

---

## Continuous-Time Models

This chapter has three objectives. The first is to introduce the concept of a Brownian motion. A Brownian motion is a random process (a variable that changes randomly over time) that evolves continuously in time and has the property that its change over any time period is normally distributed with mean zero and variance equal to the length of the time period. The “mean zero” feature means that a Brownian motion is a martingale. We will also give a different characterization (Levy’s theorem) emphasizing the “quadratic variation” process, which is a property of the paths (how the variable evolves over time, in a given state of the world) of the process.

The second objective is to explain Itô’s formula, which is the chain rule for stochastic calculus. In the Black-Scholes model, the stock price is assumed to satisfy

$$\frac{dS}{S} = \mu dt + \sigma dB ,$$

where  $B$  is a Brownian motion. In the case that the stock pays no dividend, the rate of return is its price change  $dS$  divided by the initial price  $S$ , so the model states that the expected rate of return in each instant  $dt$  is  $\mu dt$  (of course,  $t$  denotes time, so  $dt$  is the change in time). The variance of the rate of return depends on  $\sigma$ . This model can be equivalently written in terms of the natural logarithm of  $S$ , which we will write as  $\log S$ . The above equation for the rate of return is equivalent to

$$d \log S = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dB .$$

We will explain this equivalence and other similar calculations that are useful for pricing derivatives.

The third objective is to explain how, when we change numeraires, as described in the previous chapter, we can calculate the expectation in the fundamental pricing formula (1.17). The question is what effect does changing the numeraire (and hence the probability measure) have on the distribution of an asset price.

Everything in the remainder of the book is based on the mathematics presented in this chapter. For easy reference, the essential formulas have been highlighted in boxes.

## 2.1 Simulating a Brownian Motion

We begin with the fact that changes in the value of a Brownian motion are normally distributed with mean zero and variance equal to the length of the time period. Let  $B(t)$  denote the value of a Brownian motion at time  $t$ . Then for any date  $u > t$ , given the information at time  $t$ , the random variable  $B(u) - B(t)$  is normally distributed with mean zero and variance equal to  $u - t$ . Unless stated otherwise, our convention will be that a Brownian motion starts at  $B(0) = 0$ .

We can generate an approximate Brownian motion in Excel. To do so, we take a small time period  $\Delta t$  and define the value at the end of the period to be the value of the Brownian motion at the beginning plus a normally distributed variable with mean 0 and variance  $\Delta t$ . In the following procedure, the user is prompted to input the length  $T$  of the entire time period over which the Brownian motion is to be simulated and to input the number  $N$  of time periods of length  $\Delta t$  within the full interval  $[0, T]$ . The length  $\Delta t$  of each individual time period is then calculated as  $T/N$ . The quality of the approximation of this simulation to a true Brownian motion will be always be improved by increasing the number  $N$ . Plotting the output of the procedure creates a picture of what we call a “path” of the Brownian motion, which means that it shows the value taken at each time in one state of the world. Running the procedure again (for the same  $T$  and  $N$ ) will create a different plot, which can be interpreted as the values of the Brownian motion in another state of the world. In other words, the path of the Brownian motion is itself random, depending in this approximation on the numbers produced by Excel’s random number generating function.<sup>1</sup>

```
Sub Simulating_Brownian_Motion()
Dim T, dt, Sqrtdt, BrownianMotion, i, N
T = InputBox("Enter the length of the time period (T)")
N = InputBox("Enter the number of periods (N)")
dt = T / N
Sqrtdt = Sqr(dt)
```

<sup>1</sup> The generation of normally distributed random numbers in Excel is discussed in Appendix A. The function `RandN()` here is user-created (to simplify typing) to equal the function `Application.NormSInv(Rnd())` supplied in VBA. The construction `sqrtdt * z` scales the standard normal  $z$  so that its standard deviation is  $\sqrt{\Delta t}$  and hence its variance is  $\Delta t$ , as desired. The subroutine creates two columns of data below the active cell in the Excel worksheet with headings “Time” and “Brownian Motion.” To plot the path of the Brownian motion, select the two columns and insert an “XY (Scatter)” chart, with data points connected by lines.

```

ActiveCell.Value = "Time"
ActiveCell.Offset(0, 1) = "Brownian Motion"
ActiveCell.Offset(1, 0) = 0 ' beginning time
ActiveCell.Offset(1, 1) = 0 ' beginning value of Brownian motion
BrownianMotion = 0
For i = 1 To N
    ActiveCell.Offset(i + 1, 0) = i * dt ' next time
    BrownianMotion = BrownianMotion + Sqr(dt) * RandN()
    ActiveCell.Offset(i + 1, 1) = BrownianMotion ' next value
Next i
End Sub

```

## 2.2 Quadratic Variation

If we take a large number  $N$  of time steps in the simulation of the preceding section, we will see the distinctive characteristic of a Brownian motion: it jiggles rapidly, moving up and down in a very erratic way. The name “Brownian motion” derives from the botanist Robert Brown’s observations of the erratic behavior of particles suspended in a fluid. This has long been thought to be a reasonable model for the behavior of a stock price. The plot of other functions with which we may be familiar will be much smoother. This is captured in the concept of quadratic variation.

Consider a discrete partition

$$0 = t_0 < t_1 < t_2 < \cdots < t_N = T$$

of the time interval  $[0, T]$ . Let  $B$  be a Brownian motion and calculate the sum of squared changes

$$\sum_{i=1}^N [\Delta B(t_i)]^2,$$

where  $\Delta B(t_i)$  denotes the change  $B(t_i) - B(t_{i-1})$ . If we consider finer partitions with the length of each time interval  $t_i - t_{i-1}$  going to zero, the limit of the sum is called the “quadratic variation” of the process. For a Brownian motion, the quadratic variation over an interval  $[0, T]$  is equal to  $T$  with probability one.

The functions with which we are normally familiar are continuously differentiable. If  $X$  is a continuously differentiable function of time (in each state of the world), then the quadratic variation of  $X$  will be zero. A simple example is a linear function:  $X(t) = at$  for some constant  $a$ . Then, taking  $t_i - t_{i-1} = \Delta t = T/N$  for each  $i$ , the sum of squared changes is

$$\sum_{i=1}^N [\Delta X(t_i)]^2 = \sum_{i=1}^N [a \Delta t]^2 = Na^2(\Delta t)^2 = Na^2 \left(\frac{T}{N}\right)^2 = \frac{a^2 T^2}{N} \rightarrow 0$$

as  $N \rightarrow \infty$ . Essentially the same argument shows that the quadratic variation of any continuously differentiable function is zero, because such a function is approximately linear at each point.

