CHAPTER 1

A Simple Introduction to Continuous-Time Stochastic Processes

This chapter introduces continuous-time stochastic processes. Due to the applied focus of this book, we skip the proofs of the main results and provide heuristic derivations that serve to strengthen the mathematical intuition of the readers. The more rigorous proofs can be obtained from other sources.[1] Here, we simply give some of the widely applied results from stochastic calculus, a field of mathematics that combines calculus with probability, and show how these results can be used for modeling the term structure dynamics. The examples in this chapter highlight the mathematical intuition, without worrying too much about the "regularity" conditions underlying the continuous-time framework. The results in this chapter may be skipped by readers who are well versed in continuous-time mathematics, but they could be helpful to readers unfamiliar with this branch of mathematics.

Term structure modeling and derivative pricing are perhaps the most advanced areas of application of stochastic calculus in finance. Obviously, a comprehensive introduction to this mathematics is outside the scope of this chapter and this book. What we wish to accomplish here is to explain this difficult subject matter to the audience of fixed-income traders, analysts, and graduate students, using heuristic derivations and easy-to-read examples. We are well aware of the difficulty most non–mathematically inclined readers experience in comprehending advanced books in mathematical finance, and so this chapter puts together a basic toolbox of results from continuous-time stochastic processes to help those readers. More advanced mathematical ideas related to martingale valuation using absence of arbitrage are presented in Chapter 2.

We begin this chapter with an introduction to the continuous-time diffusion processes. A fundamental property of diffusion processes is that the magnitude of change over the next time interval is proportional to the length of the time interval. As the interval shrinks to zero, the change becomes continuous, but not differentiable. Since most prices and interest rates change randomly over very small time intervals (i.e., every day or minute by minute), modeling these processes as diffusion processes leads to good approximation for modeling uncertainty in the valuation process.

The basic tools of manipulating stochastic processes are Ito's lemma and the rules of stochastic differentiation and integration. Ito's lemma shows how to obtain the stochastic process of a variable that is a smooth (twice differentiable) function of another variable, whose stochastic process is known. For example, using Ito's lemma one can obtain the stochastic process of a bond price given the stochastic process of the short rate, assuming that the bond price is a twice differentiable function of the short rate. A stochastic differential equation shows how a variable changes stochastically over time. Specifically, it gives the probability distribution of the change in the variable over an infinitesimally small interval Δt. A stochastic integral is generally used to give the probability distribution of the change in the variable over a discrete interval. For example, a stochastic integral may give the distribution of the change in the bond price from time 0 to time t. We will cover some basic rules of stochastic differentiation and integration, without considering the precise mathematical regularity conditions. Most stochastic processes used in fixed income valuation generally satisfy the regularity conditions, and the reader will be alerted if violations of these conditions do occur.

Since stochastic processes are very general and may allow the instantaneous mean and variance of the underlying variable to change continuously over time, a stochastic integral can give a variety of distributions over discrete intervals, depending on the specific assumptions made about the parameters that define the stochastic process. In a multiple factor environment, modeling of stochastic processes allows dynamic conditional correlations and conditional volatilities that are more realistic than other simple approaches such as principal component analysis with a stationary variance-covariance matrix.

The final part of this chapter introduces the mixed jump-diffusion processes for modeling uncertainty. These processes use two components; one captures the diffusion element and the other captures the jump element. The jump element allows a sudden discontinuous movement in the underlying variable, the size of which is not restricted by the length of the time interval over which the jump occurs. However, the probability of the jump varies linearly with the length of the time interval, so the jumps occur with less likelihood over smaller time intervals. Many processes used in finance are a mix of a diffusion process and a jump process, such that the variable experiences a continuous change most of the time, while every once in a while the variable experiences a discontinuous change. We give examples of affine and quadratic term structure models with mixed jump-diffusion processes in Chapters 5, 6, 7, 9, 10, and 12.

CONTINUOUS-TIME DIFFUSION PROCESSES

A stochastic process is a variable whose value changes over time in a stochastic manner. If changes in the variable are measured over discrete intervals, the process is a discrete-time process. On the other hand, if the changes in the variable are measured over infinitesimally small intervals converging to zero, the process is a continuous-time process. Continuous-time diffusion processes display continuity such that the variable changes are infinitesimally small over infinitesimally small time intervals. A *Markov process* is a stochastic process in which only the present value of the variable is relevant for predicting the future evolution of the process. Most stochastic processes in finance are Markov processes, though some processes (such as the forward rate processes in the Heath, Jarrow, and Morton [1992] term structure model in Chapter 11) are non-Markovian processes.

