9

A GENTLE INTRODUCTION TO STOCHASTIC CALCULUS*

The study of mathematics, like the Nile, begins in minuteness but ends in magnificence.

–Charles Caleb Colton

9.1 INTRODUCTION

Brownian motion paths are continuous functions. Continuous functions are integrable. Integration of Brownian motion opens the door to powerful calculus-based modeling tools, such as stochastic differential equations (SDEs). Stochastic calculus is an advanced topic, which requires measure theory, and often several graduate-level probability courses. Our goal in this section is to introduce the subject by emphasizing intuition, and whet your appetite for what is possible in this fascinating field.

We will make sense of integrals such as

$$\int_0^t B_s \, ds \quad \text{and} \quad \int_0^t B_s \, dB_s.$$

In the first integral, Brownian motion is integrated over the interval $[0, t]$. Think of the integral as representing the area under the Brownian motion curve on $[0, t]$. The fact that the integrand is random means that the integral is random, hence a random variable. As a function of t, it is a *random function*, that is, a stochastic process.

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

If that is not strange enough, in the second integral Brownian motion appears in both the integrand and the *integrator*, where dB_s replaces the usual ds. Here, Brownian motion is integrated *with respect to* Brownian motion. To make sense of this will require new ideas, and even new rules, of calculus.

To start off, for $0 \leq a < b$, consider the integral

$$\int_a^b B_s \, ds.$$

For each outcome ω, $B_s(\omega)$ is a continuous function, and thus the integral

$$\int_a^b B_s(\omega) \, ds$$

is well defined in the usual sense as the limit of a Riemann sum. For a partition $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ of $[a, b]$, define the Riemann sum

$$I^{(n)}(\omega) = \sum_{k=1}^{n} B_{t_k^*}(\omega)(t_k - t_{k-1}),$$

where $t_k^* \in [t_{k-1}, t_k]$ is an arbitrary point in the subinterval $[t_{k-1}, t_k]$. The integral $\int_a^b B_s(\omega) \, ds$ is defined as the limit of the Riemann sum as n tends to infinity and the length of the longest subinterval of the partition converges to 0.

For each $n \geq 1$, the Riemann sum $I^{(n)}$ is a random variable, which is a linear combination of normal random variables. Since Brownian motion is a Gaussian process, $I^{(n)}$ is normally distributed. As this is true for all n, it is reasonable to expect that $\lim_{n\to\infty} I^{(n)}$ is normally distributed.

Let $I_t = \int_0^t B_s \, ds$, for $t \geq 0$. It can be shown that $(I_t)_{t\geq 0}$ is a Gaussian process with continuous sample paths. The mean function is

$$E(I_t) = E\left(\int_0^t B_s \, ds\right) = \int_0^t E(B_s) \, ds = 0,$$

where the interchange of expectation and integral can be justified.

For $s \leq t$, the covariance function is

$$\begin{aligned}
\mathrm{Cov}(I_s, I_t) = E(I_s I_t) &= E\left(\int_0^s B_x \, dx \int_0^t B_y \, dy\right) \\
&= \int_0^s \int_0^t E(B_x B_y) \, dy \, dx = \int_0^s \int_0^t \min\{x, y\} \, dy \, dx \\
&= \int_0^s \int_0^x y \, dy \, dx + \int_0^s \int_x^t x \, dy \, dx \\
&= \frac{s^3}{6} + \left(\frac{ts^2}{2} - \frac{s^3}{3}\right) = \frac{3ts^2 - s^3}{6}.
\end{aligned}$$

Letting $s = t$ gives $Var(I_t) = t^3/3$.

Thus, the stochastic integral $\int_0^t B_s\, ds$ is a normally distributed random variable with mean 0 and variance $t^3/3$. The integral $\int_0^t B_s\, ds$ is called *integrated Brownian motion*. See Figure 9.1 for realizations when $t = 1$.

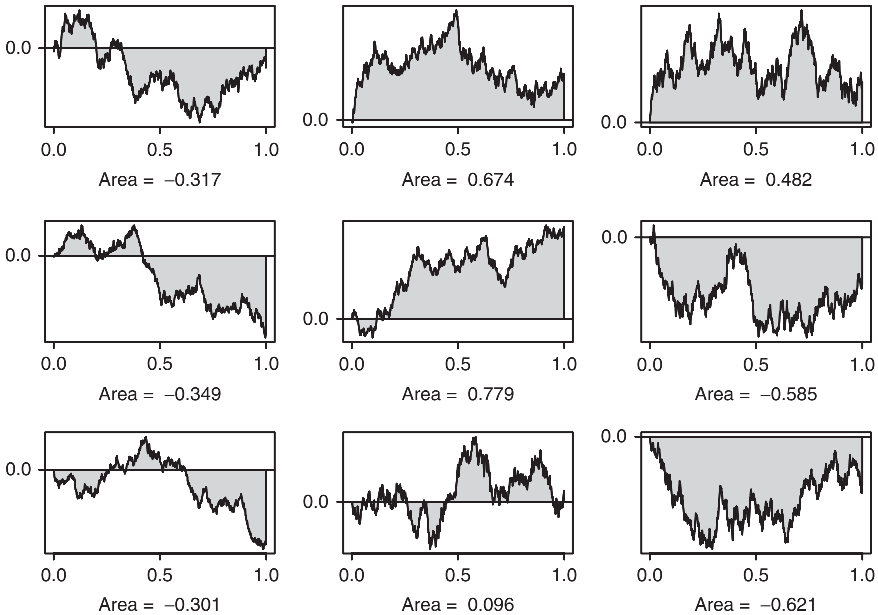

Figure 9.1 Realizations of the stochastic integral $\int_0^1 B_s\, ds$. The integral is normally distributed with mean 0 and variance $1/3$.

We next introduce the *Riemann–Stieltjes integral* of g with respect to f

$$\int_0^t g(x)df(x),$$

where f and g are continuous, and nonrandom, functions. The integral is defined as the limit, as n tends to infinity, of the approximating sum

$$\sum_{k=1}^{n} g(t_k^*) \left(f(t_k) - f(t_{k-1}) \right),$$

where $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$ is a partition of $[0, t]$, and $t_k^* \in [t_{k-1}, t_k]$. The definition generalizes the usual Riemann integral by letting $f(x) = x$. The integral can be interpreted as a *weighted* summation, or weighted average, of g, with weights determined by f.

If f is differentiable, with continuous derivative, then

$$\int_0^t g(x)df(x) = \int_0^t g(x)f'(x)\, dx,$$

which gives the usual Riemann integral. In probability, if X is a continuous random variable with density function f and cumulative distribution function F, and g is a function, the expectation $E(g(X))$ can be written as

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\, dx = \int_{-\infty}^{\infty} g(x)F'(x)\, dx = \int_{-\infty}^{\infty} g(x)dF(x),$$

that is, as a Riemann–Stieltjes integral of g with respect to the cumulative distribution function F.