Thus, the jiggling of a Brownian motion, which leads to the nonzero quadratic variation, is quite unusual. To explain exactly how unusual it is, it is helpful to introduce the concept of “total variation,” which is defined in the same way as quadratic variation but with the squared changes  $[\Delta B(t_i)]^2$  replaced by the absolute value of the changes  $|\Delta B(t_i)|$ . If the quadratic variation of a continuous function is nonzero, then its total variation is necessarily infinite, so each path of a Brownian motion has infinite total variation (with probability one). It was mentioned above that, with a large number of time steps in the simulation of the preceding section, one could see the distinctive jiggling property of a Brownian motion. This is not quite right. Any plot drawn by a pencil (or a laser printer, for that matter) must have finite total variation, because the total variation is the total distance traveled by the pencil. Hence, no matter how many time steps one uses, one will never create a continuous plot with the nonzero quadratic variation (and infinite total variation) that a Brownian path has. Another way to understand this is to consider focusing on a small segment of a plot and viewing it with a magnifying glass. If the segment is small enough, and excluding the finite number of kinks that a pencil can draw in the plot of a function, it will look approximately like a straight line under the magnifying glass (with slope equal to the derivative of the function). However, if one could view a segment of a path of a true Brownian motion under a magnifying glass, it would look much the same as the entire picture does to the naked eye—no matter how small the segment, one would still see the characteristic jiggling.

One may well question why we should be interested in this curious mathematical object. The reason is that asset pricing inherently involves martingales (variables that evolve randomly over time in such a way that their expected changes are always zero), as our fundamental pricing equation (1.17) establishes. Furthermore, continuous processes (variables whose paths are continuous functions of time) are much more tractable mathematically than are processes that can jump at some instants. More importantly, it is possible in a mathematical model with continuous processes to define perfect hedges much more readily than it is in a model involving jump processes. So, we are led to a study of continuous martingales. An important fact is that any non-constant continuous martingale must have infinite total variation! So, the normal functions with which we are familiar are left behind once we enter the study of continuous martingales.

There remains perhaps the question of why we focus on Brownian motion within the world of continuous martingales. The answer here is that any continuous martingale is really just a transformation of a Brownian motion. This is a consequence of the following important fact, which is known as Levy’s theorem:

A continuous martingale is a Brownian motion if and only if its quadratic variation over each interval  $[0, T]$  equals  $T$ .

Thus, among continuous martingales, a Brownian motion is defined by the condition that the quadratic variation over each interval  $[0, T]$  is equal to  $T$ . This is really just a normalization. A different continuous martingale may have a different quadratic variation, but it can be converted to a Brownian motion just by deforming the time scale. Furthermore, many continuous martingales can be constructed as “stochastic integrals” with respect to a Brownian motion. We take up this topic in the next section.

## 2.3 Itô Processes

An Itô process is a variable  $X$  that changes over time as

$$dX(t) = \mu(t) dt + \sigma(t) dB(t) , \quad (2.1)$$

where  $B$  is a Brownian motion, and  $\mu$  and  $\sigma$  can also be random processes. Some regularity conditions are needed on  $\mu$  and  $\sigma$  which we will omit, except for noting that  $\mu(t)$  and  $\sigma(t)$  should be known at time  $t$ . In particular, constant  $\mu$  and  $\sigma$  are certainly acceptable. When we add the changes over time, we get

$$X(T) = X(0) + \int_0^T \mu(t) dt + \int_0^T \sigma(t) dB(t)$$

for any  $T > 0$ . There are other types of random processes, in particular, processes that can jump, but we will not consider them in this book.

We will not formally define the integral  $\int_0^T \sigma(t) dB(t)$ , but it should be understood as being approximately equal to a discrete sum of the form

$$\sum_{i=1}^N \sigma(t_{i-1}) \Delta B(t_i) ,$$

where  $0 = t_0 < \dots < t_N = T$  and the time periods  $t_i - t_{i-1}$  are small. Given that we can simulate the changes  $\Delta B(t_i)$  as random normals, we can approximately simulate the random variable  $\int_0^T \sigma(t) dB(t)$  and hence we can approximately simulate  $X(T)$ .

An Itô process evolves continuously over time. We interpret  $\mu(t) dt$  as the expected change in  $X$  in an instant  $dt$ . The quantity  $\mu(t)$  is also called the “drift” of the process  $X$  at time  $t$ . The coefficient  $\sigma(t)$  is called the “diffusion” coefficient of  $X$  at time  $t$ .

If  $\mu$  and  $\sigma$  are constant, it is standard to refer to an Itô process  $X$  as a  $(\mu, \sigma)$ -Brownian motion. Of course, it is not a martingale when  $\mu \neq 0$ . For example, when  $\mu > 0$ ,  $X$  tends to increase over time. However, it has the jiggling property of a Brownian motion, scaled by the diffusion coefficient  $\sigma$ .

A very important fact is that an Itô process such as (2.1) can be a martingale only if  $\mu = 0$ . This should seem sensible, because  $\mu dt$  is the expected change in  $X$ , and a process is a martingale only if its expected change is zero.<sup>2</sup> This observation plays a fundamental role in deriving asset pricing formulas, as we will begin to see in Sect. 2.9. Conversely, if  $\mu = 0$  and

$$E \left[ \int_0^T \sigma^2(t) dt \right] < \infty \quad (2.2)$$

for each  $T$ , then the Itô process is a continuous martingale and the variance of its date- $T$  value, calculated with the information available at date 0, is:

$$\text{var}[X(T)] = E \left[ \int_0^T \sigma^2(t) dt \right] .$$

Whether  $\mu$  is zero or not, and independently of the assumption (2.2), the quadratic variation of the Itô process  $X$  is

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N [\Delta X(t_i)]^2 = \int_0^T \sigma^2(t) dt \quad (2.3)$$

with probability one. Thus we obtain (when  $\mu = 0$  and (2.2) holds) a continuous martingale with a different quadratic variation than a Brownian motion via the diffusion function  $\sigma$ .

