OPTIONS, FUTURES, AND OTHER DERIVATIVES

NINTH EDITION

JOHN C. HULL

CHAPTER 14 Wiener Processes and Itô's Lemma

Any variable whose value changes over time in an uncertain way is said to follow a *stochastic process*. Stochastic processes can be classified as *discrete time* or *continuous time*. A discrete-time stochastic process is one where the value of the variable can change only at certain fixed points in time, whereas a continuous-time stochastic process is one where changes can take place at any time. Stochastic processes can also be classified as *continuous variable* or *discrete variable*. In a continuous-variable process, the underlying variable can take any value within a certain range, whereas in a discrete-variable process, only certain discrete values are possible.

This chapter develops a continuous-variable, continuous-time stochastic process for stock prices. Learning about this process is the first step to understanding the pricing of options and other more complicated derivatives. It should be noted that, in practice, we do not observe stock prices following continuous-variable, continuous-time processes. Stock prices are restricted to discrete values (e.g., multiples of a cent) and changes can be observed only when the exchange is open for trading. Nevertheless, the continuous-variable, continuous-time process proves to be a useful model for many purposes.

Many people feel that continuous-time stochastic processes are so complicated that they should be left entirely to "rocket scientists." This is not so. The biggest hurdle to understanding these processes is the notation. Here we present a step-by-step approach aimed at getting the reader over this hurdle. We also explain an important result known as *Itô's lemma* that is central to the pricing of derivatives.

14.1 THE MARKOV PROPERTY

A *Markov process* is a particular type of stochastic process where only the current value of a variable is relevant for predicting the future. The past history of the variable and the way that the present has emerged from the past are irrelevant.

Stock prices are usually assumed to follow a Markov process. Suppose that the price of a stock is \$100 now. If the stock price follows a Markov process, our predictions for the future should be unaffected by the price one week ago, one month ago, or one year ago. The only relevant piece of information is that the price is now \$100.[1] Predictions for the future are uncertain and must be expressed in terms of probability distributions. The Markov property implies that the probability distribution of the price at any particular future time is not dependent on the particular path followed by the price in the past.

The Markov property of stock prices is consistent with the weak form of market efficiency. This states that the present price of a stock impounds all the information contained in a record of past prices. If the weak form of market efficiency were not true, technical analysts could make above-average returns by interpreting charts of the past history of stock prices. There is very little evidence that they are in fact able to do this.

It is competition in the marketplace that tends to ensure that weak-form market efficiency and the Markov property hold. There are many investors watching the stock market closely. This leads to a situation where a stock price, at any given time, reflects the information in past prices. Suppose that it was discovered that a particular pattern in a stock price always gave a 65% chance of subsequent steep price rises. Investors would attempt to buy a stock as soon as the pattern was observed, and demand for the stock would immediately rise. This would lead to an immediate rise in its price and the observed effect would be eliminated, as would any profitable trading opportunities.

14.2 CONTINUOUS-TIME STOCHASTIC PROCESSES

Consider a variable that follows a Markov stochastic process. Suppose that its current value is 10 and that the change in its value during a year is $\phi(0, 1)$, where $\phi(m, v)$ denotes a probability distribution that is normally distributed with mean m and variance v.[2] What is the probability distribution of the change in the value of the variable during 2 years?

The change in 2 years is the sum of two normal distributions, each of which has a mean of zero and variance of 1.0. Because the variable is Markov, the two probability distributions are independent. When we add two independent normal distributions, the result is a normal distribution where the mean is the sum of the means and the variance is the sum of the variances. The mean of the change during 2 years in the variable we are considering is, therefore, zero and the variance of this change is 2.0. Hence, the change in the variable over 2 years has the distribution $\phi(0, 2)$. The standard deviation of the change is $\sqrt{2}$.

Consider next the change in the variable during 6 months. The variance of the change in the value of the variable during 1 year equals the variance of the change during the first 6 months plus the variance of the change during the second 6 months. We assume these are the same. It follows that the variance of the change during a 6-month period must be 0.5. Equivalently, the standard deviation of the change is $\sqrt{0.5}$. The probability distribution for the change in the value of the variable during 6 months is $\phi(0, 0.5)$.

---

[1] Statistical properties of the stock price history may be useful in determining the characteristics of the stochastic process followed by the stock price (e.g., its volatility). The point being made here is that the particular path followed by the stock in the past is irrelevant.

[2] Variance is the square of standard deviation. The standard deviation of a 1-year change in the value of the variable we are considering is therefore 1.0.

A similar argument shows that the probability distribution for the change in the value of the variable during 3 months is $\phi(0, 0.25)$. More generally, the change during any time period of length T is $\phi(0, T)$. In particular, the change during a very short time period of length Δt is $\phi(0, \Delta t)$.

Note that, when Markov processes are considered, the variances of the changes in successive time periods are additive. The standard deviations of the changes in successive time periods are not additive. The variance of the change in the variable in our example is 1.0 per year, so that the variance of the change in 2 years is 2.0 and the variance of the change in 3 years is 3.0. The standard deviations of the changes in 2 and 3 years are $\sqrt{2}$ and $\sqrt{3}$, respectively. (Strictly speaking, we should not refer to the standard deviation of the variable as 1.0 per year.) The results explain why uncertainty is sometimes referred to as being proportional to the square root of time.

Wiener Process

The process followed by the variable we have been considering is known as a *Wiener process*. It is a particular type of Markov stochastic process with a mean change of zero and a variance rate of 1.0 per year. It has been used in physics to describe the motion of a particle that is subject to a large number of small molecular shocks and is sometimes referred to as *Brownian motion*.

Expressed formally, a variable z follows a Wiener process if it has the following two properties:

Property 1. *The change Δz during a small period of time Δt is*

$$\Delta z = \epsilon\sqrt{\Delta t} \tag{14.1}$$

where ϵ has a standard normal distribution $\phi(0, 1)$.

Property 2. *The values of Δz for any two different short intervals of time, Δt, are independent.*

It follows from the first property that Δz itself has a normal distribution with

$$\text{mean of } \Delta z = 0$$
$$\text{standard deviation of } \Delta z = \sqrt{\Delta t}$$
$$\text{variance of } \Delta z = \Delta t$$

The second property implies that z follows a Markov process.