Based on the Riemann–Stieltjes integral, we can define the integral of a function g with respect to Brownian motion

$$I_t = \int_0^t g(s)\, dB_s. \tag{9.1}$$

Technical conditions require that g be a bounded and continuous function, and satisfy $\int_0^\infty g^2(s)\, ds < \infty$. By analogy with the Riemann–Stieltjes integral, for the partition

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t,$$

let

$$I_t^{(n)} = \sum_{k=1}^{n} g(t_k^*)\left(B_{t_k} - B_{t_{k-1}}\right),$$

where $t_k^* \in [t_{k-1}, t_k]$. Since $B_{t_k} - B_{t_{k-1}}$ is normally distributed with mean 0 and variance $t_k - t_{k-1}$, the approximating sum $I_t^{(n)}$ is normally distributed for all n. It can be shown that in the limit, as $n \to \infty$, the approximating sum converges to a normally distributed random variable, which we take to be the stochastic integral of Equation (9.1). Furthermore,

$$\begin{aligned} E(I_t) &= E\left(\lim_{n\to\infty} I_t^{(n)}\right) = \lim_{n\to\infty} E\left(\sum_{k=1}^{n} g(t_k^*)\left(B_{t_k} - B_{t_{k-1}}\right)\right) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} g(t_k^*)E\left(B_{t_k} - B_{t_{k-1}}\right) = 0. \end{aligned}$$

By independent increments,

$$Var\left(I_t^{(n)}\right) = \sum_{k=1}^{n} g^2(t_k^*)Var\left(B_{t_k} - B_{t_{k-1}}\right) = \sum_{k=1}^{n} g^2(t_k^*)(t_k - t_{k-1}).$$

The last expression is a Riemann sum whose limit, as n tends to infinity, is $\int_0^t g^2(s)\, ds$. In summary,

$$\int_0^t g(s)\, dB_s \sim \text{Normal}\ \left(0, \int_0^t g^2(s)\, ds\right). \tag{9.2}$$

In fact, it can be shown that $(I_t)_{t\geq 0}$ is a Gaussian process with continuous sample paths, independent increments, mean function 0, and covariance function

$$\mathrm{Cov}(I_s, I_t) = E\left(\int_0^s g(x)\, dB_x \int_0^t g(y)\, dB_y\right) = \int_0^{\min\{s,t\}} g^2(x)\, dx.$$

Example 9.1 Evaluate

$$\int_0^t dB_s.$$

Solution With $g(x) = 1$, the integral is normally distributed with mean 0 and variance $\int_0^t ds = t$. That is, the stochastic integral has the same distribution as B_t. Furthermore, the integral defines a continuous Gaussian process with mean 0 and covariance function

$$\int_0^{\min\{s,t\}} dx = \min\{s, t\}.$$

That is, $\left(\int_0^t dB_s\right)_{t\geq 0}$ is a standard Brownian motion, and

$$\int_0^t dB_s = B_t.$$

■

Example 9.2 Evaluate

$$\int_0^t e^s\, dB_s.$$

Solution The integral is normally distributed with mean 0 and variance

$$\int_0^t (e^s)^2\, ds = \int_0^t e^{2s}\, ds = \frac{1}{2}e^{2t}.$$

■

The stochastic integral

$$\int_a^b g(s)\, dB_s$$

has many familiar properties. Linearity holds. For functions g and h, for which the integral is defined, and constants α, β,

$$\int_a^b [\alpha g(s) + \beta h(s)]\, dB_s = \alpha \int_a^b g(s)\, dB_s + \beta \int_a^b h(s)\, dB_s.$$

For $a < c < b$,

$$\int_a^b g(s)\, dB_s = \int_a^c g(s)\, dB_s + \int_c^b g(s)\, dB_s.$$

The integral also satisfies an *integration-by-parts* formula. If g is differentiable,

$$\int_0^t g(s)\, dB_s = g(t)B_t - \int_0^t B_s g'(s)\, ds.$$

By letting $g(t) = 1$, we capture the identity

$$\int_0^t dB_s = B_t. \tag{9.3}$$

Example 9.3 Evaluate

$$\int_0^t s\, dB_s$$

in terms of integrated Brownian motion.

Solution Integration by parts gives

$$\int_0^t s\, dB_s = tB_t - \int_0^t B_s\, ds.$$

See Figure 9.2 for simulations of the process. ■

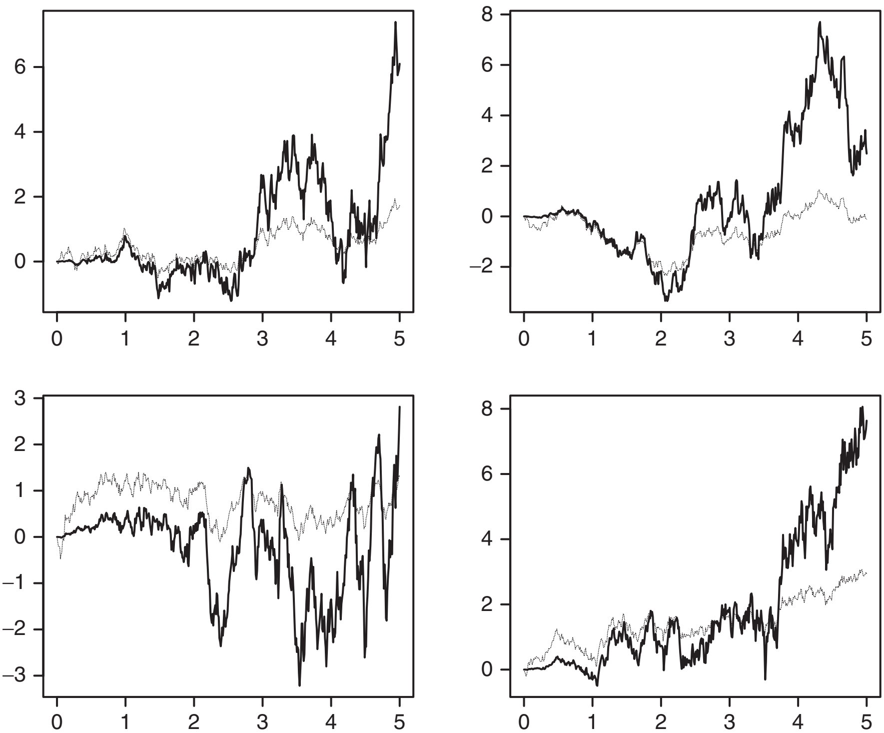

Figure 9.2 Simulations of $\left(\int_0^t s\, dB_s\right)_{0\le t\le 5}$. The light gray curve is the underlying standard Brownian motion.

White Noise

If Brownian motion paths were differentiable, Equation (9.3) could be written as

$$B_t = \int_0^t dB_s = \int_0^t \frac{dB_s}{ds}\, ds.$$

The "process" $W_t = dB_t/dt$ is called *white noise*. The reason for the quotation marks is because W_t is not a stochastic process in the usual sense, as Brownian motion derivatives do not exist. Nevertheless, Brownian motion is sometimes described as *integrated white noise*.

Consider the following formal treatment of the distribution of white noise. Letting Δ_t represent a small incremental change in t,

$$W_t \approx \frac{B_{t+\Delta_t} - B_t}{\Delta_t}.$$

The random variable W_t is approximately normally distributed with mean 0 and variance $1/\Delta_t$. We can think of W_t as the result of letting $\Delta_t \to 0$. White noise can be thought of as an *idealized* continuous-time Gaussian process, where W_t is normally distributed with mean 0 and variance $1/dt$. Furthermore, for $s \neq t$,

$$E(W_s W_t) = E\left(\frac{dB_s}{ds}\frac{dB_t}{dt}\right) = \frac{1}{ds\, dt} E\left((B_{s+ds} - B_s)(B_{t+dt} - B_t)\right) = 0,$$

by independent increments. Hence, W_s and W_t are independent, for all $s \neq t$.

It is hard to conceive of a real-world, time-indexed process in which all variables, no matter how close in time, are independent. Yet white noise is an extremely useful concept for real-world modeling, particularly in engineering, biology, physics, communication, and economics. In a physical context, white noise refers to sound that contains all frequencies in equal amounts, the analog of white light. See Figure 9.3. Applied to a time-varying signal g, the stochastic integral $\int_0^t g(s)\, dB_s$ can be interpreted as the output after the signal is transformed by white noise. For the case $g(s) = s$, see again Figure 9.2.

