# Chapter 2
# Stochastic Processes for Asset Price Modelling

**Abstract** This chapter gives an intuitive appreciation and review of many important aspects of the stochastic processes that have been used to model asset price processes. We will be interested in a probabilistic description of the time evolution of asset prices. After imposing some structure on the stochastic process for the return on the asset, this chapter introduces Markov processes, time evolution of conditional probabilities, continuous sample paths, and the Fokker–Planck and Kolmogorov equations.

## 2.1 Introduction

We shall be much concerned with how asset prices evolve over time. It was realised early in the development of the modern theory of finance that since asset prices are evolving randomly over time the best description of price behaviour would be a probabilistic one, which involves using ideas from the theory of stochastic processes. The theory of stochastic processes is not an easy theory to master, since many of its important concepts were developed roughly simultaneously in a variety of disciplines such as electrical engineering, theoretical physics and pure mathematics. The perspective taken in each of these disciplines is slightly different and the same concepts can be presented at vastly different levels of mathematical abstraction.

Our aim in this book is not at all to present a fully rigorous discussion of the theory of stochastic processes. Rather we merely attempt to give an intuitive appreciation and review of those aspects of the theory which have found application in modern finance theory. In putting together the discussion and viewpoint on stochastic processes from this chapter up to Chap. 8 we have drawn heavily on Malliaris and Brock (1982), Astrom (1970), Harrison (1990), Baxter and Rennie (1996) with some ideas from Gardiner (1985) and Horsthemke and Lefever (1984). A more complete mathematical treatment of stochastic processes and stochastic differential equations may be found in the books by Oksendal (2003), Krylov (1995) and Kijima (2002). The book by Oksendal (2003) is particularly recommended.

We will be interested in a probabilistic description of the time evolution of asset prices. Empirical examination of time series of asset prices and asset returns



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_2

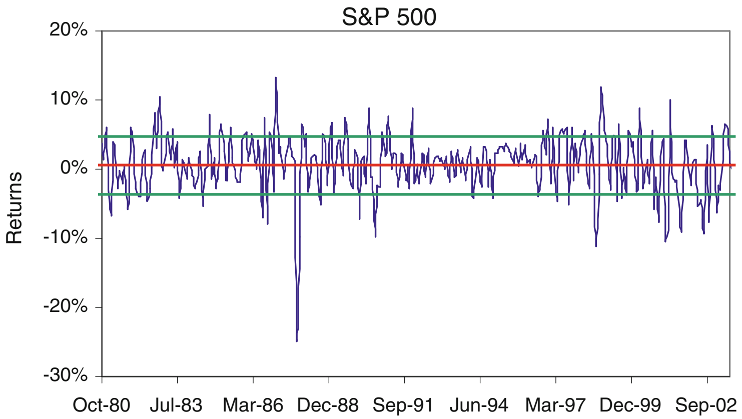


**Fig. 2.1** Intuitive notion of evolution of returns

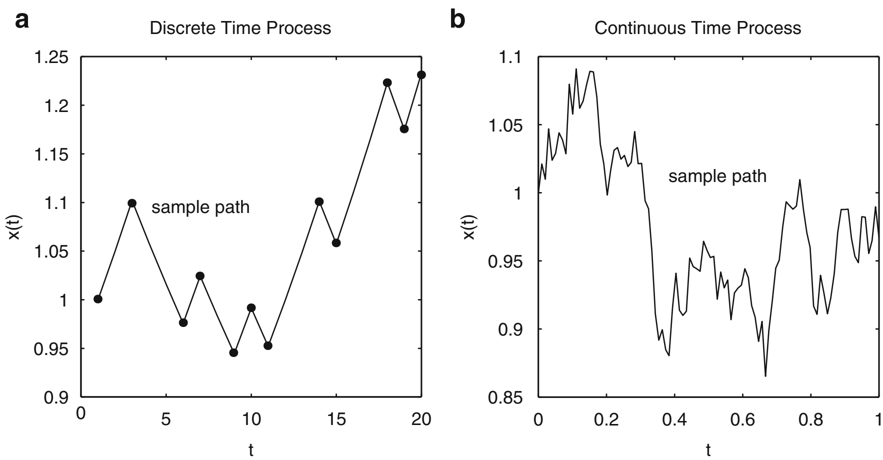


**Fig. 2.2** Sample paths of stochastic processes

suggests that at least intuitively we might initially think of the *return* on the asset as consisting of an average mean component (which we might regard as more or less certain) and some volatile, stochastic component about this mean, as shown in Fig. 2.1 which shows the time series of monthly returns on the $S\&P\,500$ from January 1980 to October 2003 together with the mean and one standard derivation band of the entire series. Already with this intuitive notion we are imposing some structure on the stochastic process for the return on the asset. Much of what we do in Chaps. 2 and 3 will be to give some mathematical description to this intuitive notion.

To make the discussion a little more formal let $x(t)$ denote either the price of or return on the asset at time $t$. In Fig. 2.2 we represent a typical sequence of prices which would be observed over a given time period. Since, depending on the application, we need to concentrate on prices or returns in both *discrete time* and *continuous time*, both are represented in Fig. 2.2.[1]

The stochastic process for the prices (or returns) may be thought of as a family of random variables $\{x(t) \mid t \in \mathscr{T}\}$, where $\mathscr{T}$ denotes the set of values to which the time parameter $t$ belongs, more formally known as the *index set*. For the *discrete time process* in Fig. 2.2a the index set is the set of non-negative integers $\mathscr{T} = \{0, 1, 2, \ldots\}$, whilst for the *continuous time process* in Fig. 2.2b the index set is the set of all $t$ between 0 and infinity, i.e. $\mathscr{T} = \{t \mid 0 \leq t < \infty\}$.

The set of values which $x(t)$ may take is known as the *sample space* and is denoted $\Omega$. For the price paths illustrated in Fig. 2.2 the sample space is all values from 0 to infinity. Generally we will be interested in prices belonging to a subset of $\Omega$, such as the set of prices illustrated in the shaded area in Fig. 1.2. In more formal discussions we may see the stochastic process denoted as $\{x(t, \omega) \mid t \in \mathscr{T}, \omega \in \Omega\}$. A set of particular values arising from the stochastic process is known as a *realisation* of the process, or a *sample path*. In the preceding formal notation, the $\omega$ refers to one sample path out of the set $\Omega$ of all possible sample paths.

A major technical problem in the theory of stochastic processes involves assigning a probability distribution or more formally a probability measure to subsets $\Omega$ of the sample space $\omega$. To do this mathematically correctly requires a great deal of measure theory, however provided the stochastic process assumed for $x(t)$ is not too "wild" then we are able to proceed with a fairly intuitive understanding of probability distributions. However to deal with more sophisticated processes (e.g. Lévy processes) then we do need to resort to a more formal mathematical description.

## 2.2 Markov Processes

In order to put some mathematical flesh on the basic notion of a stochastic process we need to introduce the concept of *joint probability density function*. As we said at the end of the previous section, proving that such a density function can be found is an intricate mathematical problem which we shall not touch on in this book. Malliaris and Brock (1982) outline some of the intricacies involved and give appropriate references (see their Chapter 1 and Section 7).

Given values $x_1, x_2, x_3, \ldots, x_k$ of the asset price $x(t)$ at times $t_1, t_2, t_3, \ldots, t_k$, we assume that we can obtain a joint probability density function

$$p(x_k, t_k; x_{k-1}, t_{k-1}; \ldots; x_1, t_1) \qquad \text{(note that } t_1 \leq t_2 \leq \cdots \leq t_k\text{)}$$

[1]In terms of concepts to be developed later, Fig. 2.2a represents the simulation of a binomial process, whilst Fig. 2.2b is the simulation of a geometric Brownian motion process starting at $x_o = 1$.

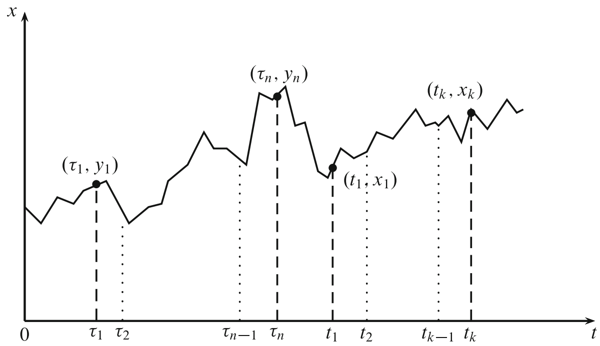


**Fig. 2.3** Conditional probability of a sample path

which measures the joint probability that $x(t_1) = x_1, x(t_2) = x_2, \dots, x(t_k) = x_k$. Using joint probability density functions, we can also define *conditional probability density functions:*

$$p(x_k, t_k; \dots; x_2, t_2; x_1, t_1 \mid y_n, \tau_n; \dots; y_2, \tau_2; y_1, \tau_1) = \frac{p(x_k, t_k; \dots; x_2, t_2; x_1, t_1; y_n, \tau_n; \dots; y_2, \tau_2; y_1, \tau_1)}{p(y_n, \tau_n; \dots; y_2, \tau_2; y_1, \tau_1)}. \tag{2.1}$$

The left-hand side of (2.1) is the probability that the price sequence $(x_1, t_1; x_2, t_2; \dots; \ x_k, t_k)$ will be observed *given that* the price sequence $(y_1, \tau_1; y_2, \tau_2; \dots; y_n, \tau_n)$ has just been observed over the previous periods $\{\tau_1, \tau_2, \dots, \tau_n\}$; see Fig. 2.3.