Wiener Process

The simplest example of a continuous-time Markovian diffusion process is the well-known Brownian motion or the *Wiener process*. The behavior of a Wiener process $Z(t)$ can be understood by considering the change in its value $\Delta Z(t)$ over an infinitesimally small time interval Δt. Two basic assumptions that $\Delta Z(t)$ must satisfy for $Z(t)$ to be a Wiener process are as follows.

Assumption 1 The process $Z(t)$ is normally distributed and is given as:

$$\Delta Z(t) = Z(t + \Delta t) - Z(t) = \varepsilon_{t+\Delta t}\sqrt{\Delta t} \tag{1.1}$$

where $\varepsilon_{t+\Delta t}$ is a standardized normal variable with mean 0 and variance 1 at time t. By definition, the conditional mean and variance of $\Delta Z(t)$ are given as follows:

$$\begin{aligned} E_t(\Delta Z(t)) &= E_t(\varepsilon_{t+\Delta t}\sqrt{\Delta t}) = \sqrt{\Delta t} \times E_t(\varepsilon_{t+\Delta t}) = 0, \text{ and} \\ V_t(\Delta Z(t)) &= V_t(\varepsilon_{t+\Delta t}\sqrt{\Delta t}) = (\sqrt{\Delta t})^2 \times V_t(\varepsilon_{t+\Delta t}) = \Delta t \end{aligned} \tag{1.2}$$

Assumption 2 The values $\Delta Z(t)$ and $\Delta Z(s)$ are independently distributed for any $t \neq s$. Assumption 2 implies that $Z(t)$ is a Markov process.

A useful property of a Wiener process is that its unconditional variance grows proportionally with time. To see this, consider the change in the variable $Z(t)$ over a discrete interval from $t = 0$ to $t = T$. Divide T into N intervals of length Δt, or

$$\Delta t = \frac{T}{N}, \text{ or } T = N\Delta t \tag{1.3}$$

Then, using equation (1.1), it follows that,

$$\Delta Z((i-1)\Delta t) = Z(i\Delta t) - Z((i-1)\Delta t) = \varepsilon_{i\Delta t}\sqrt{\Delta t}, \text{ for } i = 1, 2, \ldots, N \tag{1.4}$$

Summing up from $i = 1$ to N, and assuming that the starting value of the Wiener process $Z(0) = 0$, we get:

$$\begin{aligned}\sum_{i=1}^{N} \Delta Z((i-1)\Delta t) &= \sum_{i=1}^{N} Z(i\Delta t) - Z((i-1)\Delta t) \\ &= Z(N\Delta t) - Z(0) = Z(T) - 0 = \sum_{i=1}^{N} \varepsilon_{i\Delta t}\sqrt{\Delta t}\end{aligned} \tag{1.5}$$

Using *iterated* expectations, the mean and variance of $Z(T)$ can be given as:

$$E_0(Z(T)) = E_0\left(\sum_{i=1}^{N} \varepsilon_{i\Delta t}\sqrt{\Delta t}\right) = 0 \tag{1.6}$$

$$V_0(Z(T)) = V_0\left(\sum_{i=1}^{N} \varepsilon_{i\Delta t}\sqrt{\Delta t}\right) = \left(\sum_{i=1}^{N} 1 \times (\sqrt{\Delta t})^2\right) = N\Delta t = T \tag{1.7}$$

Note that the unconditional variance of $Z(T)$ grows proportionately with time because of the specific definition of $\Delta Z(t)$ in equation (1.1). Suppose instead of this definition we used the following more general definition:

$$\Delta Z(t) = Z(t + \Delta t) - Z(t) = \varepsilon_{t+\Delta t}(\Delta t)^{\alpha}, \text{ where } \varepsilon_{t+\Delta t} \sim N(0, 1) \tag{1.8}$$

Equation (1.8) is consistent with equation (1.1) only when $\alpha = 0.5$. Now consider a value of $\alpha \neq 0.5$ in equation (1.8). In this case equation (1.7) will change as follows:

$$\begin{aligned}V_0(Z(T)) = V_0\left(\sum_{i=1}^{N} \varepsilon_{i\Delta t}(\Delta t)^{\alpha}\right) &= \left(\sum_{i=1}^{N} 1 \times (\Delta t)^{2\alpha}\right) \\ &= N\Delta t(\Delta t)^{2\alpha-1} = T(\Delta t)^{2\alpha-1}\end{aligned} \tag{1.9}$$

If $\alpha > 0.5$ in equation (1.9), then the variance of $Z(T)$ converges to zero as Δt converges to zero. On the other hand, if $\alpha < 0.5$ in equation (1.9), then the variance of $Z(T)$ converges to infinity as Δt converges to zero. Neither a zero variance nor an infinite variance are realistic, and hence $\alpha = 0.5$ is the only value that is reasonable to use in equation (1.8).