9.2 ITO INTEGRAL

We are now ready to consider a more general stochastic integral of the form

$$I_t = \int_0^t X_s\, dB_s,$$

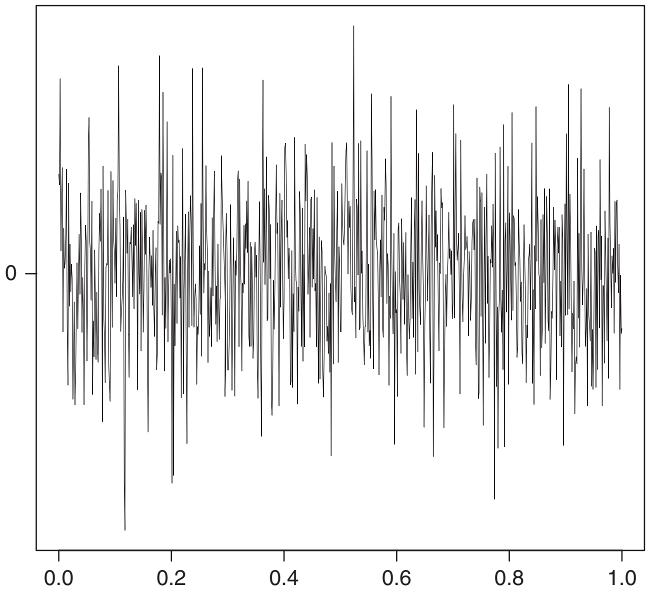

Figure 9.3 Simulation of white noise signal.

where $(X_t)_{t\geq 0}$ is a stochastic process, and $(B_t)_{t\geq 0}$ is standard Brownian motion. By analogy with what has come before, it would seem that a reasonable definition for the integral would be

$$\int_0^t X_s \, dB_s = \lim_{n\to\infty} \sum_{k=1}^{n} X_{t_k^*}\left(B_{t_k} - B_{t_{k-1}}\right), \tag{9.4}$$

for ever-finer partitions $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$, where $t_k^* \in [t_{k-1}, t_k]$. Unfortunately, the definition does not work. Unlike previous integrals, the choice of point t_k^* in the subinterval $[t_{k-1}, t_k]$ matters. The integral is not well-defined for arbitrary $t_k^* \in [t_{k-1}, t_k]$. Furthermore, the integral requires a precise definition of the meaning of the limit in Equation (9.4), as well as several regularity conditions for the process $(X_t)_{t\geq 0}$.

This brings us to the *Ito integral*, named after Kiyoshi Ito, a brilliant 20th century Japanese mathematician whose name is most closely associated with stochastic calculus. The Ito integral is based on taking each t_k^* to be the left endpoint[1] of the subinterval $[t_{k-1}, t_k]$. That is,

$$\int_0^t X_s \, dB_s = \lim_{n\to\infty} \sum_{k=1}^{n} X_{t_{k-1}}\left(B_{t_k} - B_{t_{k-1}}\right). \tag{9.5}$$

[1]A different type of stochastic integral, called the *Stratonovich integral*, is obtained by choosing $t_k^* = (t_{k-1} + t_k)/2$ to be the midpoint of the subinterval $[t_{k-1}, t_k]$.

The Ito integral requires

1. $\int_0^t E\left(X_s^2\right)\ ds < \infty$.
2. X_t does not depend on the values $\{B_s : s > t\}$ and only on $\{B_s : s \leq t\}$. We say that X_t is *adapted* to Brownian motion.
3. The limit in Equation (9.5) is defined in the *mean-square* sense. A sequence of random variables $X_0, X_1, \ldots$ is said to *converge to X in mean-square* if $\lim_{n\to\infty} E\Big((X_n - X)^2\Big) = 0$.

The Ito integral has many familiar properties, such as linearity. However, new rules of *stochastic calculus* will be needed for computations.

One of the most important properties of the Ito integral is that the process

$$\left(\int_0^t X_s\ dB_s\right)_{t\geq 0}$$

is a martingale with respect to Brownian motion.

The following properties of the Ito integral are summarized without proof.

Properties of the Ito Integral

The Ito integral

$$I_t = \int_0^t X_s\ dB_s$$

satisfies the following:

1. For processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$, and constants α, β,
$$\int_0^t \left(\alpha X_s + \beta Y_s\right)\ dB_s = \alpha \int_0^t X_s\ dB_s + \beta \int_0^t Y_s\ dB_s.$$
2. For $0 < r < t$,
$$\int_0^t X_s\ dB_s = \int_0^r X_s\ dB_s + \int_r^t X_s\ dB_s.$$
3. $$E(I_t) = 0.$$
4. $$Var(I_t) = E\left(\left(\int_0^t X_s\ dB_s\right)^2\right) = \int_0^t E\left(X_s^2\right)\ ds.$$
5. $(I_t)_{t\geq 0}$ is a martingale with respect to Brownian motion.

The Ito integral does not satisfy the usual integration-by-parts formula. Consider

$$\int_0^t B_s \, dB_s.$$

A formal application of integration by parts gives

$$\int_0^t B_s \, dB_s = B_t^2 - B_0^2 - \int_0^t B_s \, dB_s = B_t^2 - \int_0^t B_s \, dB_s,$$

which leads to $\int_0^t B_s \, dB_s = B_t^2/2$. However, this must be wrong as the Ito integral has mean 0, and thus $E\left(\int_0^t B_s \, dB_s\right) = 0$. However, $E\left(B_t^2/2\right) = t/2$.

To evaluate $\int_0^t B_s \, dB_s$, consider the approximating sum

$$\begin{aligned}
&\sum_{k=1}^n B_{t_{k-1}}\left(B_{t_k} - B_{t_{k-1}}\right)\\
&\quad = \sum_{k=1}^n \left(\frac{1}{2}\left(B_{t_k} + B_{t_{k-1}}\right) - \frac{1}{2}\left(B_{t_k} - B_{t_{k-1}}\right)\right)\left(B_{t_k} - B_{t_{k-1}}\right)\\
&\quad = \frac{1}{2}\sum_{k=1}^n \left(B_{t_k}^2 - B_{t_{k-1}}^2\right) - \frac{1}{2}\sum_{k=1}^n \left(B_{t_k} - B_{t_{k-1}}\right)^2\\
&\quad = \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{k=1}^n \left(B_{t_k} - B_{t_{k-1}}\right)^2.
\end{aligned}$$

It can be shown that $\sum_{k=1}^n \left(B_{t_k} - B_{t_{k-1}}\right)^2$ converges to the constant t in mean-square, that is,

$$\lim_{n\to\infty} E\left(\left(\sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 - t\right)^2\right) = 0.$$

This gives

$$\int_0^t B_s \, dB_s = \frac{1}{2}\left(B_t^2 - t\right),$$

which is a martingale. Recall that $B_t^2 - t$ is the quadratic martingale, shown in Example 8.19. Multiplying a martingale by a constant, does not change the martingale property.

ITO'S LEMMA

> If one disqualifies the Pythagorean Theorem from contention, it is hard to think of a mathematical result which is better known and more widely applied in the world today

> than "Ito's Lemma." This result holds the same position in stochastic analysis that Newton's fundamental theorem holds in classical analysis. That is, it is the sine qua non of the subject.
>
> –National Academy of Sciences

The most important result in stochastic calculus is Ito's Lemma, which is the stochastic version of the chain rule. It has been called the fundamental theorem of stochastic calculus.

Ito's Lemma

Let g be a real-valued function that is twice continuously differentiable. Then,

$$g(B_t) - g(B_0) = \int_0^t g'(B_s)\, dB_s + \frac{1}{2}\int_0^t g''(B_s)\, ds.$$

This is often written in shorthand differential form

$$dg(B_t) = g'(B_t)dB_t + \frac{1}{2}g''(B_t)dt.$$

Example 9.4 Let $g(x) = x^2$. By Ito's Lemma,

$$B_t^2 = \int_0^t 2B_s\, dB_s + \frac{1}{2}\int_0^t 2\, ds = 2\int_0^t B_s\, dB_s + t.$$

That is,

$$\int_0^t B_s\, dB_s = \frac{1}{2}B_t^2 - \frac{t}{2}.$$

■

Example 9.5 Evaluate $d(\sin B_t)$.