To “compute” the quadratic variation of an Itô process, we use the following simple and important rules (for the sake of brevity, we drop the “(t)” notation from  $B(t)$  here and sometimes later):

$$(dt)^2 = 0 , \quad (2.4a)$$

$$(dt)(dB) = 0 , \quad (2.4b)$$

$$(dB)^2 = dt . \quad (2.4c)$$

We apply these rules to “compute” the quadratic variation of  $X$  as follows:

<sup>2</sup> If the sources of uncertainty in the market can be modeled as Brownian motions, then in fact every martingale is an Itô process with  $\mu = 0$ . This is some justification for the assumption we will make in this book, when studying continuous-time models, that all martingales are Itô processes.

If  $dX = \mu dt + \sigma dB$  for a Brownian motion  $B$ , then

$$\begin{aligned} (dX)^2 &= (\mu dt + \sigma dB)^2 \\ &= \mu^2(dt)^2 + 2\mu\sigma(dt)(dB) + \sigma^2(dB)^2 \\ &= 0 + 0 + \sigma^2 dt . \end{aligned}$$

We integrate this from 0 to  $T$  to obtain the quadratic variation (2.3) over that time period:<sup>3</sup>

$$\int_0^T (dX(t))^2 = \int_0^T \sigma^2(t) dt . \quad (2.5)$$

## 2.4 Itô's Formula

First we recall some facts of the ordinary calculus. If  $y = g(x)$  and  $x = f(t)$  with  $f$  and  $g$  being continuously differentiable functions, then

$$\frac{dy}{dt} = \frac{dy}{dx} \times \frac{dx}{dt} = g'(x(t))f'(t) .$$

Over a time period  $[0, T]$ , this implies that

$$y(T) = y(0) + \int_0^T \frac{dy}{dt} dt = y(0) + \int_0^T g'(x(t))f'(t) dt .$$

Substituting  $dx(t) = f'(t) dt$ , we can also write this as

$$y(T) = y(0) + \int_0^T g'(x(t)) dx(t) . \quad (2.6)$$

We can contrast (2.6) with a special case of Itô's formula for the calculus of Itô processes (the more general formula will be discussed in the next section). If  $B$  is a Brownian motion and  $Y = g(B)$  for a twice-continuously differentiable function  $g$ , then

$$Y(T) = Y(0) + \int_0^T g'(B(t)) dB(t) + \frac{1}{2} \int_0^T g''(B(t)) dt . \quad (2.7)$$

<sup>3</sup> In a more formal mathematical presentation, one normally writes  $d\langle X, X \rangle$  for what we are writing here as  $(dX)^2$ . This is the differential of the quadratic variation process, and the quadratic variation through date  $T$  is

$$\langle X, X \rangle(T) = \int_0^T d\langle X, X \rangle(t) = \int_0^T \sigma^2(t) dt .$$

Thus, relative to the ordinary calculus, Itô's formula has an "extra term" involving the second derivative  $g''$ . We can write (2.7) in differential form as

$$dY(t) = \frac{1}{2}g''(B(t)) dt + g'(B(t)) dB(t).$$

Thus,  $Y = g(B)$  is an Itô process with drift  $g''(B(t))/2$  and diffusion coefficient  $g'(B(t))$ .

To gain some intuition for the "extra term" in Itô's formula, we return to the ordinary calculus. Given dates  $t < u$ , the derivative defines a linear approximation of the change in  $y$  over this time period; i.e., setting  $\Delta x = x(u) - x(t)$  and  $\Delta y = y(u) - y(t)$ , we have the approximation

$$\Delta y \approx g'(x(t)) \Delta x .$$

A better approximation is given by the second-order Taylor series expansion

$$\Delta y \approx g'(x(t)) \Delta x + \frac{1}{2}g''(x(t)) [\Delta x]^2 .$$

An interpretation of (2.6) is that the linear approximation works perfectly for infinitesimal time periods  $dt$ , because we can compute the change in  $y$  over the time period  $[0, T]$  by "summing up" the infinitesimal changes  $g'(x(t)) dx(t)$ . In other words, the second-order term  $\frac{1}{2}g''(x(t)) [\Delta x]^2$  "vanishes" when we consider very small time periods.

The second-order Taylor series expansion in the case of  $Y = g(B)$  is

$$\Delta Y \approx g'(B(t)) \Delta B + \frac{1}{2}g''(B(t)) [\Delta B]^2 .$$

For example, given a partition  $0 = t_0 < t_1 < \dots < t_N = T$  of the time interval  $[0, T]$ , we have, with the same notation we have used earlier,

$$\begin{aligned} Y(T) &= Y(0) + \sum_{i=1}^N \Delta Y(t_i) \\ &\approx Y(0) + \sum_{i=1}^N g'(B(t_{i-1})) \Delta B(t_i) + \frac{1}{2} \sum_{i=1}^N g''(B(t_{i-1})) [\Delta B(t_i)]^2 . \end{aligned} \quad (2.8)$$

If we make the time intervals  $t_i - t_{i-1}$  shorter, letting  $N \rightarrow \infty$ , we cannot expect that the "extra" term here will disappear, leading to the result (2.6) of the ordinary calculus, because we know that

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N [\Delta B(t_i)]^2 = T ,$$

whereas for the continuously differentiable function  $x(t) = f(t)$ , the same limit is zero. In fact it seems sensible to interpret the limit of  $[\Delta B]^2$  as  $(dB)^2 = dt$ . This is perfectly consistent with Itô's formula: if we take the limit in (2.8), replacing the limit of  $[\Delta B(t_i)]^2$  with  $(dB)^2 = dt$ , we obtain (2.7).



