delta t} Z_i)^2 + \frac{1}{3}(\mu \Delta t + \sigma\sqrt{\Delta t} Z_i)^3 + \ldots \Bigg) \tag{1.20}$$

By expanding out the right hand side and taking the limit as $n \to \infty$ we arrive at the following

$$\log \frac{S_T}{S_0} \sim \mu T - \frac{1}{2}\sigma^2 T + \sigma\sqrt{T} Z. \tag{1.21}$$

The details are given in the box on the next page. This shows that the logarithm of S_T/S_0 is normally distributed. Since S_0 is a constant and $\log(S_T/S_0) = \log S_T - \log S_0$, (1.21) can be written as

$$\log S_T \sim \log S_0 + \mu T - \frac{1}{2}\sigma^2 T + \sigma\sqrt{T} Z, \tag{1.22}$$

From this equation we can read off the mean and variance,[13]

$$\begin{aligned} \mathbb{E}(\log S_T) &= \log S_0 + \mu T - \frac{1}{2}\sigma^2 T \\ \mathrm{var}(\log S_T) &= \sigma^2 T \end{aligned} \tag{1.23}$$

Knowing that $\log(S_T)$ is normally distributed allows us to find the distribution of S_T itself. In general suppose X is normally distributed, $X \sim N(\alpha, \beta^2)$, and let $F_R(y)$ be the cumulative distribution function of $R = e^X$ (so that $\log R = X$).[14] We have

$$\begin{aligned} F_R(y) &= \Pr(R < y) = \Pr(\log R < \log y) \\ &= \Pr(X < \log y) = \int_{-\infty}^{\log y} \frac{1}{\beta\sqrt{2\pi}} e^{-\frac{(u-\alpha)^2}{2\beta^2}} du. \end{aligned}$$

The density function $f_R(y)$ of R is the derivative of this; so from calculus

$$\begin{aligned} f_R(y) = F_R'(y) &= \frac{1}{\beta\sqrt{2\pi}} e^{-\frac{(\log y-\alpha)^2}{2\beta^2}} \left(\frac{d\log y}{dy}\right) \\ &= \frac{1}{y\beta\sqrt{2\pi}} e^{-\frac{(\log y-\alpha)^2}{2\beta^2}}. \end{aligned} \tag{1.24}$$

[12] See Section A.1.

[13] The mean of a constant is the constant itself, the variance of a constant is zero. Of course the mean and variance of Z is 0 and 1 respectively.

[14] See Section A.3.

In summary, (1.24) gives the probability density function of a lognormally distributed random variable R having parameters α and β^2. We denote this distribution as $R \sim LN(\alpha, \beta^2)$.

The first term of the expansion (1.20) equals μT since the n-fold sum of that term is $\mu n \Delta t$. The second term can be written as

$$\sigma\sqrt{T}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i.$$

But $\mathbb{E}(Z_i) = 0$ and $\mathrm{var}(Z_i) = 1$, so the Central Limit theorem applies here and in the limit as $n \to \infty$ this term tends, in distribution, to

$$\sigma\sqrt{T}Z$$

where $Z \sim N(0,1)$.
For the next series of terms we first calculate the square,

$$-\frac{1}{2}\sum_{i=1}^{n}(\mu^2 \Delta t^2 + 2\mu\sigma(\Delta t)^{3/2} Z_i + \sigma^2 \Delta t Z_i^2).$$

This time the sum of $\mu^2\Delta t^2$ is $\mu^2 T(\Delta t)$. Hence this term will go to 0 as $\Delta t \to 0$. For the next term we invoke the Central Limit theorem just as above. But this time the limiting random variable is multiplied by Δt, so the whole term tends to 0 at $n \to \infty$,

$$-(\Delta t)\mu\sigma\sqrt{T}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i \to 0. \tag{1.25}$$

For the last term we have to deal with sums of Z_i^2 where each is an independent $N(0,1)$ random variable, say Z. Now the mean of Z^2 is also the variance of Z and so is 1. The variance of Z^2 itself is given by

$$\mathrm{var}(Z^2) = \mathbb{E}(Z^4) - \mathbb{E}^2(Z^2) = 3 - 1 = 2.$$

Here we have used the fact that the 4th moment of Z is 3 as can be verified by direct integration*. Therefore by the Central Limit theorem

$$\begin{aligned}
-\frac{1}{2}\sigma^2 \Delta t \sum_{i=1}^{n} Z_i^2 &= -\frac{1}{2}\sigma^2 \Delta t \left(\sum_{i=1}^{n} Z_i^2 - n\right) - \frac{1}{2}\sigma^2 \Delta t n \\
&= -\frac{1}{2}\sigma^2 \sqrt{\frac{2}{n}} T \left(\frac{\sum_{i=1}^{n} Z_i^2 - n}{\sqrt{2n}}\right) - \frac{1}{2}\sigma^2 T \\
&\to -\frac{1}{2}\sigma^2 T \quad \text{as } n \to \infty.
\end{aligned}$$

This is as far as we need to go in the series, the remaining terms tend to 0 as n tends to infinity. Therefore we have shown that

$$\log \frac{S_T}{S_0} \sim \mu T + \sigma\sqrt{T}Z - \frac{1}{2}\sigma^2 T.$$

$^*\mathbb{E}(Z^4) = \int_{-\infty}^{\infty} \frac{x^4}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = 3$. Use integration by parts with $u = x^3$ and $dv = xe^{-\frac{x^2}{2}}\, dx$.

The mean, $\mathbb{E}(R)$ and variance, $\mathbb{E}(R^2) - \mathbb{E}(R)^2$ of the lognormal follow directly from some lengthy integrations against the density just derived, the result is

$$\begin{aligned} \mathbb{E}(R) &= e^{\alpha + \frac{1}{2}\beta^2} \\ \mathrm{var(R)} &= (e^{\beta^2} - 1)e^{2\alpha+\beta^2}. \end{aligned} \tag{1.26}$$

We may now apply the above to $R = S_T$. Since $X \sim N(\alpha, \beta^2)$, from (1.23), it follows directly that

$$\begin{aligned} \alpha &= \log S_0 + \mu T - \frac{1}{2}\sigma^2 T \\ \beta^2 &= \sigma^2 T. \end{aligned} \tag{1.27}$$

And so from (1.26)

$$\mathbb{E}(S_T) = e^{\log S_0 + \mu T - \frac{1}{2}\sigma^2 T + \frac{1}{2}\sigma^2 T} = S_0 e^{\mu T} \tag{1.28}$$

and

$$\begin{aligned} \mathrm{var}(S_T) &= (e^{\sigma^2 T} - 1)e^{2\log S_0 + 2\mu T - \sigma^2 T + \sigma^2 T} \\ &= S_0^2 (e^{\sigma^2 T} - 1)e^{2\mu T}. \end{aligned} \tag{1.29}$$

The curve in Fig. 1.6 was drawn using these values for α and β in (1.24).

We have shown that $S_T \sim LN(\alpha, \beta^2)$ with α and β as in (1.27). Equivalently by taking $R = S_T/S_0$ we have that

$$\frac{S_T}{S_0} \sim LN(\mu T - \frac{1}{2}\sigma^2 T, \sigma^2 T), \tag{1.30}$$

that is, $\alpha = \mu T - \frac{1}{2}\sigma^2 T$ here.[15] Log normal samples may be computed by choosing $X \sim N(\alpha, \beta^2)$ and putting $R = e^X$,

$$R = e^X, \quad X = \alpha + \beta Z, \quad Z \sim N(0, 1). \tag{1.31}$$

For future reference we note that the *median* of the lognormal distribution is given by

$$\mathrm{median}(S) = e^{\alpha}.$$

This is the value for which a sample from the lognormal is equally likely to be smaller than as larger than. As applied to the GBM ending price, we have

$$\mathrm{median}(S_T) = S_0 e^{(\mu - \frac{1}{2}\sigma^2)T}. \tag{1.32}$$

Hence the median ending price is less than the mean ending price. This is due to the asymmetry or *skew* of the distribution.

[15] The density may be plotted either as $y(x) = f(x)$ with α as in (1.27) or $y(x) = (1/S_0)f(x/S_0)$ with $\alpha = (\mu - \frac{1}{2}\sigma^2)T$.

Better Models for Stock Prices

Do stock prices really follow GBM? Later on in this chapter we examine one implication of the axioms on page 7 in the light of this question – the Markov property. More generally the answer seems to be not exactly but good enough for most work. Further GBM is a starting point for more complicated models known as Lévy processes. This is a topic we take up in Chapter 6.

1.6 Binomial Lattice Approximation

The random walk technique we have explored in the previous sections gives an accurate computational tool for stock price realizations and we will have many occasions to use it throughout this text. But it is computationally intensive. Several thousands of realizations must be run in order to achieve good results. (Fortunately, today, even personal computers can do the required calculations in a few seconds.)

By contrast, there is a very simple approximate technique that also gives good results and does so quickly because it is a deterministic method. It is the *binomial model* due to Cox, Ross, and Rubinstein. The binomial model consists of a lattice structure representing prices evolving in time. It is simple to construct, easy to understand, and provides answers in virtually all cases. Moreover, as the binomial lattice is refined, its results improve. In the limit, binomial lattice prices agree with those of GBM.[16]

1.6.1 Binomial Pricing Model

In the binomial pricing model, the time horizon interval $[0, T]$ is divided into discrete periods Δt called levels. The number of such periods is $n = T/\Delta t$. We will construct a binomial lattice graph with vertices or nodes at each of the times 0, Δt, $2\Delta t$, …, $n\Delta t = T$.[17] At the first level, level 0, time equal 0 ($= 0\Delta t$), and the graph has only one node, N_0, with price S_0 as usual. From here the model postulates that the price can go up by some factor, $u \geq 1$, or down by a factor, $0 < d \leq 1$. Hence at time 1 ($= 1\Delta t$) there are two nodes, designated as $N_1(1)$ and $N_1(0)$, with the prices

$$\text{for } N_1(1){:}\;\; S_0 u, \quad \text{and for } N_1(0){:}\;\; S_0 d.$$

The construction now repeats at every level to carry prices over to the next time period. Thus at time 2 there are three nodes $N_2(2)$, $N_2(1)$ and $N_2(0)$ with corresponding prices

$$\begin{aligned} N_2(2):&\;\; S_1(u)u = S_0u^2 \\ N_2(1):&\;\; S_1(u)d = S_0du \\ N_2(0):&\;\; S_1(d)d = S_0d^2. \end{aligned}$$

[16] See Appendix C.

[17] A binomial tree graph is a directed graph with two outward edges at every node or none for leaf nodes. If the tree re-connects as here, it is called a binomial lattice.

Notice that the up price from $N_1(0)$ is $u(S_0d) = S_0ud$ and so equals the down price from $N_1(1)$; in other words the graph reconnects so that there are just 3 nodes at time 2 and not 4. This is one of the keys to success of the binomial model, as more levels are added, the lattice grows arithmetically, not geometrically. See Fig. 1.7.

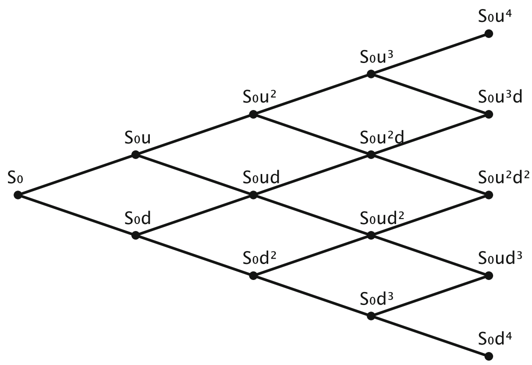


Fig. 1.7. A four step binomial tree

Along with the up and down price factors there is also the probability that the price will go up and the complementary probability that the price will go down. These are the *statistical* probabilities of the stock gotten from its recent price data as developed in Section 1.5.1. Consequently, attached to each node of the graph is a probability as well as a price. Let p be the probability of an up move and $q = 1-p$ the probability of a down move. The probability of reaching node $N_1(1)$ is p and of reaching $N_1(0)$ is q. At the end of the second period we have that node $N_2(2)$ occurs with probability p^2 and $N_2(0)$ with probability q^2. But the probability of $N_2(1)$ is $2pq$. This is because there are two paths to $N_2(1)$. In general, at the end of the kth period

$$N_k(i) \quad \text{occurs with probability} \quad \binom{k}{i} p^i q^{k-i} \tag{1.33}$$

In this model the up and down factors u and d as well as their probabilities p and q are constant over the entire time horizon.

1.6.2 Calculating the Binomial Factors

Of course the ending prices and probabilities depend crucially on u, d and p. We must fix them as appropriate for the given application, depending on μ, and σ. On the one hand, the expectation of S_1 as given by the binomial model is

$$\mathbb{E}(S_1) = puS_0 + (1-p)dS_0.$$

On the other hand, from (1.28)

$$\mathbb{E}(S_1) = S_0 e^{\mu \Delta t}$$

Equating these two we get

$$pu + (1-p)d = e^{\mu \Delta t}. \tag{1.34}$$

Now match variances. From the binomial model

$$\begin{aligned} \operatorname{var}(S_1) &= p(uS_0)^2 + (1-p)(dS_0)^2 - \mathbb{E}^2(S_1) \\ &= S_0^2 \left(pu^2 + (1-p)d^2 - e^{2\mu \Delta t}\right). \end{aligned}$$

From the continuous approach we have, using (1.29),

$$\operatorname{var}(S_1) = S_0^2(e^{\sigma^2 \Delta t} - 1)e^{2\mu \Delta t}.$$

Equating we get

$$pu^2 + (1-p)d^2 = e^{(2\mu+\sigma^2)\Delta t}. \tag{1.35}$$

Satisfying these two equations for the three parameters u, d, and p will match the statistical characteristics of the two approaches. We are thus left with an arbitrary choice for the third equation. The customary ones are either: $u = 1/d$ or $p = 1/2$. We take up both cases.

The $u = 1/d$ Case

First suppose $u = 1/d$. We start by eliminating p. Solve for p in (1.34)

$$p = \frac{e^{\mu \Delta t} - d}{u - d}. \tag{1.36}$$

Do the same in (1.35)

$$p = \frac{e^{(2\mu+\sigma^2)\Delta t} - d^2}{u^2 - d^2}$$

and equate

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

Multiply out the common factor $u - d$ and solve for $u + d$ remembering that $u = 1/d$,

$$\frac{1}{d} + d = \frac{e^{(2\mu+\sigma^2)\Delta t} - d^2}{e^{\mu \Delta t} - d}.$$

Multiplying through by the denominators and, doing some algebra, we arrive at a quadratic equation in d,

$$d^2 - \left(e^{-\mu \Delta t} + e^{(\mu+\sigma^2)\Delta t}\right) d + 1 = 0.$$

Designate the coefficient of the linear term as $2A$,

$$2A = e^{-\mu \Delta t} + e^{(\mu+\sigma^2)\Delta t}, \tag{1.37}$$

and solve the quadratic $d^2 - 2Ad + 1 = 0$. We find that

$$
\begin{aligned}
d &= A - \sqrt{A^2 - 1}, \\
u &= A + \sqrt{A^2 - 1}, \\
p &= \frac{e^{\mu \Delta t} - d}{u - d},
\end{aligned} \tag{1.38}
$$

where the second equation follows since $u = 1/d$ and the third is (1.36) repeated here for convenience. The parameter A is given in (1.37). If it works out that either $p \leq 0$ or $p \geq 1$, then this method cannot be used; try the $p = 1/2$ method instead.

The $p = 1/2$ Case

Now suppose $p = 1/2$. From (1.34) we get

$$u + d = 2e^{\mu \Delta t},$$

and from (1.35) we get

$$u^2 + d^2 = 2e^{(2\mu + \sigma^2)\Delta t}.$$

Solve the first for u and substitute into the second,

$$u = 2e^{\mu \Delta t} - d, \tag{1.39}$$

and so

$$4e^{2\mu \Delta t} - 4e^{\mu \Delta t} d + 2d^2 = 2e^{2\mu \Delta t} e^{\sigma^2 \Delta t}.$$

Again we have a quadratic in d. With some algebra we can put it into the form

$$d^2 - 2e^{\mu \Delta t} d + e^{2\mu \Delta t}(2 - e^{\sigma^2 \Delta t}) = 0.$$

By the quadratic formula we get two possibilities for d; the one we want is given by choosing the minus sign. Knowing d, use (1.39) to find u. The resulting solution for this case is

$$
\begin{aligned}
d &= e^{\mu \Delta t}\left(1 - \sqrt{e^{\sigma^2 \Delta t} - 1}\right), \\
u &= e^{\mu \Delta t}\left(1 + \sqrt{e^{\sigma^2 \Delta t} - 1}\right), \\
p &= \frac{1}{2}.
\end{aligned} \tag{1.40}
$$

In implementing this solution, d must remain positive.

The histogram in Fig. 1.8 shows the resulting distribution of S_T using the values of u, d, and p given by (1.40) with the same parameters for S_0, μ, and σ as in Fig. 1.6. Thus the two can be compared.

Besides providing the ending probabilities, the lattice can also be used to calculate the probability that a specific path is taken through the lattice. It is just the product of the probabilities along each edge of the path. For each up edge taken multiply by p and for each down edge multiply by q. For example, in Fig. 1.7, the path from node N_0 (with price S_0) to node $N_4(2)$ (with price $S_0 u^2 d^2$) by way of nodes: $N_1(1)$, $N_2(1)$, and $N_3(2)$ has probability $p^2 q^2$ of being taken as it includes 2 up edges and 2 down edges. But this is also seen in the specification of the ending price, $S_0 u^2 d^2$. In this way, a binomial lattice can be used to calculate path dependent financial instruments such as certain options.

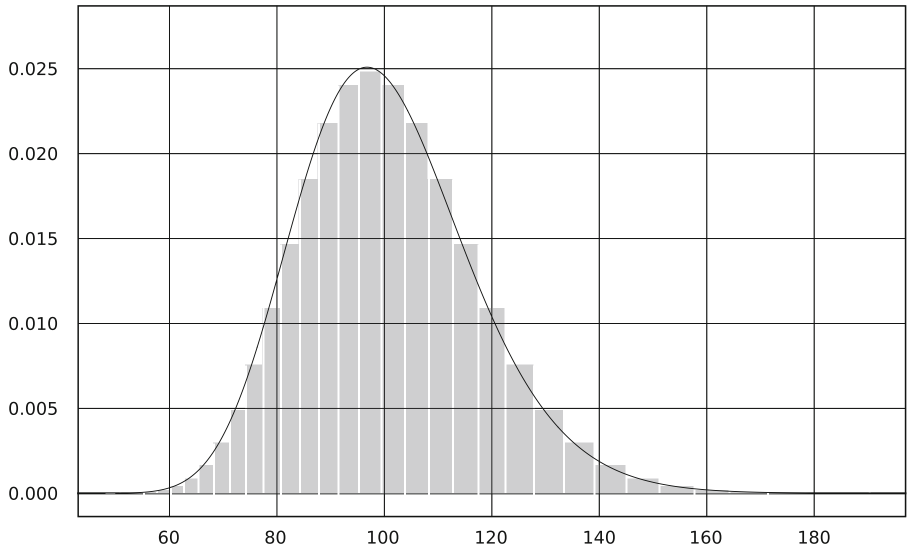


Fig. 1.8. Ending prices and probabilities for the binomial model using parameters $S_0 = 100$, $T = 60$ days, $\sigma = 0.4$. The *solid line* gives the exact lognormal distribution

1.7 Efficient Market Hypothesis

In Section 1.5 we saw that the GBM model, equation (1.15), predicts how a stock's price will behave in the future. Certainly the particular realization that will occur is not known, to this extent the future path of the price is random. But it is not without structure. The model predicts the future price will be a sample from a lognormal distribution. But to what extent is the model correct? That is the subject of this section.

The geometric random walk model asserts that from moment to moment the price will increase (or decrease depending on the sign of the drift) by a deterministic increment and that will be combined with a random increment, $S_t \sigma dW_t$, normally distributed. While the former could be impacted by factors such as the financial sector in which the company operates, management decisions, and other external factors, the latter is almost literally a coin toss. The random component has no memory. It could be up or down equally likely.

In the application of the model to real stock prices, not unlike the application of the arithmetic random walk model to real particles, moment to moment is only an approximation. It could mean minute to minute, day to day, or even week to week. In these periods of time, drift does not have much of an effect, even 10 % per year is less than 0.03 % per day. Thus over the short term, stock price movements are just ... random.

This feature of the model was known to Bachelier. Today it is one of the many assertions of the more broadly conceived *efficient market hypothesis* (EMH). Broader in that EMH applies to longer time frames than just moment to moment. The efficient market hypothesis states that financial markets are "efficient" in that prices already reflect all known information concerning a stock. Information includes not only what is currently known, but also future expectations, such as earnings and dividends.

Only new information will move stock prices significantly, and since new information is presently unknown and occurs at random, good or bad, future movements in stock prices are also unknown and thus, random.

The basis of the efficient market hypothesis is that the market consists of many rational investors who are constantly reading the news and reacting quickly to any new significant information about a security.

The EMH implies that investors cannot gain advantage by interpreting stock charts. It is no more possible to do so than to observe the past sequence of tails and heads of a fair coin toss and predict the next outcome. It is also not possible to gain advantage by scrutinizing earnings reports or other company information. Any information that might predict the future direction of the stock price is already incorporated into the current stock price. Nor can seasonal variation or other periodicity be taken advantage of because other investors know what to expect and have already tuned the current price to correct for it.

Nevertheless many investors, both professional and non-professional alike, believe there is predictive information in the charts, These investors are called *chartists* and the reasoning they use is called *technical analysis* (TA).

The several implications of the EMH are often broken out by type. The assertion that no prediction of future prices can be inferred from past prices is referred to as the *weak form* of the efficient market hypothesis. It is this form that denies the possibility of making excess profits in the market using, for example, price and volume charts, that is, from *technical analysis.* We will investigate this aspect of EMH in the balance of this section.

But first we mention the other forms of the EMH. The *semi-strong form* maintains that no prediction of future prices can be inferred from any public information whatsoever. This means not only that technical analysis cannot predict future prices, but also that a company's *fundamental information* cannot either. Fundamental information includes such things as market capitalization, price to earnings ratio, price to sales ratio, revenue per share, debt to equity ratio, and so on.

Finally the *strong form* of the EMH asserts that, along with the public information, not even insider information can give an investor an advantage. This is strong indeed!

1.7.1 Simple Moving Average

Armed with a model for calculating future price scenarios, it should be possible to test the weak form of efficiency, or, equally, the validity of technical analysis, by Monte Carlo. We first take a brief look at TA. In fact there is a vast literature on the subject.

Technical analysis is precisely the art of using historical price and volume data in an attempt to predict future price trends barring major economic news. Technical analysts attempt to identify archetypal patterns in charts such as *support*, *resistance*, *channels*, *head and shoulders*, *double tops or bottoms* and many others. Having identified such a pattern, a future price prediction is thereby implied, although not necessarily with certainty.

The most basic tool of TA is the *moving average* or sometimes called the simple moving average. It is an attempt to smooth out the short term fluctuations inherent in market prices and reveal their underlying trend. The *n-day moving average* is the sum of the last n days period of prices divided by n. Let S_i denote the stock price on day i relative to some start date; i could be negative, this would be the case if S_0 is today's price. The moving average on day t is

$$\mathrm{ma}_t = \frac{1}{n}\sum_{i=0}^{n-1} S_{t-i} = \frac{S_t + S_{t-1} + \ldots + S_{t-n+1}}{n} \tag{1.41}$$

The moving average is a weighted average with the weight $1/n$ applied to each price. The parameter n is the *window* period of the moving average because only prices appearing in a window of length n affect the average. As time moves forward, the window moves with it. The recursion equation for the moving average is obtaining by adding in the new term at the front and subtracting the last term of the old,

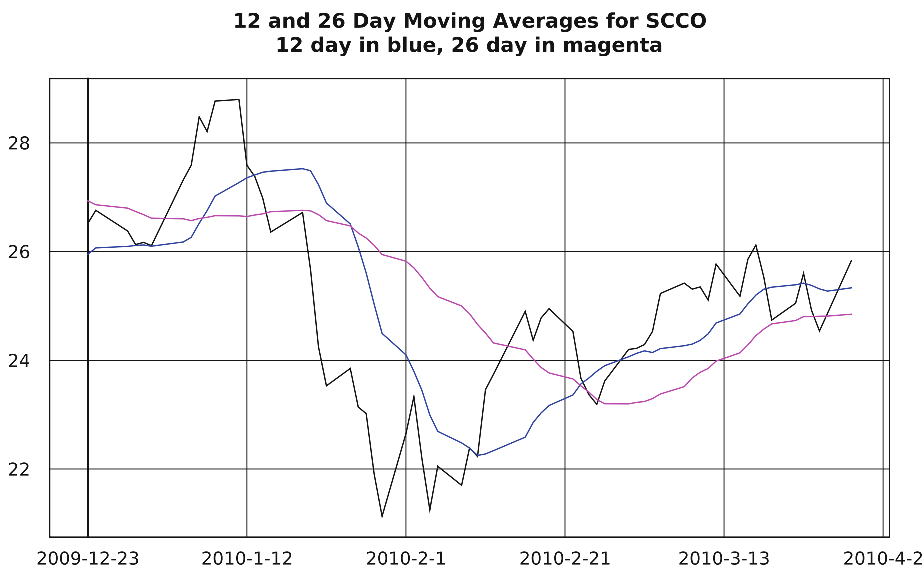


Fig. 1.9. SCCO prices with their 12 and 26 day moving averages

$$\mathrm{ma}_{t+1} = \mathrm{ma}_t + \frac{1}{n}(S_{t+1} - S_{t-n+1}). \tag{1.42}$$

In Fig. 1.9 we plot 3 months of stock prices along with their 12-day and 26-day moving averages. The 12-day average more closely tracks the raw prices while the 26-day average shows much less up and down movement. Both moving averages noticeably lag the raw prices, the 26-day more so.

An algorithm for computing a time series of n-day moving averages should take into consideration that stock prices, P_i, are usually given in reverse chronological order. Thus P_0 is today's price, P_1 is yesterday's, and so on up to P_R, the last price of the range occurring R days ago where $R > n$. It is helpful to

time			n prices				
$\longrightarrow$			acquired				today
P_R	P_{R-1}	...	P_{R-n+1}	P_{R-n}	...	P_1	P_0
S_0	S_1	...	S_{n-1}	S_n	...	S_{R-1}	S_R
			ma_{n-1}	ma_n	...	ma_{R-1}	ma_R

first put them into forward chronological order so that $S_0 = P_R$, $S_1 = P_{R-1}$, ..., and $S_R = P_0$ is today's price.

The first moving average that can be computed using a full complement of n prices is ma_{n-1}; it is based on the prices S_0 through S_{n-1}. From that point the average can be computed up to today, ma_R.

Algorithm 3. Running Simple Moving Average Calculation

```
for i = 0,...,R
   S_i = P_{R-i}
endfor

for t = n-1, n, n+1,...,R
   a = 0
   for i = 0,1,...,n-1
      a = a + S_{t-i}
   endfor
   ma_t = a/n
endfor
```

Strictly speaking the n-day moving average cannot be computed until n days after the starting date of the prices. At the penalty of accepting a slightly less smooth result initially, this is easily overcome by averaging over whatever number of terms are available until reaching the full complement of n.

1.7.2 Exponential Moving Average

As noted above, the simple moving average weights all terms in its sum equally. But it may be desirable to more heavily weight recent prices. The *exponential moving average* does just that. As time moves forward, more distant prices have less and less effect on the average. Instead of a window period, the parameter of the exponential moving average is the current price weight fraction f. To compute it, let ema_t denote today's exponential moving average, then

$$\begin{aligned} \mathrm{ema}_0 &= S_0 \\ \mathrm{ema}_t &= fS_t + (1-f)\mathrm{ema}_{t-1}, \quad t = 1, 2, \ldots \end{aligned} \tag{1.43}$$

This equation is the recursion for the exponential moving average. Writing it out explicitly gives

$$\mathrm{ema}_n = fS_n + f(1-f)S_{n-1} + f(1-f)^2 S_{n-2} + \ldots + f(1-f)^{n-1} + (1-f)^n S_0.$$

It shows that the exponential moving average includes a contribution from all prices from the beginning up to the current price. However the weight allocated to early prices falls off exponentially. In total the sum of weights is 1,[18]

$$\begin{aligned} &f + f(1-f) + f(1-f)^2 + \ldots + f(1-f)^{n-1} + (1-f)^n \\ &\quad = f\left(1 + (1-f) + \ldots + (1-f)^{n-1}\right) + (1-f)^n \\ &\quad = f\frac{1-(1-f)^n}{1-(1-f)} + (1-f)^n \\ &\quad = (1-(1-f)^n) + (1-f)^n = 1. \end{aligned}$$

To relate a simple moving average with an exponential one, the following conversion between window period and weighting fraction is customarily used:

$$\begin{aligned} \text{fraction} &= \frac{2}{\text{window}+1} \\ \text{window} &= \frac{2}{\text{fraction}} - 1. \end{aligned} \tag{1.44}$$

For example, a 5 day window simple moving average corresponds approximately to a 1/3 fraction exponential moving average.

1.7.3 Testing the EMH

Our first test of the weak form of the EMH is whether or not a stock price moves up or down equally likely from moment-to-moment. We take this to mean day-to-day. The test consists of performing a large number of the following trials:

- Select an equity at random from the list of those under test,
- Select a date at random from the period of time under test,
- Observe the closing price of the equity on that date,
- Observe the closing price of the equity on the next trading day,
- If the price increased, note that; if the price decreased, note that; otherwise discard this trial.

Results showing that the number of up days are about equal to the number of down days lend confidence in the random walk model. Contrary results would motivate a search for an explanation and further testing.

A complication in the test procedure is that for dividend paying stocks, the closing price decreases by the amount of the dividend on the ex-dividend date (independently of the normal price movement). Fortunately databases compensate for this by issuing an “adjusted closing price.” This price gives the true change across an ex-dividend day. (Note however that over extended periods of time the adjustments accumulate and an adjusted closing price several years back might be only a fraction of the actual closing price.)

[18] See Section A.1 for a refresher on the sum of geometric series.

The FIMCOM Database

Of course a central ingredient of the algorithm is access to a database of historical stock prices. At the time of this writing, historical prices are available on the internet, for example at finance.yahoo.com. However, as discussed in the preface, the FInance with Monte CarlO Methods (FIMCOM) database is also provided at the web page for this text, www.math.gatech.edu/~shenk; At the very least it can be used for testing our algorithm if not the EMH itself.

The FIMCOM data consists of daily prices and volume for 822 hypothetical symbols or tickers. The prices for each market day are: open, high, low, close, and adjusted close as discussed above. The data is available either as a zip file, *fimcom.zip*, or broken out into the individual price tables as comma separated values (csv) files.

The web page also provides two computer programs in the java programming language for downloading or accessing the database, *getWebPrices.java* and *getZipFile.java*.

A third program, *nextDayPrices.java*, makes use of the database to test the hypothesis that tomorrow's price is randomly up or down from today's as is presently under discussion. This program is also presented in the Appendix as a realization of the program sketch given above.

In the Fig. 1.10 we show the results for FIMCOM stocks over the dates 1/1990 through 6/2008, Altogether 600,000 trials, as outlined above, were performed. The chart depicts 600 batches of size 100 trials each. The mean is 50.39 and batch standard deviation is 5.20.[19]

The same test can be applied to various categories of companies, various date ranges, even days of the week. Some results of the test applied to the NYSE and NASDAQ are given in Table 1.1. Since the probabilities that the market will be

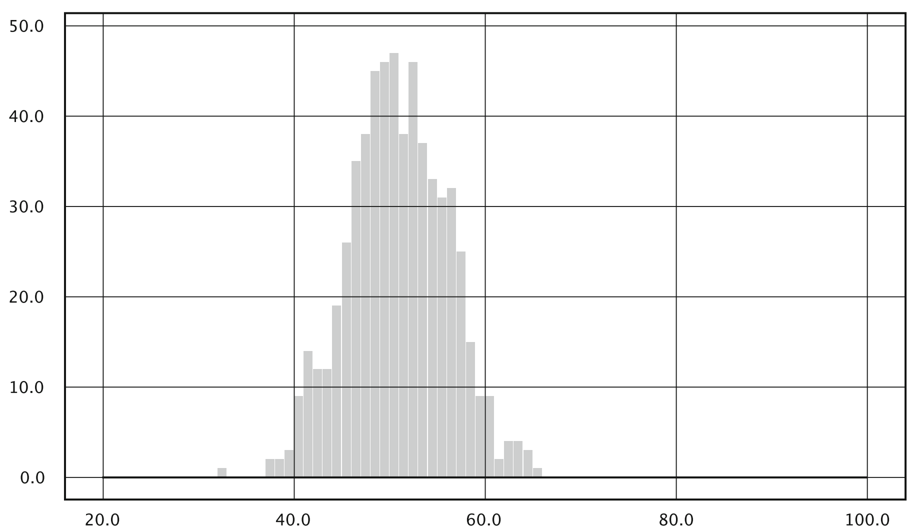


Fig. 1.10. Next day up for FIMCOM prices

[19] See Section A.8 for the rationale for batching statistical data.

up tomorrow as presented in the table are the result of Monte Carlo simulation, they are only statistical approximations to the true values as represented in the historical data. The table gives 95 % confidence intervals for the values. As can be seen, some of the values stray from 50 % somewhat. Evidently further investigation is warranted in these cases.

1.7.4 Testing an Advanced Indicator

So far we have only tested a very simple property of market prices. But the same technique can be used to test more elaborate technical analysis claims. As there are literally dozens of them, we restrict ourselves to just one for demonstration purposes and to provide an example from which others may be tested.

The *Direction Movement Indicator* (DMI) is defined as follows. The amount by which today's high price is higher than yesterday's is dm^+(or 0 if not higher). Conversely, the amount by which today's low price is below yesterday's is (numerically) dm^-. Whichever is smallest on a given day is reset to 0. Denote their 14 day (exponential) moving averages by $\mathrm{MA}(\mathrm{dm}^+)$ and $\mathrm{MA}(\mathrm{dm}^-)$ respectively.

Next calculate the 14 day moving average of the *true range*. The true range on a given day, tr_i, is the maximum of

$$H_i - L_i \quad \text{or} \quad H_i - C_{i-1} \quad \text{or} \quad C_{i-1} - L_i$$

Table 1.1. Probability that tomorrow's price is up from today's

Exchange	Sector	Dates	Up tomorrow (%)	95 % conf. interval
NYSE	All	1971–1990	50.60	±0.32
NYSE	All	1990–2008	50.02	±0.41
NYSE	Basic materials	1990–2008	50.44	±0.16
NYSE	Conglomerates	1990–2008	50.48	±0.32
NYSE	Consumer goods	1990–2008	49.91	±0.32
NYSE	Financial	1990–2008	50.62	±0.31
NYSE	Healthcare	1990–2008	48.69	±0.31
NYSE	Industrial goods	1990–2008	49.80	±0.41
NYSE	Services	1990–2008	49.35	±0.32
NYSE	Technology	1990–2008	48.62	±0.31
NYSE	Utilities	1990–2008	51.50	±0.32
NYSE	All	1996–2000	49.64	±0.16
NYSE	Basic materials	1996–2000	48.86	±0.31
NYSE	Conglomerates	1996–2000	50.66	±0.31
NYSE	Consumer goods	1996–2000	49.23	±0.32
NYSE	Financial	1996–2000	50.38	±0.32
NYSE	Healthcare	1996–2000	47.94	±0.31
NYSE	Industrial goods	1996–2000	49.46	±0.41
NYSE	Services	1996–2000	49.01	±0.32
NYSE	Technology	1996–2000	47.11	±0.32
NYSE	Utilities	1996–2000	50.85	±0.33
NASDAQ	All	1990–2008	48.86	±0.16
NASDAQ	All	1996–2000	47.75	±0.16
NASDAQ	All	2001–2003	48.74	±0.23

where H_i is today's high, L_i is today's low, and C_{i-1} is yesterday's close. If one thinks of yesterday's closing price as part of today's range of prices, then true range is just the difference between the day's high and low.

Finally today's dmi$^+$ and dmi$^-$ are the ratios

$$\begin{aligned} \mathrm{dmi}^+ &= \frac{\mathrm{MA}(\mathrm{dm}^+)}{\mathrm{MA}(\mathrm{tr})} \\ \mathrm{dmi}^- &= \frac{\mathrm{MA}(\mathrm{dm}^-)}{\mathrm{MA}(\mathrm{tr})}. \end{aligned} \tag{1.45}$$

A signal is generated when dmi$^+$and dmi$^-$cross, that is, when the difference crosses 0 either from below or from above. If dmi$^+$ crosses above dmi$^-$the prediction is that the stock price will trend up; conversely if dmi$^+$ crosses below dmi$^-$ the price is predicted to trend down.

Mathematically, begin with

$$\begin{aligned} \mathrm{dm}^+ &= \max(0, H_i - H_{i-1}) \\ \mathrm{dm}^- &= \max(0, L_{i-1} - L_i), \end{aligned} \tag{1.46}$$

followed by resetting the smallest to 0

$$\begin{aligned} \mathrm{dm}^+ &= (\mathbb{1}_{\mathrm{dm}^+ \geq \mathrm{dm}^-})\mathrm{dm}^+ \\ \mathrm{dm}^- &= (\mathbb{1}_{\mathrm{dm}^- \geq \mathrm{dm}^+})\mathrm{dm}^- \end{aligned} \tag{1.47}$$

Here the *indicator function* $\mathbb{1}_A$ is 1 when condition A is satisfied and 0 otherwise. The 14 day moving averages are given by

$$\begin{aligned} \mathrm{MA}_i(\mathrm{dm}^+) &= f\mathrm{dm}^+ + (1-f)\mathrm{MA}_{i-1}(\mathrm{dm}^+) \\ \mathrm{MA}_i(\mathrm{dm}^-) &= f\mathrm{dm}^- + (1-f)\mathrm{MA}_{i-1}(\mathrm{dm}^-) \end{aligned} \tag{1.48}$$

where $f = 2/15$. True range and its moving average are calculated as

$$\begin{aligned} \mathrm{tr}_i &= \max(H_i - L_i, H_i - C_{i-1}, C_{i-1} - L_i) \\ \mathrm{MA}_i(\mathrm{tr}) &= f\mathrm{tr}_i + (1-f)\mathrm{MA}_{i-1}(\mathrm{tr}). \end{aligned} \tag{1.49}$$

Then dmi$^+$ and dmi$^-$ are given by equation (1.45).

A sketch of the algorithm for testing dmi follows.

Algorithm 4. Testing the DMI Indicator

inputs: dmi moving average parameter f, date range to be tested R, day following signal for assessment (1 for tomorrow), exchange, sector

`for batchIndex` $= 1, \ldots, 600$
  `trialresult = 0` ▷initialize `trialresult`
  `for trials` $= 1, \ldots, 100$
  - choose a stock ticker `TKR` at random
  - read price and date data for `TKR` over R
  - compute dmi$^+$, dmi$^-$ and `dmiIndicator` $=$ dmi$^+ -$ dmi$^-$ over R

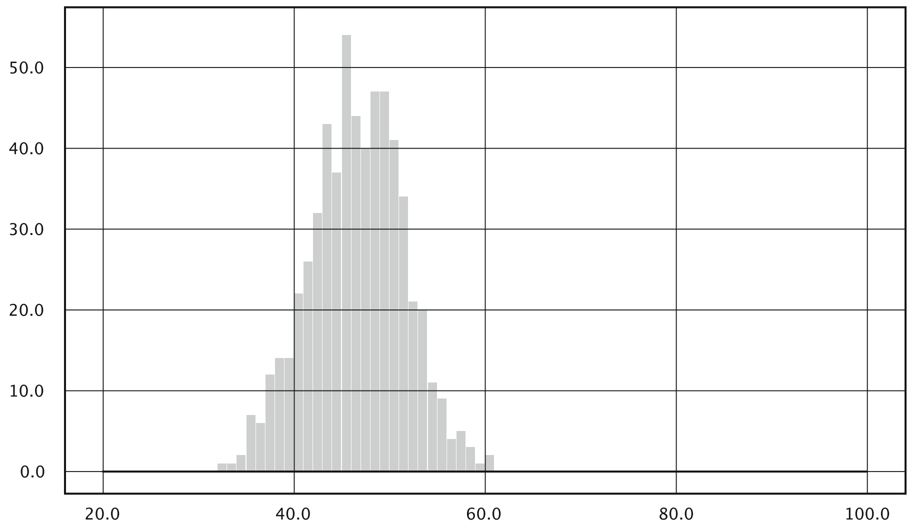


Fig. 1.11. Predictive success of DMI for FIMCOM prices

```
        • calculate the list of signal dates C when
          dmiIndicator crosses 0 and the prediction:
            up  if up crossing
            down  if down crossing
        • choose a signal date in C at random
        • compare prices on the predicted and signal dates
        • record a hit or miss in trialresult
      endfor
      • increment the histogram cell for trialresult
    endfor
```

In Fig. 1.11 we show the results of using dmi to predict tomorrow's trend for FIMCOM stocks over the dates 1/1990 through 6/2008. The trials were batched with batch size 100 as above. The mean is 46.11 and batch standard deviation is 4.92. This would seem to make dmi an anti-predictor.

In Table 1.2 we show the predictive success of dmi on NYSE and NASDAQ stocks under various conditions as specified. In the predicted day column a +1 means the prediction is whether tomorrow's price is up or down while +7 means the prediction is whether the price 7 days hence is up or down.

For more on the efficiency of the market with respect to technical analysis see [PI07].

Problems: Chapter 1

1. The following technique exploits the Central Limit Theorem[20] to create approximate samples Z from the standard normal distribution. (An exact method is given in Section A.9.) The mean of a uniformly distributed random variable U on $[0, 1]$,

[20] See Section A.6.

Table 1.2. Probability that DMI can predict a trend

Exchange	Sector	Dates	Pred. day	Predicted (%)	95 % conf. interval
NYSE	Basic materials	2001–2003	+1	47.11	±0.41
NYSE	Conglomerates	2001–2003	+1	48.50	±0.41
NYSE	Consumer goods	2001–2003	+1	48.94	±0.41
NYSE	Financial	2001–2003	+1	48.61	±0.41
NYSE	Healthcare	2001–2003	+1	49.31	±0.41
NYSE	Industrial goods	2001–2003	+1	47.56	±0.41
NYSE	Services	2001–2003	+1	48.60	±0.41
NYSE	Technology	2001–2003	+1	50.00	±0.40
NYSE	Utilities	2001–2003	+1	49.13	±0.40
NYSE	All	1990–2008	+1	50.11	±0.42
NYSE	All	1990–2008	+7	50.13	±0.39
NASDAQ	All	1990–2008	+1	49.21	±0.43
NASDAQ	All	1990–2008	+7	49.39	±0.41

denoted $U \sim U(0,1)$, is $\mu_U = 1/2$ and the variance is $\sigma_U^2 = 1/12$. Therefore by the CLT

$$Z = \frac{\sum_{i=1}^{n} U_i - n/2}{\sqrt{n/12}} \tag{1.50}$$

is approximately $N(0,1)$.

Algorithm 5. Approximate $N(0,1)$ Samples

```
inputs: n
Z = 0
for i = 1,...,n
   U ~ U(0,1)
   Z = Z + U
endfor
Z = (Z - n/2)/sqrt(n/12)
```

Generate a histogram from this algorithm with $n = 12$ and compare it with the standard normal density, (1.6) with $\mu = 0$ and $\sigma = 1$. Do the same for $n = 48$, and $n = 108$.

2. Let X_t describe a Brownian particle with parameter $\sigma = 4$ starting at $X = 0$ when $t = 0$ and moving for a time $t = T/4$. Now let Y_s, for $s = 0$ to $s = 3T/4$, describe the motion of the same particle from time $T/4$ until time T. How is $X_{T/4}$ distributed? How is $Y_{3T/4}$ distributed? How is $X_{T/4} + Y_{3T/4}$ distributed? Work this out either analytically or via simulation, for $T = 12$. This problem illustrates the inifinite divisibility property of the Wiener process.

3. (a) Suppose Brownian motion is used to model stock prices (instead of geometric Brownian motion). If $S_0 = 10$, $\mu = 0$ per year, volatility $= 1$ per square root year, and $T = 1/12$ years (about 30 days), what is the probability a stock's price S_T will be less than 0? less than 1? less than 9? less than 10?

(b) What is the probability the price went below 0 at some time before $t = T$ and ended above 0? Estimate this by simulation using several difference choices for Δt.

4. Same question as in Problem 3 but assume prices follow GBM with the same parameters. Compare the less than 10 values.

5. Starting from $S_0 = 100$, run 10,000 trials of a GRW with the following parameter sets (take $\Delta t = 1/365$ in every case). What fraction of outcomes are greater than S_0? less than S_0? equal to S_0?
(a) $\mu = 0$, $\sigma = 1$, $\texttt{nDays} = 30$ (b) $\mu = 0$, $\sigma = 1$, $\texttt{nDays} = 180$
(c) $\mu = 0$, $\sigma = 0.2$, $\texttt{nDays} = 30$ (d) $\mu = 0$, $\sigma = 0.2$, $\texttt{nDays} = 180$
(e) $\mu = 0.1$, $\sigma = 0.2$, $\texttt{nDays} = 30$ (f) $\mu = 0.1$, $\sigma = 0.2$, $\texttt{nDays} = 180$

6. Run 10,000 trials of a GRW with the following parameters: $\Delta t = 1/365, T = 60$ days, $\sigma = 40\,\%$, $\mu = 3\,\%$. What fraction of outcomes: (a) end between 105 and 115? (b) end between 95 and 100? (c) fall below 95 at some point but finish above 110? (d) rise above 105 at some point but finish below 100?

7. Answer the questions in Problem 5 by constructing a 6-step binomial tree. As in the text, Δt must equal T/n where n is the number of steps. Then all the parameters must be converted to use the same time units.

8. Answer the questions in Problem 6 by constructing a 6-step binomial tree. Again, Δt must equal T/n where n is the number of steps, so $\Delta t = 10$ days here. All the parameters must be converted to use the same time units.

9. If Z is distributed as $N(0, 1)$, how is $X = 3 + 6Z$ distributed? How is $S = e^X$ distributed? What is the mean and variance of S?

10. Let S_t, $0 \le t \le T$ be a Geometric Brownian Motion (GBM) random variable with drift μ and volatility parameter σ. Suppose $S_0 = 1$ and $\sigma^2/2 = \mu$. What is the mean of $\log S_T$? What is the mean of S_T itself? Is this sensible?

11. (a) Write a program to test the hypothesis that stock prices are up just as likely as down from one trading day to the next. Test that hypothesis on a database of stock prices of your choice.
(b) Test the hypothesis that if prices are down 2 days in a row, then they are up the third day.

12. MACD is the difference between a short term moving average and a long term moving average,
$$\text{MACD} = \text{maShort} - \text{maLong}.$$
Typically the short term average is 12 days and the long term is 26 days. When the short term exceeds the long term by the certain amount, the “overbought-oversold limit,” then it is maintained that the stock is overbought and it is predicted that the stock price will fall. Conversely when the short term is less than the long term by more than the overbought-oversold limit, then the stock is oversold and the price is predicted to rise. Write a program to test this hypothesis. (Exponential moving averages are permissible here.)

13. Research another technical analysis indicator and test it. See [Ach00] for an extensive list.

2

Return and Risk

This chapter is about the fundamentals of investment growth. It introduces important calculations with interest rates, returns, and discounting. These ideas will be needed in later chapters. It is also about investment risk, how it can be measured and how it can be minimized in the formation and maintenance of an investment portfolio. This is possible through one of the great financial breakthroughs, the *mean-variance theory* and CAPM, the capital asset and pricing model, due to Markowitz and his followers. In the 50 years since its introduction and subsequent development shortcomings of the theory have surfaced and improvements offered. Yet it is a starting point for these advanced theories and basic to a financial course of study.

Since most of the content of this chapter is deterministic, opportunities for Monte Carlo analysis are limited.

Market risk refers to the possibility of suffering a loss or even a less-than-expected return as a result of unexpected movements in some market, for example the currency, or real estate, or commodities market. However in this book our primary focus is on the stock market.

With few exceptions risk is ever present. A blue chip equity can go along smoothly for years issuing dividends on a regular basis only to succumb to unforeseen events.[1] Market prices for stocks, real estate, currencies, and precious metals rise and fall, for the most part, according to models we investigated in the last chapter. At the point of sale, which may not be at a propitious moment for the investor, a position is subject to the market price prevailing at that time. This very concrete "mark-to-market" valuation often results in a loss.

Over time risk has become better understood. Especially so upon the advent of the science of probability (see [Ber96]). Quantifying risk entails two components: fixing the amount of loss and second, its probability. Once risk became quantifiable, the investment community invented instruments for managing risk. This includes portfolio diversification, which we take up in this chapter, but also futures and option contracts which we take up in the sequel.

[1] GM was declared bankrupt in 2009.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_2,

2.1 The Risk-Free Rate

As already noted, in general all investment entails some risk. One exception to the rule, as near as possible, is an investment in government bonds. So much so that certain government securities are referred to as a *risk-free investment*.[2] U.S. bonds are one form of treasury securities; treasury bills and notes are two others.

Unlike a hard asset, a *financial instrument* is a item that derives its value from a promise to pay. If there is a well-developed market for them, then it is called a *security.* A *fixed-income security* is a security whose promise to pay is a definite amount to the holder over a given span of time. An *equity* is a security in which the investor has an ownership share, for example as in stock.

A *bond* is a fixed-income security representing the debt of the issuer, a company or government, to the holder of the bond. A bond has a stated *face* or *par* value which the issuer promises to pay the holder on the stated *maturity date.* In addition to that, a bond can have *coupons*, stated as a percentage of the face value, which the issuer pays the holder on an annual basis. (In some cases half the coupon payment is made semi-annually.) The final payment at maturity includes both the face value and the last coupon payment. A *zero-coupon bond* is one that has no coupons and just pays the face value at maturity. Originally the issuer sold the bond to borrow money. In short, a bond is an IOU for a loan with explicit payback terms.

U.S. Treasury *bills* are zero-coupon bonds. They pay no interest but sell at a discount of par value. U.S. Treasury *notes* have maturities between 2 and 10 years and have a coupon payment every 6 months. U.S. Treasury *bonds* have maturities between 20 and 30 years and a coupon payment every 6 months. These, and other, risk-free investments are centrally important throughout finance. The rate of return on a risk-free investment is called the *risk-free rate*, r_f. For investments in U.S. dollars, this is often taken as the yield rate on short-term treasury bills. These rates can be found at www.ustreas.gov/offices/domestic-finance/debt-management/interest-rate/yield.shtml.

The risk-free rate is a very important tool in use throughout finance. As we will see, it serves as a basis of comparison for all other investments. One of these is the determination of "fair" prices for many financial instruments such as futures and options. We will take up this dependence in a later chapter. It is important to understand that any rate of return exceeding the risk-free rate is considered to have risk, for example, dividend yields exceeding the risk-free rate.

The risk-free rate impacts many other investment rates throughout the financial system, for example interest rates on insured bank deposits, home mortgage rates, and corporate bond rates.

[2] This is so for US government bonds. A financial crisis was precipitated in 1998 when the Russian government defaulted on its debt.

2.2 Fixed-Income Securities Calculations

When return rates are predictable future payments can be calculated exactly. A *zero-coupon bond* is an example. It has a face value F and a maturity date T. The holder of the bond may exchange the bond for its face value on the maturity date. The original cost of the bond, P, is the investment. This fixes the rate of return.

Conversely, at the end of a financial transaction, when all the payments are known after the fact, then an analysis can be made as to the true return of the investment.

2.2.1 Simple Interest

Let P dollars be the value of an original investment, or bank deposit, and let ΔP be the gain (or loss if negative) after a period of time t in years. The value of the investment at that time is

$$A_t = P + \Delta P. \tag{2.1}$$

The *return* is the relative gain,

$$\frac{\Delta P}{P}. \tag{2.2}$$

It is usually expressed in percent. By *logarithmic return*, in brief *log-return*, we mean

$$\log(\frac{A_t}{P}) = \log(1 + \frac{\Delta P}{P}). \tag{2.3}$$

From the Taylor series for the logarithm, see (A.4), we have

$$\log(1 + \frac{\Delta P}{P}) = \frac{\Delta P}{P} - \frac{(\frac{\Delta P}{P})^2}{2} + \dots .$$

Thus, to first order, logarithmic returns and returns are the same.

The *rate of return*, or interest rate in the case of bank deposits, is

$$r = \frac{\Delta P}{Pt} \tag{2.4}$$

expressed in percent per year.[3] If the time period is 1 year, then the return and the rate of return are numerically the same.

Turning this equation around, after t years, an investment of P dollars earns the gain

$$\Delta P = Prt, \tag{2.5}$$

and the value at that time is

$$A_t = P(1 + rt). \tag{2.6}$$

[3] Some authors define the rate of return as $\frac{\Delta P}{P}$, [CZ03, Lun98]; others as we have done here, [Sch03]. As in most science and engineering applications, we prefer to reserve the term rate for changes per unit time.

The time t need not be a multiple of a year. If it is not, the amount earned is still proportional to the time invested. For example, if t is 3/2 years, then the investment returns one full years' amount and half of another.

Example 2.1. Two hundred dollars are deposited in an account paying 2 % per quarter. After 3 and 1/4th quarters the account is closed. The interest earned is $\$200(0.02)(3+1/4)=\$13$. This calculation could be put on an annual basis: 2 % per quarter is 8 % per year and 3 and 1/4 quarters is 13/16th years. Thus we have $0.02(13/4) = 0.02(4)(13/4)(1/4) = 13$. The amount of the investment when closed is \$213. □

2.2.2 Compounding

The rate r above is a *simple* earnings rate meaning the amounts accrued were not reinvested over the period; there is no interest on interest. But if accrued amounts are reinvested, the situation becomes quite different, it is called *compounding*.

Let the earnings be calculated annually, $t = 1$ year. The value of the investment after 1 year is

$$A_1 = P(1 + r).$$

If that money is reinvested, then $P(1+r)$ plays the role of the original investment and after 2 years it grows to

$$A_2 = P(1 + r)(1 + r) = P(1 + r)^2.$$

Continuing, after t years it becomes

$$A_t = P(1 + r)^t.$$

Compare this with (2.6).

Suppose an investment offers more frequent compounding periods. If r is the annual rate of return but compounding is quarterly, then the quarterly rate is $r/4$. Now quarterly periods play the role of years in the previous equations, hence after n quarters, or $t = n/4$ years, we have

$$A_t = P(1 + \frac{r}{4})^n = P(1 + \frac{r}{4})^{4t}.$$

In particular, at the end of 1 year, the amount becomes

$$A_1 = P(1 + \frac{r}{4})^4.$$

The earnings are $\Delta P = A_1 - P$ and the return rate is (since $t = 1$)

$$\begin{aligned} R &= \frac{A_1 - P}{P} = (1 + \frac{r}{4})^4 - 1 \\ &= 4(\frac{r}{4}) + 6(\frac{r}{4})^2 + 4(\frac{r}{4})^3 + (\frac{r}{4})^4. \end{aligned}$$

It is evident that $R > r$. For example, if r is 10 %, then R is 10.38 %. In other words, an annual rate of 10 % compounded quarterly earns the same as 10.38 % compounded annually. To distinguish between them, r is called the *nominal rate* and R is the *effective rate*.

Example 2.2. Two hundred dollars are deposited in an account paying 8 % per year compounded quarterly (2 % per quarter). After 3 and 1/4th quarters the account is closed. The interest earned in the first quarter is $200 * 0.02$ and this is added to the account making the value equal to $200(1 + 0.02)$. This repeats for the second quarter and third. The account value at that time is $200(1 + 0.02)^3$. Over the next 1/4 of a quarter the interest earned is this amount as principal times the interest rate for that fractional time, $200(1 + 0.02)^3(1/4)0.02$. This is 3 periods of compounding plus one-fourth period of simple interest. Hence the closing value is

$$200(1 + 0.02)^3(1 + (1/4)0.02) = 212.24(1 + 0.005) = 213.30.$$

Since $\left(1 + (\frac{1}{4})0.02\right) \neq (1 + 0.02)^{1/4}$, the exponent 3.25 does not give the exact answer.[4] This problem disappears under continuous compounding as we discuss next. □

As the number of compounding periods increases, so does the effective rate. If there are m compounding periods per year, the effective rate is given by

$$R = \left(1 + \frac{r}{m}\right)^m - 1. \tag{2.7}$$

The right-hand side is an increasing function of m, see Table 2.1.

Table 2.1 Effective rates at 10 % nominal for various compounding periods

Times/year	1	2	4	6	12	365	730
Rate	10.000	10.250	10.381	10.463	10.471	10.515	10.516

With m compounding periods per year, in t years there are $n = tm$ compounding periods. It follows that the value of the investment after t years is

$$A_t = P\left(1 + \frac{r}{m}\right)^{mt}. \tag{2.8}$$

As the number of compounding periods tends to infinity the expression on the right-hand side tends to a limit,

$$A_t = \lim_{m\to\infty} P\left(1 + \frac{r}{m}\right)^{mt} = Pe^{rt} \tag{2.9}$$

where $\boldsymbol{e}$ is a mathematical constant equal to 2.71828 accurate to 5 decimal places. This then is the accrued value for *continuous compounding* at a nominal annual rate of r. In this equation t does not have to be an exact number of years. Since compounding occurs continuously, t can be any non-negative real number.

From (2.6) the effective rate R satisfies the equation $1+R = e^r$, and therefore

$$R = e^r - 1 = r + \frac{r^2}{2!} + \frac{r^3}{3!} + \ldots;$$

[4] The exponent $3.25186\ldots = 3 + \log(1 + 0.02/4)/\log(1 + 0.02)$ is required.

the second equality follows from the power series expansion of the exponential function (A.3). For example, for a nominal rate of 10 %, the effective rate is 10.517 % under continuous compounding.

Fig. 2.1 shows in a dramatic way the effect of compounding. One unit of currency on deposit for 40 years at 8% earns about 4 times the original value under simple interest, but compounded continuous earns about 25 times its original value.

Doubling Time

An alternate way of characterizing the return rate is specifying the time required for an investment to double in value. Under continuous compounding we seek the time t_2 for which $Pe^{rt_2} = 2P$. Solving gives,

$$t_2 = \frac{\log 2}{r} \approx \frac{.7}{r}. \tag{2.10}$$

For example when r is 10 % it takes about 7 years for an investment to double.

This equation is the origin of the *Seven-Ten Rule*: Money invested at 7 % doubles in approximately 10 years and money invested at 10 % doubles in approximately 7 years.

For discrete compounding we may use (2.8),

$$t_2 = \frac{1}{m} \frac{\log(2)}{\log(1 + \frac{r}{m})}.$$

Recall m is the number of compounding periods per year. By using the effective rate R, m can be taken as 1, compare (2.7),

$$t_2 = \frac{\log(2)}{\log(1 + R)}. \tag{2.11}$$

Average Rates

As a rule return rates vary from time to time. Then the average rate becomes important. Suppose r_1 is the (simple) rate over the first compounding period, r_2 the rate over the second and so on through r_n, the nth. The amount of an investment P after this time is

$$A = P(1 + r_1)(1 + r_2) \dots (1 + r_n).$$

Therefore the average rate $\bar{r}$ over these n periods satisfies the equation $A = P(1 + \bar{r})^n$; it follows that the average is given by the nth root,

$$\bar{r} = \left((1 + r_1)(1 + r_2) \dots (1 + r_n)\right)^{1/n} - 1. \tag{2.12}$$

This is the *geometric average* of the several rates.

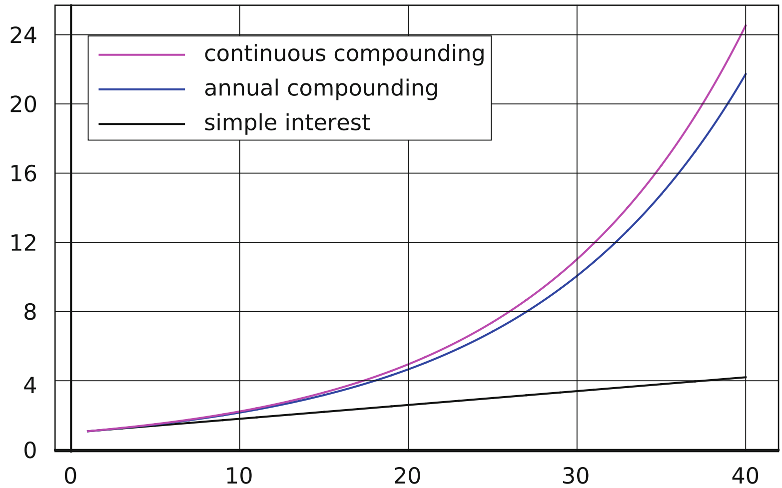


Fig. 2.1. A comparison of the growth of \$1 over 40 years at the annual rate of 8 %. Under simple interest the principal grows by about 4 times, under annual compounding, by about 22 times, and under continuous compounding by about 25 times

If compounding is continuous, the calculation is even simpler. Let r_1 be the rate for an arbitrary period of time t_1. That amount is then reinvested at the rate of r_2 over time t_2 and so on for n periods of time. Then the accrued amount is

$$A = Pe^{r_1 t_1} e^{r_2 t_2} \ldots e^{r_n t_n} = Pe^{r_1 t_1 + r_2 t_2 + \ldots + r_n t_n}.$$

Let $T = t_1 + t_2 + \ldots + t_n$ be the total time; since the average rate satisfies $A = Pe^{\bar{r}T}$, we get

$$\bar{r} = \frac{1}{T}\left(r_1 t_1 + r_2 t_2 + \ldots + r_n t_n\right), \tag{2.13}$$

the *arithmetic average* rate.

Example 2.3. What is the average rate of compounded return over four quarters if the first quarter rate is 0.01, the second is 0.02, the third is 0.03, and the fourth is 0.04?

Under discrete compounding the average quarterly rate is

$$\bar{r} = \sqrt[4]{(1+0.01)(1+0.02)(1+0.03)(1+0.04)} - 1 = 0.0249\ldots.$$

Under continuous compounding it is

$$\bar{r} = \frac{0.01+0.02+0.03+0.04}{4} = 0.025.$$

Note that the geometric average is less than the arithmetic average (slightly so here). This is always the case.[5] □

[5] Because the log function is concave down, for positive arguments, $\frac{1}{2}(\log x_1 + \log x_2) < \log(\frac{1}{2}(x_1 + x_2))$ unless $x_1 = x_2$. So $((1+x_1)(1+x_2))^{1/2} - 1 < \exp(\frac{1}{2}(\log(1+x_1) + \log(1+x_2))) - 1 < \exp(\log(\frac{1}{2}((1+x_1)+(1+x_2)))) - 1 = \frac{1}{2}(x_1+x_2)$; same argument for n terms.

Example 2.4. An amount P is put into savings certificates. The first year it earned 6 % interest and in the second 5 %. In the third year it was to have earned 8 % but the account was closed at mid-year (at no penalty). What was the average annual rate earned?

We want to find $\bar{r}$ solving the following equation

$$P(1.06)(1.05)(1+\frac{1}{2}0.08) = P(1+\bar{r})^2(1+\frac{1}{2}\bar{r}).$$

Numerical methods are required to discover that the answer is 6.0077 %. The approximate answer of 6.025 % can be gotten by approximating $(1+\bar{r}/2)$ by $(1+\bar{r})^{1/2}$.

If compounding were continuous the problem is much easier. The total time is $T = 1+1+0.5 = 2.5$, thus solving

$$Pe^{0.05+0.06+\frac{1}{2}0.08} = Pe^{\bar{r}2.5}$$

gives $\bar{r} = 6\,\%$.

This example shows why continuous compounding is often used in financial transactions. □

2.2.3 Discounting

Having P dollars today is worth more than having P dollars next week, or next month, or next year. For one thing, one could invest it at the risk-free rate r_f. Then in time t, those dollars will become

$$V = Pe^{r_f t}$$

in the case of continuous compounding. This demonstrates the time value of money.

Turning the argument around, it shows that P is today's value of a payment of V dollars at time t,

$$P = Ve^{-r_f t}. \tag{2.14}$$

This is called *discounting* future money to the present time. One must discount in this way when dealing with future payments. It then becomes possible to place transactions occurring at different times on an equal basis to compare them. We refer to P in (2.14) as the *present value* of V.

It is easier to discount under the assumption of continuous compounding as there is no need to interpolate but sometimes discrete compounding is called for. From (2.8) we get that

$$P = \frac{V}{\left(1+\frac{r}{m}\right)^{tm}} = V\left(1+\frac{r}{m}\right)^{-tm} \tag{2.15}$$

when t is an exact multiple of the compounding period $1/m$. It has the correct value when tm is an integer and interpolates for the other values of t,

Mortgages

As an example of a more intricate calculation with interest rates let us work through the problem of periodic mortgage payments. A loan of A dollars is to be paid back in n equal installments of Y dollars each. Assume the rate of interest is r per payment period (for example, if payments are monthly, then r is the annual rate divided by 12).

Let A_i be the remaining balance on the loan just after the ith payment, $A_0 = A$. At the end of the first period the balance has grown to $A(1+r)$; the payment reduces that by Y, hence the balance remaining is

$$A_1 = A(1+r) - Y.$$

For the second period, A_1 acts as the loan amount, and so

$$A_2 = A_1(1+r) - Y = A(1+r)^2 - Y(1+r) - Y.$$

Continuing, after n payments the balance is

$$\begin{aligned} A_n &= A(1+r)^n - Y\left((1+r)^{n-1} + \ldots + (1-r) + 1\right) \\ &= A(1+r)^n - Y\left(\frac{(1+r)^n - 1}{r}\right). \end{aligned} \tag{2.16}$$

But this final amount is zero, $A_n = 0$. Solving for Y we get

$$Y = \frac{Ar}{1 - \frac{1}{(1+r)^n}}$$

per period. The term

$$a = \frac{1-(1+r)^{-n}}{r}$$

is called the *annuity-immediate factor*; in this notation, $Y = A/a$.