Solution Let $g(x) = \sin x$. By Ito's Lemma,

$$d(\sin B_t) = \cos B_t\, dB_t - \frac{1}{2}\sin B_t\, dt.$$

In integral form,

$$\sin B_t = \int_0^t \cos B_s\, dB_s - \frac{1}{2}\int_0^t \sin B_s\, ds.$$

■

Example 9.6 Evaluate

$$\int_0^t B_s^2\, dB_s \quad \text{and} \quad \int_0^t (B_s^2 - s)\, dB_s.$$

Solution

(i) Use Ito's Lemma with $g(x) = x^3$. This gives

$$B_t^3 = \int_0^t 3B_s^2 \, dB_s + \frac{1}{2}\int_0^t 6B_s \, ds.$$

Rearranging gives

$$\int_0^t B_s^2 \, dB_s = \frac{1}{3}B_t^3 - \int_0^t B_s \, ds.$$

(ii) By linearity of the Ito integral,

$$\begin{aligned}\int_0^t \left(B_s^2 - s\right) \, dB_s &= \int_0^t B_s^2 \, dB_s - \int_0^t s \, dB_s \\ &= \frac{1}{3}B_t^3 - \int_0^t B_s \, ds - \left(tB_t - \int_0^t B_s \, ds\right) \\ &= \frac{1}{3}B_t^3 - tB_t.\end{aligned}$$

The second equality is by integration by parts, which is valid for stochastic integrals with deterministic integrands.

Since Ito integrals are martingales, the process $\left(\frac{1}{3}B_t^3 - tB_t\right)_{t\geq 0}$ is a martingale. ■

Here is a heuristic derivation of Ito's Lemma. For a function g, consider its Taylor series expansion

$$g(t + dt) = g(t) + g'(t)dt + \frac{1}{2}g''(t)(dt)^2 + \cdots \; .$$

Higher-order terms, starting with $(dt)^2$, are negligible. Hence,

$$dg(t) = g(t + dt) - g(t) = g'(t)dt.$$

For a given function h,

$$g(h(t) + dh(t)) = g(h(t)) + g'(h(t))dh(t) + \frac{1}{2}g''(h(t))(dh(t))^2 + \cdots \; .$$

Under suitable regularity conditions, the higher-order terms drop out, giving the usual chain rule $dg(h) = g'(h)dh$.

Replacing $h(t)$ with B_t, the Taylor series expansion is

$$g(B_t + dB_t) = g(B_t) + g'(B_t)dB_t + \frac{1}{2}g''(B_t)(dB_t)^2 + \frac{1}{6}g'''(B_t)(dB_t)^3 + \cdots \; .$$

However, what is different for Brownian motion is that the term $(dB_t)^2$ is not negligible and cannot be eliminated. Intuitively, $dB_t = B_{t+dt} - B_t$ is a stochastic element with the same distribution as B_{dt}, which is normally distributed with mean 0 and variance dt. Thus, B_{dt} takes values on the order of the standard deviation $\sqrt{dt}$. This gives $(dB_t)^2 \approx \left(\sqrt{dt}\right)^2 = dt$. Thus, the $(dB_t)^2 = dt$ term is retained in the expansion.

Higher-order terms beyond the quadratic term are dropped from the expansion, as $(dB_t)^k \approx \left(\sqrt{dt}\right)^k = (dt)^{k/2}$, which is negligible for $k > 2$. This leaves

$$dg(B_t) = g(B_t + dB_t) - g(B_t) = g'(B_t)dB_t + \frac{1}{2}g''(B_t)dt,$$

the differential form of Ito's Lemma.

Here are the heuristic stochastic calculus rules for working with stochastic differentials:

$$(dt)^2 = 0, \quad (dt)(dB_t) = 0, \quad (dB_t)^2 = dt.$$

An extended version of Ito's Lemma allows g to be a function of both t and B_t. The extended result can be motivated by considering a second-order Taylor series expansion of g.

Extension of Ito's Lemma

Let $g(t, x)$ be a real-valued function whose second-order partial derivatives are continuous. Then,

$$g(t, B_t) - g(0, B_0) = \int_0^t \left(\frac{\partial}{\partial t} g(s, B_s) + \frac{1}{2}\frac{\partial^2}{\partial x^2} g(s, B_s) \right) ds$$
$$+ \int_0^t \frac{\partial}{\partial x} g(s, B_s)\, dB_s.$$

In shorthand differential form,

$$dg = \left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial x^2} \right) dt + \frac{\partial g}{\partial x} dB_t.$$

We regret possible notational confusion in the statement of the lemma. It is common to use the letter t as the time variable, and thus t appears both as the upper limit of integration and as the dummy variable for the function g and its derivative. We trust the reader will safely navigate their way.

Example 9.7 Evaluate $d\left(tB_t^2\right)$.

Solution Let $g(t, x) = tx^2$. Partial derivatives are

$$\frac{\partial g}{\partial t} = x^2, \quad \frac{\partial g}{\partial x} = 2tx, \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} = 2t.$$

By Ito's Lemma,

$$d\left(tB_t^2\right) = \left(B_t^2 + t\right) dt + 2tB_t\, dB_t.$$

Observe that the usual product rule would give the incorrect answer

$$d\left(tB_t^2\right) = B_t^2 dt + 2tB_t\, dB_t.$$

■

Example 9.8 Use Ito's Lemma to evaluate $d\left(B_t^3\right)$ and $E\left(B_t^3\right)$.

Solution Let $g(t, x) = x^3$. By Ito's Lemma,

$$d\left(B_t^3\right) = 3B_t\, dt + 3B_t^2\, dB_t,$$

and

$$B_t^3 = 3\int_0^t B_s\, ds + 3\int_0^t B_s^2\, dB_s.$$

Taking expectations gives

$$E\left(B_t^3\right) = 3\int_0^t E(B_s)\, ds + 3E\left(\int_0^t B_s^2\, dB_s\right) = 3(0) + 0 = 0.$$

■

9.3 STOCHASTIC DIFFERENTIAL EQUATIONS

To motivate the discussion, consider an exponential growth process, be it the spread of a disease, the population of a city, or the number of cells in an organism. Let X_t denotes the size of the population at time t. The deterministic exponential growth model is described by an ordinary differential equation

$$\frac{dX_t}{dt} = \alpha X_t, \text{ and } X_0 = x_0,$$

where x_0 is the initial size of the population, and α is the growth rate. The equation says that the population growth rate is proportional to the size of the population, where the constant of proportionality is α. A *solution* of the differential equation is a function X_t, which satisfies the equation. In this case, the solution is uniquely specified

$$X_t = x_0 e^{\alpha t}, \text{ for } t \geq 0.$$

The most common way to incorporate uncertainty into the model is to add a random error term, such as a multiple of white noise W_t, to the growth rate. This gives the *stochastic differential equation*

$$\frac{dX_t}{dt} = (\alpha + \beta W_t)X_t = \alpha X_t + \beta X_t \frac{dB_t}{dt},$$

or

$$dX_t = \alpha X_t \, dt + \beta X_t \, dB_t, \tag{9.6}$$

where α and β are parameters, and $X_0 = x_0$. Equation (9.6) is really a shorthand for the integral form

$$X_t - X_0 = \alpha \int_0^t X_s \, ds + \beta \int_0^t X_s \, dB_s. \tag{9.7}$$

A solution to the SDE is a stochastic process $(X_t)_{t\geq 0}$, which satisfies Equation (9.7).