On the right-hand side of (2.1) the probability on the top line is that of observing the price sequence $(y_1, \tau_1), (y_2, \tau_2), \dots, (y_n, \tau_n), (x_1, t_1), \dots, (x_k, t_k)$, whilst the probability on the bottom line is that of observing the price sequence $(y_1, \tau_1), \dots, (y_n, \tau_n)$. To see the sense of this last formula think of the probabilities as observed frequencies and suppose the $\{y_i\}$ represent a sequence of prices growing at a rate of 3 % and the $\{x_i\}$ represent a sequence of prices growing at a rate of 4 %. Then the formula states that the probability of observing a 4 % rise, given that a 3 % rise has occurred equals the frequency of 3 % rises followed by 4 % rises divided by the frequency of 3 % rises.

A very simple kind of stochastic process that we might deal with is a completely independent one. A stochastic process is said to be completely *independent* if the probability of observing a given price at time $t$ is completely independent of the probability of observing some price at any other time. This allows us to write the joint probability density function as a product of independent probabilities, so that

$$p(x_k, t_k; \dots; x_2, t_2; x_1, t_1) = p(x_k, t_k) \dots p(x_2, t_2) p(x_1, t_1). \tag{2.2}$$

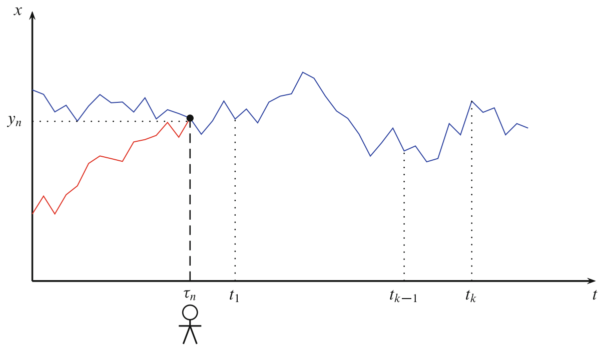


**Fig. 2.4** A Markov process

The next most simple stochastic process we might deal with is a *Markov process* in which knowledge only of the present state of the process is relevant to the future evolution of the process. Referring to the price sequences in Fig. 2.3 this idea may be expressed in terms of conditional probabilities as

$$\begin{aligned} &p(x_k, t_k; \ldots; x_2, t_2; x_1, t_1 \mid y_n, \tau_n; \ldots; y_2, \tau_2; y_1, \tau_1) \\ &\quad = p(x_k, t_k; \ldots; x_2, t_2; x_1, t_1 \mid y_n, \tau_n). \end{aligned} \tag{2.3}$$

This idea is illustrated in Fig. 2.4 where we see two possible paths arriving at $(y_n, \tau_n)$. Irrespective of which path has been followed to arrive at $(y_n, \tau_n)$, the future evolution from $\tau_n$ is only conditional on $(y_n, \tau_n)$. Markov processes are clearly related to the efficient markets concept.

The Markov assumption (i.e. Eq. (2.3)) is particularly important because it enables us to define all relevant joint probability density functions in terms of simple conditional probabilities, such as $p(x_1, t_1 \mid y_1, \tau_1)$.

To see this consider the following manipulations over the successive times $\tau_n, t_1, t_2$. By the definition of conditional probability density (Eq. (2.1)) and conditioning on $y_n, \tau_n$,

$$p(x_2, t_2; x_1, t_1; y_n, \tau_n) = p(x_2, t_2; x_1, t_1 \mid y_n, \tau_n) p(y_n, \tau_n). \tag{2.4}$$

But conditioning on $x_1, t_1; y_n, \tau_n$ we have

$$p(x_2, t_2; x_1, t_1; y_n, \tau_n) = p(x_2, t_2 \mid x_1, t_1; y_n, \tau_n) p(x_1, t_1; y_n, \tau_n). \tag{2.5}$$

Combining these last two equations we obtain

$$\begin{aligned} p(x_2, t_2; x_1, t_1 \mid y_n, \tau_n) &= \frac{p(x_2, t_2 \mid x_1, t_1; y_n, \tau_n) p(x_1, t_1; y_n, \tau_n)}{p(y_n, \tau_n)} \\ &= p(x_2, t_2 \mid x_1, t_1) p(x_1, t_1 \mid y_n, \tau_n). \end{aligned} \tag{2.6}$$

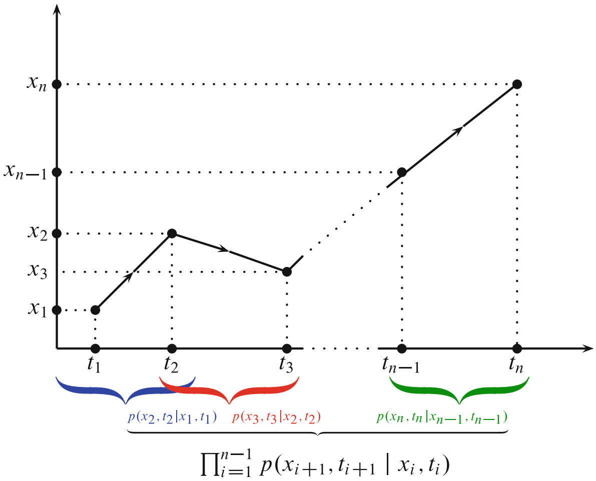


**Fig. 2.5** Illustrating the joint density as a product of conditional density functions

The first equality is due to the definition of conditional probability (2.1), the second follows by the Markovian assumption and uses again of (2.1). Substituting (2.6) to (2.4) we finally have

$$p(x_2, t_2; x_1, t_1; y_n, \tau_n) = p(x_2, t_2 \mid x_1, t_1) p(x_1, t_1 \mid y_n, \tau_n) p(y_n, \tau_n),$$

which simply states that the joint probability density function over the times $\tau_n, t_1, t_2$ is the product of the conditional probability densities over successive time intervals.

The same argument can be extended to any number of realisations of the stochastic process over successive times, to yield[2]

$$p(x_n, t_n; \ldots; x_2, t_2; x_1, t_1) = p(x_1, t_1) \prod_{i=1}^{n-1} p(x_{i+1}, t_{i+1} \mid x_i, t_i). \tag{2.7}$$

The essential feature of (2.7) from the point of view of applications is that the joint density function can be expressed as a product of conditional density functions (over successive time intervals), as illustrated in Fig. 2.5. Thus a statistical description of price and return dynamics is reduced to a description of the conditional density function.

Certainly Markov processes provide the most convenient tool for the modelling of asset prices and returns as we shall see throughout this book. On the one hand, they accord very nicely with the notion of efficiency of financial markets. On the

[2]We recall the product notation $\Pi_{i=1}^{n} X_i = X_1 X_2 \cdots X_n$. Note that in (2.7) we change slightly the notation and take $(x_1, t_1)$ as the initial point.

other hand, they allow us to make use of the highly developed theories of diffusion processes and semi-martingale integration. However it is nevertheless the case that non-Markovian processes do also play a role in the modelling of some aspects of financial markets behaviour. This is the case for instance when considering stochastic volatility models. It is more particularly the case in the modelling of interest rate sensitive derivative securities. It turns out that the most natural (and general) process for modelling the dynamic evolution of the yield curve is a non-Markovian one. We shall see in Part II that rather than working in a non-Markovian framework it turns out to be more convenient to find ways to reduce the non-Markovian process to a Markovian system of higher dimension. In this way the great mathematical convenience of Markovian processes is preserved.

## 2.3 The Time Evolution of Conditional Probabilities

As discussed in the previous sub-section, the Markov assumption implies that in order to obtain a statistical description of prices in a dynamically evolving environment we need to know how the conditional probability density functions evolve over time.

The equation which allows us to do this is the *Chapman–Kolmogorov equation* which is a simple consequence of the Markovian assumption. If $t_1 < t_2 < t_3$ then the Chapman–Kolmogorov equation states that

$$p(x_3, t_3 \mid x_1, t_1) = \int p(x_3, t_3 \mid x_2, t_2) p(x_2, t_2 \mid x_1, t_1) dx_2. \tag{2.8}$$

To see the sense of this equation consider the path $I$ in Fig. 2.6. The probability of going from $x_1$ at $t_1$ to $x_3$ at $t_3$ is the product of the probability of going from $(x_1, t_1)$ to $(x_2, t_2)$, with the probability of going from $(x_2, t_2)$ to $(x_3, t_3)$. The integral in (2.8) sums over all such probabilities by ranging over all possible paths through values $x_2$ at $t_2$. Figure 2.6 shows three such paths.