Consider the change in the value of z during a relatively long period of time, T. This can be denoted by $z(T) - z(0)$. It can be regarded as the sum of the changes in z in N small time intervals of length Δt, where

$$N = \frac{T}{\Delta t}$$

Thus,

$$z(T) - z(0) = \sum_{i=1}^{N} \epsilon_i \sqrt{\Delta t} \tag{14.2}$$

where the ϵ_i $(i = 1, 2, \dots, N)$ are distributed $\phi(0, 1)$. We know from the second property of Wiener processes that the ϵ_i are independent of each other. It follows from equation (14.2) that $z(T) - z(0)$ is normally distributed, with

$$\text{mean of } [z(T) - z(0)] = 0$$
$$\text{variance of}[z(T) - z(0)] = N\,\Delta t = T$$
$$\text{standard deviation of } [z(T) - z(0)] = \sqrt{T}$$

This is consistent with the discussion earlier in this section.

Example 14.1

Suppose that the value, z, of a variable that follows a Wiener process is initially 25 and that time is measured in years. At the end of 1 year, the value of the variable is normally distributed with a mean of 25 and a standard deviation of 1.0. At the end of 5 years, it is normally distributed with a mean of 25 and a standard deviation of $\sqrt{5}$, or 2.236. Our uncertainty about the value of the variable at a certain time in the future, as measured by its standard deviation, increases as the square root of how far we are looking ahead.

In ordinary calculus, it is usual to proceed from small changes to the limit as the small changes become closer to zero. Thus, $dx = a\,dt$ is the notation used to indicate that $\Delta x = a\,\Delta t$ in the limit as $\Delta t \to 0$. We use similar notational conventions in stochastic calculus. So, when we refer to dz as a Wiener process, we mean that it has the properties for Δz given above in the limit as $\Delta t \to 0$.

Figure 14.1 illustrates what happens to the path followed by z as the limit $\Delta t \to 0$ is approached. Note that the path is quite "jagged." This is because the standard deviation of the movement in z in time Δt equals $\sqrt{\Delta t}$ and, when Δt is small, $\sqrt{\Delta t}$ is much bigger than Δt. Two intriguing properties of Wiener processes, related to this $\sqrt{\Delta t}$ property, are as follows:

1. The expected length of the path followed by z in any time interval is infinite.
2. The expected number of times z equals any particular value in any time interval is infinite.[3]

Generalized Wiener Process

The mean change per unit time for a stochastic process is known as the *drift rate* and the variance per unit time is known as the *variance rate*. The basic Wiener process, dz, that has been developed so far has a drift rate of zero and a variance rate of 1.0. The drift rate of zero means that the expected value of z at any future time is equal to its current value. The variance rate of 1.0 means that the variance of the change in z in a time interval of length T equals T. A *generalized Wiener process* for a variable x can be defined in terms of dz as

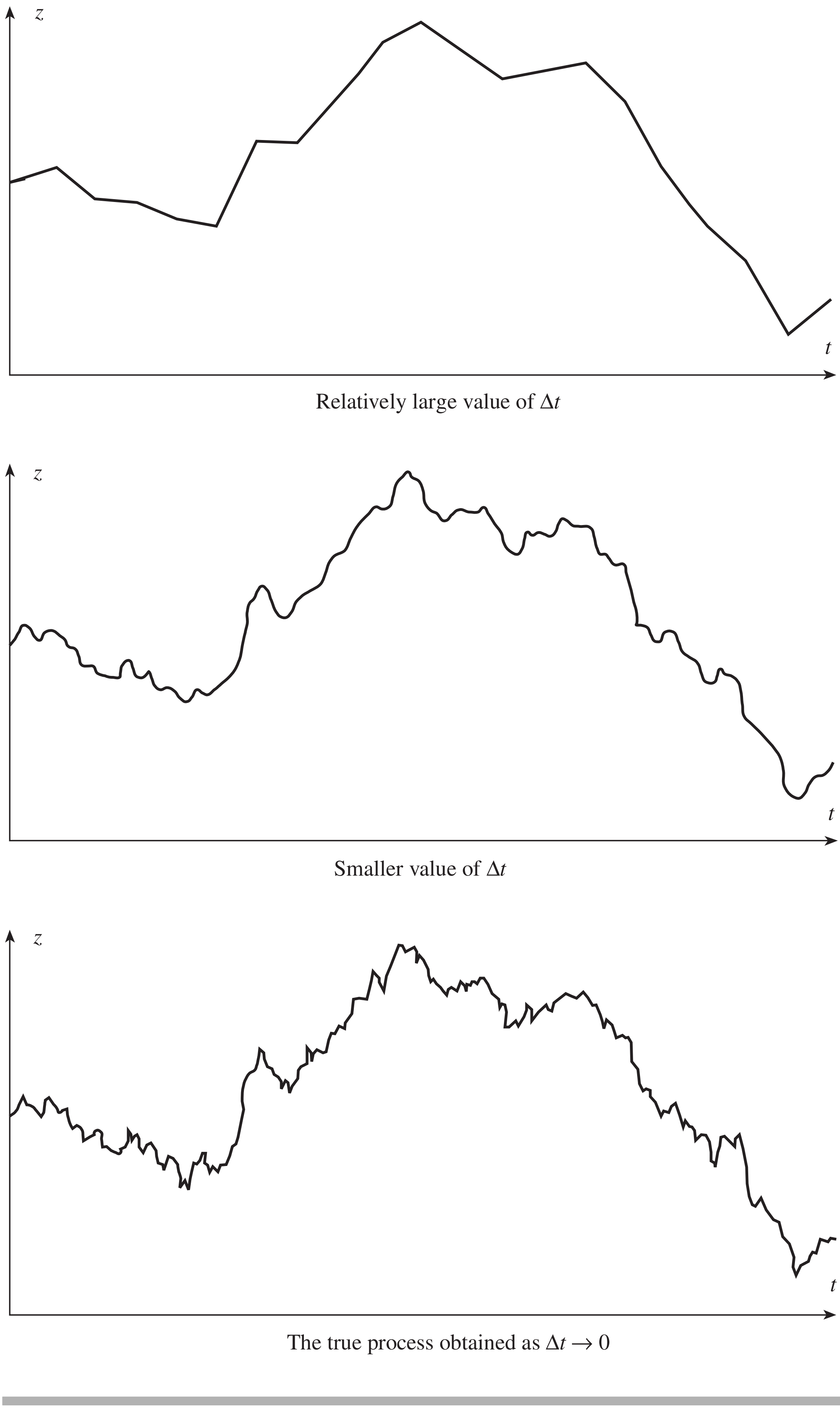
$$dx = a\,dt + b\,dz \tag{14.3}$$

where a and b are constants.