For the stochastic exponential model, we show that geometric Brownian motion defined by

$$X_t = x_0 e^{\left(\alpha - \frac{\beta^2}{2}\right)t + \beta B_t}, \text{ for } t \geq 0, \tag{9.8}$$

is a solution. Let

$$g(t, x) = x_0 e^{\left(\alpha - \frac{\beta^2}{2}\right)t + \beta x}$$

with partial derivatives

$$\frac{\partial g}{\partial x} = \beta g, \quad \frac{\partial^2 g}{\partial x^2} = \beta^2 g, \quad \text{and} \quad \frac{\partial g}{\partial t} = \left(\alpha - \frac{\beta^2}{2}\right) g.$$

By the extended version of Ito's Lemma,

$$\begin{aligned} g(t, B_t) - g(0, B_0) &= x_0 e^{\left(\alpha - \frac{\beta^2}{2}\right)t + \beta B_t} - x_0 \\ &= \left(\alpha - \frac{\beta^2}{2} + \frac{\beta^2}{2}\right) \int_0^t x_0 e^{\left(\alpha - \frac{\beta^2}{2}\right)s + \beta B_s} \, ds \\ &\quad + \beta \int_0^t x_0 e^{\left(\alpha - \frac{\beta^2}{2}\right)s + \beta B_s} \, dB_s, \end{aligned}$$

which reduces to the solution

$$X_t - X_0 = \alpha \int_0^t X_s \, ds + \beta \int_0^t X_s \, dB_s.$$

Geometric Brownian motion can be thought of as the stochastic analog of the exponential growth function.

Differential equations are the meat and potatoes of applied mathematics. Stochastic differential equations are used in biology, climate science, engineering, economics, physics, ecology, chemistry, and public health.

Example 9.9 (Logistic equation) Unfettered exponential growth is typically unrealistic for biological populations. The *logistic* model describes the growth of a self-limiting population. The standard deterministic model is described by the ordinary differential equation

$$\frac{dP_t}{dt} = rP_t\left(1 - \frac{P_t}{K}\right),$$

where P_t denotes the population size at time t, r is the growth rate, and K is the *carrying capacity*, the maximum population size that the environment can sustain.

If $P_t \approx 0$, then $dP_t/dt \approx rP_t$, and the model exhibits near-exponential growth. On the contrary, if the population size is near carrying capacity and $P_t \approx K$, then $P_t/dt \approx 0$, and the population exhibits little growth.

The solution of the deterministic equation—obtained by separation of variables and partial fractions—is

$$P_t = \frac{KP_0}{P_0 + (K - P_0)e^{-rt}}, \quad \text{for } t \geq 0. \tag{9.9}$$

Observe that $P_t \to K$, as $t \to \infty$; that is, the population size tends to the carrying capacity.

A stochastic logistic equation is described by the SDE

$$dP_t = rP_t\left(1 - \frac{P_t}{K}\right) dt + \sigma P_t \, dB_t,$$

where $\sigma > 0$ is a parameter. Let $(X_t)_{t\geq 0}$ be the geometric Brownian motion process defined by

$$X_t = e^{\left(r - \frac{\sigma^2}{2}\right)t + \sigma B_t}.$$

It can be shown that the solution to the logistic SDE is

$$P_t = \frac{P_0 K X_t}{K + P_0 r \int_0^t X_s \, ds}.$$

When $\sigma = 0$, $X_t = e^{rt}$ and the solution reduces to Equation (9.9). See Figure 9.4 for sample paths of the stochastic logistic process. ■

Example 9.10 Stochastic models are used in climatology to model long-term climate variability. These complex models are typically multidimensional and involve systems of SDEs. They are relevant for our understanding of global warming and climate change.

A simplified system that models the interaction between the atmosphere and the ocean's surface is described in Vallis (2010). Let T_t^A and T_t^S denote the atmosphere and sea surface temperatures, respectively, at time t. The system is

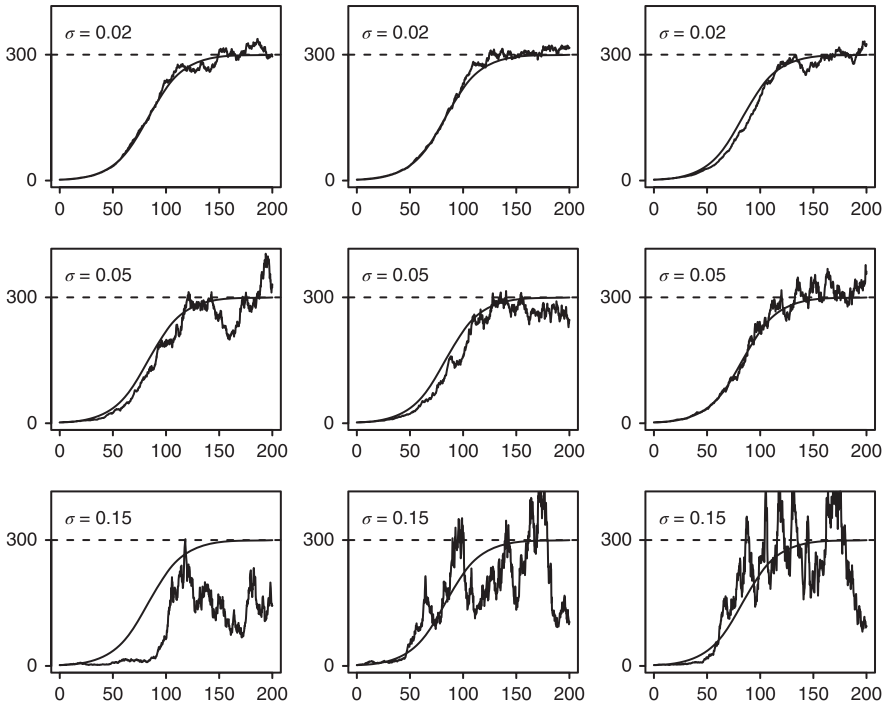

Figure 9.4 Sample paths for the logistic SDE, with $P_0 = 2$, $r = 0.06$, and $K = 300$. Smooth curve is the deterministic logistic function.

$$c_S \frac{dT^S}{dt} = aT^A - bT^S,$$
$$c_A \frac{dT^A}{dt} = cT^S - dT^A + \sigma B_t,$$

where c_A and c_S describe the heat capacity of the atmosphere and sea, respectively, and a, b, c, d, and σ are parameters. The model is based on Newton's law of cooling, by which the rate of heat loss of an object is proportional to the difference in temperature between the object and its surroundings. The Brownian motion term σB_t accounts for random fluctuations that affect the atmosphere.

Assuming the heat capacity of the ocean surface is much greater than that of the atmosphere, the model finds that rapid changes of atmospheric temperatures can affect long-term fluctuations in the ocean temperature, over possibly decades or centuries. The finding has significance in understanding how temperature changes at time scales greater than a year can occur in the earth's climate. ■

Ito's Lemma is an important tool for working with stochastic differential equations. The Lemma can be extended further to include a wide class of stochastic processes, which are solutions to SDEs of the form

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t, \tag{9.10}$$

where a and b are functions of t and X_t.

The integral form is

$$X_t - X_0 = \int_0^t a(s, X_s)\, ds + \int_0^t b(s, X_s)\, dB_s.$$

Such processes are called *diffusions* or *Ito processes*. A diffusion is a Markov process with continuous sample paths. The functions a and b are called, respectively, the *drift coefficient* and *diffusion coefficient*.

Standard Brownian motion is a diffusion with $a(t, x) = 0$ and $b(t, x) = 1$. The introductory example of this section shows that geometric Brownian motion is a diffusion with $a(t, x) = \alpha x$ and $b(t, x) = \beta x$, for parameters α, β.

Ito's Lemma for Diffusions

Let $g(t, x)$ be a real-valued function whose second-order partial derivatives are continuous. Let $(X_t)_{t\geq 0}$ be an Ito process as defined by Equation (9.10). Then

$$\begin{aligned} g(t, X_t) - g(0, X_0) = &\int_0^t \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial x}\alpha(s, X_s) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\beta^2(s, X_s) \right) ds \\ &+ \int_0^t \left(\frac{\partial g}{\partial x}\beta(s, X_s) \right) dB_s. \end{aligned}$$

In shorthand differential form,

$$dg = \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial x}\alpha(t, X_t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\beta^2(t, X_t) \right) dt + \frac{\partial g}{\partial x}\beta(t, X_t) dB_t.$$

We showed in the introductory example that geometric Brownian motion is a solution to the SDE of Equation (9.6). However, we did not solve the equation directly. Rather, we offered a candidate process and then verified that it was in fact a solution.