The equation derived in (2.16) can be used to construct a table of remaining balances, see Table 2.2. These are of considerable interest to the homeowner.

Table 2.2 Remaining mortgage balance
\$200,000 at 6 % for 15 years, monthly payment: \$1,687.71

Month	Interest	Towards principal	Principal remaining
1	1,000	687.71	199,312.29
2	996.56	691.15	198,621.13
3	993.11	694.60	197,926.53
4	989.63	698.08	197,228.45
5	986.14	701.57	196,526.87
6	982.63	705.08	195,821.79

Example 2.5. Consider financing $A = \$200,000$ over 15 years ($n = 180$ months) at a rate of 6 % annually ($r = 1/2\,\%$ monthly). From above, this requires a monthly payment of

$$Y = \frac{200{,}000 * .005}{1 - \frac{1}{1.005^{180}}} = \$1,687.71.$$

□

Example 2.6. As a second example we will solve the same problem in an entirely different way using the principle of discounting. The lender receives a stream of payments each in the amount of Y. The present value of the first is $Y/(1+r)$. The present value of the second is, following (2.15), $Y/(1+r)^2$. Continuing in this fashion, the present value of all the payments equals the amount of the loan, and so

$$\begin{aligned} A &= \frac{Y}{1+r} + \frac{Y}{(1+r)^2} + \ldots + \frac{Y}{(1+r)^n} \\ &= \frac{Y}{(1+r)^n}\left((1+r)^{n-1} + \ldots + (1+r) + 1\right) \\ &= \frac{Y}{r}\left(1 - \frac{1}{(1+r)^n}\right). \end{aligned}$$

This gives the same solution as above. □

Annuities

An *annuity* is a series of payments to a beneficiary made at fixed intervals of time. If the number of payments is known in advanced, it is an *ordinary annuity*. A *perpetuity* is an annuity in which the payments continue forever.

The cost of an annuity can be derived by calculating its present value. Let Y denote the payments and r the per period discount rate. Upon reflection one sees that this problem is exactly like the mortgage calculation above. Indeed, from the lender's position, it is an annuity. The present value of the first payment made at the end of the first period is $Y/(1+r)$, the present value of the second payment is $Y/(1+r)^2$ and so on. Therefore for an ordinary annuity having n payments

$$PV = \sum_{k=1}^{n} \frac{Y}{(1+r)^k} = \frac{Y}{r}\left(1 - \frac{1}{(1+r)^n}\right). \tag{2.17}$$

And for a perpetuity

$$\begin{aligned} PV &= \sum_{k=1}^{n} \frac{Y}{(1+r)^k} = \frac{Y}{r}\left(1 - \frac{1}{(1+r)^n}\right) \\ &\to_{n\to\infty} \frac{Y}{r}. \end{aligned} \tag{2.18}$$

Yield

The annual rate of return of a bond over its lifetime is called its *yield to maturity* (YTM). Suppose a bond with a face value of F makes m coupon payments per year in the amount C/m and there are n payments remaining until its maturity. If the bond costs P, what is its YTM?

This can be calculated by discounting to the present time all the future payments of the bond at the yield rate. Denote this annual rate by r. Again invoking (2.15), the discount factor for the kth coupon payment is $(1+(r/m))^{-k}$. The discounted amount for each is C/m; over the lifetime there will be n such payments. It remains to discount the face value F. If there were an exact number of years y remaining it would be appropriate to use $(1+r)^{-y}$ as if compounding were annually. Or, since there are exactly n coupon periods remaining, one could use $(1+(r/m))^{-n}$ and thereby invoke per period compounding. The two are not the same, $(1+r)^{-n/m} \neq (1+(r/m))^{-n}$. The custom is to use the latter. And so we have

$$P = \frac{F}{(1+(r/m))^n} + \sum_{k=1}^{n} \frac{C/m}{(1+(r/m))^k};$$

upon summing the series we obtain

$$P = \frac{F}{(1+(r/m))^n} + \frac{C}{r}\left(1 - \frac{1}{(1+(r/m))^n}\right). \tag{2.19}$$

For given values of P, F, C, m, and n, this must be solved for r. Since r cannot be solved in closed form in (2.19), numerical methods have to be used. For example the bisection method discussed in Section A.14.

Example 2.7. Find the YTM of a 10 year \$10,000 bond with coupon payments of \$400 annually. The bond costs \$9,500.

We must solve

$$9{,}500 = \frac{10{,}000}{(1+r)^{10}} + \frac{400}{r}\left(1 - \frac{1}{(1+r)^{10}}\right).$$

The value $r = 0.01$ is too low and $r = 0.06$ is too high so they can be starting points for the bisection method. The method quickly gives $r = 0.04636$. □

2.3 Risk for a Single Investment

Let us suppose an investment of S_0 dollars is initiated by the purchase of stock in a particular company. To avoid certain complications, assume the company does not pay dividends. Let us also suppose that at some time in the future, $t = T$, we sell the stock (or at least assess whether we have made a profit or a loss to date).

If we assume the dynamics of the last chapter, then at the end of the investment period the value of the asset, S_T, will be lognormally distributed. We can calculate the investment's probability that $S_T < S_0$, the probability of losing

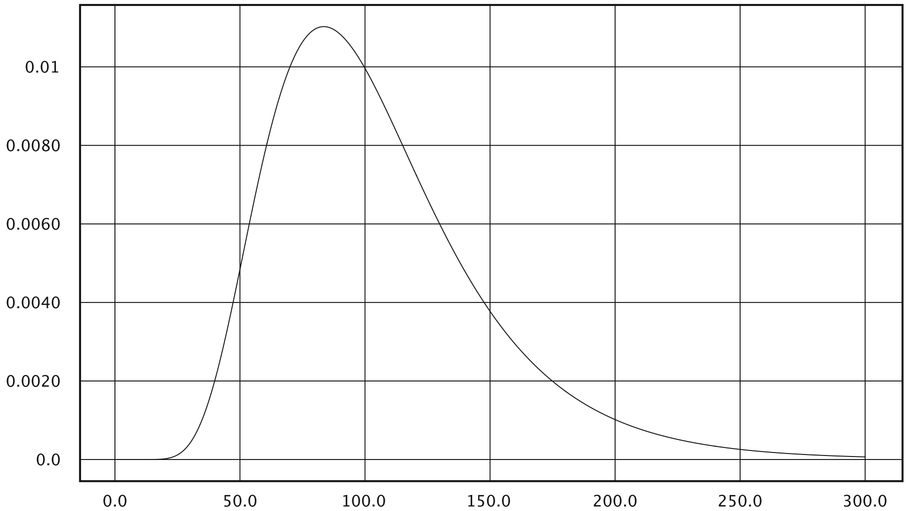


Fig. 2.2. Lognormal ending prices for: $S_0 = 100$, $\mu = 0.06$, $\sigma = 0.4$, $T = 1$ year

money, by integrating over the lognormal distribution from zero up to S_0, see Fig. 2.2.

But an easier method is available. Since S_T is lognormally distributed, $Y = \log(S_T)$ is normally distributed. From (1.23),

$$Z = \frac{1}{\sigma\sqrt{t}}\left(Y - \log(S_0) - (\mu - \frac{1}{2}\sigma^2)T\right)$$

is standard normal. It follows that for any $x > 0$,

$$\begin{aligned}
\Pr(S_T < x) &= \Pr(\log(S_T) < \log(x)) \\
&= \Pr\left(\frac{\log(S_T) - \log(S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}) < d\right) \\
&= \Pr(Z < d) \qquad (2.20)
\end{aligned}$$

where

$$d = \frac{\log(x) - \log(S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \qquad (2.21)$$

The probability can now be looked up in the cumulative normal table or, more conveniently, calculated from the cumulative normal rational approximation given in Section A.5 (or looked up online).

Example 2.8. Using the parameters as in Fig. 2.2: $S_0 = 100$, drift $\mu = 0.06$, volatility $\sigma = 0.4$, and $T = 1$ year and taking $x = S_0$ in (2.21), we get

$$d = \frac{\log(100) - \log(100) - (0.06 - \frac{1}{2}0.4^2)1}{0.4\sqrt{1}} = 0.05.$$

For this argument a cumulative normal table puts the loss probability at 52 %. □

In the example it is perhaps surprising that the probability exceeds 50 % even though the drift is positive. The explanation lies in the asymmetry of the lognormal distribution; recall equation (1.32) for the median of a lognormal. While the lognormal mean is calculated to be 106.18 according to (1.28), the probability mass extending far upscale must be compensated for by greater probabilities on the downside.

An added complication in figuring the probability of loss is accounting for the possibility that the company goes bankrupt and our investment devaluates to zero. Suppose this probability is B. From www.bloomberg.com[6] over the years 2000–2008 an average of 2.1 % of NYSE companies are *delisted* per year, so a value of $B = 0.02$ is a natural guess. The above calculation is now effected by scaling down the curve by $1 - B$, and then adding B to the result,

$$\begin{aligned}\Pr(S_T < S_0) &= \Pr(S_t < S_0 \,|\, \text{bankrupt})\Pr(\text{bankrupt}) \\ &\qquad + \Pr(S_T < S_0 \,|\, \text{not bankrupt})\Pr(\text{not bankrupt}) \\ &= B + (1 - B)\int_0^{S_0} f(s)\, ds. \end{aligned} \tag{2.22}$$

Here f is the lognormal density function with the appropriate parameter values. It is as though a histogram bar of probability B is placed at $S_T = 0$ and the rest of the histogram is scaled by $(1 - B)$.

Example 2.9. To include the probability of bankruptcy in our previous example make use of (2.22); we get

$$\Pr(\text{loss}) = 0.02 + 0.98 * 0.52 = 0.523. \tag{2.23}$$

□

2.3.1 The Solution by Simulation

We may also calculate the result through simulation. The algorithm is quite simple: run the geometric random walk algorithm on page 12 and count a hit if the end point $S_T < S_0$. Repeat this for a large number of times N (the results below are for $N = 90{,}000$) and use the number of hits divided by N as the estimate.

Example 2.10. Carrying out such a simulation with the parameters in Fig. 2.2 we get the risk of loss to be 52.2 % agreeing with the analytical calculation of Example 2.8. □

With such an algorithm in hand, we can pose and answer many relevant questions about our investment. How does the risk vary with volatility? with drift? with the time horizon? These involve making simple parameter changes in the algorithm and re-running the simulation.

[6] Search “NYSE Companies Delisted for Noncompliance” for this lengthy reference. See also www.moneycontrol.com/stocks/marketinfo/delisting/index.php

2.3.2 The Effect of Dividends

An important advantage of simulation is that more complicated situations can be easily accommodated. Such a complication is gaging the effect that issuing dividends has on the ending price distribution. The dividend payments can be fixed amounts at fixed times or amounts tied to the current stock price, or virtually any scheme.

When dividend payments are made, usually on a per share basis, the price of the stock immediately drops by the same amount. This is the result of the reduction in value of the company.

The *book value* of a company is the net worth of its tangible assets, for example the property it owns and its cash. (In particular, the talents and ideas of its employees are excluded.) Book per share is, theoretically, what a shareholder would get if the company liquidated and its proceeds were distributed to the shareholders.

While a stock's price can trade at several times its book per share value, often there is a close relationship between the two. When a company issues a dividend, its book value drops by the total amount dispersed. And so the book per share drops by that amount divided by the number of shares outstanding. This is exactly equal to a drop in price by the dividend per share.

There is a second reason why the share price must drop by the amount of the dividend. Suppose an investor buys the stock just prior to *ex-dividend day*[7] and thereby joins the rolls of those receiving a dividend. Later, maybe even the next day, the investor now sells the stock. Assuming the price does not fall by the dividend amount, this sale price will be about the same as the purchase price (possibly higher) thereby earning the investor the dividend as free money. Such a practice cannot last long. Many others will want to get in on it. The result will be to cause the stock price prior to going ex-dividend to escalate and then fall back afterwards.

To make a simulation of the problem treated above, but now with dividends, assume dividend payments are made quarterly (so the number of days between dividends is 91) with a yield of 8 % per year or 2 % per quarterly period. Assume dividends are not reinvested (otherwise there is no dividend essentially).

The modification to the algorithm consists in calculating the dividend amount on the last day of each quarter, adding this to the accumulated dividend payout, reducing the stock by the same amount and continuing. The accumulated dividends are assumed to be reinvested at the risk-free rate. At the end of the time frame, record a "hit" if the stock price plus accumulated dividends is less than S_0.

Algorithm 6. Ending Value with Dividends

```
inputs: S_0, nDays (T in days), μ, σ, daysBtwnDiv
        accumDiv, periodYield, periodRFR
dt=1/365  ▷1-day walk resolution
```

[7] The ex-dividend day and afterward is when a stock purchase does not qualify for the current dividend. Buying prior to ex-dividend day is required.

```
S = S_0  ▷initialize stock price
accumDiv = 0  ▷initialize accumulated dividends
j = 0   ▷initialize days since last dividend
for i = 1, ..., nDays
  Z ~ N(0,1)    ▷N(0,1) sample
  S = S * (1 + μΔt + σ√(Δt) Z)
  j = j + 1
  if( j == daysBtwnDiv )
      ▷grow accumulated dividends
    accumDiv = accumDiv*(1+periodRFR)
    divAmt = S*periodYield   ▷dividend this period
    accumDiv = accumDiv + divAmt
    S = S − divAmt  ▷decrease by amount of the dividend
    j = 0  ▷reset j
  endif
endfor
  ▷dividend growth over partial period
accumDiv = accumDiv*(1+periodRFR*j/daysBtwnDiv)
endValue = S + accumDiv
```

Example 2.11. Such a simulation applied to the problem of Example 2.8 with quarterly dividends at 8 % yield gives the result that the risk of losing money is the slightly reduced value 51.8 %. □

2.3.3 Stocks Follow the Market

In Chapter 1 we mentioned that a stock's price is subject to general market influences as well as the random walk fine structure. Let us now take that into account. Our goal is to calculate the risk for a particular stock under different market scenarios given the degree to which the stock follows the market. For this purpose we must learn how to generate market prices, possibly engineered to have specified characteristics, and how to generate individual stock prices in relation to the market.

Generating Market Prices

Start with the former. Of course we could generate market prices in the usual way using the algorithm on page 12. We could obtain markets with specified volatilities and drifts as desired in this way. But suppose we want more; for example we might want to specify trends, an improving market or a declining one or an oscillating one. In fact this is possible as we demonstrate by example.

Assume we want to simulate daily prices over 1 year but we want the market to have a given price each quarter, $(0, m_0)$, $(91, m_1)$, $(182, m_2)$, $(173, m_3)$ and $(364, m_4)$. We will see that in generating the stock prices, only the day to day market increments matter and therefore we take m_0 to be some convenient value such as $m_0 = 100$.

Now generate a GRW, $P[i]$, $i = 0, \ldots, 364$ with $P[0] = m_0$ and having the desired volatility, σ_M (and 0 drift for convenience).

Next let $\ell(\cdot)$ be the piecewise linear curve passing through the points $(0, m_0 - P[0])$, $(91, m_1 - P[91])$, $(182, m_2 - P[182])$, $(273, m_3 - P[273])$, and $(364, m_4 - P[364])$. The array of t-coordinates is

$$\text{tPts:} \quad 0, 91, 182, 273, 364$$

and the array of y-coordinates is

$$\text{yPts:} \quad m_0 - P[0], m_1 - P[91], m_2 - P[182], m_3 - P[273], m_4 - P[364].$$

Generically, the straight line through two points (a, A) and (b, B) is given by

$$y = \frac{1}{b-a}\Big(B(t-a) - A(t-b)\Big). \tag{2.24}$$

Letting $\mathbb{1}_{[a,b]}(t)$ denote the indicator function of the interval $[a, b]$ which is 1 for $a \le t \le b$ and 0 otherwise, put

$$\ell(t; a, b) = \frac{1}{b-a}\Big(B(t-a) - A(t-b)\Big)\mathbb{1}_{[a,b]}(t). \tag{2.25}$$

This is the line segment we want for t between a and b. Then $\ell(\cdot)$ is their sum

$$\ell(t) = \ell(t; 0, 91) + \ell(t; 91, 182) + \ell(t; 182, 273) + \ell(t; 273, 364). \tag{2.26}$$

Finally the array of market prices having the desired properties is just the sum

$$M[t] = P[t] + \ell(t), \qquad t = 0, 1, \ldots, 364.$$

Figure 2.3 portrays an example market scenario.

The function $\ell(\cdot)$ above can be implemented via the following code.

Algorithm 7. Generating piecewise linear ordinates

```
inputs: tPts, yPts
        t  ▷tPts[0] ≤ t ≤ tPts[4]
output: y = ℓ(t)
a = tPts[0]; A= yPts[0]
for i = 1,...,4
  b=tPts[i]; B=yPts[i]
  if( a ≤ t and t ≤ b )
    y=(B(t-a)-A(t-b))/(b-a)
    return y
  endif
  a=b; A=B
endfor
```

In the next section we will want to generate stock prices that correlate with the market. For this purpose we need the equivalent random increments W_i,

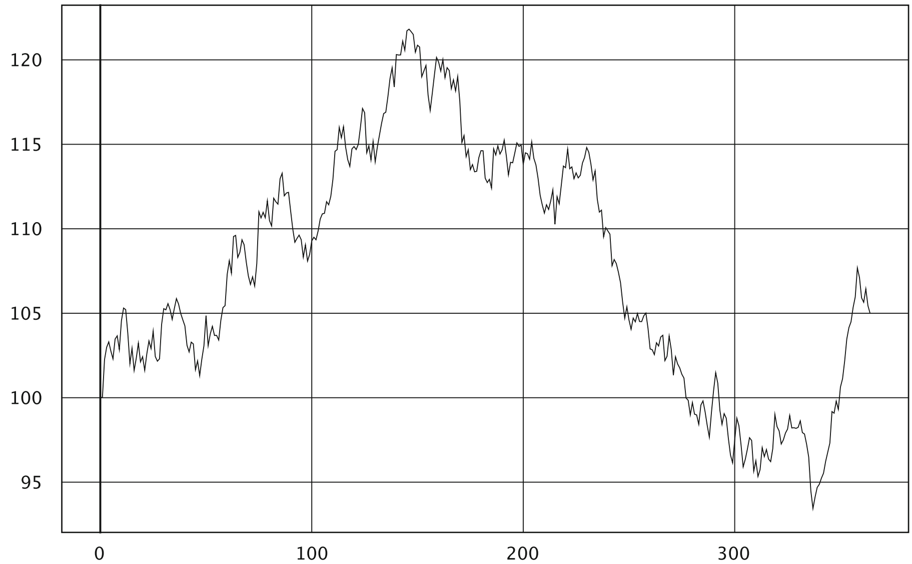


Fig. 2.3. Market prices passing through (0,100),(91,110), (182,113), (273,102) and (364,105). The volatility is 10 %

$i = 1, 2, \ldots$, from which the market could be regenerated.[8] Since $M_i = M_{i-1} + M_{i-1}\sigma_M\sqrt{dt}W_i$,

$$W_i = \frac{M_i - M_{i-1}}{M_{i-1}\sigma_M\sqrt{dt}}. \tag{2.27}$$

2.3.4 Correlated Stock Prices

Given market behavior, now we want to generate the prices of individual stocks that are influenced by the market. We consider stocks whose prices follow the market to an extent but not completely and not all the time. Specifically, let ρ, a number between -1 and 1, quantify the degree to which the stock's movement tracks the market's movement. If $\rho = 1$ then it tracks exactly in the sense of rising when the market rises and falls when the market falls. If $\rho = 1/2$ then it follows the general market about one half the time otherwise it moves independently of the market. If $\rho = 0$ then the stock moves independently all the time. A stock might even move contrary to the market, in this case ρ is negative.

Such a parameter is exemplified by the correlation coefficient defined next. First the *covariance* between two random variables X and Y is defined to be

$$\operatorname{covar}(X, Y) = \mathbb{E}((X - \mu_X)(Y - \mu_Y)), \tag{2.28}$$

where $\mu_X = \mathbb{E}(X)$ and $\mu_Y = \mathbb{E}(Y)$. The *correlation coefficient* is its normalization,

$$\rho_{XY} = \frac{\operatorname{covar}(X, Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}} = \frac{\operatorname{covar}(X, Y)}{\sigma_X\sigma_Y}. \tag{2.29}$$

[8] The calculated market path is a possible realization of a geometric random walk.

Notice from (2.28) that if X and Y both tend to be greater than their means at the same time and likewise lesser than their means at the same time, then their covariance will be a large positive value. It follows that ρ_{XY} will be near 1.[9] Conversely if Y tends to be below its mean when X is above and vice-versa, then their covariance will be a large negative value, and ρ_{XY} will be near -1. If X and Y are independent, then their correlation is 0, $\rho_{XY} = 0$.

The *covariance matrix* C for two random variables X and Y is defined as

$$\begin{aligned} C &= \begin{bmatrix} \mathrm{var}(X) & \mathrm{covar}(X,Y) \\ \mathrm{covar}(Y,X) & \mathrm{var}(Y) \end{bmatrix} \\ &= \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{YX}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}. \end{aligned} \tag{2.30}$$

We have used (2.29) to obtain the off-diagonal elements. Since $\rho_{YX} = \rho_{XY}$, the covariance matrix is symmetric meaning $C^T = C$ where superscript T designates matrix transpose.

Let $\mathbf{V}$ denote the 2×1 matrix, that is column vector, consisting of X and Y,

$$\mathbf{V} = \begin{bmatrix} X \\ Y \end{bmatrix}.$$

If $\mu_{\mathbf{V}}$ is the column vector of their means μ_X and μ_Y, and letting the expectation of a matrix mean the expectation of each of its elements, then the covariance matrix is given by

$$\begin{aligned} C &= \begin{bmatrix} \mathbb{E}((X-\mu_X)(X-\mu_X)) & \mathbb{E}((X-\mu_X)(Y-\mu_Y)) \\ \mathbb{E}((Y-\mu_Y)(X-\mu_X)) & \mathbb{E}((Y-\mu_Y)(Y-\mu_Y)) \end{bmatrix} \\ &= \mathbb{E}((\mathbf{V}-\mu_{\mathbf{V}})(\mathbf{V}-\mu_{\mathbf{V}})^T). \end{aligned} \tag{2.31}$$

Since a covariance matrix is symmetric[10] it has a *Cholesky decomposition*,

$$C = HH^T \tag{2.32}$$

where H is lower triangular. For example a 2×2 covariance matrix has the decomposition

$$\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \rho\sigma_2 \\ 0 & \sqrt{1-\rho^2}\sigma_2 \end{bmatrix}.$$

---

[9] By its definition, $-1 \le \rho_{XY} \le 1$. This follows from the well-known Cauchy-Schwarz inequality as indicated by the following. If x_i and y_i for $i = 1, \dots, n$ are empirical values of X and Y, then statistically

$$\begin{aligned} \mathrm{covar}(X,Y) &= \frac{1}{n}\sum_i (x_i-\mu_X)(y_i-\mu_Y) \\ &\le \sqrt{\frac{1}{n}\sum_i (x_i-\mu_X)^2}\sqrt{\frac{1}{n}\sum_i (y_i-\mu_Y)^2} = \sqrt{\mathrm{var}_X}\sqrt{\mathrm{var}_Y}. \end{aligned}$$

[10] It is also positive semi-definite but that is not needed for a Cholesky decomposition.

Now let Z and Z' be two independent, mean 0, unit variance random variables; their covariance matrix is therefore the identity matrix I. Put

$$\begin{bmatrix} X \\ Y \end{bmatrix} = H \begin{bmatrix} Z \\ Z' \end{bmatrix}, \tag{2.33}$$

then X and Y have covariance matrix C and therefore are correlated with coefficient ρ. The reason is that X and Y are mean 0 and, from (2.31),

$$\begin{aligned} \mathbb{E}(\begin{bmatrix} X \\ Y \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}^T) &= \mathbb{E}(H\begin{bmatrix} Z \\ Z' \end{bmatrix}(H\begin{bmatrix} Z \\ Z' \end{bmatrix})^T) \\ &= \mathbb{E}(H\begin{bmatrix} Z \\ Z' \end{bmatrix}\begin{bmatrix} Z \\ Z' \end{bmatrix}^T H^T) = H\mathbb{E}(\begin{bmatrix} Z \\ Z' \end{bmatrix}\begin{bmatrix} Z \\ Z' \end{bmatrix}^T)H^T \\ &= HIH^T = C. \end{aligned} \tag{2.34}$$

Since H is constant it can be moved outside the expectation operation.

Finally (2.33) shows how to construct correlated Gaussian random variables. Let Z and Z' be independent $N(0,1)$ random variables and put

$$\begin{aligned} X &= \sigma_1 Z \\ Y &= \rho\sigma_2 Z + \sqrt{1-\rho^2}\sigma_2 Z', \end{aligned} \tag{2.35}$$

then X and Y are correlated and normally distributed with variances σ_1 and σ_2 respectively and correlation coefficient ρ. Note that Z and Y are also correlated with coefficient ρ.

With correlated $N(0,1)$ samples in hand, to obtain correlated random walks, we simply generate them in the usual way using these samples. Let X_i and X_i', $i = 1, 2, \ldots$, be correlated $N(0,1)$ samples and set

$$\begin{aligned} S_i &= S_{i-1}(1 + \mu dt + \sigma\sqrt{dt}X_i) \\ S_i' &= S_{i-1}'(1 + \mu' dt + \sigma'\sqrt{dt}X_i'). \end{aligned}$$

Note that it is not the prices themselves that are correlated but rather the price increments. Correlating the increments is preferable because, for one thing, it is the day-to-day increments that are market correlated and, for a another, price increments are stationary in the sense of having constant means and variances, see [Mar78].

To obtain equity prices that follow the market we put the two constructions together. First construct a market scenario, M_i, engineered as desired using the techniques earlier in this section. Back out the equivalent price increments W_i, $i = 1, 2, \ldots$, given by (2.27). Finally, using a sequence of independent $N(0,1)$ samples Z_i, $i = 1, 2, \ldots$, put

$$Y_i = \rho W_i + \sqrt{1-\rho^2} Z_i$$

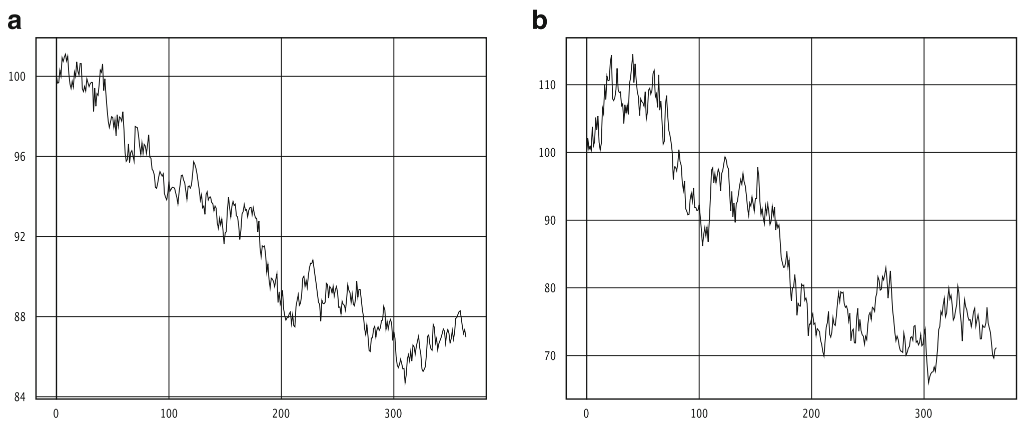
and then generate the prices as usual using the Y_i,

$$S_i = S_{i-1}(1 + \mu dt + \sigma\sqrt{dt}Y_i).$$

The technique is outlined in the following algorithm. Figure 2.4 shows an example run of the algorithm.

Algorithm 8. Stock prices correlated with the market

`inputs:` ρ, σ_m `(market volatility)`
μ_s, σ_s `(equity parameters)`
• `generate a market scenario:`
m_0, m_1, ... ▷`e.g. quarterly prices manually assigned`
▷`generate preliminary market prices`
$P_i = P_{i-1}(1 + \sigma_m \sqrt{dt} Z_i)$ ▷$Z_i \sim N(0,1)$
• `calculate the piecewise linear correction` $\ell(t)$ `(pp 48)`
• `calculate the market prices` $M_i = P_i + \ell(i)$
▷`back out the market increments`
$W_i = (M_i - M_{i-1})/(\sigma_m \sqrt{dt} M_{i-1})$
`label AA:`
▷`generate correlated` $N(0,1)$ `increments`
$Y_i = \rho W_i + \sqrt{1-\rho^2} Z_i$ ▷$Z_i \sim N(0,1)$
▷`generate the correlated stock prices using the` Y_i
$S_i = S_{i-1}(1 + \mu_s dt + \sigma_s \sqrt{dt} Y_i)$



A falling market trend

Correlated price changes, $\rho = 0.7$

Fig. 2.4. Example prices of a stock whose price movements are correlated with a market trend

It may become necessary to generate a large number of stock price histories all correlated with the same market; this is possible. In the algorithm, simply repeat starting from label AA, as many times as desired, to generate a new history.

Extension to More Variables

The technique for correlated samples given here extends to any number of random variables. For example to generate three pairwise correlated Gaussian random variables, let C be the desired covariance matrix and let HH^T be its Cholesky decomposition. If

$$C = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

Then

$$H = \begin{bmatrix} \sigma_1 & 0 & 0 \\ \rho_{12}\sigma_2 & \sqrt{1-\rho_{12}^2}\sigma_2 & 0 \\ \rho_{13}\sigma_3 & \sigma_3\frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} & \sqrt{\sigma_3^2 - h_{31}^2 - h_{32}^2} \end{bmatrix} \tag{2.36}$$

where h_{31} and h_{32} are the 31 and 32 elements of H,

$$h_{31} = \rho_{13}\sigma_3 \qquad h_{32} = \sigma_3\frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}.$$

If the h_{33} square root should result in an imaginary number, it means C is not positive semi-definite and therefore not a valid covariance matrix.[11]

Put

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = H \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix} \tag{2.37}$$

where Z_1, Z_2, and Z_3 are independent $N(0,1)$ Gaussians. Then X_1 and X_2 have correlation coefficient ρ_{12}, X_1 and X_3 have correlation coefficient ρ_{13} and X_2 and X_3 have correlation coefficient ρ_{23}.

2.4 Risk for Two Investments

Most portfolios consist of more than one investment. The interplay between several investments has a profound effect on risk. If the components of the portfolio are oppositely correlated, then the portfolio's prices tend to be more constant, avoiding large swings either way.

To see this, suppose an investment consists of equal positions in two stocks trading at about the same price. Suppose one is positively correlated and the other is negatively correlated with the market, see Fig. 2.5a and b. The day-to-day value of a 50–50 mix of these two stocks is shown in (c). Since the portfolio is the day-to-day average of the two, its value must necessarily lie halfway between them.

[11] An arbitrarily constructed real symmetric matrix is not necessarily positive-semi-definite. A little algebra shows that

$$h_{33}^2 = \frac{\sigma_3^2}{(1-\rho_{12}^2)}\left(1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + \rho_{12}\rho_{13}\rho_{23}\right).$$

This can be negative for the choices $\rho_{12} = \rho_{13} = -\rho_{23} = \frac{3}{4}$. Of course if 1 and 2 are highly correlated then 1 and 3 can't be highly correlated while 2 and 3 highly uncorrelated.

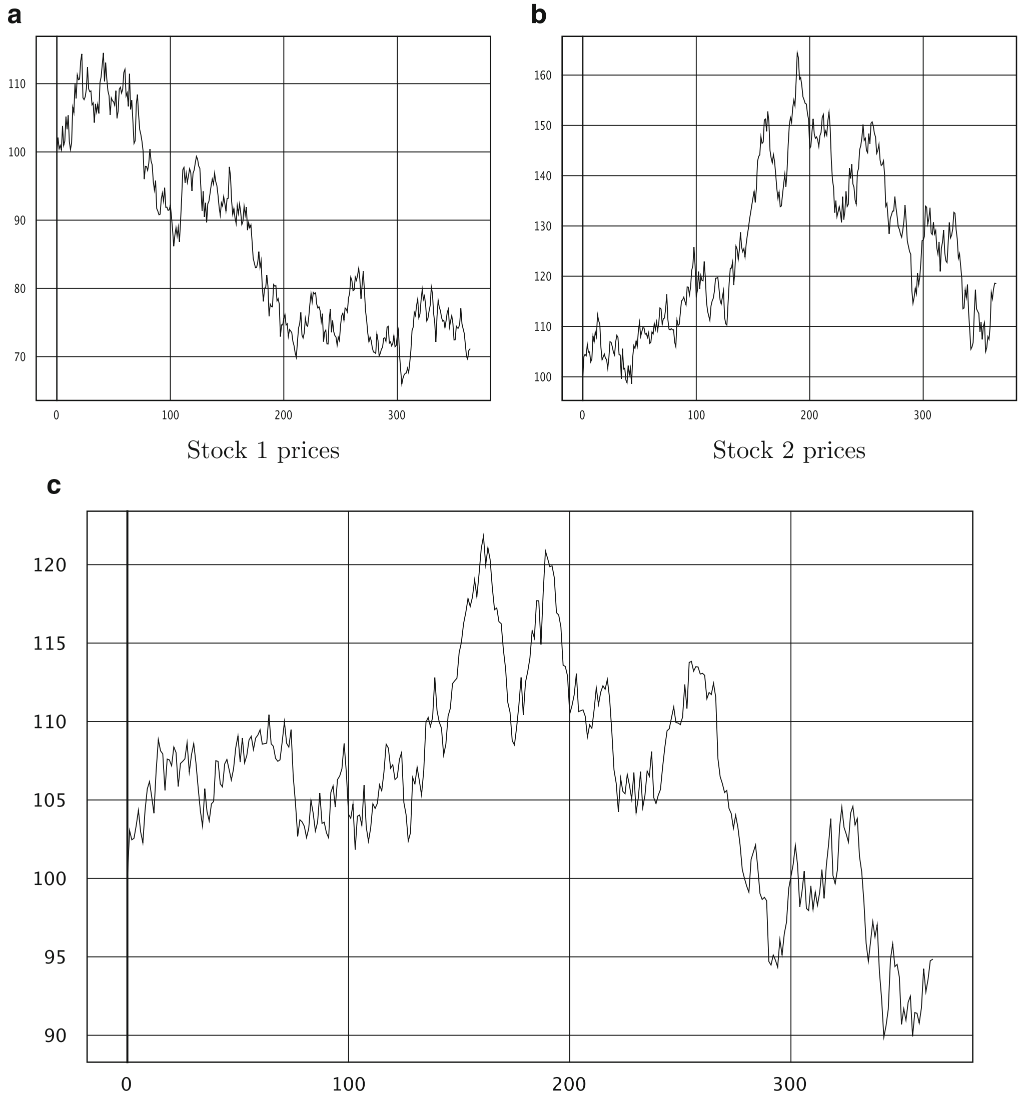


Fig. 2.5. The market for this example is that of Fig. 2.4a. The equity in (**a**) is positively correlated with the market while (**b**) is negatively correlated with the market. The individual stocks are subject to large swings but the portfolio remains between 90 and 125

The figure shows what can happen in one particular scenario of market and stock prices. To gauge the effect on the risk of such a portfolio, we must run a simulation as before over many such scenarios and count how often the result is a loss. Effectively we are integrating over the distribution of portfolio prices under the constraint of the assigned correlations.

Algorithmically the new simulation goes as follows.

Algorithm 9. Correlated Portfolio Risk

`for` $i = 1, 2, \ldots, N$
- `generate a random market scenario`

```
    • generate stock 1 prices with ρ = ρ1 (pp. 52)
    • generate stock 2 prices with ρ = ρ2
    • average to generate the portfolio prices
    • record a 'hit' if the S_T < S_0
endfor
• output (number of hits)/N
```

In this algorithm the market trend is random, therefore any difference between the risk predicted by simulations of this algorithm and that of the one-stock portfolio is due to the attributes of the portfolio.

Example 2.12. Consider the problem of Example 2.8 on page 44. Let a portfolio consist of two stocks having exactly the same financial parameters as in that example, $S_0 = 100$, $\mu = 0.06$, $\sigma = 0.4$, and $T = 1$. But let the first have correlation $\rho_1 = 0.7$ and the second have correlation $\rho_2 = -0.5$ with respect to the market. Let the market itself have drift 8 % and volatility 20 %. Then the risk of loss predicted by simulation is 41.9 %, an improvement of about 10 % from that of a stock by itself. Even if the two stocks are uncorrelated with the market or to each other there is still an improvement of about 5 %. This is due to the fact that the average of two or more values is less extreme than any of the values individually. □

Referring to Algorithm 9, shifting the market scenario generation outside the loop allows for testing portfolios under specific types of markets.

2.5 Value at Risk

The value at risk (VaR) is a measure that attempts to capture in a single number the total risk of a portfolio. The value at risk V is the maximum loss that can be expected with a given confidence over a specified period of time. For example one might assert "We are 99 % sure that over the next 30 days the portfolio will not lose more than \$10,000."

Of course the prediction is made on the basis of a model for stock prices, for example the GBM model. The prediction can also be made drawing on the pattern of historical prices for the portfolio. In this case, the model is that the economic conditions of the past and the underlying basis for stock price movement are projected to hold in the future.

For a portfolio consisting of a single stock, the GBM model predicts the future price will be lognormally distributed as in Fig. 2.2 which shows the ending price distribution after 1 year. As worked out in Section 2.3, the probability δ that the ending price will be less than $S_0 - V$ is given by the integral

$$\delta = \int_0^{S_0 - V} f(s)\, ds \tag{2.38}$$

where f is the lognormal density function with the appropriate parameter values. According to the model then, with probability $1 - \delta$ the stock price will not be less than $S_0 - V$.

Example 2.13. For a VaR confidence level of 99 %, take $\delta = 0.01$ and solve for $x = S_0 - V$. Using the parameters as shown in Fig. 2.2, from (2.20) we have

$$\begin{aligned} 0.01 = \Pr(S_T < x) &= \Pr(\log(S_T) < \log(x)) \\ &= \Pr\left(\frac{\log(S_T) - \log(S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} < d_T\right) \\ &= \Pr(Z < d_T) \end{aligned} \tag{2.39}$$

where

$$d_T = \frac{\log(x) - \log(S_0) - (\mu - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \tag{2.40}$$

Hence we want to find $d_T = \Phi^{-1}(0.01)$. From the cumulative normal table, or the rational interpolation of it, (A.9), find that $d_T = -2.3263$ and hence, from (2.40),

$$\log(x/S0) = -2.3263 * 0.4 + 0.06 - \frac{1}{2}(.4^2) = -0.95052;$$

consequently $x = 38.65$. This gives $V = 100 - 38.65 = 61.35$. Therefore with probability 99 %, the single stock portfolio is predicted to lose at most \$61.35 over the course of 1 year. (Keep in mind this is worst case (at the 99 % level); the portfolio might in fact gain in value over the year.) □

In more complex situations one can compute the VaR by Monte Carlo. We can see how it works by applying the technique to the single stock portfolio above. One simulates the price history a large number of times and notes the final price. A histogram of these produces the approximate price density but this is not what we want here.

Instead we want the cumulative price distribution. Recall that the cumulative distribution for an argument x is the integral of the density up to x. Statistically this means the sum of the number of prices that come in less than x (divided by the size of the sample). By sorting the ending prices low to high and plotting the sum of the number of sorted prices against price, the cumulative distribution is approximated, see Fig. 2.6.[12]

Having the sorted prices makes it easy to solve $\text{cdf}(x) = \delta$ for x given any δ. For example, for $\delta = 0.2 = 1/5$, the solution x is the sorted price one-fifth the way up the list, see Fig. 2.6. Let $S_T[i]$, $i = 1, \ldots, n$, be the sorted ending prices. Then the kth sorted price, $S_T[k]$, for

$$k = (\text{integer part})\delta n$$

gives the simulated solution $x = S_T[k]$, and in turn $V = S_0 - x$.

Example 2.14. For the problem in Example 2.13, a simulation gives $x = 38.87$ closely agreeing with the analytical value obtained there. □

The advantage of the Monte Carlo method is that it is easily implemented against any portfolio so long as there is a joint price model accounting for each constituent.

[12] A sorting subroutine is given in Appendix E.

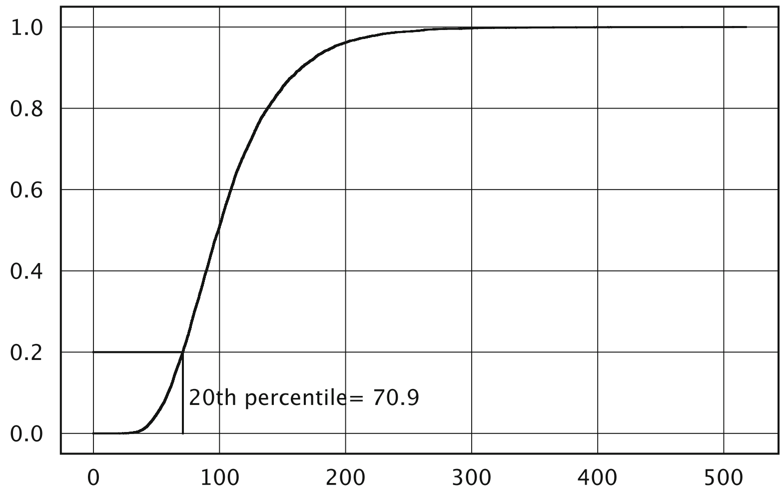


Fig. 2.6. Cumulative price distribution, $S_0 = 100$, drift$= 0.06$, $\sigma = 0.4$, $T = 1$ year. The calculation for the 20th percentile point is shown

Example 2.15. To apply the method to the two-stock portfolio in Example 2.12 of page 55 it is just a matter of sorting the price outcomes instead of recording hits. The sorted price for $\delta = 0.01$ is $x = 60.42$; hence the value at risk at the 99 % level is $V =$\$39.58; significantly less than the \$61.35 obtained above. This is due to the negative correlation among the constituents of the portfolio. □

2.5.1 Historical Simulation Method

In the technique of *historical simulation* one needs historical data for every market variable that affects the portfolio. For our simple portfolio of two equities, this means their price histories. For more complicated portfolios it could include interest rates, real estate values, exchange rates and so on.

The period of time in days, n, over which the VaR applies is called the *time horizon*. Ideally, we would like to have historical data covering several n-day periods. Usually there is not enough data for this purpose. Furthermore, the more time between the data observations and the present means the less likely it is that the economic conditions are the same. As a result, the n-day time horizon is estimated using 1-day data and the assumption that

$$n\text{-day VaR} = 1\text{-day Var} \times \sqrt{n}. \tag{2.41}$$

This assumption is correct for an arithmetical random walk (a Wiener process, see page 7) implying the data over disjoint time periods of equal length are independent and have identical normal distributions.[13]

[13] An extension formula for a portfolio consisting of a single GBM constituent can be derived from (2.40). But the day-to-day value of a portfolio consisting of several GBM constituents does not itself follow a GBM. Thus an accurate n-day VaR for such a portfolio requires a direct n-day simulation as above.

To see why, let the random variable Y_n be the portfolio value after n days starting from an initial value of Y_0. The assumption is that $Z_n = Y_n - Y_0$ is distributed as $N(0, \sigma^2 n)$. Further the 1-day VaR, V_1, is defined as $\Pr(Y_1 - Y_0 < -V_1) = \delta$ and the n-day VaR, V_n, is $\Pr(Y_n - Y_0 < -V_n) = \delta$. Since $Z_n = \sqrt{n} Z_1$,

$$\begin{aligned}\delta &= \Pr(Z_n < -V_n) = \Pr(\sqrt{n} Z_1 < -V_n) = \Pr(Z_1 < -\frac{V_n}{\sqrt{n}}) \\ &= \Pr(Z_1 < -V_1),\end{aligned}$$

provided $V_n = \sqrt{n} V_1$.

To illustrate the historical simulation method for the 360-day VaR, assume that Table 2.3 is the record of the stock prices for the two-equity portfolio over the last 360 days.[14] The historical record should include about the same number of days as the VaR to be predicted in order that the normally distributed data have approximately the same variation as expected over the VaR period.

Day 0 in the first column of the table is 360 days ago, day 360 of the table gives today's values. The relative change in the value of each market constituent is computed between each successive day; these values are calculated in columns 4 and 7 of the table. This provides 360 "experimental" observations for the change in value of each constituent.

Then, one-by-one, each observed percentage change is applied to today's constituent value to produce a possible future value for that constituent. This is shown in columns three and six of Table 2.4. Each such predicted constituent value is a possible future scenario and gives rise to a corresponding value of the portfolio, this is shown in column eight.

Table 2.3. Historical prices for the two-stock portfolio and their relative day-to-day changes

Day	c1	$\Delta c1$	$\Delta c1/c1$	c2	$\Delta c2$	$\Delta c2/c2$
364	116.80	2.225	0.01905	80.25	−1.827	−0.02276
363	114.57	−0.828	−0.00723	82.08	3.051	0.03717
362	115.40	−2.187	−0.01895	79.03	−0.259	−0.00328
361	117.59	0.944	0.00803	79.29	−4.031	−0.05084
360	116.64	−0.459	−0.00393	83.32	1.965	0.02358
⋮	⋮	⋮	⋮	⋮	⋮	⋮
4	103.96	1.948	0.01873	100.18	−0.340	−0.00339
3	102.02	−0.150	−0.00147	100.52	1.621	0.01612
2	102.17	−0.481	−0.00471	98.90	−0.412	−0.00416
1	102.65	2.648	0.02579	99.31	−0.687	−0.00692
0	100.00			100.00		

[14] We use 360 days here for illustrative and comparison purposes. In actual practice 252 "trading day" years is more likely to be used by company management. Further, the international Basel regulations specify the following VaR parameters: 10 day horizon, 99 % confidence level, and at least 1 year of historical data.

Then, as above, the 360 possible portfolio values are sorted low to high and the δ-th percentile noted. The 1-day VaR is calculated from this value and the 360-day VaR is calculated from (2.41).

Example 2.16. These particular tables were generated based upon the problem described in Example 2.15 above. From the tables we calculate that the 1-day VaR is \$3.05 and the 360-day VaR is \$58.16 by the historical method. The actual 1-day VaR from the simulation is \$2.65 which extends to a 360-day VaR of \$50.61.

Contrast these numbers with the direct 360 day simulation value of \$39.58 calculated above. □

2.6 Mean-Variance Portfolio Theory

The breakthrough that enabled mean-variance theory was the mathematical definition of risk and the attempt to deal with it through portfolio diversification. The treatment of risk is sufficiently precise that a rich theory may be developed, a theory that has proved to be useful in practice and remains the workhorse of analytical portfolio management.

Table 2.4. Three hundred and sixty 1-day price change scenarios; the third smallest portfolio value is the 1-percentile point

Scenario	Base c1	% change	c1	Base c2	% change	c2	Portfolio
1	116.80	1.905	119.02	80.25	−2.276	78.43	98.72
2	116.80	−0.723	115.95	80.25	3.717	83.24	99.59
3	116.80	−1.895	114.58	80.25	−0.328	79.99	97.29
4	116.80	0.803	117.73	80.25	−5.084	76.17	96.95
5	116.80	−0.393	116.34	80.25	2.358	82.15	99.24
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
361	116.80	1.873	118.98	80.25	−0.339	79.98	99.48
362	116.80	−0.147	116.62	80.25	1.612	81.55	99.09
363	116.80	−0.471	116.25	80.25	−0.416	79.92	98.08
364	116.80	2.579	119.81	80.25	−0.692	79.70	99.75

The assumptions of the mean-variance analysis are:

- A single period model
- At a given risk, investors prefer higher returns
- At a given return, investors prefer lower risk
- Markets are frictionless, meaning
  - Assets trade at any price and quantity
  - There are no transaction costs
  - There are no taxes

The single period model assumption means the investment is not dynamic, it does not adjust over time. All parameters of the model are fixed in advance (via estimates) and are applied as constants over the investment period. (The parameters being means, variances, and co-variances.) For example, dividends that occur over the investment period are only taken into account if incorporated into the parameters in advance. It also means that the investors preferences remain fixed over the investment period.

As mentioned, risk is central to the mean-variance analysis. The risk of the previous section is that of the actual loss of money. In his Ph.D. thesis of 1952, H. Markowitz introduced a definition of risk applicable to the *potential* for losing money. It also has the virtue of being mathematically quantifiable. For Markowitz, the risk of an investment is the variability of its returns; precisely, the standard deviation of its sequence of returns through time. (Often variance is used interchangeably with standard deviation in this context.)

A rationale for this definition stems from the fact that the greater the variance of an investment's return, the greater the uncertainty about future returns.

Using the techniques of the previous section, we can show that greater price variance aggravates the risk of actual loss as well. With parameters as in Fig. 2.2, we calculated the probability of loss to be 52 %, see page 45. If we now increase the volatility to 60 %, the probability of loss grows to 57.8 %.

2.6.1 A Two Scenario Example

The following simple example shows how effective reducing variability in a portfolio works to improve returns.

Consider a hypothetical situation in which the future value of an investment has two possible outcomes depending on which of two scenarios occur. The first scenario, ω_1, has probability 1/4 of occurring, and in this case the return on the investment will be 20 %. In the second scenario, ω_2, with probability 3/4, the return will be 5 %.[15] Under these conditions what is the risk of the investment and is it a good one?

To answer the first question we calculate the *mathematically expected* return. This is the sum of the possible outcomes each weighted by its probability. We get

$$\mu_A = \mathbb{E}(\text{return}) = \frac{1}{4}20 + \frac{3}{4}5 = 8.75\,\%.$$

The answer to the second question could depend on the other investment opportunities available. Suppose the money could be deposited in a bank account instead for a return of 8 %. The bank account is assumed safe and therefore has an expectation of 8 % as well.

From the standpoint of expected payoff the risky investment is better.

But what is the risk? As remarked above, we use the standard deviation (or variance) of the return to quantify it. For the bank account the variance of the return is 0. In the other it is

[15] Throughout this section we will measure returns in percent.

$$\text{var} = \frac{1}{4}(20 - 8.75)^2 + \frac{3}{4}(5 - 8.75)^2 = 42.1875$$

and the standard deviation is 6.5 approximately. (The standard deviation puts the value on the same numerical footing as the returns themselves.)

Risk Aversion

We have encountered a key element of investment science, *risk aversion*. Is it better to accept 8 % with certainty or take a chance on earning 20 % but with the prospect of having to settle for 5 % instead, even given that the expectation is favorable? The choice is personal and depends on one's level of willingness to gamble or not. The question of risk aversion will occur often in the sequel. In particular, if two investment returns have the same expectation, an investor is said to be *risk-neutral* upon being completely indifferent about the choice.

Aside from the question of risk aversion, a point to be made here is that the standard deviation of a return has merit as a measure of risk.

Example 2.17. It may seem that one should always choose the investment that has the biggest expected payoff. And this is a good choice if that opportunity presents itself over and over a large number of times. We will take up this topic in some detail in Chapter 7. But what if it presents itself just once?

Offered a one-time chance to win an amount of money equal in value to one's house or to lose the house altogether equally likely is not a bet most people would take.

As previously mentioned, risk-neutral means making decisions based on the best expected outcome and, if both have the same expectation, choose equally likely. For the home owner behaving in a risk neutral manner, either choice is just as good. □

Now consider another situation. Let the original investment choice, 20 % return with probability 1/4 and 5 % return with probability 3/4, be designated investment A. Suppose there is a second choice, investment B, with particulars: 2 % return under scenario 1 and 10 % return under scenario 2. These are spelled out in the following table.

Table 2.5 Two scenario risks and returns

	ω_1 1/4	ω_2 3/4	μ	Var	σ
Bank	8	8	8	0	0
A	20	5	8.75	42.18	6.50
B	2	10	8	12	3.46
I_{50-50}	11	7.5	8.375	2.3	1.52

Investment B has the same expectation, $\mu_B = 8$, as the bank account but its risk is larger and hence is inferior (as judged by the risk averse investor who measures risk by variance). Also, if an investor chooses A over the bank account, then B is likewise unattractive since B is already inferior to the bank account.

But what about a 50–50 mix of A and B? Under scenario 1 the return on such an investment is 11 % and in scenario 2 it is 7.5 %. The expectation under 50–50 can be computed as the scenario weighting of these 50–50 averaged per scenario returns,

$$\mu_{50-50} = \frac{1}{4}11 + \frac{3}{4}7.5 = 8.375.$$

For future reference we note here that it can also be calculated as the 50–50 weighted per individual returns

$$\mu_{50-50} = \frac{1}{2}8.75 + \frac{1}{2}8 = 8.375.$$

Likewise the variance can be computed as the scenario weighting of the 50–50 variances

$$\mathrm{var}_{50-50} = \frac{1}{4}(11 - 8.375)^2 + \frac{3}{4}(7.5 - 8.375)^2 = 2.2968.$$

We see that the 50–50 variance is unexpectedly low, only 2.3; much lower than either A or B alone

To see why, we calculate the variance by another method. In general, for random variables X and Y with means μ_X and μ_Y respectively, and weights α and β, we have

$$\begin{aligned}
\mathrm{var}(\alpha X + \beta Y) &= \mathbb{E}\left([(\alpha X + \beta Y) - (\alpha\mu_X + \beta\mu_Y)]^2\right) \\
&= \mathbb{E}\left([\alpha(X - \mu_X) + \beta(Y - \mu_Y)]^2\right) \\
&= \mathbb{E}\left(\alpha^2(X - \mu_X)^2 + 2\alpha\beta(X - \mu_X)(Y - \mu_Y) + \beta^2(Y - y)^2\right) \\
&= \alpha^2\mathrm{var}_X + 2\alpha\beta\mathbb{E}\left((X - \mu_X)(Y - \mu_Y)\right) + \beta^2\mathrm{var}_Y. \qquad (2.42)
\end{aligned}$$

The middle term (of the last line) is the *covariance* of X and Y we encountered in the previous section, see equation (2.28).

For these two investments the covariance is negative,

$$\mathrm{covar} = \frac{1}{4}(20 - 8.75)(2 - 8) + \frac{3}{4}(5 - 8.75)(10 - 8) = -22.5$$

because when one is greater than its mean, the other is less. These investments are *negatively correlated*. The negative covariance subtracts from the positive variances. From (2.42) with $\alpha = \beta = .5$, $X = A$, and $Y = B$, we have

$$\sigma^2_{50\text{-}50} = \left(\frac{1}{2}\right)^2(42.18) + 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)(-22.5) + \left(\frac{1}{2}\right)^2 12 = 2.30. \qquad (2.43)$$

The 50–50 investment is attractive because it has better expected return than the bank account but at the same time it has nearly zero risk (in terms of variance).

Moreover, a 50–50 split might not be the optimal split from the stand point of variance. Let w_A be the fraction of resources allocated to investment A and hence $w_B = 1 - w_A$ is allocated to B. Then as a function of w_A, return is given by

$$\mu_{w_A} = w_A \mu_A + (1 - w_A)\mu_B = 8.75 w_A + 8(1 - w_A),$$

and variance by

$$\begin{aligned} \sigma^2_{w_A} &= w_A^2 \sigma_A^2 + 2w_A(1 - w_A)\text{covar} + (1 - w_A)^2 \sigma_B^2 \\ &= 42.18 w_A^2 + 2w_A(1 - w_A)(-22.5) + 12(1 - w_A)^2. \end{aligned} \tag{2.44}$$

In Fig. 2.7 we plot return vs risk as a function of w_A. The figure encompasses all the (return, risk) pairs calculated above ($w_A = 0$ for B only, $w_A = 1$ for A only, and $w_A = 1/2$ for I_{50-50}). It also shows that for a certain value of w_A the variance can actually be brought to zero. We can find the minimum variance, be it zero or not, by differentiating the risk function with respect to w_A, setting the derivative to 0 and solving.

Example 2.18. Differentiating (2.44) gives

$$\begin{aligned} \frac{d\sigma^2_{w_A}}{dw_A} &= 2w_A \sigma_A^2 + 2\text{covar} - 4w_A \text{covar} - 2(1 - w_A)\sigma_B^2 \\ 0 &= (2\sigma_A^2 - 4\text{covar} + 2\sigma_B^2)w_A + 2\text{covar} - 2\sigma_B^2 \\ w_A &= \frac{\sigma_B^2 - \text{covar}}{\sigma_A^2 - 2\text{covar} + \sigma_B^2}. \end{aligned} \tag{2.45}$$

For $\sigma_A^2 = 42.18$, covar $= -22.5$, and $\sigma_B^2 = 12$, the minimum risk occurs for $w_A = 0.348$. As shown in the figure, the risk then is 0 for a return of 8.26 %. This is a strategy that is superior to the bank account. □

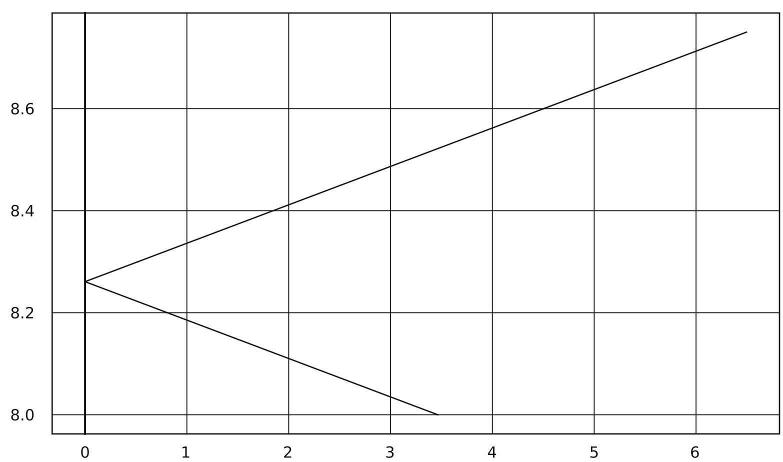


Fig. 2.7. Return vs. risk pairs (μ,σ) for two investments plotted as a function of w_A in the range $0 \le w_A \le 1$

Clearly investment B added an element of beneficial possibilities. Using Fig. 2.7, an investor can pick the allocation split satisfying his or her personal comfort of return versus risk.

In the next section we extend these ideas to portfolios of arbitrary size. But first we make the observation that two-investment portfolios generated as above are degenerate in a certain way: the product of the variances equals the covariance squared,

$$\sigma_A^2 \sigma_B^2 = \text{covar}^2. \tag{2.46}$$

Here

$$42.1875 * 12 = 22.5^2.$$

Since covar $= \sigma_A \sigma_B \rho$, it means that either $\rho = 1$ or $\rho = -1$. In this case it is the latter since the covariance is negative.

Substituting (2.46) into (2.44) yields

$$\begin{aligned}\sigma_{w_A}^2 &= w_A^2 \sigma_A^2 \pm 2 w_A (1 - w_A) \sigma_A \sigma_B + (1 - w_A)^2 \sigma_B^2 \\ &= \Big(w_A \sigma_A \pm (1 - w_A) \sigma_B \Big)^2. \end{aligned} \tag{2.47}$$

This explains why the (return, risk) plot is a straight line (broken at 0 if $\rho = -1$).

It is straightforward to show that (2.46) holds for any any assignment of returns in this two investment, two scenario example.

2.6.2 Portfolio Risk

Let V be the value of a portfolio B consisting of two equities with initial prices $S_1(0)$ and $S_2(0)$. The initial capital invested in the portfolio is $V(0)$. Let weights w_1 and w_2 be the allocation of capital to these equities respectively. The amount allocated to the first is $w_1 V(0)$ and the number of shares of this equity is

$$x_1 = \frac{w_1 V(0)}{S_1(0)}.$$

Similarly the number of shares of the second is

$$x_2 = \frac{w_2 V(0)}{S_2(0)}.$$

At any time t, the value of the portfolio depends on the prices of the equities at that time and is given by

$$V(t) = x_1 S_1(t) + x_2 S_2(t).$$

Note that the weights change as the equity prices change but the number of shares do not.

Next let K_1 and K_2 be the returns of the two securities at $t = 1$,

$$K_1 = \frac{S_1(1) - S_1(0)}{S_1(0)} \quad \text{and} \quad K_2 = \frac{S_2(1) - S_2(0)}{S_2(0)}.$$

The portfolio's return is

$$\begin{aligned}
K_B &= \frac{V(1)-V(0)}{V(0)} = \frac{x_1(S_1(1)-S_1(0))+x_2(S_2(1)-S_2(0))}{V(0)} \\
&= \frac{\frac{w_1V(0)}{S_1(0)}(S_1(1)-S_1(0)+\frac{w_2V(0)}{S_2(0)}(S_2(1)-S_2(0))}{V(0)} \\
&= w_1K_1+w_2K_2. \qquad (2.48)
\end{aligned}$$

Thus the portfolio return is linear with respect to the weights.

To calculate portfolio variance, let σ_1^2 and σ_2^2 be the variances of K_1 and K_2 respectively and let ρ_{12} be the correlation coefficient between them, see (2.29); then from (2.42),

$$\sigma_B^2 = w_1^2\sigma_1^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12} + w_2^2\sigma_2^2. \qquad (2.49)$$

since $\sigma_1\sigma_2\rho_{12} = \mathrm{covar}(K_1,K_2)$.

As above, the minimum risk for the portfolio is found by minimizing this equation under the constraint $w_1+w_2=1$. Put $s=w_2$, then $w_1=1-s$ and (2.49) becomes

$$\sigma_B^2 = (1-s)^2\sigma_1^2 + 2s(1-s)\sigma_1\sigma_2\rho_{12} + s^2\sigma_2^2. \qquad (2.50)$$

Differentiating with respect to s gives

$$\frac{d\sigma_B^2}{ds} = -2(1-s)\sigma_1^2 + 2(1-2s)\sigma_1\sigma_2\rho_{12} + 2s\sigma_2^2.$$

Setting the derivative to zero and solving for s we get

$$s_0 = \frac{\sigma_1^2-\sigma_1\sigma_2\rho_{12}}{\sigma_1^2-2\sigma_1\sigma_2\rho_{12}+\sigma_2^2} \qquad (2.51)$$

provided the denominator is not zero.

In fact, the minimum value the denominator can have is when $\rho_{12}=1$, then

$$\sigma_1^2-2\sigma_1\sigma_2\rho_{12}+\sigma_2^2 \ge \sigma_1^2-2\sigma_1\sigma_2+\sigma_2^2 = (\sigma_1-\sigma_2)^2.$$

Hence the denominator is either positive or zero. The latter occurring when $\rho_{12}=1$ and $\sigma_2=\sigma_1$. In this case, from (2.49), we have $\sigma_B=\sigma_1$. This case is not essentially different from that in which the two stocks are the same.

The zero derivative value of s given by (2.51) can be bigger than 1 or less than 0. To see this, divide the equation, numerator and denominator by $\sigma_1\sigma_2$ and let $r=\sigma_1/\sigma_2$, we get

$$s_0 = \frac{r-\rho_{12}}{(r+\frac{1}{r})-2\rho_{12}} = \frac{r-\rho_{12}}{r-\rho_{12}+(\frac{1}{r}-\rho_{12})}. \qquad (2.52)$$

Here we see that the denominator can be made arbitrarily small by choosing both r and ρ_{12} near 1. But the numerator is positive or negative depending on whether $r>\rho_{12}$ or $r<\rho_{12}$.

Example 2.19. Let $\sigma_1 = 1.04$, $\sigma_2 = 1$, and $\rho_{12} = 0.98$. Then $r = 1.04$ and, from (2.52),

$$s_0 = \frac{1.04 - 0.98}{1.04 - 0.98 + (0.9615 - 0.98)} = 1.44.$$

□

Recall that s is the weight w_2. Solutions for which s_0 is not between 0 and 1 correspond to going short, either in stock 2 if $s_0 < 0$ or stock 1 if $s_0 > 1$.

2.6.3 Efficient Frontier

The considerations of the previous section are a prototype of the general situation in which a portfolio consists of several stocks and the scenarios are the returns resulting from the infinity of possible price histories over the time horizon. While actual expected returns and variances can only be estimated, nevertheless the following theory, with its heavy emphasis on *diversification*, forms the bedrock of guiding principles for managing a portfolio.