## 2.5 Multiple Itô Processes

Now consider two Itô processes

$$dX(t) = \mu_x(t) dt + \sigma_x(t) dB_x(t) , \quad (2.9a)$$

$$dY(t) = \mu_y(t) dt + \sigma_y(t) dB_y(t) , \quad (2.9b)$$

where  $B_x$  and  $B_y$  can be different Brownian motions. The relation between the two Brownian motions is determined by their covariance or correlation. Given dates  $t < u$ , we know that both changes  $B_x(u) - B_x(t)$  and  $B_y(u) - B_y(t)$  are normally distributed with mean 0 and variance equal to  $u - t$ . There will exist a (possibly random) process  $\rho$  such that the covariance of these two normally distributed random variables, given the information at date  $t$ , is

$$E_t \left[ \int_t^u \rho(s) ds \right] .$$

The process  $\rho$  is called the correlation coefficient of the two Brownian motions, because when it is constant the correlation of the changes  $B_x(u) - B_x(t)$  and  $B_y(u) - B_y(t)$  is

$$\frac{\text{covariance}}{\text{product of standard deviations}} = \frac{\int_t^u \rho ds}{\sqrt{u-t}\sqrt{u-t}} = \frac{(u-t)\rho}{u-t} = \rho .$$

Moreover, given increasingly fine partitions  $0 = t_0 < \dots < t_N = T$  of an interval  $[0, T]$  as before, we will have

$$\sum_{i=1}^N \Delta B_x(t_i) \times \Delta B_y(t_i) \rightarrow \int_0^T \rho(t) dt$$

as  $N \rightarrow \infty$ , with probability one.

We know that

$$\sum_{i=1}^N [\Delta X(t_i)]^2 \rightarrow \int_0^T \sigma_x^2(t) dt \quad \text{and} \quad \sum_{i=1}^N [\Delta Y(t_i)]^2 \rightarrow \int_0^T \sigma_y^2(t) dt . \quad (2.10)$$

Furthermore, it can be shown that the sum of products satisfies

$$\sum_{i=1}^N \Delta X(t_i) \times \Delta Y(t_i) \rightarrow \int_0^T \sigma_x(t) \sigma_y(t) \rho(t) dt . \quad (2.11)$$

By adding the rule

$$(dB_x)(dB_y) = \rho dt \quad (2.4d)$$

to the rules (2.4a)–(2.4c), we can “compute” the limit in (2.11) as

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{i=1}^N \Delta X(t_i) \times \Delta Y(t_i) &= \int_0^T (dX)(dY) \\ &= \int_0^T (\mu_x dt + \sigma_x dB_x)(\mu_y dt + \sigma_y dB_y) \\ &= \int_0^T \sigma_x(t)\sigma_y(t)\rho(t) dt . \end{aligned} \quad (2.12)$$

The most general case of Itô’s formula that we will need is for a function  $Z(t) = g(t, X(t), Y(t))$  where  $X$  and  $Y$  are Itô processes as in (2.9). In this case, Itô’s formula is<sup>4</sup>

$$\begin{aligned} Z(T) &= Z(0) + \int_0^T \frac{\partial g}{\partial t} dt + \int_0^T \frac{\partial g}{\partial x} dX(t) + \int_0^T \frac{\partial g}{\partial y} dY(t) \\ &\quad + \frac{1}{2} \int_0^T \frac{\partial^2 g}{\partial x^2} (dX(t))^2 + \frac{1}{2} \int_0^T \frac{\partial^2 g}{\partial y^2} (dY(t))^2 \\ &\quad + \int_0^T \frac{\partial^2 g}{\partial x \partial y} (dX(t))(dY(t)) . \end{aligned} \quad (2.13)$$

In this equation, we apply the rules (2.4a)–(2.4d) to compute

$$\begin{aligned} (dX(t))^2 &= \sigma_x^2(t) dt , \\ (dY(t))^2 &= \sigma_y^2(t) dt , \\ (dX(t))(dY(t)) &= \sigma_x(t)\sigma_y(t)\rho(t) dt . \end{aligned}$$

Itô’s formula (2.13) appears a bit simpler (and easier to remember) if we write it in “differential form.” We have:

If  $Z(t) = g(t, X(t), Y(t))$  where  $X$  and  $Y$  are Itô processes as in (2.9), then

$$\begin{aligned} dZ &= \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX + \frac{\partial g}{\partial y} dY + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (dX)^2 + \frac{1}{2} \frac{\partial^2 g}{\partial y^2} (dY)^2 \\ &\quad + \frac{\partial^2 g}{\partial x \partial y} (dX)(dY) . \end{aligned} \quad (2.14)$$

<sup>4</sup> We need to assume  $g(t, x, y)$  is continuously differentiable in  $t$  and twice continuously differentiable in  $(x, y)$  for (2.13) and (2.14) to be valid. Note also that we are using a short-hand notation here. The partial derivatives of  $g$  will generally depend on  $t, X(t)$  and  $Y(t)$  just as  $g$  does.

## 2.6 Examples of Itô's Formula

The following are the applications of Itô's formula that will be used most frequently in the book. They follow from the boxed formula at the end of the previous section by taking  $g(x, y) = xy$  or  $g(x, y) = y/x$  or  $g(x) = e^x$  or  $g(x) = \log x$ .

**Products.** If  $Z = XY$ , then  $dZ = X dY + Y dX + (dX)(dY)$ . We can write this as

$$\frac{dZ}{Z} = \frac{dX}{X} + \frac{dY}{Y} + \left(\frac{dX}{X}\right) \left(\frac{dY}{Y}\right). \quad (2.15)$$

**Ratios.** If  $Z = Y/X$ , then

$$\frac{dZ}{Z} = \frac{dY}{Y} - \frac{dX}{X} - \left(\frac{dY}{Y}\right) \left(\frac{dX}{X}\right) + \left(\frac{dX}{X}\right)^2. \quad (2.16)$$

**Exponentials.** If  $Z = e^X$ , then

$$\frac{dZ}{Z} = dX + \frac{(dX)^2}{2}. \quad (2.17)$$

**Logarithms.** If  $Z = \log X$ , then

$$dZ = \frac{dX}{X} - \frac{1}{2} \left(\frac{dX}{X}\right)^2. \quad (2.18)$$

**Compounding/Discounting.** Let

$$Y(t) = \exp\left(\int_0^t q(s) ds\right)$$

for some (possibly random) process  $q$  and define  $Z = XY$  for any Itô process  $X$ . The usual calculus gives us  $dY(t) = q(t)Y(t) dt$ , and the product rule above implies

$$\frac{dZ}{Z} = q dt + \frac{dX}{X}. \quad (2.19)$$

This is the same as in the usual calculus.