To understand equation (14.3), it is useful to consider the two components on the right-hand side separately. The $a\,dt$ term implies that x has an expected drift rate of a per unit of time. Without the $b\,dz$ term, the equation is $dx = a\,dt$, which implies that

[3] This is because z has some nonzero probability of equaling any value v in the time interval. If it equals v in time t, the expected number of times it equals v in the immediate vicinity of t is infinite.

Figure 14.1 How a Wiener process is obtained when $\Delta t \to 0$ in equation (14.1).



$dx/dt = a$. Integrating with respect to time, we get

$$x = x_0 + at$$

where x_0 is the value of x at time 0. In a period of time of length T, the variable x increases by an amount aT. The $b\,dz$ term on the right-hand side of equation (14.3) can be regarded as adding noise or variability to the path followed by x. The amount of this noise or variability is b times a Wiener process. A Wiener process has a variance rate per unit time of 1.0. It follows that b times a Wiener process has a variance rate per unit time of b^2. In a small time interval Δt, the change Δx in the value of x is given by equations (14.1) and (14.3) as

$$\Delta x = a\,\Delta t + b\epsilon\sqrt{\Delta t}$$

where, as before, ϵ has a standard normal distribution $\phi(0, 1)$. Thus Δx has a normal distribution with

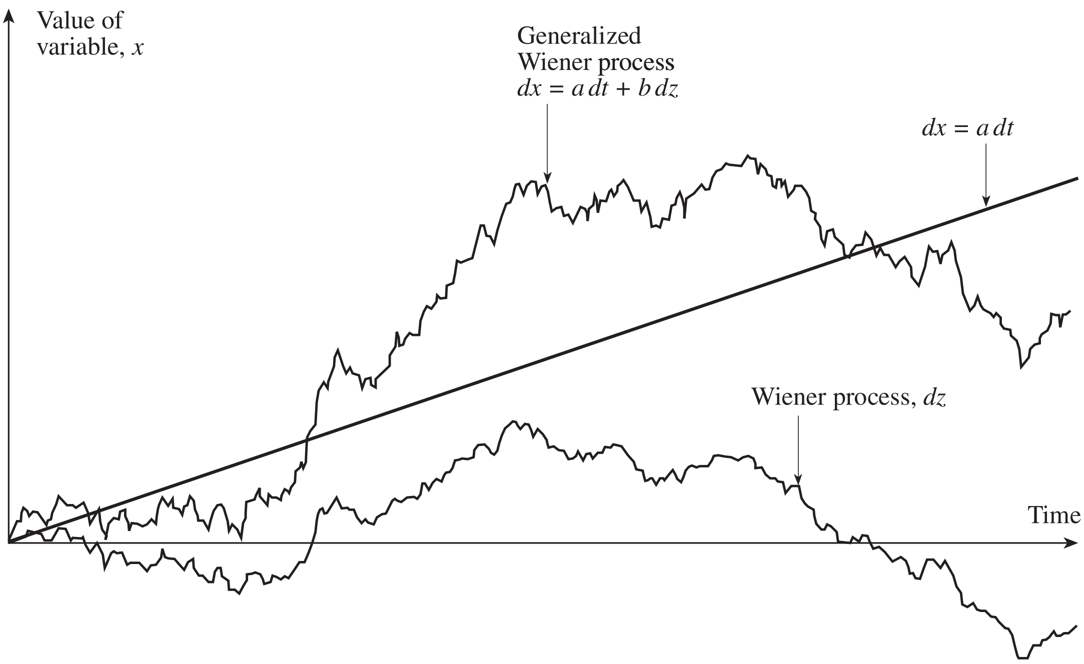
$$\text{mean of } \Delta x = a\,\Delta t$$
$$\text{standard deviation of } \Delta x = b\sqrt{\Delta t}$$
$$\text{variance of } \Delta x = b^2\Delta t$$

Similar arguments to those given for a Wiener process show that the change in the value of x in any time interval T is normally distributed with

$$\text{mean of change in } x = aT$$
$$\text{standard deviation of change in } x = b\sqrt{T}$$
$$\text{variance of change in } x = b^2T$$

To summarize, the generalized Wiener process given in equation (14.3) has an expected drift rate (i.e., average drift per unit of time) of a and a variance rate (i.e., variance per unit of time) of b^2. It is illustrated in Figure 14.2.

Figure 14.2 Generalized Wiener process with $a = 0.3$ and $b = 1.5$.



Example 14.2

Consider the situation where the cash position of a company, measured in thousands of dollars, follows a generalized Wiener process with a drift of 20 per year and a variance rate of 900 per year. Initially, the cash position is 50. At the end of 1 year the cash position will have a normal distribution with a mean of 70 and a standard deviation of $\sqrt{900}$, or 30. At the end of 6 months it will have a normal distribution with a mean of 60 and a standard deviation of $30\sqrt{0.5} = 21.21$. Our uncertainty about the cash position at some time in the future, as measured by its standard deviation, increases as the square root of how far ahead we are looking. (Note that the cash position can become negative. We can interpret this as a situation where the company is borrowing funds.)

Itô Process

A further type of stochastic process, known as an *Itô process*, can be defined. This is a generalized Wiener process in which the parameters a and b are functions of the value of the underlying variable x and time t. An Itô process can therefore be written as

$$dx = a(x, t)\, dt + b(x, t)\, dz \tag{14.4}$$

Both the expected drift rate and variance rate of an Itô process are liable to change over time. In a small time interval between t and $t + \Delta t$, the variable changes from x to $x + \Delta x$, where

$$\Delta x = a(x, t)\Delta t + b(x, t)\epsilon\sqrt{\Delta t}$$

This equation involves a small approximation. It assumes that the drift and variance rate of x remain constant, equal to their values at time t, during the time interval between t and $t + \Delta t$.