Assume then a portfolio B of several equities, $i = 1, \ldots, n$, whose expected returns $\mu_i = \mathbb{E}(K_i)$ and return variances σ_i^2 and covariances covar_{ij} are known. Each point (σ_i, μ_i) may be plotted in the risk-return plane introduced in the previous section.[16]

For a given set of weights, w_i, $i = 1, \ldots, n$, a portfolio is constructed with w_i fraction of the total investment allocated to equity i. The portfolio return K_B is given by

$$K_B = w_1 K_1 + \ldots + w_n K_n = \sum_i w_i K_i.$$

It follows that the expected return is

$$\mu_B = \sum_i w_i \mu_i. \tag{2.53}$$

And the risk is given by

$$\begin{aligned}
\sigma_B^2 &= \mathbb{E}\Big((\sum_i w_i (K_i - \mu_i))^2 \Big) = \mathbb{E}\Big(\sum_i \sum_j w_i w_j (K_i - \mu_i)(K_j - \mu_j) \Big) \\
&= \sum_i w_i^2 \mathbb{E}((K_i - \mu_i)^2) + \sum_i \sum_{j \neq i} w_i w_j \mathbb{E}((K_i - \mu_i)(K_j - \mu_j)) \\
&= \sum_i w_i^2 \sigma_i^2 + 2 \sum_i \sum_{j > i} w_i w_j \mathrm{covar}_{ij}.
\end{aligned} \tag{2.54}$$

By letting C be the covariance matrix,

[16] In this section and the next we are, more exactly, plotting the rate of return versus risk. Since the time period is fixed, the two only differ by a constant factor. In the next section we will add the risk-free rate to the diagram.

$$C = \begin{bmatrix} \sigma_1^2 & \mathrm{covar}_{12} & \dots & \mathrm{covar}_{1n} \\ \mathrm{covar}_{21} & \sigma_2^2 & \dots & \mathrm{covar}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{covar}_{n1} & \mathrm{covar}_{n2} & \dots & \sigma_n^2 \end{bmatrix}$$

we can write

$$\sigma_B^2 = \mathbf{w} \cdot C\mathbf{w} \tag{2.55}$$

where $\mathbf{w}$ is the vector of weights and the dot means the dot product of the two vectors.[17]

When there are only two investments, the subset of the risk-return plane spanned by (σ_B, μ_B) over the set of all possible weights is a curve as in Fig. 2.7. (A parabola except in degenerate cases.) But here the subset spanned is an entire region. In Fig. 2.8 we show the risk-return region spanned by three investments. The three two-investment curves are visible as parabolas within or marking the edge of the region. It can be seen that the optimal mix, for instance point `MP`, is a mix of all three investments.

The risk-return region in this figure was obtained by a very simple Monte Carlo calculation as follows:

Algorithm 10. Calculating Risk-Return Points

```
inputs: N, μ1, μ2, μ3, σ1², σ2², σ3²
        covar12, covar13, covar23
for i = 1, ..., N
  sum=0
  for j = 1,2,3
    wj ~ U(0,1)  ▷wj is a uniform [0,1) sample
    sum = sum + wj
  endfor
  for j = 1,2,3
    wj = wj/sum  ▷the weights are now normalized
  endfor
  • compute μB by (2.53)
  • compute σB using (2.54)
  • plot
endfor
```

The risk-return calculation induces a partial order on the set of investment mixtures. An investment that has the same risk as another but greater return *dominates* the latter. That is, up is better.

Likewise, an investment that has the same return as another but less risk also dominates the other. So leftwards is better. More generally, if $\mu_A \geq \mu_B$ and $\sigma_A \leq \sigma_B$, then investment A dominates investment B. Therefore the boundary of the risk-return region from the vertex of the `abc-mno` parabola running up to `xyz` consists of undominated portfolios; all others are dominated by these.

---

[17] $\mathbf{x} \cdot \mathbf{y} = (\, x_1 \quad x_2 \quad \dots \quad x_n \,) \cdot (\, y_1 \quad y_2 \quad \dots \quad y_n \,) = \sum_1^n x_i y_i = \mathbf{x}^T \mathbf{y}$.

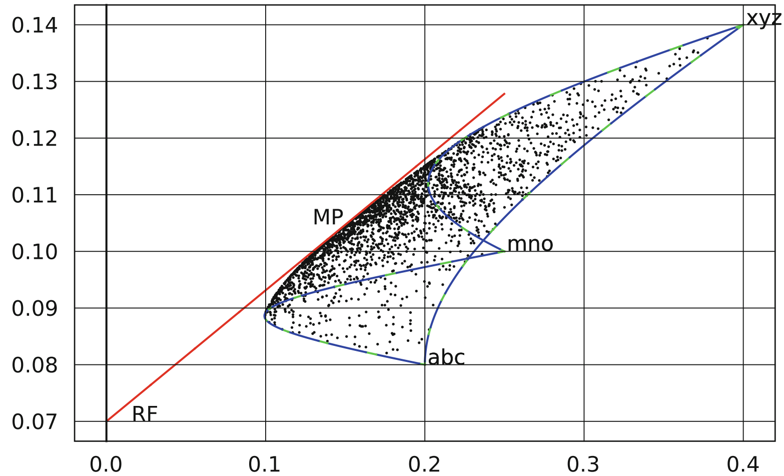


Fig. 2.8. Return vs. risk for three investments: $\mu_{abc} = 0.08$, $\sigma_{abc} = 0.2$, $\mu_{mno} = 0.10$, $\sigma_{mno} = 0.25$, $\mu_{xyz} = 0.14$, $\sigma_{xyz} = 0.4$, $\rho_{abc-mno} = -0.6$, $\rho_{abc-xyz} = 0.5$, $\rho_{mno-xyz} = -0.1$. Each *dot* is the (σ, μ) point for a mix of the three investments. The parabolas are the points for which one of the weights is zero

A portfolio is *efficient* if no other portfolio dominates it. A set of efficient portfolios among all attainable portfolios is called the *efficient frontier*. The efficient frontier of a set of individual investments is the upper left boundary of the risk-return region spanned by the set of all possible weights of those investments.

2.7 Capital Asset Pricing Model

Previously we assumed all investors are rational meaning they are risk averse and mean-variance optimizers. We now assume that all have the same information and therefore obtain the same estimates of returns, variances, and covariances. It then follows that all investors will have the same mix of stocks in the same proportion. This common portfolio is called the *market portfolio*. It is indicated as the point MP in Fig. 2.8.

The assumptions above are called the *equilibrium assumptions*. One of the implications of equilibrium is that the market portfolio consists of all the stocks in the market and in proportion to each stock's capitalization, that is, according to the total value of its shares relative to the total value of the entire market. This is because either all investors will own it or none will. If none own it, its price will be zero (or near zero). Then it will be undervalued, so now it is an attractive buy.

The next step in the CAPM development is to add the risk-free asset to the mix.

2.7.1 The Market Portfolio

Now allow portfolios to include a risk-free asset earning the risk-free return r_f. This asset appears in the risk-return plane at the point `RF` with coordinates $(0, r_f)$ in Fig. 2.8. Consider the line through $(0, r_f)$ and tangent to the efficient frontier of the risk-return region. This is called the *capital market line.* Let its point of tangency with the frontier be the point `MP` with coordinates (σ_M, μ_M). The mix of securities that gives this point on the efficient frontier is called the *market portfolio.* As mentioned, the market portfolio consists of all stocks in the market and in proportion to each stock's capitalization. In actual practice, various stock indexes such as the S&P-500 and the Russell 2000 try to approximate the market portfolio.

The slope of the line between `RF` and `MP` is

$$\frac{\mu_M - r_f}{\sigma_M} \tag{2.56}$$

and its intercept on the y-axis is r_f; therefore the capital market line has the equation

$$\mu = r_f + \frac{\mu_M - r_f}{\sigma_M}\sigma. \tag{2.57}$$

Example 2.20. Normally the capital market line must be determined numerically. But in the case of two investments simple calculus suffices.

Let A and B have mean returns μ_A and μ_B respectively, variances σ_A^2 and σ_B^2 and covariance covar. Let the risk-free rate be r_f and the market point be (σ_M, μ_M). The slope of the capital market line is given by (2.56). The plan is to equate this to the slope of the tangent to the risk-return curve. Let s play the role of w_B. The parameterization of the risk-return curve is, from (2.53)

$$\mu = (1-s)\mu_A + s\mu_B = \mu_A + (\mu_B - \mu_A)s, \tag{2.58}$$

hence

$$\frac{d\mu}{ds} = \mu_B - \mu_A.$$

And from (2.54)

$$\begin{aligned}\sigma^2 &= (1-s)^2\sigma_A^2 + 2s(1-s)\text{covar} + s^2\sigma_B^2 \\ &= \theta s^2 - 2\lambda s + \sigma_A^2\end{aligned} \tag{2.59}$$

where $\theta = \sigma_A^2 - 2\text{covar} + \sigma_B^2$ and $\lambda = \sigma_A^2 - \text{covar}$. Differentiate both sides and substitute $\sigma = \sigma_M$ at the point of tangency

$$\begin{aligned}2\sigma\frac{d\sigma}{ds} =& 2\theta s - 2\lambda \\ \frac{d\sigma}{ds} =& \sigma_M^{-1}(\theta s - \lambda).\end{aligned}$$

Therefore the slope of the tangent line is

$$\frac{d\mu}{d\sigma} = \frac{d\mu/ds}{d\sigma/ds} = \frac{(\mu_B - \mu_A)\sigma_M}{\theta s - \lambda}.$$

Equating slopes gives

$$\frac{\mu_M - r_f}{\sigma_M} = \frac{(\mu_B - \mu_A)\sigma_M}{\theta s - \lambda}. \tag{2.60}$$

Substituting for μ_M, (2.58), and σ_M, (2.59), gives

$$(\theta s - \lambda)\left((\mu_A - r_f) + (\mu_B - \mu_A)s\right) = (\mu_B - \mu_A)\left(\sigma_A^2 + (\theta s - \lambda)s - \lambda s\right)$$

The quadratic terms in s cancel; the resulting linear equation is solved for s to give

$$s = \frac{(\mu_B - \mu_A)\sigma_A^2 + (\mu_A - r_f)(\sigma_A^2 - \text{covar})}{(\mu_A - r_f)(\sigma_A^2 - 2\text{covar} + \sigma_B^2) + (\mu_B - \mu_A)(\sigma_A^2 - \text{covar})} \tag{2.61}$$

for the location of the market point.

As a numerical example, let the two-investment pair above be the equities mno and xyz in Fig. 2.8. If the portfolio consisted of these two only and the risk-free rate were 8 %, then, from the parameters given in the figure, the covariance is

$$(0.25)(0.4)(-0.1) = 0.01$$

and from (2.61)

$$\begin{aligned} s &= \frac{(0.14 - 0.1)0.25^2 + (0.1 - 0.08)(0.25^2 + 0.01)}{(0.1 - 0.08)(0.25^2 + 0.02 + 0.4^2) + (0.14 - 0.1)(0.25^2 + 0.01)} \\ &= 0.51. \end{aligned}$$

Therefore $\mu_M = 0.49(0.1) + 0.51(0.14) = 0.12$ and

$$\sigma_M = \sqrt{0.49^2 0.25^2 + 0.02(0.49)(0.51) + 0.51^2 0.4^2} = 0.23.$$

□

Since the capital market line is above and to the left of the risk-return region spanned by the risky securities, it becomes the new efficient frontier. It follows that rational investors will select a position along this line according to their personal level of risk tolerance. Hence all rational investors will have the same mix of risky securities, namely the market portfolio, differing only in their proportion as allocated between the market portfolio and a risk-free investment.

But investors may also go short. By doing so, their position on the capital market line need not be constrained between `RF` and `MP`. By going short on the risk-free asset and transferring the capital to the market portfolio, an investor's market weight will be greater than 1. The return for such a mix will exceed μ_M but at the same time, the risk will exceed σ_M.

The slope of the capital market line (2.56) is a very important investment parameter. It gives the rate at which one's level of return rises for taking on increments in the level of risk. It is called the *price of risk* or the *risk premium*.

2.7.2 Beta Factor of a Portfolio

A major result of CAPM, and perhaps a surprising one, is that the expected returns of any particular stock bears a simple relationship to that of the market portfolio.

Theorem (Capital Assets Pricing Theorem) *The expected rate of return of a portfolio, μ_B, (or an individual stock considered as a portfolio of one item) is given by*

$$\mu_B = r_f + \beta(\mu_M - r_f) \tag{2.62}$$

where μ_M is the market portfolio rate of return and β is given by

$$\beta = \frac{covar(B, M)}{\sigma_M^2}. \tag{2.63}$$

Here $\text{covar}(B, M)$ is the covariance of the portfolios sequence of periodic returns versus those of the market portfolio.

Equation (2.62) is known as the *security market line.* We will have more say about it in the next section. Beta as given by (2.63) is the *beta factor* of the portfolio or of an individual stock; it is unique to each portfolio.

The difference $\mu_B - r_f$ in (2.62) is the *excess rate of return* of the portfolio above the risk-free rate. The theorem says that it is proportional to the excess rate of return of the market portfolio itself with proportionality factor equal to β. Moreover, since β is directly proportional to its covariance with the market portfolio, the theorem says the excess rate of return of a portfolio is proportional to its covariance with the market.

This last statement may sound surprising with respect to portfolios that are uncorrelated to the market and thus whose covariance is zero. But for a large portfolio of equities each uncorrelated with the market, and each other, their combined variance will be small and therefore their combined rate of return will be the risk-free rate.

The proof of (2.62) follows along the lines of the two investment example above, Example 2.20. Let asset A of that example be the market portfolio, M. From (2.58) and (2.59), the risk-return curve for pair M and B is

$$\begin{aligned}
\mu &= s\mu_B + (1-s)\mu_M \\
\sigma^2 &= (1-s)^2\sigma_M^2 + 2s(1-s)\text{covar} + s^2\sigma_B^2 \\
&= \theta s^2 - 2\lambda s + \sigma_M^2.
\end{aligned}$$

Note that for $s = 0$ the two investments degenerate into just the market portfolio. Further, it must be that the risk-return curve for M,B is tangent to the efficient frontier at that point since it cannot cross the efficient frontier. Therefore its slope at $s = 0$ equals $(\mu_M - r_f)/\sigma_M$. Hence we have, using (2.60),

$$\begin{aligned}
\frac{\mu_M - r_f}{\sigma_M} = \frac{d\mu}{d\sigma} &= \left.\frac{(\mu_B - \mu_M)\sigma_M}{\theta s - \lambda}\right|_{s=0} \\
&= \frac{(\mu_B - \mu_M)\sigma_M}{\text{covar} - \sigma_M^2}.
\end{aligned}$$

Solving for μ_B we get

$$\begin{aligned}\mu_B &= \mu_M + \frac{(\mu_M - r_f)(\text{covar} - \sigma_M^2)}{\sigma_M^2} \\ &= \mu_M + \frac{\text{covar}}{\sigma_M^2}(\mu_M - r_f) - (\mu_M - r_f) \\ &= r_f + \beta(\mu_M - r_f),\end{aligned}$$

where β is as in (2.63).

Beta and the Line of Best Fit

In Fig. 2.9 we plot the monthly returns of two securities versus the S&P-500 over a 1 year period. Let $x_1, x_2, \ldots, x_n$ be the sequence of S&P-500 returns and $y_1, y_2, \ldots, y_n$ be those for the security. Each graph plots the pairs (x_i, y_i) for $i = 1, 2, \ldots, n$; here $n = 12$.

We wish to calculate the straight line, $y = mx + b$, that best fits these data. This line is shown superimposed on each graph.

By the method of least-squares, the well-known equations for m and b are derived in appendix Section A.7. From (A.13) we find that

$$\begin{aligned}m &= \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} \\ &= \frac{\frac{1}{n}\sum x_i y_i - \left(\frac{1}{n}\sum x_i\right)\left(\frac{1}{n}\sum y_i\right)}{\frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2} \\ &= \frac{\text{covar}(x,y)}{\text{var}(x)}.\end{aligned} \tag{2.64}$$

Thus the slope of the best fit line is β

Example 2.21. The data points for MSP in Fig. 2.9 are the following

$$(-1.8, -1.8),\ (-.9, -.25),\ (-.7, -.45),\ (.09, .44),\ (.08, .5),\ (.2, .45),$$
$$(.3, .2),\ (.6, .35),\ (.68, .45),\ (.65, .7),\ (.9, .6),\ (1.35, 1.52).$$

First calculate the means:

$$\begin{aligned}\bar{x} &= \frac{1}{12}(-1.8 + -0.9 + -0.7 + \ldots + 0.9 + 1.35) = 0.12 \\ \bar{y} &= \frac{1}{12}(-1.8 + -0.25 + -0.45 + \ldots + 0.6 + 1.52) = 0.23.\end{aligned}$$

Then the variance and covariance

$$\begin{aligned}\text{var}(x) &= \frac{1}{12}\left(-1.8^2 + -0.9^2 + \ldots + 1.35^2\right) - 0.12^2 = 0.6989 \\ \text{covar}(x,y) &= \frac{1}{12}\Big((-1.8)(-1.8) + (-.9)(-.25) + \\ &\qquad \ldots + (1.35)(1.52)\Big) - (0.12)(0.23) = 0.6038.\end{aligned}$$

Finally $\beta = 0.6038/0.6989 = 0.86$. □

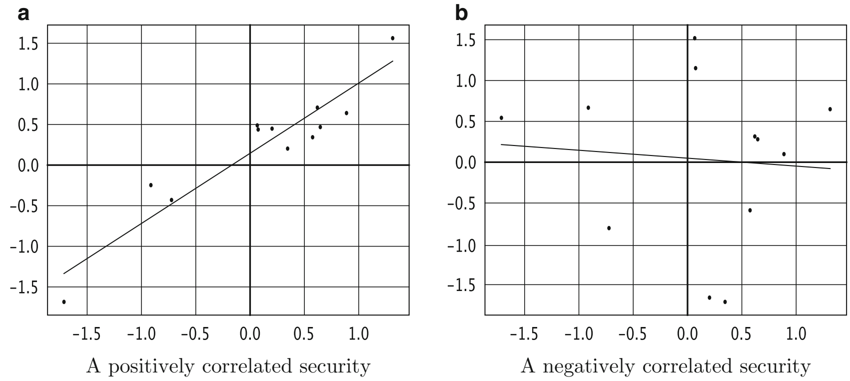


Fig. 2.9. A plot of monthly returns $\Delta S/S$ over 1 year for two securities versus those of the S&P-500 – a proxy for the entire market; MSP in (**a**) and GG in (**b**). The *straight lines* are the best least squares fit and the slope is beta in each case. Using the equations derived in the text, MSP has a beta of 0.86 and GG a beta of -0.10

2.7.3 The Security Market Line

Let $y_1, y_2, \ldots, y_n$ be a sequence of returns for an asset B and $x_1, x_2, \ldots, x_n$ those for the market portfolio. For each i let ϵ_i be the difference between the empirical return y_i and that predicted by the security market line, equation (2.62), so we can write

$$y_i = r_f + \beta(x_i - r_f) + \epsilon_i.$$

By the CAPM theorem, the expectation $\mathbb{E}(\epsilon_i) = 0$. So is the covariance,

$$\begin{aligned}\operatorname{covar}(\epsilon_i, x_i) &= \operatorname{covar}(y_i - r_f - \beta x_i + \beta r_f, x_i)\\ &= \operatorname{covar}(y_i, x_i) - \beta \operatorname{covar}(x_i, x_i)\\ &= \operatorname{covar}(y_i, x_i) - \frac{\operatorname{covar}(y_i, x_i)}{\operatorname{var}(x_i)} \operatorname{var}(x_i) = 0.\end{aligned}$$

It follows that the variance of the y_i is given by

$$\begin{aligned}\operatorname{var}(y_i) &= \operatorname{var}(r_f + \beta x_i - \beta r_f + \epsilon_i)\\ &= \beta^2 \operatorname{var}(x_i) + \operatorname{var}(\epsilon_i).\end{aligned}$$

We obtain the important relationship

$$\sigma_B^2 = \beta^2 \sigma_M^2 + \operatorname{var}(\epsilon_i). \tag{2.65}$$

Equation (2.65) shows that the risk in a portfolio has two sources. The first, $\beta^2 \sigma_M^2$, is unavoidable and is called the *systemic risk*. This risk is that of the market as a whole. This is the risk alluded to in the first chapter due to macro economic shocks arising from, for example, government policy, international economic forces, acts of nature. It cannot be diversified away.

The second source of risk is that specific to the portfolio itself. It is called *specific risk* or *diversifiable risk*. By adding more and more securities to the portfolio, this risk can be reduced to zero in the limit as the portfolio tends to the market portfolio.

Problems: Chapter 2

1. What should be the price of a 2 year \$100 zero coupon bond in order that the investment earns 6 % per year? (No compounding.)

2. An annuity starts with \$501,692 and pays out \$10,000 per month. If the remaining principle earns 4 % annual interest compounded continuously, for how many months will the annuity pay?
(Answer 55.)

3. Algorithm 6 assumes that the equity is purchased at the beginning of the dividend period. If the stock is held for an exact multiple of the dividend period, then the annualized return should be exactly the dividend yield. But what if the stock is bought or sold at mid-term intervals? For example shortly before ex-dividend day? Explore the annual return under various ownership periods with respect to the ex-dividend date.

4. Write a program to display a piecewise linear approximation of a price path as follows. Let S_0, S_1, S_2, …, S_{365} be a 1 year sequence of prices. Select a subset of these, for example monthly S_0, S_{30}, …, S_{364}, and generate the piecewise linear graph through these points, $(0, S_0)$, $(30, S_{30})$, …, $(364, S_{364})$. Such an approximation could serve as an alternative to the moving average of the prices.

5. Use Algorithm 8 to construct several figures such as Fig. 2.4 and observe the extent to which correlated prices trend together (recall that it is the increments that are correlated, not the prices themselves). For each run, calculate the correlation between the prices.

6. Investigate the probability of losing money for the investment of Example 2.8 (pp. 44) if the stock is correlated with the market and, variously, $\rho = 0.8$, $\rho = 0$, $\rho = -0.5$. Do this for various market scenarios: a rising market, a falling market, a sideways market.

7. Run Algorithm 9 (pp. 54) with various correlations between the two stocks and the market. What is the risk of loss when: (a) $\rho_1 = 1$, $\rho_2 = 1$?, (b) $\rho_1 = 1$, $\rho_2 = -1$?, (c) $\rho_1 = 0$, $\rho_2 = 0$?, (d) $\rho_1 = -1$, $\rho_2 = -1$?, (e) $\rho_1 = 0.6$, $\rho_2 = -0.6$?

8. Calculate the VaR at the 99 % level over, variously, 1 month, 3 months, and 6 months, for a stock whose initial price is \$45, whose drift is 2 %, and whose volatility is 23 %.

9. Find the VaR at the 99 % level over 2 months by simulation for a portfolio of two stocks with parameters: for the first: $S_0 = 20$, $\mu = 3\,\%$, volatility= 26 %, for the second: $S_0 = 40$, $\mu = 1\,\%$, volatility= 33 %. Assume that the stocks are correlated, variously, $\rho = 0.9$, $\rho = 0.2$, $\rho = -0.8$.

10. Find the VaR for the stocks in Problem 9 by the historical method. For their price histories, use the GBM model to generate 2 months worth of prices for the equities. Only treat the $\rho = 0.2$ case.

11. Investigate how the probability of loss in Example 2.8 (pp. 44) varies as a function of volatility. Make a graph of loss vs. volatility.

12. An investor has a choice between two ventures A and B. As the investor sees it, the future holds three possibilities: (bull) A returns 12 %, B returns 3 %, (bear) A return −4 %, B returns 4 %, or (static) A returns 6 %, B returns 0 %. Assume the probabilities are: bull 0.2, bear 0.3, static 0.5. What are the expected returns and risks (standard deviations) for each venture? Same question for a 50–50 allocation of the investor's resources. What is the allocation giving least risk?
(Answer 0.1788 : 0.8212.)

13. If the risk-free rate is 3 % in Problem 12, what is the market point?
(Answer ($\mu = 1.9411, \sigma = 1.914$).)

14. Obtain recent price data for some security from among the list: AAPL, MON, KO, F, MCD, FDX. Along with data for the S&P-500, use it to calculate daily returns over the last month and to calculate beta for the stock.

15. Use the results of Problem 14 to calculate the risk premium for that stock.

3

Forward and Option Contracts and Their Pricing

The ultimate tool for coping with risk is what we call today an option. The first to realize this was Bachelier. In his thesis of 1900 he introduced options for just this purpose. To see how it works, suppose a bank holds a few thousand shares in a security trading for $50 per share at the present time. The bank's plan is to sell the stock in 3 months but it must receive at least $45 per share at that time. To deal with its risk, the bank enters into a contract with a second party that agrees to buy its shares at $45 in 3 months no matter what the market price is then. As a further stipulation of the contract, the bank is not obligated to sell at $45; if the market price in 3 months happens to exceed this, the bank is free to sell at the market price.

The bank has entered into a *put option* contract (or just a *put*) with the shares of stock serving as the basis or *underlying* of the contract.

By means of it the bank has completely laid off its risk. Of course the party underwriting the contract charges for its service. The main topic of this chapter is deciding what the fair price of such a contract should be. This price is the exact quantification of the risk. The first to solve the problem were Fischer Black and Myron Scholes in their paper of 1973 and we will review their solution. First we look more closely at option contracts.

In another example, suppose a cookie manufacturer will have to buy a few thousand bushels of wheat after the fall harvest 6 months from now. Currently the price per bushel for fall wheat is $8.41 but many factors could intervene and escalate the price. The company would like to guarantee it pay no more than $9.50. It can do so by entering into a *call option* contract (or just a *call*). A second party agrees to sell wheat at $9.50 no matter the market price. Again, the option does not require the cookie company to buy at 9.50 if it can do better in the market. And, as above, the underwriter of the contract charges for its service.

In this example the underlying is a commodity. Option contracts have grown in popularity over the years and are now traded on a wide variety of underlyings. In a novel example, one of these is the weather. Certainly the weather is a major risk factor in many enterprises. However in this text we will confine ourselves to options on stocks.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_3,

The call and put options are basic. From combinations of these many others of widely varying characteristics can be constructed. We will learn about the possibilities a little later.

3.1 Option Payoff Diagrams

A *put* option is a contract between two parties, the *holder* or buyer of the contract and the *writer* or seller. The buyer has the right, but not the obligation, to sell an asset, known as the *underlying*, for a specified price, called the *strike price*, on a specified date, the *expiration date* or *expiry*. The seller guarantees the contract, that is, if requested, buys the asset at the strike price on the expiration date.

Specifically this is a *European put*. An *American put* differs in that the holder has the right to exercise the option at any time up to and including the expiration date. For the time being, we will work exclusively with European options

In most cases the buyer and seller never meet. Instead the transaction is handled by a third party, an exchange. The exchange sets the rules for the contract, offers the products, that is the specific stocks on which it trades puts and calls and their strike prices, maintains the minute by minute trading prices, records the sale, and enforces the contract upon expiration. Option contracts bought and sold through an exchange in this way are standardized to pertain to 100 shares of the underlying and to expire on the third Friday of the expiration month. (Recent changes now offer more frequent expiration dates.)

In Fig. 3.1, as the solid graph, we show a *payoff chart* for the value of a put option on its expiration date. The strike price in this chart is \$100. At expiration, if the stock price exceeds \$100, for example \$102, then the holder will not exercise the option. By exercising the holder receives \$100 per share, but by selling in the market, the holder receives \$102 per share. Therefore the option has no value if the stock price exceeds the strike price at expiration; it is *out-of-the-money (OTM)* and expires worthless. This is also true if the stock price is *at-the-money* (ATM), that is, exactly \$100 at expiration. Thus the solid line segment extends to the right along the x-axis from 100 in the figure.

But the situation is reversed if the stock price is less than the strike price. For example, if the stock price is \$94 per share at expiration, then the holder can buy stock at that price, exercise the put and sell it for \$100 per share. So the option has an *intrinsic* value of \$6 per share here. In general, the value of the option increases by \$1 for each \$1 the stock price is below the strike price. When an option has intrinsic value this way it is said to be *in-the-money (ITM)*. For a put option this occurs when the stock price is below the strike price. A mathematical expression for the payoff value of a put option is

$$V = \max(K - S, 0) = (K - S)^+ \tag{3.1}$$

where K is the strike price and S is the stock price at expiration. The third member of this equation is an alternate notation for the second. This is what is plotted as the solid line in the figure.

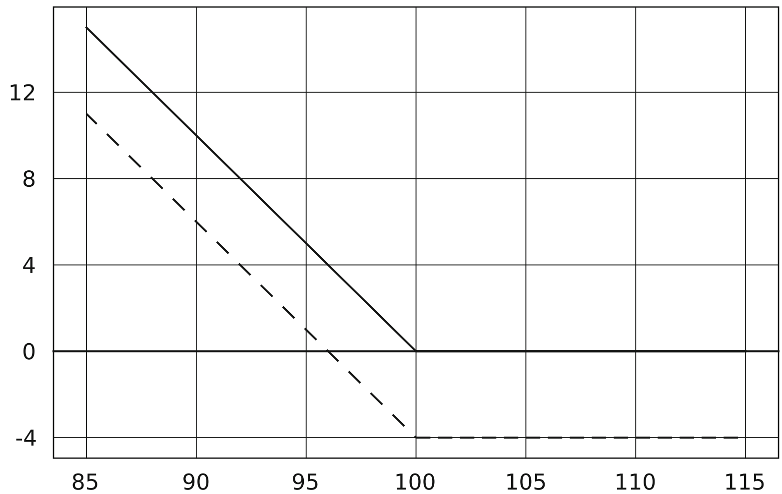


Fig. 3.1. Expiration payoff and gain for a put option with strike price $100

When an option expires in-the-money the holder rarely buys stock and puts it to the writer via the contract. This would entail commissions and unfavorable bid-ask price spreads. Instead the holder just sells the option as it has intrinsic value. On expiration day these contracts have little risk and are bought by, for example, market makers who can transact them at little cost.

The dashed line in Fig. 3.1 is the *profit curve*. It takes into account the fact that the buyer had to pay a cost for the contract; as shown in the figure this is $4. Hence the buyer breaks even when the stock price drops to $96; this is where the profit curve line crosses the x-axis.

3.1.1 Call Options

A *call* option is a contract between two parties, the *holder* or buyer of the contract and the *writer* or seller. The buyer has the right, but not the obligation, to buy an asset, the *underlying*, for a specified price, called the *strike price*, on a specified date, the *expiration date*. The seller guarantees the contract; if requested, the seller must sell the asset at the strike price on the expiration date. This is a *European call*. An *American call* differs in that the holder has the right to exercise the option at any time up to and including the expiration date.

In Fig. 3.2 we show a *payoff chart*, as the solid line segments, for the value of a call option on its expiration date. The strike price in this chart is $100. At expiration, if the stock price is less than $100, for example $96, then the holder will not exercise the option. The holder will prefer to buy stock at the market price of $96 than to call it in at the strike price of $100. Therefore the option has no value in this case; it is *out-of-the-money* and expires worthless. This is also true if the stock price is exactly $100 at expiration.

But if the stock price at expiration exceeds $100 then the call is *in-the-money*. It is in-the-money by $1 for every $1 over the strike price. This is clearly shown in the figure as a line with slope 1 starting at the strike price. A mathematical expression for the payoff value of a call option is

$$V = \max(S - K, 0) = (S - K)^+ \tag{3.2}$$

where K is the strike price and S is the stock price at expiration. Again the third member of the equation is an alternate notation for the second.

The dashed line in Fig. 3.2 is the *profit curve* and takes into account the fact that the buyer had to pay a cost for the contract; as shown in the figure, this is \$4. Hence the buyer breaks even when the stock price reaches \$104 and this is where the profit curve line crosses the x-axis.

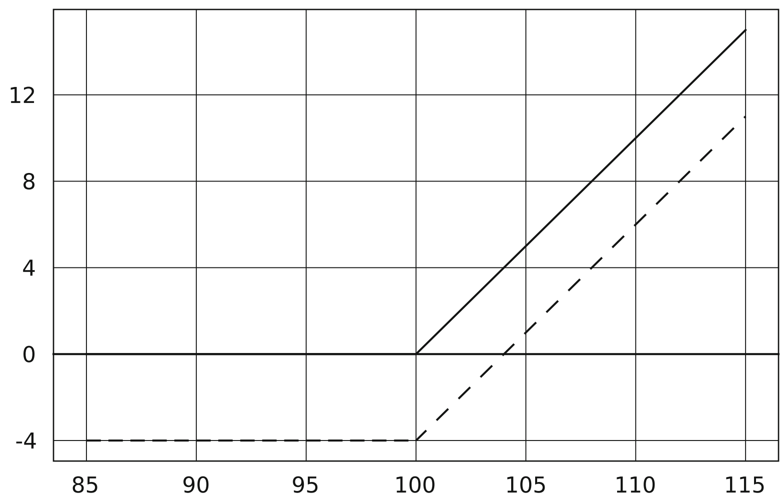


Fig. 3.2. Expiration payoff and gain for a call option with strike price \$100

3.2 Basic Assumptions

To see how to price options we must first review the fundamental assumptions upon which financial pricing mathematics rests. The most far-reaching of these is the no-arbitrage assumption.

Assumptions of Financial Mathematics

1. Random prices: future stock prices are random with no value having probability 1.
2. Positivity: all stock prices are strictly positive (a technical assumption to avoid division by zero in some derivations). This assumption does not preclude bankruptcy; such results can be obtained as limiting values.
3. Divisibility: an investor may hold fractional shares. This is not as restrictive as it may seem; mainly it allows arguments to assume a unit value for one or more items in a portfolio such as an option on one stock. In reality, options are usually on multiples of 100 stocks.

4. Liquidity: an asset can be bought or sold at any time in any amount; there is always a market.
5. Short selling: any asset may be sold, or *shorted* at any time, even if the investor does not have the asset to sell. Such a sale creates a *short position.* The shorted asset will be delivered at some future time. Buying an asset creates a *long position.*
6. No-arbitrage: a portfolio having zero value at some time cannot later have positive value with positive probability and negative value with zero probability. In other words it is not possible to make money with zero cost and no risk.

These assumptions are not perfectly implemented in practice, for example divisibility and short selling. But the unit price of most assets is small compared to the holdings of large investors such as institutions, and even many individuals, consistent with the divisibility assumption. With regard to the latter, short selling is widely available but may be restricted in some cases, for example some equities are not available for short sale or amounts may be limited. But by in large, the assumptions are approximately fulfilled in practice.

An immediate consequence of the no-arbitrage assumption is the following *Monotonicity Theorem.*

Theorem (Monotonicity) *If portfolios A and B are such that at every possible state of the market at time T, portfolio A is worth at least as much as portfolio B, then at any prior time $t < T$ portfolio A is worth at least as much as portfolio B. Moreover, if portfolio A is worth more than B in some states of the world at time T, then at any prior time $t < T$, portfolio A is worth more than B.*

Proof. For the first statement, suppose at some time $t < T$ portfolio B is worth more than A by the amount V. Let portfolio C be long portfolio A, long an amount of cash equal to V, and short portfolio B. By the hypothesis C$= 0$ at time t but has strictly positive value at time T. This is in violation of the no-arbitrage assumption, so there can be no time when B is worth more than A. A simple modification of the above argument proves the second statement as well.

3.2.1 Replication Principle

Another immediate consequence of the basic assumptions is the *replication principle.*

A portfolio is *self-financing* if, after it is initially established, no money is injected or extracted from the portfolio. Therefore all changes to the portfolio are financed by selling assets within the portfolio.

Consider two portfolios A and B over a time horizon $t = 0$ to $t = T$ and suppose the value of the two are exactly the same at time T over all eventualities at that time. Suppose also that both portfolios are self-financing and that the initial price or *set-up cost* of B is known and equal to V. Then the price of A must also equal V. Portfolio B is said to *replicate* portfolio A.

If this were not so, then there is an arbitrage opportunity. The more expensive could be sold and, with the proceeds, the cheaper bought with money left over.

At time T, the payoff of the portfolio bought is used to settle the payoff of the one sold. By the no-arbitrage principle, it follows that the initial price of both must be the same.

3.2.2 Put-Call Parity

Let C_t be the value of a European call option at time t with strike price K and expiry T. Thus the remaining time to expiration is $T-t$. Let P_t be the value of a European put with the same strike and expiry and over the same underlying and S_t be the price of the underlying at time t. Let r_f be the risk-free investment rate. Then *put-call parity* is the relationship

$$S_t + P_t = C_t + Ke^{-r_f(T-t)}. \tag{3.3}$$

This follows from the Monotonicity Theorem. At time t let portfolio A consist of one share of stock and one put option and B consist of one call option and a risk-free investment in the amount of $Ke^{-r_f(T-t)}$. At time T A is worth S_T if $S_T \geq K$ and worth K otherwise. But at time T this is what B is worth as well. So their values must be equal at all times $t < T$.

3.3 Forward Contracts

A *forward contract* is an agreement between two parties to buy/sell an asset on a fixed date in the future for a price specified in advance. (The date can be approximate in the case of farm products.) The party obliged to buy is said to be *long* the contract and the party obliged to sell is *short*. Note that no money is exchanged until the delivery date. The question is, what should be the price of the contract?

The cookie maker of the first section might have made a forward contract instead of buying a call. Through past experience, with a current price of \$8.41 for wheat, the fall price might be expected to come in at \$9.00. Going long a forward contract at this price, the cookie maker knows exactly what to expect in the fall and can make plans accordingly, the uncertainty is completely removed.

3.3.1 Pricing an Investment Forward Contract

Assets underlying forward contracts must be divided into investment assets or consumption assets. A consumption asset is one used primarily for consumption, for example commodities, wheat included. Investment assets are those held primarily for investment purposes. Generally speaking investment assets entail little or no storage or upkeep costs; stocks and bonds are examples. They are the easiest to price requiring only knowledge of the current market price and possibly other market variables such as the risk-free rate. In the following we confine ourselves to this kind of asset.

To price a forward contract, let $t = 0$ be the time the contract is signed and $t = T$ the delivery date. Let S_t be the price of the asset, a stock to be definite, at any time $0 \le t \le T$ and let F_T be the contracted forward price. Finally let r_f be the risk-free interest rate.

By the no-arbitrage principle we show that the forward price of a non-dividend paying stock must be

$$F_T = S_0 e^{r_f T}. \tag{3.4}$$

For suppose the price were higher. To invoke arbitrage, buy the under priced asset and sell the overpriced one. At time 0

- Borrow the amount S_0 at the risk-free rate;
- Buy one share for S_0;
- Short the forward position, that is agree to sell one share for F_T at time T.

At time T clear the portfolio,

- Sell the stock for F_T;
- Pay $S_0 e^{r_f T}$ to retire the loan.

This will bring a risk-free profit, at zero cost, of

$$F_T - S_0 e^{r_f T} > 0.$$

By no-arbitrage then F_T cannot be greater than $S_0 e^{r_f T}$.

On the other hand, if the forward price is less than $S_0 e^{r_f T}$, once again buy the under priced and sell the overpriced. At time 0

- Sell short one share of stock for S_0, that is, sell one share of stock now with the stock to be delivered in the future;
- Invest the proceeds at the risk-free rate;
- Buy the forward contract, that is agree to buy one share of stock for F_T at time T.

At time T settle the position

- Cash out the risk-free investment for $S_0 e^{r_f T}$;
- Exercise the forward contract and buy one share for F_T;
- Deliver the one share of shorted stock.

This will bring a risk-free profit, at zero cost, of

$$S_0 e^{r_f T} - F_T > 0.$$

Again, by no-arbitrage F_T cannot be less than $S_0 e^{r_f T}$. So the price must be as given by (3.4).

3.3.2 Forward Contracts on a Stock Awarding Dividends

Assume that during the life of a forward contract the stock earns a dividend payment of D coming at the time t_D from the beginning of the contract. Then the correct price of the forward contract is

$$F_T = (S_0 - De^{-r_f t_D})e^{r_f T}. \tag{3.5}$$

Observe that $S_0 - De^{-r_f t_D}$ is the present value of the stock at time 0. It plays the role of S_0 in the argument above.

To argue arbitrage here, first suppose F_T is bigger than the right-hand side of (3.5). Hence short the forward contract and buy the stock. To accomplish the latter, borrow S_0.

At time t_D we receive a dividend payment of D; use this to partially pay back the loan. The loan value at this time has grown to $S_0 e^{r_f t_D}$. And, after paying D on it, the new amount continuing forward is $S_0 e^{r_f t_D} - D$.

At time T hand over the stock to satisfy the contract and receive F_T. Finally pay the balance of the loan which is now

$$(S_0 e^{r_f t_D} - D)e^{r_f (T - t_D)} = (S_0 - De^{-r_f t_D})e^{r_f T}.$$

Since F_T is bigger than this, we have made a risk-free profit.

We leave it to the reader to argue the case if F_T is less than the right-hand side of (3.5).

The interpretation of (3.5) is that the initial value of the stock is reduced by the discounted dividends which occur during the life of the option. This interpretation holds for any number of dividend payments.

Example 3.1. What should be the price of an 8 month forward contract for 100 shares of a stock whose price today is \$43.44? The company has already announced it will give a \$0.77 per share dividend 3 and 6 months from now. The risk-free rate is 3 %.
Directly substituting into (3.5) we have

$$\begin{aligned} F_T &= \left(43.44 - 0.77e^{-0.03*3/12} - 0.77e^{-0.03*6/12}\right)e^{0.03*(8/12)} \\ &= (43.44 - 1.52)1.02 = 42.76 \end{aligned}$$

per share; so \$4,276 for 100 shares. □

Continuous Dividend Payments

Normally dividends are paid periodically with the return rate specified annually for comparison purposes. For example a company may have the reputation of paying dividends quarterly at the annual rate, or yield, of 8 %. Thus the company's dividend in a given quarter is $S(q/4)$ where S is the stock price on ex-dividend day and $q = 8\,\%$. Despite this, dividends are often assumed to be continuous as a simplification just as the risk-free rate is treated this way.

In Section 2.3 we saw that dividends paid out on an equity have the effect of reducing its price by the same amount. But if the dividends are reinvested in additional stock, the effects are self canceling. That is the reduction in the stock's value is exactly offset by the increase in value for having more shares. Here is the calculation. Let S_0 be the original stock price per share and let D be the dividend paid per share. The new stock price is $S_0 - D$, so the new number of shares is

$$1 + \frac{D}{S_0 - D} = \frac{S_0}{S_0 - D}. \tag{3.6}$$

The new value is the product of the new price per share times the new number of shares

$$(S_0 - D)\frac{S_0}{S_0 - D} = S_0.$$

Now assume dividends are paid continuously at the annual rate q. Divide the interval $[0, t]$ into n equal subdivisions $\Delta t = t/n$ where n is large. Over such an interval, the dividend paid is approximately $D = Sq\Delta t$ where S is the stock price at that time. From (3.6), if this is reinvested, the new number of shares is

$$\frac{1}{1 - q\Delta t} \approx \frac{1}{e^{-q\Delta t}} = e^{q\Delta t}.$$

Notice that this expression is independent of the stock price; even though the stock price may vary over the increments, the amount of the dividend paid is exactly the right amount to buy the additional stock as stated. After time t, n such re-investments have occurred and the number shares at that time is

$$\prod_{i=1}^{n} e^{q\Delta t} = e^{qn\Delta t} = e^{qt}. \tag{3.7}$$

Since this expression is independent of n, it is also its own limit as $n \to \infty$. This shows how 1 share of stock grows under continuous dividend reinvestment.

With these preliminaries we are prepared to show that the fair price of a forward contract on a stock giving continuous dividends with yield q is

$$F_T = S_0 e^{-qT} e^{r_f T}. \tag{3.8}$$

Again we use a no-arbitrage argument.

If $F_T > S_0 e^{(r_f - q)T}$, go short the forward contract and buy the stock as follows. Borrow the amount $S_0 e^{-qT}$ at the risk-free rate and buy e^{-qT} shares. As we've just seen, by continuously reinvesting the dividends on the stock, the number of shares will grow to 1 at time T. Settle the forward contract collecting F_T. Finally pay the original loan with interest for $S_0 e^{-qT} e^{r_f T}$. The difference $F_T - S_0 e^{(r_f - q)T}$ is risk-free profit. Hence F_T cannot be bigger than the right-hand side in (3.8).

Now suppose it to be smaller. In this case, go long the contract and sell short e^{-qT} shares of the stock investing the proceeds $S_0 e^{-qT}$ at the risk-free rate.

Because we have shorted the stock, we must pay the owner the dividends as they accrue. We do this by continuously shorting additional stock. The argument used to prove (3.7) applies here except that the increment in shares resulting from the dividend is now the increment shorted with the proceeds used to pay the dividend. At time T our short position has grown to 1 share.

And at that time settle the forward contract by buying one share of stock for F_T. Clear the short position with this share. Finally collect the invested made at time 0 now in the amount of $S_0 e^{-qT} e^{r_f T}$. Since $F_T < S_0 e^{(r_f - q)T}$ we have made a risk-free profit. This shows that (3.8) gives the correct contract value.

3.3.3 Valuing a Forward Contract

On the date that a forward contract is negotiated, $t = 0$, we have derived above that its contract price should be F_T; this is the price to be paid upon delivery on the forward date, when $t = T$. Let us call this the contract or delivery price and denote it by C. At this time, $t = 0$, the contract has no value because anyone could also negotiate the same forward contract for C.

But suppose the price of the stock on delivery date exceeds C. Then the party long the contract (receiving the stock) can, theoretically, sell it for the difference $S_T - C$. The contract now has value. (Of course if $S_T < C$ then the party short has the value $C - S_T$.)

Similarly, suppose that at time $t < T$, the stock price is S_t, what is the contract's value now?

The time remaining to the contract date is $T - t$, therefore the delivery price of a new contact is

$$F_{T-t} = S_t e^{r_f (T-t)}$$

if there are no further dividends, otherwise use (3.5) or (3.8) as appropriate. On delivery date the value of the old contract versus the new one is $F_{T-t} - C$. By discounting this back to time t we get the value at that time,

$$V_t = (F_{T-t} - C) e^{-r_f (T-t)}. \tag{3.9}$$

Expected Winners

As we have just seen, upon maturity of the contract, if $S_T > F_T$ then the party long can sell it for an immediate profit of the difference. On the other hand, if $S_T < F_T$, then the party short can buy the asset for the lower price and fulfill the contract for the higher price. Can either party *expect* to make a profit (in the mathematical sense)?

From Section 1.1 and in particular, equation (1.28), we have that

$$\mathbb{E}(S_T - F_T) = \mathbb{E}(S_T) - F_T = S_0 e^{\mu T} - S_0 e^{r_f T}$$

where μ is the drift of the stock. So if the drift exceeds the risk-free rate, then the buyer can expect to do better through the forward contract than waiting to pay the market price at time T.

Alternatively we could use Monte Carlo to calculate the answer. In a similar manner to what was done in Section 2.3, run the algorithm on page 12 but this time observe the buyers winnings

$$S_T - F_T.$$

If the sum of the winnings over a large number of trials is 0, then the no-arbitrage price is also the expected ending price. The Monte Carlo advantage is that it will still work even if prices do not follow a geometric random walk. The basic step by step calculation would have to be modified to adhere to an alternative model. In fact we will take up these ideas in a later chapter.

3.4 Option Pricing: Binomial Lattice Model

In this section we want to see how to fairly price basic put and call options. To simplify matters at this point, we restrict ourselves to options over non-dividend paying stocks. From what we have learned in pricing forward contracts we expect that the price will be determined by the no-arbitrage principle. This is correct. To see why, we first consider the binomial pricing approximation model introduced in Section 1.6. Initially we will be more interested in pricing principles and for this reason we start with the simple one-step binomial tree. From there we will be able to treat the general case and formulate specific pricing solutions for both European and American options.

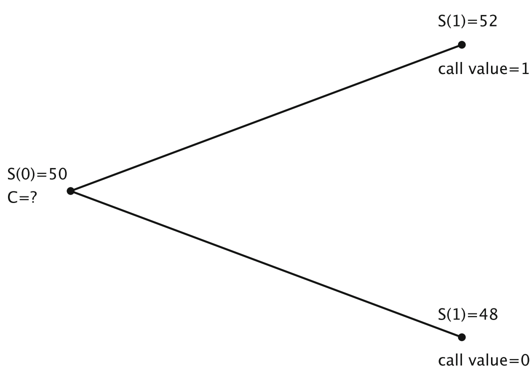


Fig. 3.3. Stock prices and call option values for a one-step binomial tree

3.4.1 Pricing for a One-Step Binomial Tree

The simplest case is that of one time step as in Fig. 3.3. Today, $t = 0$, the stock price is, say, \$50. At expiration, $t = 1$, the stock price will be either \$52 if the price goes up or \$48 if it goes down. We wish to sell a call option on the stock with strike price \$51. What should be the price C of the option?

Since we may have to deliver stock at expiration, we buy a quantity, Δ, of stock today to help cover the delivery; Δ could be fractional. Our portfolio is therefore long Δ shares of stock and short one call option; its value is

$$50\Delta - C. \tag{3.10}$$

At expiration, if the price went up, the portfolio's value is $52\Delta - 1$, \$1 being the difference between the stock price and the strike price. And if the price went down, it is 48Δ. By setting these equal and solving for Δ we can remove the uncertainty of our position. We get

$$52\Delta - 1 = 48\Delta, \qquad \text{so} \quad \Delta = \frac{1}{4}.$$

To find C, note that with Δ given as above, the portfolio's value at $t = 1$ is \$12; in fact $52(1/4) - 1 = 12$ and $48(1/4) = 12$. Following our familiarity with forward contracts we presume the $t = 0$ value should be this discounted by the risk-free rate r_f. Hence[1,2]

$$50\Delta - C = 12e^{-r_f}$$

which gives

$$C = 12.50 - 12e^{-r_f}. \tag{3.11}$$

In fact this is the no-arbitrage price. We show that C cannot be bigger than the right-hand side of (3.11) and leave it to reader to show it cannot be smaller.

Suppose $C > 12.50 - 12e^{-r_f}$. As always, we sell the more expensive and buy the cheaper. At $t = 0$

- Borrow \$12.50 and buy 1/4 share of stock;
- Sell 1 call option for C and invest the proceeds.

At $t = 1$, if the price is up

- 3/4 share may be purchased from the market for \$39, along with our 1/4 share deliver the stock for the strike price, \$51;
- Pay off the loan for $\$12.50e^{r_f}$.

Our net value will be

$$Ce^{r_f} + 51 - 39 - 12.50e^{r_f} = (C - (12.50 - 12e^{-r_f}))e^{r_f} > 0.$$

On the other hand, if the price is down

- Sell 1/4 share for \$12.

The net position in this case will be

$$Ce^{r_f} + 12 - 12.50e^{r_f} = (C - (12.50 - 12e^{-r_f}))e^{r_f} > 0$$

as in the first case. Thus C cannot be larger than the right-hand side of (3.11).

Similarly it cannot be smaller; hence (3.11) gives the no-arbitrage price.

[1] Throughout we assume r_f is given in terms of the time units used for t.

[2] For discrete discounting replace e^{-r_f} by $(1+r_f)^{-1}$. More generally, replace e^{-kr_f} by $(1+r_f)^{-k}$.

The Method of Replication

The call option can also be priced by the replication principle of Section 3.2.1. Consider a portfolio consisting of 1/4th share of stock and a loan of $12e^{-r_f}$ where r_f is the risk-free rate. At payoff, if the stock price has gone up to 52, then the value of the portfolio is $\frac{1}{4}52 - 12 = 1$. On the other hand, if the stock price has gone down to 48, then the value is $\frac{1}{4}48 - 12 = 0$. These payoffs match that of the option exactly. Since the initial value of the replicating portfolio is $\frac{1}{4}50 - 12e^{-r_f} = 12.50 - 12e^{-r_f}$, this must also be the value of the call.

Notice that the "delta" of the option, meaning its change in value with respect to a change is the stock price is $\frac{1}{4}$,

$$\frac{\Delta C}{\Delta S} = \frac{1-0}{52-48} = \frac{1}{4}, \tag{3.12}$$

and this is how much stock to buy.

Risk-Neutral Valuation

A notable aspect of the derivation above is that probability played no role.[3] Let p be the statistical probability of an up move as generated by the recent price history of the stock. This is called the *statistical distribution* or *statistical measure* of the stock's prices. One would think that if the probability of an up move were high, then the option should cost more. But as we have seen, a call premium different than that given by (3.11) results in an arbitrage opportunity. But arbitrage cannot be sustained for very long. Prices will quickly adjust so as to eliminate it.

Instead, a high up move probability results in a high profit expectation for the option holder. But there is a probability for which the expectation is zero.

Let q be the probability of an up move; then the option holder makes \$1 with probability q or \$0 with probability $1-q$. The earnings expectation is therefore q. For this, the buyer pays $12.50 - 12e^{-r_f}$ at time 0, or, equivalently, $12.50e^{r_f} - 12$ at time 1 (when the payoff occurs). The break-even probability is therefore

$$q = 12.50e^{r_f} - 12. \tag{3.13}$$

This is called the *risk-neutral* probability. For example, if the risk-free rate were 0 the risk-neutral probability would be $q = 1/2$ in this example (from (3.13)).

If the risk-neutral probability of an up move were the actual probability of an up move, then the buyer of the option has the same earnings *expectations* as having invested risk free at the risk-free rate.

Another characterization of the risk-neural probability is that it makes the payoff fair in the sense of expectation (for both sides of the option contract).

[3] More exactly, almost no role; that there are two possibilities for the future tacitly assures that neither the up branch nor down branch occurs with probability 1. Such would be a violation of financial assumption 1 (page 81). Further, if the up branch occurred with probability 1, and the time period sufficiently short, then the guaranteed \$2 return could exceed the risk-free rate; indeed, it would become the risk-free rate.

Stated differently, the expected *discounted* change in the portfolio value from $t = 0$ to $t = T$ is 0,

$$\mathbb{E}_q(\text{discounted change in portfolio value}) = 0. \tag{3.14}$$

The expectation is taken with respect to q. A probability having the property that the expected future value of a random variable is equal to it's present value is called a *martingale*.

The martingale probability provides another way to price an option. It is one of the most important principles in pricing assets that derive their value from that of an another, underlying, security

Risk-neutral option valuation principle

> An option's price is equal to the discounted expected payoff of the option, the expectation taken with respect to the risk-neutral probability.

And from above

Risk-neutral probability

> The risk-neutral probability is the probability for which the expected growth of the underlying is at the risk-free rate.

We now have 3 ways of pricing options: no-arbitrage (directly calculated), replication, and risk-neutral. However, in fact they all are based on the no-arbitrage principle.

Option Pricing Winners and Losers

As we saw in pricing forward contracts, one or the other side of the contract may enjoy an *expectation* advantage for making a profit. The same is true with regard to option pricing. In fact the risk-neutral pricing alternative spells it out explicitly since it assumes the underlying grows at the risk-free rate. But this is rarely the case. A given equity's prices grow at their own rate determined by factors including management decisions, the fortunes of the equity's sector and many others. An equity's market growth rate is quantified by the drift μ of its prices.

Again let p be the historical probability of an up move in Fig. 3.3; as before, the expected price of the stock at $t = 1$ is given by $52p + 48(1 - p)$. Using the stock's actual growth rate, $50e^{\mu}$ given by its drift, the historical probability can be determined,

$$52p + 48(1 - p) = 50e^{\mu},$$

giving the actual up move probability as

$$p = 12.50e^{\mu} - 12. \tag{3.15}$$

Since the cost of the option is C, the buyers profit is $1 - C$ in case of an up move and $-C$ for a down move. Hence

$$\begin{aligned}\text{expected profit} &= (1-C)p + (-C)(1-p) = p - C \\ &= (12.50e^{\mu} - 12) - (12.50e^{r_f} - 12) = 12.50(e^{\mu} - e^{r_f}) \\ &= \frac{1}{4}S(0)(e^{\mu} - e^{r_f});\end{aligned} \tag{3.16}$$

recall $(1/4)S(0)$ is the buyers initial investment.

Even though this looks promising for the holder of a call option on a high growth stock, there are elements that tend to counter the advantage. As we will see in the next section, option prices depend on the volatility of the underlying; as the volatility changes, the option price changes correspondingly. The result is that volatilities at any moment become market determined and, as such, vary as any other market element. VIX is the ticker symbol for the market volatility index.

3.4.2 Pricing for a Multi-step Binomial Lattice

In this section we extend the development of the one-step binomial model to an arbitrary number of steps. We continue to treat options over non-dividend paying stocks.

The development here is based on the construction in Section 1.6. As in that section, the lattice is extended by means of up and down factors $u \geq 1$ and $0 < d \leq 1$. These factors as well as their up versus down probabilities are assumed constant throughout the lattice.

As we learned in the previous section, one possibility for calculating the option price invokes the no-arbitrage principle directly. In this approach probabilities play no role. The other uses the risk-neutral principle with the lattice probabilities equal to the risk-neutral probabilities q and $1-q$. In order to contrast the two with the least distraction we first apply them to a two-step lattice.

Calculating the Call Price for a Two-Step Lattice

The starting equity price is S_0. Over the first time period Δt the price could go up to S_0u or down to S_0d. In the next time period the same thing occurs, as a result the prices at time $t = 2$ $(2\Delta t)$ are S_0u^2, S_0ud, and S_0d^2, and the probabilities are, respectively q^2, $2q(1-q)$ and $(1-q)^2$, see Fig. 3.4. And to be definite, we assume the strike price K lies between $S_0ud < K < S_0u^2$. Therefore the call payoff values at $t = 2$ are 0 except at the top node for which it is $S_0u^2 - K$.

Calculating C from the No-Arbitrage Principle Directly

We can find the option cost C using the no-arbitrage principle by analyzing all one-step two-branch subtrees in Fig. 3.4 in reverse order; that is from the right hand side of the figure to the left. Applying the techniques of the previous section to the subtree with root node $N_1(1)$ (having stock price S_0u) and branches to S_0u^2 and S_0ud calculates the value C' of the call at that root. The value at node

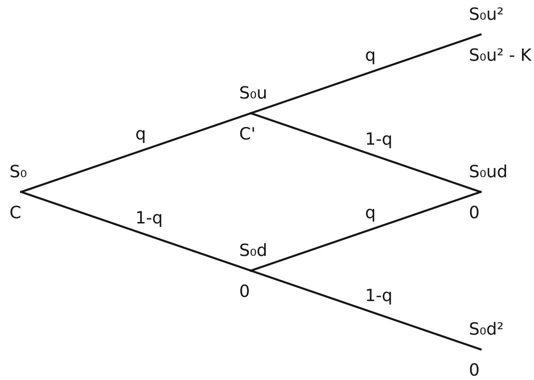


Fig. 3.4. Stock prices as the upper numbers and call option values the lower ones for a two-step binomial lattice. The strike price lies between S_0ud and S_0u^2

$N_1(0)$ (having stock price S_0d) is 0 as indicated because both its branches lead to zero payoffs. Finally, knowing C', the analysis of the one-step tree at the root node N_0 gives the call value C at time $t = 0$.

To begin, from (3.12) we find the amount of stock to buy at $N_1(1)$ to be

$$\Delta = \frac{S_0u^2 - K}{S_0u^2 - S_0ud}.$$

Thus the value of the portfolio at $N_2(1)$ is

$$S_0ud\frac{S_0u^2 - K}{S_0u^2 - S_0ud} = \frac{d}{u-d}(S_0u^2 - K).$$

Now discount this back and solve for C',

$$S_0u\left(\frac{S_0u^2 - K}{S_0u^2 - S_0ud}\right) - C' = \frac{de^{-r_f\Delta t}}{u-d}(S_0u^2 - K).$$

We obtain

$$C' = (S_0u^2 - K)\left(\frac{1 - de^{-r_f\Delta t}}{u-d}\right).$$

We leave it to the reader to repeat this program for node N_0 to derive

$$C = C'\left(\frac{1 - de^{-r_f\Delta t}}{u-d}\right) = (S_0u^2 - K)\left(\frac{1 - de^{-r_f\Delta t}}{u-d}\right)^2. \tag{3.17}$$

Calculating C from the Risk-Neutral Principle

But there is a much easier calculation using the risk-neutral principle. Furthermore, the technique extends just as easily to lattices having any number of steps.

The risk-neutral probability is easily determined by equating means as was done in Section 1.6.2. However here we must assume risk-free growth,

$$qu + (1-q)d = e^{-r_f\Delta t}$$

so

$$q = \frac{e^{-r_f \Delta t} - d}{u - d}. \tag{3.18}$$

Knowing q, at expiration the expected value of the call is

$$(S_0 u^2 - K) \times q^2 + 0 \times 2q(1-q) + 0 \times (1-q)^2.$$

Discounting this back to $t = 0$ gives the call price[4]

$$C = (S_0 u^2 - K) q^2 e^{-2r_f \Delta t} = (S_0 u^2 - K) \left(\frac{e^{r_f \Delta t} - d}{u - d} \right)^2 e^{-2r_f \Delta t},$$

and we get the same as before, (3.17).

3.5 Pricing Put and Call Options Over Non-dividend Paying Stocks by the Binomial Method

3.5.1 Extension to an n-Step Lattice

Now consider a binomial lattice with n time steps, $t = n$ representing expiry. From Section 1.6, at any time $t = k$, $0 \le k \le n$, there are $k + 1$ nodes $N_k(i)$ for $0 \le i \le k$. And the stock prices and probabilities are given in (1.33),

$$\text{price at node } N_k(i) \text{ is } \; S_0 u^i d^{k-i}, \tag{3.19}$$

and the probability of reaching this node is

$$\binom{k}{i} q^i (1-q)^{k-i}. \tag{3.20}$$

Again there are two ways to proceed. In the first, set $k = n$ in the equations above and generate the prices and probabilities at expiration. From the prices, the payoffs are calculated,

$$(S_0 u^i d^{n-i} - K)^+ \; \text{ for calls or } \quad (K - S_0 u^i d^{n-i})^+ \; \text{ for puts.}$$

Finally the sum of the probability weighted payoffs are discounted back to time $t = 0$.

In the other, again starting with the expiration payoffs on the right, one works back node by node to the root N_0 generating the option values at each intermediate node along the way. The option cost is the value calculated at N_0.

We detail both starting with the risk-neutral calculation. As we have seen, the risk-neutral calculation is more direct and computationally simpler. On the other hand the step-by-step approach has the advantage that it can be used to price other types of options such as those that are *path dependent*, that is, dependent on the price of the underlying over the course of the contract. An American option is such an example and we will illustrate this application.

[4] For discrete discounting, replace $e^{-kr_f \Delta t}$ by $(1 + r_f \Delta t)^{-k}$ throughout.

3.5.2 Risk-Neutral Binomial Pricing of Options

Let the strike price of a call option lie between the $t = n$ nodes for $i = m - 1$ and $i = m$,

$$S_0 u^{m-1} d^{n-m+1} \leq K < S_0 u^m d^{n-m}. \tag{3.21}$$

(This observation only serves to save some computational effort by summing only over those terms having a positive payoff; one could otherwise sum over all the binomial terms and use the payoff evaluation function $(S_0 u^i d^{n-i} - K)^+$ instead.) Then the call payoff is the sum over those nodes $i \geq m$ of the payoff $S_0 u^i d^{n-i} - K$ times the probability of reaching $N_n(i)$,

$$\binom{n}{i} q^i (1-q)^{n-i} (S_0 u^i d^{n-i} - K).$$

The call price is this sum discounted back n time steps to $t = 0$. In summary, we have the following for a European call option whose underlying is a non-dividend paying stock

Price for a European Call Option

$$C = e^{-r_f n \Delta t} \sum_{i=m}^{n} \binom{n}{i} q^i (1-q)^{n-i} (S_0 u^i d^{n-i} - K). \tag{3.22}$$

Example 3.2. By the binomial method calculate the call option price 3 weeks before expiration for the following: current stock price: $S_0 = 26$, strike price: $K = 26$, risk-free rate: $r_f = 3\,\%$, volatility: $\sigma = 23\,\%$.

We will use a 3-step, $u = 1/d$ type lattice, therefore $\Delta t = 7/365 = 0.019178$. Equations (1.38) may be used to find the lattice's parameters provided the drift is replaced by the risk-free rate. Then $A = 1.00050798$, $d = 0.968630$, $u = 1.032386$, and $q = 0.50106$. From (3.19) and (3.20) we calculate the ending prices as:

$$N_3(0) : 23.63 \quad N_3(1) : 25.18 \quad N_3(2) : 26.84 \quad N_3(3) : 28.61$$

and ending probabilities as:

$$N_3(0) : 0.124 \quad N_3(1) : 0.374 \quad N_3(2) : 0.376 \quad N_3(3) : 0.126.$$

See Fig. 3.5. The strike price $K = 26$ lies between nodes $N_3(1)$ and $N_3(2)$. So the sum in (3.22) extends over the two upper nodes in the figure. We have

$$\begin{aligned} C &= e^{-3(0.03*7/365)} \Big(0.376(26.84 - 26) + 0.126(28.61 - 26) \Big) \\ &= 0.6436. \end{aligned}$$

□

Example 3.3. Work the same problem as above but use the $p = 1/2$ type lattice. This time equations (1.40) calculate the lattice parameters. First $\sqrt{e^{\sigma^2 \Delta t} - 1} = 0.031859$, then $d = 0.968697$, and $u = 1.032453$. The risk-neutral probability as calculated by (3.18) is $q = 0.5$. Of course, this shows that in the $p = 1/2$ type lattice, p still satisfies (3.18).

To finish the example, from (3.19) the ending prices are:

$$N_3(0) : 23.63 \quad N_3(1) : 25.19 \quad N_3(2) : 26.85 \quad N_3(3) : 28.61$$

the ending probabilities are:

$$N_3(0) : 0.125 \quad N_3(1) : 0.375 \quad N_3(2) : 0.375 \quad N_3(3) : 0.125.$$

As before, the sum extends over the two upper nodes in the figure. We have

$$\begin{aligned} C &= e^{-3(0.03*7/365)} \Big(0.375(26.85 - 26) + 0.125(28.61 - 26) \Big) \\ &= 0.6438. \end{aligned}$$

□

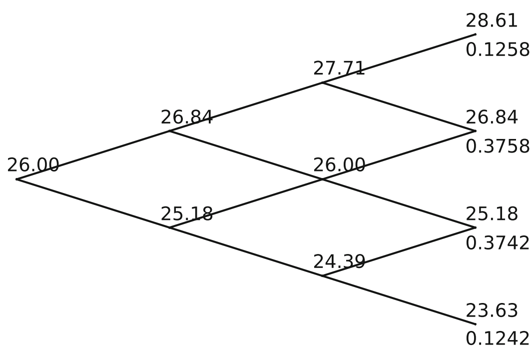


Fig. 3.5. Pricing lattice for Example 3.2. The upper numbers are the nodal prices and the lower are the nodal probabilities

It is just as easy to price put options. Just as for calls, the price is given by the discounted expected payoff. Therefore the only change is to replace the payoff function by $(K - S_0 u^i d^{k-i})^+$, or, with m as defined above in equation (3.21), sum from 0 up to $m - 1$. For a non-dividend paying underlying we have

Price for a European Put Option

$$P = e^{-r_f n \Delta t} \sum_{i=0}^{m-1} \binom{n}{i} q^i (1-q)^{n-i} (K - S_0 u^i d^{n-i}). \tag{3.23}$$

3.5.3 Node-by-Node Binomial Option Pricing

Again start with the prices at $t = n$, equation (3.19). For every pair of adjacent nodes $N_n(i)$ and $N_n(i+1)$, back calculate the option value to their root $N_{n-1}(i)$. Having calculated all the option values at $t = n - 1$, repeat the procedure for the nodes at $t = n - 2$. Continuing in this way inductively from right to left we finally calculate the option value at N_0.