## 2.7 Reinvesting Dividends

Frequently, we will assume that the asset underlying a derivative security pays a “constant dividend yield,” which we will denote by  $q$ . This means, for an asset with price  $S(t)$ , that the dividend “in an instant  $dt$ ” is  $qS(t) dt$ . If the dividends are reinvested in new shares, the number of shares will grow exponentially at rate  $q$ . To see this, consider the portfolio starting with a single share of the asset and reinvesting dividends until some date  $T$ . Let  $X(t)$  denote the number of shares resulting from this strategy at any time  $t \leq T$ . Then the dividend received at date  $t$  is  $qS(t)X(t) dt$ , which can be used to purchase  $qX(t) dt$  new shares. This implies that  $dX(t) = qX(t) dt$ , or  $dX(t)/dt = qX(t)$ , and it is easy to check (and very well known) that this equation is solved by  $X(t) = e^{qt}X(0)$ . In our case, with  $X(0) = 1$ , we have  $X(t) = e^{qt}$ .

The dollar value of the trading strategy just described will be  $X(t)S(t) = e^{qt}S(t)$ . Denote this by  $V(t)$ . This is the value of a non-dividend-paying portfolio, because all dividends are reinvested. From the Compounding/Discounting example in Sect. 2.6, we know that

$$\frac{dV}{V} = q dt + \frac{dS}{S}. \quad (2.20)$$

This means that the rate of return on the portfolio is the dividend yield  $q dt$  plus the return  $dS/S$  due to capital gains.

## 2.8 Geometric Brownian Motion

Let

$$S(t) = S(0) \exp(\mu t - \sigma^2 t/2 + \sigma B(t)) \quad (2.21)$$

for constants  $\mu$  and  $\sigma$ , where  $B$  is a Brownian motion. Using the product rule and the rule for exponentials, we obtain

$$\frac{dS}{S} = \mu dt + \sigma dB. \quad (2.22)$$

When we see an equation of the form (2.22), we should recognize (2.21) as the solution.

The process  $S$  is called a “geometric Brownian motion.” In keeping with the discussion of Sect. 2.3, we interpret (2.22) as stating that  $\mu dt$  is the expected rate of change of  $S$  and  $\sigma^2 dt$  is the variance of the rate of change in an instant  $dt$ . We call  $\mu$  the “drift” and  $\sigma$  the “volatility.” The geometric Brownian motion will grow at the average rate of  $\mu$ , in the sense that  $E[S(t)] = e^{\mu t} S(0)$ .

Taking the natural logarithm of (2.21) gives an equivalent form of the solution:

$$\log S(t) = \log S(0) + \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t). \quad (2.23)$$

This shows that  $\log S(t) - \log S(0)$  is a  $(\mu - \sigma^2/2, \sigma)$ -Brownian motion. Given information at time  $t$ , the logarithm of  $S(u)$  for  $u > t$  is normally distributed with mean  $(u - t)(\mu - \sigma^2/2)$  and variance  $(u - t)\sigma^2$ . Because  $S$  is the exponential of its logarithm,  $S$  can never be negative. For this reason, a geometric Brownian motion is a better model for stock prices than is a Brownian motion.

The differential of (2.23) is

$$d \log S(t) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB(t). \quad (2.24)$$

We conclude:

The equation

$$\frac{dS}{S} = \mu dt + \sigma dB$$

is equivalent to the equation

$$d \log S(t) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dB(t).$$

The solution of both equations is (2.21) or the equivalent formula (2.23).

Over a discrete time interval  $\Delta t$ , equation (2.24) implies that the change in the logarithm of  $S$  is

$$\Delta \log S = \left( \mu - \frac{1}{2} \sigma^2 \right) \Delta t + \sigma \Delta B . \quad (2.25)$$

If  $S$  is the price of a non-dividend-paying asset, then over the time period  $t_{i-1}$  to  $t_i$ , with  $t_i - t_{i-1} = \Delta t$ , we have

$$\Delta \log S = r_i \Delta t , \quad (2.26)$$

where  $r_i$  is the continuously compounded annualized rate of return during the period  $\Delta t$ . This follows from the definition of the continuously compounded rate of return as the constant rate over the time period  $\Delta t$  that would cause  $S$  to grow (or fall) from  $S(t_{i-1})$  to  $S(t_i)$ . To be precise,  $r_i$  is defined by

$$\frac{S(t_i)}{S(t_{i-1})} = e^{r_i \Delta t} ,$$

which is equivalent to (2.26). Thus, the geometric Brownian motion model (2.22) implies that the continuously compounded annualized rate of return over a period of length  $\Delta t$  is given by

$$r_i = \mu - \frac{1}{2} \sigma^2 + \frac{\sigma \Delta B}{\Delta t} .$$

This means that  $r_i$  is normally distributed with mean  $\mu - \sigma^2/2$  and variance  $\sigma^2/\Delta t$ . Given historical data on the rates of return, the parameters  $\mu$  and  $\sigma$  can be estimated by standard methods (see Chap. 4).

We can simulate a path of  $S$  by simulating the changes  $\Delta \log S$ . The random variable  $\sigma \Delta B$  in (2.25) has a normal distribution with zero mean and variance equal to  $\sigma^2 \Delta t$ . We simulate it as  $\sigma \sqrt{\Delta t}$  multiplied by a standard normal.

```
Sub Simulating_Geometric_Brownian_Motion()
Dim T, S, mu, sigma, dt, SigSqrt, LogS, drift, i, N
T = InputBox("Enter the length of the time period (T)")
N = InputBox("Enter the number of periods (N)")
S = InputBox("Enter the initial stock price (S)")
mu = InputBox("Enter the expected rate of return (mu)")
sigma = InputBox("Enter the volatility (sigma)")
dt = T / N
SigSqrt = sigma * Sqr(dt)
drift = (mu - 0.5 * sigma * sigma) * dt
LogS = Log(S)
ActiveCell.Value = "Time"
ActiveCell.Offset(0, 1) = "Stock Price"
ActiveCell.Offset(1, 0) = 0 ' beginning time
ActiveCell.Offset(1, 1) = S ' beginning stock price
```

```

For i = 1 To N
  ActiveCell.Offset(i + 1, 0) = i * dt      ' next time
  LogS = LogS + SigSqrtdt * RandN()
  ActiveCell.Offset(i + 1, 1) = Exp(LogS)  ' next stock price
Next i
End Sub

```

## 2.9 Numeraires and Probabilities

When we change probability measures, we cannot expect a process  $B$  that was a Brownian motion to remain a Brownian motion. The expected change in a Brownian motion must always be zero, but when we change probabilities, the expected change of  $B$  is likely to become nonzero. (Likewise, a martingale is unlikely to remain a martingale when we change probabilities.) However, the Brownian motion  $B$  will still be an Itô process under the new probability measure. In fact, every Itô process under one probability measure will still be an Itô process under the new probability measure, and the diffusion coefficient of the Itô process will be unaffected by the change in probabilities.<sup>5</sup> Changing probabilities only changes the drift of an Itô process.

