8

BROWNIAN MOTION

> Observe what happens when sunbeams are admitted into a building and shed light on its shadowy places. You will see a multitude of tiny particles mingling in a multitude of ways … their dancing is an actual indication of underlying movements of matter that are hidden from our sight. … It originates with the atoms which move of themselves [i.e., spontaneously]. Then those small compound bodies that are least removed from the impetus of the atoms are set in motion by the impact of their invisible blows and in turn cannon against slightly larger bodies. So the movement mounts up from the atoms and gradually emerges to the level of our senses, so that those bodies are in motion that we see in sunbeams, moved by blows that remain invisible.
>
> —Lucretius, *On the Nature of Things*, 60 B.C.

8.1 INTRODUCTION

Brownian motion is a stochastic process, which is rooted in a physical phenomenon discovered almost 200 years ago. In 1827, the botanist Robert Brown, observing pollen grains suspended in water, noted the erratic and continuous movement of tiny particles ejected from the grains. He studied the phenomenon for many years, ruled out the belief that it emanated from some "life force" within the pollen, but could not explain the motion. Neither could any other scientist of the 19th century.

In 1905, Albert Einstein solved the riddle in his paper, *On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat*. Einstein explained the movement by the continual bombardment of the

Introduction to Stochastic Processes with R, First Edition. Robert P. Dobrow.

immersed particles by the molecules in the liquid, resulting in "motions of such magnitude that these motions can easily be detected by a microscope." Einstein's theoretical explanation was confirmed 3 years later by empirical experiment, which led to the acceptance of the atomic nature of matter.

The description of the motion of dust particles in the classic poem *On the Nature of Things*, written by the Roman philosopher Lucretius over 2,000 years ago as an ancient proof of the existence of atoms, could have been a summary of Einstein's work!

Einstein showed that the position x of a particle at time t was described by the partial differential heat equation

$$\frac{\partial}{\partial t} f(x, t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(x, t),$$

where $f(x, t)$ represents the density (number of particles per unit volume) at position x and time t. The solution to that equation is

$$f(x, t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t},$$

which is the probability density function of the normal distribution with mean 0 and variance t.

The mathematical object we call Brownian motion is a continuous-time, continuous-state stochastic process, also called the *Wiener process*, named after the American mathematician Norbert Wiener. The British mathematician Bertrand Russell influenced Wiener to take up the theory of Brownian motion as had been studied by Einstein. In his 1956 autobiography, Wiener writes,

> The Brownian motion was nothing new as an object of study by physicists. There were fundamental papers by Einstein and Smoluchowski that covered it. But whereas these papers concerned what was happening to any given particle at a specific time, or the long-time statistics of many particles, they did not concern themselves with the mathematical properties of the curve followed by a single particle.
>
> Here the literature was very scant, but it did include a telling comment by the French physicist Perrin in his book *Les Atomes*, where he said in effect that the very irregular curves followed by particles in the Brownian motion led one to think of the supposed continuous non-differentiable curves of the mathematicians.

Standard Brownian Motion

A continuous-time stochastic process $(B_t)_{t \geq 0}$ is a *standard Brownian motion* if it satisfies the following properties:

1. $B_0 = 0$.

2. *(Normal distribution)* For $t > 0$, B_t has a normal distribution with mean 0 and variance t.
3. *(Stationary increments)* For $s, t > 0$, $B_{t+s} - B_s$ has the same distribution as B_t. That is,

$$P(B_{t+s} - B_s \leq z) = P(B_t \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}\, dx,$$

   for $-\infty < z < \infty$.
4. *(Independent increments)* If $0 \leq q < r \leq s < t$, then $B_t - B_s$ and $B_r - B_q$ are independent random variables.
5. *(Continuous paths)* The function $t \mapsto B_t$ is continuous, with probability 1.

The normal distribution plays a central role in Brownian motion. The reader may find it helpful to review properties of the univariate and bivariate normal distributions in Appendix B, Section B.4. We write $X \sim \text{Normal}(\mu, \sigma^2)$ to mean that the random variable X is normally distributed with mean μ and variance σ^2.

Brownian motion can be thought of as the motion of a particle that diffuses randomly along a line. At each point t, the particle's position is normally distributed about the line with variance t. As t increases, the particle's position is more *diffuse*; see Figure 8.1.

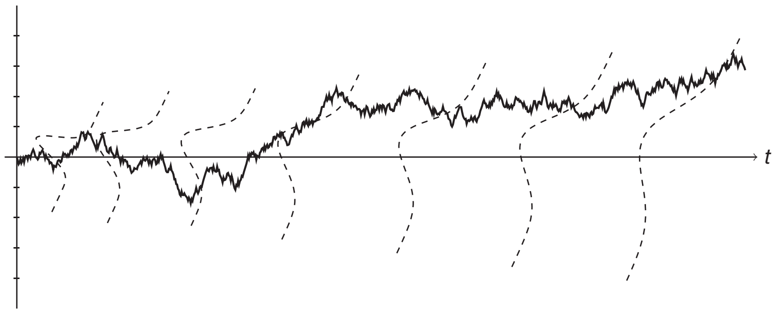

Figure 8.1 Brownian motion path. Superimposed on the graph are normal density curves with mean 0 and variance t.

It is not at all obvious that a stochastic process with the properties of Brownian motion actually exists. And Wiener's fundamental contribution was proving this existence. A rigorous derivation of Brownian motion is beyond the scope of this book, and requires measure theory and advanced analysis. The difficult part is showing the existence of a process that has stationary and independent increments together with continuous paths. The issue is touched upon at the end of the next section. First, however, we get our hands dirty with some calculations.

Computations involving Brownian motion are often tackled by exploiting stationary and independent increments. In the following examples, the reader may recognize similarities with the Poisson process, another stochastic process with stationary and independent increments. Unless stated otherwise, B_t denotes standard Brownian motion.

Example 8.1 For $0 < s < t$, find the distribution of $B_s + B_t$.

Solution Write $B_s + B_t = 2B_s + (B_t - B_s)$. By independent increments, B_s and $B_t - B_s$ are independent random variables, and thus $2B_s$ and $B_t - B_s$ are independent. The sum of independent normal variables is normal. Thus, $B_s + B_t$ is normally distributed with mean $E(B_s + B_t) = E(B_s) + E(B_t) = 0$, and variance

$$\begin{aligned} Var(B_s + B_t) &= Var(2B_s + (B_t - B_s)) = Var(2B_s) + Var(B_t - B_s) \\ &= 4Var(B_s) + Var(B_{t-s}) = 4s + (t - s) \\ &= 3s + t. \end{aligned}$$

The second equality is because the variance of a sum of independent random variables is the sum of their variances. The third equality uses stationary increments. We have that $B_s + B_t \sim \text{Normal}(0, 3s + t)$. ■

Example 8.2 A particle's position is modeled with a standard Brownian motion. If the particle is at position 1 at time $t = 2$, find the probability that its position is at most 3 at time $t = 5$.

Solution The desired probability is

$$\begin{aligned} P(B_5 \leq 3|B_2 = 1) &= P(B_5 - B_2 \leq 3 - B_2|B_2 = 1) \\ &= P(B_5 - B_2 \leq 2|B_2 = 1) \\ &= P(B_5 - B_2 \leq 2) \\ &= P(B_3 \leq 2) = 0.876. \end{aligned}$$

The third equality is because $B_5 - B_2$ and B_2 are independent. The penultimate equality is by stationary increments. The desired probability in `R` is

```
> pnorm(2,0,sqrt(3))
[1] 0.8758935
```

Note that in `R` commands involving the normal distribution are parameterized by standard deviation, not variance. ■

Example 8.3 Find the covariance of B_s and B_t.

Solution For the covariance,

$$\text{Cov}(B_s, B_t) = E(B_s B_t) - E(B_s)E(B_t) = E(B_s B_t).$$

For $s < t$, write $B_t = (B_t - B_s) + B_s$, which gives

$$\begin{aligned} E(B_s B_t) &= E(B_s(B_t - B_s + B_s)) \\ &= E(B_s(B_t - B_s)) + E\left(B_s^2\right) \\ &= E(B_s)E(B_t - B_s) + E\left(B_s^2\right) \\ &= 0 + Var(B_s) = s. \end{aligned}$$

Thus, $\text{Cov}(B_s, B_t) = s$. For $t < s$, by symmetry $\text{Cov}(B_s, B_t) = t$. In either case,

$$\text{Cov}(B_s, B_t) = \min\{s, t\}.$$

■

Simulating Brownian Motion

Consider simulating Brownian motion on $[0, t]$. Assume that we want to generate n variables at equally spaced time points, that is $B_{t_1}, B_{t_2}, \ldots, B_{t_n}$, where $t_i = it/n$, for $i = 1, 2, \ldots, n$. By stationary and independent increments, with $B_{t_0} = B_0 = 0$,

$$B_{t_i} = B_{t_{i-1}} + \left(B_{t_i} - B_{t_{i-1}}\right) \stackrel{d}{=} B_{t_{i-1}} + X_i,$$

where X_i is normally distributed with mean 0 and variance $t_i - t_{i-1} = t/n$, and is independent of $B_{t_{i-1}}$. The notation $X \stackrel{d}{=} Y$ means that random variables X and Y have the same distribution.

This leads to a recursive simulation method. Let $Z_1, Z_2, \ldots, Z_n$ be independent and identically distributed standard normal random variables. Set

$$B_{t_i} = B_{t_{i-1}} + \sqrt{t/n}Z_i, \text{ for } i = 1, 2, \ldots, n.$$

This gives

$$B_{t_i} = \sqrt{\frac{t}{n}}(Z_1 + \cdots + Z_n).$$

In R, the cumulative sum command

```
> cumsum(rnorm(n,0,sqrt(t/n)))
```

generates the Brownian motion variables $B_{t/n}, B_{2t/n}, \ldots, B_t$.

Simulations of Brownian motion on [0, 1] are shown in Figure 8.2. The paths were drawn by simulating $n = 1,000$ points in [0, 1] and then connecting the dots.

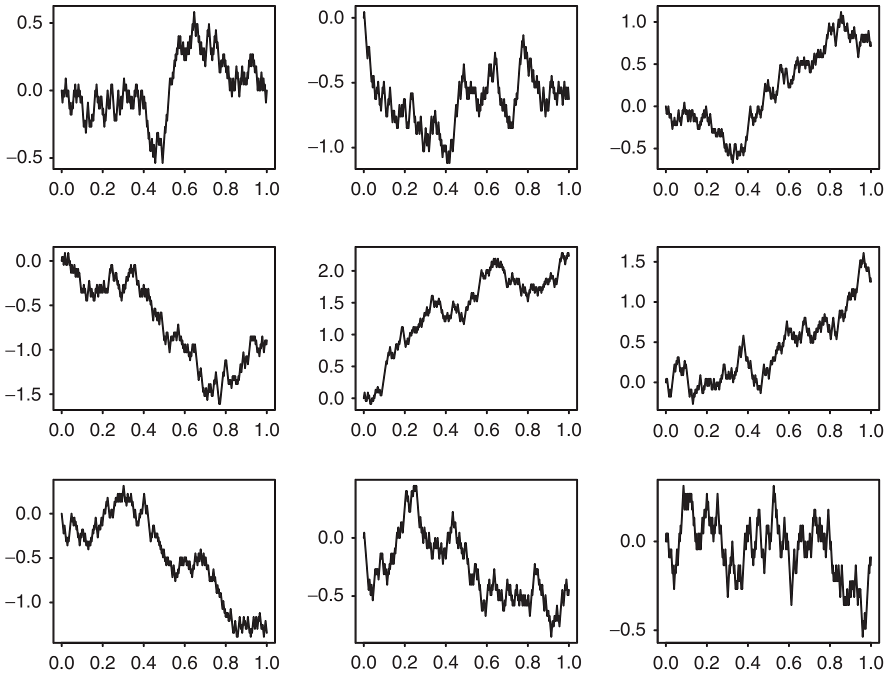

Figure 8.2 Sample paths of Brownian motion on [0, 1].

R: Simulating Brownian Motion

```
# bm.R
> n <- 1000
> t <- 1
> bm <- c(0, cumsum(rnorm(n,0,sqrt(t/n))))
> steps <- seq(0,t,length=n+1)
> plot(steps,bm,type="l")
```

More generally, to simulate $B_{t_1}, B_{t_2}, \ldots, B_{t_n}$, for time points $t_1 < t_2 < \cdots < t_n$, set

$$B_{t_i} = B_{t_{i-1}} + \sqrt{t_i - t_{i-1}} Z_i, \text{ for } i = 1, 2, \ldots, n,$$

with $t_0 = 0$.

Sample Space for Brownian Motion and Continuous Paths*

Consider the fifth defining property of Brownian motion: the function $t \mapsto B_t$ is continuous.

A continuous-time stochastic process $(X_t)_{-\infty<t<\infty}$ is a collection of random variables defined on a common sample space, or probability space, Ω. A random variable is really a function on a probability space. If $X : \Omega \to \mathbb{R}$ is a random variable, then X takes values depending on the outcome $\omega \in \Omega$. Thus, we can write $X = X(\omega)$ to emphasize the dependence on the outcome ω, although we usually suppress ω for simplicity.

In the context of stochastic processes, $X_t = X_t(\omega)$ is a function of two variables: t and ω. For fixed $t \in \mathbb{R}$, X_t is a random variable. Letting $X_t(\omega)$ vary as $\omega \in \Omega$ generates the different values of the process at the fixed time t. On the other hand, for fixed $\omega \in \Omega$, $X_t(\omega)$ is a function of t. Letting t vary generates a *sample path* or *realization*. One can think of these realizations as *random functions*.

For instance, each of the graphs in Figure 8.2, is one realization of such a random function. Using function notation, we could write $f(t) = B_t(\omega)$, for $-\infty < t < \infty$. For fixed ω, it makes sense to ask whether f is continuous, differentiable, etc.

A more precise statement of Property 5 is that

$$P(\omega \in \Omega : B_t(\omega) \text{ is a continuous function of } t) = 1.$$

Implicit in this statement is the existence of (i) the probability space Ω and (ii) a suitable probability function, or *probability measure*, P, which is consistent with the other defining properties of Brownian motion.

A probability space Ω is easy enough to identify. Let Ω be the set of all continuous functions on $[0, \infty)$. Each $\omega \in \Omega$ is a continuous function. Then, $B_t(\omega) = \omega_t$, the value of ω evaluated at t. This is the easiest way to insure that B_t has continuous sample paths. The hard part is to construct a probability function P on the set of continuous functions, which is consistent with the properties of Brownian motion. This was precisely Norbert Wiener's contribution. That probability function, introduced by Wiener in 1923, is called *Wiener measure*.

8.2 BROWNIAN MOTION AND RANDOM WALK

Continuous-time, continuous-state Brownian motion is intimately related to discrete-time, discrete-state random walk. Brownian motion can be constructed from simple symmetric random walk by suitably scaling the values of the walk while simultaneously speeding up the steps of the walk.

Let $X_1, X_2, \dots$ be an i.i.d. sequence with each X_i taking values ± 1 with probability 1/2 each. Set $S_0 = 0$ and for integer $t > 0$, let $S_t = X_1 + \cdots + X_t$. Then, $S_0, S_1, S_2, \dots$ is a simple symmetric random walk with $E(S_t) = 0$ and $Var(S_t) = t$ for $t = 0, 1, \dots$. As a sum of i.i.d. random variables, for large t, S_t is approximately normally distributed by the central limit theorem.