As a final observation, note that even though $\Delta Z(t)$ is a stochastic variable as defined in equation (1.1), *its square* $(\Delta Z(t))^2$ *is not stochastic* over an infinitesimally small time interval $\Delta t \to 0$. In fact, $(\Delta Z(t))^2$ is nonrandom and equals Δt. To understand this result we need to define what we mean by convergence.

A function of Δt is said to be of the order:

$$\begin{aligned} O(\Delta t) \quad &\text{if} \quad \lim_{\Delta t \to 0} \frac{f(\Delta t)}{\Delta t} \to \text{ constant, and} \\ o(\Delta t) \quad &\text{if} \quad \lim_{\Delta t \to 0} \frac{f(\Delta t)}{\Delta t} \to 0 \end{aligned} \tag{1.10}$$

In continuous-time mathematics, all terms that are of the order $o(\Delta t)$ can be ignored as $\Delta t \to 0$. Since these terms do not matter in convergence, from the definition of $\Delta Z(t)$ in equation (1.1), it follows that

$$E_t[(\Delta Z(t))^2] = E_t[(\varepsilon_{t+\Delta t})^2 \Delta t] = \Delta t \tag{1.11}$$

and

$$V_t[(\Delta Z(t))^2] = V_t[(\varepsilon_{t+\Delta t})^2 \Delta t] = \underbrace{[\Delta t]^2}_{o(\Delta t)} \overbrace{V[(\varepsilon_{t+\Delta t})^2]}^{\text{constant}} = o(\Delta t) \approx 0 \tag{1.12}$$

Since the variance of $(\Delta Z(t))^2$ is of the order $o(\Delta t)$, this term becomes insignificant as Δt becomes infinitesimally small, and it follows that $(\Delta Z(t))^2$ is not stochastic in the limit.[2] Since $(\Delta Z(t))^2$ is not stochastic, its value converges to its expectation given as:

$$(\Delta Z(t))^2 = E_t[(\Delta Z(t))^2] + o(\Delta t) = \Delta t + o(\Delta t) \approx \Delta t \tag{1.13}$$

Since $(\Delta Z(t))^2$ converges to Δt, it remains significant and leads to the celebrated Ito's lemma. But before we present Ito's lemma, we must define an Ito process, which is given next.

Ito Process

From now on we will consider only the limiting case when $\Delta t \to 0$, and use the notation $dZ(t)$ to represent $\Delta Z(t)$, and dt to represent $\Delta t \to 0$.

An Ito process is defined as follows:

$$dX(t) = X(t + dt) - X(t) = a(X, t)\, dt + b(X, t)\, dZ(t) \tag{1.14}$$

where $dX(t)$ gives the change in the X-process over the infinitesimal interval dt, and $a(X, t) = a(X(t), t)$ and $b(X, t) = b(X(t), t)$ are functions of the underlying variable $X(t)$ and time t. The expected value and the variance of $dX(t)$ are given as follows (using equation (1.2)):

$$E(dX(t)) = a(X,t)\,dt + b(X,t)E(dZ(t)) = a(X,t)\,dt \tag{1.15}$$

$$V(dX(t)) = b^2(X,t)V(dZ(t)) = b^2(X,t)\,dt \tag{1.16}$$

The X-process evolves with independent increments that are distributed normally with a mean equal to $a(X,t)dt$, and a conditional variance equal to $b^2(X,t)dt$ at any given time t. Since the X-process is stochastic, it could take a whole distribution of values $X(t)$ at a future time $t > 0$, given its value $X(0)$ at time 0. Obviously, at time 0, $X(t)$ is not known; and so it is a random variable with a probability distribution. This probability distribution is obtained using a *stochastic integral*. In fact, all stochastic integrals are probability distributions of some underlying variables over a discrete period.