In a sense, this should not be surprising. It was noted in Sect. 2.2 that a Brownian motion  $B$  can be defined as a continuous martingale with paths that jiggle in such a way that the quadratic variation over any interval  $[0, T]$  is equal to  $T$ . Changing the probabilities will change the probabilities of the various paths (so it may affect the expected change in  $B$ ) but it will not affect how each path jiggles. So, under the new probability measure,  $B$  should still be like a Brownian motion but it may have a nonzero drift. If we consider a general Itô process, the reasoning is the same. The diffusion coefficient  $\sigma$  determines how much each path jiggles, and this is unaffected by changing the probability measure. Furthermore, instantaneous covariances—the  $(dX)(dY)$  terms—between Itô processes are unaffected by changing the probability measure. Only the drifts are affected.

As explained in Sect. 1.5, we need to know the distribution of the underlying under probability measures corresponding to different numeraires. Let  $S$  be the price of an asset that has a constant dividend yield  $q$ , and, as in Sect. 2.7, let  $V(t) = e^{qt}S(t)$ . This is the price of the portfolio in which all dividends are reinvested, and we have

$$\frac{dV}{V} = q dt + \frac{dS}{S} .$$

Let  $Y$  be the price of another another asset that does not pay dividends. Let  $r(t)$  denote the instantaneous risk-free rate at date  $t$  and let  $R(t) =$

<sup>5</sup> To be a little more precise, this is true provided sets of states of the world having zero probability continue to have zero probability when the probabilities are changed. Because of the way we change probability measures when we change numeraires (cf. (1.11)) this will always be true for us.

$\exp\left(\int_0^t r(s) ds\right)$ . Assume

$$\begin{aligned}\frac{dS}{S} &= \mu_s dt + \sigma_s dB_s, \\ \frac{dY}{Y} &= \mu_y dt + \sigma_y dB_y,\end{aligned}$$

where  $B_s$  and  $B_y$  are Brownian motions under the actual probability measure with correlation  $\rho$ , and where  $\mu_s, \mu_y, \sigma_s, \sigma_y$  and  $\rho$  can be quite general random processes. We consider the dynamics of the asset price  $S$  under three different probability measures. In each case, we follow the same steps: (i) we note that the ratio of an asset price to the numeraire asset price must be a martingale, (ii) we use Itô's formula to calculate the drift of this ratio, and (iii) we use the fact that the drift of a martingale must be zero to compute the drift of  $dS/S$ .

### Risk-Neutral Probabilities

Under the risk-neutral measure,  $Z(t)$  defined as

$$Z(t) = \frac{V(t)}{R(t)} = \exp\left(-\int_0^t r(s) ds\right) V(t)$$

is a martingale. Using the compounding/discounting rule, we have

$$\frac{dZ}{Z} = -r dt + \frac{dV}{V} = (q - r) dt + \frac{dS}{S}.$$

For  $Z$  to be a martingale, the drift ( $dt$  part) of  $dZ/Z$  must be zero. Therefore, the drift of  $dS/S$  must be  $(r - q) dt$  under the risk-neutral measure. Because the change of measure does not affect the volatility, this implies:

$$\frac{dS}{S} = (r - q) dt + \sigma_s dB_s^*, \quad (2.27)$$

where  $B_s^*$  is a Brownian motion under the risk-neutral measure.

### Underlying as the Numeraire

When  $V$  is the numeraire, the process  $Z(t)$  defined as

$$Z(t) = \frac{R(t)}{V(t)} = \frac{\exp\left(\int_0^t r(s) ds\right)}{V(t)}$$

is a martingale. Using the rule for ratios, we have



$$\frac{dZ}{Z} = r dt - \frac{dV}{V} + \left(\frac{dV}{V}\right)^2 = (r - q + \sigma_s^2) dt - \frac{dS}{S}.$$

Because the drift of  $dZ/Z$  must be zero, this implies that the drift of  $dS/S$  is  $(r - q + \sigma_s^2) dt$ . We conclude that:

$$\frac{dS}{S} = (r - q + \sigma_s^2) dt + \sigma_s dB_s^*, \quad (2.28)$$

where now  $B_s^*$  denotes a Brownian motion when  $V(t) = e^{qt}S(t)$  is the numeraire.

### Another Risky Asset as the Numeraire

When  $Y$  is the numeraire,  $Z(t)$  defined as

$$Z(t) = \frac{V(t)}{Y(t)}$$

must be a martingale. Using again the rule for ratios, we have

$$\begin{aligned} \frac{dZ}{Z} &= \frac{dV}{V} - \frac{dY}{Y} - \left(\frac{dV}{V}\right) \left(\frac{dY}{Y}\right) + \left(\frac{dY}{Y}\right)^2 \\ &= \frac{dV}{V} - \frac{dY}{Y} - \rho\sigma_s\sigma_y dt + \sigma_y^2 dt \\ &= \frac{dS}{S} - \frac{dY}{Y} + (q - \rho\sigma_s\sigma_y dt + \sigma_y^2) dt. \end{aligned}$$

We can apply our previous example to compute the dynamics of  $Y$  when  $Y$  is the numeraire. This shows that the drift of  $dY/Y$  is  $(r + \sigma_y^2) dt$ . Because the drift of  $dZ/Z$  must be zero, it follows that the drift of  $dS/S$  is  $(r - q + \rho\sigma_s\sigma_y) dt$ . We conclude that:

$$\frac{dS}{S} = (r - q + \rho\sigma_s\sigma_y) dt + \sigma_s dB_s^*, \quad (2.29)$$

where  $B_s^*$  denotes a Brownian motion under the probability measure corresponding to the non-dividend-paying risky asset  $Y$  being the numeraire, and where  $\rho$  is the correlation of  $S$  and  $Y$ .