The Chapman–Kolmogorov equation is in fact a complex, nonlinear functional equation due to the fact that so far the nature of the stochastic process has been left very general. In order to reduce it to a form easier to deal with mathematically we need to put more restrictions on the nature of the stochastic process. In particular the magnitude and type of change that can occur in $x$ from one time period to the next. In particular it can be shown that if the price changes are small over small intervals of time in a way to be made more precise below then the Chapman–Kolmogorov equation reduces to a partial differential equation for the conditional probability which has a remarkable similarity to the partial differential equation governing stock option prices. We will eventually show how these two partial differential equations are related.

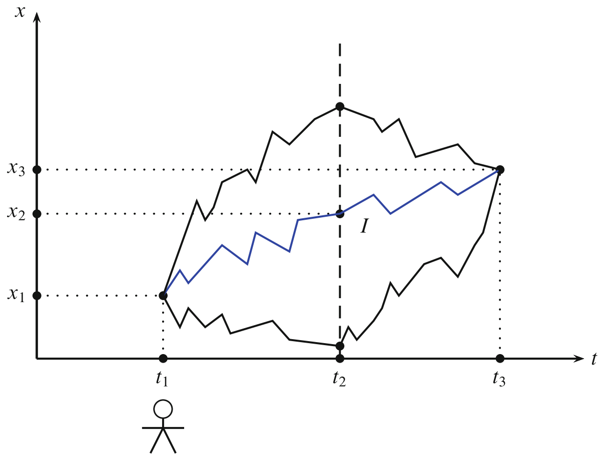


**Fig. 2.6** The Chapman–Kolmogorov equation

## 2.4 Processes with Continuous Sample Paths

As just indicated we shall for the moment focus on price (or return) processes that change by small amounts over small intervals of time. The notion that the price changes by small amounts over a small interval of time is made mathematically more precise by introducing stochastic processes having continuous sample paths.

The mathematical condition that needs to be imposed on the conditional probabilities in order that the sample paths be continuous functions of time is

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} p(x, t + \Delta t \mid z, t) dx = 0, \tag{2.9}$$

for any $\varepsilon > 0$. The sense of this condition is easily understood by referring to Fig. 2.7. Typically $\varepsilon$ would be small and the set $|x - z| > \varepsilon$ (indicated by the hashed region in Fig. 2.7) represents the set of prices $x$ at time $t + \Delta t$ which are further than a distance $\varepsilon$ from the original price of $z$ at time $t$. The probability $p(x, t + \Delta t \mid z, t)$ is the probability of observing the price $x$ at time $t + \Delta t$, given that the price at time $t$ is $z$. Thus the integral in Eq. (2.9) represents the total probability of observing at time $t + \Delta t$ a price which is further than $\varepsilon$ from the current price $z$ at time $t$. The overall condition in Eq. (2.9) states that this probability must decline more rapidly than $\Delta t$ as $\Delta t$ becomes smaller and smaller (e.g. the total probability could be proportional to $(\Delta t)^2$).

The condition in Eq. (2.9) is known as the *Lindeberg condition*, and stochastic processes whose conditional probabilities satisfy it will be expected to experience small changes in $x$ over small intervals of time. Such stochastic processes will also display continuous sample paths as has already been stated.

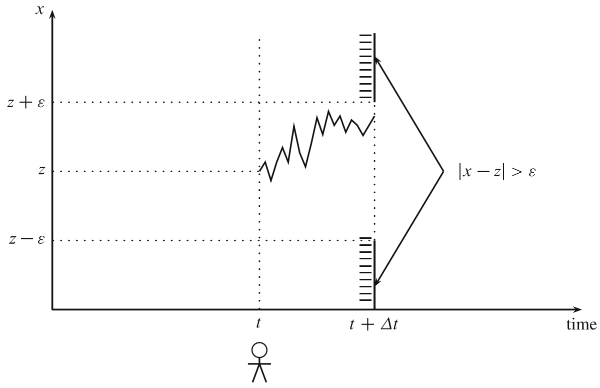


**Fig. 2.7** The Lindeberg condition; the price changes remain within the band $(z + \varepsilon, z - \varepsilon)$

At this point it may be worth considering two particular forms of conditional probability functions, one which satisfies the Lindeberg condition and one which does not.

## 2.4.1 Brownian Motion

Consider the conditional probability density given by the formula

$$p(x, t + \Delta t \mid z, t) = \frac{1}{\sqrt{2\pi \Delta t}\,\sigma} \exp\left[\frac{-(x-z)^2}{2\sigma^2 \Delta t}\right], \tag{2.10}$$

meaning that $x$ is normally distributed, $x \sim N(z, \sigma^2 \Delta t)$, centred on the current price $z$ at time $t$ and having variance $\sigma^2 \Delta t$. According to this distribution the prices expected at $t + \Delta t$ are distributed normally about the current price $z$ at $t$, as illustrated in Fig. 2.8. As $\Delta t$ becomes smaller, the distribution becomes more peaked, thereby reducing the probability that the price at $t + \Delta t$ will be very far from $z$. It is a simple (albeit tedious) exercise in integration to show that the Lindeberg condition is indeed satisfied for this distribution (the details are given in Appendix 2.2).

The stochastic process whose conditional probabilities are given by Eq. (2.10) is known as *Brownian Motion*, and is widely applied in financial economics, partly because of its mathematical tractability, and partly also because empirical studies indicate that many (though no means all) important asset prices (or returns) are reasonably well modelled by it.

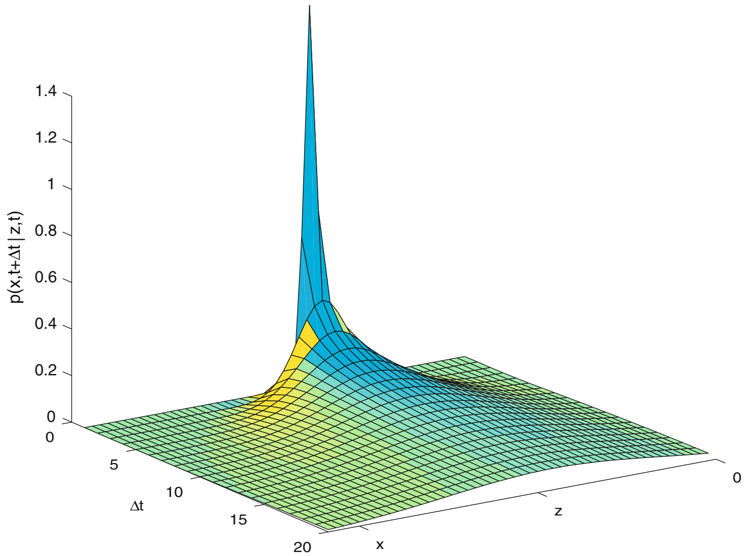


**Fig. 2.8** Conditional probability density of Brownian motion. Note how decreasing $\Delta t$ "squeezes" the distribution around the current price $z$

### 2.4.2 *The Cauchy Process*

In order that the reader not gain the impression that all bell shaped conditional probability distributions satisfy the Lindeberg condition and therefore give rise to stochastic processes having continuous sample paths, consider the distribution with conditional probability density

$$p(x, t + \Delta t \mid z, t) = \frac{\Delta t}{\pi[(x - z)^2 + (\Delta t)^2]}. \tag{2.11}$$

This distribution also has the general bell shape, shown in Fig. 2.8 for the Brownian Motion, which also becomes more peaked as $\Delta t$ becomes smaller. However it can be shown that this distribution, known as the *Cauchy distribution*, does *not* satisfy the Lindeberg condition (see Appendix 2.3) and therefore $x$ is not a continuous process. This means that as $\Delta t$ becomes small, the probability of observing a price well away from current price $z$ does not become small quickly enough, leaving a positive probability that there will be large jumps in the price from time to time.

Indeed simulations of both processes indicate that the Cauchy process exhibits large jumps not infrequently, as illustrated by the simulations in Fig. 2.9.[3] There we

[3]The simulations were calculated as follows. For the Brownian motion process, the path was calculated using $x_{i+1} = x_i + x_i \mu \Delta t + x_i \sigma \sqrt{\Delta t}\ \xi_i$, where $\mu = 0.1$, $\sigma = 0.2$, $\Delta t = 0.002$

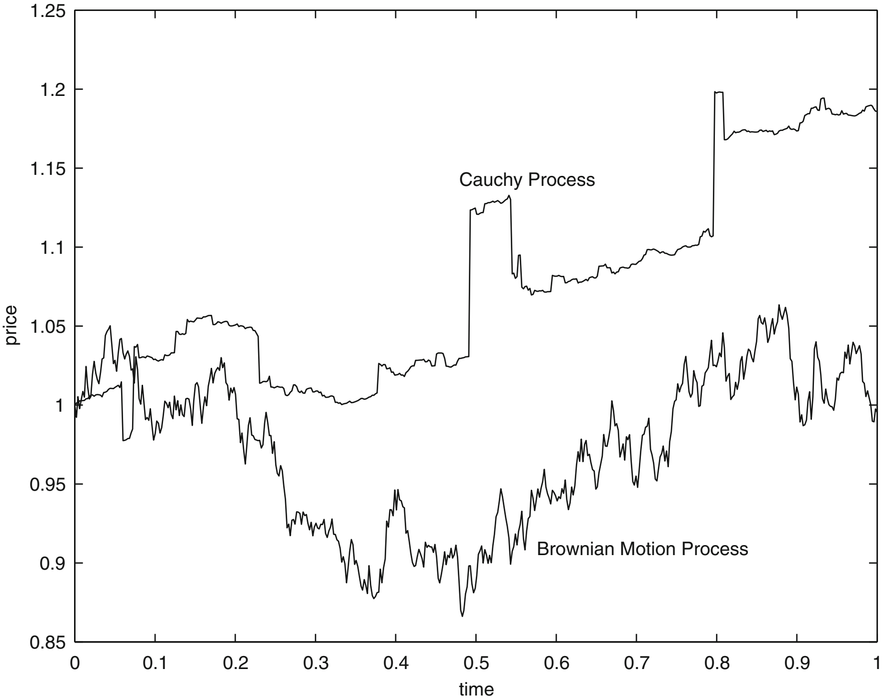


**Fig. 2.9** Typical sample paths for the Brownian motion and Cauchy process

display a typical sample path of both the Brownian motion process and the Cauchy process.