Note that the process in equation (14.4) is Markov because the change in x at time t depends only on the value of x at time t, not on its history. A non-Markov process could be defined by letting a and b in equation (14.4) depend on values of x prior to time t.

14.3 THE PROCESS FOR A STOCK PRICE

In this section we discuss the stochastic process usually assumed for the price of a non-dividend-paying stock.

It is tempting to suggest that a stock price follows a generalized Wiener process; that is, that it has a constant expected drift rate and a constant variance rate. However, this model fails to capture a key aspect of stock prices. This is that the expected percentage return required by investors from a stock is independent of the stock's price. If investors require a 14% per annum expected return when the stock price is \$10, then, *ceteris paribus*, they will also require a 14% per annum expected return when it is \$50.

Clearly, the assumption of constant expected drift rate is inappropriate and needs to be replaced by the assumption that the expected return (i.e., expected drift divided by the stock price) is constant. If S is the stock price at time t, then the expected drift rate in S should be assumed to be μS for some constant parameter μ. This means that in a short interval of time, Δt, the expected increase in S is $\mu S\, \Delta t$. The parameter μ is the expected rate of return on the stock.

If the coefficient of dz is zero, so that there is no uncertainty, then this model implies that

$$\Delta S = \mu S \, \Delta t$$

In the limit, as $\Delta t \to 0$,

$$dS = \mu S \, dt$$

or

$$\frac{dS}{S} = \mu \, dt$$

Integrating between time 0 and time T, we get

$$S_T = S_0 e^{\mu T} \tag{14.5}$$

where S_0 and S_T are the stock price at time 0 and time T. Equation (14.5) shows that, when there is no uncertainty, the stock price grows at a continuously compounded rate of μ per unit of time.

In practice, of course, there is uncertainty. A reasonable assumption is that the variability of the return in a short period of time, Δt, is the same regardless of the stock price. In other words, an investor is just as uncertain of the return when the stock price is \$50 as when it is \$10. This suggests that the standard deviation of the change in a short period of time Δt should be proportional to the stock price and leads to the model

$$dS = \mu S \, dt + \sigma S \, dz$$

or

$$\frac{dS}{S} = \mu \, dt + \sigma \, dz \tag{14.6}$$

Equation (14.6) is the most widely used model of stock price behavior. The variable μ is the stock's expected rate of return. The variable σ is the volatility of the stock price. The variable σ^2 is referred to as its variance rate. The model in equation (14.6) represents the stock price process in the real world. In a risk-neutral world, μ equals the risk-free rate r.

Discrete-Time Model

The model of stock price behavior we have developed is known as *geometric Brownian motion*. The discrete-time version of the model is

$$\frac{\Delta S}{S} = \mu \, \Delta t + \sigma \epsilon \sqrt{\Delta t} \tag{14.7}$$

or

$$\Delta S = \mu S \, \Delta t + \sigma S \epsilon \sqrt{\Delta t} \tag{14.8}$$

The variable ΔS is the change in the stock price S in a small time interval Δt, and as before ϵ has a standard normal distribution (i.e., a normal distribution with a mean of zero and standard deviation of 1.0). The parameter μ is the expected rate of return per unit of time from the stock. The parameter σ is the volatility of the stock price. In this chapter we will assume these parameters are constant.

The left-hand side of equation (14.7) is the discrete approximation to the return provided by the stock in a short period of time, Δt. The term $\mu \, \Delta t$ is the expected value of this return, and the term $\sigma \epsilon \sqrt{\Delta t}$ is the stochastic component of the return. The variance of the stochastic component (and, therefore, of the whole return) is $\sigma^2 \Delta t$. This is consistent with the definition of the volatility σ given in Section 13.7; that is, σ is such that $\sigma\sqrt{\Delta t}$ is the standard deviation of the return in a short time period Δt.

Equation (14.7) shows that $\Delta S/S$ is approximately normally distributed with mean $\mu\,\Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$. In other words,

$$\frac{\Delta S}{S} \sim \phi(\mu\,\Delta t,\ \sigma^2 \Delta t) \qquad \textbf{(14.9)}$$

Example 14.3

Consider a stock that pays no dividends, has a volatility of 30% per annum, and provides an expected return of 15% per annum with continuous compounding. In this case, $\mu = 0.15$ and $\sigma = 0.30$. The process for the stock price is

$$\frac{dS}{S} = 0.15\,dt + 0.30\,dz$$

If S is the stock price at a particular time and ΔS is the increase in the stock price in the next small interval of time, the discrete approximation to the process is

$$\frac{\Delta S}{S} = 0.15\Delta t + 0.30\epsilon\sqrt{\Delta t}$$

where ϵ has a standard normal distribution. Consider a time interval of 1 week, or 0.0192 year, so that $\Delta t = 0.0192$. Then the approximation gives

$$\frac{\Delta S}{S} = 0.15 \times 0.0192 + 0.30 \times \sqrt{0.0192}\,\epsilon$$

or

$$\Delta S = 0.00288S + 0.0416S\epsilon$$

Monte Carlo Simulation

A Monte Carlo simulation of a stochastic process is a procedure for sampling random outcomes for the process. We will use it as a way of developing some understanding of the nature of the stock price process in equation (14.6).

Consider the situation in Example 14.3 where the expected return from a stock is 15% per annum and the volatility is 30% per annum. The stock price change over 1 week was shown to be approximately

$$\Delta S = 0.00288S + 0.0416S\epsilon \qquad \textbf{(14.10)}$$

A path for the stock price over 10 weeks can be simulated by sampling repeatedly for ϵ from $\phi(0, 1)$ and substituting into equation (14.10). The expression =RAND( ) in Excel produces a random sample between 0 and 1. The inverse cumulative normal distribution is NORMSINV. The instruction to produce a random sample from a standard normal distribution in Excel is therefore =NORMSINV(RAND( )). Table 14.1 shows one path for a stock price that was sampled in this way. The initial stock price is assumed to be \$100. For the first period, ϵ is sampled as 0.52. From equation (14.10), the change during the first time period is

$$\Delta S = 0.00288 \times 100 + 0.0416 \times 100 \times 0.52 = 2.45$$

Therefore, at the beginning of the second time period, the stock price is \$102.45. The

Table 14.1 Simulation of stock price when $\mu = 0.15$ and $\sigma = 0.30$ during 1-week periods.