In general, solving an SDE may be difficult. A closed-form solution might not exist, and numerical methods are often needed. However, for the stochastic exponential growth model, the SDE can be solved exactly with the help of Ito's Lemma for diffusions.

From Equation (9.6), divide through by X_t to obtain

$$\frac{dX_t}{X_t} = \alpha\ dt + \beta\ dB_t.$$

The left-hand side suggests the function dx/x, whose integral is $\ln x$. This suggests applying Ito's Lemma with $g(t, x) = \ln x$. Derivatives are

$$\frac{\partial g}{\partial t} = 0, \quad \frac{\partial g}{\partial x} = \frac{1}{x}, \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} = -\frac{1}{x^2}.$$

This gives

$$d \ln X_t = \left(\frac{1}{X_t} \alpha X_t - \frac{1}{2X_t^2} \beta^2 X_t^2 \right) dt + \frac{1}{X_t} \beta X_t \, dB_t = \left(\alpha - \frac{\beta^2}{2} \right) dt + \beta \, dB_t.$$

Integrating gives

$$\ln X_t - \ln x_0 = \left(\alpha - \frac{\beta^2}{2} \right) t + \beta B_t,$$

with solution

$$X_t = x_0 e^{\left(\alpha - \frac{\beta^2}{2} \right) t + \beta B_t}.$$

Example 9.11 (Ornstein–Uhlenbeck process) Mathematical Brownian motion is not necessarily the best model for physical Brownian motion. If B_t denotes the position of a particle, such as a pollen grain, at time t, then the particle's position is changing over time and it must have velocity. The velocity of the grain would be the derivative of the process, which does not exist for mathematical Brownian motion.

The *Ornstein–Uhlenbeck process*, called the Langevin equation in physics, arose as an attempt to model this velocity. In finance, it is known as the Vasicek model and has been used to model interest rates. The process is called *mean-reverting* as there is a tendency, over time, to reach an equilibrium position.

The SDE for the Ornstein–Uhlenbeck process is

$$dX_t = -r(X_t - \mu)dt + \sigma B_t,$$

where r, μ, and $\sigma > 0$ are parameters. The process is a diffusion with

$$a(t, x) = -r(x - \mu) \quad \text{and} \quad b(t, x) = \sigma.$$

If $\sigma = 0$, the equation reduces to an ordinary differential equation, which can be solved by separation of variables. From

$$\frac{dX_t}{X_t - \mu} = -r \, dt,$$

integrating gives

$$\ln(X_t - \mu) = -rt + C,$$

where $C = \ln(X_0 - \mu)$. This gives the deterministic solution

$$X_t = \mu + (X_0 - \mu)e^{-rt}.$$

Observe that $X_t \to \mu$, as $t \to \infty$.

The SDE can be solved using Ito's Lemma by letting $g(t, x) = e^{rt}x$, with partial derivatives

$$\frac{\partial g}{\partial t} = re^{rt}x, \quad \frac{\partial g}{\partial x} = e^{rt}, \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} = 0.$$

By Ito's Lemma,

$$d\left(e^{rt}X_t\right) = \left(re^{rt}X_t - e^{rt}r(X_t - \mu)\right) dt + e^{rt}\sigma dB_t$$
$$= r\mu e^{rt}dt + e^{rt}\sigma\ dB_t.$$

This gives

$$e^{rt}X_t - X_0 = r\mu \int_0^t e^{rs}\ ds + \sigma \int_0^t e^{rs}\ dB_s = \mu(e^{rt} - 1) + \sigma \int_0^t e^{rs}\ dB_s,$$

with solution

$$X_t = \mu + (X_0 - \mu)e^{-rt} + \sigma \int_0^t e^{-r(t-s)}\ dB_s.$$

See Figure 9.5 for realizations of the Ornstein–Uhlenbeck process.

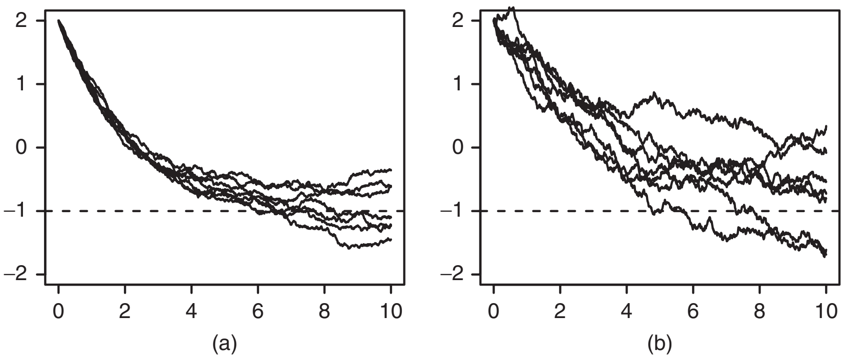

Figure 9.5 Realizations of the Ornstein–Uhlenbeck process with $X_0 = 2$ and $\mu = -1$. (a) $r = 0.5, \sigma = 0.1$. (b) $r = 0.3, \sigma = 0.2$.

If X_0 is constant, then by Equation (9.2), X_t is normally distributed with

$$E(X_t) = \mu + (X_0 - \mu)e^{-rt}$$

and

$$Var(X_t) = \sigma^2 \int_0^t e^{-2r(t-s)}\ ds = \frac{\sigma^2}{2r}\left(1 - e^{-2rt}\right).$$

The limiting distribution, as $t \to \infty$, is normal with mean μ and variance $\sigma^2/2r$. ■

Numerical Approximation and the Euler–Maruyama Method

The differential form of a stochastic differential equation lends itself to an intuitive method for simulation. Given the SDE

$$dX_t = a(t, X_t)d_t + b(t, X_t)dB_t,$$

the *Euler–Maruyama* method generates a discrete sequence $X_0, X_1, \ldots, X_n$, which approximates the process X_t on an interval $[0, T]$. The method extends the popular Euler method for numerically solving deterministic differential equations.

Partition the interval $[0, T]$ into n equally spaced points

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T,$$

where $t_i = iT/n$, for $i = 0, 1, \ldots, n$. The differential dt_i is approximated by $t_i - t_{i-1} = T/n$. The stochastic differential dB_{t_i} is approximated by $B_{t_i} - B_{t_{i-1}}$, which is normally distributed with mean 0 and variance $t_i - t_{i-1} = T/n$. Thus, dB_{t_i} can be approximated by $\sqrt{T/n}Z$, where Z is a standard normal random variable. Let

$$X_{i+1} = X_i + a(t_i, X_i)T/n + b(t_i, X_i)\sqrt{T/n}Z_i, \text{ for } i = 0, 1, \ldots, n-1,$$

where $Z_0, Z_1, \ldots, Z_{n-1}$ are independent standard normal random variables. The sequence $X_0, X_1, \ldots, X_n$ is defined recursively and gives a discretized approximate sample path for $(X_t)_{0 \le t \le T}$.