## 2.5 The Dirac Delta Function

The two examples of the proceeding section can also serve as a vehicle for another important concept that we shall need when we come to discuss the solution of option pricing equations, namely the concept of the *Dirac delta function*.

Note first of all that integrating the conditional probability density functions (2.10) and (2.11) with respect to $x$ reveals that the area under the distribution curves is equal to 1, irrespective of the value of $\Delta t$. Of course this is a fundamental requirement of conditional probability density functions. As $\Delta t \rightarrow 0$ the distribution in Fig. 2.8 becomes more and more peaked. Close to the limit $\Delta t = 0$

and $\xi_i \sim N(0,1)$, with $i = 1, 2, \cdots, 500$. For the Cauchy process, the path was calculated using $y_{i+1} = y_i + y_i \mu \Delta t + \mu_i \sigma \Delta t \cot[\pi \xi_i]$.

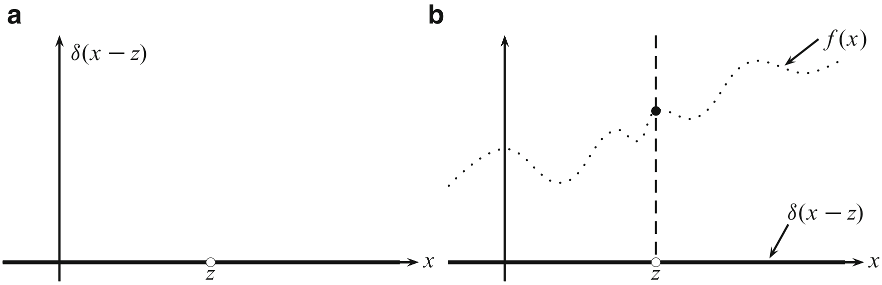


**Fig. 2.10** The sense of the definition of the Dirac delta function. (**a**) $\delta(x-z)=0, \quad x \neq z$. (**b**) $\int_{-\infty}^{\infty} \delta(x-z) f(x) dx = f(z)$

we would see a function which is almost zero everywhere except at $x = z$, and at this point its value becomes very large; all in such a way that the area beneath the distribution curve remains 1 for all $\Delta t$.

In order to formally (as opposed to mathematically rigorously) carry out mathematical operations involving the function obtained from this limiting process, the so-called *Dirac delta function* has been introduced (see e.g. Lighthill 1980). This function is usually denoted by the symbol $\delta$ and is formally defined by:

$$\text{(a)}\ \ \delta(x-z)=0, \quad x \neq z, \qquad \text{(b)}\ \ \int_{-\infty}^{\infty} \delta(x-z) f(x) dx = f(z). \tag{2.12}$$

The sense of both parts of this definition are illustrated in Fig. 2.10.

Both the Brownian motion distribution and the Cauchy distribution in the limit $\Delta t \to 0$ satisfy the formal definition of the Dirac delta function, i.e.

$$\delta(x-z) = \lim_{\Delta t \to 0} p(x, t+\Delta t \mid z, t), \tag{2.13}$$

as illustrated in Fig. 2.8.

To see the economic sense of the second condition in Eq. (2.12), suppose that $f(x)$ is some payoff on the asset at time $t + \Delta t$ if the asset price is $x$ at that point in time. Then

$$\int_{-\infty}^{\infty} p(x, t+\Delta t \mid z, t) f(x)\, dx,$$

is the expected payoff, calculated at time t when the asset price is $z$. The condition (b) of (2.12) states that as $\Delta t \to 0$, the expected payoff becomes the payoff that would be obtained at the current price $z$.

## 2.6 The Fokker–Planck and Kolmogorov Equations

The Chapman–Kolmogorov equation (2.8) tells us how the conditional probabilities are evolving over time. Up to this point in the discussion we have imposed very little structure on the stochastic process apart from the Markov assumption. We would like to reduce the Chapman–Kolmogorov equation to something which is more mathematically tractable and at the same time which involves parameters whose values we could measure from statistical observations on past and current prices.

The above aim could be achieved in a number of ways, but the simplest would be to restrict our attention to those stochastic processes having continuous sample paths, such as those discussed in Sect. 2.4 but with the possibility of sudden large jumps from time to time. In adopting this approach we shall give some mathematical precision to the intuitive notion, which we mentioned in Sect. 2.1, of the asset return consisting of a certain mean component about which there is a stochastic or volatile component.

In particular we shall restrict our attention to stochastic processes whose conditional probability density function satisfies the following three conditions. For all $\varepsilon > 0$,

$$\text{(i)} \quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} (x-z) p(x, t+\Delta t \mid z, t) dx = A(z, t), \tag{2.14}$$

$$\text{(ii)} \quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} (x-z)^2 p(x, t+\Delta t \mid z, t)\, dx = B(z, t), \tag{2.15}$$

$$\text{(iii)} \quad \lim_{\Delta t \to 0} \frac{p(x, t+\Delta t \mid z, t)}{\Delta t} = J(x \mid z, t), \tag{2.16}$$

where in (iii) the convergence is uniform in $x$, $z$ and $t$, for $\mid x - z \mid \geq \varepsilon$.

What are these conditions saying? Recall first of all that typically $\varepsilon$ is small, so that $\mid x - z \mid < \varepsilon$ refers to the set of prices $x$ at $t + \Delta t$ which have not moved very far from the current price $z$ at $t$, see Fig. 2.11, while $\mid x - z \mid \geq \varepsilon$ refers to the set of prices $x$ at $t + \Delta t$ which have moved more than a small amount from current price $z$ at $t$.

The integral in condition (i) is the mean of the small (i.e. within $\varepsilon$ of $z$) price changes over the time interval $(t, t + \Delta t)$. The condition thus states that

$$\text{the mean of small price changes over } (t, t+\Delta t) \cong A(z, t)\Delta t.$$

The choice of $A(z, t)$ may be imposed by the financial analyst but would usually be obtained from empirical analysis of past behaviour of asset prices. For example for common stock prices, the form $A(z, t) = \mu z$ where $\mu$ is the mean stock return per unit time fits fairly well with observed price behaviour. On the other hand for short term interest rates empirical evidence suggests the form $A(z, t) = \kappa(\bar{z} - z)$ where $\bar{z}$ is a long run average short term rate and $\kappa$ is a speed of adjustment constant, both of which could be determined from observed interest rate behaviour.

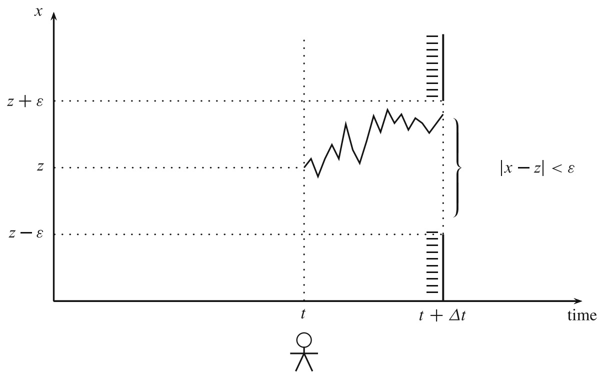


**Fig. 2.11** The price range in conditions (i) and (ii)

The integral in condition (ii) is the second moment of small price changes over $(t, t+\Delta t)$ and thus states that

$$\text{the second moment of small price changes over } (t, t+\Delta t) \cong B(z,t)\Delta t.$$

As with $A(z,t)$ the form of $B(z,t)$ would be determined from the time series behaviour of the asset of interest. Continuing the examples cited above, for common stock prices empirical studies suggest $B(z,t) = \sigma^2 z^2$ ($\sigma$ a constant), whilst for short term interest rates $B(z,t) = \sigma^2 z^{2\gamma}$ ($\sigma, \gamma$ constant, $0 < \gamma < 2$) seems to be the consensus of a range of empirical studies.