The stochastic integral of the X-process given in equation (1.14) is given as:

$$\int_0^T dX(t) = \int_0^T a(X,t)\,dt + \int_0^T b(X,t)\,dZ(t) \tag{1.17}$$

To understand the mathematical intuition of this integral (without a rigorous proof), divide the time T into N intervals as follows:

$$h = \frac{T}{N}, \text{ or } T = N \times h \tag{1.18}$$

Now taking the stochastic integral of the expression on the left-hand side (L.H.S.) of equation (1.17) gives:

$$X(T) = X(0) + \int_0^T a(X,t)\,dt + \int_0^T b(X,t)\,dZ(t) \tag{1.19}$$

The first stochastic integrals on the right-hand side (R.H.S.) of equation (1.19) are expressed as follows:

$$\int_0^T a(X,t)\,dt = \lim_{h\to 0} \left(\begin{array}{l} a(X(0),0)\,h \\ +\,a(X(h),h)\,h \\ +\,a(X(2h),\ 2h)\,h \\ \quad\vdots \\ +\,a(X(T-h),\ T-h)\,h \end{array} \right) \tag{1.20}$$

The second stochastic integral on the R.H.S. of equation (1.19) is expressed as follows:

$$\int_0^T b(X,t)\,dZ(t) = \lim_{h\to 0}\begin{pmatrix} b(X(0),\ 0)[Z(h)-Z(0)] \\ +\,b(X(h),\ h)[Z(2h)-Z(h)] \\ +\,b(X(2h),\ 2h)[Z(3h)-Z(2h)] \\ \vdots \\ +\,b(X(T-h),\ T-h)[Z(T)-Z(T-h)] \end{pmatrix} \tag{1.21}$$

Using equation (1.4), the above stochastic integral in equation (1.21) can be expressed as follows:

$$\int_0^T b(X,t)\,dZ(t) = \lim_{h\to 0}\begin{pmatrix} b(X(0),\ 0)[\varepsilon_h\sqrt{h}] \\ +\,b(X(h),\ h)[\varepsilon_{2h}\sqrt{h}] \\ +\,b(X(2h),\ 2h)[\varepsilon_{3h}\sqrt{h}] \\ \vdots \\ +\,b(X(T-h),\ T-h)[\varepsilon_T\sqrt{h}] \end{pmatrix} \tag{1.22}$$

The random variable $X(T)$ has a probability distribution at time 0, which can be obtained using the stochastic integrals on the R.H.S. of equation (1.19) and their approximations in equations (1.20) and (1.22), respectively. However, the probability distribution of the variable $X(T)$ is not always straightforward to derive analytically since both $a(X,t)$ and $b(X,t)$ are themselves stochastic and are functions of the stochastic variable $X(t)$, for $0 \le t < T$. In the latter part of this chapter we investigate some special Gaussian cases for which it is easy to solve the first two moments of the stochastic integrals analytically. In other cases, the probability distribution implied by a stochastic integral is computed numerically using trees, Monte Carlo simulation, and other techniques. We are now ready to present Ito's lemma, but without a rigorous proof.

Ito's Lemma

Consider another function $Y(t) = Y(X,t)$, which depends both on $X(t)$ and t. If $Y(t)$ is twice differentiable in $X(t)$ and once differentiable in t, then the stochastic differential equation of the Y-process can be given using Ito's lemma as follows:

$$dY(t) = \frac{\partial Y}{\partial X}\,dX(t) + \frac{1}{2}\frac{\partial^2 Y}{\partial X^2}(dX(t))^2 + \frac{\partial Y}{\partial t}\,dt \tag{1.23}$$

Note that the stochastic differential of $Y(t)$ in equation (1.23) is different from the differential of $Y(t)$ in ordinary calculus. In ordinary calculus, the term $(dX(t))^2$ goes to zero, and the differential is given as:

$$dY(t) = \frac{\partial Y}{\partial X}\, dX(t) + \frac{\partial Y}{\partial t}\, dt \tag{1.24}$$

However in stochastic calculus, the term $(dX(t))^2$ contains $(dZ(t))^2$, which is of the order $O(dt)$ (see equations (1.10) and (1.13)), and so $(dX(t))^2$ remains significant. This is the main difference between ordinary calculus and stochastic calculus. By substituting the process for $dX(t)$ from equation (1.14) into equation (1.23), we get:

$$\begin{aligned} dY(t) = {} & \frac{\partial Y}{\partial X}(a(X,t)\, dt + b(X,t)\, dZ(t)) \\ & + \frac{1}{2}\frac{\partial^2 Y}{\partial X^2}(a(X,t)\, dt + b(X,t)\, dZ(t))^2 + \frac{\partial Y}{\partial t}\, dt \end{aligned} \tag{1.25}$$

Since dt^2 and $dZ(t) \times dt$ are both of the order $o(dt)$, these terms get eliminated; but the term $(dZ(t))^2 = dt$ remains of the order $O(dt)$, and equation (1.25) can be simplified to give Ito's lemma as follows:

$$dY(t) = \left(\frac{\partial Y}{\partial t} + a(X,t)\frac{\partial Y}{\partial X} + \frac{1}{2} b^2(X,t)\frac{\partial^2 Y}{\partial X^2} \right) dt + \frac{\partial Y}{\partial X} b(X,t)\, dZ(t) \tag{1.26}$$