The random walk has independent increments. For integers $0 < q < r < s < t$, $S_t - S_s = X_{s+1} + \cdots + X_t$, and $S_r - S_q = X_{q+1} + \cdots + X_r$. Since $X_{s+1}, \dots, X_t$ is independent of $X_{q+1}, \dots, X_r$, it follows that $S_t - S_s$ and $S_r - S_q$ are independent random variables. Furthermore, for all integers $0 < s < t$, the distribution of $S_t - S_s$ and S_{t-s}

is the same since they are both a function of $t - s$ distinct X_i. Thus, the walk has stationary increments.

The simple random walk is a discrete-state process. To obtain a continuous-time process with continuous sample paths, values are connected by linear interpolation. Recall that $\lfloor x \rfloor$ is the *floor* of x, or integer part of x, which is the largest integer not greater than x. Extending the definition of S_t to real $t \geq 0$, let

$$S_t = \begin{cases} X_1 + \cdots + X_t, & \text{if } t \text{ is an integer,} \\ S_{\lfloor t \rfloor} + X_{\lfloor t \rfloor + 1} (t - \lfloor t \rfloor), & \text{otherwise.} \end{cases}$$

Observe that if k is a positive integer, then for $k \leq t \leq k + 1$, S_t is the linear interpolation of the points (k, S_k) and $(k + 1, S_{k+1})$. See Figure 8.3 to visualize the construction. We have $E(S_t) = 0$ and

$$\begin{aligned} Var(S_t) &= Var(S_{\lfloor t \rfloor} + X_{\lfloor t \rfloor + 1} (t - \lfloor t \rfloor)) \\ &= Var(S_{\lfloor t \rfloor}) + (t - \lfloor t \rfloor)^2 Var(X_{\lfloor t \rfloor + 1}) \\ &= \lfloor t \rfloor + (t - \lfloor t \rfloor)^2 \approx t, \end{aligned}$$

as $0 \leq t - \lfloor t \rfloor < 1$.

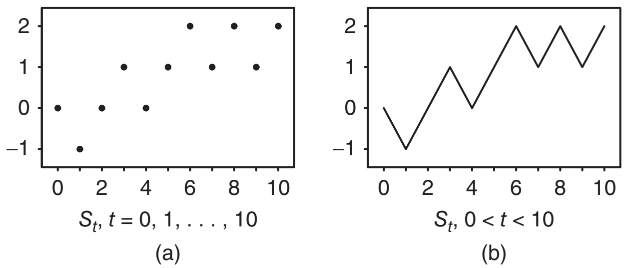

Figure 8.3 (a) Realization of a simple random walk. (b) Walk is extended to a continuous path by linear interpolation.

The process is now scaled both vertically and horizontally. Let $S_t^{(n)} = S_{nt}/\sqrt{n}$, for $n = 1, 2, \ldots$ On any interval, the new process has n times as many steps as the original walk. And the height at each step is shrunk by a factor of $1/\sqrt{n}$. The construction is illustrated in Figure 8.4.

The scaling preserves mean and variance, as $E\left(S_t^{(n)}\right) = 0$ and $Var\left(S_t^{(n)}\right) = Var(S_{nt})/n \approx t$. Sample paths are continuous and for each n, the process retains independent and stationary increments. Considering the central limit theorem, it is reasonable to think that as $n \to \infty$, the process converges to Brownian motion. This, in fact, is the case, a result proven by Monroe Donsker in 1951.

We have not precisely said what *convergence* actually means since we are not talking about convergence of a sequence of numbers, nor even convergence of random variables, but rather convergence of stochastic processes. We do not

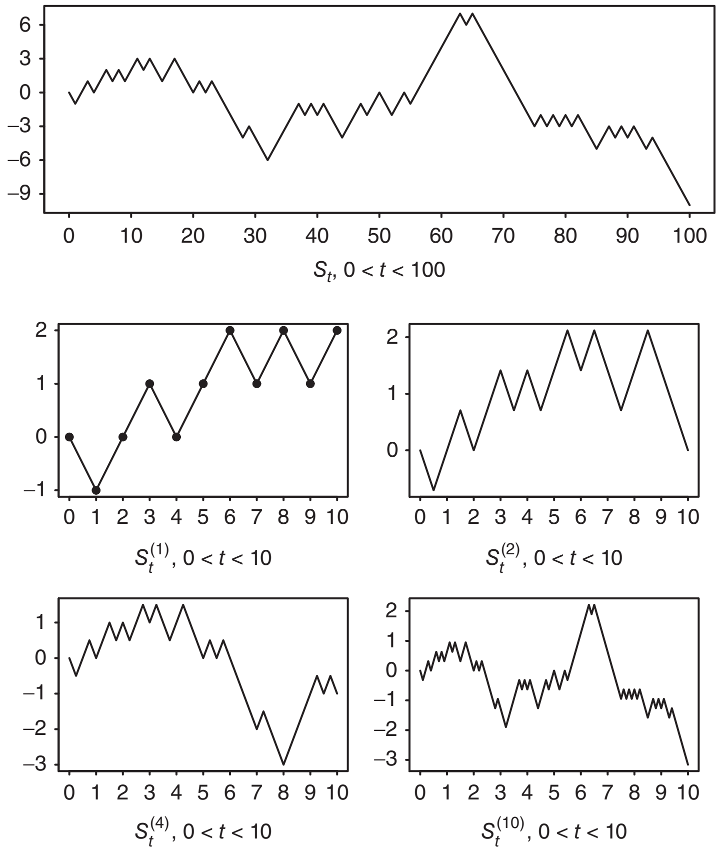

Figure 8.4 Top graph is a simple symmetric random walk for 100 steps. Bottom graphs show the scaled process $S_t^{(n)} = S_{nt}/\sqrt{n}$, for $n = 1, 2, 4, 10$.

give a rigorous statement. Nevertheless, the reader can trust their intuition that for large n, $\left(S_{nt}/\sqrt{n}\right)_{t\geq 0}$ behaves like a standard Brownian motion process $(B_t)_{t\geq 0}$.

Invariance Principle*

The construction of Brownian motion from simple symmetric random walk can be generalized so that we start with any i.i.d. sequence $X_1, \ldots, X_n$ with mean 0 and variance 1. Let $S_n = X_1 + \cdots + X_n$. Then, $S_{nt}/\sqrt{n}$ converges to B_t, as $n \to \infty$. This is known as *Donsker's invariance principle*. The word *invariance* is used because all random walks with increments that have mean 0 and variance 1, regardless of distribution, give the same Brownian motion limit.

A consequence of the invariance principle and continuity is that *functions* of the discrete process $S_{nt}/\sqrt{n}$ converge to the corresponding function of

Brownian motion, as $n \to \infty$. If g is a bounded, continuous function, whose domain is the set of continuous functions on $[0, 1]$, then $g\left(S_{nt}/\sqrt{n}\right) \approx g(B_t)$, for large n.

Functions whose domain is a set of functions are called *functionals*. The invariance principle means that properties of random walk and of functionals of random walk can be derived by considering analogous properties of Brownian motion, and vice versa.

For example, assume that f is a continuous function on $[0, 1]$. Let $g(f) = f(1)$, the evaluation of $f(t)$ at $t = 1$. Then, $g(S_{nt}/\sqrt{n}) = S_n/\sqrt{n}$ and $g(B_t) = B_1$. By the invariance principle, $S_n/\sqrt{n} \to B_1$, as $n \to \infty$. The random variable B_1 is normally distributed with mean 0 and variance 1, and we have thus recaptured the central limit theorem from Donsker's invariance principle.

Example 8.4 For a simple symmetric random walk, consider the maximum value of the walk in the first n steps. Let $g(f) = \max_{0 \le t \le 1} f(t)$. By the invariance principle,

$$\lim_{n\to\infty} g\left(\frac{S_{tn}}{\sqrt{n}}\right) = \lim_{n\to\infty} \max_{0\le t\le 1} \frac{S_{tn}}{\sqrt{n}} = \lim_{n\to\infty} \max_{0\le k\le n} \frac{S_k}{\sqrt{n}} = g\left(B_t\right) = \max_{0\le t\le 1} B_t.$$

This gives $\max_{0\le k\le n} S_k \approx \sqrt{n}\max_{0\le t\le 1} B_t$, for large n.

In Section 8.4, it is shown that the random variable $M = \max_{0\le t\le 1} B_t$ has density function

$$f_M(x) = \sqrt{\frac{2}{\pi}} e^{-x^2/2}, \text{ for } x > 0.$$

Mean and standard deviation are

$$E(M) = \sqrt{\frac{2}{\pi}} \approx 0.80 \quad \text{and} \quad SD(M) = \frac{\pi - 2}{\pi} \approx 0.60.$$

With these results, we see that in the first n steps of simple symmetric random walk, the maximum value is about $(0.80)\sqrt{n}$ give or take $(0.60)\sqrt{n}$. In $n = 10,000$ steps, the probability that a value greater than 200 is reached is

$$\begin{aligned} P\left(\max_{0\le k\le n} S_k > 200\right) &= P\left(\max_{0\le k\le n} \frac{S_k}{100} > 2\right) \\ &= P\left(\max_{0\le k\le n} \frac{S_k}{\sqrt{n}} > 2\right) \\ &\approx P(M > 2) \\ &= \int_2^\infty \sqrt{\frac{2}{\pi}} e^{-x^2/2}\, dx = 0.0455. \end{aligned}$$

■

R: Maximum for Simple Symmetric Random Walk

```
# maxssrw.R
> n <- 10000
> sim <- replicate(10000,
  + max(cumsum(sample(c(-1,1),n,replace=T))))
> mean(sim)
[1] 79.7128
> sd(sim)
[1] 60.02429
> sim <- replicate(10000,
  + if(max(cumsum(sample(c(-1,1),n,rep=T)))>200)
  1 else 0)
> mean(sim) # P(max > 200)
[1] 0.0461
```

8.3 GAUSSIAN PROCESS

The normal distribution is also called the *Gaussian distribution*, in honor of Carl Friedrich Gauss, who introduced the distribution more than 200 years ago. The bivariate and multivariate normal distributions extend the univariate distribution to higher finite-dimensional spaces. A *Gaussian process* is a continuous-time stochastic process, which extends the Gaussian distribution to infinite dimensions. In this section, we show that Brownian motion is a Gaussian process and identify the conditions for when a Gaussian process is Brownian motion.

(Gauss, considered by historians to have been one of the greatest mathematicians of all time, first used the normal distribution as a model for measurement error in celestial observations, which led to computing the orbit of the planetoid Ceres. The story of the discovery of Ceres, and the contest to compute its orbit, is a mathematical page-turner. See *The Discovery of Ceres: How Gauss Became Famous* by Teets and Whitehead (1999).)

We first define the multivariate normal distribution and establish some of its key properties.

Multivariate Normal Distribution

Random variables $X_1, \dots, X_k$ have a *multivariate normal distribution* if for all real numbers $a_1, \dots, a_k$, the linear combination

$$a_1X_1 + \cdots + a_kX_k$$

has a univariate normal distribution. A multivariate normal distribution is completely determined by its *mean vector*

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k) = (E(X_1), \ldots, E(X_k))$$

and *covariance matrix* $\boldsymbol{V}$, where

$$V_{ij} = \text{Cov}(X_i, X_j), \text{for } 1 \leq i, j \leq k.$$

The joint density function of the multivariate normal distribution is

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{k/2}|\boldsymbol{V}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{V}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right),$$

where $\boldsymbol{x} = (x_1, \ldots, x_k)$ and $|\boldsymbol{V}|$ is the determinant of $\boldsymbol{V}$.

The multivariate normal distribution has the remarkable property that all marginal and conditional distributions are normal. If $X_1, \ldots, X_k$ have a multivariate normal distribution, then the X_i are normally distributed, joint distributions of subsets of the X_i have multivariate normal distributions, and conditional distributions given subsets of the X_i are normal.

If $X_1, \ldots, X_k$ are independent normal random variables, then their joint distribution is multivariate normal. For jointly distributed normal random variables, independence is equivalent to being uncorrelated. That is, if X and Y are jointly distributed normal random variables, then X and Y are independent if and only if $E(XY) = E(X)E(Y)$.

A *Gaussian process* extends the multivariate normal distribution to stochastic processes.

Gaussian Process

A *Gaussian process* $(X_t)_{t\geq 0}$ is a continuous-time stochastic process with the property that for all $n = 1, 2, \ldots$ and $0 \leq t_1 < \cdots < t_n$, the random variables $X_{t_1}, \ldots, X_{t_n}$ have a multivariate normal distribution.

A Gaussian process is completely determined by its *mean function* $E(X_t)$, for $t \geq 0$, and *covariance function* $\text{Cov}(X_s, X_t)$, for $s, t \geq 0$.

Standard Brownian motion is a Gaussian process. The following characterization gives conditions for when a Gaussian process is a standard Brownian motion.

Gaussian Process and Brownian Motion

A stochastic process $(B_t)_{t\geq 0}$ is a standard Brownian motion if and only if it is a Gaussian process with the following properties:

1. $B_0 = 0$.
2. *(Mean function)* $E(B_t) = 0$, for all t.
3. *(Covariance function)* $\mathrm{Cov}(B_s, B_t) = \min\{s,t\}$, for all s, t.
4. *(Continuous paths)* The function $t \mapsto B_t$ is continuous, with probability 1.

Proof. Let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Consider $(B_{t_1}, B_{t_2}, \ldots, B_{t_k})$, for $0 < t_1 < t_2 < \cdots < t_k$. For constants $a_1, a_2, \ldots, a_k$, we need to show that $a_1 B_{t_1} + a_2 B_{t_2} + \cdots + a_k B_{t_k}$ has a univariate normal distribution. By independent increments, $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}}$ are independent normal random variables, whose joint distribution is multivariate normal. Write

$$\begin{aligned} &a_1 B_{t_1} + a_2 B_{t_2} + \cdots + a_k B_{t_k} \\ &\quad = (a_1 + \cdots + a_k) B_{t_1} + (a_2 + \cdots + a_k)(B_{t_2} - B_{t_1}) \\ &\qquad + \cdots + (a_{k-1} + a_k)(B_{t_{k-1}} - B_{t_{k-2}}) + a_k (B_{t_k} - B_{t_{k-1}}), \end{aligned}$$

which is a linear combination of $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}}$, and thus has a univariate normal distribution.

The mean and covariance functions for Brownian motion are

$$E(B_t) = 0 \quad \text{and} \quad \mathrm{Cov}(B_s, B_t) = \min\{s,t\}.$$

It follows that standard Brownian motion is the unique Gaussian process with these mean and covariance functions.

Conversely, assume that $(B_t)_{t\geq 0}$ is a Gaussian process that satisfies the stated properties. We need to show the process has stationary and independent increments.

Since the process is Gaussian, for $s, t \geq 0$, $B_{t+s} - B_s$ is normally distributed with mean $E(B_{t+s} - B_s) = E(B_{t+s}) - E(B_s) = 0$, and variance

$$Var(B_{t+s} - B_s) = Var(B_{t+s}) + Var(B_s) - 2\mathrm{Cov}(B_{t+s}, B_s) = (t+s) + s - 2s = t.$$

Thus, $B_{t+s} - B_s$ has the same distribution as B_t, which gives stationary increments.

For $0 \leq q < r \leq s < t$,

$$\begin{aligned} E((B_r - B_q)(B_t - B_s)) &= E(B_r B_t) - E(B_r B_s) - E(B_q B_t) + E(B_q B_s) \\ &= \mathrm{Cov}(B_r, B_t) - \mathrm{Cov}(B_r, B_s) - \mathrm{Cov}(B_q, B_t) + \mathrm{Cov}(B_q, B_s) \\ &= r - r - q + q = 0. \end{aligned}$$

Thus, $B_r - B_q$ and $B_t - B_s$ are uncorrelated. Since $B_r - B_q$ and $B_t - B_s$ are normally distributed, it follows that they are independent.