The third condition concentrates on price changes which are not small (i.e. are more than a distance $\varepsilon$ from the current price $z$). Condition (iii) states that

$$\text{the probability of large changes over } (t, t+\Delta t) \cong J(x \mid z,t)\Delta t.$$

In essence the quantity $J(x \mid z,t)$ captures the probability that the price will jump from $z$ to $x$ at time $t$. As with the $A$ and $B$ functions, the $J$ function would be obtained from observations on past price movements.

If we assume, or confirm by observation, that $J = 0$, then the asset price will only exhibit small price changes over small time intervals. The stochastic process is then known as a *diffusion process*. We know from the discussion in Sect. 2.4 that the sample paths of such processes are continuous.

If we allow $J \neq 0$ then we are admitting the possibility of sudden large jumps in the asset price. The frequency and magnitude of these jumps determine the functional form of $J$. For instance, in order to model the large jumps in the prices of some of the assets in which we will be interested (e.g. foreign exchange rates), there is empirical evidence that the Poisson process is an appropriate form for $J$.

The $A(z,t)$ term is referred to as *the drift term* of the stochastic process, the $B(z,t)$ term is referred to as *the diffusion term* of the stochastic process and the $J(x \mid z,t)$ term is referred to as *the jump term* of the stochastic process. A stochastic process in which all of these terms are present is known as a *jump-diffusion process*.

The sample paths of such a stochastic process are not continuous in general (unless of course the jump component is zero). Figure 2.9 illustrates the difference between the sample paths of a diffusion process (the Brownian motion process) and a jump-diffusion process (the Cauchy process). For most applications in option pricing we will be concerned with pure diffusion processes. However in some of our applications we will need to allow for jump terms.

Before proceeding to discuss the Fokker–Planck equation we should point out that conditions (i) and (ii) imply that all the higher order moments of the conditional probability density function of the stochastic process vanish, i.e.

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} (x-z)^k p(x, t + \Delta t \mid z, t) dx = 0$$

*for all* $k \geq 3$. See Appendix 2.4 for an outline of how this result may be proved. *Thus a pure diffusion process is completely specified by the drift and diffusion terms* $A(z,t)$ *and* $B(z,t)$.

Let us assume for the moment that the jump term is zero (i.e. $J = 0$) so that asset prices are following a pure diffusion process. It can be shown that conditions (i) and (ii) in (2.14) and (2.15) reduce the Chapman–Kolmogorov equation for the evolution of conditional probabilities to the partial differential equation (see Appendix 2.5 for details)

$$\frac{\partial}{\partial t} p(z,t \mid y,\tau) = -\frac{\partial}{\partial z}[A(z,t)p(z,t \mid y,\tau)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[B(z,t)p(z,t \mid y,\tau)], \tag{2.17}$$

for the conditional probability $p(z,t \mid y,\tau)$ of observing asset price $z$ at time $t$ given that the current price is $y$ at time $\tau$. This equation is known as the *Fokker–Planck equation* and must be solved for $t \geq \tau$ subject to the initial time condition

$$p(z,\tau \mid y,\tau) = \delta(y-z). \tag{2.18}$$

It is also necessary to impose some *boundary* conditions (e.g. in the case of interest rates, the probability of these becoming negative must be zero) but we will discuss these as they arise in particular applications. It is important to emphasise that the viewpoint adopted with the Fokker–Planck equation is a *forward* one, i.e., we take the current price $y$ and time $\tau$ as fixed and consider the conditional probability as a forward evolving function of the price $z$ at the later time $t$.

However in order to price options we need to adopt the alternative perspective in which the final time is fixed (i.e. the maturity date of the option) and the initial time is varying. In other words we hold $z$ and $t$ fixed and allow $y$ and $\tau$ to vary. In this case it can be shown that under conditions (i) and (ii) the Chapman–Kolmogorov equation becomes (see Appendix 2.5 for details)

$$\frac{\partial}{\partial \tau} p(z,t \mid y,\tau) = -A(y,\tau)\frac{\partial}{\partial y} p(z,t \mid y,\tau) - \frac{1}{2}B(y,\tau)\frac{\partial^2}{\partial y^2} p(z,t \mid y,\tau), \tag{2.19}$$

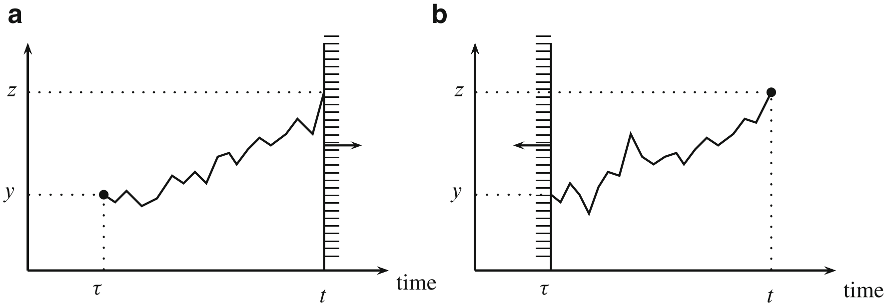


**Fig. 2.12** (**a**) Fokker–Planck equation for forward evolving probability; $(y, \tau)$ fixed. (**b**) Kolmogorov equation for backward evolving probability; $(z, t)$ fixed

which is known as the *Kolmogorov backward equation*, and must also be solved for $\tau \leq t$ subject to the final time condition

$$p(z, t \mid y, t) = \delta(y - z). \tag{2.20}$$

Now $p(z, t \mid y, \tau)$ is a backward evolving function of the price, $y$, at the earlier time $\tau$.

The different viewpoints of the partial differential equations (2.17) and (2.19), i.e. probability evolving forwards or backwards are illustrated in Fig. 2.12. In this figure, the reader should view the hashed wall as moving forward in Fig. 2.12a and backward in Fig. 2.12b.

Since we shall be referring frequently to both the Fokker–Planck equation and the Kolmogorov equation it is useful to introduce some short-hand notation for writing them down. Thus in relation to diffusion processes with drift function $A$ and diffusion function $B$ we introduce the partial differential operators $\mathscr{F}$ and $\mathscr{K}$, defined by

$$\mathscr{F}p = -\frac{\partial}{\partial z}[A(z, t)p(z, t \mid y, \tau)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[B(z, t)p(z, t \mid y, \tau)], \tag{2.21}$$

and

$$\mathscr{K}p = A(y, \tau)\frac{\partial}{\partial y}p(z, t \mid y, \tau) + \frac{1}{2}B(y, \tau)\frac{\partial^2}{\partial y^2}p(z, t \mid y, \tau). \tag{2.22}$$

The Fokker–Planck equation may then be written succinctly as

$$\frac{\partial p}{\partial t} - \mathscr{F}p = 0, \tag{2.23}$$

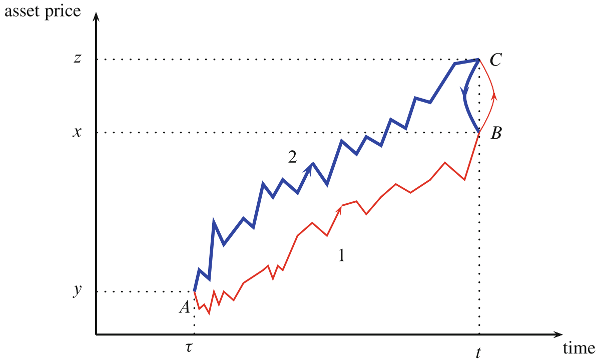


**Fig. 2.13** Explaining the jump term in the Fokker–Planck equation

whilst the Kolmogorov equation as

$$\frac{\partial p}{\partial \tau} + \mathscr{K} p = 0. \tag{2.24}$$

Of course when using this more succinct notation we must be a little cautious in keeping track of whether it is the final point (i.e. $z, t$ in the Fokker–Planck equation) or the initial point (i.e. $y, \tau$ in the Kolmogorov equation) which is varying.

The forward Fokker–Planck equation and the backward Kolmogorov equation are equivalent to each other in the sense that they yield the same conditional probability. They differ only as to whether the initial point or the final point is held fixed. Which form we use depends on the application at hand. In the technical language of the theory of partial differential equations the partial differential operators $\mathscr{F}$ and $\mathscr{K}$ are said to be adjoint operators.