Example 1.1 Reconsider the X-process given in equation (1.14), with $a(X,t) = aX(t)$, and $b(X,t) = bX(t)$, where a and b are two constants. The X-process can be given as:

$$dX(t) = aX(t)\, dt + bX(t)\, dZ(t) \tag{1.27}$$

The above process is the well-known geometric Brownian motion followed by the stock price under the Black and Scholes [1973] option pricing model. Now consider a function $Y(t) = \ln(X(t))$. Since $Y(t)$ is only a function of $X(t)$, and not a function of t, we have:

$$\frac{\partial Y}{\partial t} = 0, \quad \frac{\partial Y}{\partial X} = \frac{\partial \ln X}{\partial X} = \frac{1}{X}, \quad \frac{\partial^2 Y}{\partial X^2} = \frac{\partial}{\partial X}\left(\frac{1}{X}\right) = -\frac{1}{X^2}$$

Substituting the above values of partial derivatives, and $a(X,t) = aX(t)$ and $b(X,t) = bX(t)$, in equation (1.26), the Y-process can be given as follows:

$$dY(t) = \left(\frac{aX(t)}{X(t)} - \frac{1}{2}\frac{(bX(t))^2}{X^2(t)} \right) dt + \frac{bX(t)}{X(t)}\, dZ(t) \tag{1.28}$$

or

$$dY(t) = d\ln(X(t)) = (a - b^2/2)\, dt + b\,dZ(t) \tag{1.29}$$

Simple Rules of Stochastic Differentiation and Integration

Two rules of stochastic differentiation and stochastic integration are given as follows. Both these rules hold regardless of the specification of the lower boundary of these integrals, even though we have put zero as the lower boundary.

Rule 1 Consider the following stochastic integral:

$$\int_0^t dX(v) = \int_0^t a(X,\ v)\, dv + \int_0^t b(X,\ v)\, dZ(v) \tag{1.30}$$

The stochastic differential equation corresponding to equation (1.30) is given as:

$$dX(t) = a(X, t)\, dt + b(X, t)\, dZ(t) \tag{1.31}$$

Rule 2 Consider the following stochastic integral:

$$\int_0^t dX(v) = \int_0^t a(X,\ v,\ t)\, dv + \int_0^t b(X,\ v,\ t)\, dZ(v) \tag{1.32}$$

The stochastic differential equation corresponding to equation (1.32) is given as:

$$\begin{aligned} &\left[a(X,\ t,\ t) + \int_0^t \frac{\partial a(X,\ v,\ t)}{\partial t}\, dv + \int_0^t \frac{\partial b(X,\ v,\ t)}{\partial t}\, dZ(v)\right] dt \\ &\quad + b(X,\ t,\ t)\, dZ(t) \end{aligned} \tag{1.33}$$

Though most finance applications require rule 1, certain applications such as the derivation of the short rate process under the Heath, Jarrow, Morton [1992] term structure model require rule 2.

Obtaining Unconditional Mean and Variance of Stochastic Integrals under Gaussian Processes

Many financial valuation problems require the computation of risk-neutral expectations of an underlying asset price. Since asset price distributions can be represented as stochastic integrals, computation of the expectations can be accomplished by knowing the probability distribution implied by the stochastic integral. For the case of Gaussian (i.e., normally distributed) processes, it is possible to obtain the mean and variance of the underlying variable by applying rules that give the mean and variance of a stochastic integral. These rules are given as follows:

Rule 3 Let $Z(t)$ be a Wiener process and $b(t)$ be a deterministic (i.e., nonrandom) function of time t. Then for all $T \geq t$, the stochastic integral:

$$g(t,\ T) = \int_t^T b(v)\, dZ(v) \tag{1.34}$$

is a Gaussian process with time t mean and variance given as follows:

$$\begin{aligned} E_t[g(t,\ T)] &= 0 \\ V_t[g(t,\ T)] &= \int_t^T b^2(v)\, dv \end{aligned} \tag{1.35}$$