We have shown that $(B_t)_{t\geq 0}$ is a standard Brownian motion. ■

Example 8.5 For $a > 0$, let $X_t = B_{at}/\sqrt{a}$, for $t \geq 0$. Show that $(X_t)_{t\geq 0}$ is a standard Brownian motion.

Solution For $0 \leq t_1 < \cdots < t_k$ and real numbers $a_1, \ldots, a_k$,

$$\sum_{i=1}^{k} a_i X_{t_i} = \sum_{i=1}^{k} \frac{a_i}{\sqrt{a}} B_{at_i},$$

which has a univariate normal distribution, since $(B_t)_{t\geq 0}$ is a Gaussian process. Thus, $(X_t)_{t\geq 0}$ is a Gaussian process. Clearly, $X_0 = 0$. The mean function is

$$E(X_t) = E\left(\frac{1}{\sqrt{a}} B_{at}\right) = \frac{1}{\sqrt{a}} E(B_{at}) = 0.$$

The covariance function is

$$\begin{aligned}\operatorname{Cov}(X_s, X_t) &= \operatorname{Cov}\left(\frac{1}{\sqrt{a}} B_{as}, \frac{1}{\sqrt{a}} B_{at}\right) = \frac{1}{a}\operatorname{Cov}(B_{as}, B_{at}) \\ &= \frac{1}{a}\min\{as, at\} = \min\{s, t\}.\end{aligned}$$

Finally, path continuity of $(X_t)_{t\geq 0}$ follows from the path continuity of standard Brownian motion, as the function $t \mapsto B_{at}/\sqrt{a}$ is continuous for all $a > 0$, with probability 1. ■

Nowhere Differentiable Paths

The property illustrated in Example 8.5 shows that Brownian motion preserves its character after rescaling. For instance, given a standard Brownian motion on $[0, 1]$, if we look at the process on, say, an interval of length one-trillionth ($= 10^{-12}$) then after resizing by a factor of $1/\sqrt{10^{-12}} = 1{,}000{,}000$, what we see is indistinguishable from the original Brownian motion!

This highlights the invariance, or *fractal*, structure of Brownian motion sample paths. It means that the jagged character of these paths remains jagged at all time scales. This leads to the remarkable fact that Brownian motion sample paths are *nowhere differentiable*. It is hard to even contemplate a function that is continuous at every point on its domain, but not differentiable at any point. Indeed, for many years, mathematicians believed that such a function was impossible, until Karl Weierstrass, considered the founder of modern analysis, demonstrated their existence in 1872.

The proof that Brownian motion is nowhere differentiable requires advanced analysis. Here is a heuristic argument. Consider the formal derivative

$$\frac{d}{dt}B_t = \lim_{h\to 0} \frac{B_{t+h} - B_t}{h}.$$

By stationary increments, $B_{t+h} - B_t$ has the same distribution as B_h, which is normal with mean 0 and variance h. Thus, the difference quotient $(B_{t+h} - B_t)/h$ is normally distributed with mean 0 and variance $1/h$. As h tends to 0, the variance tends to infinity. Since the difference quotient takes arbitrarily large values, the limit, and hence the derivative, does not exist.

8.4 TRANSFORMATIONS AND PROPERTIES

In addition to invariance under rescaling, Brownian motion satisfies numerous reflection, translation, and symmetry properties.

Transformations of Brownian Motion

Let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Then, each of the following transformations is a standard Brownian motion.

1. *(Reflection)* $(-B_t)_{t\geq 0}$.
2. *(Translation)* $(B_{t+s} - B_s)_{t\geq 0}$, for all $s \geq 0$.
3. *(Rescaling)* $(a^{-1/2}B_{at})_{t\geq 0,}$ for all $a > 0$.
4. *(Inversion)* The process $(X_t)_{t\geq 0}$ defined by $X_0 = 0$ and $X_t = tB_{1/t}$, for $t > 0$.

We leave the proofs that reflection and translation are standard Brownian motions as exercises. Rescaling was shown in Example 8.5. For inversion, let $t_1 < \cdots < t_k$. For constants $a_1, \ldots, a_k$,

$$a_1X_{t_1} + \cdots + a_kX_{t_k} = a_1t_1B_{1/t_1} + \cdots + a_kt_kB_{1/t_k}$$

is normally distributed and thus $(X_t)_{t\geq 0}$ is a Gaussian process. The mean function is $E(X_t) = E(tB_{1/t}) = tE(B_{1/t}) = 0$, for $t > 0$. Covariance is

$$\begin{aligned}\text{Cov}(X_s, X_t) &= \text{Cov}(sB_{1/s}, tB_{1/t}) = E(sB_{1/s}tB_{1/t})\\ &= stE(B_{1/s}B_{1/t}) = st\,\text{Cov}(B_{1/s}, B_{1/t})\\ &= st\min\{1/s, 1/t\} = \min\{t, s\}.\end{aligned}$$

Continuity, for all $t > 0$, is inherited from $(B_t)_{t\geq 0}$. What remains is to show the process is continuous at $t = 0$. We do not prove this rigorously. Suffice it to show that

$$0 = \lim_{t\to 0} X_t = \lim_{t\to 0} tB_{1/t}, \text{ with probability 1}$$

is equivalent to

$$0 = \lim_{s\to\infty} \frac{B_s}{s}, \text{ with probability 1},$$

and the latter is intuitive by the strong law of large numbers.

For real x, the process defined by $X_t = x + B_t$, for $t \geq 0$, is called *Brownian motion started at* x. For such a process, $X_0 = x$ and X_t is normally distributed with mean function $E(X_t) = x$, for all t. The process retains all other defining properties of standard Brownian motion: stationary and independent increments, and continuous sample paths.

See Figure 8.5 for examples of transformations of Brownian motion.

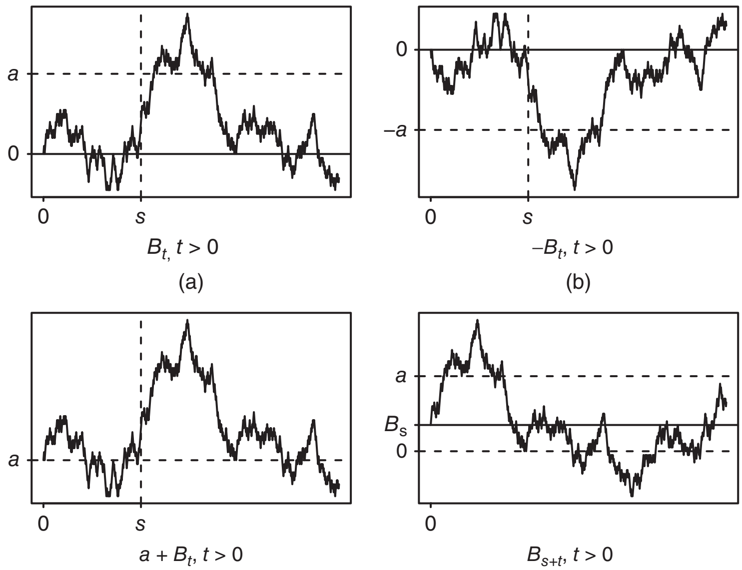

Figure 8.5 Transformations of Brownian motion. (a) Standard Brownian motion. (b) Reflection across t-axis. (c) Brownian motion started at a. (d) Translation.

Example 8.6 Let $(X_t)_{t\geq 0}$ be a Brownian motion process started at $x = 3$. Find $P(X_2 > 0)$.

Solution Write $X_t = B_t + 3$. Then,

$$P(X_2 > 0) = P(B_2 + 3 > 0) = P(B_2 > -3) = \int_{-3}^{\infty} \frac{1}{\sqrt{4\pi}} e^{-x^2/4}\, dx = 0.983.$$

In R, type

```
> 1-pnorm(-3,0,sqrt(2))
[1] 0.9830526
```

■

Markov Process

Brownian motion satisfies the Markov property that conditional on the present, past and future are independent. This is a consequence of independent increments. Brownian motion is an example of a *Markov process*. A continuous-state stochastic process $(X_t)_{t\geq 0}$ is a Markov process if

$$P(X_{t+s} \leq y | X_u, 0 \leq u \leq s) = P(X_{t+s} \leq y | X_s), \tag{8.1}$$

for all $s, t \geq 0$ and real y. The process is *time-homogeneous* if the probability in Equation (8.1) does not depend on s. That is,

$$P(X_{t+s} \leq y | X_s) = P(X_t \leq y | X_0).$$

For a continuous-state Markov process, the *transition function*, or *transition kernel*, $K_t(x, y)$ plays the role that the transition matrix plays for a discrete-state Markov chain. The function $K_t(x, \cdot)$ is the conditional density of X_t given $X_0 = x$.

If $(X_t)_{t\geq 0}$ is Brownian motion started at x, then X_t is normally distributed with mean x and variance t. The transition kernel is

$$K_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/2t}.$$

The transition kernel of a Markov process satisfies the Chapman–Kolmogorov equations. For continuous-state processes, integrals replace sums. The equations are

$$K_{s+t}(x, y) = \int_{-\infty}^{\infty} K_s(x, z) K_t(z, y)\, dz, \text{ for all } s, t, \tag{8.2}$$

as

$$\begin{aligned}
\int_{-\infty}^{y} K_{s+t}(x, w)\, dw &= P(X_{s+t} \leq y | X_0 = x) \\
&= \int_{-\infty}^{\infty} P(X_{s+t} \leq y | X_s = z, X_0 = x) K_s(x, z)\, dz \\
&= \int_{-\infty}^{\infty} P(X_t \leq y | X_0 = z) K_s(x, z)\, dz \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{y} K_t(z, w)\, dw \right) K_s(x, z)\, dz \\
&= \int_{-\infty}^{y} \left(\int_{-\infty}^{\infty} K_s(x, z) K_t(z, w)\, dz \right) dw.
\end{aligned}$$

Taking derivatives with respect to y gives Equation (8.2).

First Hitting Time and Strong Markov Property

Brownian motion satisfies the strong Markov property. Recall that Brownian motion translated left or right by a constant is also a Brownian motion. For $s > 0$, with $B_s = x$, the process $(B_{t+s})_{t \geq 0}$ is a Brownian motion started at x.

By the strong Markov property this also holds for some random times as well. If S is a *stopping time*, $(B_{t+S})_{t \geq 0}$ is a Brownian motion process. See Section 3.9 to reference the strong Markov property for discrete-time Markov chains.

A common setting is when Brownian motion first hits a particular state or set of states. Let $T_a = \min\{t : B_t = a\}$ be the *first hitting time* that Brownian motion hits level a. See Figure 8.6. The random variable T_a is a stopping time. Moving forward from T_a, the translated process is a Brownian motion started at a.

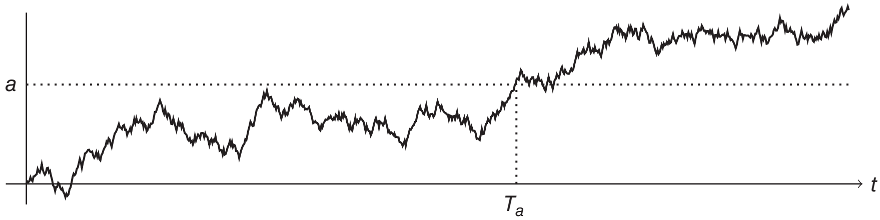

Figure 8.6 T_a is the first time Brownian motion hits level a.

The strong Markov property is used to find the distribution of T_a. Consider standard Brownian motion. At any time t, B_t is equally likely to be above or below the line $y = 0$. Assume that $a > 0$. For Brownian motion started at a, the process is equally likely to be above or below the line $y = a$. This gives,

$$P(B_t > a | T_a < t) = P(B_t > 0) = \frac{1}{2},$$

and thus,

$$\frac{1}{2} = P(B_t > a | T_a < t) = \frac{P(B_t > a, T_a < t)}{P(T_a < t)} = \frac{P(B_t > a)}{P(T_a < t)}.$$

The last equality is because the event $\{B_t > a\}$ implies $\{T_a < t\}$ by continuity of sample paths. We have that

$$\begin{aligned} P(T_a < t) &= 2P(B_t > a) && (8.3) \\ &= 2 \int_a^\infty \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \, dx \\ &= 2 \int_{a/\sqrt{t}}^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx. && (8.4) \end{aligned}$$

If $a < 0$, the argument is similar with $1/2 = P(B_t < a | T_a < t)$. In either case,

$$P(T_a < t) = 2\int_{|a|/\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx.$$

Differentiating with respect to t gives the probability density function of the first hitting time.

First Hitting Time Distribution

For a standard Brownian motion, let T_a be the first time the process hits level a. The density function of T_a is

$$f_{T_a}(t) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-a^2/2t}, \text{ for } t > 0. \tag{8.5}$$

Example 8.7 A particle moves according to Brownian motion started at $x = 1$. After $t = 3$ hours, the particle is at level 1.5. Find the probability that the particle reaches level 2 sometime in the next hour.

Solution For $t \geq 3$, the translated process is a Brownian motion started at $x = 1.5$. The event that the translated process reaches level 2 in the next hour, is equal to the event that a standard Brownian motion first hits level $a = 2 - 1.5 = 0.5$ in $[0, 1]$. The desired probability is

$$P(T_{0.5} \leq 1) = \int_0^1 \frac{0.5}{\sqrt{2\pi t^3}} e^{-(0.5)^2/2t}\, dt = 0.617.$$

■

The first hitting time distribution has some surprising properties. Consider

$$\begin{aligned} P(T_a < \infty) &= \lim_{t\to\infty} P(T_a < t) = \lim_{t\to\infty} 2\int_{|a|/\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \\ &= 2\int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = 2\left(\frac{1}{2}\right) = 1. \end{aligned} \tag{8.6}$$

Brownian motion hits level a, with probability 1, for all a, no matter how large.

On the contrary,

$$E(T_a) = \int_0^{\infty} \frac{t|a|}{\sqrt{2\pi t^3}} e^{-a^2/2t}\, dt = +\infty.$$

The expected time to first reach level a is infinite. This is true for all a, no matter how small!

Reflection Principle and the Maximum of Brownian Motion

Brownian motion reflected at a first hitting time is a Brownian motion. This property is known as the *reflection principle* and is a consequence of the strong Markov property.

For a standard Brownian motion, and first hitting time T_a, consider the process $\left(\tilde{B}_t\right)_{t\geq 0}$ defined by

$$\tilde{B}_t = \begin{cases} B_t, & \text{if } 0 \leq t \leq T_a, \\ 2a - B_t, & \text{if } t \geq T_a. \end{cases}$$

This is called *Brownian motion reflected at* T_a. The construction is shown in Figure 8.7. The reflected portion $a - (B_t - a) = 2a - B_t$ is a Brownian motion process started at a by the strong Markov property and the fact that $-B_t$ is a standard Brownian motion. Concatenating the front of the original process $(B_t)_{0\leq t\leq T_a}$ to the reflected piece preserves continuity.