To complete the discussion we show how the forward and backward equations need to be modified in order to allow for a jump component, i.e. when $J \neq 0$. The technical details are also included in Appendix 2.5 where it is shown that the Fokker–Planck equation becomes

$$\frac{\partial p}{\partial t} - \mathscr{F} p = \int_{-\infty}^{\infty} [J(z \mid x, t) p(x, t \mid y, \tau) - J(x \mid z, t) p(z, t \mid y, \tau)] dx. \tag{2.25}$$

Recalling that $J(z \mid x, t)$ essentially measures the probability that the price will jump from $x$ to $z$ at time $t$, the sense of the integral term on the right-hand side of (2.25) can be understood by referring to Fig. 2.13. The jump term, $J$, allows two additional types of events which are not possible under the pure diffusion process described by the operator $\mathscr{F}$. Firstly the price may follow a path to the value $x$ at $t$, where $x$ is not "close" to $z$, and then jump to the value $z$. The product $J(z \mid x, t) p(x, t \mid y, \tau)$ is the probability of going from $A$ to $B$, then $B$ to $C$ along such a path, and the integral of this probability with respect to $x$ measures the probability

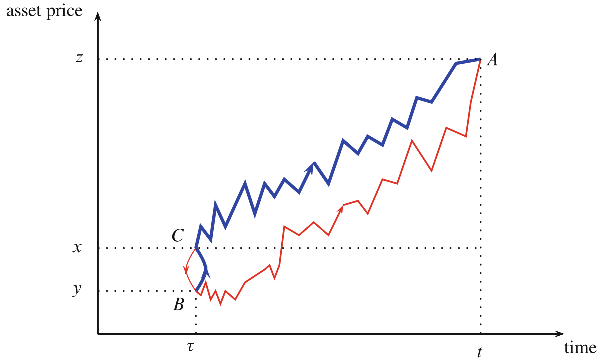


**Fig. 2.14** Explaining the jump term in the Kolmogorov equation

of going from $A$ to $C$ along all such paths. Secondly the price may follow a path to the value $z$ at $t$, and then jump to a value $x$ not "close" to $z$, this would be the path 2 in Fig. 2.13. The product $J(x \mid z, t)p(z, t \mid y, \tau)$ is the probability of going from $A$ to $C$, then $C$ to $B$ along such a path and the integral of this probability with respect to $x$ measures the probability over all such paths of the price reaching $z$ at time $t$ and then immediately jumping away to some other price. The overall term on the right-hand side of (2.25) is the probability of the price being at $z$ at time $t$ once the probabilities of price jumps at $t$ (both to and away from $z$) are fully accounted for.

Inclusion of the jump term in the Kolmogorov equation leads to

$$\frac{\partial p}{\partial \tau} + \mathscr{K} p = \int_{-\infty}^{\infty} J(x \mid y, \tau)[p(z, t \mid y, \tau) - p(z, t \mid x, \tau)]dx, \tag{2.26}$$

as illustrated in Fig. 2.14.

In Chap. 13 we shall consider option pricing when the underlying asset follows a jump-diffusion process. We shall see that the option price is determined by an integro-partial differential equation of the form (2.26). The techniques of the theory of stochastic differential equations that we develop in Chaps. 4, 6 and 8 allow us to conveniently arrive at a specification of the $J(z \mid x, t)$ term.

## 2.7 Appendix

### Appendix 2.1 Probability Density Functions

A function $P(x)$ is a cumulative distribution function if $P$ is a non-decreasing function and satisfies the properties

$$P(x_{\min}) = 0, \quad P(x_{\max}) = 1,$$

where $x_{\min}$ and $x_{\max}$ are respectively the minimum and maximum values attainable by $x$. If $X$ denotes the random variable of interest (e.g. a stock price) then

$$P(x) = \text{Prob.}\{X \le x\}.$$

Typically we will deal with cumulative distribution functions which are differentiable and we write

$$P'(x) = p(x)$$

so that the cumulative distribution function can be written

$$P(x) = \int_{x_{\min}}^{x} p(\xi)d\xi.$$

With this notation we can readily see that

$$p(x)dx = dP(x) = \text{Prob.}\left\{x < X \le x + dx\right\}.$$

The function $p(x)$ is known as the probability density function and since $P(x_{\max}) = 1$ it satisfies the property

$$\int_{x_{\min}}^{x_{\max}} p(\xi)d\xi = 1. \tag{2.27}$$

Sometimes it is convenient to transform to a new variable $y$ related to $x$ by

$$x = g(y),$$

where $g$ is increasing and differentiable. Making the change of variable $\xi = g(\zeta)$ Eq. (2.27) becomes

$$\int_{y_{\min}}^{y_{\max}} p(g(\zeta))g'(\zeta)d\zeta = 1,$$

where

$$y_{\max} = g^{-1}(x_{\max}), \quad y_{\min} = g^{-1}(x_{\min}).$$

If we set $\pi(y) = p(g(y))$ then the last equation becomes

$$\int_{y_{\min}}^{y_{\max}} \pi(\zeta))g'(\zeta)d\zeta = 1,$$

from which we see that in the new co-ordinates the density function is $\pi(y)g'(y)$. The corresponding c.d.f. is given by

$$\Pi(y) = \int_{y_{\min}}^{y} \pi(\zeta)g'(\zeta)d\zeta.$$

Thus the rule for transforming from $x$ to $y$ is

$$p(x) = \pi(y)g'(y)$$

which in terms of $x$ may be written

$$p(x) = \pi(g^{-1}(x)) \cdot g'(g^{-1}(x)). \tag{2.28}$$

## Appendix 2.2 Brownian Motion is a Continuous Process

We seek to verify that Brownian motion (2.10) satisfies (2.9), that is, Brownian motion is continuous. We need to use the following facts:

(i) In general, for $\alpha > 0$,

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}.$$

(ii) For $\delta$ large enough,

$$\int_{|x|<\delta} e^{-x^2} dx \approx \sqrt{\pi}\sqrt{1-e^{-\delta^2}}.$$

For the conditional probability density function of the Brownian motion

$$p(x, t + \Delta t|z, t) = \frac{1}{\sqrt{2\pi\Delta t}\sigma} e^{-\frac{(x-z)^2}{2\sigma^2\Delta t}},$$

we have

$$I := \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} p(x, t + \Delta t|z, t)dx = \frac{1}{\Delta t} \frac{1}{\sqrt{2\pi\Delta t}\sigma} \int_{|x-z|>\varepsilon} e^{-\frac{(x-z)^2}{2\sigma^2\Delta t}} dx.$$

Let

$$y = \frac{x-z}{\sqrt{2\Delta t}\sigma}$$

then

$$
\begin{aligned}
I &= \frac{1}{\Delta t}\frac{1}{\sqrt{\pi}}\int_{|y|>\frac{\varepsilon}{\sqrt{2\Delta t}\sigma}} e^{-y^2}dy = \frac{1}{\Delta t}\frac{1}{\sqrt{\pi}}\left[\int_{|y|<\infty} e^{-y^2}dy - \int_{|y|<\frac{\varepsilon}{\sqrt{2\Delta t}\sigma}} e^{-y^2}dy\right] \\
&\approx \frac{1}{\Delta t}\frac{1}{\sqrt{\pi}}\left[\sqrt{\pi} - \sqrt{\pi}\sqrt{1-e^{-\frac{\varepsilon^2}{2\Delta t\sigma^2}}}\right] = \frac{1}{\Delta t}\left[1-\sqrt{1-e^{-\frac{\varepsilon^2}{2\Delta t\sigma^2}}}\right].
\end{aligned}
$$

Let $\alpha = \frac{1}{\Delta t}$ and $A = \frac{\varepsilon^2}{2\sigma^2}$, then

$$
\begin{aligned}
\lim_{\Delta t\to 0} I &= \lim_{\alpha\to\infty} \alpha\left[1-\sqrt{1-e^{-A\alpha}}\right] = \lim_{\alpha\to\infty}\frac{1-\sqrt{1-e^{-A\alpha}}}{1/\alpha} \\
&= \lim_{\alpha\to\infty}\frac{\frac{-Ae^{-A\alpha}}{2\sqrt{1-e^{-A\alpha}}}}{-1/\alpha^2} \quad \text{(applying l'Hôpital's Rule)} \\
&= \lim_{\alpha\to\infty}\frac{A}{2\sqrt{1-e^{-A\alpha}}}\frac{\alpha^2}{e^{A\alpha}} = 0.
\end{aligned}
$$

Hence (2.10) satisfies (2.9).

## Appendix 2.3 The Cauchy Process is Not Continuous

We seek to verify that (2.11) does not satisfy (2.9). For the Cauchy process,

$$
p(x, t+\Delta t|z, t) = \frac{\Delta t}{\pi((x-z)^2 + (\Delta t)^2)},
$$

consider

$$
I = \frac{1}{\Delta t}\int_{|x-z|>\varepsilon} p(x, t+\Delta t|z, t)dx = \int_{|x-z|>\varepsilon}\frac{1}{\pi((x-z)^2+(\Delta t)^2)}dx.
$$

Let $y = x - z$, then $dy = dx$ and

$$
I = \int_{|y|>\varepsilon}\frac{1}{\pi(y^2+(\Delta t)^2)}dy.
$$

Let $y = \Delta t\tan\theta$ so that $\theta = \tan^{-1}(\frac{y}{\Delta t})$ and $dy = \Delta t\sec^2\theta d\theta$. Hence

$$
I = \int_{|y|>\varepsilon}\frac{\Delta t\sec^2\theta}{\pi(\Delta t)^2\sec^2\theta}d\theta
$$

$$= \frac{1}{\pi \Delta t}\left[\tan^{-1}\left(\frac{y}{\Delta t}\right)\Big|_{y=-\infty}^{\infty} - \tan^{-1}\left(\frac{y}{\Delta t}\right)\Big|_{y=-\varepsilon}^{\varepsilon}\right]$$
$$= \frac{1}{\pi \Delta t}\left[\pi - 2\tan^{-1}\left(\frac{\varepsilon}{\Delta t}\right)\right].$$

Consider

$$\lim_{\Delta t \to 0} I = \lim_{\Delta t \to 0} \frac{1}{\pi} \frac{\pi - 2\tan^{-1}\left(\frac{\varepsilon}{\Delta t}\right)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{1}{\pi}\left(-2\frac{1}{\sec^2\left(\frac{\varepsilon}{\Delta t}\right)}\right)\left(-\frac{\varepsilon}{(\Delta t)^2}\right) = \infty.$$

Here we use

$$\sec^2\left(\frac{\varepsilon}{\Delta t}\right) = 1 + \tan^2\left(\frac{\varepsilon}{\Delta t}\right) \to 1 + \left(\frac{\pi}{2}\right)^2$$

as $\Delta t \to 0$. Hence (2.11) does not satisfy (2.9).