Stock price at start of period	*Random sample for ϵ*	*Change in stock price during period*
100.00	0.52	2.45
102.45	1.44	6.43
108.88	−0.86	−3.58
105.30	1.46	6.70
112.00	−0.69	−2.89
109.11	−0.74	−3.04
106.06	0.21	1.23
107.30	−1.10	−4.60
102.69	0.73	3.41
106.11	1.16	5.43
111.54	2.56	12.20

value of ϵ sampled for the next period is 1.44. From equation (14.10), the change during the second time period is

$$\Delta S = 0.00288 \times 102.45 + 0.0416 \times 102.45 \times 1.44 = 6.43$$

So, at the beginning of the next period, the stock price is \$108.88, and so on.[4] Note that, because the process we are simulating is Markov, the samples for ϵ should be independent of each other.

Table 14.1 assumes that stock prices are measured to the nearest cent. It is important to realize that the table shows only one possible pattern of stock price movements. Different random samples would lead to different price movements. Any small time interval Δt can be used in the simulation. In the limit as $\Delta t \to 0$, a perfect description of the stochastic process is obtained. The final stock price of 111.54 in Table 14.1 can be regarded as a random sample from the distribution of stock prices at the end of 10 weeks. By repeatedly simulating movements in the stock price, a complete probability distribution of the stock price at the end of this time is obtained. Monte Carlo simulation is discussed in more detail in Chapter 21.

14.4 THE PARAMETERS

The process for a stock price developed in this chapter involves two parameters, μ and σ. The parameter μ is the expected return (annualized) earned by an investor in a short period of time. Most investors require higher expected returns to induce them to take higher risks. It follows that the value of μ should depend on the risk of the return from the stock.[5] It should also depend on the level of interest rates in the economy. The higher the level of interest rates, the higher the expected return required on any given stock.

[4] In practice, it is more efficient to sample $\ln S$ rather than S, as will be discussed in Section 21.6.

[5] More precisely, μ depends on that part of the risk that cannot be diversified away by the investor.

Fortunately, we do not have to concern ourselves with the determinants of μ in any detail because the value of a derivative dependent on a stock is, in general, independent of μ. The parameter σ, the stock price volatility, is, by contrast, critically important to the determination of the value of many derivatives. We will discuss procedures for estimating σ in Chapter 15. Typical values of σ for a stock are in the range 0.15 to 0.60 (i.e., 15% to 60%).

The standard deviation of the proportional change in the stock price in a small interval of time Δt is $\sigma\sqrt{\Delta t}$. As a rough approximation, the standard deviation of the proportional change in the stock price over a relatively long period of time T is $\sigma\sqrt{T}$. This means that, as an approximation, volatility can be interpreted as the standard deviation of the change in the stock price in 1 year. In Chapter 15, we will show that the volatility of a stock price is exactly equal to the standard deviation of the continuously compounded return provided by the stock in 1 year.

14.5 CORRELATED PROCESSES

So far we have considered how the stochastic process for a single variable can be represented. We now extend the analysis to the situation where there are two or more variables following correlated stochastic processes. Suppose that the processes followed by two variables x_1 and x_2 are

$$dx_1 = a_1\,dt + b_1\,dz_1 \quad \text{and} \quad dx_2 = a_2\,dt + b_2\,dz_2$$

where dz_1 and dz_2 are Wiener processes.

As has been explained, the discrete-time approximations for these processes are

$$\Delta x_1 = a_1\,\Delta t + b_1\,\epsilon_1\sqrt{\Delta t} \quad \text{and} \quad \Delta x_2 = a_2\,\Delta t + b_2\,\epsilon_2\sqrt{\Delta t}$$

where ϵ_1 and ϵ_2 are samples from a standard normal distribution $\phi(0, 1)$.

The variables x_1 and x_2 can be simulated in the way described in Section 14.3. If they are uncorrelated with each other, the random samples ϵ_1 and ϵ_2 that are used to obtain movements in a particular period of time Δt should be independent of each other.

If x_1 and x_2 have a nonzero correlation ρ, then the ϵ_1 and ϵ_2 that are used to obtain movements in a particular period of time should be sampled from a bivariate normal distribution. Each variable in the bivariate normal distribution has a standard normal distribution and the correlation between the variables is ρ. In this situation, we would refer to the Wiener processes dz_1 and dz_2 as having a correlation ρ.

Obtaining samples for uncorrelated standard normal variables in cells in Excel involves putting the instruction "=NORMSINV(RAND))" in each of the cells. To sample standard normal variables ϵ_1 and ϵ_2 with correlation ρ, we can set

$$\epsilon_1 = u \quad \text{and} \quad \epsilon_2 = \rho u + \sqrt{1-\rho^2}\,v$$

where u and v are sampled as uncorrelated variables with standard normal distributions.

Note that, in the processes we have assumed for x_1 and x_2, the parameters a_1, a_2, b_1, and b_2 can be functions of x_1, x_2, and t. In particular, a_1 and b_1 can be functions of x_2 as well as x_1 and t; and a_2 and b_2 can be functions of x_1 as well as x_2 and t.

The results here can be generalized. When there are three different variables following correlated stochastic processes, we have to sample three different ϵ's. These have a trivariate normal distribution. When there are n correlated variables, we have n different ϵ's and these must be sampled from an appropriate multivariate normal distribution. The way this is done is explained in Chapter 21.

14.6 ITÔ'S LEMMA

The price of a stock option is a function of the underlying stock's price and time. More generally, we can say that the price of any derivative is a function of the stochastic variables underlying the derivative and time. A serious student of derivatives must, therefore, acquire some understanding of the behavior of functions of stochastic variables. An important result in this area was discovered by the mathematician K. Itô in 1951,[6] and is known as *Itô's lemma*.