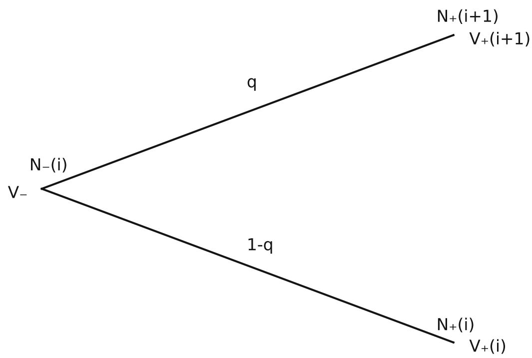


Fig. 3.6. A generic two branch subtree. The option values $V_+(i)$ and $V_+(i+1)$ are known as is the probability q. The root value V_- can be inferred from them. See the text

In Fig. 3.6 we show a generic two branch subtree. Between the node on the left and those on the right is a single time increment Δt. The option values $V_+(i)$ and $V_+(i+1)$ are known. To backcalculate the option value V_- we will use the risk-neutral principle. In this way we do not have to maintain the stock price at each node.

To calculate the expected option value over these two + nodes we must use the probabilities q and $(1-q)$ and not their absolute probabilities. This is because reaching either of these nodes is conditioned on first reaching their root node $N_-(i)$. The conditional expected value of the option is therefore $qV_+(i+1) + (1-q)V_+(i)$. By the risk-neutral principle the option value at the root is this discounted at the risk-free rate, hence

$$V_- = e^{-r_f \Delta t} \left(qV_+(i+1) + (1-q)V_+(i) \right). \tag{3.24}$$

Example 3.4. To illustrate the foregoing techniques, we work through the put price calculation example shown in Fig. 3.7.

Assume the current stock price is \$50, expiration is in 4 days, the strike price is \$49, the risk-free rate is 26 % and the volatility is 40 %. We will use the $q = 1/2$ system. In terms of years, $\Delta t = 1/365 = 0.002740$. From (1.40) we find that

$$\begin{aligned} d &= e^{r_f \Delta t} \left(1 - \sqrt{e^{\sigma^2 \Delta t} - 1} \right) \\ &= 1.00071 \left(1 - \sqrt{1.000438 - 1} \right) = .97976. \end{aligned}$$

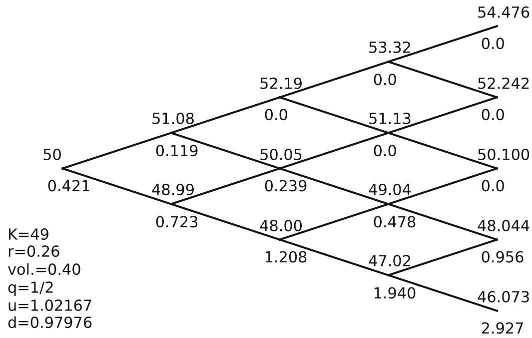


Fig. 3.7. Back calculation method for a put option. Stock prices are the upper numbers and option values the lower ones for a four-step binomial lattice

Similarly

$$u = 1.02167.$$

We may now calculate the expiration prices: $S_0u^4 = 54.476$, $S_0u^3d = 52.242$, $S_0u^2d^2 = 50.100$, $S_0ud^3 = 48.044$, and $S_0d^4 = 46.073$. Since the strike price is 49, the put payoffs are, respectively: 0, 0, 0, 0.956, and 2.927.

Using (3.24), for the first back calculation at node $N_3(0)$ we get

$$V = e^{-0.26*0.002740}\left(\frac{1}{2}0.956 + \frac{1}{2}2.927\right) = 1.940.$$

In a similar fashion work back through the other 9 nodes. The last two-branch subtree gives the price of the option

$$V = e^{-0.26*0.002740}\left(\frac{1}{2}0.119 + \frac{1}{2}0.723\right) = 0.421.$$

We can corroborate our answer with the risk-neutral method, equation (3.23). Since the branch probabilities are $q = 1/2$, the expiry probabilities are easily seen to be: 1/16, 4/16, 6/16, 4/16, and 1/16. Hence the expected payoff is

$$\frac{1}{16}2.927 + \frac{4}{16}0.956 = 0.4219$$

Discounting this 4 days gives

$$e^{-4r_f\Delta t}0.4219 = 0.9971 * 0.4219 = 0.421.$$

Monte Carlo Solution

The risk-neutral method discussed above can also be implemented via Monte Carlo by simulating paths through the binomial lattice. As usual the Monte Carlo calculation will take more computational time than its deterministic counterpart but, by its very nature, it can be applied when other methods fail. Recall the assumptions on the binomial lattice method are that volatility and the risk-free rate remain constant. If they are not, then neither are the up and down factors. When that happens the lattice does not reconnect, even in the second time step $S_0 u_0 d_1 \neq S_0 u_1 d_0$ where u_k and d_k are the up and down factors at step k. Then, technically, it becomes a binomial tree and the the number of nodes grows exponentially with the number of steps. The method becomes untenable at some point.

In contrast the Monte Carlo method can allow the up and down factors and the branch probabilities to be nodal dependent with no appreciable computational cost.

Starting at the root node with $S = S_0$, simulate a path through the tree by selecting up or down at each subsequent node according to the up probability q. Upon reaching the right hand side of the lattice, S holds the price of the underlying at expiration for the particular path taken. The option payoff for this path is $V = (S-K)^+$ for a call or $V = (K-S)^+$ for a put as usual. Repeat this calculation N times accumulating the payoffs, for example, let E be their sum. Then the estimate for the expected payoff is E/N and the option price is this discounted back to time 0.

Algorithm 11. Monte Carlo Method for Traversing a Binomial lattice

```
inputs: S_0, K, nDays, r, σ, N
  ▷calculate u, d, q according to (1.38) or(1.40)
Δt = 1/365.0    ▷one day time increments
E = 0
for i = 1,...,N
  S = S_0
  for t = 1,...,nDays
    U ~ U(0,1)
    if( U < q ) S = Su else S = Sd
  endfor
  V = (S − K)^+ for calls or V = (K − S)^+ for puts
  E = E+V
endfor
E = E/N
price = e^{−nDays∗r_f∗Δt} E
```

3.5.4 American Options

An *American option* works exactly like its European counterpart with the exception that the option can be exercised at any time up to and including the expiration date. Thus the American option bestows more privileges than its European counterpart and so its value at any time must be at least as great.

To price an American option consider the binomial pricing technique as implemented node-by-node. The value of the option is calculated at each node as for a European option. But since the option can be exercised at any time, this value must be compared to the payoff if the option were to be exercised at that moment. If the early exercise value is larger, then it becomes the true value at that node. This is the only modification required. Referring to Fig. 3.6, the value at node V_- is the greater of

V_- as given by discounting its child nodes, (3.24), or
the in-the-money value of the option.

To illustrate these considerations, we rework the previous example but now for an American put. See Fig. 3.8.

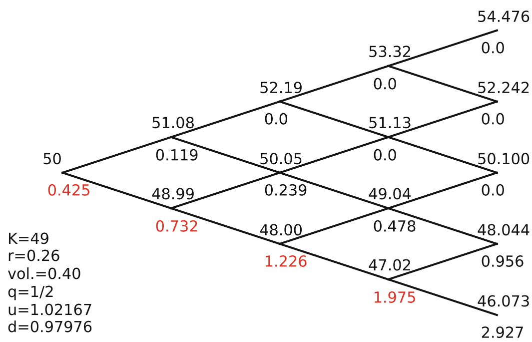


Fig. 3.8. Node-by-node calculation method for an American put option. Stock prices are the upper numbers and option values the lower ones for a four-step binomial lattice. The nodal option values increased by early exercise are shown in *red*

Example 3.5. As before the current stock price is \$50, expiration is in 4 days, the strike price is \$49, the risk-free rate is 26 % and the volatility is 40 %. And as before, the back calculated option value at node $N_3(0)$ is \$1.940. But since the stock price is $S_0 d^3 = 47.025$ at that node, the intrinsic value is $49 - 47.025 = 1.975$, thus exceeding the European option value. Hence the holder should exercise the option. For our calculation, we must replace 1.940 by 1.975 and continue with this larger number. Note that this 35 cent difference is propagated along but by approximately half at each new step. In the end, the cost of the American version of the option is a modest 4 cents higher in this example. □

An American Call Has No Excess Value for Non-dividend Paying Stocks

As we have just seen by example, the price of an American put can be larger than that of its European counterpart. But this is not the case for an American call on a non-dividend paying stock. The reason is that, prior to expiration, the value of such a call always exceeds its intrinsic value. To see this, let C_t, P_t, and S_t be the time t values of a call option, a put option with the same strike and expiry, and the underlying. Obviously if $S_t \le K$ there is no incentive to exercise at time t; so we may assume $S_t > K$. Since $K > Ke^{-r_f(T-t)}$, we have the following from put-call parity at time t (the remaining time to expiry is $T - t$)

$$S_t - K < S_t - Ke^{-r_f(T-t)} = C_t - P_t < C_t.$$

The difference $C_t - (S_t - K)$ is the *time value* of call. In other words, it is better to sell the option and capture the time value than to exercise early.

3.6 Option Pricing: Integrating the Lognormal

The binomial lattice pricing methodology is very powerful, simple to understand and implement, and usually adequately accurate given a sufficient subdivision of the time horizon. However it does have its shortcomings. The ending prices for the underlying are discrete and only an approximation to the continuous nature of actual prices. Further the ending price distribution is binomial, only an approximation to the lognormal distribution of actual prices. Perhaps its biggest shortcoming is that it does not provide a simple formula for computing an option price, only a recursive process.

Knowing that under the GBM model, a stock's future price is lognormal, it should be possible to exploit that knowledge for the purpose of generating option prices. In this section we show how to do so.

3.6.1 Black–Scholes Pricing Formula

Recall from Section 1.5 that the logarithm of the ending price S_T is normally distributed with mean $\alpha = \log S_0 + (\mu - \frac{1}{2}\sigma^2)T$ and variance $\beta^2 = \sigma^2 T$. See equations (1.22) and (1.23). By replacing the drift μ with the risk-free rate r_f we have the distribution for the risk-free growth of the asset. The ending price itself is lognormally distributed with density $g(y)$ given by (1.24),

$$g(y) = \frac{1}{y\beta\sqrt{2\pi}} e^{-\frac{(\log y - \alpha)^2}{2\beta^2}}.$$

The expected payoff is the integral of the payoff function $G()$ against this density,

$$\mathbb{E}(\text{payoff}) = \int_0^\infty G(y)g(y)dy. \tag{3.25}$$

Black–Scholes Price of a European Call

For a call option with strike price K the payoff function is $G(S) = \max(S-K, 0)$. Thus

$$\begin{aligned}\mathbb{E}(\text{payoff}) &= \int_0^\infty \max(y-K, 0)g(y)dy = \int_K^\infty (y-K)g(y)dy \\ &= \int_K^\infty yg(y)dy - K\int_K^\infty g(y)dy. \end{aligned} \tag{3.26}$$

We see the integral can be done in two parts. The first simplifies to

$$S_0 e^{r_f T}\Phi(d_1)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal and

$$d_1 = \frac{\log(S_0/K) + (r_f + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \tag{3.27}$$

The derivation is given in the details box on the next page. The second integral simplifies to

$$K\Phi(d_2)$$

where

$$d_2 = \frac{\log(S_0/K) + (r_f - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}. \tag{3.28}$$

Altogether the expected payoff is

$$\mathbb{E}(\text{payoff}) = S_0 e^{r_f T}\Phi(d_1) - K\Phi(d_2).$$

Discounting back to time 0 gives the call price

$$C = S_0\Phi(d_1) - Ke^{-r_f T}\Phi(d_2). \tag{3.29}$$

Black–Scholes Price of a European Put

Similarly, replacing $G(S)$ by $\max(K-S,0)$ and integrating we get the put price. First the expected payoff,

$$\begin{aligned}\mathbb{E}(\text{payoff}) &= \int_0^\infty \max(K-y,0)g(y)dy = \int_0^K (K-y)g(y)dy \\ &= K\int_0^K g(y)dy - \int_0^K yg(y)dy. \end{aligned} \tag{3.30}$$

For the first integral in (3.26) make the change of variable

$$z = \frac{\log y - (\alpha+\beta^2)}{\beta}$$

where α and β are as above. Then

$$y = e^{\beta z + (\alpha+\beta^2)}, \quad \text{and} \quad dy = y\beta dz. \tag{3.31}$$

Hence the first integral becomes

$$\begin{aligned}\int_K^\infty y\frac{1}{y\beta\sqrt{2\pi}}e^{-\frac{(\log y-\alpha)^2}{2\beta^2}}dy &= \int_d^\infty e^{\beta z+(\alpha+\beta^2)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2\beta^2}(z\beta+\beta^2)^2}dz \\ &= e^{\alpha+\beta^2}\int_d^\infty \frac{e^{\beta z}}{\sqrt{2\pi}}e^{-\frac{1}{2}(z+\beta)^2}dz.\end{aligned}$$

The lower limit of integration was $y = K$; under the change of variable it becomes

$$d = \frac{\log K - (\alpha+\beta^2)}{\beta}.$$

Combining the exponentials under the integral we have

$$e^{\alpha+\beta^2}e^{-\frac{1}{2}\beta^2}\int_d^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz = e^{\alpha+\frac{1}{2}\beta^2}(1-\Phi(d)). \tag{3.32}$$

Since $\alpha + \frac{1}{2}\beta^2 = \log S_0 + r_f T$ and noting that $1 - \Phi(d) = \Phi(-d)$, the first term in (3.26) becomes

$$S_0 e^{r_f T}\Phi(d_1)$$

where

$$d_1 = -d = \frac{\log(S_0/K) + (r_f + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}.$$

For the second integral in (3.26) use the change of variable

$$z = \frac{\log y - \alpha}{\beta}, \quad \text{so} \quad dz = \frac{1}{y\beta}dy.$$

The integral becomes

$$\begin{aligned} K\int_K^\infty \frac{1}{y\beta\sqrt{2\pi}} e^{-\frac{(\log y-\alpha)^2}{2\beta^2}} dy &= K\int_{d'}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \\ &= K(1 - \Phi(d')) = K\Phi(d_2) \end{aligned} \tag{3.33}$$

where

$$d_2 = -d' = -\frac{\log K - \alpha}{\beta} = \frac{\log(S_0/K) + (r_f - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}.$$

Notice that these integrals are just the complementary probabilities of those above for a call (3.26). Hence we can just read off the values. From (3.33) the first integral is

$$\int_0^K g(y)dy = \Phi(d').$$

And from (3.32) the second is

$$\int_0^K yg(y)dy = e^{\alpha + \frac{1}{2}\beta^2}\Phi(d) = S_0 e^{r_f T}\Phi(d).$$

Therefore from (3.27) and (3.28) the put price is

$$P = Ke^{-r_f T}\Phi(-d_2) - S_0\Phi(-d_1). \tag{3.34}$$

3.6.2 Probability the Option Ends ITM

In reviewing the derivation of these price we note that the second integral in the call price (3.26) is exactly the probability that the ending price S_T is above K, that is, in-the-money. Hence

$$\Pr(\text{call finishes ITM}) = \Phi(d_2). \tag{3.35}$$

Similarly the first integral in (3.30) is the probability that S_T comes in below K and so is the probability that the put is in-the-money,

$$\Pr(\text{put finishes ITM}) = \Phi(-d_2). \tag{3.36}$$

In either case it is the coefficient of the discounted K term in the option price.

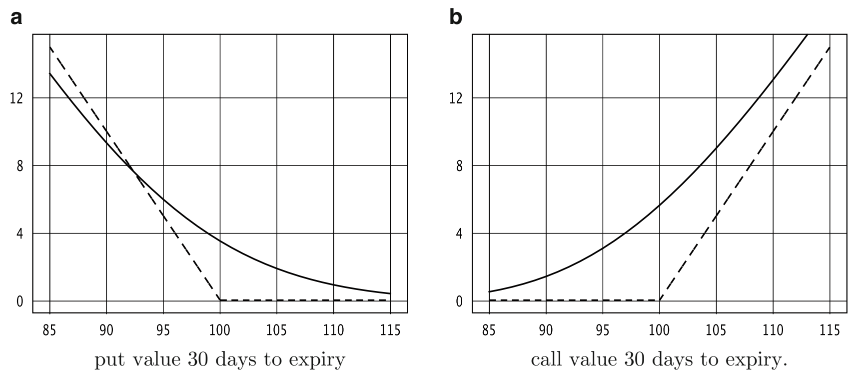


Fig. 3.9. Black-Scholes put and call values as a function of the price of the underlying, parameters: strike 100, volatility 40 %, risk-free rate 26 %

In Fig. 3.9 we show the values of a put and a call 30 days prior to expiration as determined by the Black-Scholes formulas above. The dashed curve is the payoff at expiration in both cases. These figures are quite informative.

The call price as seen in (b) lies entirely above the payoff showing that a call always has positive time value illustrating what we previously proved. By contrast, as shown in (a), a put can have negative time value. If this were an American option, the point at which the put value crosses the payoff curve indicates when the option should be exercised. The high risk-free rate of 26 % used in the figure is mainly responsible for causing the negative time value; a high rate makes an early payoff with reinvestment more advantageous than sticking with the option.

Graphically it is clear that the greatest time value for both puts and calls is at-the-money.

Figure 3.10 shows how options behave as time to expiration decays away. Three different and constant values of stock price are depicted, 5 % ITM, 5 % OTM, and ATM.

3.6.3 Interpretation of d_1 and d_2

All formulas deriving from the Black-Scholes analysis make use of the collection of terms d_1 and d_2 defined in (3.27) and (3.28). First we notice by straightforward substitution that

$$d_2 = d_1 - \sigma\sqrt{T}. \tag{3.37}$$

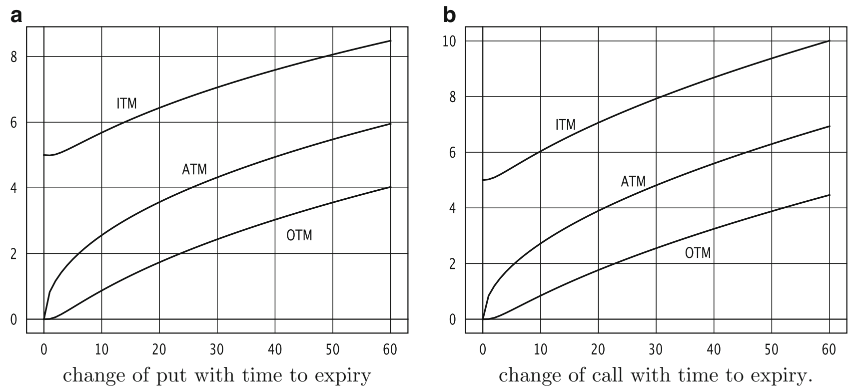


Fig. 3.10. Black-Scholes put and call values as a function of the time to expiration. Time runs down from 60 days on the *right* to 0 on the *left*

Next, among the terms for d_1 and d_2, the term $r_f T$ is the deterministic growth in the value of the stock with time (assumed to be at the risk-free rate). Similarly the term $\frac{1}{2}\sigma^2 T$ can be interpreted as the maximum stochastic growth in time while $-\frac{1}{2}\sigma^2 T$ is the maximum stochastic decline in value. Thus the terms $(r_f + \frac{1}{2}\sigma^2)T$ and $(r_f - \frac{1}{2}\sigma^2)T$ are the upper and lower envelopes for the growth in the value of the stock in some sense. On the other hand the term $\log(S_0/K)$ is the amount by which a call is initially in-the-money in geometric terms. Hence we may interpret d_1 as the upper envelope for the projected amount by which a call is ITM normalized by standard deviation. In the same way, d_2 is the lower envelope.

For a put, $-d_1$ and $-d_2$ play the same roles.

3.6.4 Implied Volatility

The volatility used in the Black-Scholes formula should be the volatility of the underlying over the interval $[0, T]$ which is assumed to be constant in the derivation. Of course this is unknown and, equally important, is unlikely to remain constant. The latter problem can be treated by using the average volatility over the time horizon. Even so, the volatility to use is generally unavailable.

One possibility is to compute the volatility of the underlying over a recent period of time, see (1.17). This is called *historical volatility.* With historical volatility in hand, the initial price of an option can be established. However, after its introduction, an option is traded through an exchange and the price is determined by the market. It then becomes possible to ask what volatility gives the current market price? This is a kind of instantaneous volatility; it is called the *implied volatility* (IV). It is the market's prediction of the volatility over the time horizon.

While there is no closed form for inverting Black-Scholes to solve for volatility in terms of the other variables, including the option's price, the formula can easily be solved numerically. A simple and adequate method is *bisection*; it is discussed in Section A.14. The method starts with an estimate of volatility that is too low,

σ_1, and an estimate that is too high, σ_2. The computed option price for σ_1 is less than the actual price but for σ_2 it is greater. In this way the root we are seeking, σ_0, is bracketed. Next the option price is calculated for the mid-volatility

$$\sigma_m = \frac{1}{2}(\sigma_1 + \sigma_2).$$

Then either σ_1 or σ_2 is set equal to σ_m depending on whichever brackets the root. The process is now iterated until the desired accuracy is attained.

3.6.5 Option Prices for Dividend Paying Stocks

In Section 2.3 we saw that dividends paid out on an equity have the effect of reducing its price by the same amount. For the holder of an option, this tends to lower the price of the stock at expiration with the result that call options should be less expensive and put options more expensive than otherwise.

Consider the case of an option over an underlying that pays dividends continuously at the fixed rate of q percent per year, see Section 3.3.2. The annual dividend rate for an equity is called the *dividend yield*. If the underlying is an index, continuous payment is a good approximation. Even if dividends are paid at discrete times, for example quarterly, assuming continuous dividends is a reasonable first order approximation, especially if the time horizon is fairly long.

To price the option we can appeal to simulation. Along these lines see Section 2.3.2 and in particular Algorithm 6. In performing the simulation we must reduce the stock's price by the amount $q\,dt$ on each time step. In fact it works the same as negative drift. But, over the entire time horizon 0 to T, reducing the price by this amount on each step has the same result as reducing the value of the stock at the start by e^{-qT} once and for all.

In both cases the ending price distribution is lognormal. From (1.28) and (1.29) we can calculate the mean and variance in each case. For a drift of $-q$ and a starting value of S_0,

$$\mathbb{E}(S_T) = S_0 e^{-qT}, \qquad \mathrm{var}(S_T) = S_0^2(e^{\sigma^2 T} - 1)e^{-2qT}.$$

For a drift of 0 but a starting value of $S_0 e^{-qT}$ we have

$$\begin{aligned}\mathbb{E}(S_T) &= (S_0 e^{-qT})e^{0T} = S_0 e^{-qT} \\ \mathrm{var}(S_T) &= (S_0 e^{-qT})^2(e^{\sigma^2 T} - 1)e^{0T} = S_0^2(e^{\sigma^2 T} - 1)e^{-2qT}.\end{aligned}$$

Hence their means and variances are the same.

This means we can use the Black-Scholes equations simply by replacing S_0 with $S_0 e^{-qT}$. We have, from (3.29) for a call,

$$C = S_0 e^{-qT}\Phi(d_1) - Ke^{-r_f T}\Phi(d_2). \tag{3.38}$$

and, from (3.34) for a put,

$$P = Ke^{-r_f T}\Phi(-d_2) - S_0 e^{-qT}\Phi(-d_1). \tag{3.39}$$

Since

$$\log(\frac{S_0e^{-qT}}{K}) = \log(\frac{S_0}{K}) - qT,$$

we have for d_1 and d_2,

$$\begin{aligned} d_1 &= \frac{\log(S_0/K) + (r_f - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \\ d_2 &= \frac{\log(S_0/K) + (r_f - q - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}. \end{aligned} \tag{3.40}$$

Put-Call Parity with Dividends

These ideas also allow us to upgrade put-call parity to account for dividends. From (3.3) we have, for a non-dividend paying stock,

$$S_t + P_t = C_t + Ke^{-r_f(T-t)}.$$

To account for dividends it is a matter of substituting for S_t as above

$$S_te^{-q(T-t)} + P_t = C_t + Ke^{-r_f(T-t)}. \tag{3.41}$$

This equation can be interpreted as saying that at time t portfolio A consisting of $e^{-q(T-t)}$ shares of stock and 1 put option is replicated by portfolio B consisting of 1 call option and $Ke^{-r_f(T-t)}$ cash invested at the risk-free rate. By reinvesting the dividends continuously as they accrue in additional shares of stock, at time T portfolio A will consist of 1 share of stock and 1 put contract. As in Section 3.2.2, A will be worth K if $S_T < K$ and S_T if $S_T \geq K$, same as portfolio B.

3.7 The Monte Carlo Method for Option Pricing

The analytical solutions derived in the previous section only apply to the simplest cases, European puts and calls. These are often called *vanilla options*. Since these equations are easily invoked, they are the first and most often to be used, sometimes under conditions at the fringe of their applicability.

By contrast, Monte Carlo is always applicable notwithstanding that in some cases its solution demands a degree of cleverness. One of these occurs when knowledge of the future is required in order to make decisions at the present time. This is the case for American options which we encounter in this section and shout options of Chapter 4. We will see how simulation can overcome the problem of advanced knowledge.

However most non-vanilla options are merely path-dependent, not requiring advanced knowledge, and can be handled straightforwardly. Before taking up exotic options, as they are called, we first consider the plain vanilla case.

If the option's price is not path dependent then it is just a matter of simulating an ending price S_T, valuing the option for that price, and discounting the value back to time 0. This being one sample, the option's price is the average of such values over a sufficiently large number of samples.

3.7.1 European Options

In fact there are two techniques for doing so: numerical integration and simulation.

Numerical Integration

From first principles, the no-arbitrage price of an option is equal to the discounted expected payoff value of the option. In turn, the expected payoff is the integral of the expiration value of the option, $G(\cdot)$, integrated against the expiration density, recall (3.25). By definition, the empirical expectation of a function $G(y)$ with respect to a random variable Y is the sum $\sum_i G(Y_i)/N$ over N samples of Y. By the *Law of Large Numbers*, this calculation is exact as $N \to \infty$.

To carry out this calculation for the option price, start with α and β; repeating (1.27)

$$\begin{aligned} \alpha &= \log S_0 + \mu T - \frac{1}{2}\sigma^2 T \\ \beta^2 &= \sigma^2 T. \end{aligned} \tag{3.42}$$

Recall that μ here is taken as the risk-free rate and σ is the volatility. Now let $Z \sim N(0,1)$ be a standard normal sample and let Y be given by

$$Y = e^{\beta Z + \alpha}; \tag{3.43}$$

then Y is the required lognormal sample. It is also an instance of the expiration stock price S_T. For a call option with strike price K the payoff function is $G(Y) = (Y - K)^+$ and for a put it is $(K - Y)^+$. Altogether the algorithm is as follows.

Algorithm 12. Numerical integration pricing algorithm

inputs: S_0, μ, σ, T, r, N
$\alpha = \log S_0 + \mu T - \frac{1}{2}\sigma^2 T$ ▷see (3.42)
$\beta^2 = \sigma^2 T$
$E = 0$
for $i = 1, \ldots, N$
  $Z \sim N(0,1)$
  $Y = e^{\beta Z + \alpha}$ ▷see (3.43)
  $E = E + G(Y)$
end for
$E = E/N$
option price = e^{-rT}E

Simulation

The difference between this method and that above is that we use the GRW to generate the ending stock price S_T. Starting with the price at S_0, generate a GRW over the time horizon to calculate an instance of S_T. Recall that we must assume the drift of the equity is the risk-free rate. Apply the payoff function to S_T to obtain the payoff for that instance. Sum these payoffs over a statistically large number of trials N and divide by N to get the expected payoff. The option price is the expected payoff discounted back to the present. As above, the payoff function is given by, $G(S_T) = (S_T - K)^+$ for a call and $G(S_T) = (K - S_T)^+$ for a put.

Algorithm 13. Monte Carlo continuous pricing algorithm

```
inputs: S_0, K, T, r, σ, N
E = 0
for i = 1,...,N
   S = S_0
   ▷use Algo.2 page 12 to generate S_T
   E = E + G(S_T)
end for
option price = e^{-rT}E/N
```

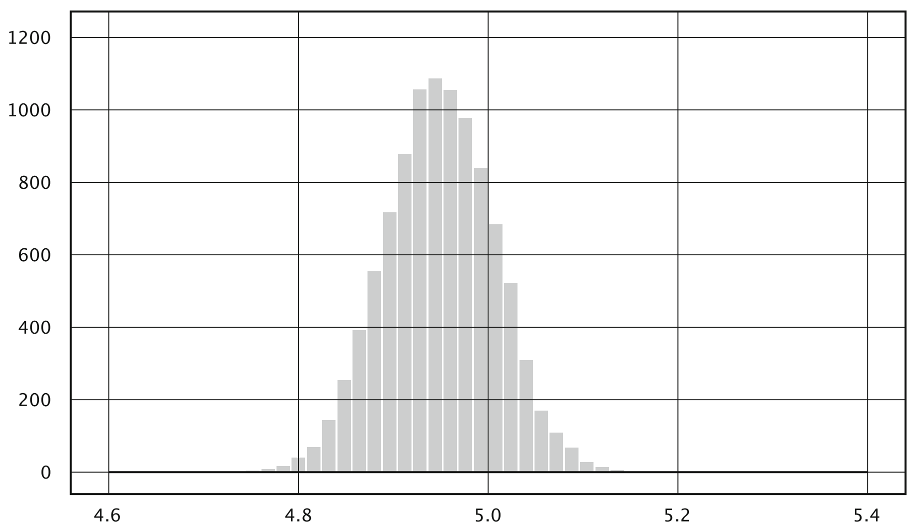


Fig. 3.11. The distribution of 10,000 price evaluations for an American put in conjunction with an optimized exercise boundary. The mean and standard deviation of these prices is 4.9479 and 0.0563 respectively. The option is the same as in Fig. 3.12. The exercise boundary parameters are also the same as in that figure

As usual, the Monte Carlo method is computationally slower than analytical methods, but it is easy to program and can be used when other methods can not.

An additional complication is that, being a stochastic method, the computed price is also stochastic. The calculated option price will depend on the exact random choices made throughout the run. A second price calculation will most likely be different. In fact, a single run of the price calculation is a sample from some probability distribution. By the Central Limit Theorem the distribution is approximately normal, see Fig. 3.11. The price we want is the mean of this price distribution. By performing several runs and reporting the average, a better estimate is provided. In addition, by calculating the variance of the trials, a confidence interval for the option prices can be worked out. These matters are treated in Section A.8.

3.7.2 The Monte Carlo Method for American Options

As we have seen, the basic binomial method requires only a straightforward and modest modification to enable it to calculate American option prices. This is because it is a backward method, option prices are first determined at expiration and then worked backwards to the starting date. Therefore at every point where an early exercise is possible, the choice is easy to make because the future value of the option is known if the option is not exercised. Further the future value in question is only between two possibilities.

However a forward method has no such future information. Consider the problem confronting a GRW at some point in the walk where the option is deep in the money. Even if the discounted expiration value of the option were known at that point, for example by using the Black-Scholes formula, the information would be insufficient since there could be further early exercise opportunities prior to expiration. Thus the value of the option at the point in question is not given accurately by its discounted expiration value.

Nevertheless, there is a region in which, if the option were sufficiently deep in the money, depending on the remaining time to expiry, then it should be exercised early. This idea gives rise to the concept of an *exercise boundary.* Figure 3.12 shows the exercise boundary for a American put option. On the ordinate is plotted the relative extent to which the option is in the money,

$$\frac{K - S_t}{K}, \tag{3.44}$$

against the time to expiration, τ, along the abscissa. At expiration itself, $\tau = 0$, the boundary point is the strike price. That is, if $K - S_T \geq 0$ the option should be exercised. As time to expiration increases, the boundary point tends to have a logarithmic character. (Although no analytical formula for it is known.) A forward method could be made to work if the exercise boundary were known or even well approximated.

If the exercise boundary were known, then a Monte Carlo method for pricing an American option would go just as for its European counterpart with the exception that the option is exercised any time the walk touches or crosses the exercise boundary. The contribution to the option price expectation is the

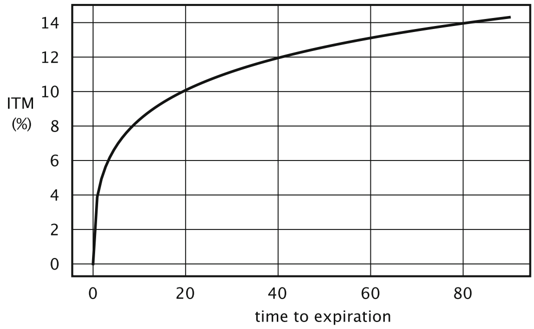


Fig. 3.12. Calculated exercise boundary for an American put option with parameters: $S = 100$, $K = 100$, $T = 90$ days, $\sigma = 30\,\%$, $r_f = 10\,\%$; option price= 5.00 (cf. European option price= 4.73). The boundary parameters for the figure are: $a_1 = 0.0243$, $b_1 = 0.6336$, $c_1 = 0.9960$, $a_2 = 0.0181$, $b_2 = 3.3632$, $c_2 = 0.4120$. See the text

exercise value discounted back to $t = 0$. The sum of such calculations over N trials divided by N is the estimated price. Note that in this calculation the discounting is applied in each trial rather than for all trials at the end.

Algorithm 14. Monte Carlo pricing algorithm for an American put

```
inputs: S_0, K, T, Δt, r, σ, N
        and the exercise boundary B(t)
E = 0; n = T/Δt
for i = 1,...,N
  S = S_0
  for t = Δt, 2Δt, ..., nΔt = T
    Z ~ N(0,1)
    S_t = S_{t-1}(1 + rΔt + σ√(Δt) Z)
    if K - S_t ≥ K * B(T - t)
      E = E + e^{-rt}(K - S_t)
      go to next i
    endif
  endfor t
  E = E + e^{-rT}(K - S_t)
endfor i
E = E/N
option price = E
```

Although pricing the option is an easy matter once the exercise boundary is known, finding the boundary itself is difficult. One approach is to treat it as a free-boundary problem in the numerical solution of the Black-Scholes partial differential equation (B.17), see [Mey98]. Sophisticated algorithms are required for the solution but they do work well in plain vanilla cases (only one underlying with no dividends). Even then, very fine meshes are needed in order to obtain good accuracy of the boundary. One confounding aspect in the nature of the problem is that good approximations of the option's price can result from even quite inaccurate boundaries.

Alternatively the boundary can be computed by the Monte Carlo method. One way to proceed is to observe that the boundary is the solution of a maximization problem. This is because the option holder's decision as when to exercise is driven by attempting to maximizing the option's value. Therefore the objective of the maximization problem is maximizing the option's value.

The method starts by proposing a parametrized form for the boundary and proceeds by maximizing on the parameters. The algorithm above can be used to calculate the option price for any choice of such parameters. The choice has to be made carefully because, in addition to the fact, as noted above, that the objective is nearly flat with respect to boundary perturbations, the Monte Carlo estimate of the option's price is stochastic. The number of paths used to calculate the option's price must be large in order that the variance of the calculation be small. Therefore it is important to start with basis functions having approximately the right shape to start with.

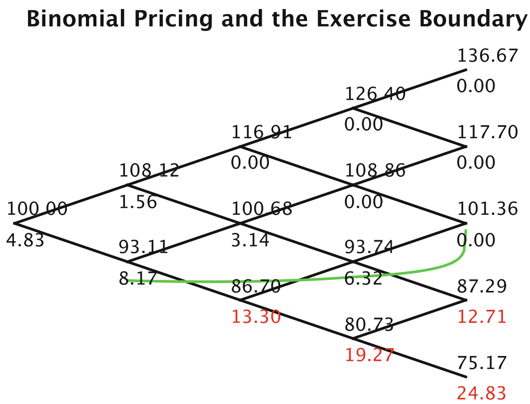


Fig. 3.13. The exercise boundary calculated by a genetic algorithm for finding the price of an American put with particulars as in Fig. 3.12. The resulting exercise boundary is the *green curve* superimposed on a 4-step binomial tree for the same problem. Early exercise nodes, all of which lie below the curve, are shown in *red*

By doing a preliminary binomial lattice calculation, as shown in Fig. 3.13, the general description of the boundary can be envisioned. It is seen to have logarithmic character with a steep slope at $\tau = 0$, followed by a sharp bend, then a very slight, nearly constant slope.

Consistent with these attributes is the function

$$a \log(bt^c + 1). \tag{3.45}$$

The parameter a effects vertical scaling and parameters b and c stretch and shrink horizontal scaling. The function of the 1 in the argument is so when $t = 0$, the function value is zero.

With the choice of two such basis functions, we try the following 6 parameter exercise boundary function

$$B(\tau) = \frac{K - S_\tau}{K} = a_1 \log(b_1 \tau^{c_1} + 1) + a_2 \log(b_2 \tau^{c_2} + 1). \quad 0 \le \tau \le T \tag{3.46}$$

Here τ is the time remaining to expiration.

In a splendid twist of synergy, the maximization problem itself can also be done by Monte Carlo, for example using simulated annealing or genetic algorithms. Figure 3.13 depicts the result obtained by a genetic algorithm for calculating the put price of a 90 day American option. Although the result is shown overlaid on a 22.5 day binomial tree, the algorithm generated prices based on continuous geometric random walks with daily increments. The \$4.83 calculated by the tree is an approximation to the \$5.00 price calculated by the continuous walk. Figure 3.14 shows that the estimated exercise boundary is accurate to the resolution of one day increments.

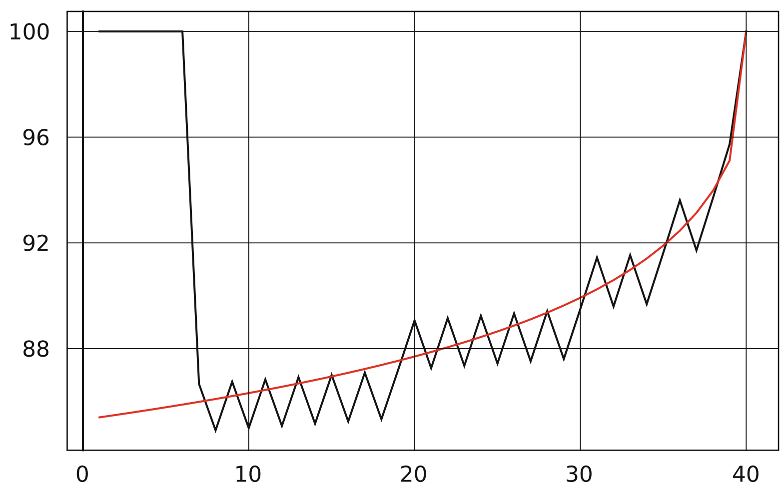


Fig. 3.14. Successive points on either side of the exercise boundary as calculated by a binomial lattice are shown in *black* for an American put. Shown in *red* is the interpolated boundary calculated by optimizing parameterized logarithmic basis functions

A Brief Look at Simulated Annealing

A genetic algorithm for American puts is provided on the web page for this text. We give here the outline of a simulated annealing algorithm. First a bit of background.

A *state* of the system in this case is a 6–component vector of positive real numbers,

$$x = (\, a_1 \quad b_1 \quad c_1 \quad a_1 \quad b_2 \quad c_2 \,)\,. \tag{3.47}$$

These components are used in (3.46) to form the exercise boundary. The *energy* of the system is a function of x and can be, for example,

$$\mathrm{E}_x = 1/\text{option price}. \tag{3.48}$$

The option price is calculated by the algorithm on page 110 using the exercise boundary given by (3.46). The option price will be maximized by minimizing the system's energy.

An important parameter of an anneal is the system temperature T. The temperature cools from some high value at the start of the algorithm to a low value at the end. Thus T is a function of algorithm time t as measured by iterations in the algorithm's main loop. Two common cooling schedules are *geometric*

$$T = ab^t, \quad a > 0, \quad 0 < b < 1, \tag{3.49}$$

and *logarithmic*

$$T = \frac{c}{\log(1+t)} \quad c > 0. \tag{3.50}$$

The parameters a and b or c are parameters of the anneal. Usually preliminary trials of the algorithm are performed in order to get an idea as to the size of ΔE and, working from there together with knowing how many trials will be allocated, appropriate values for them are determined. For example, theoretically c should be equal to the largest value ΔE can have. Knowing the number of trials determines the size of p near the end of the run; p should be less than 5 % at the end. (If geometric cooling is used, b should be quite close to 1, e.g. $b = 0.9999$ or larger, otherwise p tends to zero too quickly.)

In addition, each state x must define a neighborhood $N(x)$ about itself defining what it means to be "close" to x. New trials of the algorithm use points close to the current trial. An appropriate neighborhood for this problem is the set of vectors $y = (\, a_1' \quad b_1' \quad c_1' \quad a_2' \quad b_2' \quad c_2' \,)$ such that $\|y - x\| < \epsilon$ for some ϵ and some vector norm, for example L_1,

$$|a_1' - a_1| + |b_1' - b_1| + \ldots + |c_2' - c_2| < \epsilon. \tag{3.51}$$

With that background, we present an algorithm using geometric cooling.

Algorithm 15. Simulated Annealing Algorithm

```
inputs: a, b, neighborhoods N(x), and
        the number of iterations to perform, n
initialize x randomly and evaluate E_x
initialize temperature T
for t = 1,...,n
  from N(x) choose y at random and evaluate E_y
  put ΔE = E_y − E_x
  with probability p = e^{−ΔE/T} replace x by y
  otherwise keep x
  update the temperature, T = ab^t
endfor
report the option price 1/E_x
(and the maximizing state x if desired)
```

The key step in the algorithm is the line in which the old state x is replaced by the new one y with probability p (as achieved by drawing a computer random number r and replacing if $r < p$). If the new energy E_y is less than the old, then

p as calculated is greater than 1 and the new state is accepted. But the new state can also be accepted even if the new state is worse. And this will happen with greater probability at high temperature than low. In this way the algorithm does not become trapped in local minima.

Another complication of the present application is that the energy calculation (i.e. the price calculation) is stochastic, see Fig. 3.11. What we want to maximize on is the mean of the option prices. A way to do so is to make the reported option price an average over several individual runs. By calculating the variance of the average, a confidence interval for the option prices can be worked out. For more information on this, see Section A.8.

Problems: Chapter 3

1. Use the Monotonicity Theorem to show that all risk-free assets must have the same return rate.

2. A portfolio consists of a long 55 put and a short 50 put. Show how to replicate this portfolio using stocks and calls. Comparing the payoff graphs of the two puts, which put has greater value no matter what the stock price (i.e. is more expensive to buy)? Same question for the replicating calls.

3. In Example 3.1 on page 84 what is the value of the forward contract at 5 months if the stock price at that time is \$48?
(Ans. 4.79.)

4. A company whose current stock price is \$43.44 has been paying a quarterly dividend of \$0.77 per quarter. What is the annual yield? What should be the price of an 8 month forward contract on this stock if it is now assumed that dividends will accrue continuously at this yield? The risk-free rate is 3 %. What is the value of this forward contract after 5 months if the stock price is \$48 then.

5. Let $S_0 = 20$, $\sigma = 0.3$, $r = 0.06$ and $\Delta t = 1$ week. Construct a 2-week binomial tree (2 steps), use the $p = 1/2$ method, and calculate the no-arbitrage call price C for $K = 21$. If the market price C_M were $3/2$ times C, explain in detail how that could be exploited to make a risk-free profit.

6. Simulate the tree in Problem 5. (Start with $S = S_0$ and randomly choose "up" with probability p or "down" with probability $1 - p$, do this twice. Note the ending price S_T and note the payoff $(S_T - K)^+$.) Answer the following by Monte Carlo. Letting C' denote the actual price of the call, what is the expected gain to the option holder if $C' = C$? If $C' = 3C/2$? If $C' = C$ but $p = 0.6$?

7. (a) Use a 4-step binomial tree to price a call option with these particulars: $S_0 = 36$, $K = 34$, $r = 0.02$, $\sigma = 0.3$, $T = 4$ weeks (28 days).
(b) What is the Black-Scholes price? What is the probability the option finishes in-the-money?

8. (a) Price the option in Problem 7 by Monte Carlo. Use the Numerical Integration algorithm on page 107.
(b) Same question using Algorithm 13. How many trials are required to get 3 correct digits in both? What are the runs times for both?

9. (a) Price the call option of Problem 7 under the assumption that the volatility increases by 10 % each week.
(b) Same question under the assumption the volatility decreases by 10 % each week.

10. (a) Price the option in Problem 7 if the company has announced it will give a \$0.66 per share dividend in 7 days.
(b) Same question except the option is an index fund whose dividend yield is 7.3 %.

11. A 90 day call option with strike price \$100 is valued at \$8.23. The stock price is \$102, and the risk-free rate is $r = 3\,\%$. Write a bisection or other numerical solver to find the implied volatility.

12. A 120 day put option with strike price \$80 is selling for \$5.33. The risk-free rate is 3 % and the current stock price is \$82. Presently the VIX shows volatility at 22.6 %. Is the option over priced? under priced? or just about right?

13. Use a 4-step binomial tree to price an American put option with these financial parameters: $S_0 = 60$, $K = 60$, $r = 0.10$, $\sigma = 0.3$, $T = 90$ days. Compare with the Black-Scholes price. What is the probability that the option ends ITM (by Monte Carlo calculation)?

14. This problem is for gaining experience with simulated annealing. Write an annealer to find the minimum value of the following function defined for $0 < x < 10$,
$$f(x) = 20 - \frac{1}{3(x-3)^2 + 0.18} - \frac{1}{(x-7)^2 + .25} - \frac{6}{(x-9)^2 + 2}.$$
This is a 1 variable problem in x. Try both geometric and logarithmic cooling. Experiment with neighborhood size ϵ
$$N(x) = \{y : x - \epsilon < y < x + \epsilon, 0 < y < 10\}.$$
(Ans. the minimum occurs at $x = 3$.)

15. Use an annealer or a genetic algorithm or other stochastic optimizer to solve the American put problem stated in the caption of Fig. 3.12 page 110. (A genetic algorithm is available at the web page for this text, www.math.gatech.edu/~shenk; it is set up to use the pricing algorithm pp. 110 under the name amerputExerBoundary (for the reader to supply)).

4

Pricing Exotic Options

European put and call options are valued according to the expected price of the underlying on the expiration date of the option. This makes it easy all around to price the option at any time. The Black-Scholes formula does exactly that. The history of prices of the underlying plays no role in determining the option value.

But this is the exception as far as options go. Already the American option has an associated exercise boundary; the option is exercised if the path of prices touches it. And there are even more exotic options yet. Most of them are path dependent.

In this chapter we review some of these exotic options and show how they can be priced by Monte Carlo methods. Pricing options that depend on the price history of the underlying is a major theoretical challenge for analytical methods. In many cases Monte Carlo is the only practical solution.

The following is a partial list of exotic options along with their brief descriptions. The options marked by an asterisk have analytical pricing formulas (at least for their European version). The reference for the analysis is given in parentheses.

Asian the payoff is determined by the average price of the underlying over some pre-set period of time.

Barrier* if a trigger price is crossed it causes a pre-determined option to come into existence (knock-in) or go out of existence (knock-out) [Hull11].

Basket the underlying is a weighted average of several assets.

Bermuda the buyer has the right to exercise at a discrete set of times.

Binary* the payoff is a fixed amount of some asset or nothing at all, also called a digital option [Hull11].

Chooser* gives the holder a fixed period of time in which to decide whether the option will be a European put or call [Hull11].

Compound* an option on an option; the exercise payoff of a compound option is determined by the value of another option [Hull11].

Exchange* the holder gets the best performing out of two underlying assets at expiration [Mar78].

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_4,

Extendible* allows the holder or writer to choose, on the expiration date, to extend the life of the option by a specified amount [Lon90].

Lookback* the holder has the right to buy (sell) the underlying at its lowest (highest) price over some preceding period [Hull11].

Shout during the life of the option the holder can, at any time, "shout" to the seller that he or she is locking-in the current price, if this gives a better deal than the payoff at maturity the asset price on the shout date may be used instead of that on the expiration date.

Spread* its underlying is the difference between two specific assets [CD03].

We will discuss some of these in the following sections in terms of their Monte Carlo solutions. Even for those having analytical formulas, that solution requires their financial parameters be constant, such as the risk-free rate. But they can be solved as well by Monte Carlo under less stringent, non-constant, conditions.

In pricing exotic options by Monte Carlo, the random number generator must be of high quality.

4.1 Asian Options

In place of the price of the underlying at exercise, an Asian option uses the average price of the underlying over some pre-set period of time. For example the entire life of the option or perhaps the last 30 days before expiration. A reason for preferring Asian options in certain cases is to provide protection from price manipulation as the option nears expiration. This is a risk for thinly traded assets. Asian options also avoid the vagaries of volatility in the market. And they are cheaper than their European counterparts because, by averaging the price of the underlying, the effective volatility is much less.

The algorithm for pricing an Asian option is only mildly different from our standard pricing algorithm, Algorithm 13 on page 108. It is noteworthy that to obtain accurate results, dt must be taken to be a very small increment of time, on the order of one one-hundredth of a day or about $dt = 2.74 \times 10^{-5}$ in years. This greatly increases the run time of the GRW. The following algorithm takes about 90 seconds for $N = 100{,}000$ trials on contemporary equipment. For techniques that reduce the run time see Chapter D.

Algorithm 16. Pricing algorithm for Asian options

```
inputs: S_0, K, T, Δt, r, σ, N
E = 0  ▷expected option value
n = T/Δt  ▷number of walk steps
A = 0  ▷A = average price over entire walk
for  i = 1,...,N
   S = S_0  ▷starting price
   for  t = 1,...,n
      Z ~ N(0,1)
      dS = S(rΔt + σ√(Δt) Z)
      S = S + dS
      A = A + S
```

```
    endfor
    A = A/n  ▷average price
    E = E + G(A)
  endfor
  E = E/N
  option price = e^{-rT}E
```

Option payoffs are as usual,

$$\text{for calls} \quad G(A) = \max(A - K, 0), \qquad \text{for puts} \quad G(A) = \max(K - A, 0).$$

The algorithm relies on discrete arithmetical averaging

$$A = \frac{1}{n} \sum_{i=1}^{n} S_i,$$

but other types of averaging are also used. These include continuous (in analytical calculations) and geometric averaging, respectively

$$A = \frac{1}{T} \int_0^T S(t)dt$$
$$A = \exp\left(\frac{1}{T} \int_0^T \log(S(t))dt \right)$$

There are analytical formulas for calculating Asian options under geometric averaging.

In Table 4.1 we compare various Asian option prices with their European counterparts. One immediate observation is that as the averaging period becomes shorter at the end of the life of the option, the Asian price increases up to that of the European.

Table 4.1 Asian versus European option prices

$S_0 = 100$, $r_f = 3\,\%$, $\sigma = 20\,\%$, $\Delta t(\text{days}) = 0.01$

Type	Strike	Expiry(days)	Avg. period	Asian	European
Call	100	60	Entire	1.99	3.48
Call	100	60	Last 30 days	2.82	3.48
Call	100	60	Last 15 days	3.16	3.48
Call	95	60	Entire	5.50	6.61
Call	105	60	Entire	0.41	1.54
Put	100	60	Entire	1.75	2.99

4.1.1 Floating Strike Asian Option

The option described above is known as the *fixed strike Asian option*. There is also a variant in which it is the strike price that is averaged. In this case the put and call payoffs are, respectively,

$$\max(A - S_T, 0), \quad \text{and} \quad \max(S_T - A, 0).$$

As usual, S_T is the underlying price at expiration while A is the average underlying price over the designated period. We leave it to the reader to explore this case.

4.2 Barrier Options

In addition to the strike price, a barrier option specifies a second price as well, the barrier. The barrier can function to engage the option or to nullify it depending on the type. In the case of a "knock-out" barrier, if the barrier price is crossed, the option becomes valueless. The opposite occurs for a "knock-in" barrier, the option comes into existence.

Evidently the price of a knock-out type plus the price of a knock-in type equals the price of a plain vanilla European option. This implies that the price of a barrier option is always less than that of its European counterpart. Their reduced cost is one attraction of a barrier option.

It also implies that given the price of one of the options, say the knock-out variant, then the price of the knock-in can be easily calculated by subtracting from the price of the vanilla option as determined by the Black-Scholes formula.

In calculating the value of a knock-out barrier option by simulation there is a fundamental problem. We may and do simulate the stock price at the nodes, $t_i = i\Delta t$, $i = 1, \ldots, n$ and therefore know if the barrier is crossed at those points, but what about between the nodes? Fortunately there is a way to decide, probabilistically, whether the barrier has been crossed in this manner. The technique is called *Brownian bridges*. Let $X_t = \mu t + \sigma W_t$ be a Wiener process with drift which has the value x_{i-1} at t_{i-1} and x_i at t_i both less than the barrier B. Then the probability the process does not cross the barrier between these bridge points is given by, see [BS02]

$$\Pr(X_\tau < B,\, t_{i-1} < \tau < t_i) = 1 - e^{-2(B-x_i)(B-x_{i-1})/(\sigma^2 \Delta t)}. \tag{4.1}$$

One sees from (4.1) that in the limit as $x_i \to B$ (or $x_{i-1} \to B$) the probability of not crossing tends to 0.

The following algorithm runs the simulation, reports the ending stock price and whether the barrier was crossed or not. From (4.1), the barrier is crossed between end points with probability

$$e^{-2(B-x_i)(B-x_{i-1})/(\sigma^2 \Delta t)}. \tag{4.2}$$

But if the barrier is crossed at one of the end points, then the product $(B - x_i)(B - x_{i-1})$ is negative[1] and the exponential (4.2) is greater than 1. Hence the Brownian bridge check may be combined with the end-point check.

Algorithm 17. Pricing algorithm for a barrier option

```
inputs: S_0, K, B, T, r, σ, N, dt
E = 0; n = T/dt;
for  i = 1,...,N
    S = S_0; barrierCrossed = false;
    diff1= B − S_0;
    for  t = 1,...,n
        Z ~ N(0,1)
        dS = S(rdt + σ√dt Z)
        S = S + dS
        diff2 = B − S
        U ~ U(0,1)
        if( U < e^(−2diff1·diff2/σ²·dt) )  ▷barrier crossed
            barrierCrossed = true
        endif
        diff1 = diff2;
    endfor
    if( barrierCrossed == false )
        E = E + G(S)   ▷knock out type
    endif
endfor
E = E/N
option price = e^(−rT) E
```

Again, to obtain accurate results, dt must be taken to be a very small increment of time.

Some example barrier prices are presented in Table 4.2.

Table 4.2 Barrier versus European option prices

$S_0 = 100$, $K = 100$, $r_f = 3\,\%$

Type	Barrier	Expiry(days)	Vol.(%)	Δt(days)	Barrier price	European
Call	95	60	20	0.05	3.09	3.48
Call	95	60	40	0.005	4.04	6.70
Call	95	90	20	0.005	3.53	3.32
Call	90	60	20	0.05	3.48	3.48
Put	105	60	20	0.05	2.57	2.99

[1] Only one is negative the first time.

4.3 Basket Options

The payoff of a basket option is the weighted average of two or more underlying assets – what would be called a "basket" of assets. For example, a European style basket option has a specified strike price K and an expiration date T. The payoff is $(S_T - K)^+$ for a call or $(K - S_T)^+$ for a put where

$$S_t = \sum_{k=1}^{n} w_k S_t^k, \quad 0 \le t \le T;$$

$S_t^1, \ldots, S_t^n$ are the prices of the underlying assets, n of them in this case, and $w_1, \ldots, w_n$ are the weights, $\sum w_k = 1$.

The complication in evaluating a basket option is that the underlying assets are almost always correlated. Fortunately correlation is no problem for the Monte Carlo method. Refer back to Section 2.3.4 for a discussion on the matter.

To illustrate, we will work through a three basket problem. Let ρ_{12} be the correlation coefficient between assets 1 and 2. Similarly let ρ_{13} and ρ_{23} be the correlations between assets 1 and 3 and 2 and 3 respectively. According to (2.36) and (2.37) we may use the lower triangular matrix

$$H = \begin{bmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ \rho_{13} & h_{32} & \sqrt{\sigma_3^2 - \rho_{13}^2 - h_{32}^2} \end{bmatrix} \tag{4.3}$$

where

$$h_{32} = \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}$$

to generate the required correlated standard normal random variables. Let Z_1, Z_2, and Z_3 be independent $N(0,1)$ samples, then X_1, X_2, and X_3 given by

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = H \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix} \tag{4.4}$$

serve as the increments to the GRW.

Note that perfect correlations $\rho = 1$ or perfect anti-correlations, $\rho = -1$ must be worked out as special cases. For example, if $\rho_{12} = 1$, then $\rho_{13} = \rho_{23}$. The terms of matrix H will be $h_{11} = h_{21} = 1$, $h_{31} = \rho_{13}$, $h_{32} = \sqrt{1-\rho_{13}^2}$, and $h_{22} = h_{33} = 0$ in this case.

Algorithm 18. Pricing algorithm for a 3-basket option

```
inputs: $S_0^k$, $w_k$, $\sigma_k$, $k = 1,2,3$,
        $H$, $K$, $T$, $r$, $N$, $\Delta t$
$E = 0$; $n = T/\Delta t$;
for  $i = 1, \ldots, N$
   $S^k = S_0^k$, $k = 1,2,3$
```

```
    for   t = 1,...,n
      Z_k ~ N(0,1),  k = 1,2,3
      [X_1  X_2  X_3]^T = H [Z_1  Z_2  Z_3]^T   ▷eqn. (4.4)
      dS^k = S^k(rΔt + σ_k √Δt X_k),  k = 1,2,3
      S^k = S^k + dS^k,  k = 1,2,3
    endfor
    S_T = w_1 S^1 + w_2 S^2 + w_3 S^3
    E = E + G(S_T)
  endfor
  E = E/N
  option price = e^{-rT} E
```

In Table 4.3 we give some basket option values. There are many possible combinations to explore; only a small subset can be undertaken here. In the last column of the table we give the Black-Scholes value of an option having volatility equal to the weighted average of that of the three assets. This is for reference purposes only, the basket option is not expected to equal it.

An exception is the first table entry. Here two perfectly correlated assets with the same volatility constituting the entire portfolio should behave as a single underlying, and does. In the next row, the two assets are perfectly anti-correlated. The result is that the option value is very small. This is because the volatility of the portfolio is now nearly zero, when one asset is heading up, the other is heading down. But the reverse directions of the two do not cancel because both have positive drift, the risk-free rate. This example shows that options over portfolios of assets should cost less. And the next row shows this for arbitrary correlations.

In the next row we see that if the assets are perfectly correlated then they give the same as Black-Scholes even if they have different volatilities. On the other hand, if the assets are uncorrelated, then their option cost is suitably reduced from Black-Scholes.

Table 4.3 A sampling of basket option prices

Call, $S_0^1 = S_0^2 = S_0^3 = 100$, $K = 100$, $r_f = 3\,\%$, $T = 60\,\text{days}$

σ_1	σ_2	σ_3	ρ_{12}	ρ_{13}	ρ_{23}	w_1	w_2	w_3	Basket	BlkSch.
0.2	0.2	0.2	1	0	0	0.5	0.5	0.0	3.48	3.48
0.2	0.2	0.2	−1	0	0	0.5	0.5	0.0	0.49	3.48
0.2	0.2	0.2	0.7	0.3	−0.1	0.33	0.33	0.33	2.62	3.48
0.2	0.3	0.4	1	1	1	0.33	0.33	0.33	5.09	5.09
0.2	0.3	0.4	0	0	0	0.33	0.33	0.33	3.14	5.09

4.4 Exchange Options

The payoff of an *exchange option* is the amount by which one asset outperforms another. If the contract matches asset A versus B, then at expiration the payoff is

$$\max(A_T - B_T, 0). \tag{4.5}$$

Another way of thinking about it is that the holder is allowed to exchange one share of asset B for one share of A at expiration if A is worth more (otherwise B is retained).

Exchange options are also called *Margrabe options* after the person who first studied them or *outperformance options*.

From the standpoint of asset A, the option is a European call with exercise price equal to B_T. But from the standpoint of B, it is a European put with exercise price A_T. From the first interpretation it is not surprising that during the life of the option it never has value 0 and therefore the price of an American exchange option is the same as the European one.

The payoff of an exchange option does not depend on the path of prices of the underlying giving rise to the hope that an analytical expression can be found to price them. One elegant way to proceed is by change of *numeraire*. Numeraire refers to the basis for measuring the value of things. Normally currency is used for this purpose, but here, following [Der96], we will use shares of B.

Let $C_\$$ be the value of the exchange option in terms of dollars and C_B the value in terms of shares of B. Similarly let $A_\$(0) = A(0)$ denote the value of one share of A in dollars at the time the contract is made and let $A_B(0)$ denote the value of one share of A in terms of shares of B at that time. The exchange rate between B-shares and dollars is $B_\$ = B(0)$, that is $B_\$$ is dollars per share of B. To convert a value in B-shares to dollars, multiply by $B_\$$.

In terms of B-shares, the option contract is to exchange 1 share of B for 1 share of A at expiration, in other words the payoff is

$$\max(A_B(T) - 1, 0).$$

Therefore the value of the contract denominated in B-shares is given by the Black-Scholes call formula, a function of S_0, K, T, r_f, and σ, see Section 3.6,

$$C = BS(S_0, K, T, r_f, \sigma) = S_0 \Phi(d_1) - K e^{-r_f T} \Phi(d_2).$$

In terms of B-shares the parameters are as follows: the starting value of A is $A_B(0)$, and the strike price is 1. Let the time to expiration be T as usual. The risk-free rate must be taken in terms of B-shares – it is the dividend yield for B, denote it q_B. Finally, for the volatility, we must use the volatility of A in terms of B-shares, denote it by $\sigma_B(A)$. We will calculate this below.

From Black-Scholes then

$$C_B = BS(A_B(0), 1, T, q_B, \sigma_B(A)) = A_B(0)\Phi(d_1) - e^{-q_B T}\Phi(d_2). \tag{4.6}$$

and in terms of dollars

$$C_\$ = B_\$ C_B = B_\$ A_B(0)\Phi(d_1) - B_\$ e^{-q_B T}\Phi(d_2)$$
$$= A(0)\Phi(d_1) - B(0)e^{-q_B T}\Phi(d_2).$$

It remains only to accommodate the dividend yield of A by replacing $A(0)$ by $A(0)e^{-q_A T}$ throughout (here and in d_1 and d_2 below), see Section 3.6.5. Thus

$$C_\$ = A(0)e^{-q_A T}\Phi(d_1) - B(0)e^{-q_B T}\Phi(d_2). \tag{4.7}$$

The combination of terms comprising d_1 and d_2 refer to B-shares as the numeraire, for example

$$d_1 = \frac{\log(A_B(0)/1) + (q_B + \frac{1}{2}\sigma_B(A)^2)T}{\sigma_B(A)\sqrt{T}}.$$

Since $B_\$ A_B(0) = A(0)$, it follows that $A_B(0) = A(0)/B(0)$. Then, accounting for the A dividend rate, we have

$$\begin{aligned} d_1 &= \frac{\log(A(0)/B(0)) + (q_B - q_A + \frac{1}{2}\sigma_B(A)^2)T}{\sigma_B(A)\sqrt{T}} \\ d_2 &= \frac{\log(A(0)/B(0)) + (q_B - q_A - \frac{1}{2}\sigma_B(A)^2)T}{\sigma_B(A)\sqrt{T}}. \end{aligned} \tag{4.8}$$

As mentioned above, $\sigma_B(A)$ is the volatility of A with respect to B; it is the square root of the variance of A/B. It can be shown that this is given by[2]

$$\sigma_B(A) = \sqrt{\sigma_A^2 + \sigma_B^2 - 2\rho_{AB}\sigma_A\sigma_B}. \tag{4.9}$$

Notice that the calculation of the exchange option price does not make use of the risk-free rate. This is because the risk-neutral requirement has both equities growing at that rate and therefore the effect cancels out.

4.4.1 Non-constant Correlation

Equation (4.7) assumes the correlation coefficient ρ_{AB} is constant. When this is not expected to be a valid assumption, Monte Carlo can be used. For example it may be anticipated that the two assets will become less correlated over the time horizon of the option. An arbitrary dependence on time, $\rho_{AB} = \rho_{AB}(t)$ can be accommodated or even a dependence on relative prices. The following algorithm incorporates a time profile.