## Appendix 2.4 The Higher Order Moment Condition

Consider for example the third order moment (i.e. $k = 3$). Note that

$$\left|\int_{|x-z|<\varepsilon} (x-z)^3 p(x, t+\Delta t \mid z, t)\, dx\right| \le \int_{|x-z|<\varepsilon} |x-z|^3 \cdot p(x, t+\Delta t \mid z, t)\, dx$$
$$\le \int_{|x-z|<\varepsilon} |x-z| \cdot |x-z|^2 \cdot p(x, t+\Delta t \mid z, t)\, dx$$
$$\le \varepsilon \int_{|x-z|<\varepsilon} |x-z|^2 \cdot p(x, t+\Delta t \mid z, t)\, dx.$$

Hence

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\left|\int_{|x-z|<\varepsilon} (x-z)^3 p(x, t+\Delta t \mid z, t)\, dx\right|$$
$$\le \varepsilon \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} |x-z|^2 \cdot p(x, t+\Delta t \mid z, t) dx$$
$$\le \varepsilon B(z, t).$$

Since we may choose $\varepsilon$ as small as we like, this last term tends to zero. A similar argument applies for $k > 3$.

## Appendix 2.5 Derivation of Fokker–Plank and Kolmogorov Equations

The derivation is based on Gardiner (1985, Sect. 3.4). Consider the Chapman–Kolmogorov equation in the form (see Fig. 2.15)

$$p(z, t + \Delta t \mid y, \tau) = \int p(z, t + \Delta t \mid x, t) p(x, t \mid y, \tau) dx, \tag{2.29}$$

from which we subtract $p(z, t \mid y, \tau)$. Hence

$$\begin{aligned} & p(z, t + \Delta t \mid y, \tau) - p(z, t \mid y, \tau) \\ &\quad = \int p(z, t + \Delta t \mid x, t) p(x, t \mid y, \tau) dx - p(z, t \mid y, \tau). \end{aligned} \tag{2.30}$$

Now introduce an arbitrary function $R(x)$, which together with all of its derivatives, vanishes as $x$ approaches the extremities of the range of interest (e.g. 0 and $\infty$ for stock prices). Then multiplying both sides of (2.30) by $R(z)/\Delta t$ and integrating with respect to $z$,

$$\begin{aligned} & \int R(z) \frac{[p(z, t + \Delta t \mid y, \tau) - p(z, t \mid y, \tau)]}{\Delta t} dz \\ &\qquad = \frac{1}{\Delta t} \int R(z) \left( \int p(z, t + \Delta t \mid x, t) p(x, t \mid y, \tau)\, dx \right) dz \\ &\qquad\quad - \frac{1}{\Delta t} \int R(z) p(z, t \mid y, \tau)\, dz. \end{aligned} \tag{2.31}$$

The function $R$ is next expanded in a Taylors series about the point $x$, i.e.

$$R(z) = R(x) + \frac{dR(x)}{dx}(z - x) + \frac{1}{2}\frac{d^2 R(x)}{dx^2}(z - x)^2 + o(z - x)^3, \tag{2.32}$$

and the integrals are considered over regions $\mid z - x \mid < \varepsilon$ and $\mid z - x \mid \geq \varepsilon$.

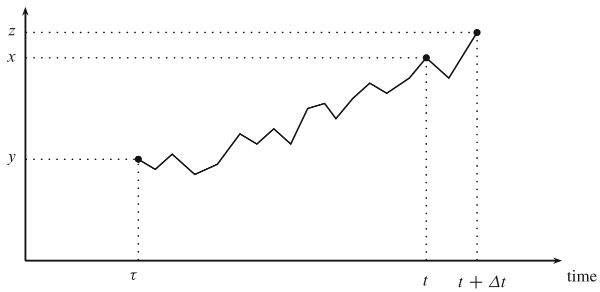


**Fig. 2.15** Evolution from $\tau$ to $t + \Delta t$ via $t$

If we assume a pure diffusion process, i.e. $J = 0$, then the Lindeberg condition allows us to ignore integrals in the region $| z - x | \geq \varepsilon$. The result in Appendix 2.4 allows us to ignore the $o(z-x)^3$ term in integrals over the region $| z - x | < \varepsilon$. We are left with integrals involving terms up to $(z-x)^2$ over the region $| z - x | < \varepsilon$. In the limit $\Delta t \to 0$ the term on the left hand side of (2.31) becomes

$$\int R(z) \frac{\partial p}{\partial t} dz.$$

On the right hand side, after substituting (2.32) and interchanging the order of the integrations, we are left with

$$\begin{aligned} &\frac{1}{\Delta t} \int R(x) p(x,t \mid y,\tau) \left( \int p(z, t+\Delta t \mid x,t) dz \right) dx \\ &+ \sum_{n=1}^{2} \frac{1}{n!} \int \frac{d^n R(x)}{dx^n} p(x,t \mid y,\tau) \left( \frac{1}{\Delta t} \int (z-x)^n p(z,t+\Delta t \mid x,t) dz \right) dx \\ &- \frac{1}{\Delta t} \int R(z) p(z,t \mid y,\tau) dz, \end{aligned} \tag{2.33}$$

where all the integrals are taken over the region $| z - x | < \varepsilon$. Using conditions (i) and (ii) in Eqs. (2.14) and (2.15) and the result in Appendix 2.4 we see that in the limit $\Delta t \to 0$ the terms in the big bracket on the middle line tend to $A(x,t)$ (when $n = 1$), $B(x,t)$ (when $n = 2$) and 0 (when $n \geq 3$). Using the fact that

$$\int p(z, t+\Delta t \mid x,t) dz = 1,$$

Eq. (2.31) now reduces to

$$\int \left[ R(x) \frac{\partial p}{\partial t} - \frac{dR}{dx} A(x,t) p(x,t \mid y,\tau) - \frac{1}{2} \frac{d^2 R}{dx^2} B(x,t) p(x,t \mid y,\tau) \right] dx = 0. \tag{2.34}$$

The final step is to perform an integration by parts on the last two terms so that

$$\int \frac{dR}{dx} A(x,t) p(x,t \mid y,\tau)\, dx = - \int \frac{\partial}{\partial x} (A(x,t) p(x,t \mid y,\tau)) R(x)\, dx$$

and

$$\int \frac{d^2 R}{dx^2} B(x,t) p(x,t \mid y,\tau)\, dx = \int \frac{\partial^2}{\partial x^2} (B(x,t) p(x,t \mid y,\tau)) R(x)\, dx.$$

Note that in performing these integrations by parts we have invoked the properties that R and all of its derivatives vanish at the extremities of the range of interest. So Eq. (2.34) finally becomes

$$\int \left[ \frac{\partial p}{\partial t} + \frac{\partial}{\partial x}(A(x,t)p(x,t \mid y,\tau)) - \frac{1}{2}\frac{\partial^2}{\partial x^2}(B(x,t)p(x,t \mid y,\tau)) \right] R(x)dx = 0.$$

Since this last expression holds for an arbitrary function $R(x)$, the term in the squared bracket must be zero, i.e.

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(A(x,t)p(x,t \mid y,\tau)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(B(x,t)p(x,t \mid y,\tau)),$$

which is the Fokker–Plank forward equation.

If we are dealing with a jump-diffusion process then $J \neq 0$ and we cannot invoke the Lindeberg condition to ignore integrals in the region $\mid x - z \mid \geq \varepsilon$. In this case the expansion at Eq. (2.31) will involve the extra terms

$$\lim_{\varepsilon \to 0} \frac{1}{\Delta t} \left( \int \int_{|x-z| \geq \varepsilon} R(z)p(z,t+\Delta t \mid x,t)p(x,t \mid y,\tau)dzdx \right.$$
$$\left. - \int \int_{|x-z| \geq \varepsilon} R(x)p(z,t+\Delta t \mid x,t)p(x,t \mid y,\tau)dxdz \right).$$

Recalling the definition of $J(x \mid z,t)$ from Eq. (2.16) we can write these terms when $\Delta t \to 0$ as

$$\int \int_{|x-z| \geq \varepsilon} R(z)J(z \mid x,t)p(x,t \mid y,\tau)dzdx$$
$$- \int \int_{|x-z| \geq \varepsilon} R(x)J(z \mid x,t)p(x,t \mid y,\tau)dxdz.$$

Interchanging the roles of $x$ and $z$ in the first integral the two terms can be combined to yield

$$\int \int_{|x-z| \geq \varepsilon} R(x) \left[ J(x \mid z,t)p(z,t \mid y,\tau) - J(z \mid x,t)p(x,t \mid y,\tau) \right] dzdx.$$

Incorporating this term into Eq. (2.34) and proceeding to the limit $\varepsilon \to 0$ we obtain Eq. (2.25) of the main text.