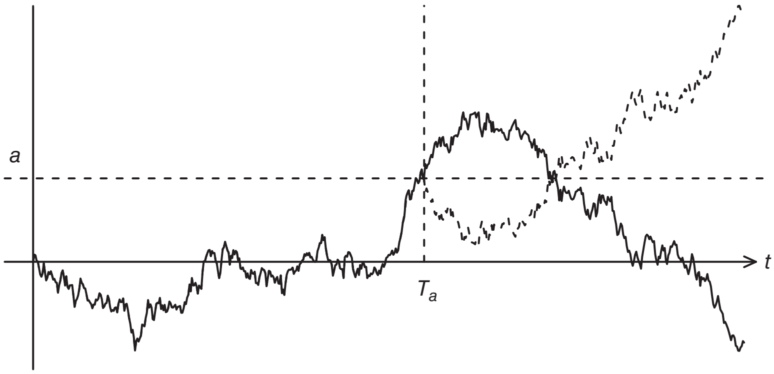

Figure 8.7 Reflection principle.

Another way of thinking of the reflection principle is that it establishes a one-to-one correspondence between paths that exceed level a at time t and paths that are below a at time t and have hit a by time t.

The reflection principle is applied in the following derivation of the distribution of $M_t = \max_{0\leq s\leq t} B_s$, the maximum value of Brownian motion on $[0, t]$.

Let $a > 0$. If at time t, B_t exceeds a, then the maximum on $[0, t]$ is greater than a. That is, $\{B_t > a\}$ implies $\{M_t > a\}$. This gives

$$\begin{aligned} \{M_t > a\} &= \{M_t > a, B_t > a\} \cup \{M_t > a, B_t \leq a\} \\ &= \{B_t > a\} \cup \{M_t > a, B_t \leq a\}. \end{aligned}$$

As the union is disjoint, $P(M_t > a) = P(B_t > a) + P(M_t > a, B_t \leq a)$.

Consider a sample path that has crossed a by time t and is at most a at time t. By the reflection principle, the path corresponds to a reflected path that is at least a at time t. This gives that $P(M_t > a, B_t \leq a) = P\left(\tilde{B}_t \geq a\right) = P(B_t > a)$. Thus,

$$P(M_t > a) = 2P(B_t > a) = \int_a^\infty \sqrt{\frac{2}{\pi t}} e^{-x^2/2t}\, dx, \text{ for } a > 0.$$

The distribution of M_t is equal to the distribution of $|B_t|$, the absolute value of a normally distributed random variable with mean 0 and variance t.

Since we have already found the distribution of T_a, a different derivation of the distribution of M_t is possible, using the fact that $M_t > a$ if and only if the process hits a by time t, that is $T_a < t$. This gives

$$P(M_t > a) = P(T_a < t) = \int_0^t \frac{a}{\sqrt{2\pi s^3}} e^{-a^2/2s}\, ds \tag{8.7}$$

$$= \int_a^\infty \sqrt{\frac{2}{\pi t}} e^{-x^2/2t}\, dx, \text{ for } a > 0. \tag{8.8}$$

The last equality is achieved by the change of variables $a^2/s = x^2/t$.

Example 8.8 A laboratory instrument takes annual temperature measurements. Measurement errors are assumed to be independent and normally distributed. As precision decreases over time, errors are modeled as standard Brownian motion. For how many years can the lab be guaranteed that there is at least 90% probability that all errors are less than 4 degrees?

Solution The problem asks for the largest t such that $P(M_t \leq 4) \geq 0.90$. We have

$$0.90 \leq P(M_t \leq 4) = 1 - P(M_t > 4) = 1 - 2P(B_t > 4) = 2P(B_t \leq 4) - 1.$$

This gives

$$0.95 \leq P(B_t \leq 4) = P\left(Z \leq \frac{4}{\sqrt{t}}\right),$$

where Z is a standard normal random variable. The 95th percentile of the standard normal distribution is 1.645. Solving $4/\sqrt{t} = 1.645$ gives

$$t = \left(\frac{4}{1.645}\right)^2 = 5.91 \text{ years.}$$

■

Zeros of Brownian Motion and Arcsine Distribution

Brownian motion reaches level x, no matter how large x, with probability 1. Brownian motion also returns to the origin infinitely often. In fact, on any interval $(0, \epsilon)$, no matter how small ϵ, the process crosses the t-axis infinitely many times.

The times when the process crosses the t-axis are the *zeros* of Brownian motion. That Brownian motion has infinitely many zeros in any interval of the form $(0, \epsilon)$ is a consequence of the following.

Zeros of Brownian Motion

Theorem 8.1. *For $0 \leq r < t$, let $z_{r,t}$ be the probability that standard Brownian motion has at least one zero in (r, t). Then,*

$$z_{r,t} = \frac{2}{\pi} \arccos\left(\sqrt{\frac{r}{t}}\right).$$

With $r = 0$, the result gives that standard Brownian motion has at least one zero in $(0, \epsilon)$ with probability

$$z_{0,\epsilon} = (2/\pi) \arccos(0) = (2/\pi)(\pi/2) = 1.$$

That is, $B_t = 0$, for some $0 < t < \epsilon$. By the strong Markov property, for Brownian motion restarted at t, there is at least one zero in (t, ϵ), with probability 1. Continuing in this way, there are infinitely many zeros in $(0, \epsilon)$.

The *arcsine distribution* arises in the proof of Theorem 8.1 and other results related to the zeros of Brownian motion. The distribution is a special case of the beta distribution, and has cumulative distribution function

$$F(t) = \frac{2}{\pi} \arcsin\left(\sqrt{t}\right), \text{ for } 0 \leq t \leq 1. \tag{8.9}$$

The arcsine density function is

$$f(t) = F'(t) = \frac{1}{\pi\sqrt{t(1-t)}}, \text{ for } 0 \leq t \leq 1.$$

The density is bimodal and symmetric, as shown in Figure 8.8.

Proof of Theorem 8.1. Conditioning on B_r,

$$\begin{aligned} z_{r,t} &= P(B_s = 0 \text{ for some } s \in (r,t)) \\ &= \int_{-\infty}^{\infty} P(B_s = 0 \text{ for some } s \in (r,t) | B_r = x) \frac{1}{\sqrt{2\pi r}} e^{-x^2/2r}\, dx. \end{aligned} \tag{8.10}$$

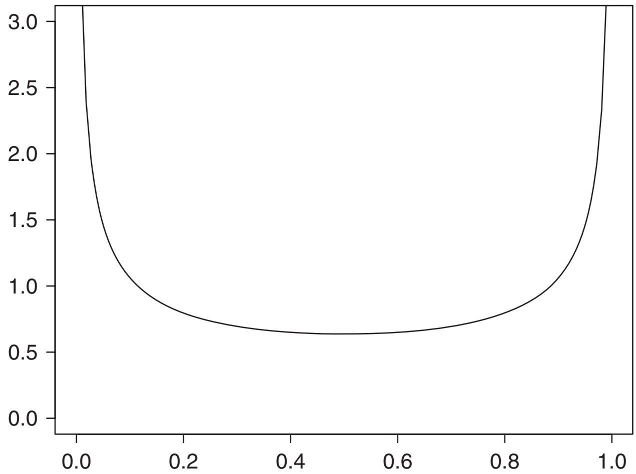

Figure 8.8 Arcsine distribution.

Assume that $B_r = x < 0$. The probability that $B_s = 0$ for some $s \in (r, t)$ is equal to the probability that for the process started in x, the maximum on $(0, t - r)$ is greater than 0. By translation, the latter is equal to the probability that for the process started in 0, the maximum on $(0, t - r)$ is greater than x. That is,

$$P(B_s = 0 \text{ for some } s \in (r, t)|B_r = x) = P(M_{t-r} > x).$$

For $x > 0$, consider the reflected process $-B_s$ started in $-x$. In either case, with Equations (8.7) and (8.10),

$$\begin{aligned}
z_{r,t} &= \int_{-\infty}^{\infty} P(M_{t-r} > |x|)\frac{1}{\sqrt{2\pi r}}e^{-x^2/2r}\,dx \\
&= \int_{-\infty}^{\infty}\int_{0}^{t-r} \frac{1}{\sqrt{2\pi s^3}}|x|e^{-x^2/2s}\,ds\frac{1}{\sqrt{2\pi r}}e^{-x^2/2r}\,dx \\
&= \frac{1}{\pi}\int_0^{t-r}\frac{1}{\sqrt{rs^3}}\int_0^{\infty} xe^{-x^2(r+s)/2rs}\,dx\,ds \\
&= \frac{1}{\pi}\int_0^{t-r}\frac{1}{\sqrt{rs^3}}\int_0^{\infty} e^{-z(r+s)/rs}\,dz\,ds \\
&= \frac{1}{\pi}\int_0^{t-r}\frac{1}{\sqrt{rs^3}}\left(\frac{rs}{r+s}\right)\,ds \\
&= \frac{1}{\pi}\int_{r/t}^{1}\frac{1}{\sqrt{x(1-x)}}\,dx.
\end{aligned}$$

The last equality is by the change of variables $r/(r+s) = x$. The last expression is an arcsine probability, which, by Equation (8.9), gives

$$z_{r,t} = \frac{2}{\pi}\left(\arcsin\left(\sqrt{1}\right) - \arcsin\left(\sqrt{\frac{r}{t}}\right)\right)$$
$$= 1 - \frac{2}{\pi}\arcsin\left(\sqrt{\frac{r}{t}}\right) = \frac{2}{\pi}\arccos\left(\sqrt{\frac{r}{t}}\right).$$

■

Last Zero Standing

Corollary 8.2. *Let L_t be the last zero in $(0, t)$. Then,*

$$P(L_t \le x) = \frac{2}{\pi}\arcsin\left(\sqrt{\frac{x}{t}}\right), \textit{ for } 0 < x < t.$$

Proof of corollary. The last zero occurs by time $x < t$ if and only if there are no zeros in (x, t). This occurs with probability

$$1 - z_{x,t} = 1 - \frac{2}{\pi}\arccos\left(\sqrt{\frac{x}{t}}\right) = \frac{2}{\pi}\arcsin\left(\sqrt{\frac{x}{t}}\right).$$

■

Example 8.9 (Fluctuations in Coin Tossing)

> We shall encounter theoretical conclusions which not only are unexpected but actually come as a shock to intuition and common sense. They will reveal that commonly accepted notions concerning chance fluctuations are without foundation and that the implications of the law of large numbers are widely misconstrued.
>
> —William Feller

Feller, the author of the quotation and of the classic and ground-breaking probability textbook *An Introduction to Probability Theory and Its Applications*, was discussing fluctuations in coin tossing and random walk. As a discrete process random walk is often studied with counting and combinatorial tools. Because of the connection between random walk and Brownian motion many discrete results can be obtained by taking suitable limits and invoking the invariance principle.

Consider a fair coin-tossing game between two players, A and B. If the coin lands heads, A pays B one dollar. If tails, B pays A. The coin is flipped $n = 10,000$ times. To test your "intuition and common sense," when would you guess is the *last* time the players are even? That is, at what point will the remaining duration of the game see one player always ahead?

Perhaps you think that in an evenly matched game there will be frequent changes in the lead and thus the last time the players are even is likely to be close to n, near the end of the game?

Let $\tilde{L}_n$ be the last time, in n plays, that the players are tied. This is the last zero for simple symmetric random walk on $\{0, 1, \ldots, n\}$. Before continuing, we invite the reader to sketch their guesstimate of the distribution of $\tilde{L}_n$.

We simulated the coin-flipping game 5,000 times, generating the histogram of $\tilde{L}_n$ in Figure 8.9. The distribution is symmetric. It is equally likely that the last zero of the random walk is either k or $n - k$. Furthermore, the probabilities near the ends are the greatest. There is a surprisingly large probability of one player gaining the lead early in the game, and keeping the lead throughout. After just 20% of the game, there is almost 30% probability that one player will be in the lead for the remainder.

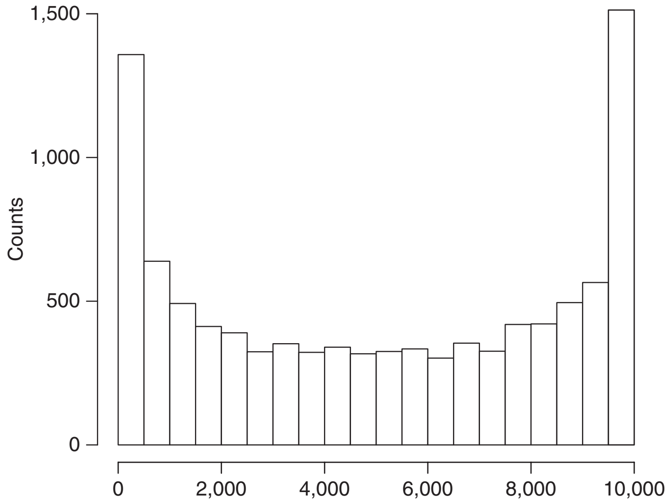

Figure 8.9 Last time players are tied in 10,000 coin flips.

It is no accident, of course, that the histogram bears a striking resemblance to the arcsine density curve in Figure 8.8. Let $0 < t < 1$. The probability that the last zero of the random walk occurs by step tn, that is, after $100t$ percent of the walk, is approximately the probability that the last zero of Brownian motion on $[0, 1]$ occurs by time t. For large n,

$$P\left(\tilde{L}_n \leq tn\right) \approx P(L_1 \leq t) = \frac{2}{\pi} \arcsin\left(\sqrt{t}\right).$$

Simulated probabilities for the random walk and theoretical values for Brownian motion are compared in Table 8.1.

TABLE 8.1 Random Walk and Brownian Motion Probabilities for the Last Zero ($n = 10{,}000$)

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$P\left(\tilde{L}_n \leq tn\right)$	0.207	0.297	0.367	0.434	0.499	0.565	0.632	0.704	0.795
$P(L_1 \leq t)$	0.205	0.295	0.369	0.436	0.500	0.564	0.631	0.705	0.795

R: Random Walk and Coin Tossing

```
# coinflips.R
trials <- 10000
simlist <- numeric(trials)
for (i in 1:trials) {
 rw <- c(0,cumsum(sample(c(-1,1),(trials-1),
 replace=T)))
 simlist[i] <- tail(which(rw==0),1)
  }
mean(simlist)
hist(simlist,xlab="",ylab="Counts",main="")
```

■

8.5 VARIATIONS AND APPLICATIONS

Standard Brownian motion is often too simple a model for real-life applications. Many variations arise in practice. Brownian motion started in x has a constant mean function. A common variant of Brownian motion has linear mean function as well as an additional variance parameter.

Brownian Motion with Drift

For real μ and $\sigma > 0$, the process defined by

$$X_t = \mu t + \sigma B_t, \text{ for } t \geq 0,$$

is called *Brownian motion with drift parameter μ and variance parameter σ^2*.

Brownian motion with drift is a Gaussian process with continuous sample paths and independent and stationary increments. For $s, t > 0$, $X_{t+s} - X_t$ is normally distributed with mean μs and variance $\sigma^2 s$.

Example 8.10 Find the probability that Brownian motion with drift parameter $\mu = 0.6$ and variance $\sigma^2 = 0.25$ takes values between 1 and 3 at time $t = 4$.

Solution Write $X_t = (0.6)t + (0.5)B_t$. The desired probability is

$$P(1 \leq X_4 \leq 3) = P(1 \leq (0.6)4 + (0.5)B_4 \leq 3) = P(-2.8 \leq B_4 \leq 1.2)$$
$$= \int_{-2.8}^{1.2} \frac{1}{\sqrt{8\pi}} e^{-x^2/8}\, dx = 0.645.$$

■

Example 8.11 (Home team advantage) A novel application of Brownian motion to sports scores is given in Stern (1994). The goal is to quantify the home team advantage by finding the probability in a sports match that the home team wins the game given that they lead by l points after a fraction $0 \le t \le 1$ of the game is completed. The model is applied to basketball where scores can be reasonably approximated by a continuous distribution.