Example 9.12 (Ornstein–Uhlenbeck process) To simulate the solution of the Ornstein–Uhlenbeck SDE

$$dX_t = -r(X_t - \mu)dt + \sigma \, dB_t, \text{ for } 0 \le t \le T,$$

let

$$X_{i+1} = X_i - r(X_i - \mu)T/n + \sigma\sqrt{T/n}Z_i, \text{ for } i = 0, 1, \ldots, n-1.$$

With $n = 1000$, we generate the process with $X_0 = 2$, $\mu = -1$, $r = 0.5$, and $\sigma = 0.1$. Realizations are shown in Figure 9.5(a). ■

R : Ornstein–Uhlenbeck Simulation

```
# ornstein.R
> mu <- -1
> r <- 0.5
> sigma <- 0.1
> T <- 10
> n <- 1000
> xpath <- numeric(n+1)
> xpath[1] <- 2  # initial value
> for (i in 1:n) {
  +   xpath[i+1] <- xpath[i]-r*(xpath[i]-mu)*T/n
  +    + sigma*sqrt(T/n)*rnorm(1) }
> plot(seq(0,T,T/n),xpath,type="l")
```

To simulate the random variable X_T, for fixed T, it is not necessary to store past outcomes X_t, for $t < T$. To generate one outcome of X_T the code simplifies.

```
> x <- 2  # initial value
> for (i in 1:n) {
  + x <- x - r*(x-mu)*T/n + sigma*sqrt(T/n)*rnorm(1) }
> x
[1] -0.9404498
```

Here, we simulate the mean of X_{10} based on 10,000 trials.

```
> trials <- 10000
> simlist <- numeric(trials)
> for (k in 1: trials) {
> x <- 2
> for (i in 1:n) {
  + x <- x - r*(x-mu)*T/n + sigma*sqrt(T/n)*rnorm(1) }
> simlist[k] <- x }
> mean(simlist)
[1] -0.9978892
```

From Example 9.11, the exact mean is

$$E(X_{10}) = \mu + (X_0 - \mu)e^{-r(10)} = -1 + 3e^{-5} = -0.9798.$$

Example 9.13 (Random genetic drift) The SDE

$$dX_t = \sqrt{X_t(1 - X_t)}\, dB_t, \text{ for } 0 \le t \le 1$$

arises as a model for *random genetic drift*. It is a continuous version of the discrete-time Wright–Fisher Markov chain introduced in Example 2.6. The latter is a model for the evolution of a population of N genes consisting of two alleles A and a. In the discrete-time process, the number of A alleles is obtained by drawing from replacement from the gene population. Given i A alleles at time n, the number of A alleles at time $n + 1$ has a binomial distribution with parameters $2N$ and $p = i/2N$. The Markov chain is absorbing with absorbing states 0 and $2N$.

The discrete-time process extends to a continuous-time diffusion $(X_t)_{t\ge 0}$ by a suitable scaling of time and space, where X_t denotes the *proportion* of A alleles in the gene population at time t. The diffusion process is absorbing with absorbing states 0 and 1.

Solving the SDE exactly is beyond the scope of this book. However, numerical methods are used (i) to approximate the sample paths of the process on the time interval $[0, 1.5]$ and (ii) to simulate the probability density function of X_t, for $t = 0.1, 0.2, 0.4, 1$. See Figures 9.6 and 9.7.[2]

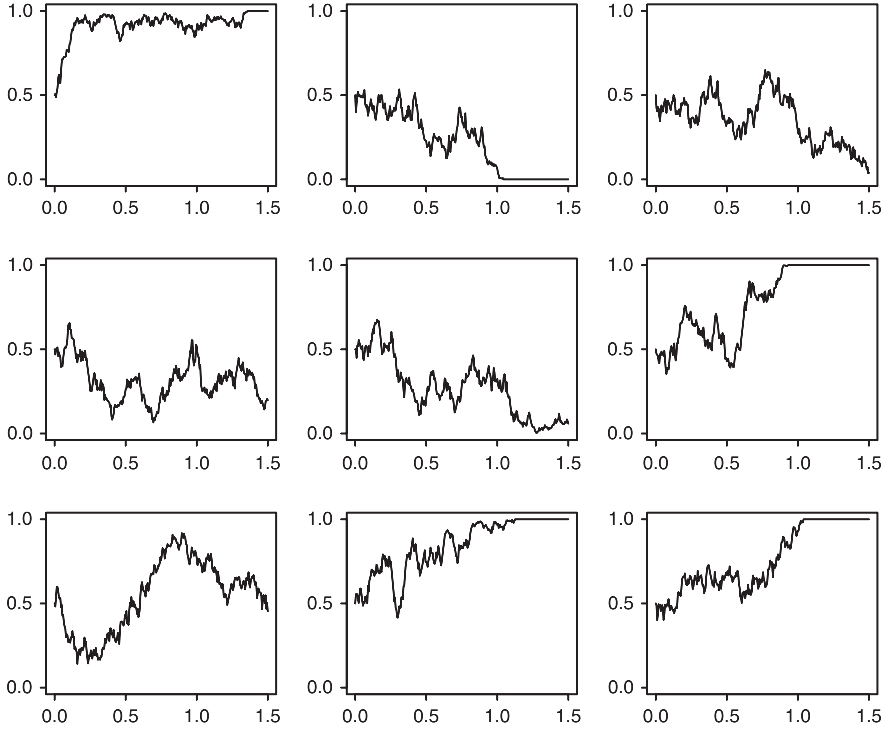

Figure 9.6 Sample paths for the solution of the random genetic drift SDE.

R: Euler–Maruyama Method for Simulating SDE

The following code generates the histograms in Figure 9.7.

```
# drift.R
> par(mfrow=c(2,2))
> times <- c( 0.1, 0.2,0.35,1)
> n = 100  # number of subintervals
> trials <- 10000
> for (k in 1:4) {
> t = times[k]
> simlist <- numeric(trials)
```

[2]In the Euler–Maruyama R code, to insure that the argument to the square foot function is non-negative, the absolute value of $x(1 - x)$ is taken. This gives an equivalent model to the original SDE.

```
> for (j in 1:trials) {
> x <- 1/2  # initial state
> for (i in 2:n) {
  x <- x + sqrt(abs(x*(1-x)))*sqrt(t/n)*rnorm(1)  }
> simlist[j] <- x }
> hist(simlist,freq=F,main="", col="gray") }
```

■

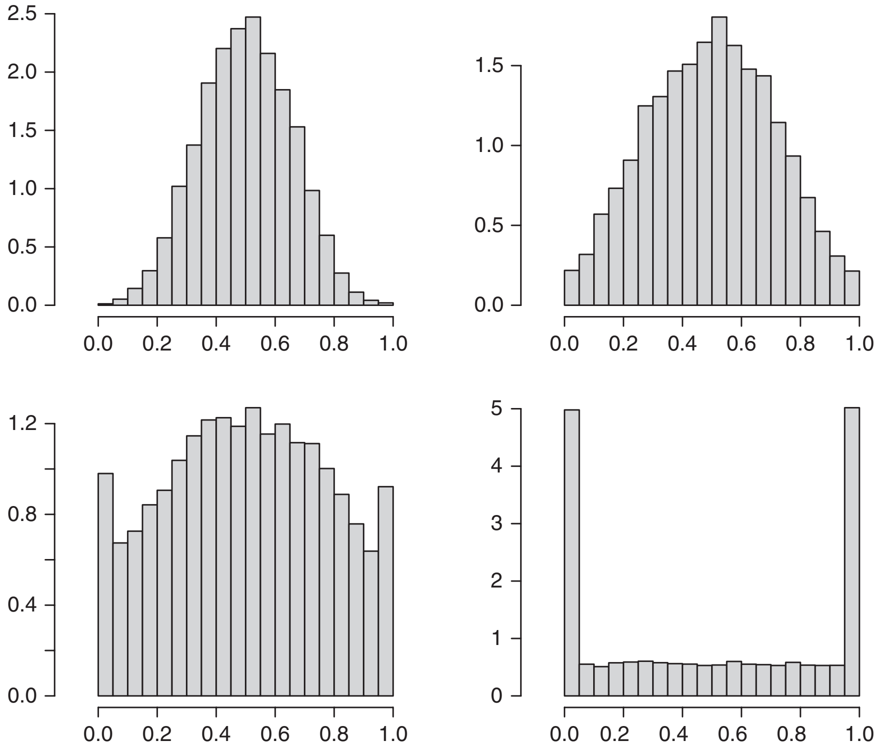

Figure 9.7 Simulating the distribution of X_t in the random genetic drift model, for $t = 0.1, 0.2, 0.35, 1.0$ (top-left to bottom-right).