Algorithm 19. Pricing algorithm for an Exchange Option

```
inputs: A_0, q_A, σ_A, B_0, q_B, σ_B
        ρ(t), T, r, N, Δt
E = 0; n = T/Δt;
for  i = 1,...,N
```

[2] Expand the function $f(a, b) = a/b$ in a Taylor series through first order terms about the means μ_A and μ_B and take expectation.

```
A = A_0; B = B_0;
for j = 1, ..., n
   ρ = ρ(j Δt)
   Z_1 ~ N(0,1); Z_2 ~ N(0,1);
   X_A = Z_1; X_B = ρZ_1 + sqrt(1 - ρ^2) Z_2;
   dA = A((r - q_A)Δt + σ_A sqrt(Δt) X_A)
   dB = B((r - q_B)Δt + σ_B sqrt(Δt) X_B)
   A = A + dA; B = B + dB;
endfor
E = E + max(A - B, 0)
endfor
E = E/N
option price = e^{-rT} E
```

In Table 4.4 we show the results of a few runs of the algorithm. The first three use a constant correlation profile and hence, for them, the solution derived above should equal that of the simulation, and it does. We notice that as the assets are more correlated, the smaller is the option value. In the third case, the value of B starts out greater than that of A, thus A does not often exceed B at expiration and the option cost is low. In the fourth run the correlation decreases over the life of the option (from 0.8 down to 0.2). The result is that the option price behaves more like the correlation was the lower value than the upper one. In the last run the correlation increases over the life of the option. Again the result is that the option price is more like that for the higher correlation.

Table 4.4 Sample exchange option prices

$q_A = 8\,\%$, $q_B = 6\,\%$, $r = 3\,\%$, $T = 90$ days

A_0	B_0	σ_A	σ_B	Correlation	Exchange	BlkSch.
100	100	0.2	0.2	Const. at 0.6	3.24	3.24
100	100	0.2	0.2	Const. at 0.0	5.27	5.26
100	102	0.2	0.2	Const. at 0.6	1.88	1.88
100	100	0.2	0.2	Decr. 0.8 to 0.2	2.52	2.70
100	100	0.2	0.2	Incr. 0.2 to 0.8	3.47	2.70

4.5 Bermudan Options

A *Bermudan option* is one that can be exercised at any one of a set of specified times, the last one being the expiration date of the option. A Bermudan option is in this sense in-between an American and a European option.

A Bermudan option can be priced by either of the methods used for American options: the binomial tree method or maximization over a parameter set controlling an exercise boundary. Refer to Sections 3.5.4 and 3.7.2 respectively.

4.5.1 The Binomial Tree Method

The only change in the binomial method for a Bermudan option from its application in the American case is that the test for early exercise is only made at the designated exercise times, for all other nodes the tree is valued exactly as a European option. Of course the early exercise times should be among the nodes of the tree.

In Fig. 4.1 we show the binomial tree for a 20 day put option with possible exercise on days 10 and 20. The time between nodes on the tree is 5 days so the early exercise test must be made at nodal step 2. The Bermudan value of this option is \$1.788; the European is \$1.785. To get this kind of accuracy the step period for the binomial method must be on the order of 0.02 days or 500 steps per exercise period. The example as shown was chosen for illustration purposes only.

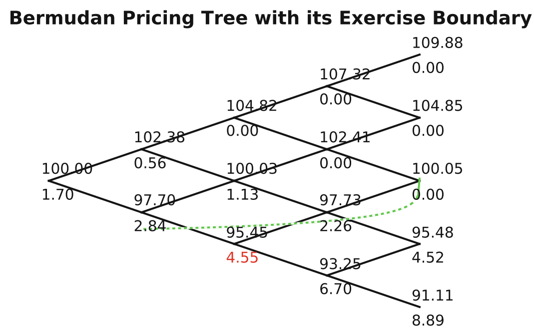


Fig. 4.1. Binomial pricing tree for a 2-exercise period put option. The 20 day Bermudan option can be exercised on the 10th day or otherwise at expiration. The binomial step size is 5 days. Those nodes for which early exercise is advantageous expresses the option's value in *red*. Superimposed on the graph is the early exercise boundary

4.5.2 The Exercise Boundary Method

As above, the only difference here from the American option case is that the test for early exercise is only made on the permissible exercise days. Additionally there are some special considerations that come into play in the Bermudan case.

Since there are only a finite number of exercise opportunities, and usually a small number, the parametrized analytical formula for the exercise boundary can be replaced by parameters giving the early exercise prices directly on the exercise days, either relativized (i.e. in the form $(K - S)/K$) or absolute. Thus for the problem in Fig. 4.1 there will be only one optimization parameter, the early exercise price on the 10th day.

Another consideration relates to the accuracy of the expectation estimates. Recall that, having fixed a trial set of parameters, whether or not they produce the maximum option value is determined by simulating a large number of random walks in order to calculate the expected payoff based on those parameters. But as these are only Monte Carlo estimates, there is inherent variance in them. Since the prices of the European versus the Bermudan options can be fairly close, it is desirable that the variance be of a smaller order of magnitude. Fortunately there are some simple remedies.

First, the random walk should not be carried out in small steps. Instead, the walk should jump from exercise day to exercise day by sampling from the appropriate lognormal distribution. This modification speeds up the simulation by many orders of magnitude. As a result, many more trials can be included toward determining the payoff expectation.

Secondly, a more discriminating objective can be used in place of the expected payoff, namely the expected payoff raised to some power. As mentioned above, in the example of Fig. 4.1 the (discounted) expected payoff is \$1.79. But the difference between 1.76 and 1.79, for example, does not discriminate between parameters sufficiently well to drive a simulated annealer or a genetic algorithm toward optimization. On the other hand $1.76^{10} = 285$ versus $1.79^{10} = 337$ has better effect.

Algorithm 20. Pricing a Bermudan Option Given an Exercise Boundary

```
inputs: S_0, K, T, r, σ, N, exercise dates t_j and the
        exercise boundary on those dates B_j, j = 1,...,n
E = 0
for  i = 1,...,N
  for  j = 1,2,...,n  ▷t_n = T
    ▷jump to next price S_j (cf. Algo. 12 page 107)
    Δt = t_j − t_{j−1}
    β = σ√Δt
    α = log(S_{j−1}) + (r − σ²/2)Δt
    Z ∼ N(0,1)
    S_j = e^{βZ+α}
    if  K − S_j ≥ K ∗ B_j
      E = E + e^{−r t_j}(K − S_j)  ▷exercise
      go to next i
    endif
  endfor j
endfor i
E = E/N
option price = E
```

In conjunction with the algorithm, as in the American case, we may use a simulated annealer or genetic algorithm to optimize the B_j's.

In Table 4.5 we make a comparison between a European, an American, and a Bermudan option; the latter calculated by both methods discussed above. The option is a 90 day 5 exercise opportunity put. With an exercise opportunity every 18 days the value is closer to that of an American versus a European option.

Table 4.5 Bermudan option comparison prices

$S_0 = 100$, $K = 100$, $r_f = 6\,\%$, $\sigma = 40\,\%$, #periods= 5, $T = 90$ days

European	American	Bermudan (binomial)	Bermudan (optimization)
7.14	7.28	7.23	7.23

4.6 Shout Options

During the life of a *shout option* the holder may lock in the current stock price for the purpose of recalculating the payoff value of the option at expiration. This is called *shouting* and the associated price is the *shout price*. At one time the shout price S_H was used in place of the expiration price S_T if it led to a bigger payoff. In such a case the payoff value for a call is

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

Thus the holder attempts to shout at the maximum price of the underlying over the option's life.

In recent times it is more common to use the shout price to replace the strike price. This is called a *reset strike shout option*. In this case the payoff value of a call is

$$\text{payoff} = \begin{cases} \max(S_T - S_H, 0), & \text{if shouting occurs} \\ \max(S_T - K, 0), & \text{if no shouting occurs.} \end{cases} \tag{4.10}$$

Now the holder of the option attempts to shout at the lowest price of the underlying for a call. The holder does not shout when the asset price is above the original strike price.

Of course the S_T and S_H or K are reversed in (4.10) for a put as usual. For a reset strike put, the holder tries to shout at the maximum underlying price over the life of the option and does not shout when the asset price is below the strike price.

In the following we shall address the problem of pricing the reset strike version of the option. This is a very difficult problem for exact solution by analytical methods because a forward method can not specify a condition for shouting since the ending price of the stock is not known, and a backward method must know if shouting occurred earlier in the course of the price history in order to calculate the ending value of the option. The author knows of no such analytical method. Instead we will solve the problem by a two phase technique similar to that of the American put option: by estimating a "shout boundary" and subsequent simulation. The boundary calculation is lengthy and we make no attempt here to shorten it, but the subsequent option valuation is very fast.

4.6.1 Maximizing Over a Shout Boundary

Once again, let the time from inception to expiration, 0 to T, be divided into n equal time steps of interval Δt. At each time step $t_i = i\Delta t$, $i = 0, \ldots, n$, let b_i be the relativized boundary point at that time (3.44). Then the actual boundary point, B_i, is given by

$$B_i = K(1 \pm b_i)$$

where the plus sign applies for a put because the boundary is above the strike price in this case, and the minus sign applies for a call, because the boundary is below the strike for a call. To simplify the subsequent discussion assume we are pricing a put option.

As was the case for an American option, we proceed in reverse order and first consider the *conditional* boundary point at t_{n-1}, meaning the boundary point given that shouting has not previously occurred. It is easy to see that here the boundary point must be the strike price K. If the stock price at this step exceeds K, shouting resets the strike to a higher value which can only increase the payoff and make a positive payoff more likely. Further there is no penalty for doing so here. At earlier times, the restraint for shouting is that the stock price might go higher before expiration; that does not apply on this penultimate time step. Hence the relativized boundary point at $i = n - 1$ is $b_{n-1} = 0$.

Now consider the next time step proceeding in reverse order and again we assume shouting has not yet occurred. The higher the stock price, the more valuable to shout; if shouting at the price S leads to an improved expected payoff and $S' > S$, then shouting at S' leads to a bigger one. Hence the minimum (technically infinimum) of all those points where shouting leads to an improved expected payoff is the boundary point.

Finding this point is a straightforward one variable unimodal optimization problem. If the boundary point is set too low, then, stochastically, subsequent stock prices can allow for a later shout with an even better expected payoff. Similarly if the boundary point is set too high then the stock price might reach this level too infrequently to have a larger expected payoff than a lower value. The effect is shown in Fig. 4.2. This is a plot of expected payoff as a function of various trial locations of the boundary point all at the same time before expiration. Although the data has considerable stochastic variability, it clearly shows a unimodal maximum in the vicinity of $S = 102$.

Finding the maximum of such data numerically is problematic. It becomes much easier if the data is smoothed as shown in Fig. 4.3. The smoothing used in the figure is a 11 window central moving average, each smoothed value m_i is the sum of 5 prior values, the current value, and 5 future values all divided by 11,[3]

$$m_i = \frac{1}{11}(p_{i-5} + \ldots + p_{i-1} + p_i + p_{i+1} + \ldots + p_{i+5}).$$

The simulation of stock prices from the present step to expiration uses the boundary points that have already been calculated. The objective calculation is shown in Algorithm 21. Note that in the algorithm we use the absolute boundary values B_i. Further the order of the boundary values is reversed from that in the discussion above; B_0 is the boundary value at expiration and B_n is the value at $t = 0$. The algorithm first calculates the number of steps m to expiration; τ is the remaining time to expiration. Lognormal samples will be used to assign stock prices from step to step and the parameters α and β are calculated; α must remain a function of stock price and be recalculated from step to step.

In each trial, the starting stock price is drawn from a range of possibilities above and below the strike price. More exactly, the range should extend

[3] This is a discrete example of *convolution* smoothing. The general form is $m(t) = \int_{-\infty}^{\infty} f(t-s)k(s)ds$ where $f()$ is the function to be smoothed, $m()$ is the smoothed version and $k()$ is the smoothing kernel. In the discrete analogy here $k(s) = 1$ for $-5 \leq s \leq 5$ and 0 otherwise.

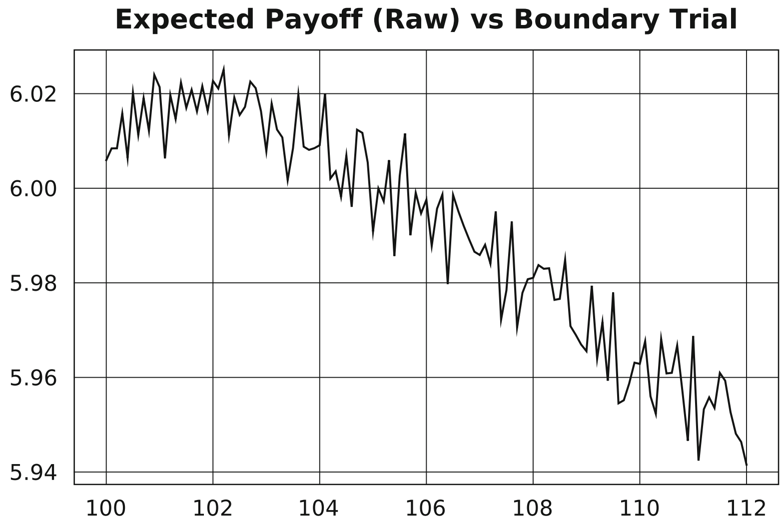


Fig. 4.2. Expected payoff as a function of increasing the conditional shout boundary point at $t = 18$ days for a put option. The option's particulars are: $K = 100$, $T = 36$, $r = 3\,\%$ and $\sigma = 20\,\%$. These are raw simulation results over the remaining 18 days to expiry. Each plotted point represents 1,000,000 simulations

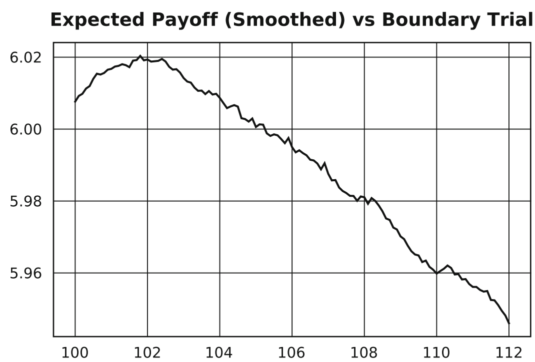


Fig. 4.3. This figure uses the same data as in Fig. 4.2 but here the raw data has been smoothed using a 11 window central moving average. It is much easier to determine the trial boundary value at which the maximum occurs

above and below the boundary point being tested. If paths do not encounter the boundary, then their payoff will mimic that of a European put.

Each new trial also begins with a "noshoutyet" variable set to true and the shout strike K_s set to K. Upon encountering the boundary, and only for the first time, K_s is reset to the current stock price, otherwise it remains at K. In either case the put payoff is calculated as $\max(K_s - S_T, 0)$ as usual.

Having determined the maximizing boundary value at step m from expiration, in like manner processing continues to step $m + 1$ and finally ends with the boundary at $t = 0$. At this point the option value itself can be calculated. Algorithm 21 can also be used for this provided the starting range is collapsed to 0 around S_0.

Algorithm 21. Monte Carlo objective calculation for a shout put

```
inputs: K, τ, Δt, r, σ, nTrials
          and the shout boundary B_i
m = τ/Δt      ▷number of steps to expiry
β = σ√Δt      ▷for lognormal samples
α(·) = log(·) + (r − ½σ²)Δt      ▷α = α(S)
V = 0;
for k = 1,..., nTrials      ▷loop over trials
  S ~ uniform over a starting range
  K_s = K      ▷set the shouting strike equal to K initially
  noShoutYet = true      ▷keep track of shouting
  for i = 0, 1, ...
    if noShoutYet
      if S≥ B_{m−i}      ▷reset the strike
        K_s = S; noShoutYet = false;
      endif
    endif
      ▷take the next step
    S ~ LN(α(S), β)   ▷lognormal sample
    i = i + 1
    if i == m, break out of loop      ▷expiry stock price
  endfor i
  V = V + max(K_s − S, 0)      ▷payoff for this trial
endfor
V = V/nTrials
```

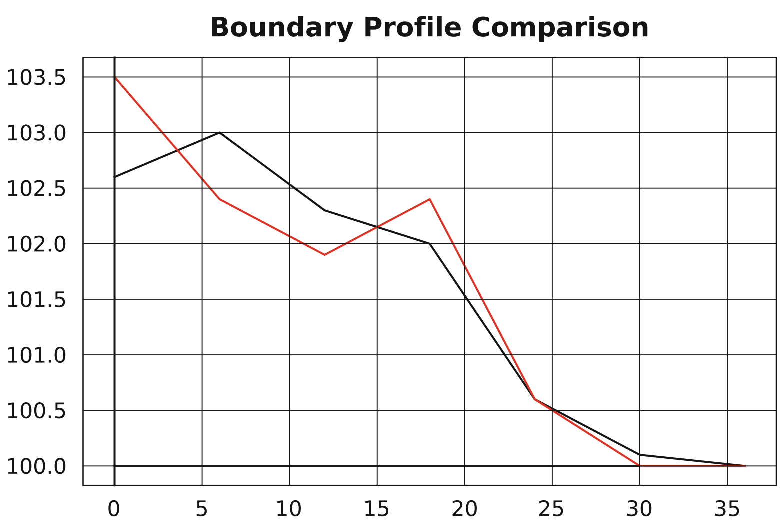


Fig. 4.4. Shout boundary as calculated by two runs of the method described in the text. The time horizon is divided into six periods

Two boundary calculation runs are shown in Fig. 4.4 for the $S_0 = K = 100$, $T = 36$ day shout put option with six division periods. As seen there, the calculated boundary points have considerable variance but despite that the option value is stable and has low variance. This phenomenon was previously noted in Section 3.7.2. In Table 4.6 we give some shout option prices calculated by the boundary method described above along with their statistical standard deviations. These are compared with the Black-Scholes prices for their European counterparts.

Table 4.6 Sample shout option prices

$S_0 = 100$, $K = 100$, $r_f = 3\,\%$, $\sigma = 20\,\%$, $T = 36$ days

Put/ call	Option value (standard deviation (10 trials)) versus number of time steps				
	3	6	12	24	European
Put	2.986(0.005)	3.059(0.004)	3.091(0.003)	3.104(0.005)	2.36
Call	3.253(0.005)	3.329(0.006)	3.375(0.005)	3.374(0.005)	2.65

Problems: Chapter 4

1. Write a program to calculate Asian options. Try it out for a 60 day ATM call option with $S_0 = 100$, and $r = 3\,\%$. Let the averaging take place over the last 30 days. Plot the option price as a function of volatility.

2. Repeat Problem 1 for a floating strike Asian option.

3. Write a program to calculate correlated basket options. Extend the results of Table 4.3 to $T = 90$ days.

4. Price a 90 day 100 strike Bermudian option with 15 day early exercise periods. Assume $r = 1\,\%$ and $\sigma = 20\,\%$. Use the binomial tree solution method. Plot the price of the option versus originating stock price. Compare the graph with that of its European counterpart.

5. Same question as Problem 4 but use the exercise boundary method.

6. Find the price of a 365 day exchange option between equities A and B. Assume $r = 6\,\%$, $\sigma_B = 20\,\%$ and the current price of B is \$60. Plot the price as a function of the current price of A for $\sigma_A = 15$, 30, and 45 %. Assume that neither A nor B issues dividends.

   In the following problems, create a calculator for the given exotic option and use it to calculate a table of prices for various option parameters.

7. A barrier option.

8. A binary option.

9. A chooser option.
10. A lookback option.
11. A spread option.

5

Option Trading Strategies

Although originally conceived as a tool to alleviate risk, once introduced, options quickly became a vehicle for speculative investment. For the trader who believes himself or herself capable of predicting the direction of the market, either short or long term, puts and calls offer *financial leverage*. For a fraction of the cost of owning 100 shares of stock, an option contract bestows a kind of temporary ownership suitable for profiting should the prediction prove out.

Moreover, combinations of the basic options provide various and unique investment possibilities. For example, certain combinations take advantage of changes in volatility independent of the direction of the market. In another instance, known as *delta hedging*, by holding a carefully calculated balance of stocks and options, the trader makes money by capturing the drift of the underlying without risk (theoretically, unfortunately trading costs outpace the profits). In fact we used delta hedging in Section 3.4 to derive the arbitrage free price of a call for a one step binomial model.

Delta hedging shows the importance of the *Greeks*. These are the mathematical derivatives of the value of a portfolio with respect to the Black-Scholes variables: stock price S, time to expiration τ, volatility σ, and the risk-free rate r. The Greeks are an important tool for trading in options.

Also important are the predictions as to: whether an option combination will expire ITM, the expected profit, and the strategy's risk. The answers can be found by simulations based on the GBM model for stock prices.

5.1 Related Option Trades

As previously explained, an option is a trade between two parties. Except for the commission (a payment to the broker) and the bid-ask spread (essentially a payment to the market maker and the exchange), what one party makes from the trade, the other party pays. Therefore related to every trade is the same trade with the roles reversed. We refer to this as the *reverse trade*.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_5,

5.1.1 Reverse Trades

Even a complicated trade consisting of any number of puts and calls with arbitrary strikes and including the underlying can be easily reversed: simply interchange all buys and sells and, for the underlying, interchange long and short.

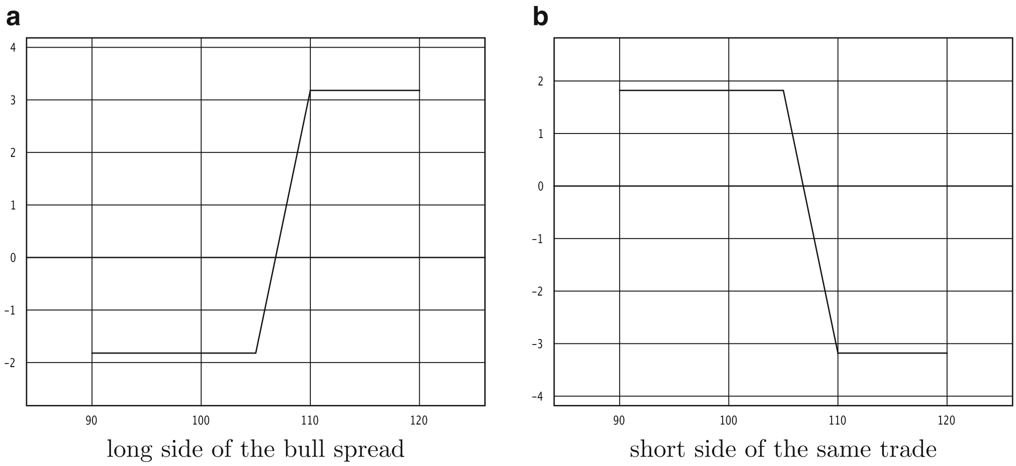


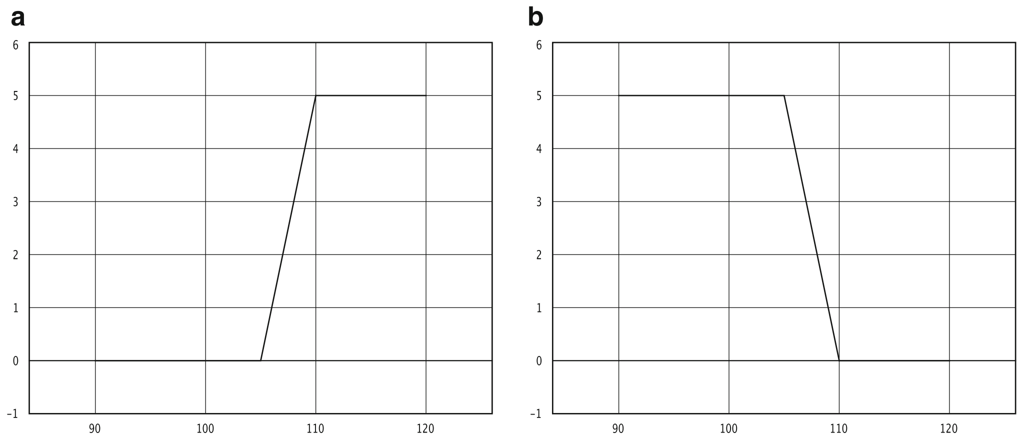
Fig. 5.1. Net payoff for a debit call spread with strikes \$105 and \$110. The price of the \$105 call is \$2.74 and the price of the \$110 call is \$0.92

For example, a *debit call spread* is the trade in which one *buys* a call at some strike price K_ℓ and *sells* a call at a higher strike price $K_h > K_\ell$. The net payoff (including the net premium) at expiration for this trade is shown in Fig. 5.1a. Notice that for ending prices less than \$105, the graph is \$1.82 below the horizontal axis, the S-axis. And for ending prices above \$110, the payoff is $-1.82+5.00 = 3.18$. The reverse trade consists in *selling* a call at K_ℓ and *buying* one at K_h. This is the other side of the trade. In all likelihood it will not be a single person taking the other side of a given trade, instead the broker puts multiple trades together from multiple sources. But it is sometimes helpful to imagine that a “virtual person” has done so. The payoff graph for the reverse trade taken by the virtual person is shown in Fig. 5.1b. Notice that it is the reflection of the original trade in the *horizontal* axis. For $S < 105$, the payoff is \$1.82 and for $S > 110$ the payoff is $\$-3.18$.

This property is true in general: the payoff graph of the reverse trade is the reflection in the horizontal axis (the S-axis) of the original payoff graph.

5.1.2 Dual Trades

Another possibility for a related trade is to swap puts and calls. With careful handling of the strike prices, this leads to a payoff graph that is vertically symmetric with the original provided option costs are excluded and provided all options expire at the same time. We call this the *dual trade*. (Option costs can be included by using the overall cost of the trade. We leave it to the reader to see how this is done.)



Debit Call spread Debit Put Spread

Fig. 5.2. Payoff graphs for debit call and put spreads with strikes \$105 and \$110 (excluding option costs). The line of symmetry was taken at $S = 107.5$

Pick a point A on the S-axis and reflect all the strike prices across this point. Thus the strike K is reflected to $K' = 2A - K$. A must be chosen so that all reflected strikes are positive. The point $S = 0$ is reflected to $2A$ (and serves as a value larger than any strike; this point is useful in proving the symmetry). Now to create the dual, every call becomes a put with the reflected strike price and conversely, every put becomes a call with the reflected strike price. For the underlying, swap long for short and conversely and reflect S_0 to S_0'. Then the payoff graph of the dual trade is the reflection of the original in the vertical line at $S = A$. In Fig. 5.2a, b we show the graphs of the debit call spread from above and its dual.

Vertical symmetry is not possible under duality for option combinations with differing expiration dates because calls always have positive time value but puts can have negative time value. Nevertheless we extend the notion of duality to include combinations with differing expiration dates.

5.2 The Greeks

The Black-Scholes formulas show that the arbitrage free value of a put or call depends on: the price of the underlying S, the time remaining until expiration τ, the volatility of the underlying σ, the risk-free rate r, and the strike price K. Once an option contract has been made, all of those variables except the strike price change from moment to moment. Consequently so does the value of the option. The *Greeks*, collectively, are the mathematical derivatives of the option price with respect to these variables. Therefore each Greek predicts how strongly the option price will change for a unit change in its variable and in which direction. In calculating each Greek, the position is assumed to be long the option. If one's position is short the option, then the Greek is oppositely valued.

5.2.1 Delta (Δ), the Derivative with Respect to S

Delta is the most important Greek partly because price changes in the underlying are relatively larger and more frequent than for the other variables and partly because its effect on the option price is greater. An analytical formula for delta for a put follows directly by differentiating (3.34) to get

$$\Delta_P = \frac{\partial P}{\partial S} = -\Phi(-d_1) = \Phi(d_1) - 1 \tag{5.1}$$

and (3.29) for a call

$$\Delta_C = \frac{\partial C}{\partial S} = \Phi(d_1) \tag{5.2}$$

($\Phi(\cdot)$ is the normal cumulative distribution function and d_1 is defined in (3.27)). Note that delta is the *partial* derivative of the option price; thus its prediction pertains when there is a change in stock price only, the other variables being fixed.

It is immediately seen that

$$\Delta_P = \Delta_C - 1. \tag{5.3}$$

This result also follows directly from put-call parity.

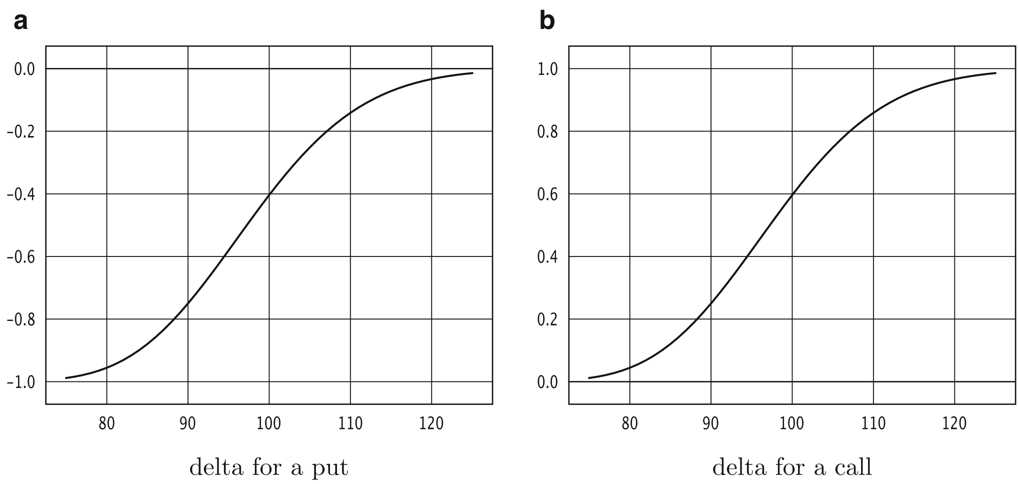


Fig. 5.3. Delta as a function of stock price 30 days before expiration. The strike price is $K = 100$ in both figures. Delta for the put is negative while for a call it is positive

From a traders point of view, the value of delta tells by how much the option will change give a unit change in the underlying price; that is by a dollar if the option price is given in dollars. Under the conditions pertaining in Fig. 5.3, when the stock price is 100, the delta of a long call is 0.6; the call increases by 60 cents for every dollar increase in the stock price, approximately.

It is easy to get a good sense about delta since it is the slope of the option payoff graph by definition. For example, delta for a put must be negative because as the price of the underlying increases, the payoff of a put decreases. By contrast, as the price of the underlying increases, the value of a call increases. Thus delta for a call is positive.

Continuing along these lines, from Fig. 3.9a (page 103) we see that for a deep ITM put, the slope is asymptotically equal to that of the expiration payoff graph, namely -1. At the other end of the figure, for a deep OTM put the slope of the graph is asymptotically 0. So as a function of S, delta for a put increases from -1 to 0.

From Fig. 3.9b we see that for a deep ITM call the slope tends to $+1$ and for a deep OTM call the slope tends to 0. Hence delta for a call increases from 0 to $+1$. This is exactly what we see in the graphs of delta versus underlying price shown in Fig. 5.3a, b.

Making a Portfolio Delta Neutral

The delta of a share of stock is 1 since $\partial S/\partial S = 1$. But by adding an option, delta of the portfolio can be adjusted. Since the Greeks are mathematical derivatives, it follows that the delta of a portfolio is just the sum of the deltas of each constituent weighted according to its numbers. For example, under the conditions pertaining in Fig. 5.3, when the stock price is 100, a portfolio consisting of 1 long call contract (representing 100 shares of stock) and short 60 shares of the underlying is *delta neutral* meaning its delta is 0.

The delta of even the most complicated portfolio can be calculated provided the delta of each constituent is available. Then, as above, shares can be bought or sold short as required to achieve delta neutrality.

Achieving delta neutrality is an important goal for many traders because it insulates the portfolio against (small) changes in the price of the underlying constituents in the same way that zero derivative points are the stationary points of a graph. Recall that in Section 3.4 we derived the arbitrage free price of an option by considering a simple delta neutral portfolio. After a single time step the value of the portfolio was the same independent of whether the stock price went up or down.

Example 5.1. The price of an ATM 30 day put contract with strike price 60 is 2.69 per share; this for a volatility of 40 % and risk-free rate of 3 %. To find delta first figure d_1,

$$d_1 = \frac{\log(60/60) + (0.02 + \frac{1}{2}0.4^2)(30/365)}{0.4\sqrt{30/365}} = 0.071673.$$

then from (5.1)

$$\Phi(0.071673) - 1 = 0.5285 - 1 = -0.4714.$$

The purchase of 47 shares of stock will make this portfolio delta neutral.

Suppose that after two days the stock price increases to \$61. The value of the option is now down to 2.16 from 2.69. Without the stock, this is a loss of 0.53 times 100 or \$53. The stock has gained \$1 per share or \$47 for the portfolio, so with the stock, the net position is a loss of \$6.

Suppose instead the stock price decreases to \$59. The option value becomes 3.11, a gain of 0.42 or \$42 for the contract. Meanwhile the stock has lost \$1 per share or \$47 to the portfolio. Together with the stock the portfolio's net change is −\$5. (A loss in both cases because 2 days worth of time value slipped away.) With the stock, the swing in price is about \$5 instead of \$50 in both cases. □

When stock prices change continuously, delta neutrality does not last very long. It becomes necessary to re-balance the portfolio by buying or selling stock until it is regained. This is called *delta hedging*. If it were possible to delta hedge cost free, then, as we saw in Section 3.4, the portfolio would earn at the rate of the drift of its constituents without risk.

5.2.2 Gamma (Γ), the Second Derivative with Respect to S

The Greek parameter gamma is the second derivative of the portfolio value with respect to S. Being the second derivative, gamma is therefore the first derivative of delta with respect to S and predicts how delta will change with stock price. Deriving the put and call formulas for gamma is a straightforward exercise; and they are equal, both being the normal density function evaluated at d_1 times the derivative of d_1,

$$\Gamma = \frac{\partial \Delta}{\partial S} = \frac{\phi(d_1)}{S\sigma\sqrt{\tau}} = \frac{e^{-d_1^2/2}}{S\sigma\sqrt{2\pi\tau}}. \tag{5.4}$$

It can be seen from this that gamma is always positive (if long the option, negative if short the option). This means that if S increases, then delta also increases for both puts and calls. This is seen in Fig. 5.3a, b in which the curves have positive slopes at every point.

Although delta is dimensionless, gamma has units of reciprocal units of currency.

Achieving Gamma and Delta Neutrality

The importance of gamma is that it predicts how quickly delta will change with respect to stock price and therefore how frequently the portfolio will have to be rebalanced to remain delta neutral. Best possible is if the portfolio is both delta neutral and gamma neutral. It is possible to achieve such a thing but only by adding a position in another option, either long or short as necessary, because the gamma of a stock is 0 since a stock position has a constant delta of 1 (or −1 if short).

Denote by Γ and Δ the gamma and delta of the portfolio we want to bring to delta–gamma neutrality. Let the gamma and delta of an option we will use to

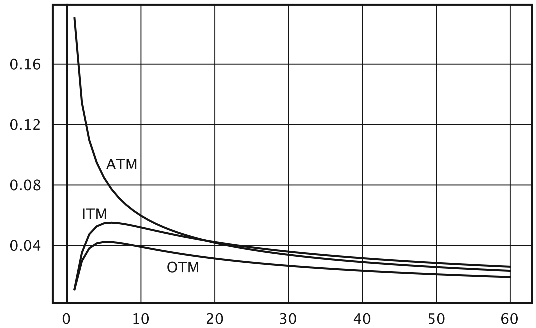


Fig. 5.4. Variation of gamma vs time to expiration depending on whether the underlying is in-the-money, out of-the-money, or at-the-money

achieve neutrality be Γ_U and Δ_U respectively. We will have to go long w units of the option (or short if w is negative) where w solves the equation

$$w\Gamma_U + \Gamma = 0, \qquad \text{thus} \quad w = -\Gamma/\Gamma_U. \tag{5.5}$$

But now the delta of our portfolio changes to $\Delta + w\Delta_U$. We can correct this by going short this many shares of stock.

Example 5.2. In Example 5.1 above the gamma of the put option is, from (5.4),

$$\Gamma_P = \frac{e^{-0.071673^2/2}}{60 * 0.4\sqrt{2\pi(30/365)}} = 0.057832.$$

So the gamma of the portfolio is 5.78.

Let us add a short 60 day ATM call contract with strike price 60. Its gamma is

$$d_1 = \frac{0 + (0.02 + \frac{1}{2}0.4^2)(60/365)}{0.4\sqrt{60/365}} = 0.1014$$
$$\Gamma_C = \frac{e^{-0.1014^2/2}}{60 * 0.4\sqrt{2\pi 60/365}} = 0.040789.$$

From (5.5)

$$w = -\frac{0.057832 \times 100}{0.040789} = 142;$$

the short call contract should be for 142 shares to create a gamma neutral position.

Use (5.2) to figure the delta of the call.

$$\Delta_C = \Phi(0.1014) = 0.5404.$$

Therefore the combined delta of the two options is

$$-0.4714 \times 100 - 0.5404 \times 142 = -128.88.$$

Going long 129 shares of stock will create a delta-gamma neutral portfolio.

As above, suppose that after 2 days the stock price increases to \$61. The value of the put option falls to 2.16 from 2.69 or by 0.53. The value of a long call increases by 0.49 to 4.46 from 3.97. Since the stock value increases by 129, the net portfolio change is

$$-0.53 \times 100 - 0.49 \times 142 + 129 = 6.42.$$

If instead the stock price decreases to \$59, the put option gains 0.42 per share as before, the call option falls 0.59 to 3.38 from 3.97, and the stock has lost \$1 per share; the portfolio's change will be

$$0.42 \times 100 + 0.59 \times 142 - 129 = -3.22.$$

□

Gamma Asymptotics

As expiration draws closer, gamma tends to 0 for ITM or OTM positions. This is because delta tends to either ± 1 or 0 in these cases. But for ATM options gamma tends to infinity. This is because the delta of a put (call) option changes from -1 $(+1)$ ITM to 0 OTM very quickly near expiration for S ATM. See Fig. 5.4.

5.2.3 Theta (Θ), the Derivative with Respect to τ

Theta is the derivative of the option price with respect to time to expiration. Since expiration time runs down as real time moves forward, the convention is to put theta equal to the negative of the derivative. Thus while the slopes in Fig. 3.10 (page 104) are positive, theta itself is negative in those examples. This is in line with the fact that most options lose value as expiration approaches (via loss of time value).[1] Thus theta expresses the *time decay* of a portfolio.

Time is usually measured in years when using the Black-Scholes formula. For example the risk-free rate and the volatility are expressed in annual terms. Therefore theta gives the change in portfolio value per year. To get a per day value, the per year value must be divided by the number of days in a year. (A calendar year is 365 days of course but sometimes the number of trading days in a year is used, that number is approximately 252 depending on the year).

Theta may be calculated as follows: for a put

$$\Theta_P = -\frac{\partial P}{\partial \tau} = -\frac{S\sigma\phi(d_1)}{2\sqrt{\tau}} + rKe^{-r\tau}\Phi(-d_2) \tag{5.6}$$

and for a call

$$\Theta_C = -\frac{\partial C}{\partial \tau} = -\frac{S\sigma\phi(d_1)}{2\sqrt{\tau}} - rKe^{-r\tau}\Phi(d_2). \tag{5.7}$$

[1] As we saw in Section 3.5.4 and in Fig. 3.9 (page 103) a possible exception is in-the-money put options whose time-value may actually be negative. Equation (5.6) shows exactly when theta is positive.

The differentiation to obtain these equations is somewhat involved and requires use of the algebraic identity

$$S\phi(d_1) = Ke^{-r\tau}\phi(d_2). \tag{5.8}$$

We leave this as an exercise for the reader.

In Fig. 5.5 we show the variation of theta with time to expiration. The most important feature of the graphs is that, for ITM and OTM options, time decay reaches a maximum about 5 days before expiration. For an option trader *selling* options in order to capture time value, the trade should be made before this critical time.

To understand these graphs, refer to Fig. 3.10 (page 104). As time to expiration approaches zero, both ITM options and OTM options become unlikely of finishing much differently than their current intrinsic value. As a result, their time value rapidly tends to zero. Thus theta makes a large negative move. After this, theta turns around and tends to its limiting value as $\tau \to 0$. For an OTM option, this is 0. For a put (call), theta tends to rK $(-rK)$.

To obtain these limits, first note that as $\tau \to 0$, d_1 and d_2 asymptotically equal $\log(S/K)/(\sigma\sqrt{\tau})$. Since, for $c > 0$, $\lim_{\tau\to 0} e^{-c\tau}/\sqrt{\tau} = 0$, it follows that the term $S\sigma\phi(d_1)/(2\sqrt{\tau}) \to 0$ as $\tau \to 0$. The second term for theta in each case tends to the limits noted above.

For ATM options the situation is much different. We notice from Fig. 3.10 that the graph becomes quite steep at 0. This is because even small changes in S could move the option from expiring worthless to expiring with a positive value. The steep slope of the option's value is due to the as yet high probability of ending in positive territory, but all the while the remaining time is nearly zero. Analytically, since $S = K$ if ATM, $\log(S/K) = 0$, so both d_1 and d_2 tend to 0. In turn $\phi(d_1)$ and $\Phi(\pm d_2)$ tend to the finite limits $1/\sqrt{2\pi}$ and $1/2$ respectively.

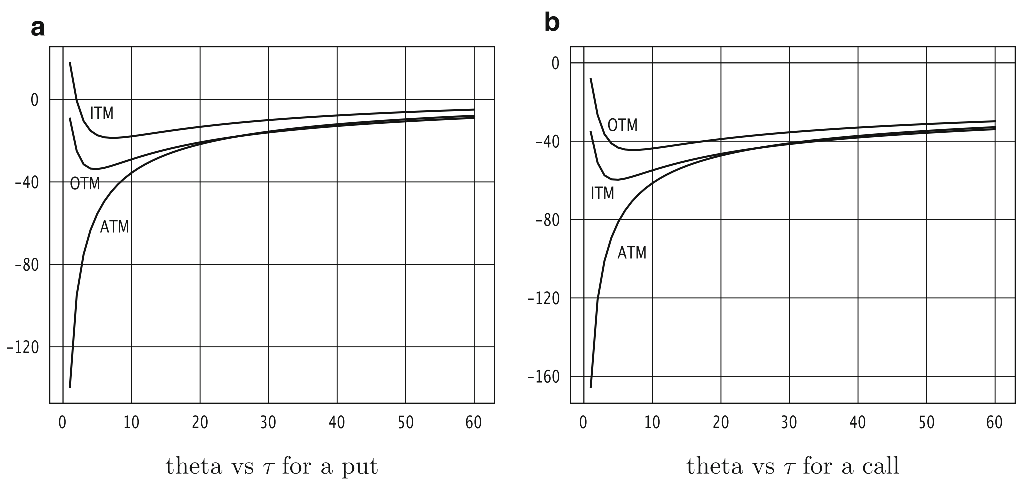


Fig. 5.5. The behavior of theta as time to expiration tends to zero for in-the-money, out of-the money, and at-the-money puts and calls. The parameters here are the same as those prevailing in Fig. 3.9 on page 103

It follows that $S\sigma\phi(d_1)/(2\sqrt{\tau}) \to \infty$ in the limit at 0 while the second term for theta remains bounded.

5.2.4 Vega (ν), the Derivative with Respect to σ

Vega is the derivative of the option price with respect to volatility. Actually volatility rarely moves by one full point; therefore the convention is to express vega as the change in option price for a 1 % change in volatility. Hence vega is usually reported as one one-hundredth times the derivative.

To get an expression for vega for a call, differentiate (3.29) with respect to σ

$$\nu = \frac{\partial C}{\partial \sigma} = S\phi(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{-r\tau}\phi(d_2)\frac{\partial d_2}{\partial \sigma}.$$

From (3.27) we get

$$\frac{\partial d_1}{\partial \sigma} = \sqrt{\tau} - \frac{d_1}{\sigma^2\sqrt{\tau}}$$

and, since $d_2 = d_1 - \sigma\sqrt{\tau}$,

$$\frac{\partial d_2}{\partial \sigma} = -\frac{d_1}{\sigma^2\sqrt{\tau}}.$$

Substituting

$$\begin{aligned}\nu =& S\phi(d_1)(\sqrt{\tau} - \frac{d_1}{\sigma^2\sqrt{\tau}}) - Ke^{-r\tau}\phi(d_2)(-\frac{d_1}{\sigma^2\sqrt{\tau}}) \\ =& S\phi(d_1)\sqrt{\tau} - \left(S\phi(d_1) - Ke^{-r\tau}\phi(d_2)\right)\frac{d_1}{\sigma^2\sqrt{\tau}}.\end{aligned}$$

But from the identity (5.8) the second term vanishes. Hence vega for a call is

$$\nu = S\sqrt{\tau}\phi(d_1) = S\sqrt{\frac{\tau}{2\pi}}e^{-d_1^2/2}. \tag{5.9}$$

By direct calculation, this is also vega for a put; the two are the same. Alternatively it follows immediately from put-call parity that they are the same.

An immediate fact about vega is that it is always positive; thus all option prices go up as volatility increases. In Fig. 5.6 we show how vega varies as time to expiration tends to zero. The figure shows that vega is largest for ATM options and falls off rapidly as the underlying falls away from the strike on either side. The figure also shows that vega tends to zero in all cases with time to expiration. In the case of an ATM option, a high vega value is maintained right to the end and then drops off rapidly over the last few days.

Unsurprisingly, it can be desirable to guard a portfolio against adverse movement in vega. Just as in the case of gamma neutrality, we can only change a portfolio's vega by buying or selling an option; the vega of the underlying is zero since the stock price and volatility are independent variables; the stock price does not vary as a function of volatility.

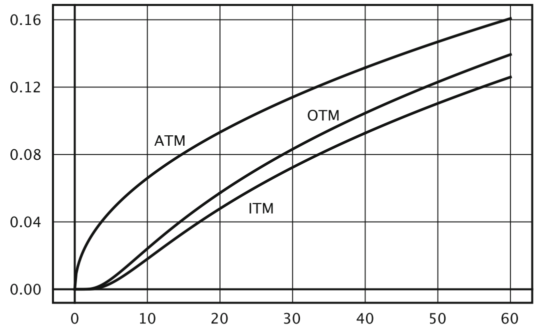


Fig. 5.6. Variation of vega vs time to expiration for put and call options at-the-money, 10 % in-the-money and 10 % out of-the-money as labeled

Suppose our current portfolio has a vega of ν_0 and we are willing to take a negative position in another option whose per unit vega is ν_A. Shorting w units of the addition makes the combined portfolio vega neutral if

$$w = \frac{\nu_0}{\nu_A}.$$

But, as before, this addition will change the delta of our portfolio. Suppose originally the portfolio was delta neutral. With the addition, the new portfolio has a delta of

$$-w\Delta_A$$

where Δ_A is the delta of the added option (possibly negative). We could buy $w\Delta_A$ shares of A (sell if Δ_A is negative) to make the new portfolio both delta and vega neutral.

Example 5.3. The vega of the put option in Example 5.1 is, from (5.9),

$$\nu_P = 60\sqrt{\frac{30/365}{2\pi}}e^{-0.071673^2/2} = 6.844.$$

The reported vega is 0.068. This means that the value of the option will increase by 6.8 cents for every 1 % increase in volatility.

The vega of the call option in Example 5.2 is

$$\nu_C = 60\sqrt{\frac{60/365}{2\pi}}e^{-0.1014^2/2} = 9.655$$

for a reported value of 0.097. The number of these calls to short to achieve vega neutrality is given by

$$-w(0.09655) + 0.06844 = 0, \quad w = 0.709 \quad \text{per put}.$$

For the 100 puts this comes to 71 calls. The combined delta for these two options is

$$-0.4714 \times 100 - 0.5404 \times 71 = -85.51.$$

So we should go long 86 shares to make delta zero too. □

5.2.5 Rho (ρ), the Derivative with Respect to r

Rho is the mathematical derivative of the option value with respect to the risk-free rate. Thus rho measures option sensitivity to the interest rate. Normally rho is the least important of the Greeks.

Rho is typically expressed as the amount of money, per share, that the value an option will gain as the risk-free rate rises by 1 % so it too is divided by 100 for the reported value.

By differentiating the Black-Scholes equations for a put and call, and noting the identity (5.8), we find that rho for a put is

$$\frac{\partial P}{\partial r} = -\tau K e^{-r\tau} \Phi(-d_2) \tag{5.10}$$

and for a call is

$$\frac{\partial C}{\partial r} = \tau K e^{-r\tau} \Phi(d_2). \tag{5.11}$$

We see that rho is negative for a put and positive for a call. To explain that, we also note the equations show that rho impacts options through the discounted strike price. Consider an at-the-money option and imagine that the risk-free rate is 0. Under the no arbitrage requirement for figuring fair option prices, we must assume the growth rate of the underlying is also zero. Hence the ending price is purely random around the present price. The buyer of a put or call would not be willing to pay much in this case. For the same reason the seller would be willing to sell at a low price. Hence the value of such a call is small.

But things change if the risk-free rate is high. Now the stock would be expected to gain considerable value over the time to expiry. For a call, this means it will more likely end in-the-money; for a put is means it will more likely end out of-the-money. Hence the value of a put decreases with increasing risk-free rate and the value of a call rises.

It is evident from (5.10) and (5.11) that rho tends to 0 as expiration approaches.

5.3 Setting Stops: Maximum Variables

Consider the following dilemma. A trader sells a 3 month call option struck at \$100. The underlying spot price is also currently at \$100. At the same time the trader sets an automatic stop to exit the trade if the spot price reaches \$105. What is the probability the stop will take out what would have been a profitable trade? In other words, what is the probability the underlying price will rise to

105 sometime during the life of the contract but still finish below 100? This is sometimes called the problem of *touching*, in this case, touching 105.

In another case, suppose a 2 month put contract was bought with a strike price of \$97.50 when the underlying was at \$100. The trader has an automatic stop set to sell back the option and cut losses if the underlying price reaches \$103. What is the probability the spot price will rise to 103 but still fall back below 97.50 before expiration?

Both of these questions can be easily answered by Monte Carlo methods but at the same time there are approximating analytical formulas. These derive from the theory of *maximum variables* for arithmetic random walks, see [KT75]. If X_t is an Wiener process with no drift, diffusion parameter 1, and $X(0) = 0$ then

$$\Pr(\max_{0\le t\le T} X_t > a \text{ and } X_T < b) = 1 - \Phi(\frac{2a-b}{\sqrt{T}}), \tag{5.12}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Similarly

$$\Pr(\min_{0\le t\le T} X_t < a \text{ and } X_T > b) = 1 - \Phi(\frac{b-2a}{\sqrt{T}}). \tag{5.13}$$

Now let S_t be GBM with starting value S_0, no drift, and volatility σ. Then

$$X_t = \frac{\log(S_t) - \log(S_0)}{\sigma} \tag{5.14}$$

is an Wiener process as above. Working through the substitutions in (5.12) and (5.13) yield the following

$$\Pr(\max_{0<t<T} S_t > m \text{ and } S_T < x) = 1 - \Phi(\frac{\log(m^2/(xS_0))}{\sigma\sqrt{T}}), \tag{5.15}$$

and

$$\Pr(\min_{0\le t\le T} S_t < m \text{ and } S_T > x) = 1 - \Phi(\frac{\log(xS_0/m^2)}{\sigma\sqrt{T}}). \tag{5.16}$$

Let us apply (5.15) to the first dilemma above; $S_0 = 100$, $T = 3/12$, $m = 105$, and $x = 100$. Suppose the underlying has drift equal to 3 % and volatility 20 %. The probability in question is calculated as

$$\begin{aligned}\Pr(\max S_t > 105, S_T < 100) &= 1 - \Phi(\frac{\log(105^2/(100^2))}{0.2\sqrt{0.25}}) \\ &= 1 - \Phi(0.9758) = 0.164.\end{aligned}$$

The analytical calculation is an approximation because it assumes the underlying has no drift – there is no exact formula when there is drift. But we can get a numerical approximation in this case by Monte Carlo.

Algorithm 22. Maximum Variables

```
inputs: S_0, T, μ, σ, m, x, N, Δt
hits=0; n = T/Δt;
for i = 1,...,N
```

```
    M = S = S_0   ▷M = maximum price over this scenario
    for t = 1,...,n
      Z ~ N(0,1)
      ΔS = S(μΔt + σ√(Δt) Z)
      S = S + ΔS
      M = max(M, S)
    endfor
    if M ≥ m and S ≤ x
      hits = hits + 1
    endif
  endfor
  P = hits/N
```

The time increment Δt must be taken very small for accuracy, on the order of one five thousandth of T. Running the algorithm on the first dilemma gives a probability of 0.160.

Using automatic stops is a static strategy. A dynamic strategy would be to consider closing the position day by day. For example, suppose the stock's price rises to \$105 with 1 month to go on the call contract. Now what is the probability the contract will expire out of the money (and the seller keeps the premium)? This has exactly the same answer as the question "what is the probability that a put option struck at 100 with 1 month to go will finish ITM if the current underlying is \$105?" From Section 3.6 the answer is given by (3.36),

$$\Phi(-d2) = \Phi(-\frac{\log(105/100) + (0.03 - \frac{1}{2}0.2^2)\frac{1}{12}}{0.2\sqrt{\frac{1}{12}}}) = 0.195.$$

5.4 Some Popular Option Trades

In this section we will investigate several well known option trade combinations. They have differing characteristics to take advantage of rising, falling, or stationary markets, of high or low volatility markets and so on. Most importantly we will introduce mathematical tools for assessing and comparing the various trades.

Traders are attracted to options for several reasons. Among them are *leverage*, limited loss, and flexibility. Leverage arises because one option contract controls 100 shares of stock but at just a fraction of the cost, often by a factor of 10 or more. In many cases the maximum loss of an option trade is limited to the cost of the options themselves, which, via leverage, is small relative to ownership of the underlying. In fact options enjoy the flexibility to tailor combinations of option trades so as to pre-engineer costs, return expectations, and risks (but not necessarily simultaneously).

Even the simplest trade such as buying a call entails choices: should the trade be ITM, ATM, or OTM? How far out in time from expiration? How long should it be held? The choices multiply with the complexity of the trade. In order to make rational decisions about these trade-offs we must arm ourselves with facts. We want to know:

- What is the risk of the trade – how much can be lost?
- What is the probability of the maximum loss?
- What is the probability of finishing with a gain?
- What is the overall expected gain?

While the amount at risk in a trade is usually easy to assess, the other questions are harder to figure. Fortunately they can be answered by simple Monte Carlo. Below we present a generic program for computing these critical variables. In the algorithm, the function $G(\cdot)$ is the payoff function dependent on the ending price S_T and is assumed to be supplied separately.

Algorithm 23. Trade Analysis Program

```
inputs: netProceeds of the trade (negative for a debit trade)
        maxLoss
gainCount = 0;
lossCount = 0;
expectedPO = 0;  ▷expected payoff
for i = 0,..., nTrials
  using the GRW algo., page 12, generate S_T
  gain = G(S_T) + netProceeds
  if gain > 0
    gainCount = gainCount + 1;
  else if gain == maxLoss
    lossCount = lossCount + 1;
  endif
  expectedPO = expectedPO + gain;
endfor
maxlossProbability = lossCount/nTrials;
gainProbability = gainCount/nTrials;
gainExpectation = expectedPO/nTrials;
```

In the following we will use this algorithm to analyze the trades. Important inputs to the GRW simulation itself are the option prices and the drift and volatility parameters. For prices, one should use the ask price in buying an asset and the bid price in selling. For volatility the implied volatility should be used. This is the market's assessment of the volatility of the equity at the present moment. For drift, it seems that recent experience is the best possible. That is, the drift over the past few months or few weeks or few days. Recall the US Treasury web site for up-to-date rates was given on page 34.

A caveat of the methodology is that real markets entail changing attributes from day to day. Optimal would be to program in the exact drift and volatility rates the same way. But these future value are, of course, unknown. However arbitrary scenarios can be programmed as desired.

For our demonstrations here we will use modest drifts and volatilities, usually 4–8 % drift, plus or minus, and 20–40 % volatility. We will also study the effect of changing volatilities during the course of the simulation in some cases.

However, we will use parameters in the simulation that favor the trade under study in order to demonstrate the strength of that particular type of trade and because we will assume the trade was chosen because conditions are favorable for it. This cannot be over emphasized in interpreting the results, *the GRW parameters used in the simulations are not necessarily typical but instead favor the trade under study.* Also it should be remembered that the simulations are for academic purposes only, meant to demonstrate techniques and should not be interpreted as trade recommendations.

5.4.1 Buying Puts and Calls

The simplest option trade is buying a put or a call; these are dual trades. To be definite, let us suppose a call. When making an option trade, the relevant graph is somewhat the reverse of the usual one for now it is the stock price that is fixed while the trader has the choice of strike price. For example, the stock price might be \$100 and the available strikes are \$90, \$95, \$100, \$105, and \$110. Figure 5.7 depicts a call option from this point of view. It shows that selecting a higher strike price makes it an out-of-the-money call while choosing one of the low strikes results in one in-the-money.

There are differing expectations among these choices. An OTM call is the least expensive but has the lowest probability of yielding a payoff. An ATM call improves the chances of a payoff, but it is the most expensive in terms of the cost of its time value. Finally, an ITM call is the most expensive because it has both intrinsic value and time value but it has the greatest chance of avoiding total loss.

Holding a deep ITM call is something like temporary ownership of the underlying itself. The deeper the call is in the money, the greater is its delta and the more it moves up and down to the same degree as the stock, see Fig 5.3. With respect to stock ownership, a deep ITM option has tremendous *leverage*. For example, with FDX at \$95.65, owning 100 shares of stock will cost \$9,565. But the 40 day 90 call might cost \$6.95 per share or \$695 for the contract; this is substantially less than outright ownership. This call has a delta in the neighborhood of 0.77. If the price of FDX rises by \$1, then the stock gain is 1/95.65 or about 1 %. But the option gain is $0.77/6.95 = 11\,\%$.

The downside is that the call earns no dividend and the "ownership" ends on the expiration date whether desirable or not.

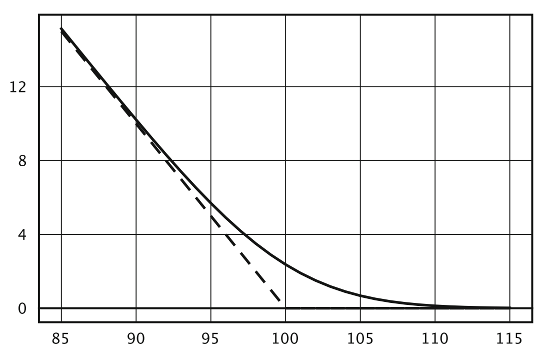


Fig. 5.7. The value of a call option from the point of view of a fixed stock price, $S = 100$, as a function of the strike price. The *solid graph* is 30 days prior to expiration while the *dashed graph* is at expiration

The maximum loss for put or call buying is restricted to the cost of the option itself.

Table 5.1 Gain expectation for put and call options

for the calls: $S_0 = \$80$, $T = 20$(days), $\mu = 8\,\%$, $r = 1\,\%$, $\sigma = 20\,\%$

Trade (strike p/c)	Strike vs stock	Price	Time value	Amt at risk	Prob. of total loss(%)	Prob. of a gain (%)	Expected gain	Gain rate (%)
75 call	ITM	5.18	0.18	5.18	8	51	0.283	100
80 call	ATM	1.52	1.52	1.52	47	37	0.161	193
85 call	OTM	0.18	0.18	0.18	89	10	0.034	344
for the puts: $S_0 = \$80$, $T = 20$(days), $\mu = -4\,\%$, $r = 1\,\%$, $\sigma = 20\,\%$								
85 put	ITM	5.14	0.14	5.14	9	51	0.188	67
80 put	ATM	1.47	1.47	1.47	47	37	0.110	137
75 put	OTM	0.14	0.14	0.14	91	9	0.019	249

In Table 5.1 we show the Monte Carlo analysis for typical put and call trades. Note that the drift rates are different between the put and call runs. Buying a call is a good strategy when the market is rising and we have used 8 % for these simulations. In contrast, buying a put is a strategy when the market is falling. These simulations use −4 %.

The greatest expected earnings is 28.3 cents for the ITM call but we must risk \$5.18 for it. Therefore the expected earnings rate is $0.283/5.18 = 5.46\,\%$ over 20 days or $5.46 * 365/20 = 99.6\,\%$ annually. By contrast, although the OTM call only generates an expected 3.4 cents of income, it does so on an 18 cent investment. Its annual earnings rate works out to 344 %.

Although the OTM call has the greatest growth rate, that trade does experience total loss 89 % of the time. As in the chapter on risk, it comes down to individual preferences between higher return rates vs greater safety.

Finally we can use prices from actual market data to test put-call parity, $S + P = C + Ke^{-rT}$. As an actual example, with 16 days remaining to expiration, APC stock sold between \$75.32(bid) and \$75.36(ask). The $K = 77.50$ calls sold for \$1.26(bid) to \$1.30(ask) and the puts for \$3.35(bid) to \$3.50(ask). Using the bid ask mid-points, and the 3 month Treasure rate of 0.03 %, we calculate

$$
\begin{aligned}
75.34 + 3.43 &\approx 1.28 + 77.50e^{-0.0003(16/365)} \\
78.77 &\approx 78.778,
\end{aligned}
$$

so within 1 cent; very close agreement. It has to be so or else there are arbitrage opportunities.

In passing we observe that put-call parity can be turned around to get the market's snapshot of the risk-free rate,

$$
r = -\log((S + P - C)/K)/T
$$

as a kind of "implied risk-free rate."

5.4.2 Selling Naked Puts and Calls

The reverse strategy to that above is selling puts and calls. However these trades entail great risk. In the case of a naked put, if the stock price goes to zero, the amount at risk is equal to the strike price (less the premium received). This could be quite substantial. The case for naked calls is even worse, the amount at risk is theoretically unlimited since the stock price could rise to any level. Most brokers do not allow these trades without approval.

Through put-call parity, at expiration, the payoff graph for a naked put is the same as that of the trade consisting of selling the call and buying the underlying stock, $-P = -C + (S - K)$. This trade is called a *covered call* and we will take it up next. The chart for a covered call is shown in Fig. 5.8b. In the same way, put-call parity written as $-C = -(S - K) - P$ shows that a naked call has the same payoff as shorting the stock and selling the put. This is called a *protective put* and is shown in Fig. 5.8a.

Selling a put can have its place in certain cases. If one is quite sure a particular equity is not going bankrupt and is looking to buy the stock at a low price, naked puts will do the job. If the price remains high, the option maker pockets the premium. If the price falls below the strike, the maker buys stock at the strike price, less the option premium.

5.4.3 Covered Calls and Protective Puts

The repair for naked put and call selling is covered calls and protective puts. Covered call writing consists of selling a call on a stock and buying stock in the amount to cover it, 100 shares per contract. Its dual is selling a put and shorting stock to cover it. The payoff diagrams are shown in Fig. 5.8.

Although these trades have the same payoff charts as do naked puts and calls, there is a very big difference – the stock purchase occurs up front. Consider the covered call. First the stock is purchased and then the call is sold (or better yet, the trade is made *atomically*, i.e. as one trade). If the stock experiences an unexpected increase, the stock is already in hand to cover the exercise. If instead the stock price dramatically falls, possibly even to zero, no additional monies are required. This is not the case for a naked put where the money to fulfill the exercise must be forthcoming at expiration.

Covered calls are widely used for generating income. If one already owns the stock, and this is where the risk originates, then selling calls on it produces an income stream so long as the stock price finishes below the strike price at expiration. The downside to the strategy occurs when the stock makes a run up. Then the trader only participates to the extent of the strike price and loses the stock upon exercise. Therefore covered calls are a good strategy for generating income on a stock very unlikely to lose value and not gain much in value either.

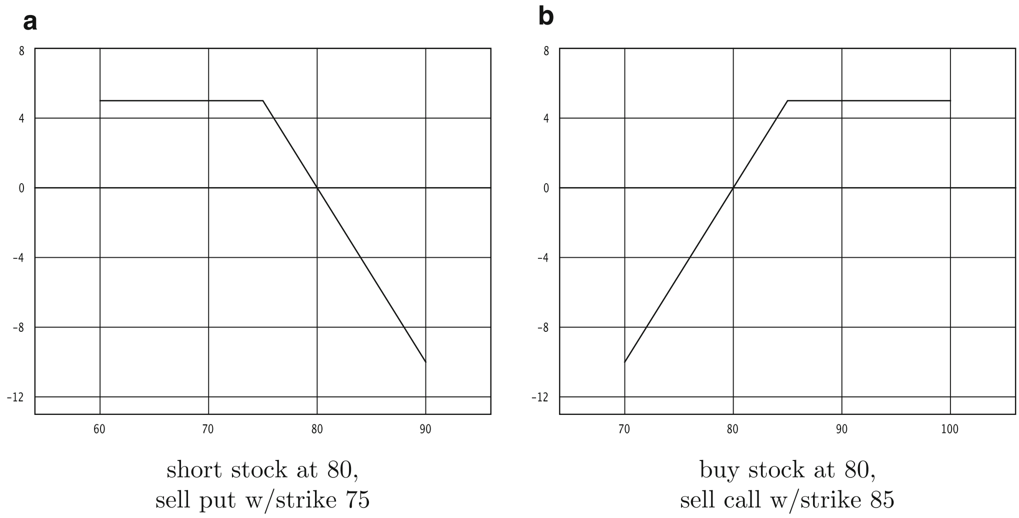


Fig. 5.8. Payoff diagram for the dual trades, a protective put in (**a**) and a covered call in (**b**). Through put-call parity these are the same as the graphs for a naked call and naked put respectively

In Table 5.2 we show simulation results for some covered call trades.