For $0 \le t \le 1$, let X_t denote the difference in scores between the home and visiting teams after $100t$ percent of the game has been completed. The process is modeled as a Brownian motion with drift, where the mean parameter μ is a measure of home team advantage. The probability that the home team wins the game, given that they have an l point lead at time $t < 1$, is

$$\begin{aligned} p(l,t) &= P(X_1 > 0 | X_t = l) = P(X_1 - X_t > -l) \\ &= P(X_{1-t} > -l) = P(\mu(1-t) + \sigma B_{1-t} > -l) \\ &= P\left(B_{1-t} < \frac{l + \mu(1-t)}{\sigma}\right) \\ &= P\left(B_t < \frac{\sqrt{t}[l + \mu(1-t)]}{\sigma\sqrt{1-t}}\right). \end{aligned}$$

The last equality is because B_t has the same distribution as $\sqrt{t/(1-t)}B_{1-t}$.

The model is applied to the results of 493 National Basketball Association games in 1992. Drift and variance parameters are fit from the available data with estimates $\hat{\mu} = 4.87$ and $\hat{\sigma} = 15.82$.

Table 8.2 gives the probability of a home team win for several values of l and t. Due to the home court advantage, the home team has a greater than 50% chance of winning even if it is behind by two points at halftime ($t = 0.50$). Even in the last

TABLE 8.2 Table for Basketball Data Probabilities $p(l,t)$ that the Home Team Wins the Game Given that they are in the Lead by l Points After a Fraction t of the Game is Completed

	Lead						
Time t	$l = -10$	$l = -5$	$l = -2$	$l = 0$	$l = 2$	$l = 5$	$l = 10$
0.00				0.62			
0.25	0.32	0.46	0.55	0.61	0.66	0.74	0.84
0.50	0.25	0.41	0.52	0.59	0.65	0.75	0.87
0.75	0.13	0.32	0.46	0.56	0.66	0.78	0.92
0.90	0.03	0.18	0.38	0.54	0.69	0.86	0.98
1.00	0.00	0.00	0.00		1.00	1.0	1.0

Source: Stern (1994).

five minutes of play ($t = 0.90$), home team comebacks from five points are not that unusual, according to the model, with probability 0.18.

We recommend this paper to the mathematically inclined sports fan. It is both accessible and readable. The author discusses model assumptions and limitations, the extent to which theoretical predictions follow empirical results, and an interesting extension from basketball to baseball. ■

Brownian Bridge

The two ends of a bridge are both secured to level ground. A Brownian bridge is a Brownian motion process conditioned so that the process ends at the same level as where it begins.

> **Brownian Bridge**
>
> From standard Brownian motion, the conditional process $(B_t)_{0 \le t \le 1}$ given that $B_1 = 0$ is called a *Brownian bridge*. The Brownian bridge is *tied down* to 0 at the endpoints of $[0, 1]$.

Examples of Brownian bridge are shown in Figure 8.10. Let $(X_t)_{t \ge 0}$ denote a Brownian bridge. For $0 \le t \le 1$, the distribution of X_t is equal to the conditional distribution of B_t given $B_1 = 0$. Since the conditional distributions of a Gaussian process are Gaussian, it follows that Brownian bridge is a Gaussian process. Continuity of sample paths, and independent and stationary increments are inherited from standard Brownian motion.

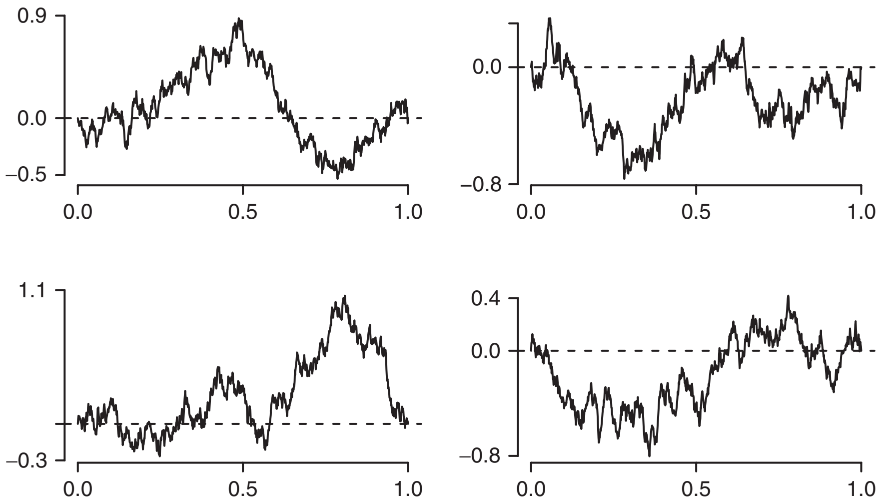

Figure 8.10 Brownian bridge sample paths.

To find the mean and covariance functions results are needed for bivariate normal distributions. We encourage the reader to work through Exercise 8.5(b) and show that for $0 < s < t$,

$$E(B_s|B_t = y) = \frac{sy}{t} \quad \text{and} \quad Var(B_s|B_t = y) = \frac{s(t-s)}{t}.$$

This gives the mean function of Brownian bridge

$$E(X_t) = E(B_t|B_1 = 0) = 0, \text{ for } 0 \le t \le 1.$$

For the covariance, $\text{Cov}(X_s, X_t) = E(X_sX_t)$. By the law of total expectation,

$$\begin{aligned} E(X_sX_t) &= E(E(X_sX_t)|X_t) = E(X_tE(X_s|X_t)) \\ &= E\left(X_t\frac{sX_t}{t}\right) = \frac{s}{t}E\left(X_t^2\right) = \frac{s}{t}E\left(B_t^2|B_1 = 0\right) \\ &= \frac{s}{t}Var(B_t|B_1 = 0) = \left(\frac{s}{t}\right)\frac{t(1-t)}{1} = s - st. \end{aligned}$$

By symmetry, for $t < s$, $E(X_sX_t) = t - st$. In either case, the covariance function is

$$\text{Cov}(X_s, X_t) = \min\{s, t\} - st.$$

Example 8.12 Let $X_t = B_t - tB_1$, for $0 \le t \le 1$. Show that $(X_t)_{0\le t\le 1}$ is a Brownian bridge.

Solution The process is a Gaussian process since $(B_t)_{t\ge 0}$ is a Gaussian process. Sample paths are continuous, with probability 1. It is suffice to show that the process has the same mean and covariance functions as a Brownian bridge.

The mean function is $E(X_t) = E(B_t - tB_1) = E(B_t) - tE(B_1) = 0$. The covariance function is

$$\begin{aligned} \text{Cov}(X_s, X_t) &= E(X_sX_t) = E((B_s - sB_1)(B_t - tB_1)) \\ &= E(B_sB_t) - tE(B_sB_1) - sE(B_tB_1) + stE(B_1^2) \\ &= \min\{s, t\} - ts - st + st = \min\{s, t\} - st, \end{aligned}$$

which is the covariance function of Brownian bridge.

The construction described in this example gives a direct method for simulating a Brownian bridge used to draw the graphs in Figure 8.10.

R: Simulating Brownian Bridge

```
# bbridge.R
> n <- 1000
> t <- seq(0,1,length=n)
```

```
> bm <- c(0,cumsum(rnorm(n-1,0,1)))/sqrt(n)
> bb <- bm - t*bm[n]
> plot(t,bb,type="l")
```

■

Example 8.13 (Animal tracking) In *Analyzing Animal Movements Using Brownian Bridges*, ecologists Horne et al. (2007) develop a Brownian bridge model for estimating the expected movement path of an animal. The model is based on a two-dimensional Brownian motion, where $Z_t^{\boldsymbol{a},\boldsymbol{b}} = Z_t$ is defined to be the position in $\mathbb{R}^2$ of an animal at time $t \in [0, T]$, which starts at **a** and ends at **b**. Each Z_t is normally distributed with mean vector $E(Z_t) = \boldsymbol{a} + \frac{t}{T}(\boldsymbol{b} - \boldsymbol{a})$ and covariance matrix

$$\sigma_t^2 = \frac{t(T-t)}{T}\sigma_m^2 \boldsymbol{I},$$

where $\boldsymbol{I}$ is the identity matrix and σ_m^2 is a parameter related to the mobility of the animal. The probability that the animal is in region A at time t is $P(Z_t \in A)$. The model is applied to animal location data, often obtained through global positioning system telemetry, which allows for monitoring animal movements over great distances.

An objective of the researchers is to estimate the frequency of use of a region over the time of observation. Let $I_A(x)$ be the usual indicator function, which takes the value 1, if $x \in A$, and 0, otherwise. The *occupation time* for region A is defined as the random variable

$$\int_0^T I_A(Z_t)\, dt.$$

The expected fraction of time an animal occupies A is then

$$E\left(\frac{1}{T}\int_0^T I_A(Z_t)\, dt\right) = \frac{1}{T}\int_0^T P(Z_t \in A)\, dt,$$

a quantity used to estimate the home range of a male black bear in northern Idaho and the fall migration route of 11 caribou in Alaska.

The authors argue that the Brownian bridge movement model (BBMM) has the advantage over other methods in that BBMM assumes successive animal locations are not independent and explicitly incorporates the time between locations into the model. ■

Example 8.14 (Kolmogorov–Smirnov statistic) Following are 40 measurements, which take values between 0 and 1. We would like to test the claim that they are an i.i.d. sample from the uniform distribution on (0, 1). The Brownian bridge arises in the analysis of the *Kolmogorov–Smirnov test*, a common statistical method to test such claims.

0.100	0.296	0.212	0.385	0.993	0.870	0.070	0.815	0.123	0.588
0.332	0.035	0.758	0.362	0.453	0.047	0.134	0.389	0.147	0.424
0.060	0.003	0.800	0.011	0.085	0.674	0.196	0.715	0.342	0.519
0.675	0.799	0.768	0.721	0.315	0.009	0.109	0.835	0.044	0.152

Given a sample $X_1, \ldots, X_n$, define the *empirical distribution function*

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le t\}},$$

where $I_{\{X_i \le t\}}$ is the indicator function equal to 1, if $X_i \le t$, and 0, otherwise. The empirical distribution function gives the proportion of values in the sample that are at most t. If the data are a sample from a population with cumulative distribution function F, then $F_n(t)$ is an estimate of $P(X_i \le t) = F(t)$.

If the data in our example is an i.i.d. sample from the uniform distribution on $(0, 1)$, then $F(t) = t$, for $0 \le t \le 1$, and we would expect $F_n(t) \approx t$. Figure 8.11 shows the empirical distribution function for these data plotted alongside the line $y = t$.

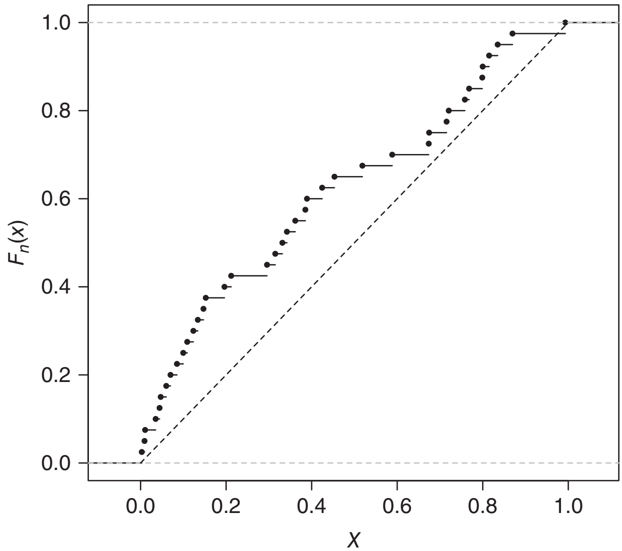

Figure 8.11 Empirical distribution function for sample data.

For a given cdf F, the *Kolmogorov–Smirnov statistic* is

$$D_n = \max_{0 \le t \le 1} |F_n(t) - F(t)|,$$

the maximum absolute distance between the empirical cdf and F. If the data are a sample from F, we expect $F_n(t) \approx F(t)$ and the value of D_n to be close to 0. Furthermore, large values of D_n are evidence against the hypothesis that the data are a sample from F. For the test of uniformity, $F(t) = t$, and

$$D_n = \max_{0 \le t \le 1} |F_n(t) - t|.$$

For our data, the Kolmogorov–Smirnov test statistic is $D_{40} = 0.223$. Is this large or small? Does it support or contradict the uniformity hypothesis? A statistician would ask: if the data do in fact come from a uniform distribution, what is the probability that D_{40} would be as large as 0.223? The probability $P(D_{40} > 0.223)$ is the *P-value* of the test. A small P-value means that it is unusual for D_{40} to be as large as 0.223, which would be evidence against uniformity. The distribution of D_n is difficult to obtain, which leads one to look for a good approximation.

If $X_1, \dots, X_n$ is an i.i.d. sample from the uniform distribution on $(0, 1)$, then for $0 < t < 1$,

$$F_n(t) - t = \sum_{i=1}^{n} \frac{I_{\{X_i \le t\}} - t}{n}$$

is a sum of i.i.d. random variables with common mean

$$E\left(\frac{I_{\{X_i \le t\}} - t}{n}\right) = \frac{P(X_i \le t) - t}{n} = \frac{t - t}{n} = 0$$

and variance

$$Var\left(\frac{I_{\{X_i \le t\}} - t}{n}\right) = \frac{1}{n^2} Var(I_{\{X_i \le t\}}) = \frac{P(X_i \le t)(1 - P(X_i \le t))}{n^2} = \frac{t(1-t)}{n^2}.$$

Thus, $F_n(t) - t$ has mean 0 and variance $t(1-t)/n$. For fixed $0 < t < 1$, the central limit theorem gives that $\sqrt{n}(F_n(t) - t)/\sqrt{t(1-t)}$ converges to a standard normal random variable, as $n \to \infty$. That is, for all real x,

$$\lim_{n \to \infty} P\left(\sqrt{n}(F_n(t) - t) \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi t(1-t)}} e^{-z^2/(2t(1-t))}\, dz = P(Y \le x),$$

where $Y \sim \text{Normal}(0, t(1-t))$.

By Donsker's invariance principle, it can be shown that the process $\sqrt{n}(F_n(t) - t)$, for $0 \le t \le 1$, converges to $B_t - tB_1$. The limiting process $(B_t - tB_1)_{0 \le t \le 1}$ is a Brownian bridge. The invariance principle further gives that

$$\sqrt{n} D_n = \sqrt{n} \max_{0 \le t \le 1} |F_n(t) - t|$$

converges to the maximum of a Brownian bridge. For large n, the distribution of D_n is approximately that of $M/\sqrt{n}$, where M is the maximum of a Brownian bridge.

We simulated the maximum of a Brownian bridge to find the P-value for our data, which is found to be

$$P\left(D_{40} > 0.223\right) \approx P(M/\sqrt{40} > 0.223) = P(M > 1.41) \approx 0.018.$$

The exact distribution of the maximum of a Brownian bridge is known. By using that distribution, the P-value is 0.0157. The P-value for the Kolmogorov–Smirnov test is obtained in R with the command `ks.test`.

The interpretation is that if the data were in fact uniformly distributed then the probability that D_{40} would be as large as 0.223 is less than 2%. Since the P-value is so small, we reject the claim and conclude that the data do not originate from a uniform distribution.