Notice that the formula (2.29), while more complicated, is also more general than the others. In fact, it includes the formulas (2.27) and (2.28) as special cases: (i) if  $Y$  is the price of the instantaneously risk-free asset, then  $\sigma_y = 0$  and (2.29) simplifies to (2.27), and (ii) if  $Y = V$ , then  $\sigma_y = \sigma_s$  and  $\rho = 1$ , so (2.29) simplifies to (2.28).

### Further Discussion

It would be natural for one to ask at this point: “what is the Brownian motion  $B_s^*$  and where did it come from?” We have argued that once we know the drift, and the fact that the volatility does not change, we can immediately write down, for example,

$$\frac{dS}{S} = (r - q) dt + \sigma_s dB_s^*$$

for a Brownian motion  $B_s^*$  under the risk-neutral measure. To answer this question, we will give here the definition of  $B_s^*$  under the risk-neutral measure. The definition shows that we are justified in writing down (2.27)–(2.29), but we will not repeat the definition each time we make a statement of this sort.

We showed that  $Z$  is a martingale under the risk-neutral measure, where  $Z$  satisfies

$$\frac{dZ}{Z} = (q - r) dt + \frac{dS}{S} = (q - r + \mu_s) dt + \sigma_s dB_s . \quad (2.30)$$

Define  $B_s^*(0) = 0$  and

$$dB_s^* = \left( \frac{q - r + \mu_s}{\sigma_s} \right) dt + dB_s . \quad (2.31)$$

Then

$$dB_s^* = \frac{1}{\sigma_s} \left( \frac{dZ}{Z} \right)$$

and hence is a continuous martingale under the risk-neutral measure. We can compute its quadratic variation as

$$(dB_s^*)^2 = \left( \frac{q - r + \mu_s}{\sigma_s} \right)^2 (dt)^2 + 2 \left( \frac{q - r + \mu_s}{\sigma_s} \right) (dt)(dB_s) + (dB_s)^2 = dt .$$

Therefore, by Levy’s theorem (Sect. 2.2),  $B_s^*$  is a Brownian motion under the risk-neutral measure. From (2.30) and (2.31) we have

$$(q - r) dt + \frac{dS}{S} = \sigma_s dB_s^* \iff \frac{dS}{S} = (r - q) dt + \sigma_s dB_s^* ,$$

as in (2.27).

## 2.10 Tail Probabilities of Geometric Brownian Motions

For each of the numeraires discussed in the previous section, we have

$$d \log S = \alpha dt + \sigma dB , \quad (2.32)$$

for some  $\alpha$  and  $\sigma$ , where  $B$  is a Brownian motion under the probability measure associated with the numeraire. Specifically,  $\sigma = \sigma_s$ ,  $B = B_s^*$ , and

- (1) for the risk-neutral measure,  $\alpha = r - q - \sigma_s^2/2$ ,
- (2) when  $e^{qt}S(t)$  is the numeraire,  $\alpha = r - q + \sigma_s^2/2$ ,
- (3) when another risky asset price  $Y$  is the numeraire,  $\alpha = r - q + \rho\sigma_s\sigma_y - \sigma_s^2/2$ .

We will assume in this section that  $\alpha$  and  $\sigma$  are constants. The essential calculation in pricing options, as we will see in the next chapter and in Chap. 8, is to compute  $\text{prob}(S(T) > K)$  and  $\text{prob}(S(T) < K)$  for a constant  $K$  (the strike price of an option), where  $\text{prob}$  denotes the probabilities at date 0 (the date we are pricing an option) associated with a particular numeraire.

Equation (2.32) gives us

$$\log S(T) = \log S(0) + \alpha T + \sigma B(T) .$$

Given this, we deduce

$$\begin{aligned} S(T) > K &\iff \log S(T) > \log K \\ &\iff \sigma B(T) > \log K - \log S(0) - \alpha T \\ &\iff \frac{B(T)}{\sqrt{T}} > \frac{\log K - \log S(0) - \alpha T}{\sigma\sqrt{T}} \\ &\iff -\frac{B(T)}{\sqrt{T}} < \frac{\log S(0) - \log K + \alpha T}{\sigma\sqrt{T}} \\ &\iff -\frac{B(T)}{\sqrt{T}} < \frac{\log\left(\frac{S(0)}{K}\right) + \alpha T}{\sigma\sqrt{T}} . \end{aligned} \tag{2.33}$$

The random variable on the left-hand side of (2.33) has the standard normal distribution—it is normally distributed with mean equal to zero and variance equal to one. As is customary, we will denote the probability that a standard normal is less than some number  $d$  as  $N(d)$ . We conclude:

Assume  $d \log S = \alpha dt + \sigma dB$ , where  $B$  is a Brownian motion. Then, for any number  $K$ ,

$$\text{prob}(S(T) > K) = N(d) , \tag{2.34}$$

where

$$d = \frac{\log\left(\frac{S(0)}{K}\right) + \alpha T}{\sigma\sqrt{T}} . \tag{2.35}$$

The probability  $\text{prob}(S(T) < K)$  can be calculated similarly, but the simplest way to derive it is to note that the events  $S(T) > K$  and  $S(T) < K$  are “complementary”—their probabilities sum to one (the event  $S(T) = K$  having zero probability). Therefore  $\text{prob}(S(T) < K) = 1 - N(d)$ . This is the probability that a standard normal is greater than  $d$ , and by virtue of the symmetry of the standard normal distribution, it equals the probability that a standard normal is less than  $-d$ . Therefore, we have:

Assume  $d \log S = \alpha dt + \sigma dB$ , where  $B$  is a Brownian motion. Then, for any number  $K$ ,

$$\text{prob}(S(T) < K) = N(-d) , \quad (2.36)$$

where  $d$  is defined in (2.35).

## 2.11 Volatilities

As mentioned in Sect. 2.8, when we encounter an equation of the form

$$\frac{dS}{S} = \mu dt + \sigma dB$$

where  $B$  is a Brownian motion, we will say “ $\sigma$  is the volatility of  $S$ .” For example, in the Black-Scholes model, the most important assumption is that the volatility of the underlying asset price is constant. We will occasionally need to compute the volatilities of products or ratios of random processes. These computations follow directly from Itô’s formula.