Table 5.2 Gain expectation for covered call trades

$S_0 = \$80.00$, $r = 1\,\%$, $\mu = 4\,\%$, $\sigma = 20\,\%$

Trade	Time to expir.	Price	Amt at risk	Total loss prob(%)	Gain prob (%)	Ex-pected gain	Gain rate (%)
82.50 call	15	0.43	79.57	0+	56	0.110	4
82.50 call	30	0.89	79.11	0+	59	0.200	3
82.50 call	45	1.27	78.73	0+	61	0.295	3
85 call	15	0.10	79.90	0+	52	0.121	4
85 call	30	0.36	79.64	0+	54	0.242	4
85 call	45	0.64	79.36	0+	56	0.337	3

Delta Hedging Case Study

In this paragraph we track the fate of a call delta hedged versus a straight covered call. In this study, the stock price falls from \$84.71 42 days from expiration to \$74.58 at expiration. The call cost \$3.80 per share or \$380 for 1 contract (by not counting fees and by splitting the bid/ask).

For the covered call the initial outlay is \$8,471 minus the \$380 from the call for a net position of $\$-8{,}091.00$. At the end the option expired worthless so the stocks were sold back for \$74.58 per share. The final position is $\$-633.00$.

For the delta hedge the initial delta is 0.5126, thus 51 shares are bought for an initial outlay of \$3,909.61. The hedge was rebalanced once a week until expiration. The result, presented in Table 5.3, shows that the covered call lost \$633 dollars while the hedge gained \$49.97. This came from the call premium.

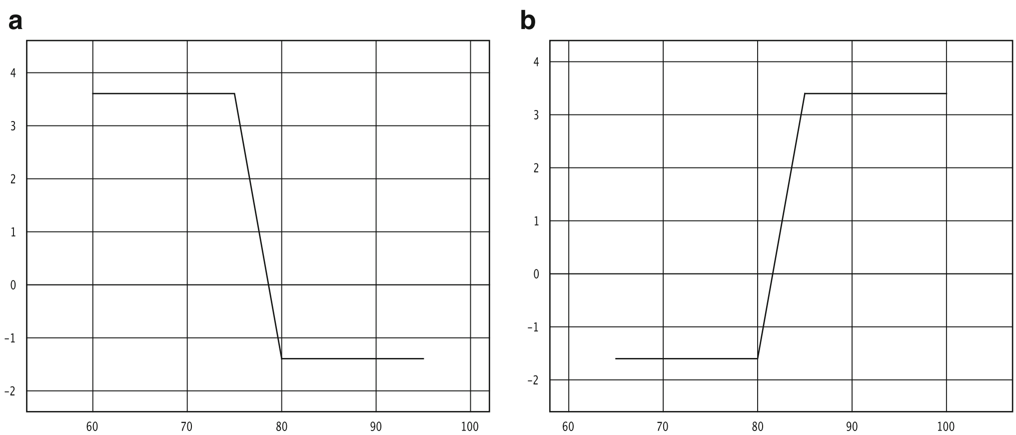
Table 5.3 A delta hedge case study

equity at \$84.71, 85 call at \$3.80, $r = 0.01$, $v = 34\,\%$

Days to expiration	Stock price	Delta	b/s shares	Proceeds	Shares held	Net proceeds	Covered call
42	84.71	0.5126	b 51	−4,289.61	51	−3,909.61	−8,091.00
35	80.42	0.3201	s 19	1,527.98	32	−2,381.63	−8,091.00
28	79.07	0.2377	s 8	632.56	24	−1,749.07	−8,091.00
21	78.54	0.1953	s 4	300.56	20	−1,448.51	−8,091.00
14	75.14	0.0355	s 16	1,202.24	4	−246.27	−8,091.00
7	74.06	0.0022	s 4	296.24	0	49.97	−8,091.00
0	74.58	0.0000	s 0	0	0	49.97	−633

5.4.4 Debit Spreads

A debit spread, either with puts or calls, is like buying a put or a call and financing it, to an extent, by selling the same type further out-of-the-money. For example, a *debit call spread* entails buying a call at one strike price K_ℓ and selling a call, on the same underlying, at a higher strike price, $K_h > K_\ell$. Usually the long option is at or near-the-money. From Fig. 5.7 we see that the higher the strike of a call, the lesser is the option's price. Therefore the option bought costs more than the option sold and the trade is for a net debit.



put spread with strikes 80 and 75 call spread with strikes 80 and 85

Fig. 5.9. Debit put and call spreads. The stock price was \$80 when the trades were executed

A debit put spread means buying a put at one strike, K_h and selling a second put at a lower strike, $K_\ell < K_h$. This is the dual trade to the debit call spread. Figure 5.9 shows the payoff graphs of both.

As mentioned, selling the deeper OTM option brings in premium to offset the main purchase. The downside is that it limits the possible gain to the difference between the strikes minus the cost of the trade. For example, in the figure the cost of the 80 call is \$2.04 and the premium for the 85 call is \$0.48; the net cost for the trade is therefore \$1.56. The holder of the trade loses this amount if the stock comes in below \$80 at expiration. From that point the trade gains 1 for 1 for each increase in expiration price S_T. The trade breaks even at $S_T = 81.56$. And it continues to profit to a maximum of $5 - 1.56 = 3.48$ for any ending price over \$85.

The maximum loss of a debit spread is restricted to the net cost of the trade.

Table 5.4 Gain expectation for debit spreads

$S_0 = \$80$, $r = 1\,\%$, $T = 20$(days), $\sigma = 20\,\%$ $\mu = 8\,\%$ for calls, $-4\,\%$ for puts

Trade buy strike/ sell strike	Net price (debit)	Amt at risk	Simulation drift	Prob. of total loss (%)	Gain prob (%)	Expected gain	Gain rate (%)
80/85 calls	(1.33)	1.33	0.08	47	39	0.129	177
82.50/85 calls	(0.41)	0.41	0.08	72	25	0.054	240
82.50/87.50 calls	(0.55)	0.55	0.08	72	23	0.079	263
80/75 puts	(1.34)	1.34	−0.04	48	38	0.084	115
77.50/75 puts	(0.40)	0.40	−0.04	73	24	0.036	165
77.50/72.50 puts	(0.51)	0.51	−0.04	73	23	0.051	181

In Table 5.4 we show the Monte Carlo analysis for some debit spread trades. In buying a debit spread, the holder is hoping the market moves "towards" the trade, that prices increase for the call spread and decrease for the put spread. For this reason a debit call spread is also known as a *Bull call spread* and a debit put spread as a *Bear put spread*. Accordingly we have set the drift to $8\,\%$ for the call spreads and $-4\,\%$ for the puts spreads in these simulations.

5.4.5 Credit Spreads

The *credit spreads* are the reverse trades of their debit spread counterparts. A *credit call spread* means selling a call with strike K_ℓ and buying a call with strike $K_h > K_\ell$. The function of the call purchased is to provide protection against an unexpectedly large rise in the price of the underlying. Without it, the trade is a naked call. Similarly, to create a *credit put spread*, one buys a put at one strike K_h and sells another at a lower strike $K_\ell < K_h$. Again the function of the option purchased is protection.

As advertised, the trade is for a credit; it brings in income immediately. However, if the market moves against the trade, the loss can be as large as the distance between the strikes less the initial credit. See Fig. 5.10.

As opposed to the debit spreads, the holder of a credit spread is expecting the stock to move away from the trade or at least not move toward it. For a credit call spread this means anticipating the price to stay the same or fall. Thus

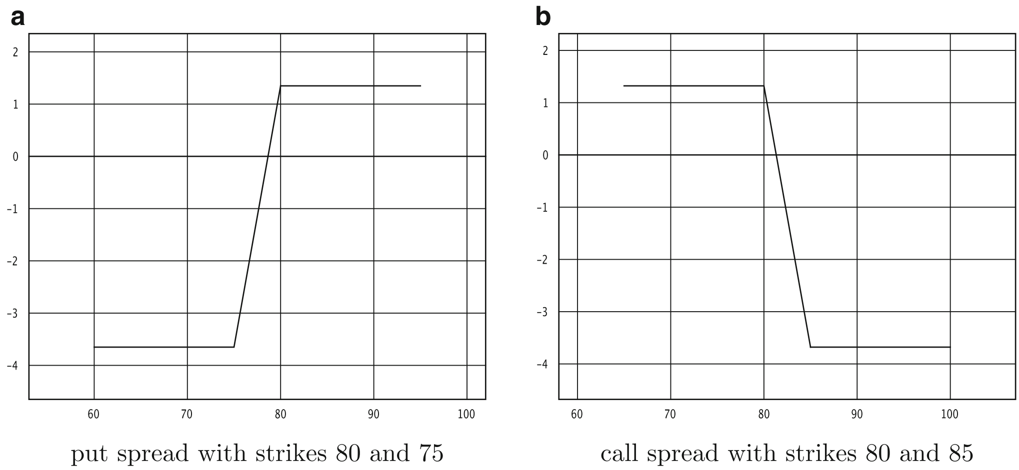


Fig. 5.10. Credit put and call spreads. With the stock price at \$80 when the trades were executed, the call spread brought in a net income of \$1.35 while the put spread brought in \$1.32

a credit call spread is also known as a *Bear call spread*. Accordingly we use a drift of $-4\,\%$ in the simulations. On the other hand the holder of a credit put spread counts on the price staying the same or increasing. So it is called a *Bull put spread*. We will use a drift of $8\,\%$ in the analysis. The results are presented in Table 5.5.

Table 5.5 Gain expectation for credit spreads

$S_0 = \$80$, $r = 1\,\%$, $T = 20$(days), $\sigma = 20\,\%$

Trade sell strike/ buy strike	Net price (credit)	Amt at risk	Simu-lation drift	Prob. of total loss (%)	Gain prob (%)	Ex-pected gain	Gain rate (%)
80/85 calls	1.33	3.67	−0.04	9	66	0.083	41
82.50/85 calls	0.41	2.09	−0.04	9	80	0.034	29
82.50/87.50 calls	0.55	4.45	−0.04	2	81	0.049	20
80/75 puts	1.34	3.66	0.08	8	67	0.125	62
77.50/75 puts	0.40	2.10	0.08	8	80	0.045	39
77.50/72.50 puts	0.51	4.49	0.08	2	81	0.065	27

5.4.6 Calendar Spreads

The *calendar spread* is a trade in which a near term expiration option is sold and a longer term option on the same stock and for the same strike price is bought. With more time to expiration, the long term option is more expensive so the trade is for a net debit. This is a trade that takes advantage of the behavior of theta. If the stock remains about the same price, then the time value of the near term option will tend to zero more rapidly than that of the long term option. In the optimal case the near term option expires worthless. At some point along the way the trade becomes profitable.

In Fig. 5.11 we show the payoff picture of a calendar spread done with puts in (a) and with calls in (b). These are dual trades.

With regard to panel (b), when created, the short call was ATM and had 39 days to expiration for a Black-Scholes price of \$1.66. The long call was 67 days out and cost \$2.21. The net debit was \$0.55. Upon expiration of the near term call, the long term call still has 28 days of time value. If the stock price is \$60 at that time, then the short option expires worthless and the long option has a Black-Scholes value of \$1.39. The net gain of the trade in that event is \$0.84 and this is the maximal gain. For lower ending prices of the underlying the long term option becomes successively more out-of-the-money and loses value. But for higher ending prices the near term option becomes successively more in-the-money and again the trade loses net value.

A similar analysis holds for the calendar put spread.

The maximum loss of a calendar spread with calls is restricted to the net cost of the options. But for puts the maximum loss can be bigger. This is due to the fact that puts can have negative time value, see Fig. 3.9a (page 103).

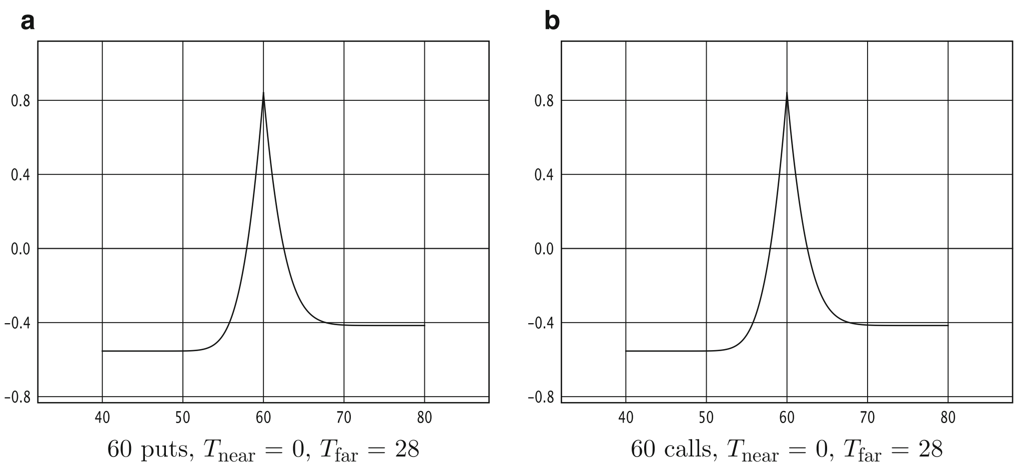


Fig. 5.11. Calendar spread with: (**a**) puts, (**b**) calls; strike price 60. Symmetry is broken because calls always have positive intrinsic value, puts do not. As a result, the maximum loss of a put calendar spread exceeds the net cost of the constituent options, here 55 cents vs 41 cents

As is seen in Fig 5.11 above, the gain region is rather narrow and centered around the strike price of the two options. In establishing this trade, the trader is expecting the underlying to stay relatively constant in price.

The Monte Carlo analysis for this trade is presented in Table 5.6. For this simulation we have added a column headed IV (implied volatility) and another headed SV (simulation volatility). The first is the volatility used for calculating the Black-Scholes price for the trade. The second is the volatility used for the simulation.

Table 5.6 Gain expectation for calendar spreads

$S_0 = \$60$, $r = 1\,\%$, IV = implied vol., SV = simulation vol.

Trade	T (days) near/far	Price = amt at risk	Drift (%)	IV (%)	SV (%)	Total loss prob	Gain prob (%)	Expected gain	Gain rate (%)
60 call	10/38	0.89	4	23	23	0+	51	0.001	4
60 call	10/38	1.52	0	40	40	0+	51	0.001	3
62.50 call	10/38	1.41	6	40	40	0+	50	0.012	31
60 call	10/38	0.77	4	20	23	0+	60	0.114	537

From the first line of the table we see that under typical rates of drift and volatility, the trade tends to finish with a small profit or loss. Even with 0 or slight drift and high volatility the trade essentially breaks even. The third lines shows that if the drift is favorable, it might take the trade into the region of maximum payoff. But the fourth line shows that the profitability of calendar spreads is very sensitive to changes in volatility during the period of the contract. The expected gain increases substantially as volatility increases. This shows the effect of vega in a dramatic way.

5.4.7 Straddles

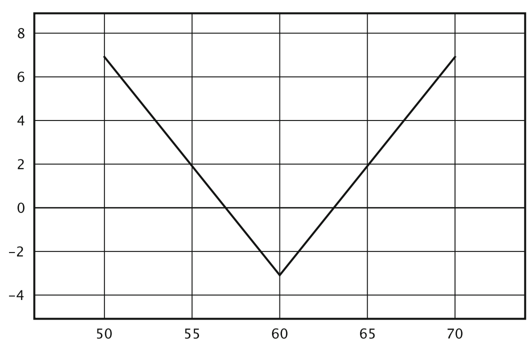


Fig. 5.12. The payoff graph of a straddle. The trade makes a gain if the underlying moves sufficiently either way

A *straddle* entails buying both a put and a call on the same underlying at the same strike price and with the same expiration. The payoff diagram for a straddle is illustrated in Fig. 5.12. This trade is self dual.

A straddle makes a gain if the underlying moves enough either up or down. Evidently the trader believes that the underlying will make a big move in some direction but does not know which. For example an important announcement about the company is due out; it may be good news or bad news. The straddle also is a winner if the volatility should increase over the duration of the option. If volatility is high when the trade is placed, then the options are expensive to start with. Rather it is an increase in volatility that is required. A profit is also possible if the drift is high either positively or negatively.

Table 5.7 Gain expectation for straddles

$S_0 = \$60$, buy 60 call, buy 60 put

Trade expiry (days)	Price = amt at risk	Drift (%)	IV (%)	SV (%)	Total loss prob	Gain prob (%)	Expected gain	Gain rate (%)
10	1.58	2	20	20	1	42	−0.002	−5
20	2.24	2	20	20	1	43	0.003	2
10	1.58	2	20	23	1	49	0.241	556
20	2.24	2	20	23	1	49	0.341	278
10	1.58	4	20	20	1	43	0.003	7
20	2.24	4	20	20	1	43	0.008	6

As with all long option trades, the risk is restricted to the option premiums.

In Table 5.7 we show the results of a Monte Carlo analysis. As was the case with calendar spreads, straddles greatly benefit if the volatility increases over the time period of the trade. As above, IV is the volatility when the option was placed and SV is the volatility at which the simulation was conducted. The table also shows that profitability improves as the stock's drift rate increases.

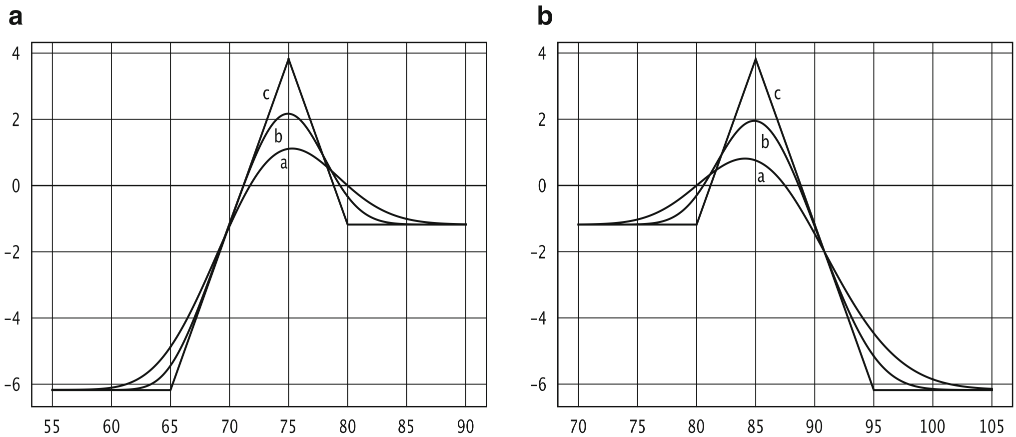
The opposite trade, going short a straddle, is obviously very dangerous. It has the danger of both a naked call and a naked put combined. Being the opposite of a long straddle, a short straddle is profitable when the underlying stays at about the same price over the duration of the option. But the risk is very high. To garner the benefits of a short straddle but with limited risk, motivates our next trade, the butterfly.

5.4.8 Butterflies

A *butterfly* is a 3 option trade, either with 3 calls or, for the dual, with 3 puts. For the call butterfly, a call is bought at the low strike, K_1, then two calls are sold at the middle strike, K_2, and finally, a call is bought at the high strike, K_3,

$$K_1 < K_2 < K_3.$$

In some ways the butterfly is the opposite of a straddle as is evident in its payoff chart, graphs (c) Fig 5.13. The butterfly has a peak of profitability centered around its middle strike and falls off in either direction from that. Buying the options on the two sides lowers the income from the option sold,



with puts, strikes: 80/75/65 with calls, strikes: 80/85/95

Fig. 5.13. Black-Scholes values of a skip strike butterfly, (**a**) with puts, (**b**) with calls. In each case curve (**a**) is 21 days from expiration, (**b**) is 7 days from expiration, and (**c**) is at expiration

even to the point of a net debit, but provides crucial insurance limiting loss should the stock make a large move up or down.

In the butterflies depicted in the figure, the first two strikes are sequential as offered by the exchange, but the next strike is skipped; the final strike is the next after that. This *skip-strike butterfly*, as it is called, can be designated 1:2::1. Obviously there are lots of possibilities, e.g. 1:2:1 (traditional), 1:3::2, and so on. The standard skip-strike call butterfly is often established when the stock price is near the low strike expecting it to rise modestly. Conversely, the put butterfly would be used when a slight decline in price is expected. Alternatively, in a declining market the call butterfly can be set up with the current stock price at the last strike expecting the negative drift to carry the stock price into the profit region. Similar remarks apply to the put butterfly in a growing market.

A butterfly will generally be for a net debit because, if put on when the stock price is near the first strike, the most expensive option will be that one. The maximum gain of the trade occurs when the expiration stock price falls on the middle strike. The maximum loss depends on how far the insurance options are, in terms of their strikes, from the middle strike. Therefore it is case of reward vs risk. (For the call butterfly in the figure, the maximum loss is $6.18.)

In Table 5.8 we give the results of a Monte Carlo analysis of some butterfly trades. The table shows that the most favorable trade is at the high strike in a declining market.

Note that these results do not include trading commissions or the bid-ask bias. As option combinations involve more options, 3 for the butterfly, the trading costs likewise escalate.

Table 5.8 Gain expectation for butterflies

$T = 21$, $r = 0.03$

Trade: calls at 80:85:90:95	Stock price S_0	Price (net debt)	Max loss	Drift (%)	Total loss prob	Gain prob (%)	Ex-pected gain	Gain rate (%)
1:2:1:	80	(1.19)	1.19	4	50	37	0.019	28
1:2:1:	85	(2.18)	2.18	2	22	51	0.003	2
1:2:1:	90	(1.21)	1.21	−4	48	40	0.097	140
1:2::1	80	(1.18)	6.18	4	0+	37	0.016	5
1:2::1	85	(1.94)	6.94	2	1	55	0.010	3
1:2::1	90	0.28	4.72	−4	12	55	0.228	84
1:3::2	80	(0.97)	15.97	4	0+	37	0.013	1
1:3::2	85	(0.25)	15.25	2	1	60	0.034	4
1:3::2	90	5.34	9.66	−4	12	55	0.511	92

5.4.9 Iron Condors

An *iron condor* is a two spread trade with four strike prices, $K_1 < K_2 < K_3 < K_4$. It's payoff is depicted in Fig 5.14. The rationale for the trade is that it consists of two spreads and at least one of them is guaranteed to finish in-the-money; usually both.

The K_1, K_2 spread can be either a debit call spread or a credit put spread. And the K_3, K_4 spread is either a credit call spread or a debit put spread. Mix or match; thus there are four possible iron condors: (A) with credit spreads, (B) with calls, (C) with puts, and (D) with debit spreads. For example, the maximum up front net credit is (A), arranged by choosing a credit put spread at K_1, K_2 and a credit call spread at K_3 and K_4. In this case the trade consists of:

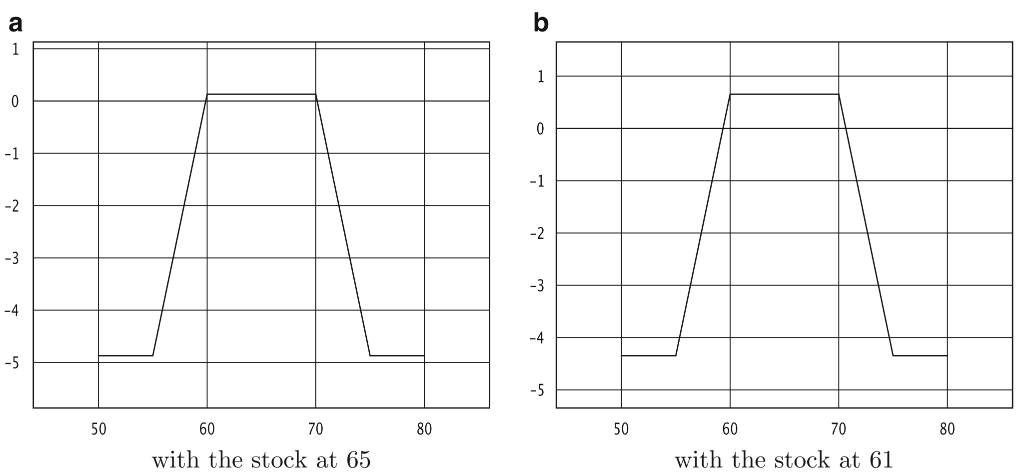


Fig. 5.14. An iron condor with strikes 55/60/70/75. There is a broad region over which it makes a small profit. But the amount at risk is large relative to it

- Buying a put at K_1
- Selling a put at K_2
- Selling a call at K_3
- Buying a call at K_4

This trade is self-dual. So is trade (D). The dual of (B), with calls, is (C), with puts and conversely.

The iron condor is attractive because: it is a non-directional, it has high probability of finishing in-the-money, and it has limited risk. On the downside, the strategy does involve trading four options.

The most striking feature of the iron condor's payoff graph is its small reward relative to its high potential loss. Since the trade is a combination of spreads, the maximum gains and maximum losses are just the sum of those for the constituent spreads.

In Fig. 5.14a the stock price is at 65, the credit put spread brings in a net 5 cents and the credit call spread brings in a net 8 cents. Therefore the maximum gain is 13 cents. The maximum loss is that minus \$5 or \$4.87.

In panel (b) the stock price is at 61, the credit put spread brings in a net 65 cents and the credit call spread brings in less than 1 cent. Hence the maximum gain is 65 cents and the maximum loss is 4.35. The difference in the gains and losses between the trade with the stock at 65 versus the stock at 61 is due to the non-linear nature of Black-Scholes pricing with respect to stock price S.

While the gains are small and the losses relatively big, the former comes with high probability while the latter at a very small probability. These trade-off situations are exactly where Monte Carlo expectation analysis excels. Some results are presented in Table 5.9. As before, IV refers to the volatility at which the options are purchased and SV is the simulation volatility.

In the first series of runs the initial stock price is in the middle of the range and small or zero drifts are assumed. The results show that under these conditions there are virtually no gains, even despite a wide range of volatilities.

In the second series of tests the initial stock price was set at the bottom of the range and fair or large drifts are assumed. Under these conditions the condor produces modest grow rates in step with drift. In the last of that series, the symmetric condition was tested with the stock at the top of the range and a small negative drift assumed. The results are about the same.

Next a series was tried using condors with calls instead of credit spreads. There is virtually no difference in performance.

Lastly conditions of changing volatility were tested with dramatic results. If volatility increases over the course of the contract as compared to the volatility prevailing at its purchase (SV $>$ IV), then the condor sustains big losses. If the reverse holds, volatility decreases during the contract, then profits get a major boost.

Once again this example shows the major effect that vega has on certain option trades.

Table 5.9 Gain expectation for iron condors

$T = 20$, $r = 0.01$, costs and losses as noted in the text

Trade: strikes at 55:60:70:75	Stock price S_0	Drift (%)	IV (%)	SV (%)	Total loss prob	Gain prob (%)	Ex-pected gain	Gain rate (%)
w/cred.spds	65	2	20	20	0+	91	−0.002	−1
w/cred.spds	65	0	10	10	0+	99.9	0.000	0
w/cred.spds	65	0	40	40	10	69	−0.002	−1
w/cred.spds	61	4	40	40	15	63	0.030	16
w/cred.spds	61	8	40	40	15	64	0.058	31
w/cred.spds	69	−2	40	40	18	63	0.019	10
w/calls	61	4	40	40	15	63	0.027	14
w/calls	61	8	40	40	15	64	0.057	30
w/calls	69	−2	40	40	18	63	0.021	11
w/cred.spds	63	6	20	40	11	59	−0.998	−382
w/cred.spds	63	6	40	20	0+	93	1.035	503

Problems: Chapter 5

Many of the problems for this Chapter ask that you analyze option trades. For these, please create a program similar to Algorithm 23 on page 149. Use the GBM and Black-Scholes models in the simulations (e.g. for setting option prices).

1. At the present time BAC is selling for 9.52. Its 18 day 10 dollar calls have these Greeks: $\varDelta = 0.3366$, $\varGamma = 0.3509$, $\nu = 0.0077$ (directly from a brokers web site). (a) What are the corresponding Greeks of the 18 day 10 dollar puts? The 46 day 10 dollar calls have these Greeks: $\varDelta = 0.388$, $\varGamma = 0.277$, $\nu = 0.0129$. (b) Set up a delta-gamma neutral portfolio in BAC. (c) Set up a delta-vega neutral portfolio in BAC. (d) Can the implied volatility be calculated from these data? Explain.

2. In addition to the Greeks as in Problem 1, for the 46 day 7.50 puts delta is −0.0528, gamma is 0.0738, and vega is 0.0036. Can you set up a delta-gamma-vega neutral position? Do so if possible.
(Ans. long 843 stock, long 100 18 day calls, short 1,538 46 day calls, long 5,298 46 day puts.)

3. Conduct a delta hedge similar to that in Section 5.4.3. In one case assume the stock price increases modestly over the 42 days until expiration. In another, generate a GRW 42 day price sequence.

4. Analyze the strategy of selling covered calls 5 or 6 days before expiration. Experiment with different volatilities, drifts, and stock/strike relationships (i.e. ITM, ATM, etc.).

5. Answer the second "dilemma" in the section on maximum variables, Section 5.3. If the volatility is 20 % and the time to expiration is 2 months, what is the probability that the stock price starting from 100 will rise above 103 over the term of the

put option but nevertheless finish below 97.50? In the first case, assume the drift is zero. Write a program to answer the question if the drift is 6 %, if the drift is −4 %.

6. Using the parameters in the header of Table 5.1, except for the drift, and analyze the OTM call trade (85 strike) for various values of drift; for example $\mu = 0.02$, $\mu = 0.04$, and $\mu = 0.06$.

7. Using the parameters in Table 5.2 for the 45 day 82.50 covered call, analyze the following stop outs: exit the trade if the stock price falls to 79, to 78, to 77. What is the gain rate in each case?

8. Using the parameters in the header of Table 5.4 except for the starting stock price S_0, analyze the 80/75 put spread for various starting prices ITM; for example $S_0 = 79$, $S_0 = 77$, $S_0 = 75$.

9. Using the parameters in the header of Table 5.5, investigate the effect on the gain rate of the 80/85 spread if the volatility: becomes 30 % just after the trade is established, becomes 10 % under the same conditions. (The option prices are still determined by the stated 20 % volatility.)

10. Analyze the dual trade to the third butterfly trade in Table 5.8. Thus go long a put at 80, short 2 puts at 75, and long a put at 70. Let the starting price be $S_0 = 70$ and assume the drift is −2 %.

6

Alternatives to GBM Prices

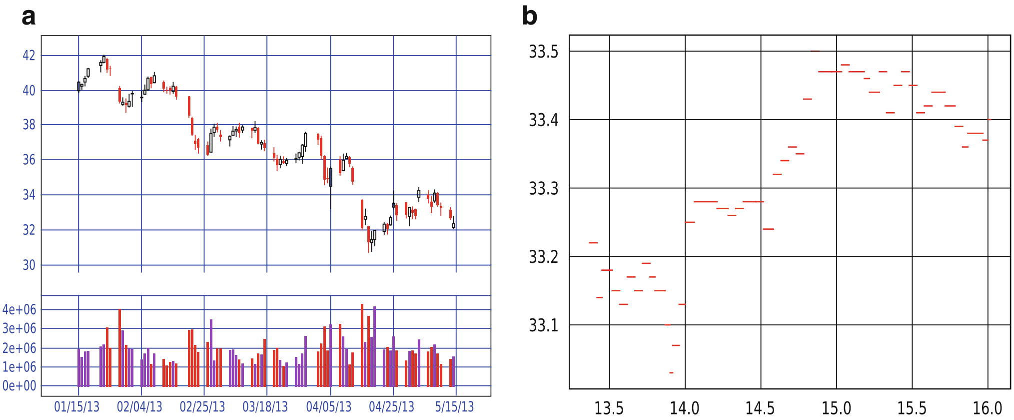
The developments of the previous chapters have built a financial edifice upon the normal distribution, the Gaussian, primarily through the Wiener process. But there is much evidence that the world is not Gaussian, that Gaussian is only an approximation to reality. Some evidence that it is not is seen in Fig. 6.1. These depictions should be compared with Fig. 1.1 on page 2. The earlier graph portrays a stock's price through time as being continuous. But by magnifying the time scale and viewing prices over a few months we see that stock prices are occasionally discontinuous, they can suddenly change from one value to another without going through the values in-between. This often occurs between days as seen in Fig. 6.1a. By expanding the scale to the level of hours, one sees that the prices are possibly nothing but jumps, many of them small as in panel (b).

Other evidence comes from the phenomenon of *volatility smile*. According to Black-Scholes theory, for a fixed time to maturity T, the price of all options on a given stock as a function of strike price, should be calculated using the same volatility. Namely, it should be the volatility that prevails over the time horizon of the option (or at least the average of such). But this is not what is observed. Implied volatilities for puts are greater than those for calls; and the lower the strike, the greater the volatility.

And there is more. Events that should happen only rarely or, practically speaking, never, instead occur two or three times a generation. This indicates that the Gaussian is the wrong distribution, that rare events should have a higher probability of occurring. It indicates that the tails of a more accurate distribution should have more probability mass than does the Gaussian.

In this chapter we study price processes that are not Gaussian; processes that have jumps and "heavy tails." However there is a concomitant downside, namely that options can no longer be hedged, and therefore have no unique price. The term used is *incomplete market*, and it is here that we start.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_6,



daily open and close intraday prices

Fig. 6.1. Stock price history for SCCO over short periods of time. In (**a**) 4 months of prices are illustrated. Each candle's top is the day's high while the bottom is the day's low. If the price fell during the day, the candle is shown in *red*. It can be seen that frequently 1 day's candle does not overlap the next; thus the price jumped by at least the gap between the two. The *lower* part of the figure shows the volume or the number of stocks traded during the day. In (**b**) SCCO's intra day's prices are shown over a 3 h period during an afternoon. It is clearly seen that the price jumps almost minute by minute

6.1 Martingale Measures

Up to now we have lived in a discrete time world. Our techniques have exploited binomial lattices and GBM implemented discretely over finite increments in time. But a Wiener process is a continuous time theory. Its major accomplishment lies in showing how to define a probability or *measure* to Brownian motion paths; the object our random walks attempt to simulate. Of course, any single path has probability 0; there are, for any finite interval of time $0 \leq t \leq T$, an uncountable infinity of continuous paths X_t. But it makes sense to talk about the probability of sets of paths. For example, all paths whose Brownian particle lies between $x = 0$ and $x = 1$ when $t = T$, or in another example, all paths that were less than $x = -1$ at some time $t < T$ but finished bigger than $x = 5$ when $t = T$. And there are sets of paths for which we can assign a probability from first principles, those determined by their position at any fixed time t. We can do so because, by axiom, W_t is normally distributed with mean 0 and variance t at this time.

As we noted in Section 1.3, a Wiener process is a martingale. In this chapter we consider price processes which are not based on the Wiener process. The paths $X = \{X_t, t \geq 0\}$ of such a process must belong to some universal set Ω. And there must also be a measure or probability function defined for subsets of Ω as discussed above; the class of subsets must be closed under countable set operations (set complement, countable unions, countable intersections). And as we have seen earlier, there can be more than one probability function defined on paths. For example, for price paths there can be a historical or real-world probability P and there can be a risk-neutral probability Q. Two probabilities are said to be *equivalent* if their sets of probability zero are identical (hence also their sets of probability one).

A stochastic process $X = \{X_t, t \geq 0\}$ in such a space Ω is a *martingale* with respect to a given measure Q if the expected future value of X at any time s is X_s,

$$\mathbb{E}_Q(X_{s+t}|\text{the information about } X \text{ up to time } s) = X_s. \tag{6.1}$$

A measure for which the process is a martingale is called a *martingale measure*. A martingale process is something like a fair game in that a players expected fortune at the end of the game is the same as his fortune at the start, see Chapter 7.

The importance of martingale measures is made clear in the Fundamental Theorem of Asset Pricing. In order to state it we need to define a few terms. A financial *derivative* or *contingent claim* is a security whose value depends on the value of other more basic underlying securities. Options and forwards are two examples of derivatives. A *complete market* is one for which every contingent claim has a self-financing replicating portfolio.

Theorem (Fundamental Theorem of Asset Pricing) *A discrete time pricing model has no arbitrage opportunities if and only if it has a measure for which discounted prices are a martingale. Further, the model is complete if and only if the martingale measure is unique.*

For a proof see [Rom12]. For continuous time pricing models, the theorem breaks down in that the conditions are no longer if and only if. It is still true however that the existence of a martingale measure implies there are no arbitrage opportunities and the uniqueness of the measure implies market completeness.

6.2 Incomplete Markets

In Section 3.4 we began our study of option pricing by applying the principle of no-arbitrage to a one-step price tree. Suppose now there had been three possible prices of the stock at expiry instead of just two, see Fig. 6.2.

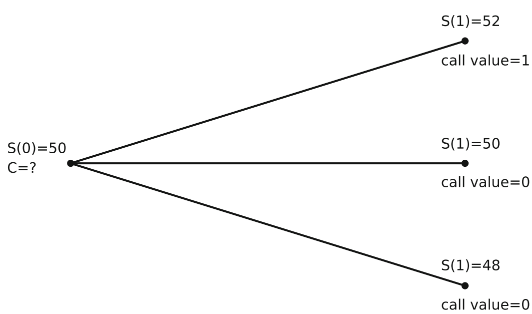


Fig. 6.2. A one step price tree with three possible prices at expiry

As before, consider a portfolio consisting of Δ shares of stock and short 1 call option struck at $K = 51$ costing C. To simplify the calculation, assume the risk-free rate is 0; this just means money can be borrowed with no interest but still has to be paid back. From the previous analysis the number of shares to hold in the portfolio is $\Delta = 1/4$ and from (3.11) the value of the call is $12.50 - 12e^{-r_f} = 0.50$. So the initial value of the portfolio is $(\frac{1}{4})50 - 0.50 = 12$.

And in the present case this is the value of the portfolio if the stock goes up to 52 or down to 48 as before (recall $r_f = 0$). But if the final price is 50, the value of the portfolio is 12.50. Now an arbitrage is possible: borrow \$12 to set up the portfolio, pay it back if the price goes up or down, but if the price stays the same, the portfolio makes \$0.50 after retiring the loan. Thus with some positive probability, the probability of the middle branch of the tree, the portfolio makes a positive profit with no chance of losing money.

This happens because no value of Δ makes the expiry value of the portfolio equal for all three branches of the tree. So there is no one value to discount back to time 0 in order to find C.

Suppose the call is \$0.40 instead of \$0.50. Now to set up the portfolio \$12.10 will have to be borrowed. If the price goes up to 52 the stock will be worth 13, satisfying the call will cost 1 leaving only 12 to pay back the loan. Therefore the portfolio loses \$0.10. Likewise it loses the same if the price goes down to 48. If the expiry price is 50, then the stock is worth 12.50 and, after repaying the loan, the portfolio makes \$0.40. So if the call price is \$0.40 the portfolio's expectation, positive or negative, depends on the probabilities of the three outcomes.

What are those probabilities? Perhaps we can proceed by calculating the risk-neutral probability, from that find the expected payoff and discount back to get C.

Since there are three branches, the risk-neural probability will in fact be a probability density: q_1 that the price rises, q_2 that it stays the same, and q_3 that it falls. We must have $q_1 + q_2 + q_3 = 1$ and no probability can be zero. Recall that the risk-neutral density is the one for which the expected growth of the underlying equals the risk-free rate, see page 90. Combining this expectation balance with the total density summing to 1, we have the system

$$\begin{aligned} 52q_1 + 50q_2 + 48q_3 &= 50 \\ q_1 + q_2 + q_3 &= 1. \end{aligned} \tag{6.2}$$

There is no one solution; solving in terms of q_3 we have

$$q_1 = q_3, \quad q_2 = 1 - 2q_3, \quad 0 < q_3 < \frac{1}{2}. \tag{6.3}$$

The bounds on q_3 assure that all three probabilities will be positive. To say that the expected price grows according to the risk-free rate is equivalent to saying the discounted expectation of S_t is a martingale; in this example, it will be so for any q_3 between 0 and 1/2.

For example, choosing $q_3 = 0.4$, entails $q_1 = 0.4$ and $q_2 = 0.2$. This makes the expected call payoff equal to \$0.40, and, discounting back with $r_f = 0$, puts the price of the call at \$0.40. As analyzed above, there is no risk free profit for this value of the call.

Thus we have encountered an example of an incomplete market.

6.2.1 Pricing in an Incomplete Market

In an incomplete market there is no unique no-arbitrage price; instead there are many. In the example above, q_1 and q_2 given by (6.3) along with any choice $0 < q_3 < 1/2$ produces a martingale and with it, a no-arbitrage price for the call. The decision as to which value of q_3 to use becomes a subjective matter; a risk-averse investor would want the real-world payoff to exceed the martingale payoff.

Finally, what is the point of pricing vanilla options by a mathematical model anyway; the market already prices them. Instead, the prevailing thought is to use the price of vanilla options to determine what measure the market is using and apply those parameters in the models to calculate prices for exotic options which are only thinly traded.

The choice of measure also impacts hedging. In a complete market, continuous delta-hedging is perfect at all times and the variance of the hedge is 0. In an incomplete market, zero variance is not possible. Two possible choices are to hedge to minimize final variance or to minimize the day-to-day variance, see [Jos03] Section 15.5.

6.3 Lévy Processes

As was the case in Section 1.2, it makes sense to start with arithmetic random walks and define price processes as their geometric counterparts. In order for an arithmetic process $X = \{X_t\}$ to serve it must satisfy a very special condition, one we have used repeatedly for Brownian motion. Namely, we must be able to: (1) divide the fundamental interval, $[0, T]$ into arbitrary subintervals $\Delta t = T/n$, (2) simulate identical and independent random increments ΔX_i on each subinterval, and (3) add the increments together, $X = \sum_{i=1}^{n} \Delta X_i$, and get the same result statistically, that is in terms of probability density, as for any other subdivision. Such a process is said to be *infinitely divisible*.

Lévy processes are exactly those that are infinitely divisible. A Wiener process is an example of a Lévy process. Like a Wiener process, a Lévy process $L = \{L_t\}$ satisfies $L_0 = 0$ and the axioms of independent and stationary increments:

1. Every increment $L_{t+h} - L_t$ depends only on L_t and not on $L = \{L_s, 0 \le s \le t\}$.
2. The distribution of $L_{t+h} - L_t$ does not depend on t, it has the same distribution as L_h.[1]

As we will see, a Lévy process can have jumps. By a jump we mean $\Delta L_t = \lim_{\epsilon \downarrow 0} L_{t+\epsilon} - \lim_{\epsilon \downarrow 0} L_{t-\epsilon}$. But the probability of a jump at any given value of t is 0. Note that one can always assume that a Lévy process is right continuous

[1] These conditions imply the infinite divisibility property since $L_t = L_{\frac{t}{n}} + (L_{\frac{2t}{n}} - L_{\frac{t}{n}}) + (L_{\frac{3t}{n}} - L_{\frac{2t}{n}}) + \ldots + (L_t - L_{\frac{(n-1)}{n}t})$.

and has left limits at every point, $L_t = lim_{\epsilon \downarrow 0} L_{t+\epsilon}$ and $\lim_{\epsilon \downarrow 0} L_{t-\epsilon}$ exists, the latter may be denoted as L_{t-}. Sometimes this requirement is called the *cadlag property.*[2] The reason for this choice is that, given a specific time in the future, say t_1, the value of the process at t_1 cannot be predicted with complete confidence from its values at times $t < t_1$ leading up to t_1, the process might undergo a jump at that time. If left continuity were made the choice, it would be possible to make the stated prediction.

6.3.1 The Poisson Process

Besides Wiener processes there are several known Lévy processes. The simplest is pure drift, $L_t = \mu t$. This and the Wiener process are the only two that are continuous, all others have jumps. The simplest non-continuous Lévy process is the *Poisson process* $\mathrm{Po}(\lambda)$ (here we have put $t = 1$, because of the infinite divisibility condition the Poisson parameter for an arbitrary time t is λt). The Poisson random variable $N_t \sim \mathrm{Po}(\lambda t)$ denotes the number of events, in our case jumps, which occur in the interval $[0, t]$. N_t is non-negative integer valued, $N_t = 0, 1, \ldots$; λ is called the *intensity* parameter.

The probability density for N_t is given by

$$\Pr(N_t = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \tag{6.4}$$

where k is the number of jumps. The expectation, that is mean, of the Poisson random variable is λt. The variance is also λt.

The events themselves arrive at increments of time Δt according to the exponential distribution $E(\lambda)$ where λ, the same λ as above, is the *event rate*. The cumulative distribution function of $E(\lambda)$ is

$$F(t) = 1 - e^{-\lambda t}. \tag{6.5}$$

In fact N_t can be simulated by sampling $E(\lambda)$ until the time increments sum to t, a sample $N_t = k$ is returned as the greatest integer k such that

$$\sum_{i=1}^{k} \Delta t_i < t \quad \text{where } \Delta t_i \sim E(\lambda). \tag{6.6}$$

The Δt_i are called the *inter-arrival times*. The event arrival times themselves, t_i, are given by

$$t_i = \sum_{j=1}^{i} \Delta t_j, \quad i = 1, 2, \ldots, k. \tag{6.7}$$

[2] From the French for the same phrase, ‘continue à droite et limites à gauche’. In French the spelling is càdlàg.

A sample $\Delta t_i \sim E(\lambda)$ is obtained as follows, (see (A.16))

$$\Delta t_i = \frac{-1}{\lambda}\log(1-U) \quad \text{where} \quad U \sim U(0,1).^{3}$$

With these preliminaries in hand, the *Poisson process with drift* μ is defined by

$$L_t = \mu t + \sum_{k=1}^{N_t} J \tag{6.8}$$

where J is the fixed jump size. This is an infinitely divisible process because if $X \sim \mathrm{Po}(\lambda_1)$ and $Y \sim \mathrm{Po}(\lambda_2)$, then $X + Y \sim \mathrm{Po}(\lambda_1 + \lambda_2)$, (see the Exercises). The Poisson process is always nondecreasing (if $J > 0$), that is, stays the same value or increases. In order to make the drift meaningful, one can subtract the jump size times the expected number of jumps; this gives rise to the *compensated* Poisson process with drift,

$$L_t = \mu t + \sum_{k=1}^{N_t} J - \lambda J t. \tag{6.9}$$

In Fig. 6.3a we show an instance of a compensated Poisson process. This is an *event-to-event* simulation in that time moves forward from one event to the next thus highlighting the jumps. The events are generated according to (6.7), see Algorithm 24.

Algorithm 24. Compensated Event-to-Event Simulation

```
inputs: t, λ, J (jumpsize)
X = 0;
simTime = 0;
plot(simTime,X);
arrivalArray = poissonArrivals(λt);   ▷use (6.7)
N_t = arrivalArray length;   ▷number of jumps
j = 0;   ▷N_t could be 0
for ...   ▷infinite loop
  j = j+1;   ▷update event index
  if j > N_t break out of loop
    ▷increment simTime
  ΔsT = arrivalArray[j] - simTime;
  simTime = simTime + ΔsT;   ▷move to next jump
  ΔX = −JλΔsT;   ▷pro-rated compensation
  X = X + ΔX;
  plot(simTime,X)   ▷before jump
  X = X+J;   ▷add in the jump
```

[3] Note that $U < 1$ for uniform random number generators, but $U = 0$ is possible. Since $1 - U$ is uniform if U is, it is tempting to save an operation and use the latter; but this comes with the risk of computer overflow in the middle of a calculation.

```
  plot(simTime,X)   ▷vertical gap
endfor
ΔsT = t - simTime;   ▷since final jump
X = X − JλΔsT;
simTime = t;   ▷final time
plot(simTime,X);
```

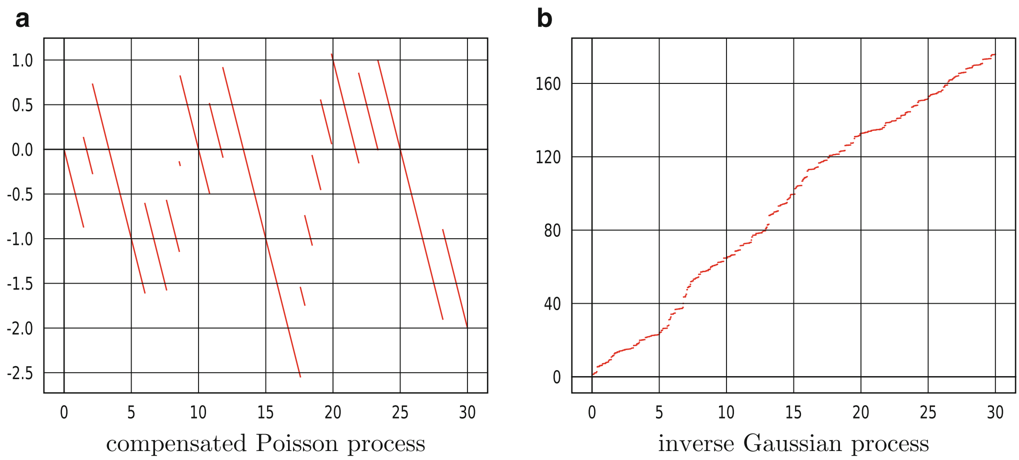


Fig. 6.3. Event to event simulations of Lévy pure jump processes

6.3.2 The Inverse Gaussian Process

The inverse Gaussian distribution has two parameters denoted by a and b. The first is a shifting parameter and has units of reciprocal time; larger a shifts the density to the right. The second is a spreading parameter, smaller b widens the density. The density itself is given by

$$f_{IG}(x;a,b) = \frac{ae^{ab}}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}\left(\frac{a^2}{x}+b^2x\right)}, \quad x > 0. \tag{6.10}$$

Figure 6.4 shows the density function for two sets of parameters. The mean and variance of an inverse Gaussian are

$$\mu_{IG} = \frac{a}{b} \qquad \mathrm{var}_{IG} = \frac{a}{b^3}. \tag{6.11}$$

Since the density is only defined on the positive real line, the IG process is always nondecreasing; such a process is called a *subordinator*. However a process may be defined by differences of two independent inverse Gaussians to have both positive and negative jumps. We investigate this possibility in Section 6.8.

The Lévy process defined by the inverse Gaussian is a pure jump process, see Fig. 6.3b. We discuss this in the next section.

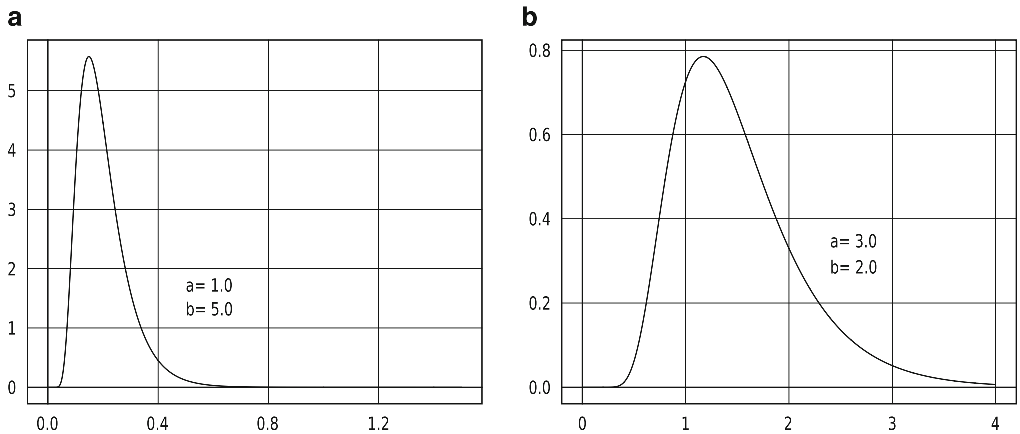


Fig. 6.4. The inverse Gaussian density for two parameter sets. Parameter a shifts the density to the right, parameter b narrows the density (for larger b). Note that the density is only defined for $x > 0$. It follows that an ARW based on this density can only move to the right

The density (6.10) gives the distribution of the end points of the process, that is at the end of the random walk, much as the normal distribution gives the end point distribution of a Brownian motion.

A random walk based on the inverse Gaussian is simulated exactly as before: the interval $[0, T]$ is subdivided into, say, n subintervals, and the update goes subinterval by subinterval, see Algorithm 25. This is a point-to-point simulation; a point-to-point path will not show jumps as they occur between the steps of the walk. Example runs of the algorithm are shown in Fig. 6.5.

Algorithm 25. Arithmetic Random Walk IG

`inputs:` $X_0 = 0$, T, Δt, a, b
$n = T/\Delta t$ ▷number of iterations in time T
`for` $t = 1, \ldots, n$
  $I \sim IG(a\Delta t, b)$ ▷`an IG sample, see A.10`
  $\Delta X_t = I$
  $X_t = X_{t-1} + \Delta X_t$
`endfor`
  ▷the last X_t is an outcome of $IG(aT, b)$

6.4 Lévy Measures

Associated with each Lévy process is a unique set valued function $\nu(A)$ called the *Lévy measure*. The meaning of the measure is

$\nu(A)$ is the intensity (arrival rate) of the Poisson process for jumps of sizes in A for the path L_t, $0 \leq t \leq 1$.

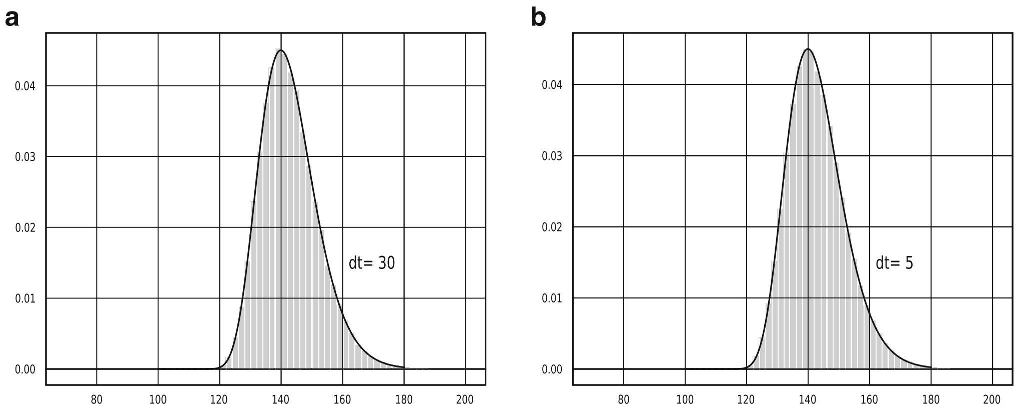


Fig. 6.5. These figures demonstrate the infinite divisibility property of the inverse Gaussian density. In (**a**) a $T = 30$ process is simulated in one step using Algorithm 25. In (**b**) the process is simulated adding six steps of size $dt = 5$. In each case the density $f_{IG}(x; 30a, b)$ is overlaid on the histogram, $a = 1$, $b = 0.7$

In particular if the measure is given by a density

$$\nu(dx) = h(x)\,dx \tag{6.12}$$

then $h(x)$ is the intensity for jumps of size x.

A Lévy measure has the same properties as a probability distribution except that it must have zero mass at the origin and its total mass may be infinite. The latter would be due to having a countable infinity of jumps of very small size. If the total mass is infinite,

$$\nu(\mathbb{R}) = \int_{-\infty}^{\infty} \nu(dx) = \infty,$$

the Lévy process has *infinite activity*. In this case there are infinitely many jumps on every interval (closed and bounded). Even if the process has infinite activity, it is always the case that it is *square summable* in the following sense

$$\int_{\mathbb{R}} \min(1, |x|^2)\nu(dx) < \infty. \tag{6.13}$$

The Lévy measure for the pure drift process and the Wiener process is null. For the Poisson process it is given by

$$\begin{aligned}\nu(A) &= \begin{cases} \lambda & \text{if } J \in A, \\ 0 & \text{otherwise} \end{cases} \\ &= \lambda \mathbb{1}_A(J). \end{aligned} \tag{6.14}$$

The Lévy measure for the inverse Gaussian process is given by a density

$$\nu_{IG}(dx) = \frac{a}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}b^2 x} dx. \tag{6.15}$$

In Fig. 6.3b we show an instance of an inverse Gaussian process. This is an event-to-event simulation and is somewhat complicated. An approximation can be made as follows. Given $\epsilon > 0$, chose positive numbers

$$\epsilon = c_0 < c_1 < \ldots < c_{d+1}. \tag{6.16}$$

For each interval $[c_i, c_{i+1})$, $i = 0 \ldots, d$, let $\mathrm{Po}_i(\lambda_i)$ be an independent Poisson process with intensity given by the Lévy measure of the interval,

$$\lambda_i = \nu([c_i, c_{i+1})) = \int_{c_i}^{c_{i+1}} h(x)\, dx. \tag{6.17}$$

The jump size J_i should be chosen so that the variance of the Poisson process Po_i matches that part of the variance of the Lévy process for that interval,

$$J_i^2 \lambda_i = \int_{c_i}^{c_{i+1}} x^2 \nu(dx). \tag{6.18}$$

To carry out the simulation, the event times for all d processes are sampled in advance. They are then combined but with each identified to its corresponding jump size, and sorted from early to late.[4] Then the simulation may proceed event-to-event as in Algorithm 24. When each event comes due, increment the process X_t using that event's corresponding jump size.

The above does not account for jumps of smaller size than ϵ. They may be handled, if necessary, by approximating all the small jumps by a Wiener process with drift. The parameter $\sigma(\epsilon)$ is given by

$$\sigma^2(\epsilon) = \int_0^{\epsilon} x^2 \nu(dx) \tag{6.19}$$

and the drift is given by

$$\mu(\epsilon) = \int_0^{\epsilon} x \nu(dx). \tag{6.20}$$

6.5 Jump-Diffusion Processes

By combining a Wiener process with a jump process we have what is called a jump-diffusion process. Let $F(\cdot)$ be a probability distribution (not necessarily a Lévy process) and let $J \sim F$ denote its samples. We may define a Lévy process by

$$L_t = \mu t + \sum_{k=1}^{N_t} J_k - t\lambda \mathbb{E}(J). \tag{6.21}$$

This is called a (compensated) *compound Poisson* process with drift. Just as in the compensated Poisson process of (6.9), the arrival times of the jump events

[4] A simple technique is to maintain two arrays, one with times, t and the other with jump size, J. Now sort the times array via a rank permutation r_i so that $t_{r_1} < t_{r_2} < \ldots$, see Chapter E. Then jump size J_{r_i} corresponds to time t_{r_i}.

are exponential; the only difference here is that the jumps can vary in size according to F. Since the jump sizes may vary, the compensation is determined by the average or expected jump size as shown in (6.21). The Lévy measure for a compound Poisson process is $\lambda F(dx)$.

In financial applications F is often taken to be the normal distribution. Such an example is shown in Fig. 6.6a (uncompensated in this example).

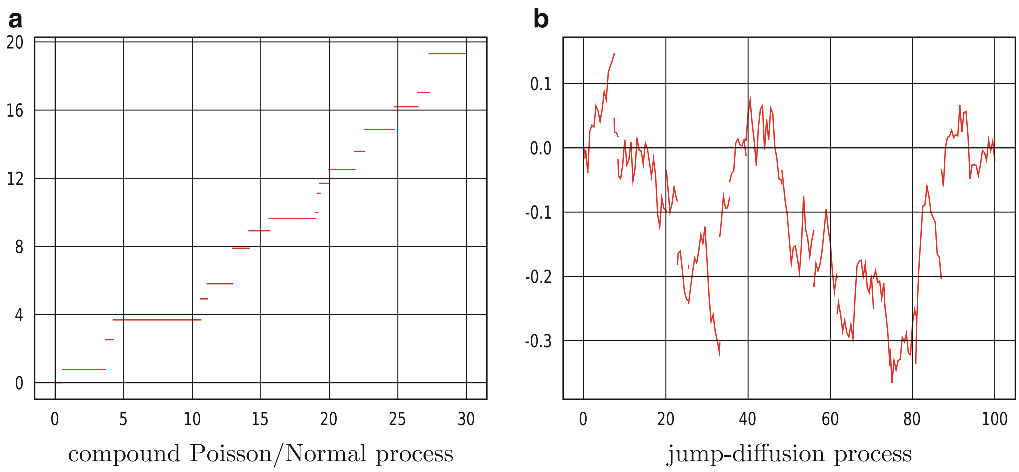


Fig. 6.6. (**a**) Shows an instance of an uncompensated compound Poisson process with normally distributed jump sizes. A jump-diffusion process is shown in (**b**), the jump sizes are random normal variates (with independent mean and variance from that of the diffusion process)

The most general Lévy process is obtained by combining all four types of processes into one: drift, diffusion (Wiener), compensated compound Poisson, and an infinite activity pure jump. The combination of the first three of these is called a *jump-diffusion* process.

$$L_t = \mu t + \sigma W_t + \left(\sum_{k=1}^{N_t} J_k - t\lambda \mathbb{E}(J_k) \right). \tag{6.22}$$

A jump-diffusion process always has finite activity. Further it is a martingale if and only if $\mu = 0$.

A jump-diffusion path is simulated from event-to-event exactly as in Algorithm 24. However, there is the additional step that the jump size be drawn from F before adding the jump to X,

$$J \sim F$$
$$X = X + J.$$

The additional difference from the cited algorithm is that, if compensation is used, the jump size to use for it is the constant expected jump size, $\mathbb{E}_F(J)$.

6.6 Application to Asset Pricing

As we learned in Chapter 1, an arithmetic random walk is an inadequate model for asset prices; a geometric walk is required. Therefore it is the log returns of the asset that must be modeled by the Lévy process

$$\frac{dS_t}{S_{t-}} = dL_t, \tag{6.23}$$

(S_{t-} is the left limit of S at t; by the cadlag property it always exists). By conducting simulations of L_t, $0 \le t \le T$, as described in the previous sections, and using (6.23) we obtain a histogram approximation of the maturity price distribution and statistical information on the paths of the process leading to maturity. Of considerable importance in this regard is the *Martingale preserving Property* stating that if $(L_t)_{t\ge 0}$ is a martingale, then so is $(S_t)_{t\ge 0}$.

Generating asset prices via (6.23) is called the *stochastic exponential* method. It is the method we will use. An alternative is the *exponential-Lévy* model given by

$$S_t = S_0 e^{L_t}.$$

The two approaches are equivalent and are related by the Itô Lemma (B.11).

As was the case for diffusion increments, jump increments are taken proportional to the current asset price S. For example, $S_{\text{new}} = S_{\text{old}} J$. Then $\Delta S = S_{\text{new}} - S_{\text{old}} = S_{\text{old}}(J-1)$. If $J > 1$ then the increment is positive. If $0 < J < 1$ then the increment is negative. And if $J < 0$ then the new price is negative; downward jumps must not exceed the current stock price. An alternative is to put $\Delta S = S_{\text{old}}(e^J - 1)$. Since $e^J > 0$ for all J, the non-negativity requirement is automatically fulfilled. By the series expansion for the exponential function, to first order, $e^J - 1 = J$. In this section we follow Merton, [Mer76], and put $\Delta S = S(J-1)$.

Recall that, for the drift-diffusion process of Chapter 1, we were able to derive the maturity distribution analytically, see (1.18). In that case S_T is distributed lognormally. However things are not so easy for an arbitrary Lévy processes. In general the maturity distribution is the solution of the stochastic differential equation (SDE) for $\log(S_t)$ where S_t is as in (6.23). The differential of $\log(S_t)$ is given by Itô's Lemma, see (B.11), page 227. In appendix Section B.2 we solve this for the drift-diffusion process (Wiener process with drift) obtaining the lognormal as its solution. Solving it for the several processes described in the previous sections is beyond the scope of this text. Thus we will content ourselves with the simulation of the end point via small steps. In that way we generate the paths too; as we have seen, they are needed in any case for several of the exotic options.

6.6.1 Merton's Model

Besides drift-diffusion there is another process for which the end point distribution may be determined, namely for jump-diffusion processes. Let L_t be an uncompensated jump-diffusion process and consider the product $S\,dL_t$ term by term. The drift and diffusion terms are $S\mu dt$ and $\sigma S dW_t$ as usual. If there is no jump at t, then the contribution from the jump term is 0. If $t = t_k$ is one of the jump event times then S jumps to SJ so the increment is $dS = SJ - S = S(J-1)$. Therefore we have

$$dS_t = S_t \mu dt + \sigma S_t dW_t + S_t \sum_{k=1}^{N_t} (J_k - 1)\delta_{t_k}(dt) \tag{6.24}$$

where the singular measure $\delta_{t_k}(dt)$ is equal to 1 if $t = t_k$ and 0 otherwise. Only one term of the sum will be non-zero for any t. Note that the value of S_t used as the multiplier for the jumps in (6.24) is the limiting value of S from the left, S_{t-}; at an event time t_k itself, S jumps to S_{t_k}.

By an extended version of Itô's Lemma, (B.11), the differential of $\log S_t$ is given by

$$d(\log S_t) = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t + d\left(\sum_{k=1}^{N_t} \log J_k \mathbb{1}_{t_k}(t)\right). \tag{6.25}$$

The last term signifies the following: it calculates that a difference in the sum, which can not be infinitesimal, at t is $\log J_k$ if $t = t_k$ and 0 otherwise. Again, only one term is non-zero for any value of t. Integrating (6.25) we get

$$\log S_t - \log S_0 = (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t + \sum_{k=1}^{N_t} \log J_k. \tag{6.26}$$

Upon exponentiation we arrive at the exponential-Lévy formulation

$$\begin{aligned} S_t &= S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t + \sum_{k=1}^{N_t} \log J_k} \\ &= S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \prod_{k=1}^{N_t} J_k. \end{aligned} \tag{6.27}$$

In Fig. 6.7a we show a typical path for a jump-diffusion simulation using lognormally distributed jump sizes, (b) depicts the maturity distribution. These figures were made using Algorithm 26.