It is easy to see why the expected value of $g(t,\ T)$ equals zero. However, to derive the expression of the variance of $g(t,\ T)$, divide $T - t$ into N intervals of length h, such as $(T - t)/N = h$. Using equation (1.4), we have,

$$g(t,\ T) = \int_t^T b(v)\, dZ(v) = \lim_{h \to 0} \begin{pmatrix} b(t)[\varepsilon_{t+h}\sqrt{h}] \\ +\, b(t+h)[\varepsilon_{t+2h}\sqrt{h}] \\ +\, b(t+2h)[\varepsilon_{t+3h}\sqrt{h}] \\ \vdots \\ +\, b(T-h)[\varepsilon_T\sqrt{h}] \end{pmatrix} \tag{1.36}$$

where the N error terms ε on the R.H.S. of equation (1.36) are all standard normal variables with mean 0 and variance 1. Since all of the mean-zero error terms are independent, the variance of $g(t, T)$ is given by taking the time t expectation, given as follows:

$$\begin{aligned} V_t[g(t,\ T)] &= \lim_{h \to 0} E_t \begin{pmatrix} b^2(t)[\varepsilon_{t+h}\sqrt{h}]^2 \\ +\, b^2(t+h)[\varepsilon_{t+2h}\sqrt{h}]^2 \\ +\, b^2(t+2h)[\varepsilon_{t+3h}\sqrt{h}]^2 \\ \vdots \\ +\, b^2(T-h)[\varepsilon_T\sqrt{h}]^2 \end{pmatrix} \\ &= \lim_{h \to 0} \begin{pmatrix} b^2(t)h \\ +\, b^2(t+h)h \\ +\, b^2(t+2h)h \\ \vdots \\ +\, b^2(T-h)h \end{pmatrix} = \int_t^T b^2(v)\, dv \end{aligned} \tag{1.37}$$

Rule 4 Let $Z(t)$ be a Wiener process and $b(t)$ and $c(t)$ be deterministic functions of time t. Define a new variable $h(t,\ T)$ as follows:

$$h(t,\ T) = \int_t^T c(v) g(t,\ v)\, dv$$

where $g(t,\ T)$ is a Gaussian variable defined in equation (1.34). Then $h(t,\ T)$ is a Gaussian process with time t mean and variance given as follows:

$$\begin{aligned} E_t[h(t,\ T)] &= 0 \\ V_t[h(t,\ T)] &= \int_t^T b^2(v) \left(\int_v^T c(u)\, du \right)^2 dv \end{aligned} \tag{1.38}$$

The proof of equation (1.38) is more involved, but can be given using the same technique used for proving rule 3.

Rule 5 If a variable Y follows a Gaussian distribution with mean $E(Y)$ and variance $V(Y)$, then the mean of the variable $X = e^Y$ is given as follows:

$$E(X) = e^{E(Y)+(1/2)V(Y)} \tag{1.39}$$

Examples of Gaussian Stochastic Integrals

As mentioned earlier, the probability distributions of stochastic integrals are not easy to solve analytically. For example, as shown in equation (1.19), since both $a(X, t)$ and $b(X, t)$ are stochastic and depend upon $X(t)$, the probability distribution of the stochastic integral is not easy to obtain. However, for certain specific functional forms of $a(X, t)$ and $b(X, t)$, transformations of the function X can be represented as stochastic integrals with Gaussian distributions. The mean and variance of these Gaussian-distributed stochastic integrals can be derived using the results in the previous section. The following two examples demonstrate this transform technique. The first example obtains the stochastic integral of the log of the stock price when stock price follows a geometric Brownian motion (as in the model of Black and Scholes [1973]), and the second example obtains the stochastic integral of the short rate, when the short rate follows the Ornstein-Uhlenbeck process under the term structure model of Vasicek [1977].

Example 1.2 This example derives the log stock price distribution under the geometric Brownian motion. Let the variable $X(t)$ represent the price of a stock in equation (1.27) in Example 1.1. Consider the stochastic integral of the X-process given as follows:

$$X(T) = X(t) + \int_t^T a\, X(v)\, dv + \int_t^T b\, X(v)\, dZ(v) \tag{1.40}$$

As mentioned earlier, since the X-variable appears as a part of the integrand, an analytical solution for the probability distribution of this integral does not seem feasible. However, we show that the log transform $Y(T) = \ln(X(T))$ is Gaussian, and hence the mean and variance of $Y(T)$ can be computed using the results given in equations (1.34) and (1.35). The stochastic differential equation of the Y-process was obtained in equation (1.29) in Example 1.1, using Ito's lemma. Taking the stochastic integral of the Y-process, we get:

$$Y(T) = Y(t) + \int_t^T (a - b^2/2)\, dv + \int_t^T b\, dZ(v) \tag{1.41}$$

or

$$Y(T) = Y(t) + (a - b^2/2)(T - t) + \int_t^T b\, dZ(v) \tag{1.42}$$

Applying equations (1.34) and (1.35) to the stochastic integral on the R.H.S. of equation (1.42), the time t mean and variance of the $Y(T)$ are given as follows:

$$E_t(Y(T)) = Y(t) + (a - b^2/2)(T - t) \tag{1.43}$$

$$V_t(Y(T)) = \int_t^T b^2\, dv = b^2(T - t) \tag{1.44}$$

The variable $Y(T) = \ln(X(T))$ is distributed normally, implying that the stock price $X(T)$ is distributed *lognormally*. The mean of the stock price $X(T)$ can be computed as follows: Since $Y(T) = \ln(X(T))$, it follows that $X(T) = e^{Y(T)}$. Further, since $Y(T)$ is a Gaussian process, the mean of $X(T)$ can be given using rule 5 (equation (1.39)) as follows:

$$\begin{aligned} E_t(X(T)) &= e^{E_t(Y(T)) + (1/2)V_t(Y(T))} \\ &= e^{Y(t) + (a - b^2/2)(T-t) + (1/2)b^2(T-t)} = e^{Y(t) + a(T-t)} = X(t)e^{a(T-t)} \end{aligned} \tag{1.45}$$

Hence, the current value of the stock is given as its future expected value, discounted by the drift rate a, or:

$$X(t) = \frac{E_t(X(T))}{e^{a(T-t)}} \tag{1.46}$$

Example 1.3 This example derives the short rate distribution under the Ornstein-Uhlenbeck process. Let the variable $r(t)$ represent the instantaneous short rate following the Ornstein-Uhlenbeck process given by Vasicek [1977] as follows:

$$dr(t) = \alpha(m - r(t))\,dt + \sigma\,dZ(t) \tag{1.47}$$

where the parameters α, m, and σ are constants. The stochastic integral of equation (1.47) is given as:

$$r(T) = r(t) + \int_t^T \alpha(m - r(v))\,dv + \int_t^T \sigma\,dZ(v) \tag{1.48}$$

To remove the variable r from the first integrand, consider the transform $Y(t) = r(t)\exp(\alpha t)$. Using Ito's lemma, the stochastic differential equation of Y-process can be obtained. To apply Ito's lemma, we need the following three partial derivatives:

$$\frac{\partial Y}{\partial t} = \alpha e^{\alpha t} r(t), \quad \frac{\partial Y}{\partial r} = e^{\alpha t}, \quad \frac{\partial^2 Y}{\partial r^2} = \frac{\partial}{\partial r}(e^{\alpha t}) = 0 \tag{1.49}$$

Using the above partial derivatives and Ito's lemma given in equation (1.26), we get:

$$\begin{aligned} dY(t) &= e^{\alpha t}\,dr(t) + \alpha e^{\alpha t} r(t)\,dt \\ &= e^{\alpha t}(\alpha(m - r(t))\,dt + \sigma\,dZ(t)) + \alpha e^{\alpha t} r(t)\,dt \\ &= e^{\alpha t}\alpha m\,dt + e^{\alpha t}\sigma\,dZ(t) \end{aligned} \tag{1.50}$$

The stochastic integral of $Y(T)$ can be written as:

$$\begin{aligned} Y(T) &= Y(t) + \int_t^T (e^{\alpha v}\alpha m)\,dv + \int_t^T (e^{\alpha v}\sigma)\,dZ(v) \\ &= Y(t) + m(e^{\alpha T} - e^{\alpha t}) + \int_t^T (e^{\alpha v}\sigma)\,dZ(v) \end{aligned} \tag{1.51}$$

Substituting $Y(T) = e^{\alpha T} r(T)$ and $Y(t) = e^{\alpha t} r(t)$ in equation (1.51) and simplifying we get:

$$r(T) = e^{-\alpha(T-t)} r(t) + m(1 - e^{-\alpha(T-t)}) + e^{-\alpha T}\sigma \int_t^T e^{\alpha v}\,dZ(v) \tag{1.52}$$

Applying equation (1.34) to equation (1.52), the mean of $r(T)$ is given as follows:

$$E_t(r(T)) = e^{-\alpha(T-t)} r(t) + m(1 - e^{-\alpha(T-t)}) \tag{1.53}$$

As T goes to infinity, the long-term mean of $r(T)$ converges to m. By applying equation (1.35) to the stochastic integral on the R.H.S. of equation (1.52), the variance of $r(T)$ is given as follows:

$$\begin{aligned} V_t(r(T)) &= e^{-2\alpha T}\int_t^T (e^{2\alpha v}\sigma^2)\,dv = e^{-2\alpha T}\sigma^2\left(\frac{e^{2\alpha T}-e^{2\alpha t}}{2\alpha}\right) \\ &= \sigma^2\left(\frac{1-e^{-2\alpha(T-t)}}{2\alpha}\right) \end{aligned} \tag{1.54}$$