Suppose that the value of a variable x follows the Itô process

$$dx = a(x, t)\,dt + b(x, t)\,dz \tag{14.11}$$

where dz is a Wiener process and a and b are functions of x and t. The variable x has a drift rate of a and a variance rate of b^2. Itô's lemma shows that a function G of x and t follows the process

$$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right)dt + \frac{\partial G}{\partial x}b\,dz \tag{14.12}$$

where the dz is the same Wiener process as in equation (14.11). Thus, G also follows an Itô process, with a drift rate of

$$\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2$$

and a variance rate of

$$\left(\frac{\partial G}{\partial x}\right)^2 b^2$$

A completely rigorous proof of Itô's lemma is beyond the scope of this book. In the appendix to this chapter, we show that the lemma can be viewed as an extension of well-known results in differential calculus.

Earlier, we argued that

$$dS = \mu S\,dt + \sigma S\,dz \tag{14.13}$$

with μ and σ constant, is a reasonable model of stock price movements. From Itô's lemma, it follows that the process followed by a function G of S and t is

$$dG = \left(\frac{\partial G}{\partial S}\mu S + \frac{\partial G}{\partial t} + \tfrac{1}{2}\frac{\partial^2 G}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial G}{\partial S}\sigma S\,dz \tag{14.14}$$

Note that both S and G are affected by the same underlying source of uncertainty, dz. This proves to be very important in the derivation of the Black–Scholes–Merton results.

[6] See K. Itô, "On Stochastic Differential Equations," *Memoirs of the American Mathematical Society*, 4 (1951): 1–51.

Application to Forward Contracts

To illustrate Itô's lemma, consider a forward contract on a non-dividend-paying stock. Assume that the risk-free rate of interest is constant and equal to r for all maturities. From equation (5.1),

$$F_0 = S_0 e^{rT}$$

where F_0 is the forward price at time zero, S_0 is the spot price at time zero, and T is the time to maturity of the forward contract.

We are interested in what happens to the forward price as time passes. We define F as the forward price at a general time t, and S as the stock price at time t, with $t < T$. The relationship between F and S is given by

$$F = Se^{r(T-t)} \tag{14.15}$$

Assuming that the process for S is given by equation (14.13), we can use Itô's lemma to determine the process for F. From equation (14.15),

$$\frac{\partial F}{\partial S} = e^{r(T-t)}, \qquad \frac{\partial^2 F}{\partial S^2} = 0, \qquad \frac{\partial F}{\partial t} = -rSe^{r(T-t)}$$

From equation (14.14), the process for F is given by

$$dF = \left[e^{r(T-t)}\mu S - rSe^{r(T-t)}\right] dt + e^{r(T-t)}\sigma S\, dz$$

Substituting F for $Se^{r(T-t)}$ gives

$$dF = (\mu - r)F\, dt + \sigma F\, dz \tag{14.16}$$

Like S, the forward price F follows geometric Brownian motion. It has an expected growth rate of $\mu - r$ rather than μ. The growth rate in F is the excess return of S over the risk-free rate.

14.7 THE LOGNORMAL PROPERTY

We now use Itô's lemma to derive the process followed by $\ln S$ when S follows the process in equation (14.13). We define

$$G = \ln S$$

Since

$$\frac{\partial G}{\partial S} = \frac{1}{S}, \qquad \frac{\partial^2 G}{\partial S^2} = -\frac{1}{S^2}, \qquad \frac{\partial G}{\partial t} = 0$$

it follows from equation (14.14) that the process followed by G is

$$dG = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\, dz \tag{14.17}$$

Since μ and σ are constant, this equation indicates that $G = \ln S$ follows a generalized Wiener process. It has constant drift rate $\mu - \sigma^2/2$ and constant variance rate σ^2. The change in $\ln S$ between time 0 and some future time T is therefore normally distributed, with mean $(\mu - \sigma^2/2)T$ and variance $\sigma^2 T$. This means that

$$\ln S_T - \ln S_0 \sim \phi\left[\left(\mu - \frac{\sigma^2}{2}\right)T,\ \sigma^2 T\right] \tag{14.18}$$

or

$$\ln S_T \sim \phi\left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T,\ \sigma^2 T\right] \tag{14.19}$$

where S_T is the stock price at time T, S_0 is the stock price at time 0, and as before $\phi(m, v)$ denotes a normal distribution with mean m and variance v.

Equation (14.19) shows that $\ln S_T$ is normally distributed. A variable has a lognormal distribution if the natural logarithm of the variable is normally distributed. The model of stock price behavior we have developed in this chapter therefore implies that a stock's price at time T, given its price today, is lognormally distributed. The standard deviation of the logarithm of the stock price is $\sigma\sqrt{T}$. It is proportional to the square root of how far ahead we are looking.

SUMMARY

Stochastic processes describe the probabilistic evolution of the value of a variable through time. A Markov process is one where only the present value of the variable is relevant for predicting the future. The past history of the variable and the way in which the present has emerged from the past is irrelevant.

A Wiener process dz is a Markov process describing the evolution of a normally distributed variable. The drift of the process is zero and the variance rate is 1.0 per unit time. This means that, if the value of the variable is x_0 at time 0, then at time T it is normally distributed with mean x_0 and standard deviation $\sqrt{T}$.

A generalized Wiener process describes the evolution of a normally distributed variable with a drift of a per unit time and a variance rate of b^2 per unit time, where a and b are constants. This means that if, as before, the value of the variable is x_0 at time 0, it is normally distributed with a mean of $x_0 + aT$ and a standard deviation of $b\sqrt{T}$ at time T.

An Itô process is a process where the drift and variance rate of x can be a function of both x itself and time. The change in x in a very short period of time is, to a good approximation, normally distributed, but its change over longer periods of time is liable to be nonnormal.

One way of gaining an intuitive understanding of a stochastic process for a variable is to simulate the behavior of the variable. This involves dividing a time interval into many small time steps and randomly sampling possible paths for the variable. The future probability distribution for the variable can then be calculated. Monte Carlo simulation is discussed further in Chapter 21.

Itô's lemma is a way of calculating the stochastic process followed by a function of a variable from the stochastic process followed by the variable itself. As we shall see in Chapter 15, Itô's lemma plays a very important part in the pricing of derivatives. A key point is that the Wiener process dz underlying the stochastic process for the variable is exactly the same as the Wiener process underlying the stochastic process for the function of the variable. Both are subject to the same underlying source of uncertainty.