R: Test for Uniformity: Finding the *P*-value

```
# kstest.R
> trials <- 10000
> n <- 1000
> simlist <- numeric(trials)
> for (i in 1:trials) {
+ t <- seq(0,1,length=n)
# Brownian motion
+ bm <- c(0,cumsum(rnorm(n-1,0,1)))/sqrt(n)
+ bb <- bm-t*bm[n] # Brownian bridge
+ z <- max(bb) # maximum of Brownian bridge on [0,1]
+ simlist[i] <- if (z > 0.223*sqrt(40)) 1 else 0
}
> mean(simlist) # P-value = P(Z>1.41)
[1] 0.018
> ks.test(data,"punif",0,1)$p. value # exact P-value
[1] 0.015743
```

■

Geometric Brownian Motion

Geometric Brownian motion is a nonnegative process, which can be thought of as a stochastic model for exponential growth or decay. It is a favorite tool in mathematical finance, where it is used extensively to model stock prices.

Geometric Brownian Motion

Let $(X_t)_{t\geq 0}$ be a Brownian motion with drift parameter μ and variance parameter σ^2. The process $(G_t)_{t\geq 0}$ defined by

$$G_t = G_0 e^{X_t}, \text{ for } t \geq 0,$$

where $G_0 > 0$, is called *geometric Brownian motion*.

Taking logarithms, we see that $\ln G_t = \ln G_0 + X_t$ is normally distributed with mean

$$E(\ln G_t) = E(\ln G_0 + X_t) = \ln G_0 + \mu t$$

and variance

$$Var(\ln G_t) = Var(\ln G_0 + X_t) = Var(X_t) = \sigma^2 t.$$

A random variable whose logarithm is normally distributed is said to have a *lognormal distribution*. For each $t > 0$, G_t has a lognormal distribution.

We leave as an exercise the derivation of mean and variance for geometric Brownian motion

$$E(G_t) = G_0 e^{t(\mu+\sigma^2/2)} \text{ and } \quad Var(G_t) = G_0^2 e^{2t(\mu+\sigma^2/2)}(e^{t\sigma^2} - 1). \tag{8.11}$$

The exponential mean function shows that, on average, geometric Brownian motion exhibits exponential growth with growth rate $\mu + \sigma^2/2$.

Geometric Brownian motion arises as a model for quantities which can be expressed as the *product* of independent random multipliers. For $s, t \geq 0$, consider the ratio

$$\frac{G_{t+s}}{G_t} = \frac{G_0 e^{\mu(t+s)+\sigma X_{t+s}}}{G_0 e^{\mu t+\sigma X_t}} = e^{\mu s+\sigma(X_{t+s}-X_t)},$$

which has the same distribution as $e^{\mu s+\sigma X_s} = G_s/G_0$, because of stationary increments for the $(X_t)_{t\geq 0}$ process. For $0 \leq q < r \leq s < t$,

$$\frac{G_t}{G_s} = e^{\mu(t-s)+\sigma(X_t-X_s)} \text{ and } \quad \frac{G_r}{G_q} = e^{\mu(r-q)+\sigma(X_r-X_q)}$$

are independent random variables, because of independent increments for $(X_t)_{t\geq 0}$.

Let $Y_k = G_k/G_{k-1}$, for $k = 1, 2, \ldots$ Then, $Y_1, Y_2, \ldots$ is an i.i.d. sequence, and

$$G_n = \left(\frac{G_n}{G_{n-1}}\right)\left(\frac{G_{n-1}}{G_{n-2}}\right)\cdots\left(\frac{G_2}{G_1}\right)\left(\frac{G_1}{G_0}\right) G_0 = G_0 Y_1 Y_2 \cdots Y_{n-1} Y_n.$$

Example 8.15 Stock prices are commonly modeled with geometric Brownian motion. The process is attractive to economists because of several assumptions.

Historical data for many stocks indicate long-term exponential growth or decline. Prices cannot be negative and geometric Brownian motion takes only positive values. Let Y_t denote the price of a stock after t days. Since the price on a given day is probably close to the price on the next day (assuming normal market conditions), stock prices are not independent. However, the *percent changes in price* from day to day Y_t/Y_{t-1}, for $t = 1, 2, \ldots$ might be reasonably modeled as independent and identically distributed. This leads to geometric Brownian motion. In the context of stock prices, the standard deviation σ is called the *volatility*.

A criticism of the geometric Brownian motion model is that it does not account for extreme events like the stock market crash on October 19, 1987, when the world's stock markets lost more than 20% of their value within a few hours.

Assume that XYZ stock currently sells for \$80 a share and follows a geometric Brownian motion with drift parameter 0.10 and volatility 0.50. Find the probability that in 90 days the price of XYZ will rise to at least \$100.

Solution Let Y_t denote the price of XYZ after t years. Round 90 days as 1/4 of a year. Then,

$$\begin{aligned} P(Y_{0.25} \geq 100) &= P\left(80e^{\mu(0.25)+\sigma B_{0.25}} \geq 100\right) \\ &= P((0.1)(0.25) + (0.5)B_{0.25} \geq \ln 1.25) \\ &= P(B_{0.25} \geq 0.396) = 0.214. \end{aligned}$$

In R, type

```
> x <- (log(100/80)-(0.1)/4)/0.5
> 1-pnorm(x,0,sqrt(1/4))
[1] 0.214013
```

Simulations of the stock price are shown in Figure 8.12. ■

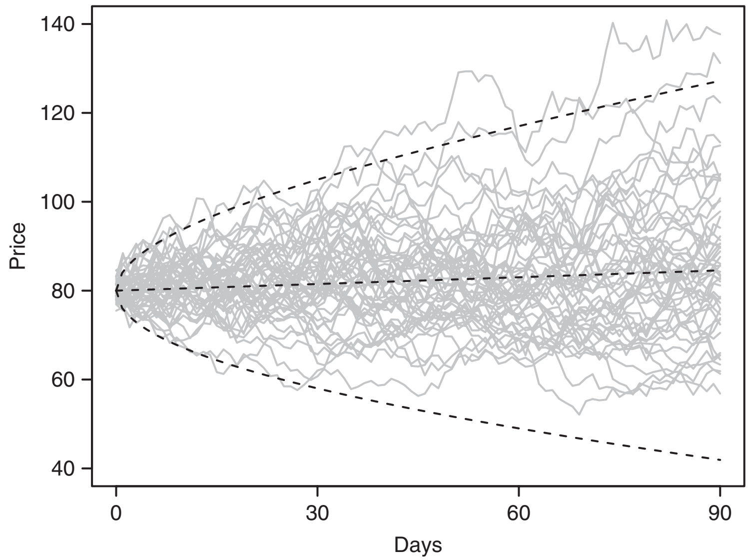

Figure 8.12 Fifty sample paths of a stock's price over 90 days modeled as geometric Brownian motion. Dotted lines are drawn at the mean function, and the mean plus or minus two standard deviations.

Example 8.16 (Financial options) An *option* is a contract that gives the buyer the right to buy shares of stock sometime in the future at a fixed price. In Example 8.13, we assumed that XYZ stock currently sells for \$80 a share. Assume that an XYZ option is selling for \$10. Under the terms of the option, in 90 days you may buy a share of XYZ stock for \$100.

If you decide to purchase the option, consider the payoff. Assume that in 90 days the price of XYZ is greater than \$100. Then, you can *exercise* the option, buy the stock for \$100, and turn around and sell XYZ for its current price. Your payoff would be $G_{90/365} - 100$, where $G_{90/365}$ is the price of XYZ in 90 days.

On the other hand, if XYZ sells for less than \$100 in 90 days, you would not exercise the option, and your payoff is nil. In either case, the payoff in 90 days is $\max\{G_{90/365} - 100, 0\}$. Your final profit would be the payoff minus the initial \$10 cost of the option.

Find the future expected payoff of the option, assuming the price of XYZ follows a geometric Brownian motion.

Solution Let G_0 denote the current stock price. Let t be the *expiration date*, which is the time until the option is exercised. Let K be the *strike price*, which is how much you can buy the stock for if you exercise the option. For XYZ, $G_0 = 80$, $t = 90/365$ (measuring time in years), and $K = 100$.

The goal is to find the expected payoff $E(\max\{G_t - K, 0\})$, assuming $(G_t)_{t\geq 0}$ is a geometric Brownian motion. Let $f(x)$ be the density function of a normal distribution with mean 0 and variance t. Then,

$$
\begin{aligned}
E(\max\{G_t - K, 0\}) &= E\left(\max\{G_0 e^{\mu t+\sigma B_t} - K, 0\}\right) \\
&= \int_{-\infty}^{\infty} \max\{G_0 e^{\mu t+\sigma x} - K, 0\} f(x)\, dx \\
&= \int_{\beta}^{\infty} (G_0 e^{\mu t+\sigma x} - K) f(x)\, dx \\
&= G_0 e^{\mu t} \int_{\beta}^{\infty} e^{\sigma x} f(x)\, dx - KP\left(Z > \frac{\beta}{\sqrt{t}}\right),
\end{aligned}
$$

where $\beta = \left(\ln\,(K/G_0) - \mu t\right)/\sigma$, and Z is a standard normal random variable.

By completing the square, the integral in the last expression is

$$
\begin{aligned}
\int_{\beta}^{\infty} e^{\sigma x} f(x)\, dx &= \int_{\beta}^{\infty} e^{\sigma x} \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}\, dx \\
&= e^{\sigma^2 t/2} \int_{\beta}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-(x-\sigma t)^2/2t}\, dx
\end{aligned}
$$

$$= e^{\sigma^2 t/2} \int_{(\beta-\sigma t)/\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$

$$= e^{\sigma^2 t/2} P\left(Z > \frac{\beta - \sigma t}{\sqrt{t}} \right).$$

This gives

$$E(\max\{G_t - K, 0\})$$
$$= G_0 e^{t(\mu+\sigma^2/2)} P\left(Z > \frac{\beta - \sigma t}{\sqrt{t}} \right) - KP\left(Z > \frac{\beta}{\sqrt{t}} \right). \tag{8.12}$$

Assume that XYZ stock follows a geometric Brownian motion with drift parameter $\mu = 0.10$ and variance $\sigma^2 = 0.25$. Then,

$$\beta = \frac{\ln\ (K/G_0) - \mu t}{\sigma} = \frac{\ln\ (100/80) - (0.10)(90/365)}{0.5} = 0.3970,$$

which gives

$$E(\max\{G_{90/365} - 100, 0\})$$
$$= 80e^{(90/365)(0.10+0.25/2)} P\left(Z > \frac{0.397 - 0.5(90/365)}{\sqrt{90/365}} \right)$$
$$- 100P\left(Z > \frac{0.397}{\sqrt{90/365}} \right)$$
$$= 1.788.$$

Given that the initial cost of the option is \$10, your expected profit is $1.788 - 10 < 0$. So you can expect to lose money.

For this example, we set an arbitrary initial cost of the option. A fundamental question in finance is how such an option should be priced. This leads to the Black–Scholes model for option pricing, which is introduced in the next section. ■

8.6 MARTINGALES

A martingale is a stochastic process that generalizes the notion of a *fair game*. Assume that after n plays of a gambling game your winnings are x. Then, by *fair*, we mean that your expected future winnings should be x regardless of past history.

Martingale

A stochastic process $(Y_t)_{t\geq 0}$ is a *martingale*, if for all $t \geq 0$,

1. $E(Y_t|Y_r, 0 \leq r \leq s) = Y_s$, for all $0 \leq s \leq t$.
2. $E(|Y_t|) < \infty$.

A discrete-time martingale $Y_0, Y_1, \ldots$ satisfies

1. $E(Y_{n+1}|Y_0, \ldots, Y_n) = Y_n$, for all $n \geq 0$.
2. $E(|Y_n|) < \infty$.

A most important property of martingales is that they have constant expectation. By the law of total expectation,

$$E(Y_t) = E(E(Y_t|Y_r, 0 \leq r \leq s)) = E(Y_s),$$

for all $0 \leq s \leq t$. That is,

$$E(Y_t) = E(Y_0), \text{ for all } t.$$

Example 8.17 (Random walk) Show that simple symmetric random walk is a martingale.

Solution Let

$$X_i = \begin{cases} +1, & \text{with probability } 1/2, \\ -1, & \text{with probability } 1/2, \end{cases}$$

for $i = 1, 2, \ldots$ For $n \geq 1$, let $S_n = X_1 + \cdots + X_n$, with $S_0 = 0$. Then,

$$\begin{aligned} E(S_{n+1}|S_0, \ldots, S_n) &= E(X_{n+1} + S_n|S_0, \ldots, S_n) \\ &= E(X_{n+1}|S_0, \ldots, S_n) + E(S_n|S_0, \ldots, S_n) \\ &= E(X_{n+1}) + S_n = S_n. \end{aligned}$$

The third equality is because X_{n+1} is independent of $X_1, \ldots, X_n$, and thus independent of $S_0, S_1, \ldots, S_n$. The fact that $E(S_n|S_0, \ldots, S_n) = S_n$ is a consequence of a general property of conditional expectation, which states that if X is a random variable and g is a function, then $E(g(X)|X) = g(X)$.

The second part of the martingale definition is satisfied as

$$E(|S_n|) = E(\left(\left|\sum_{i=1}^{n} X_i\right|\right) \leq E\left(\sum_{i=1}^{n} |X_i|\right) = \sum_{i=1}^{n} E(|X_i|) = n < \infty.$$

■

Since simple symmetric random walk is a martingale, the next example should not be surprising.

Example 8.18 (Brownian motion) Show that standard Brownian motion $(B_t)_{t\geq 0}$ is a martingale.

Solution We have that

$$\begin{aligned} E(B_t|B_r, 0 \leq r \leq s) &= E(B_t - B_s + B_s|B_r, 0 \leq r \leq s) \\ &= E(B_t - B_s|B_r, 0 \leq r \leq s) + E(B_s|B_r, 0 \leq r \leq s) \\ &= E(B_t - B_s) + B_s = B_s, \end{aligned}$$

where the second equality is because of independent increments. Also,

$$E(|B_t|) = \int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}\, dx = \int_0^{\infty} x\sqrt{\frac{2}{\pi t}} e^{-x^2/2t}\, dx = \sqrt{\frac{2t}{\pi}} < \infty.$$

■

The following extension of the martingale definition is used frequently.

Martingale with Respect to Another Process

Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be stochastic processes. Then, $(Y_t)_{t\geq 0}$ is a *martingale with respect to* $(X_t)_{t\geq 0}$, if for all $t \geq 0$,

1. $E(Y_t|X_r, 0 \leq r \leq s) = Y_s$, for all $0 \leq s \leq t$.
2. $E(|Y_t|) < \infty$.

The most common application of this is when Y_t is a function of X_t. That is, $Y_t = g(X_t)$ for some function g. It is useful to think of the conditioning random variables $(X_r)_{0\leq r\leq s}$ as representing past information, or history, of the process up to time s.

Following are several examples of martingales that are functions of Brownian motion.

Example 8.19 (Quadratic martingale) Let $Y_t = B_t^2 - t$, for $t \geq 0$. Show that $(Y_t)_{t\geq 0}$ is a martingale with respect to Brownian motion. This is called the *quadratic martingale*.