Example 9.14 (Stochastic resonance) Stochastic resonance is a remarkable phenomenon whereby a signal, which is too weak to be detected, can be boosted by adding noise to the system. The idea is counter-intuitive, since we typically expect that noise (e.g., random error) makes signal detection more difficult. Yet the theory has found numerous applications over the past 25 years in biology, physics, and engineering, and has been demonstrated experimentally in the operation of ring lasers and in the neurons of crayfish.

The phenomenon was first introduced by Roberto Benzi in 1980 in the context of climate research, where it was proposed as a mechanism to explain how dramatic climactic events such as the almost periodic occurrence of the ice ages might be caused by minute changes in the earth's orbit around the sun. The theory has prompted discussions of whether rapid climate change is a hallmark of human impact (e.g., noise in the system).

As explained in Benzi (2010), stochastic resonance can be observed by considering the SDE

$$dX_t = (X_t - X_t^3 + A \sin t)dt + \sigma \, dB_t.$$

Think of the sinusoidal term, called a *periodic forcing*, as representing a weak, external signal, with amplitude A. We are interested in studying the effect of the noise parameter σ on detection of the forcing signal.

The process is simulated using the Euler–Maruyama method.

R : Stochastic Resonance

```
# stochresonance.R
> T <- 100
> n <- 10000
> A <- 0.3
> sigma <- 0.2
> w <- 2*pi/40
> xpath <- numeric(n+1)
> xpath[1]<- 0
> for (i in 2:(n+1)) {
+  x[i]<-x[i-1]+(x[i-1]-x[i-1]^3+A*sin(w*T*(i-1)/n))
+     *T/n+ sigma*sqrt(T/n)*rnorm(1)
> plot(seq(0,T,T/n),x,type="l",ylim=c(-2.8,2.8),
+     xaxt="n",xlab="",ylab="",yaxt="n",lwd=0.5)
> axis(2,c(-1,0,1))
> axis(1,c(0,25,50,75,100))
> curve(A*sin(w*x),0,100,lty=2,add=TRUE)
```

The process $(X_t)_{t\geq 0}$ has two *stable* points, at ± 1. For *small* σ (little noise), paths tend to stay near one of these values, although jumps may occur from one stable point to another. Three sample paths are shown in Figure 9.8 for $\sigma = 0.2$. The periodic forcing function (dashed curve) is not detectable. For this example, the amplitude of the sine function is $A = 0.3$, which is significantly smaller than the distance between the two stable points.

The effect of a relatively *large* random error term, with $\sigma = 2.0$, is apparent in Figure 9.9. The noise swamps any underlying structure. Again, the periodic forcing function is not detectable.

For Figure 9.10, an *optimal* value of σ is chosen at $\sigma = 0.8$. The *hidden* periodic forcing is now apparent. The added noise is sufficient for paths of the process to intersect with the range of the sine wave, which facilitates switching states. The system exhibits stochastic resonance.

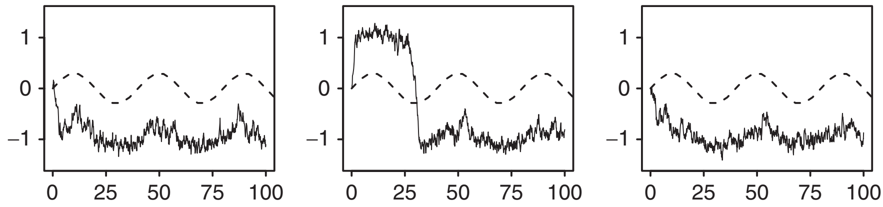

Figure 9.8

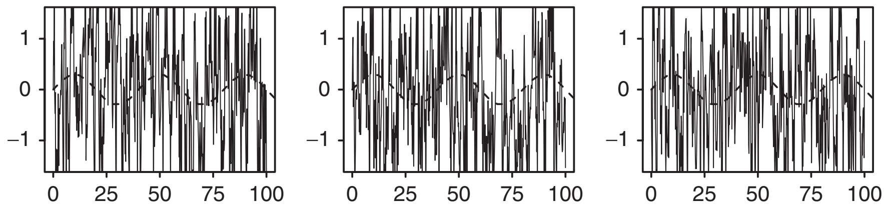

Figure 9.9

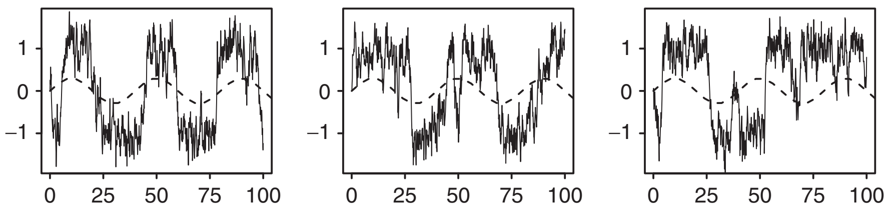

Figure 9.10

■

EXERCISES

9.1 Find the distribution of the stochastic integral $I_t = \int_0^t sB_s\, ds$.

9.2 Show that Brownian motion with drift coefficient μ and variance parameter σ^2 is a diffusion.

9.3 Find $E\left(B_t^4\right)$ by using Ito's Lemma to evaluate $d\left(B_t^4\right)$.

9.4 Use Ito's Lemma to show that

$$E\left(B_t^k\right) = \frac{k(k-1)}{2} \int_0^t E\left(B_s^{k-2}\right) ds, \text{ for } k \geq 2.$$

Use this result to find $E\left(B_t^k\right)$, for $k = 1, \ldots, 8$.

9.5 Use the methods of Example 9.6 to derive a martingale that is a fourth-degree polynomial function of Brownian motion.

9.6 Consider the stochastic differential equation

$$dX_t = (1 - 2X_t)dt + 3\ dB_t.$$

(a) Use Ito's Lemma to find $d(e^{rt}X_t)$.

(b) For suitable choice of r, simplify the drift coefficient in the resulting SDE. Solve the SDE and find the mean of X_t, and the asymptotic mean of the process.

9.7 Consider the SDE for the *square root process*

$$dX_t = dt + 2\sqrt{X_t}\ dB_t.$$

With $X_0 = x_0$, show that $X_t = (B_t + x_0)^2$ is a solution.

9.8 R : Show how to use the Euler–Maruyama method to simulate geometric Brownian motion started at $G_0 = 8$, with $\mu = 1$ and $\sigma^2 = 0.25$.

(a) Generate a plot of a sample path on $[0, 2]$.

(b) Simulate the mean and variance of G_2. Compare with the theoretical mean and variance.

9.9 R Use the Euler–Maruyama method to simulate the square root process of Exercise 9.7 with $x_0 = 1$.

(a) Estimate $E(X_3)$, $Var(X_3)$, and $P(X_3 < 5)$.

(b) Using the fact that $X_t = (B_t + x_0)^2$ is a solution to the SDE, compare your simulations in (a) with the exact results.

9.10 R: The random drift model of Example 9.13 is an absorbing process with two absorbing states. Use the Euler-Marayuma method to estimate the expectation and standard deviation of the time until absorption.

9.11 R : The *Cox–Ingersoll–Ross (CIR) model*

$$dX_t = -r(X_t - \mu)dt + \sigma\sqrt{X_t}dB_t$$

has been used to describe the evolution of interest rates. The diffusion has the same drift coefficient as the Ornstein–Uhlenbeck process and is also mean-reverting. The CIR model has the advantage over the Ornstein–Uhlenbeck process as a model for interest rates since, unlike the latter, the process is non-negative. However, unlike that process, the CIR model has no closed-form solution.

With $X_0 = 0$, $\mu = 1.25$, $r = 2$, and $\sigma = 0.2$, simulate the CIR process. Estimate the asymptotic mean and variance by taking $t = 100$. Demonstrate that the process is mean-reverting.