The stochastic process usually assumed for a stock price is geometric Brownian motion. Under this process the return to the holder of the stock in a small period of time is normally distributed and the returns in two nonoverlapping periods are independent. The value of the stock price at a future time has a lognormal distribution. The Black–Scholes–Merton model, which we cover in the next chapter, is based on the geometric Brownian motion assumption.

FURTHER READING

On Efficient Markets and the Markov Property of Stock Prices

Brealey, R. A. *An Introduction to Risk and Return from Common Stock*, 2nd edn. Cambridge, MA: MIT Press, 1986.

Cootner, P. H. (ed.) *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press, 1964.

On Stochastic Processes

Cox, D. R., and H. D. Miller. *The Theory of Stochastic Processes*. London: Chapman & Hall, 1977.

Feller, W. *Introduction to Probability Theory and Its Applications*. New York: Wiley, 1968.

Karlin, S., and H. M. Taylor. *A First Course in Stochastic Processes*, 2nd edn. New York: Academic Press, 1975.

Shreve, S. E. *Stochastic Calculus for Finance II: Continuous-Time Models*. New York: Springer, 2008.

Practice Questions (Answers in Solutions Manual)

14.1. What would it mean to assert that the temperature at a certain place follows a Markov process? Do you think that temperatures do, in fact, follow a Markov process?

14.2. Can a trading rule based on the past history of a stock's price ever produce returns that are consistently above average? Discuss.

14.3. A company's cash position, measured in millions of dollars, follows a generalized Wiener process with a drift rate of 0.5 per quarter and a variance rate of 4.0 per quarter. How high does the company's initial cash position have to be for the company to have a less than 5% chance of a negative cash position by the end of 1 year?

14.4. Variables X_1 and X_2 follow generalized Wiener processes, with drift rates μ_1 and μ_2 and variances σ_1^2 and σ_2^2. What process does $X_1 + X_2$ follow if:
(a) The changes in X_1 and X_2 in any short interval of time are uncorrelated?
(b) There is a correlation ρ between the changes in X_1 and X_2 in any short time interval?

14.5. Consider a variable S that follows the process

$$dS = \mu\, dt + \sigma\, dz$$

For the first three years, $\mu = 2$ and $\sigma = 3$; for the next three years, $\mu = 3$ and $\sigma = 4$. If the initial value of the variable is 5, what is the probability distribution of the value of the variable at the end of year 6?

14.6. Suppose that G is a function of a stock price S and time. Suppose that σ_S and σ_G are the volatilities of S and G. Show that, when the expected return of S increases by $\lambda\sigma_S$, the growth rate of G increases by $\lambda\sigma_G$, where λ is a constant.

14.7. Stock A and stock B both follow geometric Brownian motion. Changes in any short interval of time are uncorrelated with each other. Does the value of a portfolio consisting of one of stock A and one of stock B follow geometric Brownian motion? Explain your answer.

14.8. The process for the stock price in equation (14.8) is

$$\Delta S = \mu S\,\Delta t + \sigma S\epsilon\sqrt{\Delta t}$$

where μ and σ are constant. Explain carefully the difference between this model and each of the following:

$$\Delta S = \mu\,\Delta t + \sigma\epsilon\sqrt{\Delta t}$$
$$\Delta S = \mu S\,\Delta t + \sigma\epsilon\sqrt{\Delta t}$$
$$\Delta S = \mu\,\Delta t + \sigma S\epsilon\sqrt{\Delta t}$$

Why is the model in equation (14.8) a more appropriate model of stock price behavior than any of these three alternatives?

14.9. It has been suggested that the short-term interest rate r follows the stochastic process

$$dr = a(b - r)\,dt + rc\,dz$$

where a, b, c are positive constants and dz is a Wiener process. Describe the nature of this process.

14.10. Suppose that a stock price S follows geometric Brownian motion with expected return μ and volatility σ:

$$dS = \mu S\,dt + \sigma S\,dz$$

What is the process followed by the variable S^n? Show that S^n also follows geometric Brownian motion.

14.11. Suppose that x is the yield to maturity with continuous compounding on a zero-coupon bond that pays off \$1 at time T. Assume that x follows the process

$$dx = a(x_0 - x)\,dt + sx\,dz$$

where a, x_0, and s are positive constants and dz is a Wiener process. What is the process followed by the bond price?

14.12. A stock whose price is \$30 has an expected return of 9% and a volatility of 20%. In Excel, simulate the stock price path over 5 years using monthly time steps and random samples from a normal distribution. Chart the simulated stock price path. By hitting F9, observe how the path changes as the random samples change.

Further Questions

14.13. Suppose that a stock price has an expected return of 16% per annum and a volatility of 30% per annum. When the stock price at the end of a certain day is \$50, calculate the following:
(a) The expected stock price at the end of the next day
(b) The standard deviation of the stock price at the end of the next day
(c) The 95% confidence limits for the stock price at the end of the next day.

14.14. A company's cash position, measured in millions of dollars, follows a generalized Wiener process with a drift rate of 0.1 per month and a variance rate of 0.16 per month. The initial cash position is 2.0.

(a) What are the probability distributions of the cash position after 1 month, 6 months, and 1 year?

(b) What are the probabilities of a negative cash position at the end of 6 months and 1 year?

(c) At what time in the future is the probability of a negative cash position greatest?

14.15. Suppose that x is the yield on a perpetual government bond that pays interest at the rate of \$1 per annum. Assume that x is expressed with continuous compounding, that interest is paid continuously on the bond, and that x follows the process

$$dx = a(x_0 - x)\,dt + sx\,dz$$

where a, x_0, and s are positive constants, and dz is a Wiener process. What is the process followed by the bond price? What is the expected instantaneous return (including interest and capital gains) to the holder of the bond?

14.16. If S follows the geometric Brownian motion process in equation (14.6), what is the process followed by

(a) $y = 2S$

(b) $y = S^2$

(c) $y = e^S$

(d) $y = e^{r(T-t)}/S$.

In each case express the coefficients of dt and dz in terms of y rather than S.

14.17. A stock price is currently 50. Its expected return and volatility are 12% and 30%, respectively. What is the probability that the stock price will be greater than 80 in 2 years? (*Hint*: $S_T > 80$ when $\ln S_T > \ln 80$.)