Solution For $0 \le s < t$,

$$\begin{aligned}
E(Y_t|B_r, 0 \le r \le s) &= E\Big(B_t^2 - t|B_r, 0 \le r \le s\Big) \\
&= E\left((B_t - B_s + B_s)^2|B_r, 0 \le r \le s\right) - t \\
&= E\left((B_t - B_s)^2 + 2(B_t - B_s)B_s + B_s^2|B_r, 0 \le r \le s\right) - t \\
&= E\left((B_t - B_s)^2\right) + 2B_sE(B_t - B_s) + B_s^2 - t \\
&= (t - s) + B_s^2 - t = B_s^2 - s = Y_s.
\end{aligned}$$

Furthermore,

$$E(|Y_t|) = E\Big(|B_t^2 - t|\Big) \le E\Big(B_t^2 + t\Big) = E\Big(B_t^2\Big) + t = 2t < \infty.$$

■

Example 8.20 Let $G_t = G_0e^{X_t}$ be a geometric Brownian motion, where $(X_t)_{t\ge0}$ is Brownian motion with drift μ and variance σ^2. Let $r = \mu + \sigma^2/2$. Show that $e^{-rt}G_t$ is a martingale with respect to standard Brownian motion.

Solution For $0 \le s < t$,

$$\begin{aligned}
E\left(e^{-rt}G_t|B_r, 0 \le r \le s\right) &= e^{-rt}E\left(G_0e^{\mu t+\sigma B_t}|B_r, 0 \le r \le s\right) \\
&= e^{-rt}E\left(G_0e^{\mu(t-s)+\sigma(B_t-B_s)}e^{\mu s+\sigma B_s}|B_r, 0 \le r \le s\right) \\
&= e^{-rt}e^{\mu s+\sigma B_s}E\left(G_0e^{\mu(t-s)+\sigma(B_t-B_s)}\right) \\
&= e^{-rt}e^{\mu s+\sigma B_s}E(G_{t-s}) \\
&= e^{-t(\mu+\sigma^2/2)}e^{\mu s+\sigma B_s}G_0e^{(t-s)(\mu+\sigma^2/2)} \\
&= e^{-s(\mu+\sigma^2/2)}G_0e^{\mu s+\sigma B_s} \\
&= e^{-rs}G_s.
\end{aligned}$$

Also,

$$E(|e^{-rt}G_t|) = e^{-rt}E(G_t) = e^{-t(\mu+\sigma^2/2)}G_0e^{t(\mu+\sigma^2/2)} = G_0 < \infty, \text{ for all } t.$$

■

Example 8.21 (Black–Scholes) In Example 8.16, the expected payoff for a financial option was derived. This leads to the *Black–Scholes formula* for pricing options, a fundamental formula in mathematical finance.

The formula was first published by Fisher Black and Myron Scholes in 1973 and then further developed by Robert Merton. Merton and Scholes received the 1997 Nobel Prize in Economics for their work. The ability to price options and other financial *derivatives* opened up a massive global market for trading ever-more complicated financial instruments. Black–Scholes has been described both as a formula which "changed the world" and as "the mathematical equation that caused the banks to crash." See Stewart (2012).

There are several critical assumptions underlying the Black–Scholes formula. One is that stock prices follow a geometric Brownian motion. Another is that an investment should be *risk neutral*. What this means is that the expected return on an investment should be equal to the so-called risk-free rate of return, such as what is obtained by a short-term U.S. government bond.

Let r denote the risk-free interest rate. Starting with P dollars, because of compound interest, after t years of investing risk free your money will grow to $P(1+r)^t$ dollars. Under continuous compounding, the future value is $F = Pe^{rt}$. This gives the future value of your present dollars. On the other hand, assume that t years from now you will be given F dollars. To find its *present value* requires *discounting* the future amount by a factor of e^{-rt}. That is, $P = e^{-rt}F$.

Let G_t denote the price of a stock t years from today. Then, the present value of the stock price is $e^{-rt}G_t$. The Black–Scholes risk neutral assumption means that the discounted stock price process is a *fair game*, that is, a martingale. For any time $0 < s < t$, the expected present value of the stock t years from now given knowledge of the stock price up until time s should be equal to the present value of the stock price s years from now. In other words,

$$E(e^{-rt}G_t|G_x, 0 \leq x \leq s) = e^{-rs}G_s. \tag{8.13}$$

In Example 8.20, it was shown that Equation (8.13) holds for geometric Brownian motion if $r = \mu + \sigma^2/2$, or $\mu = r - \sigma^2/2$. The probability calculations for the Black–Scholes formula are obtained with this choice of μ. In the language of Black–Scholes, this gives the *risk-neutral probability* for computing the options price formula. The Black–Scholes formula for the price of an option is then the present value of the expected payoff of the option under the risk-neutral probability.

See Equation (8.12) for the future expected payoff of a financial option. The present value of the expected payoff is obtained by multiplying by the discount factor e^{-rt}. Replace μ with $r - \sigma^2$ to obtain the Black–Scholes option price formula

$$e^{-rt}E(\max\{G_t - K, 0\}) = G_0 P\left(Z > \frac{\alpha - \sigma t}{\sqrt{t}}\right) - e^{-rt}KP\left(Z > \frac{\alpha}{\sqrt{t}}\right),$$

where

$$\alpha = \frac{\ln\ (K/G_0) - (r - \sigma^2/2)t}{\sigma}.$$

For the XYZ stock example, $G_0 = 80$, $K = 100$, $\sigma^2 = 0.25$, and $t = 90/365$. Furthermore, assume $r = 0.02$ is the risk-free interest rate. Then,

$$\alpha = \frac{\ln\ (100/80) - (0.02 - 0.25/2)(90/365)}{0.5} = 0.498068,$$

and the Black–Scholes option price is

$$80P\left(Z > \frac{\alpha - 0.50(90/365)}{0.5}\right) - e^{-0.02(90/365)}(100)P\left(Z > \frac{\alpha}{0.5}\right) = \$2.426.$$

Remarks:

1. For a given expiration date and strike price, the Black–Scholes formula depends only on volatility σ and the risk-free interest rate r, and not on the drift parameter μ. In practice, volatility is often estimated from historical price data. The model assumes that σ and r are constant.
2. The original derivation of the Black–Scholes formula in the paper by Black and Scholes was based on deriving and solving a partial differential equation

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} = rV,$$

   where $V = V(S, t)$ is the value of the option, a function of stock price S and time t. This is known as the *Black–Scholes equation*, and can be reduced to the simpler heat equation.
3. The Black–Scholes formula in the example is for a European call option, in which the option can only be exercised on the expiration date. In an American call option, one can exercise the option at any time before the expiration date. There are many types of options and financial derivatives whose pricing structure is based on Black–Scholes. ∎

Optional Stopping Theorem

A martingale $(Y_t)_{t\geq 0}$ has constant expectation. For all $t \geq 0$, $E(Y_t) = E(Y_0)$. The property holds for all fixed, deterministic times, but not necessarily for *random times*. If T is a random variable, which takes values in the index set of a martingale, it is not necessarily true that $E(Y_T) = E(Y_0)$. For instance, let T be the first time that a standard Brownian motion hits level a. Then, $B_T = a = E(B_T)$. However, $E(B_t) = 0$, for all $t \geq 0$.

The optional stopping theorem gives conditions for when a random time T satisfies $E(Y_T) = E(Y_0)$. While that might not sound like a big deal, the theorem is remarkably powerful. As mathematicians David Aldous and Jim Fill write in *Reversible Markov chains and random walks on graphs*, "Modern probabilists regard the martingale optional stopping theorem as one of the most important results in their subject." (Aldous and Fill, 2002.)

In Section 3.9, we introduced the notion of a stopping time in the context of discrete-time Markov chains. For a stochastic process $(Y_t)_{t\geq 0}$, a nonnegative random variable T is a stopping time if for each t, the event $\{T \leq t\}$ can be determined from $\{Y_s, 0 \leq s \leq t\}$. That is, if the outcomes of Y_s are known for $0 \leq s \leq t$, then it can be determined whether or not $\{T \leq t\}$ occurs.

On the interval [0, 1], the first time Brownian motion hits level a is a stopping time. Whether or not a is first hit by time t can be determined from $\{B_s, 0 \leq s \leq t.\}$ On the other hand, the *last time a* is hit on [0, 1] is not a stopping time. To determine whether or not a was last hit by time $t < 1$ requires full knowledge of B_s, for all $s \in [0, 1]$.

Optional Stopping Theorem

Let $(Y_t)_{t\geq 0}$ be a martingale with respect to a stochastic process $(X_t)_{t\geq 0}$. Assume that T is a stopping time for $(X_t)_{t\geq 0}$. Then, $E(Y_T) = E(Y_0)$ if one of the following is satisfied.

1. T is bounded. That is, $T \leq c$, for some constant c.
2. $P(T < \infty) = 1$ and $E(|Y_t|) \leq c$, for some constant c, whenever $T > t$.

The proof of the optional stopping theorem is beyond the scope of this book. There are several versions of the theorem with alternate sets of conditions.

In the context of a fair gambling game, one can interpret the optional stopping theorem to mean that a gambler has no reasonable strategy for increasing their initial stake. Let Y_t denote the gambler's winnings at time t. If a gambler strategizes to stop play at time T, their expected winnings will be $E(Y_T) = E(Y_0) = Y_0$, the gambler's initial stake.

The word *martingale* has its origins in an 18th century French gambling strategy. Mansuy (2009) quotes the dictionary of the Académie Française, that "To play the martingale is to always bet all that was lost."

Consider such a betting strategy for a game where at each round a gambler can win or lose one franc with equal probability. After one turn, if you win you stop. Otherwise, bet 2 francs on the next round. If you win, stop. If you lose, bet 4 francs on the next round. And so on. Let T be the number of bets needed until you win. The random variable T is a stopping time. If the gambler wins their game after k plays, they gain

$$2^k - (1 + 2 + \cdots + 2^{k-1}) = 2^k - (2^k - 1) = 1 \text{ franc,}$$

and, with probability 1, the gambler will eventually win some bet. Thus, T seems to be a winning strategy.

However, if Y_n denotes the gambler's winnings after n plays, then $E(Y_T) = 1 \neq 0 = E(Y_0)$. The random variable T does not satisfy the conditions of the optional sampling theorem. The reason this martingale gambling strategy is not *reasonable* is that it assumes the gambler has infinite capital.

The optional stopping theorem gives elegant and sometimes remarkably simple solutions to seemingly difficult problems.

Example 8.22 Let $a, b > 0$. For a standard Brownian motion, find the probability that the process hits level a before hitting level $-b$.

Solution Let p be the desired probability. Consider the time T that Brownian motion first hits either a or $-b$. That is, $T = \min\{t : B_t = a \text{ or } B_t = -b\}$. See Figure 8.13.

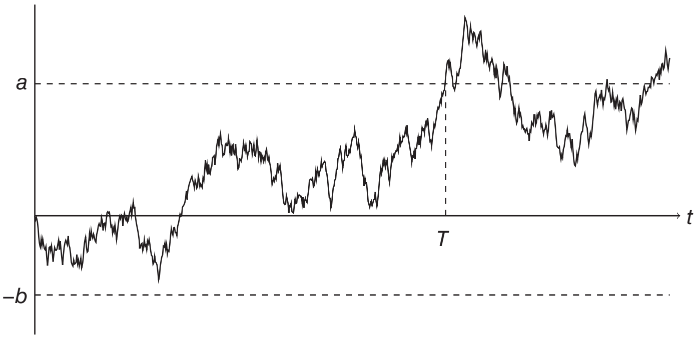

Figure 8.13 First hitting time T that Brownian motion hits level a or $-b$.

The random variable T is a stopping time. Furthermore, it satisfies the conditions of the optional stopping theorem. From Equation (8.6), the first hitting time T_a is finite with probability 1. By the strong Markov property, from a, the first time to hit $-b$ is also finite with probability 1. Thus, the first part of condition 2 is satisfied. Furthermore, for $t < T$, $B_t \in (-b, a)$. Thus, $|B_t| < \max\{a, b\}$, and the second part of condition 2 is satisfied.

Observe that $B_T = a$, with probability p, and $B_T = -b$, with probability $1 - p$. By the optional stopping theorem

$$0 = E(B_0) = E(B_T) = pa + (1 - p)(-b).$$

Solving for p gives $p = b/(a + b)$. ∎

Example 8.23 (Expected hitting time) Apply the optional stopping theorem with the same stopping time as in Example 8.22, but with the quadratic martingale $B_t^2 - t$. This gives

$$E\left(B_T^2 - T\right) = E\left(B_0^2 - 0\right) = 0,$$

from which it follows that

$$E(T) = E\left(B_T^2\right) = a^2\left(\frac{b}{a+b}\right) + b^2\left(\frac{a}{a+b}\right) = ab.$$

We have thus discovered that the expected time that standard Brownian motion first hits the boundary of the region defined by the lines $y = a$ and $y = -b$ is ab. ∎

Example 8.24 (Gambler's ruin) It was shown that discrete-time simple symmetric random walk $S_0, S_1, \ldots$ is a martingale. As with the quadratic martingale, the process $S_n^2 - n$ is a martingale. The results from the last two examples for Brownian motion can be restated for gambler's ruin on $\{-b, -b+1, \ldots, 0, \ldots, a-1, a\}$ starting at the origin. This gives the following:

1. The probability that the gambler gains a before losing b is $b/(a+b)$.
2. The expected duration of the game is ab. ■

Example 8.25 (Time to first hit the line $y = a - bt$) For $a, b > 0$, let $T = \min\{t : B_t = a - bt\}$ be the first time a standard Brownian motion hits the line $y = a - bt$. The random variable T is a stopping time, which satisfies the optional stopping theorem. This gives

$$0 = E(B_0) = E(B_T) = E(a - bT) = a - bE(T).$$

Hence, $E(T) = a/b$. For the line $y = 4 - (0.5)t$ in Figure 8.14, the mean time to first hit the line is $4/(0.5) = 8$. ■

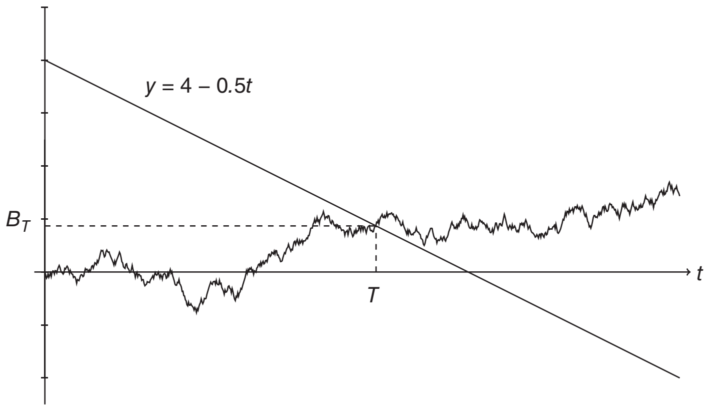

Figure 8.14 First time T Brownian motion hits the line $y = 4 - 0.5t$.

Example 8.26 (Time to reach a for Brownian motion with drift) Assume that $(X_t)_{t\geq 0}$ is a Brownian motion process with drift parameter μ and variance parameter σ^2, where $\mu > 0$. For $a > 0$, find the expected time that the process first hits level a.

Solution Let $T = \min\{t : X_t = a\}$ be the first time that the process hits level a. Write $X_t = \mu t + \sigma B_t$. Then, $X_t = a$ if and only if

$$B_t = \frac{a - \mu t}{\sigma} = \frac{a}{\sigma} - \left(\frac{\mu}{\sigma}\right) t.$$

Applying the result of Example 8.25, $E(T) = (a/\sigma)/(\mu/\sigma) = a/\mu$. ■

Example 8.27 (Variance of first hitting time) Assume that $(X_t)_{t\geq 0}$ is a Brownian motion process with drift μ and variance σ^2. Let $T = \min\{t : X_t = a\}$ be the first hitting time to reach level a. In the last example, the expectation of T was derived. Here, the variance of T is obtained using the quadratic martingale $Y_t = B_t^2 - t$.