Suppose

$$\frac{dX}{X} = \mu_x dt + \sigma_x dB_x \quad \text{and} \quad \frac{dY}{Y} = \mu_y dt + \sigma_y dB_y ,$$

where  $B_x$  and  $B_y$  are Brownian motions with correlation  $\rho$ , and  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\rho$  may be quite general random processes.

### Products

If  $Z = XY$ , then (2.15) gives us

$$\frac{dZ}{Z} = (\mu_x + \mu_y + \rho\sigma_x\sigma_y) dt + \sigma_x dB_x + \sigma_y dB_y . \quad (2.37)$$

The instantaneous variance of  $dZ/Z$  is calculated, using the rules for products of differentials, as

$$\begin{aligned} \left( \frac{dZ}{Z} \right)^2 &= (\sigma_x dB_x + \sigma_y dB_y)^2 \\ &= (\sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y) dt . \end{aligned}$$

As will be explained below, the volatility is the square root of the instantaneous variance (dropping the  $dt$ ). This implies:

The volatility of  $XY$  is

$$\sqrt{\sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y} . \quad (2.38)$$

## Ratios

If  $Z = Y/X$ , then (2.16) gives us

$$\frac{dZ}{Z} = (\mu_y - \mu_x - \rho\sigma_x\sigma_y + \sigma_x^2) dt + \sigma_y dB_y - \sigma_x dB_x. \quad (2.39)$$

The instantaneous variance of  $dZ/Z$  is therefore

$$\begin{aligned} \left(\frac{dZ}{Z}\right)^2 &= (\sigma_y dB_y - \sigma_x dB_x)^2 \\ &= (\sigma_x^2 + \sigma_y^2 - 2\rho\sigma_x\sigma_y) dt. \end{aligned}$$

This implies:

The volatility of  $Y/X$  is

$$\sqrt{\sigma_x^2 + \sigma_y^2 - 2\rho\sigma_x\sigma_y}. \quad (2.40)$$

## Further Discussion

To understand why taking the square root of  $(dZ/Z)^2$  (dropping the  $dt$ ) gives the volatility, consider for example the product case  $Z = XY$ . Define a random process  $B$  by  $B(0) = 0$  and

$$dB = \frac{\sigma_x}{\sigma} dB_x + \frac{\sigma_y}{\sigma} dB_y, \quad (2.41)$$

where  $\sigma$  is the volatility defined in (2.38). Then we can write (2.37) as

$$\frac{dZ}{Z} = (\mu_x + \mu_y + \rho\sigma_x\sigma_y) dt + \sigma dB. \quad (2.42)$$

From the discussion in Sect. 2.3, we know that  $B$  is a continuous martingale. We can compute its quadratic variation from

$$\begin{aligned} (dB)^2 &= \left(\frac{\sigma_x dB_x + \sigma_s dB_s}{\sigma}\right)^2 \\ &= \frac{(\sigma_x^2 + \sigma_s^2 + 2\rho\sigma_x\sigma_s) dt}{\sigma^2}, \\ &= dt. \end{aligned}$$

By Levy's theorem (see Sect. 2.2), any continuous martingale with this quadratic variation is necessarily a Brownian motion. Therefore, (2.42) shows that  $\sigma$  is the volatility of  $Z$  as defined at the beginning of the section.

## Problems

**2.1.** Consider a discrete partition  $0 = t_0 < t_1 < \cdots t_N = T$  of the time interval  $[0, T]$  with  $t_i - t_{i-1} = \Delta t = T/N$  for each  $i$ . Consider the function

$$X(t) = e^t .$$

Create a VBA subroutine, prompting the user to input  $T$  and  $N$ , which computes and prints  $\sum_{i=1}^N [\Delta X(t_i)]^2$ , where

$$\Delta X(t_i) = X(t_i) - X(t_{i-1}) = e^{t_i} - e^{t_{i-1}} .$$

Hint: The sum can be computed as follows.

```
sum = 0
For i = 1 To N
    DeltaX = Exp(i/N)-Exp((i-1)/N)
    sum = sum + DeltaX * DeltaX
Next i
```

**2.2.** Repeat the previous problem for the function  $X(t) = t^3$ . In both this and the previous problem, what happens to  $\sum_{i=1}^N [\Delta X(t_i)]^2$  as  $N \rightarrow \infty$ ?

**2.3.** Repeat the previous problem to compute  $\sum_{i=1}^N [\Delta B(t_i)]^2$ , where  $B$  is a simulated Brownian motion. For a given  $T$ , what happens to the sum as  $N \rightarrow \infty$ ?

**2.4.** Repeat the previous problem, computing instead  $\sum_{i=1}^N |\Delta B(t_i)|$  where  $|\cdot|$  denotes the absolute value. What happens to this sum as  $N \rightarrow \infty$ ?

**2.5.** Consider a discrete partition  $0 = t_0 < t_1 < \cdots t_N = T$  of the time interval  $[0, T]$  with  $t_i - t_{i-1} = \Delta t = T/N$  for each  $i$ . Consider a geometric Brownian motion

$$\frac{dZ}{Z} = \mu dt + \sigma dB .$$

An approximate path  $\tilde{Z}(t)$  of the geometric Brownian motion can be simulated as

$$\Delta \tilde{Z}(t_i) = \tilde{Z}(t_{i-1}) [\mu \Delta t + \sigma \Delta B] . \quad (2.43)$$

The subroutine `Simulating_Geometric_Brownian_Motion` simulates a path  $Z$  of a geometric Brownian motion. Modify the subroutine to prompt the user to input  $T$ ,  $N$ ,  $\sigma$ ,  $\mu$ , and  $Z(0)$  and to generate both a path  $Z(t)$  and an approximate path  $\tilde{Z}(t)$  according to (2.43), using the same  $\Delta B$  for both paths and taking  $\tilde{Z}(0) = Z(0)$ . Plot both paths in the same figure. How well does the approximation work for large  $N$ ? Warning: For  $N$  larger than about  $100T$ , the approximation will look perfect—you won't be able to tell that there are two plots in the figure.