Algorithm 26. Jump-Diffusion GRW, Point-to-Point Simulation

```
inputs: T, dt, λ, F(·) (jumpsize distribution)
        S_0, μ, σ
S = S_0;
n = T/dt;   ▷number of steps
simTime = 0;
arrivalArray = poissonArrivals(λT);   ▷use (6.6)
N_T = arrivalArray length;   ▷number of jumps
sDX=1;   ▷step index, point at next step
jDX = 1;   ▷jump index, point at next jump
```

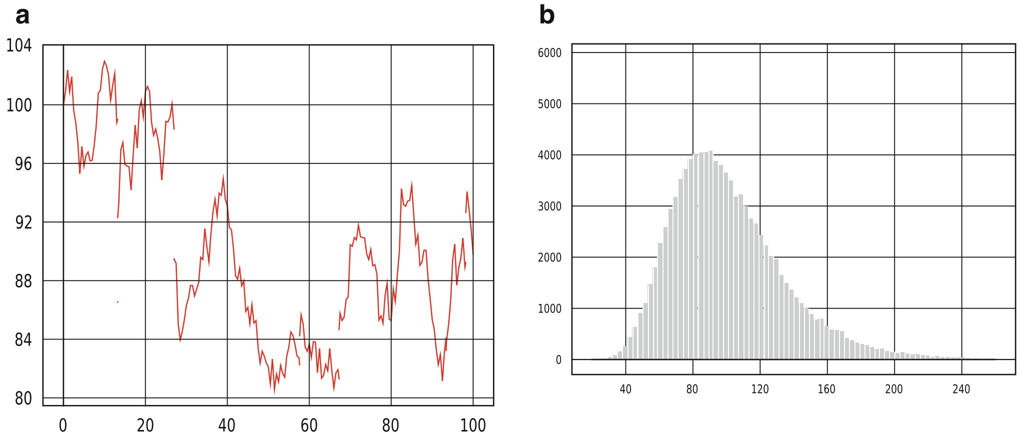


Fig. 6.7. Characteristics of a jump-diffusion geometric random walk. A typical path is shown in (**a**) and the end-point distribution is shown in (**b**). The drift-diffusion parameters are: $\mu = 3\,\%$ and $\sigma = 40\,\%$. The jumps are distributed as $LN(-0.0032, 0.08^2)$ with event rate $\lambda = 0.1$ per day

```
for ...   ▷infinite loop
  if( jDX > N_T or sDX * dt < arrivals[jDX] )
    dst = sDX*dt - simTime; //increment in simtime
        ▷do a diffusion
    dS = S(μ * dst + σ√(dst) Z);   ▷Z ~ N(0,1)
    S = S + dS;
    simTime = simTime+dst;   ▷update simTime
    sDX = sDX+1;   ▷point at next step
  else   ▷jump event
        ▷do a diffusion since last step
    dst = arrivals[jDX]-simTime;   ▷increment in simTime
    dS = S(μ * dst + σ√(dst) Z);   ▷Z ~ N(0,1)
    S = dS + S;
    simTime = arrivals[jDX];   ▷update simTime
    S = S*J;   ▷J ~ F(·), after jump price
    jDX= jDX + 1;   ▷point to next jump time
  endif
    ▷check if done
  if sDX > n break out of loop
endfor
S_T = S;
```

6.6.2 Jump-Diffusion Risk-Free Growth

In order to use a Lévy process for market predictions, the process must be a martingale. It is possible to achieve this by adjusting the drift of the process, this is a consequence of the *Girsarnov Theorem*, [Bjo04]. However, except for the Poisson pure jump and Wiener processes, the martingale measure is not unique. From the discussion in Section 6.1, this means the market is incomplete and there is no one no-arbitrage price. Notwithstanding uniqueness, next we show how to calculate a no-arbitrage drift for the jump-diffusion process.

The infinitesimal growth rate of the jump-diffusion model may be calculated from (6.24). The drift term being constant, its expected value is itself $\mu S_t dt$, and the expected value of the diffusion term is 0 because the expected value of a Wiener process is that. With regard to the jump term, the expected value is the expected jump size times the expected arrival rate of the jumps. Since the latter arise according to a Poisson distribution with intensity λ, we may write

$$\mathbb{E}\left(S_t \sum_{k=1}^{N_t} (J_k - 1)\delta_{t_k}(dt) \right) = S_t \mathbb{E}(J-1)\,(\lambda dt)\,. \tag{6.28}$$

For example, if the jumps are distributed according to $N(\mu_J, \sigma_J^2)$, then

$$\mathbb{E}(dS_t) = \mu S dt + (\mu_J - 1) S_t \lambda dt. \tag{6.29}$$

And if they are distributed according to $LN(\alpha, \beta^2)$, then

$$\mathbb{E}(dS_t) = \mu S dt + (e^{\alpha + \frac{1}{2}\beta^2} - 1) S_t \lambda dt. \tag{6.30}$$

On the other hand, in order to be risk-neutral, the expected growth rate should be $S_t r dt$ where r is the risk-free rate. Hence in the case of normally distributed jumps

$$r = \mu + \lambda(\mu_J - 1)$$

so that

$$\mu = r - \lambda(\mu_J - 1). \tag{6.31}$$

And in the case of lognormally distributed jumps,

$$\mu = r - \lambda(e^{\alpha + \frac{1}{2}\beta^2} - 1). \tag{6.32}$$

Using these drifts in the price simulations for these jump diffusion processes is the equivalent of using the risk-free rate in the GBM simulations. One still has to discount back the option payoffs.

6.6.3 Calculating Prices for Vanilla Options

We will use the Monte Carlo method to obtain option prices by simulating the jump-diffusion model. The only change to Algorithm 26 is to add the option payoff function $G(S_T)$ (for path independent options) at the end of the loop and then discount this back to $t = 0$, see Algorithm 27.

If the jump sizes are to be normally or lognormally distributed, use (6.31) or (6.32) as appropriate for the drift.

In Fig. 6.8 we compare jump-diffusion ending price distributions for both normal and lognormal jump sizes against that of geometric Brownian motion.

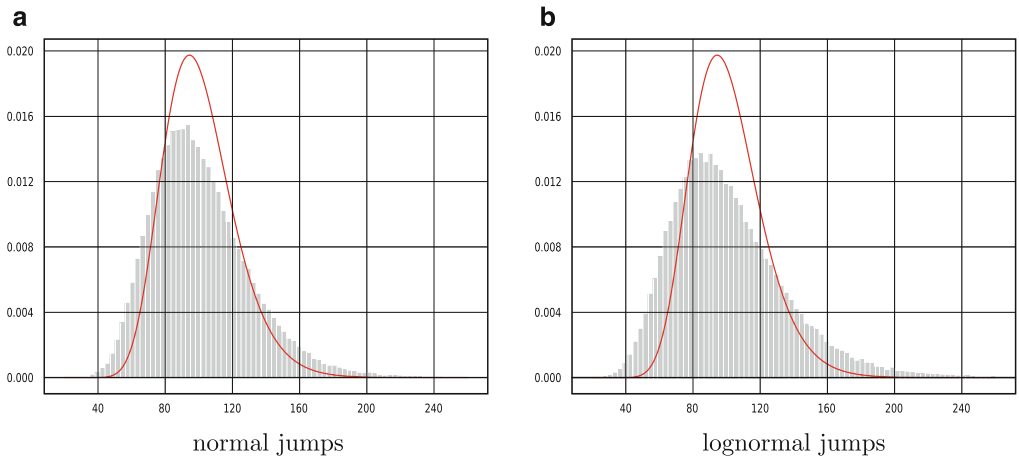


Fig. 6.8. Comparisons between GBM maturity distribution (in *red*) and a jump diffusion maturity histogram. For the GBM $r = 3\,\%$, vol $= 40\,\%$, T $= 100$ days. The Poisson process is $\lambda = 0.1$ per day. In (**a**) the jumps are $N(1, 0.06^2)$; in (**b**) they are $LN(-0.0032, 0.08^2)$

In each case the jump-diffusion prices show greater spread and so we can expect higher option prices as if the volatility were greater.

Algorithm 27. Monte Carlo Jump-Diffusion Pricing Algorithm

```
inputs: S_0, K, T, r, σ, nTrials
E = 0
for i = 1,...,nTrials
   S = S_0
   ▷use Algorithm 26 to generate S_T
   E = E + G(S_T)
end for
option price = e^{-rT} E/nTrials
```

Figure 6.9 illustrates a comparison between option prices under the Black-Scholes model and those of a jump diffusion model. As previously mentioned, the jump diffusion model is incomplete and therefore there is no unique no-arbitrage price. In the figure the risk-neutral value of (6.31) was used.

Exotic Options

Many of the exotic options go just as discussed in Chapter 4 since we are able to simulate instances of the price paths for Lévy processes. However others require some care in the use of Lévy jump process. In the case of a barrier option, a jump can carry the underlying's price across the barrier triggering the corresponding action. And again, in our shout boundary approach to shout options, a jump

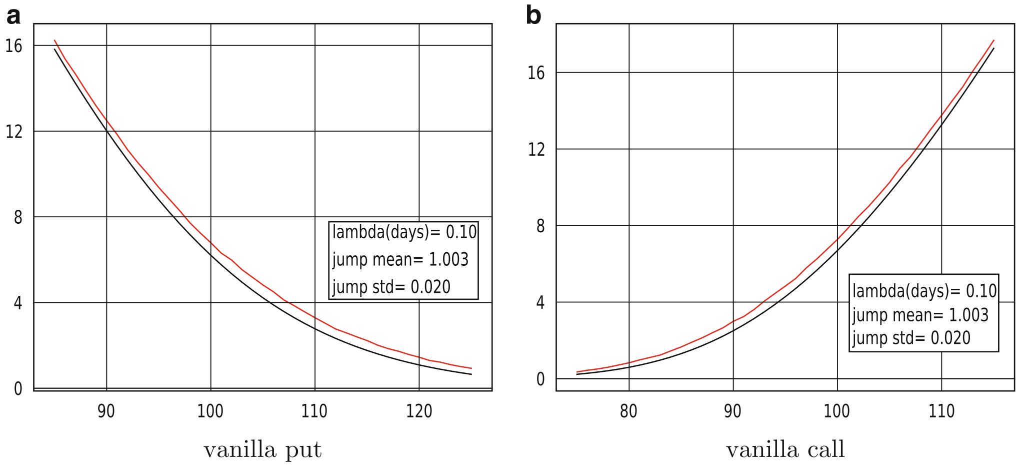


Fig. 6.9. Black-Scholes put and call values (*black*) as compared with those for the jump diffusion model (*red*) using normal distribution jumps plotted against stock price S. The option characteristics are: $K = 100$, $T = 60$ (days), $r_f = 3\,\%$, vol$=40\,\%$. The jump parameters are as indicated. The jump diffusion ATM put costs 6.75 vs 6.20 for Black-Scholes, a 9 % increase. The jump diffusion ATM call costs 7.21 vs 6.70 for Black-Scholes, an 8 % increase

can carry the price across the boundary calling for a shout. These options must be simulated event-to-event and Brownian bridges must be considered between events (4.1), see [CT04].

As previously reported, simulation prices of vanilla options for a range of parameter values, for example, drifts in the jump-diffusion model, when compared to their market prices can be used to determine the exact (current) parameter values applicable. Then these values are used to calculate exotic option prices.

6.7 Time Shifted Processes

In Section 6.3 we encountered an example of a subordinator, a process that is either constant or increasing. One of the main uses of such a process is to replace the smoothly moving calendar time by the subordinator process. In this way an entirely new class of Lévy processes can be generated. In finance such a process is used to simulate *business time* since businesses tend to operate from event to event. If τ_t is a subordinator and X_t an overlying Lévy process, then the *subordinated* or *time changed* process is

$$L_t = X_{\tau_t}. \tag{6.33}$$

Often a Wiener process is used as the overlying process.

Figure 6.10 shows a typical Gaussian subordinator path in (a) and the end point histogram in (b). The simulation, done via Algorithm 28, proceeds in regular time increments Δt as usual. But the Wiener process is based on an inverse Gaussian time step. The path of such a process often has large movement as shown but they are not exactly jumps since they occur Δt time units apart. These can occur when there is a long time period commanded by the subordinator, then the Gaussian step has a chance to be large. The end point distribution is depicted in (b). It shows a narrow peak but a very wide base. A small number of large jumps in the same direction can account for this phenomenon.

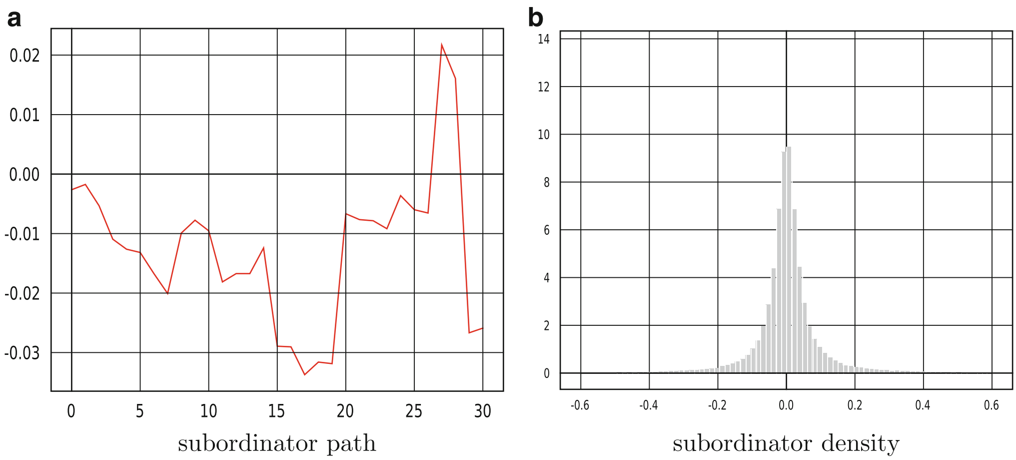


Fig. 6.10. Illustrated in (**a**) is a typical Gaussian subordinator path. Illustrated in (**b**) is an end point histogram. It shows a narrow peak but a very wide base. A small number of large jumps in the same direction can account for this phenomenon

Algorithm 28. Gaussian Subordinator Simulation

inputs: T, dt, X_0, μ, σ, a, b
$X = X_0$;
$n = T/dt$; $\triangleright$number of steps
for $i = 1, 2, \ldots, n$
  $\tau \sim IG(a * dt, b)$ $\triangleright$incr. in time via subordinator
  $\Delta X = \mu\tau + \sigma\sqrt{\tau}Z$; $\triangleright Z \sim N(0, 1)$
  $X = X + \Delta X$;
endfor
$X_T = X$;

6.8 Heavytailed Distributions

The normal distribution is widely used in finance, but often it is only an approximation to the actual distribution of the circumstance. One piece of evidence for this is that events which should only occur once in thousands of years, instead occur 2 or 3 times in 40. Some have labeled these as “six-sigma events” since, if the normal distribution applied, their probability of occurring would be that in the upper tail six standard deviations from the mean. It can be inferred that the actual distribution governing, for example price movements, has greater probability for extreme events than is accounted for by the normal distribution. That is to say, the tails of the distribution should be fatter.

It is for this reason that financial mathematicians study *heavytailed* distributions, densities decaying more slowly in the tails than the normal. In this section we examine two examples. The first is the widely known family of *t distributions*, sometimes known as the *Student t*. For the second example, we show that a heavytailed distribution can be constructed as the difference between two independent subordinators.

6.8.1 Student's t-Distribution

The t distribution has a single parameter, $\nu > 0$, known as the *degrees of freedom* (dof). Some members of the family are shown in Fig. 6.11.

The t probability density function is given by

$$f_\nu(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}. \tag{6.34}$$

In this $\Gamma(\cdot)$ is the *gamma function* defined by the integral

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt.$$

In (6.34) the gamma terms are just constants contributing to normalization.

The gamma function is an extension of the factorial function. Using integration by parts, it is easy to see that it satisfies the recursion

$$\Gamma(z+1) = z\Gamma(z).$$

And by direct integration we get

$$\Gamma(1) = 1.$$

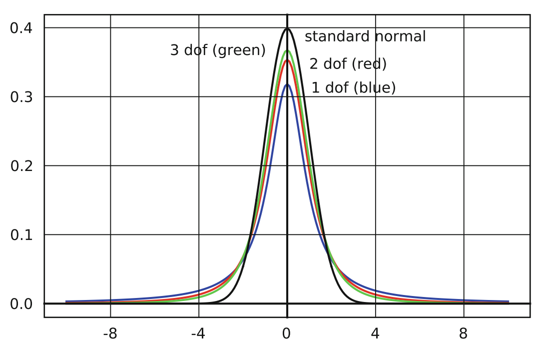


Fig. 6.11. Student t densities for degrees of freedom equal to 1, 2, 3, and infinity (the standard normal)

From these two facts it is easy to see that, for integers, gamma is the factorial function,

$$\Gamma(n) = (n-1)! \qquad n \text{ a positive integer.} \tag{6.35}$$

The only other commonly needed value of gamma is for $z = 1/2$ and that value is well-known,

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}. \tag{6.36}$$

With these preparations, we may write the first five members of the t_ν family (for integral ν)

$$\begin{aligned}
f_1(x) &= \frac{1}{\pi(1+x^2)} \\
f_2(x) &= \frac{1}{2\sqrt{2}}(1+x^2/2)^{-3/2} \\
f_3(x) &= \frac{2}{\pi\sqrt{3}}(1+x^2/3)^{-2} \\
f_4(x) &= \frac{3}{8}(1+x^2/4)^{-5/2} \\
f_5(x) &= \frac{8}{3\pi\sqrt{5}}(1+x^2/5)^{-3}.
\end{aligned}$$

The $\nu = 1$ density is also known as *Cauchy's density.* As $\nu \to \infty$ the t_ν distribution tends to the standard normal density. As seen in Fig. 6.11 the tails become less heavy as ν increases.

The t densities are symmetric about $x = 0$ and hence have mean equal to 0. The $\nu = 1$ and $\nu = 2$ densities do not have bounded square integrals and therefore their variances are infinite. But for $\nu > 2$ the variances are finite,

$$\text{var}(t_\nu) = \frac{\nu}{\nu - 2} \qquad \nu = 3, 4, \ldots. \tag{6.37}$$

Sampling from the Student-t

The most widely used method for sampling from the t_ν distribution is due to Baily [Bai94]. It is valid for all $\nu > 0$ The Baily algorithm is a simple modification of the Marsaglia-Bray algorithm, Algorithm 2, for the standard normal.

Algorithm 29. Baily's Algorithm for t_ν Samples

```
repeat
  U ~ U(0,1); U = 2U-1;  ▷uniform on -1 to 1
  V ~ U(0,1); V = 2V-1;  ▷a point in the sqr.
until W=U^2+V^2<=1
C=U^2/W; R = ν(W^(-2/ν) - 1)
T = √(RC);
if( U~ U(0,1) <.5 )
  return T
else
  return -T
```

6.8.2 Difference Subordinator Densities

Although we have only studied two Lévy densities besides the normal, namely the Poisson and the inverse Gaussian, there are many known. And we have shown how Lévy processes can be constructed as compound Poisson processes or by time change. Many of the Lévy process densities are heavytailed. Here we show another method for constructing a heavytailed density guaranteed to be infinitely divisible.

Let X_t and Y_t be subordinators and put Z_t equal to their difference,

$$Z_t = X_t - Y_t. \tag{6.38}$$

Then Z_t is infinitely divisible and therefore a Lévy process. For example, X and Y can be the same subordinator.

By independence, the mean of Z_t is just the difference of that of X_t and Y_t, and the variance is the sum $\mathrm{var}(Z_t) = \mathrm{var}(X_t) + \mathrm{var}(Y_t)$. In Fig. 6.12 we show the difference between two inverse Gaussians for two different parameter sets. Also shown is the normal density having the same mean and variance. At about 2σ the difference density exceeds the normal showing that its tails are heavy.

Application to Asset Prices

An alternate formulation of (6.23) derives from including the risk-free rate in the stochastic exponential separately,

$$dS_t = r_f S_{t-} dt + S_{t-} dZ_t. \tag{6.39}$$

This formulation is completely general and holds for any Lévy process Z_t. From the martingale preserving property, in this formulation the process $e^{-r_f t} S_t$ is a martingale if and only if $\mathbb{E}(Z_1) = 0$.

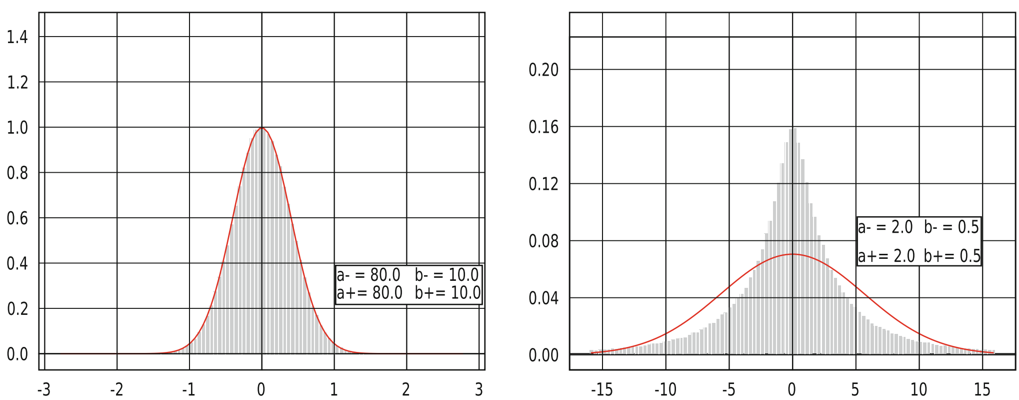


Fig. 6.12. Difference inverse Gaussian histograms for two quite different parameter sets. Overlaying each figure is the normal distribution with the same mean and variance. Both inverse Gaussians are heavytailed

Problems: Chapter 6

1. Show that if $X \sim \mathrm{Po}(\lambda_1)$ and $Y \sim \mathrm{Po}(\lambda_2)$, then $X + Y \sim \mathrm{Po}(\lambda_1 + \lambda_2)$. Hint:

$$\begin{aligned}\Pr(X+Y \le z) &= \sum_{y=0}^{z} \Pr(X \le z-y | Y=y) \Pr(Y=y) \\ &= \sum_{y=0}^{z} \frac{\lambda_1^{z-y} e^{-\lambda_1}}{(z-y)!} \frac{\lambda_2^{y} e^{-\lambda_2}}{y!}.\end{aligned}$$

2. The *skew* of a random variable X is defined as

$$\mathrm{skew} = \mathbb{E}\big((X-\mu_X)^3\big)/\mathrm{std}_X^3. \tag{6.40}$$

Given data $x_1, x_2, \ldots, x_n$ an estimator for skew is

$$\overline{\mathrm{skew}} = \frac{\sum_1^n (x_i - \overline{x})^3}{n\overline{s}^3}$$

where $\overline{x}$ and $\overline{s}$ are empirical mean and standard deviation. Being a symmetric distribution the normal has 0 skew. Calculate the empirical skew of the log returns ($\log \frac{S_{i+1}}{S_i}$), for 3 stock equities of your choice using daily prices over the last 2 years. (Use the FIMCOM database or finance.yahoo for the prices, see Section 1.7.3 page 25.)

3. The *kurtosis* of a random variable X is defined as

$$\mathrm{kurtosis} = \mathbb{E}\big((X-\mu_X)^4\big)/\mathrm{std}_X^4. \tag{6.41}$$

Given data $x_1, x_2, \ldots, x_n$ an estimator for kurtosis is

$$\overline{\mathrm{kurtosis}} = \frac{\sum_1^n (x_i - \overline{x})^4}{n\overline{s}^4}$$

where $\overline{x}$ and $\overline{s}$ are empirical mean and standard deviation. The kurtosis of the normal distribution is 3, cf. the footnote on page 15. Calculate the empirical kurtosis of the log returns ($\log \frac{S_{i+1}}{S_i}$), for 3 stock equities of your choice using daily prices over the last 2 years. (Use the FIMCOM database or finance.yahoo for the prices, see Section 1.7.3 page 25.)

4. (a) From market price data make a graph of implied volatility σ versus strike price K for call options on the S&P-500 for expiration maturities of T on the order of 30 days (near as possible). Do the same for $T = 60$ and 90 days. You now have a *volatility surface*, implied volatility versus strike and time.
(b) Do the same for put options.

5. The Gamma distribution, $G(\alpha, \lambda)$ has density given by

$$f_G(x : \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0. \tag{6.42}$$

Here $\Gamma(\alpha)$ is the gamma function of Section 6.8 and equals $(\alpha - 1)!$ if α is a positive integer. Show that the Gamma is infinitely divisible (empirically) by showing that the histogram for the sum of six samples of $G(1,\lambda)$ has the same density as $G(6,\lambda)$. Note that, for α a positive integer, then

$$W = \frac{-1}{\lambda}\log(\prod_1^{\alpha} U_i) \quad U_i \sim U(0,1) \tag{6.43}$$

is a sample from $G(\alpha,\lambda)$, [SM09].

6. (a) Make a chart similar to Fig. 6.9 showing the price of a put option using the jump diffusion model with lognormal jumps for stock prices versus the GBM model. In order to compare the results with jump diffusion using $N(\mu_J,\sigma_J^2)$ jumps, find α and β to match the mean and variance,

$$\mu_j = \mu_{LN} = e^{\alpha+\frac{1}{2}\beta^2}, \quad \sigma_J^2 = (e^{\beta^2}-1)\mu_{LN}^2.$$

   (b) Do the same for calls.

7. (a) Make a chart similar to Fig. 6.9 showing the price of a put option using a difference IG model for stock prices versus the GBM model. Use $a- = a+ = 41$ and $b- = b+ = 8$. What are the mean and variance of the difference IG process? (b) Do the same for calls.

8. (a) Work the Bermuda option Problem 5 of Chapter 4 assuming prices follow a jump diffusion with normal sized jumps. Be sure to report your jump parameters. (b) Repeat (a) using lognormal sized jumps.

9. Work the Bermuda option Problem 5 of Chapter 4 assuming prices follow a symmetric differential IG model, use equation (6.39). Be sure to report your model's parameters.

10. Recalculate Table 4.2 page 121 for barrier options assuming prices follow a jump diffusion model with normal sized jumps. Recall that the simulation must go event-by-event.

11. Recalculate Table 4.1 page 119 for Asian options assuming prices follow a differential IG process.

12. A portfolio consists of 100 shares each of stock A: $S_0 = 60$, $\mu = 8\,\%$, $\sigma = 40\,\%$; and B: $S_0 = 40$, $\mu = 3\,\%$, $\sigma = 20\,\%$. Their correlation is $\rho = 0.3$. After 6 months what is the probability of losing money and the expected gain of the portfolio if (a) prices follow a Gaussian GBM model? (b) a jump diffusion with normal jumps?

13. Work the VaR Problem 9 of Chapter 2 assuming prices follow jump diffusion with normal sized jumps.

   In the following assume a jump diffusion model for prices with normal jumps.

14. Analyze covered calls as in Table 5.2.

15. Analyze creditspreads as in Table 5.5.

In the following assume a (symmetric) difference IG model for prices.

16. Analyze iron condors as in Table 5.9.

17. Analyze the straddle strategy as in Table 5.7.

7

Kelly's Criterion

The nature of markets in finance is one of almost constant ups and downs. Against this environment investors make every effort to succeed in their investments relying on studies of equity fundamentals, market dynamics, experience and other factors to assure success. Fortunately most of the time their efforts are rewarded thus growing their investment capital. But not always. Individual stocks may plunge in value only to remain suppressed for a long period of time or even to never recover. And from time to time it gets worse; this can happen to the entire market. These are called *Bear markets*.

Thus investing is a proposition in which an investment generally pays off but occasionally takes a loss, possibly even a complete loss.[1] Mathematically we would say investing is a probabilistic enterprise with a positive expectation. In this regard, investing is like gambling with the difference being that gambling usually has a negative expectation.

Suppose an investor has P capital and places it all in investment A. This is usually not very wise. Assuming that investment A could completely fail, even though the probability might be very small, the investor could lose everything and be left with no investment capital at all, game over. The question with which we concern ourselves in this chapter is, what fraction of P should be invested in A? What criteria should be used to decide?

Working at Bell Labs in the early 1950s, John Kelly considered the problem of deciding how much money to risk on probabilistic outcomes that are nonetheless in your favor, [Kel56]. His motivation was speculation about gambling tips over noisy communication channels, but the work has been much more widely interpreted. Among others, it is the basis of a strategy for asset allocation in the stock market, the *Kelly Criterion*.

The criterion applies to investments having a finite time horizon upon which the investment is settled. Furthermore, such investments are to be repeated indefinitely. As a class, option trades fit these conditions very well. In the following we derive Kelly's Criterion and then investigate its application to option trading.

[1] The law provides that stockholders are not liable for the company's debts limiting their loss to 0.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7_7,

7.1 Kelly's Formula for a Simple Game

We start with a simple example. A trader with long experience in credit spreads determines that 40 % of the trades have failed and lost their complete investment. But the remaining 60 % have succeeded and earned an average return of 88 %. Under these conditions, what fraction of the trader's investment capital should be risked on a continuing series of these credit spreads?

It may be clearer to recast the problem as a gamble. The bet is on the outcome of a weighted coin that lands Heads with probability $p = 0.6$ (and Tails with probability $q = 1 - p = 0.4$). On Tails, the entire bet is lost; but on Heads the net gain is $\gamma = 88\,\%$ of the bet.

Gamblers describe payoffs in terms of *odds* (and often, probabilities too); for example $(a : b)$ means that winning increases one's bankroll by $\$a$ while losing decreases one's bankroll by $\$b$. On a per unit basis this is $(\frac{a}{b} : 1)$, that is, for \$1 bet, the gain is $\$(a/b)$ in case of a win.

In either formulation, we calculate the expected gain per unit bet as

$$E_{\text{gain}} = \gamma p - q, \tag{7.1}$$

or $E_{\text{gain}} = 0.128$ for the particulars given. The expectation is positive so the gamble is in the players favor and, over the long haul, the gambler expects to make 12.8 % per play. Therefore the gambler plans to repeat the bet indefinitely.

But how much to risk on each bet?

Since the bets will be made repeatedly, it would a mistake to risk it all each time – one loss leaves the gambler, or trader, with nothing. Instead, we suppose the amount bet on each play is a fixed fraction f of the gambler's bankroll. Letting F_0 denote the gambler's initial fortune, after one play the new fortune will be

$$F_1 = \begin{cases} \text{if win} & F_0 + \gamma f F_0 = (1 + \gamma f)F_0, \\ \text{if lose} & F_0 - f F_0 = (1 - f)F_0, \end{cases}$$

After N plays the fortune will be the product

$$F_N = (1 + \gamma f)^W (1 - f)^L F_0 \tag{7.2}$$

where W is the number of wins and L is the number of losses in the N plays. Kelly's key insight is this:

we should strive to maximize the expected growth rate of our fortune.

The growth rate after N plays, G_N, satisfies the equation $e^{G_N N} F_0 = F_N$, see (2.9),[2] and therefore

$$G_N = \frac{1}{N} \log \frac{F_N}{F_0} = \frac{W}{N} \log(1 + \gamma f) + \frac{L}{N} \log(1 - f). \tag{7.3}$$

[2] This is the effective growth rate for this discrete process; the actual growth rate follows from the equation $2^{g_N N} F_0 = F_N$.

Since W and L are random variables, so is G_N. As N increases indefinitely, the ratio W/N tends to p and L/N tends to q with probability 1; hence in the limit we obtain the expected growth rate

$$G = p\log(1+\gamma f) + q\log(1-f). \tag{7.4}$$

Maximize G with respect to f by setting the derivative to zero,

$$0 = \frac{p\gamma}{1+\gamma f} - \frac{q}{1-f} = \frac{p\gamma(1-f) - q(1+\gamma f)}{(1+\gamma f)(1-f)},$$

for a solution of

$$f = \frac{\gamma p - q}{\gamma} = \frac{E_{\text{gain}}}{\gamma}. \tag{7.5}$$

Note the simple form of the answer, the expectation over the gain, or sometimes referred to in brief as the edge over the odds. This is *Kelly's formula.*

A notable aspect of Kelly's Criterion is that, since only a fraction of the fortune is bet on each play, there is always a fraction held back – one's fortune can never be zero.

In Fig. 7.1 we show $N = 260$ plays of the simple game with $p = 0.6$, $q = 0.4$, $\gamma = 0.88$, and for various fractions f (with differing ordinate scales). From (7.5) the optimal betting fraction is

$$f = 0.6 - 0.4/0.88 = 0.145.$$

Therefore if the old fortune is F_0, the new fortune F_1 is

$$F_1 = \begin{cases} \text{if win} & (1 + 0.88 * 0.145)F_0 = 1.128F_0, \\ \text{if lose} & (1 - 0.145)F_0 = 0.855F_0. \end{cases} \tag{7.6}$$

The maximal growth rate, from (7.4) with f given by (7.5), is

$$\begin{aligned} G &= p\log(1+\gamma p - q) + q\log(1 - \frac{\gamma p - q}{\gamma}) \\ &= p\log(p(1+\gamma)) + q\log(\frac{q}{\gamma}(1+\gamma)) \end{aligned} \tag{7.7}$$

or, with the present numbers,

$$G = 0.0096.$$

After $N = 260$ plays the expected fortune is

$$\mathbb{E}(F_{260}) = F_0 e^{260*0.0096} = 12.15F_0,$$

a 12 fold increase.

The runs were obtained by simulation using the following code, and therefore each is only a single possible fortune history. Of course, projected out indefinitely, the $f = 0.145$ graph will have the greatest gain rate; this is the point of the maximization. Otherwise, the most important feature demonstrated by the figure is that the volatility of the fortunes markedly increase with increasing f. Even the maximizing fraction is accompanied by large swings in fortune.

This is a cautionary aspect of the Kelly Criterion. We take up this topic again in Section 7.5.

Algorithm 30. 60/40 Game simulation

```
Fortune=1; f=0.145; gain = 0.88
for i=1 to N
  wager = f*Fortune
  U~ U(0,1)  ▷uniform sample on 0 to 1
  if U < 0.6
    Fortune = Fortune + gain*wager
  else
    Fortune = Fortune - wager
  endif
endfor
  ▷ending Fortune after N plays
```

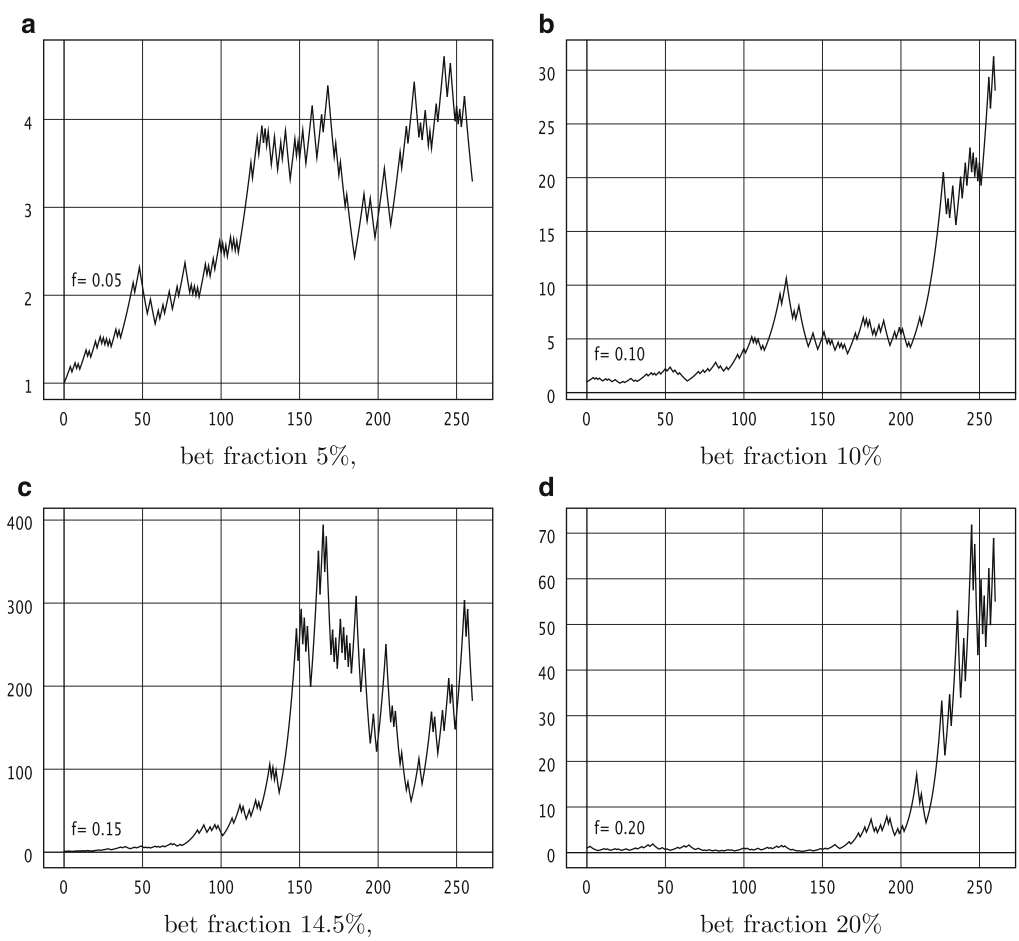


Fig. 7.1. Instances of fortune histories for the 60/40 game for various betting fractions

7.2 The Simple Game with Catastrophic Loss

Consider the example of an option trader who kept careful records of his or her last 75 butterfly trades and notes these statistics:

- 49 gained money, 26 lost money
- Average gain per success, $889.12;
- Average loss of the worst two, $2,637.00;
- Average loss of the remaining 24, $1,366.32.

The payoff expectation of this trade has been

$$E_{\text{gain}} = 889.12 * \frac{49}{75} - 1{,}366.32 * \frac{24}{75} - 2{,}637.00 * \frac{2}{75} = 73.35;$$

a gain of $73.35 per trade.

As above, we ask what fraction f of our investment capital should be risked on this trade going forward. Start by taking the average (intermediate) loss as the unit "bet". Regard the loss suffered by the worst two trades as a "catastrophic" loss. Per unit bet the gains and losses are:

$$\gamma = \frac{889.12}{1{,}366.32} = 0.651, \quad \lambda = 1, \quad \mu = \frac{2{,}637.00}{1{,}366.32} = 1.930.$$

and their probabilities are

$$p_1 = \frac{49}{75} = 0.653, \quad p_2 = \frac{24}{75} = 0.32, \quad p_3 = \frac{2}{75} = 0.026.$$

On a per unit basis the expected payoff is

$$E_{\text{gain}} = 0.651 * 0.653 - 0.32 - 1.930 * 0.026 = 0.0536.$$

The important difference here from the simple game is that f cannot be as big as 1; because of the catastrophic loss, we could lose nearly twice the bet. In fact to avoid losing more than our fortune we must have

$$f < \frac{1}{1.93} = 0.518.$$

To find the maximizing fraction, compute the expected gain rate G as before. After N trades our fortune will be

$$F_N = (1+\gamma f)^W (1-f)^L (1-\mu f)^C F_0,$$

where W, L, C are the number of wins, intermediate losses, and catastrophic losses in the N trades. The expected gain becomes

$$G = \lim_{N\to\infty} \frac{1}{N} \log \frac{F_N}{F_0} = p_1 \log(1+\gamma f) + p_2 \log(1-f) + p_3 \log(1-\mu f).$$

The maximizing fraction is given by the root of

$$0 = G' = \frac{p_1 \gamma}{1 + f\gamma} - \frac{p_2}{1 - f} - \frac{p_3 \mu}{1 - f\mu}.$$

Combining these fractions results in a quadratic polynomial in the numerator to solve for f. Using the numbers for the butterfly application, the maximizing fraction is

$$f = 0.078.$$

An instance of 260 iterations of this investment is shown in Fig. 7.2. The maximal growth rate and expect fortune after 260 identical investments are

$$G = 0.019, \qquad \mathbb{E}(F_{260}) = 151.17.$$

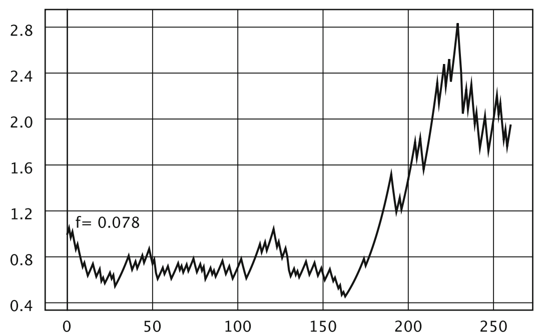


Fig. 7.2. An example fortune history of a butterfly trader. The size of each trade is 7.8 % of the traders investment capital. At the end of the period the trader has doubled his fortune on the basis of these trades in this example. But after 160 trades the fortune was down about half

7.3 The Optimal Allocation Problem When Only One Winner

The foregoing was concerned with optimally allocating our resources to a single venture. Now we want to allow for the possibility of several ventures. At first suppose that only one of them will succeed and our entire investment allocated to the others is lost. This could happen for example by buying calls on each of several firms vying for a lucrative contact. We anticipate that our call on the company winning the contract will pay off handsomely but those on the others will probably expire worthless. As before, this problem can equally well be posed in gambling terms. In fact Kelly's original formulation was that of bets on a horse race at the track.

When framed in terms of an allocation problem, it is more convenient to work with *payouts* rather than gains and losses. By the payout α of a bet we mean the multiplier per unit bet, or per unit investment, upon success. There is a simple relationship between payouts and gains. For a \$1 bet, if the gain is γ, then the payout is

$$\alpha = \gamma + 1. \tag{7.8}$$

because the payout includes the \$1 paid in.

We know that a gamble is favorable if its expectation is positive, which is to say $\mathbb{E} = p\gamma - (1-p) > 0$. In terms of payouts, this is equivalent to the condition

$$p\alpha > 1 \tag{7.9}$$

since $p\alpha - 1 = p(\gamma+1) - 1 = p\gamma - (1-p)$. It follows that the odds are fair when $p\alpha = 1$.

7.3.1 A Four Choice Allocation Problem

To illustrate ideas, we will work through a specific example. Suppose there are four investments under consideration, A, B, C, and D. Through careful research assume that their probabilities of success and associated payoffs are estimated to be as given in Table 7.1. Payouts are higher for the long shots, possibly because other investors have the same idea we have. Denote by f_i the fraction of our bankroll allocated to investment i, $i = A, B, C, D$, and let b denote the fraction held back, that is not invested. Thus

$$b + f_A + f_B + f_C + f_D = 1. \tag{7.10}$$

Table 7.1 Four choice probabilities and payoffs

Choice	Win probability	Payoff multiplier	Fraction
A	0.5	2.1	f_A
B	0.3	3.2	f_B
C	0.1	10.8	f_C
D	0.1	8.5	f_D

Before continuing, observe that the sum of the reciprocal payouts is an important quantity,

$$\tau = \sum_i \frac{1}{\alpha_i}. \tag{7.11}$$

It is easiest to explain this in terms of gambling. Let I_i denote the total amount of money bet on outcome i. Then $I = I_A + I_B + I_C + I_D$ is the total amount the "house" receives from the bets on this play. If A succeeds, the "house" pays back $\alpha_A I_A$. Similarly for B, C, and D. If each of these payouts equals the total paid in, meaning $\alpha_i I_i = I$, then the sum τ equals 1,

$$\frac{1}{\alpha_A} + \frac{1}{\alpha_B} + \frac{1}{\alpha_C} + \frac{1}{\alpha_D} = \frac{I_A}{I} + \frac{I_B}{I} + \frac{I_C}{I} + \frac{I_D}{I} = 1.$$

It means that all money paid in is paid back out. In the investment interpretation it means there are no commissions or other losses, for example, via the bid/ask spread. If $\tau > 1$ it means that less money is paid back than paid in and consequently there are commissions or other losses. On the other hand, if $\tau < 1$, it means money is being created over all by the venture.

Now to the solution. As before, let F_0 denote the initial size of the investor's fortune, or the gambler's bankroll. If A wins, the fortune will then be bF_0 held over plus $\alpha_A f_A F_0$ as the earnings of the investment on A; altogether $(b + \alpha_A f_A)F_0$. Similar expressions hold in case B, C, or D win. The fortune after N such cycles will be

$$F_N = (b + \alpha_A f_A)^{N_A}(b + \alpha_B f_B)^{N_B}(b + \alpha_C f_C)^{N_C}(b + \alpha_D f_D)^{N_D} F_0$$

where N_i is the number of times outcome i, $i = A, B, C, D$, occurred in the N plays. In the limit as $N \to \infty$, the growth rate tends to

$$\begin{aligned} G &= \lim_{N\to\infty} \frac{1}{N} \log \frac{F_N}{F_0} \\ &= p_A \log(b + \alpha_A f_A) + p_B \log(b + \alpha_B f_B) \\ &\qquad + p_C \log(b + \alpha_C f_C) + p_D \log(b + \alpha_D f_D). \end{aligned} \tag{7.12}$$

Finally, the fractions we seek, including the fraction held back, are those that maximize G. But this time the constraint (7.10) applies. By the LaGrange method for constrained optimization we derive the following system with λ as the LaGrange multiplier,[3]

$$\begin{aligned} \frac{\partial G}{\partial f_i} &= \frac{p_i \alpha_i}{b + \alpha_i f_i} = \lambda, \quad i = A, B, C, D \\ \frac{\partial G}{\partial b} &= \sum_i \frac{p_i}{b + \alpha_i f_i} = \lambda \\ \text{and the constraint} \quad & b + f_A + f_B + f_C + f_D = 1. \end{aligned} \tag{7.13}$$

Equations (7.13) comprise the archetypal system for these problems. Problems with fewer or more allocation choices than four entail the obvious modifications. Note that the system (7.13) only applies to solutions interior to the constraint region.

Unfortunately the general solution of this system is not straightforward. Therefore before tackling it, we first gain familiarity by considering some simpler examples.

7.3.2 A Two Choice Problem with $\tau \le 1$

First consider a two choice problem in which $\tau = 1$. For example let $p_A = 0.6$, $p_B = 0.4$, $\alpha_A = 1.88$, and $\alpha_B = 2.136$; $1/1.88 + 1/2.136 = 1$. The gain for betting on A here is $\gamma = \alpha - 1 = 0.88$ and we see that this is the simple game of Section 7.1 in disguise.

Since $\tau = 1$, meaning that the "house" pays back everything it takes in, we reason that there is no need to hold money back; we can effectively hold money back by placing canceling bets. So put $b = 0$. From (7.13), with $b = 0$, we have that $p_A = f_A\lambda$ and $p_B = f_B\lambda$. Together these show that $\lambda = 1$,

[3] See Section A.13.

$$1 = p_A + p_B = f_A\lambda + f_B\lambda = (f_A + f_B)\lambda = \lambda. \tag{7.14}$$

But then we get as a solution $f_A = p_A = 0.6$ and $f_B = p_B = 0.4$.

With this solution, if the old fortune is F_0, the new fortune F_1 is

$$F_1 = \begin{cases} \text{if A} & 1.88 * 0.6F_0 = 1.128F_0, \\ \text{if B} & 2.136 * 0.4F_0 = 0.855F_0, \end{cases}$$

the same as before, (7.6) above.

By placing canceling bets (only one of A or B will win) we effectively bet only a fraction of our bankroll. In the worst case, that of B winning, we get back 85.5 % of our bankroll; this is effectively the fraction not risked. Therefore $0.145 = 1 - 0.855$ is the fraction at risk.

The solution is not unique. In fact every choice of b, f_A, and f_B for which $\lambda = 1$ is a solution; for example $b = 0.1$, $f_A = 0.547$, and $f_B = 0.353$. It is because, with this choice, the new fortune upon a win for A will be $p_A\alpha_A F_0$, and upon a win for B will be $p_B\alpha_B F_0$, the same as before.

If $\tau < 1$ everything is the same. Again choices of b, f_A, and f_B for which $\lambda = 1$ is a solution.

7.3.3 A Two Choice Problem with $\tau > 1$

In this case b cannot be zero because the "house" extracts a payment for each play. Again assume $p_A = 0.6$ and $p_B = 0.4$. Since the Kelly Principle does not apply unless we have a positive expectation, we take α_A big enough for this to happen, for example $\alpha_A = 2$. For $\alpha_B = 1.5$ we will have $\tau = 1/2 + 1/1.5 = 7/6 > 1$.

Because outcome A is advantageous, from what we have seen so far, we must have (from the first equation of (7.13) with $\lambda = 1$)

$$\frac{p_A\alpha_A}{b + \alpha_A f_A} = 1 \tag{7.15}$$

maximizing the gain for this outcome. On the other hand, since outcome B is disadvantageous, we expect that f_B should be zero. This is indeed the case, see the details box on the next page.

7.3.4 Solution to the Four Choice Problem

Now return to the four choice problem of Table 7.1. We first rank the choices in terms of best payoff prospects, $p\alpha$, this works out to be: C with $0.1*10.8 = 1.08$, A with $0.5 * 2.1 = 1.05$, B with $0.3 * 3.2 = 0.96$, and D with $0.1 * 8.5 = 0.85$. Because C and A have positive expectation we know they must satisfy

$$\frac{p_C\alpha_C}{b + \alpha_C f_C} = 1 \qquad \text{and} \qquad \frac{p_A\alpha_A}{b + \alpha_A f_A} = 1. \tag{7.16}$$

The resulting growth rate equation becomes

$$\begin{aligned} G = p_A \log(p_A \alpha_A) + p_C \log(p_C \alpha_C) \\ + p_B \log(b + \alpha_B f_B) + p_D \log(b + \alpha_D f_D); \end{aligned} \tag{7.17}$$

now a function of three variables b, f_B, and f_D. Using (7.16) the constraint reduces to the following

$$\begin{aligned} 1 &= b + f_A + f_C + f + B + f_D \\ &= b + (p_A - \frac{b}{\alpha_A}) + (p_C - \frac{b}{\alpha_C}) + f_B + f_D \\ 1 - (p_A + p_C) &= b(1 - (\frac{1}{\alpha_A} + \frac{1}{\alpha_C})) + f_B + f_D. \end{aligned} \tag{7.18}$$

Given (7.15), the optimization problem we are trying to solve can be reduced. After substituting the partial solution into (7.13), the first term contains only known values. Subtracting that out leaves the simpler problem

$$H = G - p_A \log(p_A \alpha_A) = p_B \log(b + \alpha_B f_B) \tag{7.19}$$

for H as a function of b and f_B and subject to the constraint that $b + f_A + f_B = 1$. But again calling on (7.15), $f_A = p_A - b/\alpha_A$, and so the constraint becomes

$$b(1 - \frac{1}{\alpha_A}) + f_B = 1 - p_A. \tag{7.20}$$

Maximizing H is equivalent to maximizing $b + \alpha_B f_B$ and this occurs at the largest possible values of b and f_B subject to (7.20). But because $1/\alpha_B > 1 - 1/\alpha_A$, this occurs when $f_B = 0$ and, from (7.20),

$$b = \frac{1 - p_A}{1 - \frac{1}{\alpha_A}}. \tag{7.21}$$

For the numbers given above we get

$$b = \frac{1 - 0.6}{1 - \frac{1}{2}} = 0.8,$$

and so $f_A = 0.2$ along with $f_B = 0$. For the degenerate fraction f_B the solution is not from the interior of the solution domain and (7.15) does not necessarily hold for B,

$$\frac{p_B \alpha_B}{b + \alpha_B f_B} = \frac{0.4(\frac{3}{2})}{0.8} = \frac{3}{4} \neq 1.$$

It would seem that since investing in B or D is disadvantageous, we should put $f_B = f_D = 0$. On the other hand, no need to guess, these assignments should be derivable. We first see if adding a fraction of B will increase the growth rate on its own by taking $f_D = 0$. With this assignment, the growth rate equation from (7.17) is

$$\begin{aligned} H &= G - p_A \log(p_A \alpha_A) - p_C \log(p_C \alpha_C) \\ &= p_B \log(b + \alpha_B f_B) + p_D \log(b). \end{aligned} \tag{7.22}$$

Again H is maximized by choosing the largest possible values of b and f_B; but at the same time they must satisfy (7.18) (with $f_D = 0$).

It may be that H is maximized on the boundary of its domain, that would be at $f_B = 0$. But this is not the case. To find the interior maximizing values of b and f_B we must again call on the LaGrange method. With LaGrange multiplier μ, we have from (7.22) and (7.18)

$$\begin{aligned} \frac{\partial H}{\partial f_B} &= \frac{p_B \alpha_B}{b + \alpha_B f_B} = \mu \\ \frac{\partial H}{\partial b} &= \frac{p_B}{b + \alpha_B f_B} + \frac{p_D}{b} = \mu(1 - \tau_2) \end{aligned} \tag{7.23}$$

where

$$\tau_2 = \frac{1}{\alpha_C} + \frac{1}{\alpha_A}. \tag{7.24}$$

By eliminating μ between equations (7.23) and substituting f_B from (7.18) we find the solution to be

$$\begin{aligned} b &= \frac{1 - p_3}{1 - \tau_3} \\ f_B &= p_B - \frac{b}{\alpha_B} \\ \text{and} \qquad \mu &= 1 \end{aligned} \tag{7.25}$$

where

$$\begin{aligned} p_3 &= p_C + p_A + p_B \\ \tau_3 &= \frac{1}{\alpha_C} + \frac{1}{\alpha_A} + \frac{1}{\alpha_B}. \end{aligned}$$

Perhaps surprisingly we find that, once again,

$$\frac{p_B \alpha_B}{b + \alpha_B f_B} = 1$$

maximizing the effect of f_B as $p_B \log(p_B \alpha_B)$. To complete the solution, we should reduce G even further by this choice of f_B and optimize over b and f_D. We leave it to the reader to see that indeed $f_D = 0$ for optimality and so b is given by (7.25).

Using the numbers of Table 7.1 we find that

$$b = \frac{1 - (0.1 + 0.5 + 0.3)}{1 - (\frac{1}{10.8} + \frac{1}{2.1} + \frac{1}{3.2})} = 0.8423.$$

and

$$\begin{aligned} f_A &= 0.5 - \frac{0.8423}{2.1} = 0.099, \quad f_B = 0.3 - \frac{0.8423}{3.22} = 0.037, \\ f_c &= 0.1 - \frac{0.8423}{10.8} = 0.022 \quad f_D = 0. \end{aligned}$$

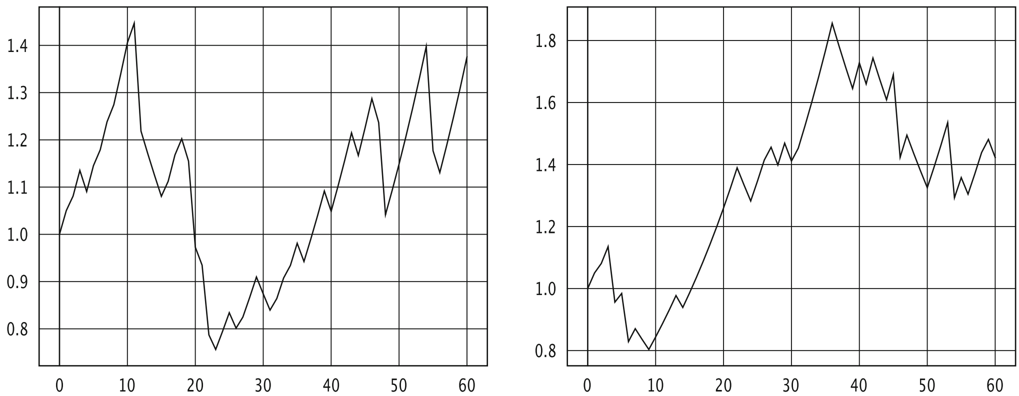


Fig. 7.3. Fortune histories for the four choice investment problem

The maximum growth rate is

$$\begin{aligned} G &= p_A \log(p_A \alpha_A) + p_B \log(p_B \alpha_B) + p_C \log(p_C \alpha_C) + p_D \log(b) \\ &= 0.00268. \end{aligned}$$

Note that by not taking $f_B = 0$ the contribution of that term is $p_B \log(p_B \alpha_B) = -0.0122$ as opposed to $p_B \log(b) = -0.0515$ otherwise.

After 60 and 260 iterations respectively under this allocation the expected fortunes are

$$\mathbb{E}(F_{60}) = 1.17, \qquad \mathbb{E}(F_{260}) = 2.00.$$

In Fig. 7.3 we show two instances of 60 iterations of this investment problem.

7.4 The Optimal Allocation Problem with Multiple Winners

In the previous section our capital was allocated among several investments for which there could be only one winner. But more often the investments tend to be independent; some fraction of them, or even all of them, can be successful. Or conversely, they may all lose money. We take up this sort of optimal allocation problem next. Again we demonstrate the method with the help of an example.

Table 7.2 Probabilities and payoffs for two investments

Choice	Success probability	Payoff multiplier	Fraction
A	$p_A = 0.6$	$\alpha_A = 2.1$	f_A
B	$p_B = 0.5$	$\alpha_B = 2.2$	f_B

Suppose we have two investment opportunities A and B with particulars as in Table 7.2. Both investments have positive expectation but since they could both lose, we must not risk our entire capital; again let b be the fraction held back.

The growth rate equation this time must contains terms for the possibility that both investments succeed, with probability $p_A p_B$, that one succeeds while the other fails, with probabilities $p_A(1-p_B)$ and $(1-p_A)p_B$, or that both fail, $(1-p_A)(1-p_B)$. In the usual way we derive the growth rate to be

$$\begin{aligned} G = p_A p_B \log(b + \alpha_A f_A + \alpha_B f_B) + p_A(1-p_B)\log(b+\alpha_A f_A) \\ +(1-p_A)p_B \log(b+\alpha_B f_B) + (1-p_A)(1-p_B)\log(b). \end{aligned} \tag{7.26}$$

The unknowns b, f_A, and f_B are constrained in that $b + f_A + f_B = 1$.

Using the techniques developed in the previous sections, the solution reduces to solving for w in the following equation

$$\frac{p_A p_B}{w + \tau_2 - 1} + \frac{p_A(1-p_B)}{w + \frac{1}{\alpha_B} - 1} + \frac{(1-p_A)p_B}{w + \frac{1}{\alpha_A} - 1} + \frac{(1-p_A)(1-p_B)}{w} = 0.$$

The details are given in the box on the next page. While it would seem there is no symbolic solution, this equation is easily solved numerically. Using the parameters in Table 7.2 there are three solutions of which only one, $w = 0.29113$, gives admissible values of b, f_A, and f_B. Then from (7.29)

$$u = 1 - \frac{1}{2.2} - 0.29113 = 0.2543 \qquad v = 1 - \frac{1}{2.1} - 0.29113 = 0.2327$$

and from (7.28)

$$b = 0.687, \qquad f_A = 0.2345, \qquad f_B = 0.0784.$$

The sum $b + f_A + f_B = 1$ confirming that λ is indeed 1. The maximal growth rate and expected fortune for 60 iterations are

$$G = 0.0347 \qquad \mathbb{E}(F_{60}) = 8.02.$$

Two 60 iteration instances of this investment are shown in Fig. 7.4.

7.4.1 Allocation for Correlated Investments

In the previous section the investments were taken to be independent. But this is probably not realistic. We have already called attention to the fact that, to an extent, the market moves in concert both up and down, see Section 2.3.3. If our investments are correlated, how does that affect the Kelly Criterion? To see how, we re-examine investments A and B of the previous section in this new light.

Start by assuming that A and B are correlated with a given correlation coefficient ρ, see page 49. Recall that ρ quantifies the degree to which A and B move together. But we need to specify more than just the correlation coefficient, we also need the joint probabilities for the two investments.

By the method of LaGrange, with multiplier $\lambda = 1$, the system of equations for the unknowns is

$$\begin{aligned}
\frac{\partial G}{\partial f_A} &= \frac{p_A p_B \alpha_A}{b + \alpha_A f_A + \alpha_B f_B} + \frac{p_A(1-p_B)\alpha_A}{b + \alpha_A f_A} = 1 \\
\frac{\partial G}{\partial f_B} &= \frac{p_A p_B \alpha_B}{b + \alpha_A f_A + \alpha_B f_B} + \frac{(1-p_A)p_B\alpha_B}{b + \alpha_B f_B} = 1 \\
\frac{\partial G}{\partial b} &= \frac{p_A p_B}{b + \alpha_A f_A + \alpha_B f_B} \\
&\quad + \frac{p_A(1-p_B)}{b + \alpha_A f_A} + \frac{(1-p_A)p_B}{b + \alpha_B f_B} + \frac{(1-p_A)(1-p_B)}{b} = 1
\end{aligned} \tag{7.27}$$

Temporarily make the following assignments in the fractions above

$$\begin{aligned}
t &= \tfrac{p_A p_B}{b+\alpha_A f_A+\alpha_B f_B} & u &= \tfrac{p_A(1-p_B)}{b+\alpha_A f_A} \\
v &= \tfrac{(1-p_A)p_B}{b+\alpha_B f_B} & w &= \tfrac{(1-p_A)(1-p_B)}{b}.
\end{aligned} \tag{7.28}$$

Solving (7.27) as an under-determined linear system we get the following solutions in terms of w

$$\begin{aligned}
t &= w + \tau_2 - 1 \\
u &= 1 - \frac{1}{\alpha_B} - w \\
v &= 1 - \frac{1}{\alpha_A} - w
\end{aligned} \tag{7.29}$$

where $\tau_2 = \frac{1}{\alpha_A} + \frac{1}{\alpha_B}$. Now reciprocate each equation (and add b to the first)

$$\begin{aligned}
2b + \alpha_A f_A + \alpha_B f_B &= \frac{p_A p_B}{w + \tau_2 - 1} + b \\
b + \alpha_A f_A &= \frac{p_A(1-p_B)}{1 - \frac{1}{\alpha_B} - w} \\
b + \alpha_B f_B &= \frac{(1-p_A)p_B}{1 - \frac{1}{\alpha_A} - w}
\end{aligned} \tag{7.30}$$

Subtract the second and third equations from the first, and noting that $b = (1 - p_A)(1-p_B)/w$, we get

$$0 = \frac{p_A p_B}{w + \tau_2 - 1} + \frac{p_A(1-p_B)}{w + \frac{1}{\alpha_B} - 1} + \frac{(1-p_A)p_B}{w + \frac{1}{\alpha_A} - 1} + \frac{(1-p_A)(1-p_B)}{w} \tag{7.31}$$

to be solved for w.

Let X be the random variable which is 1 if A succeeds and 0 if A fails. This is a Bernoulli random variable; it is easy to see that

$$\mathbb{E}(X) = p_A, \quad \text{and} \quad \operatorname{var}(X) = p_A(1-p_A). \tag{7.32}$$

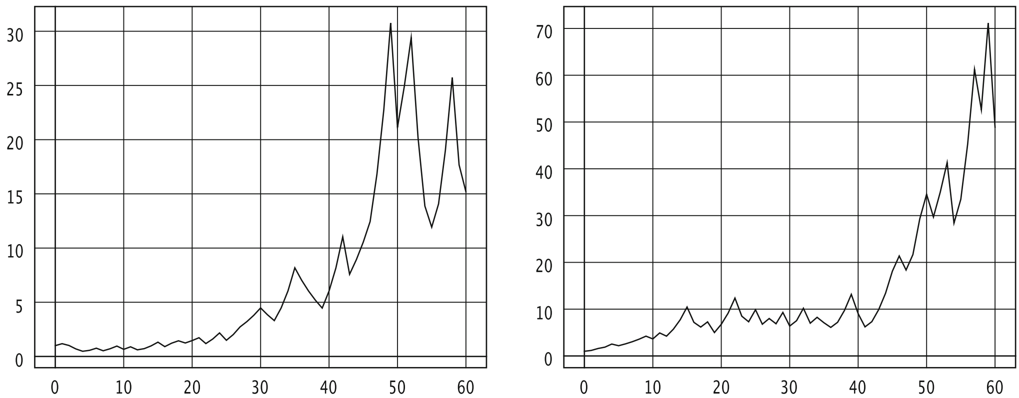


Fig. 7.4. Two instances of fortune histories for the two investment allocation using the optimal investment fractions, $b = 0.687$, $f_A = 0.2345$, $f_B = 0.0784$

Similarly let Y be the same thing for B, $Y = 1$ with probability p_B and $Y = 0$ with probability $1 - p_B$.

The joint density, $d(\cdot, \cdot)$, for X and Y is defined by: $d(1,1)$ is the probability both succeed, $d(0,0)$ is the probability they both fail, $d(1,0)$ is the probability A succeeds and B fails and $d(0,1)$ is the other way around. Since the joint random variable XY is 1 only if both A and B succeed and 0 otherwise, we have

$$\mathrm{covar}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = d(1,1) - p_A p_B. \tag{7.33}$$

But from (2.29)

$$\rho = \frac{\mathrm{covar}(X,Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}} = \frac{d(1,1) - p_A p_B}{\sigma_A \sigma_B},$$

where σ_A and σ_B are the standard deviations of X and Y respectively. Therefore

$$d(1,1) = p_A p_B + \rho \sigma_A \sigma_B. \tag{7.34}$$

It says the joint probability that both A and B succeed at the same time is increased by the product of the correlation coefficient with the product of the two standard deviations.