MIXED JUMP-DIFFUSION PROCESSES

As mentioned earlier, diffusion processes display continuity such that the variable changes are infinitesimally small over infinitesimally small time intervals. Jump processes display discontinuity such that the variable changes are non-infinitesimally large even over infinitesimally small time intervals. Jump processes experience no change with a probability $1-\lambda\,dt$ and a discontinuous change with an infinitesimally small probability λdt, over the infinitesimally small time interval dt, where λ defines the intensity of the process. Many processes used in finance are a mix of a diffusion process and a jump process, such that the variable experiences a continuous change most of the time, while every once in a while the variable experiences a discontinuous change.

The Jump-Diffusion Process

Consider the mixed jump-diffusion process given as follows:

$$dX(t) = a(X,t)\,dt + b(X,t)\,dZ(t) + h(J,\ X,t)\,dN(\lambda) \tag{1.55}$$

Equation (1.55) is identical to the diffusion process given in equation (1.14), except for the last term. This term has two components. The term $dN(\lambda)$ represents the Poisson process, which is independent of $dZ(t)$, and is given as follows:

$$dN(\lambda) = \left\{\begin{array}{l} 0,\ \text{with probability } 1-\lambda\,dt,\ \text{and} \\ 1,\ \text{with probability } \lambda\,dt \end{array}\right\} \tag{1.56}$$

When $dN(\lambda)$ equals 1 (with probability $\lambda\,dt$), then the magnitude of jump equals $h(J,X,t)\,dN(\lambda) = h(J,X,t)$. The variable $h(J,X,t)$ is a function of the random variable J that can have some arbitrarily specified distribution (i.e., binomial, Gaussian, etc.), $X(t)$, and t. The variable J is independent of both $dZ(t)$ and $dN(\lambda)$.

Ito's Lemma for the Jump-Diffusion Process

Let $Y(t) = Y(X, t)$ be a twice-differentiable function of $X(t)$ (where the X-process is given by equation (1.55)) and once differentiable function of t. Then, the stochastic differential equation of Y-process is given as follows:

$$\begin{aligned} dY(t) = &\left(\frac{\partial Y}{\partial t} + a(X,t)\frac{\partial Y}{\partial X} + \frac{1}{2}b^2(X,t)\frac{\partial^2 Y}{\partial X^2}\right) dt + \frac{\partial Y}{\partial X} b(X,t)\, dZ(t) \\ &+ [Y((X + h(J,\ X,\ t)),\ t) - Y(X,t)]\, dN(\lambda) \end{aligned} \tag{1.57}$$

Equation (1.57) generalizes Ito's lemma to jump-diffusion processes. The top line on the R.H.S. of equation (1.57) gives the terms of Ito's lemma with respect to the diffusion process. The expression in the bottom line of the R.H.S. of equation (1.57) gives the additional term due to the jump component.

Example 1.4 Reconsider the X-process given in equation (1.27), but with an added jump component as follows:

$$dX(t) = aX(t)\, dt + bX(t)\, dZ(t) + JX(t)\, dN(\lambda) \tag{1.58}$$

Equation (1.58) is used by Merton [1976] to derive option prices under the jump-diffusion process. Now, consider the stochastic differential of the variable $Y(t) = \ln(X(t))$. Using equations (1.55) and (1.57), the Y-process is given as:

$$\begin{aligned} dY(t) = d\ln(X(t)) = &\ (a - b^2/2)\, dt + b dZ(t) \\ &+ \big(\ln(X(t) + JX(t)) - \ln X(t)\big)\, dN(\lambda) \end{aligned} \tag{1.59}$$

Since $\ln(X(t)+JX(t)) = \ln(X(t)(1+J)) = \ln(X(t)) + \ln(1+J)$, equation (1.59) simplifies to the following equation:

$$dY(t) = d\ln(X(t)) = (a - b^2/2)\, dt + b dZ(t) + \ln(1 + J)\, dN(\lambda) \tag{1.60}$$

The Ito's lemma for jumps given in equation (1.57) is very useful and is applied in Chapters 5, 6, 7, and 10 of this book to obtain the stochastic jump-diffusion process for the bond price when the interest rate or the state variable is driven by a jump-diffusion process.

NOTES

1. For example, Duffie [2001].
2. It is easy to show that skewness, kurtosis, and all other higher moments of $(\Delta Z(t))^2$ also are of the order $o(\Delta t)$.