14.18. Stock A, whose price is \$30, has an expected return of 11% and a volatility of 25%. Stock B, whose price is \$40, has an expected return of 15% and a volatility of 30%. The processes driving the returns are correlated with correlation parameter ρ. In Excel, simulate the two stock price paths over 3 months using daily time steps and random samples from normal distributions. Chart the results and by hitting F9 observe how the paths change as the random samples change. Consider values for ρ equal to 0.25, 0.75, and 0.95.

APPENDIX

DERIVATION OF ITÔ'S LEMMA

In this appendix, we show how Itô's lemma can be regarded as a natural extension of other, simpler results. Consider a continuous and differentiable function G of a variable x. If Δx is a small change in x and ΔG is the resulting small change in G, a well-known result from ordinary calculus is

$$\Delta G \approx \frac{dG}{dx}\,\Delta x \tag{14A.1}$$

In other words, ΔG is approximately equal to the rate of change of G with respect to x multiplied by Δx. The error involves terms of order Δx^2. If more precision is required, a Taylor series expansion of ΔG can be used:

$$\Delta G = \frac{dG}{dx}\,\Delta x + \tfrac{1}{2}\frac{d^2G}{dx^2}\,\Delta x^2 + \tfrac{1}{6}\frac{d^3G}{dx^3}\,\Delta x^3 + \cdots$$

For a continuous and differentiable function G of two variables x and y, the result analogous to equation (14A.1) is

$$\Delta G \approx \frac{\partial G}{\partial x}\,\Delta x + \frac{\partial G}{\partial y}\,\Delta y \tag{14A.2}$$

and the Taylor series expansion of ΔG is

$$\Delta G = \frac{\partial G}{\partial x}\,\Delta x + \frac{\partial G}{\partial y}\,\Delta y + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}\,\Delta x^2 + \frac{\partial^2 G}{\partial x\,\partial y}\,\Delta x\,\Delta y + \tfrac{1}{2}\frac{\partial^2 G}{\partial y^2}\,\Delta y^2 + \cdots \tag{14A.3}$$

In the limit, as Δx and Δy tend to zero, equation (14A.3) becomes

$$dG = \frac{\partial G}{\partial x}\,dx + \frac{\partial G}{\partial y}\,dy \tag{14A.4}$$

We now extend equation (14A.4) to cover functions of variables following Itô processes. Suppose that a variable x follows the Itô process

$$dx = a(x,t)\,dt + b(x,t)\,dz \tag{14A.5}$$

and that G is some function of x and of time t. By analogy with equation (14A.3), we can write

$$\Delta G = \frac{\partial G}{\partial x}\,\Delta x + \frac{\partial G}{\partial t}\,\Delta t + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}\,\Delta x^2 + \frac{\partial^2 G}{\partial x\,\partial t}\,\Delta x\,\Delta t + \tfrac{1}{2}\frac{\partial^2 G}{\partial t^2}\,\Delta t^2 + \cdots \tag{14A.6}$$

Equation (14A.5) can be discretized to

$$\Delta x = a(x,t)\,\Delta t + b(x,t)\epsilon\sqrt{\Delta t}$$

or, if arguments are dropped,

$$\Delta x = a\,\Delta t + b\epsilon\sqrt{\Delta t} \tag{14A.7}$$

This equation reveals an important difference between the situation in equation (14A.6) and the situation in equation (14A.3). When limiting arguments were used to move from equation (14A.3) to equation (14A.4), terms in Δx^2 were ignored because they were second-order terms. From equation (14A.7), we have

$$\Delta x^2 = b^2\epsilon^2\Delta t + \text{terms of higher order in } \Delta t \tag{14A.8}$$

This shows that the term involving Δx^2 in equation (14A.6) has a component that is of order Δt and cannot be ignored.

The variance of a standard normal distribution is 1.0. This means that

$$E(\epsilon^2) - [E(\epsilon)]^2 = 1$$

where E denotes expected value. Since $E(\epsilon) = 0$, it follows that $E(\epsilon^2) = 1$. The expected value of $\epsilon^2\Delta t$, therefore, is Δt. The variance of $\epsilon^2\Delta t$ is, from the properties of the standard normal distribution, $2\Delta t^2$. We know that the variance of the change in a stochastic variable in time Δt is proportional to Δt, not Δt^2. The variance of $\epsilon^2\Delta t$ is therefore too small for it to have a stochastic component. As a result, we can treat $\epsilon^2\Delta t$ as nonstochastic and equal to its expected value, Δt, as Δt tends to zero. It follows from equation (14A.8) that Δx^2 becomes nonstochastic and equal to $b^2 dt$ as Δt tends to zero. Taking limits as Δx and Δt tend to zero in equation (14A.6), and using this last result, we obtain

$$dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial t}dt + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2 dt \tag{14A.9}$$

This is Itô's lemma. If we substitute for dx from equation (14A.5), equation (14A.9) becomes

$$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right)dt + \frac{\partial G}{\partial x}b\,dz.$$

Technical Note 29 at www.rotman.utoronto.ca/~hull/TechnicalNotes provides proofs of extensions to Itô's lemma. When G is a function of variables $x_1, x_2, \dots, x_n$ and

$$dx_i = a_i\,dt + b_i\,dz_i$$

we have

$$dG = \left(\sum_{i=1}^{n}\frac{\partial G}{\partial x_i}a_i + \frac{\partial G}{\partial t} + \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 G}{\partial x_i\,\partial x_j}b_i b_j \rho_{ij}\right)dt + \sum_{i=1}^{n}\frac{\partial G}{\partial x_i}b_i\,dz_i \tag{14A.10}$$

Also, when G is a function of a variable x with several sources of uncertainty so that

$$dx = a\,dt + \sum_{i=1}^{m} b_i\,dz_i$$

we have

$$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \tfrac{1}{2}\frac{\partial^2 G}{\partial x^2}\sum_{i=1}^{m}\sum_{j=1}^{m} b_i b_j \rho_{ij}\right)dt + \frac{\partial G}{\partial x}\sum_{i=1}^{m} b_i\,dz_i \tag{14A.11}$$

In these equations, ρ_{ij} is the correlation between dz_i and dz_j (see Section 14.5).