To get the other joint probabilities we proceed in a similar manner. For example, for $d(1,0)$, let Y^C be the complementary random variable to Y: Y^C is 1 when B fails and 0 when B succeeds. Then the correlation between X and Y^C is the negative of that between X and Y. Since XY^C is non-zero only when A succeeds and B fails we have

$$-\rho = \frac{\mathbb{E}(XY^C) - \mathbb{E}(X)\mathbb{E}(Y^C)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y^C)}} = \frac{d(1,0) - p_A(1-p_B)}{\sigma_A \sigma_B}$$

and so

$$d(1,0) = p_A(1-p_B) - \rho \sigma_A \sigma_B. \tag{7.35}$$

The other joint probabilities are

$$d(1,0) = p_A(1-p_B) - \rho\sigma_A\sigma_B, \quad d(0,0) = (1-p_A)(1-p_B) + \rho\sigma_A\sigma_B. \quad (7.36)$$

The only modification to the growth rate equation (7.26) is that now the coefficients are the modified probabilities,

$$\begin{aligned} G = d(1,1)\log(b+\alpha_A f_A+\alpha_B f_B) + d(1,0)\log(b+\alpha_A f_A) \\ +d(0,1)\log(b+\alpha_B f_B) + d(0,0)\log(b). \end{aligned} \quad (7.37)$$

With the same algebraic manipulation as before we again arrive at (7.31) only with the numerators replaced by $d(1,1), \ldots, d(0,0)$,

$$0 = \frac{d(1,1)}{w+\tau_2-1} + \frac{d(1,0)}{w+\frac{1}{\alpha_B}-1} + \frac{d(0,1)}{w+\frac{1}{\alpha_A}-1} + \frac{d(0,0)}{w}. \quad (7.38)$$

Consider the problem of the previous section and assume the correlation coefficient between A and B is 0.3. Equation (7.38) becomes

$$0 = \frac{0.3727}{w-0.0692} + \frac{0.2272}{w-0.5454} + \frac{0.1272}{w-0.5238} + \frac{0.2727}{w}.$$

The relevant root is $w = 0.3621$ giving the solution

$$b = 0.755, \quad f_A = 0.229, \quad f_B = 0.012, \quad \text{for } \rho = 0.3.$$

With this correlation, the fraction held back has increased mostly at the expense of the fraction to B. The growth rate for this allocation is

$$G = 0.0309.$$

High Correlation Between Investments

What if the correlation is much bigger? With steady increase in ρ the fraction allocated to B will decrease to 0. This occurs at $\rho = 0.3488\ldots$. Then the solution will no longer lie in the interior of the constraint region and at that point the LaGrange equations will not be valid. Instead the solution will be found on the boundary, either $b = 0$, or $f_A = 0$ or $f_B = 0$.

Consider the problem of the previous section and assume the correlation coefficient between A and B is 0.8. Among the two expected payouts, that for A is bigger since $p_A\alpha_A = 1.26 > 1.1 = p_B\alpha_B$; a reason to try the $f_B = 0$ boundary. But then (7.37) reduces to

$$\begin{aligned} G &= (d(1,1)+d(1,0))\log(b+\alpha_A f_A) + (d(0,1)+d(0,0))\log(b) \\ &= p_A\log(b+\alpha_A f_A) + (1-p_A)\log(b), \end{aligned} \quad (7.39)$$

since

$$d(1,1)+d(1,0) = (p_A p_B + \rho\sigma_A\sigma_B) + (p_A(1-p_B) - \rho\sigma_A\sigma_B) = p_A$$

similarly $d(0,1) + d(0,0) = 1 - p_A$. The problem has now reduced to the simple game of Section 7.1; its solution is given by

$$f_A = \frac{p_A \gamma_A - (1 - p_A)}{\gamma_A} = \frac{0.6 * 1.1 - 0.4}{1.1} = 0.236.$$

Of course $b = 1 - 0.236 = 0.764$. The corresponding growth rate is

$$G = 0.0308.$$

Anti-correlation Between Investments

What if A and B are negatively correlated? Assume $\rho = -0.3$. The solution to (7.38) for this correlation is $w = 0.2223$. The resulting fractions are

$$b = 0.569, \quad f_A = 0.279, \quad f_B = 0.153.$$

The growth rate for these fractions is

$$G = 0.0449.$$

We see once again the advantage of negatively correlated investments.

7.5 Taming the Kelly Volatility

A review of the several growth histories presented in this chapter for allocations fixed by the Kelly Criterion reveal a characteristic feature – high volatility. In fact these fortune histories are a kind of geometric random walk with binomially instead of normally distributed increments. One of their properties is that for any fraction m of one's initial fortune F_0, given enough time the walk will be less than mF_0. If q is the probability of an adverse outcome, one that will reduce the fortune, then for any $n > 0$, the probability of n consecutive adverse outcomes is $q^n > 0$. There are other possibilities, for example the probability of n adverse outcomes with one intervening beneficial outcome is $\binom{n-1}{1} q^n (1-q)$. As all these probabilities are positive, given enough time a catastrophic drawdown will happen.

Furthermore, while in the limit the Kelly Criterion will produce maximum growth, this is of little value to the investor since investors are bound by finite time horizons. A more important question is, "What will the fortune be after a specific number of plays?"

There are attempts to ameliorate these shortcomings of the Kelly Criterion by fixing allocations at smaller fractions than the maxmizers, for example "half-Kelly." Figure 7.1 shows some fortune histories for sub-maximal fractions.

Even better are schemes that dynamically adjust the allocation fractions depending on the current fortune growth rate. Such schemes are hard to analyze theoretically but can be easily evaluated by Monte Carlo methods.

Consider one possible scheme: reduce the betting fraction if the growth rate up to the present time is bigger than expected and increase the fraction if the growth rate is too small. For example in the 60/40 game with gain $\gamma = 1$ the betting fraction is, equation (7.5),

$$f = p - q = 0.2.$$

and the maximal growth rate is, from (7.7),

$$G_{\max} = 0.6 \log(0.6 * 2) + 0.4 \log(0.4 * 2) = 0.020. \tag{7.40}$$

The more the present growth rate exceeds $G_{\max}$, the more f is reduced. On the other hand, if the present growth rate is much smaller than $G_{\max}$, then f is allowed to approach 0.2.

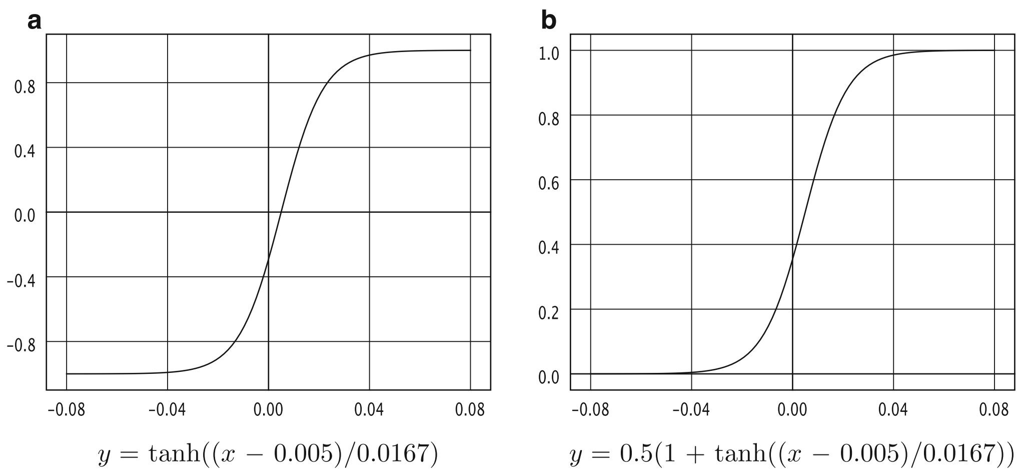


Fig. 7.5. Graphs of the hyperbolic tangent function horizontally scaled and shifted; (**b**) is also vertically scaled and shifted to vary from 0 to 1

To implement such a plan we use the *hyperbolic tangent* function

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{7.41}$$

This function lies between -1 and 1, tends to -1 as $x \to -\infty$, and increases asymptotically to 1 as $x \to \infty$. It can be horizontally scaled and shifted as desired using parameters s and h respectively,

$$y = \tanh((x - h)/s).$$

The parameter h shifts the graph to the right by h units and s scales the argument so that what used to happen at $x = 1$ now happens at $x = s$. Figure 7.5a is the graph of

$$y = \tanh((x - 0.005)/0.0167).$$

The range of this function is -1 to 1. But we require the range to be 0 to 1 so we use the function in panel (b) which is vertically scaled by 1/2 and shifted up by 1.

Now, given minimum and maximum fractions, say `minbet` and `maxbet` respectively, the dynamically determined fraction is given by

$$f = \texttt{minbet} + y(\texttt{maxbet - minbet}).$$

Finally, in the application we take x to be the difference between the target growth rate and the actual growth rate, denoted G_t and G_a respectively. The complete system is

$$\begin{aligned} y &= 0.5(1 + \tanh((G_t - G_a - h)s)) \\ f &= \texttt{minbet} + y(\texttt{maxbet - minbet}). \end{aligned} \tag{7.42}$$

With the numbers above, if the actual growth rate equals the target, then the ratio y is 0.35, so f is about a third of the way between the minimum and maximum fractions, this can be seen is Fig. 7.5b. Also seen in this figure, when the actual growth rate lags behind the target by 0.02, then the ratio y is approximately 0.8, and when the actual exceeds the target by 0.2, the ratio is approximately 0.05, very close to the minimum bet.

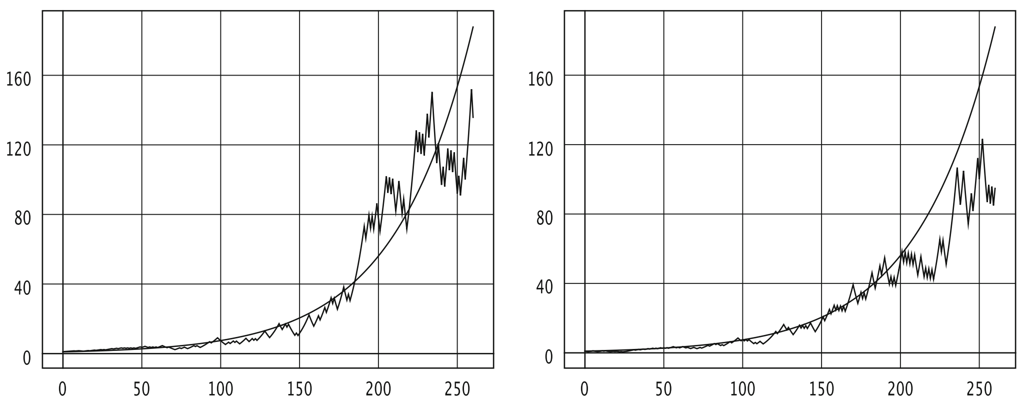


Fig. 7.6. Two runs of the dynamically adjusted 60/40 game

The implementation of this scheme for dynamically adjusting Kelly bets is given below.

Algorithm 31. Dynamic Kelly growth

inputs: p, γ, h, s, G_t, minbet, maxbet
        nTrials, nBets ▷number of times to play the game
avgGR = 0 ▷average growth rate over nTrials
for $i = 1, \dots,$ nTrials
  $F = 1$ ▷each trial starts with unit fortune
  for n=$1, \dots,$ nBets
    $G_a = (\log F)/n$ ▷actual current growth rate
    $y = 0.5(1 + \texttt{tanh}((G_t - G_a - h)/s))$
    f = minbet+y(maxbet- minbet)
    if $U \sim U(0,1) \ < \ p$ ▷win or lose this bet?
      $F = (1 + \gamma f)F$

```
        else
          F = (1 − f)F
        endif
      endfor
      avgGR = avgGR + F  ▷F is ending fortune
    endfor
    avgGR = avgGR/nTrials;
```

Example 7.1. In an example run of the algorithm for the 60-40 game, the target growth rate was taken to be $G_t = 0.02$, the Kelly maximum, the minimum bet to be 0.05, equal to one fourth the Kelly fraction, and the maximum bet to be the Kelly fraction 0.20. The results are that over 10,000 trials the average growth rate was 0.0162, compare with (7.40). Two of the runs are shown in Fig. 7.6. The smooth curve is fortune growth at the target rate.

Problems: Chapter 7

1. Implement the 60-40 game over 300 iterations and show the fortune of the game for several runs. Assume the return upon success is 88 %. Experiment with several betting fractions above and below the growth rate maximizer. Plot the average return rate as a function of betting fraction. Be sure to make enough runs so that these averages are good to 2 places.

2. In Table 5.1 on page 151 the second line gives the gain expectation for buying ATM calls as 0.161 with the probability of a gain as 0.37. Assuming complete loss of investment is the complementary probability, what is the average return on this investment? What is the size of the Kelly bet?
(Answer 2.14, 7.5 %.)

3. In Table 5.2 on page 153 the fifth line gives the gain expectation for \$5 OTM 30 day covered calls as 0.242 with probability of a gain as 54 %. Since the gain could be anywhere between 0 and 5.36, assume it is the mid-point, 2.68. Further assume that the trade loses \$1 36 % of the time. How much does it lose the remaining 10 % of the time? Under these conditions, what is the Kelly fraction of ones portfolio to be allocated to this investment?
(Answer 8.45, 1.9 %)

4. What fractions of a portfolio should be allocated to the following independent investments: A has 70 % chance of succeeding and payoff multiplier $\alpha = 1.5$; B has 40 % chance of succeeing and payoff multiplier $\alpha = 2.6$.

5. Same question as Problem 4 if A and B are correlated with coefficient $\rho = 0.2$.

6. Throughout this chapter (and in Problem 1 above) it has been assumed that the return upon success was fixed, e.g. 88 %. But what if the return is only an average? Rework Problem 1 as follows: upon each success, let the return be either 100 % or 76 % equally likely. How does this change your average return rate graph?

7. Implement Algorithm 30 of page 194 on the investment of Problem 2. Show the fortune after 30 iterations of the strategy, after 60. Estimate the mean and variance of these ending fortunes. Use at least 10,000 trials.

8. Repeat Problem 7 but now use a dynamic allocation strategy such as Algorithm 31.

A

Some Mathematical Background Topics

A.1 Series Identities

A.1.1 Geometric Series

Direct multiplication confirms the algebraic identity $(x-y)(x^n + x^{n-1}y + \ldots + xy^{n-1} + y^n) = x^{n+1} - y^{n+1}$ for any x and y and n a positive integer. If $x \neq y$, then

$$x^n + x^{n-1}y + \ldots + xy^{n-1} + y^n = \frac{x^{n+1} - y^{n+1}}{x-y}.$$

In particular

$$1 + r + r^2 + \ldots + r^{n-1} = \frac{1-r^n}{1-r}. \tag{A.1}$$

If $|r| < 1$ the right hand side converges as $n \to \infty$ and this becomes

$$1 + r + r^2 + \ldots = \frac{1}{1-r}. \tag{A.2}$$

A.1.2 Arithmetic Series

By adding $1 + 2 + 3 + \ldots + n$ to itself backwards one gets n terms each equal to $n+1$. Hence

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}.$$

A.1.3 Taylor's Series

An infinitely differentiable function $y = f(x)$ can be expanded in a power series about a given point $x = a$ according to

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \ldots.$$

In particular, the exponential function can be expanded about 0 to give

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots. \tag{A.3}$$

Likewise the log function can be expanded about 1. Using the change of variable $u = t - 1$, from the definition of the log function

$$\begin{aligned}\log(1+x) &= \int_1^{1+x} \frac{dt}{t} = \int_0^x \frac{du}{1+u} \\ &= \int_0^x (1 - u + u^2 - u^3 + \dots)du = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \end{aligned} \tag{A.4}$$

valid for $|x| < 1$.

A.2 Histograms

A histogram is a special kind of bar chart for depicting a set of values, v_1, v_2, …, v_N, numbering, say, N in total. A convenient subdivision of the x-axis is created containing the values, for example by means of the points x_0, x_1, x_2, …, x_K; $x_0 \le v_i \le x_K$, $i = 1, 2, \dots, N$. They establish intervals, or *bins*, $[x_0, x_1)$, $[x_1, x_2)$, …, $[x_{K-1}, x_K)$. A count is made of how many of the values lie in each bin, for example n_1 in $[x_0, x_1)$, n_2 in $[x_1, x_2)$ and so on. Finally a rectangle of that height n_k is drawn standing on bin $[x_{k-1}, x_k)$, $k = 1, \dots, K$. This is called a *frequency histogram.*

Altogether the area under to bars is $A = \sum_{k=1}^K n_k(x_k - x_{k-1})$. Redrawing the figure and making the height on the kth subinterval equal to n_k/A, produces a *density histogram* or just *histogram* for short.

A density histogram is an approximation of the probability density of the process generating the original values.

A.3 Probability Distributions and Densities

A *random variable* X is the specific (real-valued) outcome of a trial of a process whose outcomes are unpredictable (in exact value). By means of a histogram it is possible to see with what frequency the various outcomes occur.

For a *discrete* random variable, one with only finitely many outcomes, the frequency p_i of each outcome, x_i, is its *probability*, $\Pr(X = x_i) = p_i$, and the function of these probabilities, $f(x_i) = p_i$, is its *probability density function* or pdf.

For every real number x, the sum of the probabilities less than or equal to x is called the *cumulative distribution function* or cdf,

$$F(x) = \sum\{f(x_i) : x_i \le x\}.$$

The cumulative distribution function is 0 for x less than the smallest outcome of the process, is 1 for x larger than the largest outcome, and is otherwise constant except for jumps at each outcome x_i by the amount p_i. It follows that for each x, $F(x)$ is the probability that a trial of the process will be less than or equal to x.

Likewise for a *continuous* random variable, its cumulative distribution function $F(x)$ is the probability that a trial of the process will be less than or equal to x. However for a continuous random variable, the cdf is continuous, that is, has no jumps. But nevertheless it is monotone increasing (if $y > x$, then $F(y) \geq F(x)$) and tends to 0 as $x \to -\infty$ and 1 as $x \to \infty$.

The probability density function $f(\cdot)$ for a continuous random variable is the derivative of its cdf. Therefore the probability a trial of the process will lie between two real values is given by the integral of its pdf,

$$\Pr(a < X < b) = \int_a^b f(x)\,dx.$$

A.4 Expectations

The *expectation* of a function $g(X)$ of a random variable is, in the discrete case,

$$\mathbb{E}(g(X)) = \sum_i g(X_i) f(x_i)$$

and in the continuous case

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\,dx.$$

The *mean* is the expectation of X itself

$$\mu = \mathbb{E}(X)$$

and the *variance* is the expectation of the squared differences from the mean

$$\mathrm{var}(X) = \mathbb{E}\left[(X - \mu)^2\right] = \mathbb{E}(X^2) - \mu^2.$$

The third member is an equivalent expression for the second.

If a distribution is tightly clustered about its mean, its variance is small.

By the Law of Large Numbers, expectations can be approximated empirically. Let X_1, X_2, …, X_n be the outcomes of n trials of the process. The estimate of the expectation $\mathbb{E}(g(X))$ is

$$\mathbb{E}(g(X)) \approx \frac{1}{n} \sum_{i=1}^{n} g(X_i).$$

This tends to the exact value as $n \to \infty$.

Let X and Y be two random variables defined over the same probability space, their *covariance* is defined as

$$\mathrm{covar}(X,Y) = \mathbb{E}\Big((X-\mu_X)(Y-\mu_Y)\Big). \tag{A.5}$$

The *correlation* between X and Y is defined as

$$\rho_{XY} = \frac{\mathrm{covar}(X,Y)}{\sigma_X \sigma_Y}. \tag{A.6}$$

If X and Y are independent then $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$ for functions of X and Y respectively and so $\mathrm{covar}(X,Y) = \rho_{XY} = 0$. Further, if X and Y are independent, then $\mathrm{var}(f(X)+g(Y)) = \mathrm{var}(f(X)) + \mathrm{var}(g(Y))$.

A.5 The Normal Distribution

Among probabilistic processes, one of the most important is the *normal distribution*. It is a continuous process with density given by

$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty.$$

The density has two parameters, μ and σ^2, and these are its mean and variance respectively. A notation for this distribution is $N(\mu, \sigma^2)$.

There is no closed form expression for the cdf of the normal distribution in terms of familiar functions, but there are several accurate rational approximations. That due to Abramowitz and Stegun is as follows.

Let $\Phi(\cdot)$ denote the cumulative distribution function and let

$$t = 1/(1 + a|x|),$$

$|x|$ indicates the absolute value of x. Then

$$\Phi(x) \approx 1 - \phi(x)(b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5). \tag{A.7}$$

The constants are:

$$\begin{aligned} &a = 0.2316419 \\ &b_1 = 0.319381530, \quad b_2 = 0.356563782, \quad b_3 = 1.781477937 \\ &b_4 = 1.821255978, \quad b_5 = 1.330274429 \end{aligned} \tag{A.8}$$

There is also an rational approximation for the inverse of the cumulative distribution function. The following is from [BFS83]

Let $x = \Phi^{-1}(u)$ for $0.5 \le u < 1$ and put

$$y = \sqrt{-\log((1-u)^2)}.$$

Then

$$x = y + \frac{p_0 + p_1 y + p_2 y^2 + p_3 y^3 + p_4 y^4}{q_0 + q_1 y + q_2 y^2 + q_3 y^3 + q_4 y^4}. \tag{A.9}$$

If $0 < u < 0.5$, by symmetry, $\Phi^{-1}(u) = -\Phi^{-1}(1-u)$. The constants are:

p_0=−0.322232431088	q_0=0.099348462606
p_1=−1	q_1= 0.588581570495
p_2=−0.342242088547	q_2=0.531103462366
p_3=−0.0204231210245	q_3=0.10353775285
p_4=−0.0000453642210148	q_4=0.0038560700634

A.6 The Central Limit Theorem

Theorem A.1. *(Central limit theorem) Let X_1, X_2, ..., X_n be independent random samples from a distribution with mean μ and finite variance σ^2. Then*

$$Y = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}}$$

has a limiting distribution as $n \to \infty$ and it is $N(0,1)$, normal with mean 0 and variance 1.

A.7 Least Squares

Assume variables y and x are linearly related with slope m and intercept b. And assume we have n empirical data points testing that relationship, (x_1, y_1), (x_2, y_2), ..., (x_n, y_n). Let e_i be the difference between the empirical value y_i and the predicted value, $mx_i + b$. The sum of the squares of these differences is

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - (mx_i + b))^2 . \tag{A.10}$$

We seek to find the values of m and b which minimize this sum.

Start by differentiating (A.10) with respect to m and b and set the derivatives to zero. First with respect to m

$$\begin{aligned} 0 &= 2\sum_{i=1}^n (y_i - (mx_i + b))(-x_i) \\ 0 &= \sum_{i=1}^n x_i y_i - m\sum_{i=1}^n x_i^2 - b\sum_{i=1}^n x_i, \end{aligned} \tag{A.11}$$

a -2 has been divided out since what's left will still be zero.

Then with respect to b

$$
\begin{aligned}
0 &= 2\sum_{i=1}^{n}(y_i - (mx_i + b))\,(1) \\
0 &= \sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i - nb.
\end{aligned}
\tag{A.12}
$$

The resulting system of two linear equations in two unknowns is, from (A.11) and (A.12),

$$
\begin{array}{rcl}
m\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i & = & \sum_{i=1}^{n} x_i y_i \\
m\sum_{i=1}^{n} x_i + \quad nb & = & \sum_{i=1}^{n} y_i.
\end{array}
$$

The solution by Cramer's Rule, see below, is

$$
\begin{aligned}
m &= \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \\
b &= \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}
\end{aligned}
\tag{A.13}
$$

A.8 Error Estimates for Monte Carlo Simulations

Suppose we are trying to use Monte Carlo to estimate some value, call it θ.

Let $X_1, X_2, \ldots, X_n$ be n estimates of θ as derived from the outcome of the simulation. If the X_i are independent and identically distributed with mean θ, then by the central limit theorem their sample average $\bar{X}$ is approximately normally distributed with mean equal to θ and variance equal to σ_X^2/n, where σ_X^2 is the (unknown) variance of the X_i. In this case

$$
Y = \frac{\bar{X} - \theta}{\sqrt{\sigma_X^2/n}}
$$

is approximately $N(0,1)$ distributed. From a $N(0,1)$ table, e.g. Equation (A.7), notice that with probability 0.954 a normal sample lies within two standard deviations of the mean (above and below)

$$
2(\Phi(2) - 0.5) = 0.954.
$$

Hence

$$
\Pr\left(-2 < \frac{\bar{X} - \theta}{\sqrt{\sigma_X^2/n}} < 2\right) = 0.954.
$$

In other words, with probability 0.954, θ lies in the interval

$$\bar{X} - 2\sqrt{\sigma_X^2/n} < \theta < \bar{X} + 2\sqrt{\sigma_X^2/n}.$$

Now, given a value for σ_X^2, we may calculate probabilistic error bounds for θ, or *confidence intervals* as they are called.

In practice there are three problems with this program. First, usually σ_X^2 must itself be estimated from the data. Second, the X_i may not be identically distributed; the simulation may suffer start-up effects, for example. And third, the X_i may be correlated.

The second and third issues may be dealt with by *batching*. Divide the n trials into m batches each of size J:

$$X_1 \quad \ldots \quad X_J \ \mid \ X_{J+1} \quad \ldots \quad X_{2J} \ \mid \ \ldots \ \mid \ X_{(m-1)J+1} \quad \ldots \quad X_{mJ}.$$

Thus there are $m = n/J$ batches. By the CLT the batch random variables

$$B_i = \frac{1}{J} \sum_{j=(i-1)J+1}^{iJ} X_j, \qquad i = 1, \ldots, m,$$

tend to be independent and identically distributed. Now we may apply the development above to the B_i in place of the X_i. Thus

$$\hat{\theta} = \bar{B} = \frac{1}{m} \sum_{i=1}^{m} B_i$$

is an estimator $\hat{\theta}$ for θ. And the random variable

$$Y = \frac{\hat{\theta} - \theta}{\sqrt{\sigma_B^2/m}}$$

is approximately $N(0, 1)$.

If we knew the variance σ_B^2 of the batch random variables, then we could use the normal distribution itself to make error bounds as was done above. However, σ_B^2 is generally not known and must itself be estimated from the B_i data. The *sample variance* for σ_B^2 is given by

$$s_B^2 = \frac{1}{m-1} \sum_{i=1}^{m} (B_i - \hat{\theta})^2 \tag{A.14}$$

and is itself a random variable (gamma distributed).

Thus in place of Y we have

$$t = \frac{\hat{\theta} - \theta}{(s_B/\sqrt{m})}, \tag{A.15}$$

the quotient of a normal random variable by a gamma random variable. Such a combination is a *Student-t* random variable. The Student-t is well known and has one parameter, its *degrees-of-freedom* or dof, see Section 6.8.1. Tables giving ordinates of the Student-t are included in the appendix of most statistics books. We provide an abbreviated one as well, Table A.2. So we must use the *t-statistic* to obtain the confidence intervals we seek.

For example, given α, t_α is defined by

$$\Pr(-t_\alpha < t < t_\alpha) = \alpha.$$

These values t_α can be looked up in or derived from a t-table. If the table gives cumulative values instead of two-sided values, as does the table below, a conversion is required. Since the t distribution is symmetric, $\Pr(-t_\alpha{<}t{<}t_\alpha){=}\alpha$ is equivalent to

$$\Pr(t < t_\alpha) = \alpha + \frac{1-\alpha}{2} = \frac{1+\alpha}{2}.$$

Thus if one wants, say, a 95 % confidence interval for t, use the cumulative table with $\beta = (1+0.95)/2 = 0.975$.

The t distribution

The table gives t_β vs dof where t_β is defined by $\Pr(t < t_\beta) = \beta$.

dof	1	2	3	4	5	6	8	10	15	20	30	40	60	120	∞
t_β	6.31	2.92	2.35	2.13	2.02	1.94	1.86	1.81	1.75	1.73	1.70	1.68	1.67	1.66	1.65

Table A.1. Cumulative t values for $\beta = 0.95$

dof	1	2	3	4	5	6	8	10	15	20	30	40	60	120	∞
t_β	12.71	4.30	3.18	2.78	2.57	2.45	2.31	2.23	2.13	2.09	2.04	2.02	2.00	1.98	1.96

Table A.2. Cumulative t values for $\beta = 0.975$

A.9 Drawing Normal Samples

The *Box–Muller Algorithm* generates exact $N(0,1)$ samples.

Algorithm 32. Box–Muller Algorithm

$U_1, U_2 \sim U(0,1)$ ▷`draw two independent uniform samples`
$Z_1 = \cos(2\pi U_1)\sqrt{-2\ln U_2}$
$Z_2 = \sin(2\pi U_1)\sqrt{-2\ln U_2}$
▷Z_1, Z_2 `are two independent` $N(0,1)$ `samples`

The *Marsaglia-Bray algorithm* is an alternative that also produces exact samples.

Algorithm 33. Marsaglia-Bray Algorithm

```
repeat
  U ∼ U(0,1); U = 2U-1;  ▷uniform on -1 to 1
  V ∼ U(0,1); V = 2V-1;  ▷a point in the sqr.
until W=U²+V²<= 1
```
$Z_1 = U\sqrt{(-2/W)\ln W}$
$Z_2 = V\sqrt{(-2/W)\ln W}$
  ▷Z_1, Z_2 are two independent $N(0,1)$ samples

A.10 Drawing Inverse Gaussian Samples

The inverse Gaussian density is, from (6.10),

$$f_{IG(x;a,b)} = \frac{ae^{ab}}{\sqrt{2PIx^3}} e^{-\frac{1}{2}(\frac{a^2}{x}+b^2x)}, \quad x > 0.$$

Samples may be drawn as follows:

Algorithm 34. Samples from IG(a,b)

$Z \sim N(0,1)$ ▷standard normal sample, see A.9
$y = Z^2$
$x = (a/b) + (y - \sqrt{4aby + y^2})/(2b^2)$
$U \sim U(0,1)$
if $U < a/(a+bx)$ then return x;
else return $a^2/(b^2x)$;

A.11 Inverting the CDF

A sample X from a distribution having cdf $F(\cdot)$ is obtaining by solving $U = F(X)$ for X where $U \sim U(0,1)$. This derives from the following argument

$$\begin{aligned} U \sim U(0,1) \quad &\text{iff} \quad \Pr(U < u) = u \\ &\text{iff} \quad \Pr(F(X) < F(x)) = F(x) \\ &\text{iff} \quad \Pr(X < x) = F(x). \end{aligned}$$

The last line shows $X \sim F(\cdot)$.

For example, since the cdf of the exponential is $F(t) = 1 - e^{-\lambda t}$, one easily obtains an exponential sample by solving $U = F(T) = 1 - e^{-\lambda T}$,

$$T = \frac{-1}{\lambda}\log(1-U). \tag{A.16}$$

A.12 Cramer's Rule

A linear system with n equations in n unknowns is a system as follows,

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & r_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & r_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & r_n \end{array} .$$

The *discriminant* is the determinant of the matrix of coefficients,

$$\Delta = \det A = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} .$$

The solution for the ith unknown is the ratio of determinants

$$x_i = \frac{\det A_i}{\Delta}$$

where A_i is A with the ith column replaced by the right hand side,

$$A_i = \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} & r_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & r_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & r_n & a_{n,i+1} & \cdots & a_{nn} \end{bmatrix} .$$

For a 2 by 2 system this gives

$$x_1 = \frac{\det \begin{bmatrix} r_1 & a_{12} \\ r_2 & a_{22} \end{bmatrix}}{\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}} \qquad x_2 = \frac{\det \begin{bmatrix} a_{11} & r_1 \\ a_{21} & r_2 \end{bmatrix}}{\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}$$

as can be verified by direct substitution.

A.13 LaGrange Optimization

Let $y = f(\mathbf{x})$ be a real–valued function of a vector variable $\mathbf{x}$. We want to maximize (or minimize) f subject to the condition that $g(\mathbf{x}) = 0$ where g is a second real–valued function of $\mathbf{x}$ called the constraint. LaGrange noticed that f continues to increase (or decrease) as $\mathbf{x}$ varies along the path enforced by g until the tangents of the two curves are parallel. But then their gradients are parallel too (see (B.5)). Thus the optimum occurs when

$$\mathrm{grad} f = \lambda\, \mathrm{grad} g$$

where λ is a real number called the *LaGrange multiplier*. This equation along with the constraint constitute the system to be solved for the optimizers of the system.

The *LaGrangian* is the difference

$$\Lambda = f(\mathbf{x}) - \lambda g(\mathbf{x}).$$

The system of equations above can also be generated by setting to 0 the partial derivative of the LaGrangian with respect to each component of $\mathbf{x}$. Setting to 0 the partial derivative of the LaGrangian with respect to λ gives back the constraint equation.

A.14 Bisection Method for Roots

A simple and adequate method for numerical root solving is the *bisection* method. Suppose one wants the value of x which solves $y = f(x)$ for a given function f and value of y. This is the same problem as finding the root $g(x) = 0$ of g where $g(x) = f(x) - y$.

The method starts with an estimate of x that is too low, x_1, and an estimate that is too high, x_2. For example $g(x_1) < 0$ and $g(x_2) > 0$ (or conversely); thus $g(x_1)g(x_2) < 0$.

In this way the root we are seeking, x_0, is bracketed. Next the value of g is calculated for the mid-point

$$x_m = \frac{1}{2}(x_1 + x_2).$$

Then either x_1 or x_2 is set equal to x_m depending on whichever maintains a bracket on the root. The process is now iterated until the desired accuracy is attained.

B

Stochastic Calculus

In earlier chapters we have derived equations and algorithms based on the binomial tree and other discrete models. And we have alluded to the notion that these models converge to the correct prices as they become more refined, which is to say, in the limit as $n \to \infty$, or equivalently, as $\Delta t \to 0$. Just as in ordinary calculus, such a limit leads to a calculus of stochastic processes, the Itô calculus. In this chapter we introduce and analyze the continuous counterparts of the random walk models with which we have become familiar. In the next chapter we take up limits in the binomial model. Our objective lies in validating the limiting processes of the discrete models we have been using and not in analyzing the resulting continuous models per se. We also hope to acquaint the reader with the continuous approach.

B.1 The Itô Integral

The stochastic calculus we require is based on the Wiener Process W_t and hence on Brownian motion and arithmetical random walks. From Algorithm 1 on page 10 increments ΔX of the walk are given by $\Delta X = \mu \Delta t + \sigma \sqrt{\Delta t} Z_t$ where $Z_t \sim$ N(0,1) is a standard normal sample. Summing these increments we obtain the end point random variable X_T,

$$X_T = X_0 + \sum_{i=1}^{n} \left(\mu \Delta t + \sigma \sqrt{\Delta t} Z_i \right). \tag{B.1}$$

By letting $\Delta t \to 0$ we would expect to get in the limit the integral

$$X_T = X_0 + \int_0^T \mu dt + \int_0^T \sigma dW_t. \tag{B.2}$$

But the third term of this *stochastic integral* is quite unlike ordinary Riemann or Lebesgue integrals. How does one integrate with respect to the stochastic process W_t? In the following we hope to make this precise within the constraints of the scope and level of this text.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

The infinitesimal counterpart to (B.2) is the stochastic differential equation[1]

$$dX_t = \mu dt + \sigma dW_t. \tag{B.3}$$

We start by defining stochastic differentials dX_t.

Since all our modeling is referenced back to the Wiener Process, we are concerned with a calculus based on differentiating functions of W_t, for example $df(W_t)$ or possibly $df(t, W_t)$. To define such differentials we start where it all began, with the derivative of a real valued function of a real variable,

$$f'(t) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}.$$

To extend its application, it is helpful to rewrite the definition as

$$df = f(t+h) - f(t) = f'(t)h + o(h). \tag{B.4}$$

In this, *little oh* of h stands for terms that converge to zero as $h \to 0$ when divided by h. In this form the derivative, $f'(t)$, is seen as the best linear approximation to f at t.

For example, if f is a real valued function of a vector variable, $\mathbf{t} = [\,t_1 \quad t_2\,]^T$, then $f'(\mathbf{t})$ is the gradient,[2] the 1×2 matrix

$$\left[\, \tfrac{\partial f}{\partial t}_1 \quad \tfrac{\partial f}{\partial t}_2 \,\right] \tag{B.5}$$

since

$$f(t_1 + h_1, t_2 + h_2) - f(t_1, t_2) \approx \left[\, \tfrac{\partial f}{\partial t}_1 \quad \tfrac{\partial f}{\partial t}_2 \,\right] \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \frac{\partial f}{\partial t}_1 h_1 + \frac{\partial f}{\partial t}_2 h_2$$

and the difference is $o(h_1, h_2)$.

In the case at hand, defining the derivative of $df(W_t)$, the difference $f(W_{t+h}) - f(W_t)$ is a random variable and has a probability distribution. We will require that, in the limit, the mean be proportional to h, thus μh where μ is not a function of h, and the variance be proportional to increments of W_t. Thus the differential $d(f(W_t))$ is defined by

$$df(W_t) = \mu h + \sigma(W_{t+h} - W_t) + o(h) \tag{B.6}$$

which means that the difference between $df(W_t)$ and the first two terms on the right hand side is a random variable whose mean and variance are both $o(h)$.
Example. Calculate the differential $d(W_t^2)$.

$$\begin{aligned} W_{t+h}^2 - W_t^2 &= (W_{t+h} - W_t)(W_{t+h} + W_t) \\ &= (W_{t+h} - W_t)(2W_t + W_{t+h} - W_t) \\ &= 2W_t(W_{t+h} - W_t) + (W_{t+h} - W_t)^2. \end{aligned}$$

---

[1] In some texts (B.2) is taken as the definition of (B.3).

[2] More exactly, the gradient is a vector, in this case a 2×1 matrix, while the derivative is a linear functional (a 1×2 matrix). The gradient acts on its argument via dot product while the derivative acts via matrix multiplication; both yield the same result.

By the definition of a Wiener process, see Section 1.3, the term $W_{t+h} - W_t$ is a normally distributed random variable with mean 0 and variance h, that is $W_{t+h} - W_t = \sqrt{h} Z_h$ where $Z_h \sim$ N(0,1) (and may depend on h). Therefore the mean of this random variable squared is h,

$$\mathbb{E}((W_{t+h} - W_t)^2) = \text{var}(W_{t+h} - W_t) = h. \tag{B.7}$$

We may now calculate the variance of $(W_{t+h} - W_t)^2 - h$ as

$$\mathbb{E}((W_{t+h} - W_t)^2 - h)^2 = \mathbb{E}(hZ_h^2 - h)^2 = h^2 \mathbb{E}(Z_h^2 - 1)^2. \tag{B.8}$$

Since the expectation in the last member is finite (see Section 1.5.2), when divided by h this term goes to 0. Therefore we have calculated

$$d(W_t^2) = h + 2W_t(W_{t+h} - W_t),$$

that is to say, $\mu = 1$ and $\sigma = 2W_t$ in (B.6). In terms of infinitesimals,

$$d(W_t^2) = dt + 2W_t dW_t. \tag{B.9}$$

The surprise is the unexpected term dt. This is called the *Itô term*. It will appear again as we take up Itô's Lemma below. Another surprise in this example is that the Itô term is deterministic. As seen in (B.8), this is because the variance of each term goes to zero on the square of h, much faster than the mean, compare (B.7). As infinitesimals, these terms have no variance.

Before moving on we may exploit what we have derived in the Example by now integrating this infinitesimal equation to get

$$2\int_0^T W_t dW_t = \int_0^T d(W_t^2) - \int_0^T dt = W_T^2 - T$$

since $W_0 = 0$.

B.2 Itô's Lemma

Itô's Lemma can be thought of as the chain rule for stochastic calculus. In the sequel we assume that X_t is a stochastic process satisfying the stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \tag{B.10}$$

where W_t is a Wiener process. These are often referred to as *Itô processes*. As shown in (B.10), the coefficients μ and σ are allowed to be functions of both t and the process X_t itself. In the sequel we will notationally suppress this dependence and just write μ and σ.

Let $f(t, x)$ be a twice continuously differentiable function. Then Itô's Lemma asserts that

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma \frac{\partial f}{\partial x} dW_t. \tag{B.11}$$

Why this should be so can be seen from a second order Taylor expansion of f (higher order terms are all $o(h)$ and need not be considered). The first variable of f is an ordinary real valued variable and gives rise to the term $\frac{\partial f}{\partial t}dt$ in (B.11). We omit it from the discussion below and focus on the second variable in f. By Taylor's expansion

$$\begin{aligned} f(X_{t+h}) - f(X_t) &= \frac{\partial f}{\partial x} dX + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX)^2 + \dots \\ &= \frac{\partial f}{\partial x}(\mu dt + \sigma dW_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX)^2 \end{aligned} \tag{B.12}$$

We have dropped the higher order terms here and substituted for dX from (B.10). We consider $(dX)^2$ separately next. Again from (B.10)

$$\begin{aligned} (dX)^2 &= \mu^2 dt^2 + 2\mu dt\sigma dW + \sigma^2 (dW)^2 \\ &= \mu^2 dt^2 + 2\mu dt\sigma dW + \sigma^2((dW)^2 - dt) + \sigma^2 dt. \end{aligned}$$

The first term is clearly $o(dt)$ and has no variance. The second term has 0 mean. From (B.7), the variance of the second term is $4\mu^2\sigma^2(dt)^3$ and therefore this term has $o(dt)$ variance. Hence the first two terms can be ignored. Again from (B.7) the third term has mean 0 and from (B.8) it has variance equal to $\sigma^4(dt)^2 E(Z^2-1)^2$ where $Z \sim N(0,1)$. So the variance of this term is also $o(dt)$. Hence the only term that must be kept is the last one.

Returning to (B.12) with these calculations, we obtain (B.11). Note that the Itô term is $\frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}$.

For a quick example use of (B.11) take $dX = dW$ (so $\mu = 0$ and $\sigma = 1$ in (B.10)) and $f(t,x) = x^2$; verify (B.9).

An important application of (B.11) occurs for lognormal stock prices. Let $f(t,x) = \log(x)$ and $dS_t = \mu S dt + \sigma S dW_t$. The coefficients in (B.10) in this application are μ becomes μS and σ becomes σS. From (B.11)

$$\begin{aligned} d(\log S) &= \left(0 + \mu S\frac{1}{S} + \frac{1}{2}\sigma^2 S^2 \frac{-1}{S^2}\right) dt + \sigma S \frac{1}{S} dW \\ &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t. \end{aligned} \tag{B.13}$$

We may solve (B.13) by integrating

$$\begin{aligned} \log(S_T) - \log(S_0) &= \int_0^T (\mu - \frac{1}{2}\sigma^2)dt + \int_0^T \sigma dW_t \\ &= (\mu - \frac{1}{2}\sigma^2)T + \sigma W_T \\ &= (\mu - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z \end{aligned} \tag{B.14}$$

for $Z \sim N(0,1)$. In this way we verify (1.22) in an entirely different way.

B.3 Black-Scholes Differential Equation

In this section we derive the differential equation for a call option. This can be made to be an ordinary differential equation by constructing a portfolio consisting of both the option and some of the underlying stock because both are affected by the same stochastic process. As we have previously seen, the right amount of stock is that given by the delta of the option; we denote that amount by α. The derivation here is incomplete in that within the scope of this text we can not explain why it is that we are allowed to treat α as a constant in the derivation. For an explanation of this matter see for example [Jos03].

Let C be the price of a call option and let S be the price of the underlying stock. Consider a portfolio A consisting of one option and α shares of stock, $A = C + \alpha S$. The portfolio is a function of time and stock price, $A = A(t, S)$ and S follows geometric Brownian motion

$$dS = \mu S dt + \sigma S dW_t.$$

From Itô's Lemma (B.11) we have

$$dA = \left(\frac{\partial C}{\partial t} + \mu S \left(\frac{\partial C}{\partial S} + \alpha \right) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) dt + \sigma S \left(\frac{\partial C}{\partial S} + \alpha \right) dW_t.$$

The objective at this point is to eliminate the stochastic component of the portfolio; this can be achieved by choice of α:

$$\alpha = -\frac{\partial C}{\partial S}. \tag{B.15}$$

Then dA reduces to the ordinary differential

$$dA = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) dt. \tag{B.16}$$

Invoking arbitrage pricing theory, the portfolio must grow at the risk free rate r, so

$$\begin{aligned} \frac{dA}{dt} &= rA = r(C + \alpha S) \\ &= r(C - \frac{\partial C}{\partial S} S) \end{aligned}$$

Substituting for dA in (B.16) and transposing the partial derivative to the right hand side gives the *Black-Scholes differential equation*

$$rC = \frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}. \tag{B.17}$$

We must leave the development at this point except to say that, through a series of substitutions, (B.17) reveals itself to be a form of the well-known Heat Equation and can be solved analytically. The solution is (3.29).
Exercise:

Check that the expression for C given in (3.29) is a solution of (B.17).

C

Convergence of the Binomial Method

Consider now the binomial model with parameters p, u, and d. To adjust the model to the given application we match means and variances, see (1.34) and (1.35). One solution to these equations can be found by setting $p = 1/2$. With this choice, we show that the solution for u and d are[1]

$$\begin{aligned} u &= 1 + \mu\Delta t + \sigma\sqrt{\Delta t} + o(\Delta t) \\ d &= 1 + \mu\Delta t - \sigma\sqrt{\Delta t} + o(\Delta t). \end{aligned} \tag{C.1}$$

From (1.34) and the choice of p above

$$\begin{aligned} \frac{1}{2}u + \frac{1}{2}d &= \frac{1}{2}(1 + \mu\Delta t + \sigma\sqrt{\Delta t}) + \frac{1}{2}(1 + \mu\Delta t - \sigma\sqrt{\Delta t}) + o(\Delta t) \\ &= 1 + \mu\Delta t + \frac{(\mu\Delta t)^2}{2!} + \frac{(\mu\Delta t)^3}{3!} + \ldots = e^{\mu\Delta t}. \end{aligned}$$

Then for (1.35) we have

$$\begin{aligned} \frac{1}{2}u^2 + \frac{1}{2}d^2 &= \frac{1}{2}(1 + \mu\Delta t + \sigma\sqrt{\Delta t})^2 + \frac{1}{2}(1 + \mu\Delta t - \sigma\sqrt{\Delta t})^2 + o(\Delta t) \\ &= (1 + \mu\Delta t)^2 + \sigma^2\Delta t - \sigma^2\Delta t^2 + \frac{(2\mu + \sigma^2)^2(\Delta t)^2}{2!} + \ldots \\ &= e^{(2\mu+\sigma^2)\Delta t}. \end{aligned}$$

Let B be a Bernoulli random variable taking the value 1 with probability 1/2 and -1 with probability 1/2; the mean of B is $\mathbb{E}(B) = 0$ and its variance is $\mathrm{var}(B) = 1$. Then, up to terms of order $o(\Delta t)$, the binomial method advances one step with $S_{i+1} = S_i u$ if an up-step or $S_{i+1} = S_i d$ if a down-step; thus

[1] Recall that $o(\Delta t)$ refers to terms, which when divided by Δt, go to 0 as $\Delta t \to 0$.

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

$$S_{i+1} = S_i(1 + \mu \Delta t + \sigma \sqrt{\Delta t} B_i).$$

And hence

$$S_n = S_0 \prod_{i=1}^{n} \left(1 + \mu \Delta t + \sigma \sqrt{\Delta t} B_i\right). \tag{C.2}$$

We may now apply the development of Section 1.5.2 with the B_i in place of the Z_i. Just as in that case, the limiting end-point distribution for S_T, is given by (1.21) and repeated here

$$\log \frac{S_T}{S_0} \sim \mu T + \sigma \sqrt{T} Z - \frac{1}{2}\sigma^2 T.$$

Note that in the derivation on page 15, since $B^2 = 1$ with probability 1, the sum

$$-\frac{1}{2}\sigma^2 \Delta t \sum_{i=1}^{n} B_i^2 = -\frac{1}{2}\sigma^2 n \Delta t = -\frac{1}{2}\sigma^2 T,$$

directly without the need to invoke the Central Limit Theorem.

D

Variance Reduction Techniques

Three applications provided inspiration for the invention of the digital computer: the code breaking work at Bletchley Park, the ballistic artillery calculations at the Aberdeen Proving Grounds in Maryland, and the neutron flux calculations at Los Alamos. The Los Alamos challenge required no less than a unique and ingenious solution and thus the Monte Carlo method was born, and named, by Stan Ulam and John Von Neumann.

Initially a great deal of expectation was held out for the method. But soon it became clear that Monte Carlo was just too slow. Despite that, the method continued on, solving those problems that could only be solved by Monte Carlo. Meanwhile research went into the problem of speeding up Monte Carlo calculations. Being stochastic by its very nature, at the heart of the problem is that answers generated by Monte Carlo vary from run to run. So how can the variance be reduced?

One way is to perform more iterations because the variance is reduced by the factor $1/\sqrt{n}$, this means to halve the variance requires four times the iterations. Fortunately for Monte Carlo, computer speed has revived the method. For most of the algorithms given in this text, a few hundred thousand iterations can be conducted in under one second. Then too, often the accuracy required is to the nearest penny, three decimal places.

However a handful of the algorithms we encountered do indeed require a great deal of time. And so we seek methods for reducing the variance other than more iterations. In the following we give a brief introduction of the main techniques used to reduce variance in Monte Carlo applications. For a more comprehensive treatment, with financial applications, see [Gla03].

D.1 Antithetic Sampling

Virtually all of our simulations use either uniform random samples $U \sim U(0,1)$ or standard normal random samples $Z \sim N(0,1)$. In the case of a uniform sample U, then also $1-U$ is also a uniform sample. In the case of Z, then also $-Z$ is a standard normal sample. The pairs U and $1-U$ or Z and $-Z$ are called

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

antithetic variates. The idea behind antithetic variates is that if $C = G(S)$ is a calculation based on a random variable S, for example C is the value of a call option based on the expiration price S, then $C' = G(S')$ is likely to be negatively correlated with C if the simulation of S' is based on antithetic variates to those generating S.

This can be carried out quite easily. In the loop over $i = 1, 2, \ldots, n$ where each step S_i is generated, simultaneously generate S'_i by using Z for the former and $-Z$ for the latter. The option value estimate is based on all $2n$ observations

$$\bar{C} = \frac{1}{2n}\left(\sum_{k=1}^{n} C_k + \sum_{k=1}^{n} C'_k\right) = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{C_k + C'_k}{2}\right). \tag{D.1}$$

Although the $2n$ observations C_k and C'_k, $k = 1, \ldots, n$ are not independent, the n observations $(C_k + C'_k)/2$ are. Therefore the error estimates of Section A.8 apply to them.

D.2 Control Variates

Let C_k, $k = 1, 2, \ldots, n$ be the results of n replication of a simulation and we want to estimate $\mathbb{E}(C)$. Suppose on each replication we calculate another output R_k and that the expectation $\mathbb{E}(R)$ is known. Then for any constant β,

$$C_k(\beta) = C_k - \beta(R_k - \mathbb{E}(R)) \tag{D.2}$$

is a *control variate estimator*.

The expectation is given by

$$\bar{C}(\beta) = \bar{C} - \beta(\bar{R} - \mathbb{E}(R)) = \frac{1}{n}\sum_{k=1}^{n}\Big(C_k - \beta(R_k - \mathbb{E}(R))\Big). \tag{D.3}$$

The variance is given by

$$\begin{aligned} \mathrm{var}C(\beta) &= \mathrm{var}(C - \beta(R - \mathbb{E}(R)) \\ &= \mathrm{var}(C) - 2\beta\mathrm{covar}(C, R) + \beta^2\mathrm{var}(R). \end{aligned} \tag{D.4}$$

Thus, depending on the choice of β, the control variate estimator can have very much smaller variance then $\mathrm{var}(C)$. The optimal value is

$$\beta = \frac{\mathrm{covar}(C, R)}{\mathrm{var}(R)}. \tag{D.5}$$

Although knowing this value is equivalent to knowing $\mathbb{E}(C)$, β can be statistically estimated from the runs themselves.

As an example consider estimating the cost of a path dependent option $\mathbb{E}(C)$. The discounted stock prices S_i can be used as a control variate since, by the martingale property, $\mathbb{E}(e^{-rT}S_T) = S_0$.

D.3 Importance Sampling

Suppose we want to determine the expectation of a function of a random variable X and the function is only non-zero for rare values of X. Without compensating in some way, most of the values contributing to the expectation will be zero. This is as it should be of course, but at the same time it is inefficient. The idea in *importance sampling* is to change the probability density so that the rare events occur more often. We encountered an example of this problem when calculating the boundary of a shout option; only those paths that venture near the boundary area are effective at determining the boundary.

To treat the problem in general terms, assume we want to estimate the expectation

$$\theta = \mathbb{E}(h(X)) = \int h(x) f(x)\, dx \tag{D.6}$$

where $f(x)$ is the density of X. Let $g(x)$ be another density function which is positive wherever f is positive. The integral (D.6) can be written

$$\begin{aligned}\theta &= \int h(x) \frac{f(x)}{g(x)} g(x)\, dx \\ &= \mathbb{E}_g\left[h(X)\frac{f(X)}{g(X)}\right].\end{aligned} \tag{D.7}$$

The expectation here is with respect to the density g.

In order to gain insight as to the choice of g, suppose h is a positive function, then for some constant c, $\frac{1}{c}h(x)f(x)$ is a density. If now g were taken to be that density, then $hf/g = c$ and then the variance of the estimator hf/g is zero. (And from (D.7) $c = \theta$.) This shows that g should be chosen as close to the product hf as possible.

For example suppose that h is the indicator function of some set A, $h(x) = \mathbb{1}_A(x)$. Then $\theta = \Pr(X \in A)$. In this case the optimal choice for $g = hf/\Pr(X \in A)$ is the conditional density of X given $X \in A$. This means choosing g to make the event $X \in A$ more likely.

E

Shell Sort

The following program creates a permutation array r_i so that the doubly subscripted array x_{r_i} is sorted low to high,

$$x_{r_0} \leq x_{r_1} \leq \ldots \leq x_{r_{n-1}}.$$

```
void sort(double[] x, int[] rankPerm) {
 boolean done;
 int swap, gap, len=x.length;
 for( int i=0; i<len; i++) rankPerm[i] = i;

 gap=len/2; //integer arithmetic
 while( gap>=1 ) {
   do {
     done=true;
     for( int i=0; i<len-gap; i++) {
       if( x[rankPerm[i]] > x[rankPerm[i+gap]] )
       {
         swap = rankPerm[i];
         rankPerm[i] = rankPerm[i+gap];
         rankPerm[i+gap] = swap;
         done = false;
       } //end if
     } //end for
   } while(!done); //end of do.while
   gap = gap/2;
 }//end while
}
```

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

F

NextDayPrices Program

```
import java.io.*;
import java.util.*;
import static java.lang.System.err;

public class
nextDayPrices
{
  static int histoTrials = 200;
  static int trialsPerHisto = 50;
  static double binWidth = 1.0;
  static double binStart = 30.0;
  static double binEnd = 71.0;
  static int[] histo;

  static ArrayList rDate = new ArrayList();
  static ArrayList rOpen = new ArrayList();
  static ArrayList rHigh = new ArrayList();
  static ArrayList rLow = new ArrayList();
  static ArrayList rClose = new ArrayList();
  static ArrayList rVolume = new ArrayList();
  static ArrayList rAdjClose = new ArrayList();

  //===
  static int
  getPrices(String tkr)
  {
    FileInputStream fis=null;
    String line;
    StringBuffer ipBuff = new StringBuffer();
    try {
      //Open a reader on the file
      fis= new FileInputStream(tkr+".csv");
```

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

```
      BufferedReader reader =
      new BufferedReader(new InputStreamReader(fis));

      //Read the file and store in buffer
      while( (line=reader.readLine()) !=null) {
        ipBuff.append(line+"\n");
      }
      //Close the file
      reader.close();
      fis.close();

    } catch( IOException ioe ) {
      err.println("Exc. reading prices, error= "+ioe);
      err.println("Quitting...");
      System.exit(2);
    }
    return parseDataTable(ipBuff);
  }
  //===
  /**
  * return size of the lists or an error code
  * -1 no such element, -2 array index OOB,
  * -3 list sizes mis-match
  */
  static int
  parseDataTable(StringBuffer ipBuff)
  {
    String text = ipBuff.toString();
    String strTokens = ",\n";
    StringTokenizer st = new StringTokenizer(text,strTokens,false);
    String tokA;

    //first 7 are the table headers
    try {
      for( int i=0;i<7;++i) tokA = st.nextToken();
    } catch( NoSuchElementException nsee ) {
      err.println(" parseDataTable: (parsing table header) "+nsee);
      return -1;
    }

    //reset the arraylists
    rDate.clear(); rOpen.clear(); rHigh.clear();
    rLow.clear(); rClose.clear(); rVolume.clear();
    rAdjClose.clear();

    int j=0, k=st.countTokens();
```

```
    try{
      //pick the values from the web page (rows of 7)
      for( ; j<k/7; ++j) //over lines
      {
        rDate.add(st.nextToken()+" "); //format: yyyy-mm-dd
        rOpen.add(st.nextToken()+" ");
        rHigh.add(st.nextToken()+" ");
        rLow.add(st.nextToken()+" ");
        rClose.add(st.nextToken()+" ");
        rVolume.add(st.nextToken()+" ");
        rAdjClose.add(st.nextToken()+" ");
      }
    }catch( ArrayIndexOutOfBoundsException aioob) {
      err.println("parseDataTable@ got aioob exception, j= "+j);
      err.println("#tokens "+k+", so #lines "+(k/7));
      return -2;
    }
    if( j != rAdjClose.size() )
    {
      err.println("parseTable@ size mis-match among lists");
      return -3;
    }
    return j;
  }
  //===
  static void
  printHisto()
  {
    int maxHeight = 0;
    int nTrials=0;
    for( int i=0; i<histo.length; ++i)
    {
      nTrials += histo[i];
      maxHeight = Math.max(maxHeight,histo[i]);
    }

    //80 char rows, minus 14 for other material in the row
    double charRate = (80-14)/(double)maxHeight;

    System.out.println("\tNumber of histogram trials: "+nTrials);
    System.out.println(
    "\thistogram depicts fraction of cells at the given percentage\n");
    for( int i=0; i<histo.length; ++i )
    {
      System.out.format("%.2f |",binStart+i*binWidth);
      for( int j=0; j<(int)(charRate*histo[i]); ++j)
```

```
    System.out.print("*");
    System.out.format("%6.3f\n",histo[i]/(double)nTrials);
  }
}
//===
static void
removeCSVfiles(ArrayList tList)
{
  for( int i=0; i<tList.size(); ++i)
  {
    String fileName = (String)tList.get(i);
    File f = new File(fileName+".csv");
    if( !f.exists() ) continue;
    boolean success = f.delete();
    if( !success )
    err.println("nextDayPrices@ could not delete "+fileName);
  }
}
//mmm===
public static void
main(String[] args)
{
  long seed = System.currentTimeMillis();
  //seed = 17; //for debugging
  Random rng = new Random(seed);

  boolean getNames = false;
  boolean getData = true;
  getZipFile gzf = null;

  try{
    gzf = new getZipFile(getNames,getData);
  }catch( IOException ioe) {
    err.println(
    "Exception getting web based symbols files, error= "+ioe);
    err.println("Quitting.");
    System.exit(1);
  }
  System.out.println(""); //clear the dots

  int nBins = (int)Math.rint((binEnd-binStart)/binWidth);
  histo = new int[nBins];

  int nTickers = gzf.tickersList.size();
  int upCnt=0, dwnCnt=0, oaUp=0, oaDwn=0; //overall Up/Dwn
  int nPrices, i, j, bin;
```

```
    double p0, p1;
    double upAvg = 0;
    for( int h=1; h<=histoTrials; ++h)
    {
      upCnt = dwnCnt = 0;
      for( int n=0; n<trialsPerHisto; ++n)
      {
        i = rng.nextInt(nTickers); //ticker index
        nPrices = getPrices((String)gzf.tickersList.get(i));

        j = 1 + rng.nextInt(nPrices-1); //date, j=0 has no successor
        p0 = Double.parseDouble((String)rAdjClose.get(j));
        p1 = Double.parseDouble((String)rAdjClose.get(j-1));
        //err.format("date0= %s, p0= %.2f; date1= %s, p1= %.2f\n",
        //(String)rDate.get(j),p0,(String)rDate.get(j-1),p1);
        if( p1>p0 ) { ++upCnt; ++oaUp;}
        else if( p1<p0 ) {++dwnCnt; ++oaDwn; }
      }
      upAvg = 100*upCnt/(double)(upCnt+dwnCnt); //in percent
      bin = (int)((upAvg-binStart)/binWidth);
      if( 0<=bin && bin < nBins ) ++histo[bin];
      else
      System.out.format("outlier: %.2f\n",upAvg);
      if( h%10 == 0 ) System.out.print("*");
    }
    int totalTrials = histoTrials*trialsPerHisto;
    int totalCounted = oaUp+oaDwn;
    System.out.format("\n"+
    "Out of %d trials, up= %d (%.2f%%), dwn=%d (%.2f%%) nochange=%d\n",
    totalTrials,oaUp,100*oaUp/(double)totalCounted, oaDwn,
    100*oaDwn/(double)totalCounted,(totalTrials-totalCounted));
    printHisto();
    removeCSVfiles(gzf.tickersList);
  }
}
```

References

[Ach00] Achelis, S.B.: Technical Analysis from A to Z. McGraw-Hill, New York (2000)

[Bai94] Bailey, R.W.: Polar generation of random variaes with the t-distribution. Math. Comput. **62**(206), 779–781 (1994)

[Ber96] Bernstein, P.L.: Against the Gods. Wiley, New York (1996)

[Bjo04] Björk, T.: Arbitrage Theory in Continuous Time. Oxford University Press, New York (2004)

[BS02] Borokin, A., Salminsen, P.: Handbook of Brownian Motion – Facts and Formulae. Birkhäuser Verlag, Basel (2002)

[BFS83] Bratley, P., Fox, B.L., Schrage, L.E.: A Guide to Simulation. Springer, New York (1983)

[CZ03] Capiński, M., Zastawniak, T.: Mathematics for Finance, an Introduction to Financial Engineering, Springer, New York (2003)

[CD03] Carmona, R., Durrleman, V.: Pricing and hedging spread options. SIAM Rev. **45**(4), 627–685 (2003)

[CT04] Cont, R., Tankov, P.: Financial Modeling with Jump Processes. Chapman & Hall, New York (2004)

[Der96] Derman, E.: Outperformance options. In: Nelken, I. (ed.) The Handbook of Exotic Options. Irwin, Chicago (1996). Chapter Nine

[Fri07] Fries, C.: Mathematical Finance. Wiley–Interscience, New York (2007)

[Gla03] Glasserman, P.: Monte Carlo Methods in Financial Engineering. Springer, New York (2003)

[Gra76] Graybill, F.: Theory and Application of the Linear Model. Wadsworth Publishing, Belmont (1976)

[Hau99] Haugen, R.A.: The New Finance, the Case Against Efficient Markets. Prentice Hall, Upper Saddle River (1999)

[Hull11] Hull, J.C.: Options, Futures, and Other Derivatives. Prentice Hall, Upper Saddle River (2011)

[Jos03] Joshi, M.: The Concepts and Practice of Mathematical Finance. Cambridge University Press, New York (2003)

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,

[KT75] Karlin, S., Taylor, H.M.: A First Course in Stochastic Processes. Academic, New York (1975)

[Kel56] Kelly, J.L. Jr.: A new interpretation of information rate. Bell Syst. Tech. J. **35**, 917–926 (1956)

[KV90] Kemna, A.G.Z., Vorst, A.C.F.: A Pricing Method for Options Based on Average Asset Values. Econometrisch Instituut, Erasmus University, Rotterdam (1990). http://ideas.repec.org/a/eee/jbfina/v14y1990i1p113-129.html

[Lon90] Longstaff, F.A.: Pricing options with extendible maturities: analysis and applications. J. Financ. **XLV**(3), 935–957 (1990)

[LS01] Longstaff, F.A., Schwartz, E.S.: Valuing American options by simulation: a simple least squares approach. Rev. Financ. Stud. **14**(4), 557–583 (2001)

[Lun98] Luenberger, D.G.: Investment Science. Oxford University Press, New York (1998)

[Mal03] Malkiel, B.G.: A Random Walk Down Wall Street. W.W. Norton, New York (2003)

[Mar78] Margrabe, W.: The value of an option to exchange one asset for another. J. Financ. **33**, 177–186 (1978)

[Mer76] Merton, R.: Option pricing when the underlying stock returns are discontinous. J. Financ. Econ. **3**, 124–144 (1976)

[Mey98] Meyer, G.H.: Pricing options with transaction costs with the method of lines. In: Otani, M. (ed.) Nonlinear Evolution Equations and Applications. Kokyuroku, vol. 1061. RIMS Kyoto University (1998)

[Nel96] Nelken, I. (ed.), Handbook of Exotic Options. McGraw-Hill, New York (1996)

[Pap05] Papapantoleon, A.: An introduction to Lévy processes with applications in finance. Web document http://page.math.tu-berlin.de/~papapan/papers/introduction.pdf (2005)

[PI07] Park, C.-H., Irwin, S.H.: What do we know about the profitability of technical analysis. J. Econ. Surv. **21**(4), 786–826 (2007)

[Pou05] Poundstone, W.: Fortunes Formula. Hill and Wang, New York (2005)

[Rip87] Ripley, B.D.: Stochastic Simulation. Wiley, New York (1987)

[Rom12] Roman, S.: Introduction to the Mathematics of Finance. Springer, New York (2012)

[Sch03] Schoutens, W.: Lévy Processes in Finance, Pricing Financial Derivatives. Wiley, New York (2003)

[SM09] Shonkwiler, R., Mendivil, F.: Explorations in Monte Carlo Methods. Springer, New York (2009)

[Yud10] Yudaken, L.: Numerical pricing of shout options. MSc thesis, Trinity College, University of Oxford, Oxford, UK (2010)

Index

R.W. Shonkwiler, *Finance with Monte Carlo*, Springer Undergraduate Texts in Mathematics and Technology, DOI 10.1007/978-1-4614-8511-7,