Solution Since T is a stopping time with respect to B_t,

$$0 = E(Y_0) = E(Y_T) = E\left(B_T^2 - T\right) = E\left(B_T^2\right) - E(T).$$

Thus, $E\left(B_T^2\right) = E(T) = a/\mu$. Write $X_t = \mu t + \sigma B_t$. Then, $X_T = \mu T + \sigma B_T$, giving

$$B_T = \frac{X_T - \mu T}{\sigma} = \frac{a - \mu T}{\sigma}.$$

Thus,

$$\begin{aligned} Var(T) &= E\left((T - E(T))^2\right) = E\left(\left(T - \frac{a}{\mu}\right)^2\right) \\ &= \frac{1}{\mu^2}E\left((\mu T - a)^2\right) = \frac{\sigma^2}{\mu^2}E\left(\left(\frac{a-\mu T}{\sigma}\right)^2\right) \\ &= \frac{\sigma^2}{\mu^2}E\left(B_T^2\right) = \frac{\sigma^2}{\mu^2}\left(\frac{a}{\mu}\right) = \frac{a\sigma^2}{\mu^3}. \end{aligned}$$

■

R: Hitting Time Simulation for Brownian Motion with Drift

Let $(X_t)_{t\geq 0}$ be a Brownian motion with drift $\mu = 0.5$ and variance $\sigma^2 = 1$. Consider the first hitting time T of level $a = 10$. Exact results are

$$E(T) = a/\mu = 20 \quad \text{and } Var(T) = a\sigma^2/\mu^3 = 80.$$

Simulated results are based on 1,000 trials of a Brownian motion process on $[0, 80]$. See Figure 8.15 for several realizations.

```
# bmhitting.R
> mu <- 1/2
> sig <- 1
> a <- 10
> simlist <- numeric(1000)
> for (i in 1:1000) {
    t <- 80
    n <- 50000
    bm <- c(0,cumsum(rnorm(n,0,sqrt(t/n))))
```

```
    xproc <- mu*seq(0,t,t/n) + sig*bm
    simlist[i] <- which(xproc >= a)[1] * (t/n) }
> mean(simlist)
[1] 20.07139
> var(simlist)
[1] 81.31142
```

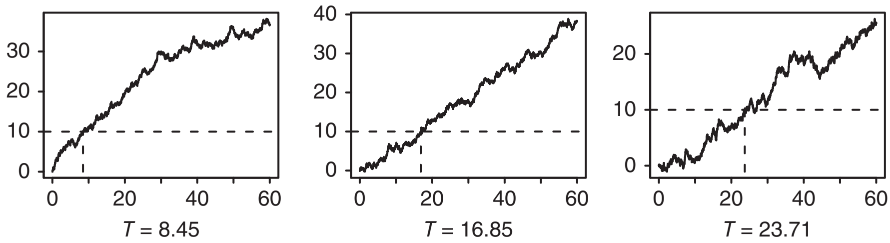

Figure 8.15 First time T to hit $a = 10$ for Brownian motion with drift $\mu = 0.5$ and variance $\sigma^2 = 1$.

EXERCISES

8.1 Show that

$$f(x,t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)},$$

satisfies the partial differential heat equation

$$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}.$$

8.2 For standard Brownian motion, find

(a) $P(B_2 \leq 1)$

(b) $E(B_4|B_1 = x)$

(c) $\text{Corr}(B_{t+s}, B_s)$

(d) $Var(B_4|B_1)$

(e) $P(B_3 \leq 5|B_1 = 2)$.

8.3 For standard Brownian motion started at $x = -3$, find

(a) $P(X_1 + X_2 > -1)$

(b) The conditional density of X_2 given $X_1 = 0$.

(c) $\text{Cov}(X_3, X_4)$

(d) $E(X_4|X_1)$.

8.4 In a race between Lisa and Cooper, let X_t denote the amount of time (in seconds) by which Lisa is ahead when $100t$ percent of the race has been completed. Assume that $(X_t)_{0\leq t\leq 1}$ can be modeled by a Brownian motion with drift parameter 0 and variance parameter σ^2. If Lisa is leading by $\sigma/2$ seconds after three-fourths of the race is complete, what is the probability that she is the winner?

8.5 Consider standard Brownian motion. Let $0 < s < t$.

(a) Find the joint density of (B_s, B_t).

(b) Show that the conditional distribution of B_s given $B_t = y$ is normal, with mean and variance

$$E(B_s|B_t = y) = \frac{sy}{t} \quad \text{and } Var(B_s|B_t = y) = \frac{s(t-s)}{t}.$$

8.6 For $s > 0$, show that the translation $(B_{t+s} - B_s)_{t\geq 0}$ is a standard Brownian motion.

8.7 Show that the reflection $(-B_t)_{t\geq 0}$ is a standard Brownian motion.

8.8 Find the covariance function for Brownian motion with drift.

8.9 Let $W_t = B_{2t} - B_t$, for $t \geq 0$.

(a) Is $(W_t)_{t\geq 0}$ a Gaussian process?

(b) Is $(W_t)_{t\geq 0}$ a Brownian motion process?

8.10 Let $(B_t)_{t\geq 0}$ be a Brownian motion started in x. Let

$$X_t = B_t - t(B_1 - y), \text{ for } 0 \leq t \leq 1.$$

The process $(X_t)_{t\geq 0}$ is a Brownian bridge with start in x and end in y. Find the mean and covariance functions.

8.11 A standard Brownian motion crosses the t-axis at times $t = 2$ and $t = 5$. Find the probability that the process exceeds level $x = 1$ at time $t = 4$.

8.12 Show that Brownian motion with drift has independent and stationary increments.

8.13 Let $(X_t)_{t\geq 0}$ denote a Brownian motion with drift μ and variance σ^2. For $0 < s < t$, find $E(X_s X_t)$.

8.14 A Brownian motion with drift parameter $\mu = -1$ and variance $\sigma^2 = 4$ starts at $x = 1.5$. Find the probability that the process is positive at $t = 3$.

8.15 See Example 8.11 on using Brownian motion to model the home team advantage in basketball. In Stern (1994), Table 8.3 is given based on the outcomes of 493 basketball games played in 1992.

Here, X_t is the difference between the home team's score and the visiting team's score after $t(100)$ percent of the game is played. The data show the mean

TABLE 8.3 Results by Quarter of 493 NBA Games

Quarter	Variable	Mean	Standard Deviation
1	$X_{0.25}$	1.41	7.58
2	$X_{0.50} - X_{0.25}$	1.57	7.40
3	$X_{0.75} - X_{0.50}$	1.51	7.30
4	$X_1 - X_{0.75}$	0.22	6.99

and standard deviation of these differences at the end of each quarter. Why might the data support the use of a Brownian motion model? What aspects of the data give reason to doubt the Brownian motion model?

8.16 Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be independent, standard Brownian motions. Show that $Z_t = a(X_t - Y_t)$ defines a standard Brownian motion for some a. Find a.

8.17 For $a > 0$, show that for standard Brownian motion the first hitting time T_a has the same distribution as $1/X^2$, where X is a normal random variable with mean 0 and variance $1/a^2$.

8.18 Show that the first hitting time T_a has the same distribution as $a^2 T_1$.

8.19 Find the mean and variance of the maximum value of standard Brownian motion on $[0, t]$.

8.20 Use the reflection principle to show

$$P(M_t \geq a, B_t \leq a - b) = P(B_t \geq a + b), \text{ for } a, b > 0. \tag{8.14}$$

8.21 From standard Brownian motion, let X_t be the process defined by

$$X_t = \begin{cases} B_t, & \text{if } t < T_a, \\ a, & \text{if } t \geq T_a, \end{cases}$$

where T_a is the first hitting time of $a > 0$. The process $(X_t)_{t\geq 0}$ is called *Brownian motion absorbed at a*. The distribution of X_t has discrete and continuous parts.

(a) Show

$$P(X_t = a) = \frac{2}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/2t}\, dx.$$

(b) For $x < a$, show

$$P(X_t \leq x) = P(B_t \leq x) - P(B_t \leq x - 2a) = \frac{1}{\sqrt{2\pi t}} \int_{x-2a}^{x} e^{-z^2/2t}\, dz.$$

Hint: Use the result of Exercise 8.20.

8.22 Let Z be the smallest zero of standard Brownian motion past t. Show that

$$P(Z \leq z) = \frac{2}{\pi} \arccos \sqrt{\frac{t}{z}}, \text{ for } z > 0.$$

8.23 Let $0 < r < s < t$.

(a) Assume that standard Brownian motion is not zero in (r, s). Find the probability that standard Brownian motion is not zero in (r, t).

(b) Assume that standard Brownian motion is not zero in $(0, s)$. Find the probability that standard Brownian motion is not zero in $(0, t)$.

8.24 Derive the mean and variance of geometric Brownian motion.

8.25 The price of a stock is modeled with a geometric Brownian motion with drift $\mu = -0.25$ and volatility $\sigma = 0.4$. The stock currently sells for \$35. What is the probability that the price will be at least \$40 in 1 year?

8.26 For the stock price model in Exercise 8.25, assume that an option is available to purchase the stock in six months for \$40. Find the expected payoff of the option.

8.27 Assume that $Z_0, Z_1, \ldots$ is a branching process whose offspring distribution has mean μ. Show that Z_n/μ^n is a martingale.

8.28 An urn contains two balls—one red and one blue. At each discrete step, a ball is chosen at random from the urn. It is returned to the urn along with another ball of the same color. Let X_n denote the number of red balls in the urn after n draws. (Thus, $X_0 = 1$.) Let $R_n = X_n/(n+2)$ be the proportion of red balls in the urn after n draws. Show that $R_0, R_1, \ldots$ is a martingale with respect to $X_0, X_1, \ldots$ The process is called *Polya's Urn*.

8.29 Show that $\left(B_t^3 - 3tB_t\right)_{t\geq 0}$ is a martingale with respect to Brownian motion.

8.30 A Brownian motion with drift $\mu = 2$ and variance $\sigma^2 = 4$ is run until level $a = 3$ is first hit. The process is repeated 25 times. Find the approximate probability that the average first hitting time of 25 runs is between 1 and 2.

8.31 Consider standard Brownian motion started at $x = -3$.

(a) Find the probability of reaching level 2 before -7.

(b) When, on average, will the process leave the region between the lines $y = 2$ and $y = -7$?

8.32 Let $(N_t)_{t\geq 0}$ be a Poisson process with parameter λ. Let $X_t = N_t - \lambda t$, for $t \geq 0$. Show that X_t is a martingale with respect to N_t.

8.33 Let $X_1, X_2, \ldots$ be i.i.d. random variables with mean $\mu < \infty$. Let $Z_n = \sum_{i=1}^{n}(X_i - \mu)$, for $n = 0, 1, 2, \ldots$

(a) Show that $Z_0, Z_1, \ldots$ is a martingale with respect to $X_0, X_1, \ldots$

(b) Assume that N is a stopping time that satisfies the conditions of the optional stopping theorem. Show that

$$E\left(\sum_{i=1}^{N} X_i\right) = E(N)\mu.$$

This result is known as *Wald's equation*.

8.34 Let $(N_t)_{t\geq 0}$ be a Poisson process with parameter λ.

(a) Find the quantity $m(t)$ such that $M_t = (N_t - \lambda t)^2 - m(t)$ is a martingale.

(b) For fixed integer $k > 0$, let $T = \min\{t : N_t = k\}$ be the first time k arrivals occur for a Poisson process. Show that T is a stopping time that satisfies the conditions of the optional stopping theorem.

(c) Use the optional stopping theorem to find the standard deviation of T.

8.35 For $a > 0$, let T be the first time that standard Brownian motion exits the interval $(-a, a)$.

(a) Show that T is a stopping time that satisfies the optional stopping theorem.

(b) Find the expected time $E(T)$ to exit $(-a, a)$.

(c) Let $M_t = B_t^4 - 6tB_t^2 + 3t^2$. Then, $(M_t)_{t\geq 0}$ is a martingale, a fact that you do not need to prove. Use this to find the standard deviation of T.

8.36 Let $(X_t)_{t\geq 0}$ be a Brownian motion with drift $\mu \neq 0$ and variance σ^2. The goal of this exercise is to find the probability p that X_t hits a before $-b$, for $a, b > 0$.

(a) Let $Y_t = e^{-tc^2/2+cB_t}$, where B_t denotes standard Brownian motion. Show that $(Y_t)_{t\geq 0}$ is a martingale for constant $c \neq 0$.

(b) Let $T = \min\{t : X_t = a \text{ or } -b\}$. Then, T is a stopping time that satisfies the optional stopping theorem. Use (a) and appropriate choice of c to show

$$E\left(e^{-\frac{2\mu X_T}{\sigma^2}}\right) = 1.$$

(c) Show that

$$p = \frac{1 - e^{2\mu b/\sigma^2}}{e^{-2\mu a/\sigma^2} - e^{2\mu b/\sigma^2}}. \tag{8.15}$$

8.37 Consider a Brownian motion with drift μ and variance σ^2. Assume that $\mu < 0$. The process tends to $-\infty$, with probability 1. Let M be the maximum value reached.

(a) See Exercise 8.36. By letting $b \to \infty$ in Equation (8.15), show

$$P(M > a) = e^{2\mu a/\sigma^2}, \text{ for } a > 0.$$

Conclude that M has an exponential distribution.

(b) A particle moves according to a Brownian motion with drift $\mu = -1.6$ and variance $\sigma^2 = 0.4$. Find the mean and standard deviation of the largest level reached.

8.38 Let $T = \min\{t : B_t \notin (-a, a)\}$. Show that

$$E\left(e^{-\lambda T}\right) = \frac{2}{e^{a\sqrt{2\lambda}} + e^{-a\sqrt{2\lambda}}} = \frac{1}{\cosh(a\sqrt{2\lambda})}.$$

(a) Apply the optional stopping theorem to the exponential martingale in Exercise 8.36(a) to show that

$$E\left(e^{\sqrt{2\lambda}B_T}e^{-\lambda T}\right) = 1.$$

(b) Show that

$$P(B_T = a, T < x) = P(B_T = -a, T < x) = \frac{1}{2}P(T < x)$$

and conclude that B_T and T are independent, and thus establish the result.

8.39 `R`: Simulate a Brownian motion $(X_t)_{t\geq 0}$ with drift $\mu = 1.5$ and variance $\sigma^2 = 4$. Simulate the probability $P(X_3 > 4)$ and compare with the exact result.

8.40 `R`: Use the script file **bbridge.R** to simulate a Brownian bridge $(X_t)_{t\geq 0}$. Estimate the probability $P(X_{3/4} \leq 1/3)$. Compare with the exact result.

8.41 `R`: The price of a stock is modeled as a geometric Brownian motion with drift $\mu = -0.85$ and variance $\sigma^2 = 2.4$. If the current price is \$50, find the probability that the price is under \$40 in 2 years. Simulate the stock price, and compare with the exact value.

8.42 `R`: Simulate the mean and standard deviation of the maximum of standard Brownian motion on $[0, 1]$. Compare with the exact values.

8.43 `R`: Write a function for pricing an option using the Black–Scholes formula. Option price is a function of initial stock price, strike price, expiration date, interest rate, and volatility.

(a) Price an option for a stock that currently sells for \$400 and has volatility 40%. Assume that the option strike price is \$420 and the expiration date is 90 days. Assume a risk-free interest rate of 3%.

(b) For each of the five parameters in (a), how does varying the parameter, holding the other four fixed, effect the price of the option?

(c) For the option in (a), assume that volatility is not known. However, based on market experience, the option sells for \$30. Estimate the volatility.