To obtain the Kolmogorov backward equation we consider

$$p(z,t|y,\tau) = \int p(z,t|x,\tau+\Delta t)p(x,\tau+\Delta t|y,\tau)dx, \tag{2.35}$$

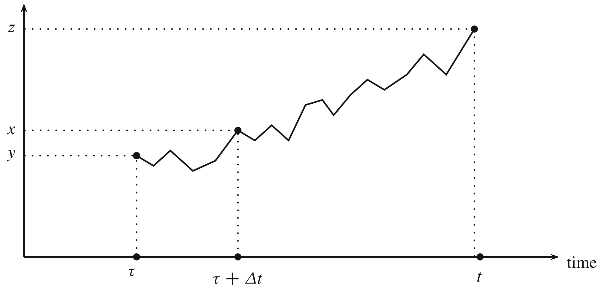


**Fig. 2.16** Evolution from $\tau$ to $t$ via $\tau + \Delta t$

the sense of which is illustrated in Fig. 2.16.

Note that

$$\frac{\partial}{\partial \tau} p(z,t|y,\tau) = \lim_{\Delta t \to 0} \frac{1}{\Delta t}[p(z,t|y,\tau + \Delta t) - p(z,t|y,\tau)]$$

Consider

$$\begin{aligned} I &= \frac{1}{\Delta t}[p(z,t|y,\tau + \Delta t) - p(z,t|y,\tau)] \\ &= \frac{1}{\Delta t}\left[p(z,t|y,\tau + \Delta t) - \int p(z,t|x,\tau + \Delta t)p(x,\tau + \Delta t|y,\tau)dx\right] \\ &= \frac{1}{\Delta t}\int \left[p(z,t|y,\tau + \Delta t) - p(z,t|x,\tau + \Delta t)\right]p(x,\tau + \Delta t|y,\tau)dx \\ &= I_1 + I_2, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \frac{1}{\Delta t}\int_{|x-y|<\varepsilon} \left[p(z,t|y,\tau + \Delta t) - p(z,t|x,\tau + \Delta t)\right]p(x,\tau + \Delta t|y,\tau)dx, \\ I_2 &= \frac{1}{\Delta t}\int_{|x-y|>\varepsilon} \left[p(z,t|y,\tau + \Delta t) - p(z,t|x,\tau + \Delta t)\right]p(x,\tau + \Delta t|y,\tau)dx. \end{aligned}$$

Assume first that there is no jump, then $I_2 = 0$ as $\Delta t \to 0$ and

$$\begin{aligned} I_1 &= \frac{1}{\Delta t}\int_{|x-y|<\varepsilon} \left[p(z,t|y,\tau + \Delta t) - p(z,t|x,\tau + \Delta t)\right]p(x,\tau + \Delta t|y,\tau)dx \\ &\approx -\frac{1}{\Delta t}\int_{|x-y|<\varepsilon} \left[\frac{\partial p(z,t|y,\tau+\Delta t)}{\partial y}(x-y) + \frac{1}{2}\frac{\partial^2 p}{\partial y^2}(x-y)^2\right] p(x,\tau+\Delta t|y,\tau)dx \end{aligned}$$

$$= -\frac{\partial p(z,t|y,\tau+\Delta t)}{\partial y}\frac{1}{\Delta t}\int_{|x-y|<\varepsilon}(x-y)p(x,\tau+\Delta t|y,\tau)dx$$
$$-\frac{1}{2}\frac{\partial^2 p(z,t|y,\tau+\Delta t)}{\partial y^2}\frac{1}{\Delta t}\int_{|x-y|<\varepsilon}(x-y)^2 p(x,\tau+\Delta t|y,\tau)dx.$$

Then

$$\lim_{\Delta t\to 0} I_1 = -A(y,\tau)\frac{\partial p(z,t|y,\tau)}{\partial y} - \frac{1}{2}B(y,\tau)\frac{\partial^2 p}{\partial y^2}.$$

Hence we have established that

$$\frac{\partial p}{\partial \tau}(z,t|y,\tau) = -A(y,\tau)\frac{\partial p(z,t|y,\tau)}{\partial y} - \frac{1}{2}B(y,\tau)\frac{\partial^2 p(z,t|y,\tau)}{\partial y^2}.$$

In the presence of jump, the only difference is that

$$\lim_{\Delta t\to 0} I_2 = \int_{-\infty}^{\infty} J(x|y,\tau)[p(z,t|y,\tau)-p(z,t|x,\tau)]dx.$$

Hence $p(z,t|y,\tau)$ satisfies

$$\frac{\partial}{\partial \tau}p(z,t|y,\tau) = -A(y,\tau)\frac{\partial p(z,t|y,\tau)}{\partial y} - \frac{1}{2}B(y,\tau)\frac{\partial^2 p(z,t|y,\tau)}{\partial y^2}$$
$$+\int_{-\infty}^{\infty} J(x|y,\tau)[p(z,t|y,\tau)-p(z,t|x,\tau)]dx,$$

which is the result given in (2.26).

## Appendix 2.6 The Mean Value Theorem

Suppose that $f(x)$ is a continuous and non-negative function on the interval $[a,b]$. There exists a value $\xi$ satisfying $a \le \xi \le b$ such that

$$\int_a^b f(x)dx = (b-a)f(\xi).$$

Basically this result says that we can always find $\xi$ such that the area under the rectangle shown in the figure equals the area under the curve (Fig. 2.17).

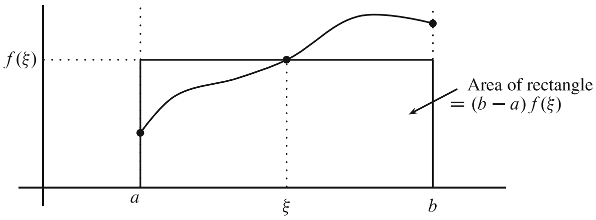


**Fig. 2.17** Illustrating the mean value theorem for integrals

## 2.8 Problems

### Problem 2.1

(a) Consider the Brownian notion whose transition probability density is given by

$$p(x,t|y,\tau) = \frac{1}{\sqrt{2\pi(t-\tau)}\sigma} \exp\left[\frac{-(x-y)^2}{2\sigma^2(t-\tau)}\right].$$

By direct differentiation show that $p$ satisfies the partial differential equations

$$\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial y^2} + \frac{\partial p}{\partial \tau} = 0,$$

and

$$-\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2} + \frac{\partial p}{\partial t} = 0.$$

What is the boundary condition for both of these partial differential equations?

(b) Consider the Cauchy distribution whose transitional partial differential function is given by

$$p(x,t|y,\tau) = \frac{(t-\tau)}{\pi\left[(x-y)^2 + (t-\tau)^2\right]}.$$

Show that the first and second moments of this distribution are given by

$$\int_{-\infty}^{\infty} (x-y)p(x,t|y,\tau)dx = 0,$$

$$\int_{-\infty}^{\infty} (x-y)^2 p(x,t|y,\tau)dx = \infty.$$

Explain why these results indicate that it would not be possible to obtain Kolmogorov or Fokker–Planck equations as in (a).

**Problem 2.2** Consider the function $\delta_\varepsilon(x)$ defined by

$$\delta_\varepsilon(x) = \begin{cases} 0, |x| > \varepsilon/2, \\ 1/\varepsilon, |x| \le \varepsilon/2. \end{cases}$$

(a) Sketch this function.
(b) Show that

$$\int_{-\infty}^{\infty} \delta_\varepsilon(x)dx = 1$$

for all $\varepsilon$.

Consider a function $f$, defined on $\mathbb{R}$, which is continuous.

(c) Sketch the function $f(x)\delta_\varepsilon(x)$.
(d) Use the mean value theorem of integral calculus to show that

$$\int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(\theta),$$

where $-\varepsilon/2 \le \theta \le \varepsilon/2$.
(e) Explain the intuition of this result in the sketch you have just drawn. Hence show that

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(0).$$

(f) Explain how to use the foregoing arguments to establish the result

$$\lim_{\varepsilon\to 0} \delta_\varepsilon(x) = \delta(x),$$

where $\delta(x)$ is the Dirac delta function.

**Problem 2.3** Consider the transitional probability density function

$$p(x,t|y,\tau) = \frac{1}{\sqrt{2\pi(t-\tau)}\sigma}\exp\left[-\frac{(x-y-a(t-\tau))^2}{2\sigma^2(t-\tau)}\right].$$

Show by direct differentiation that $p$ satisfies the partial differential equations

$$\frac{1}{2}\sigma^2\frac{\partial^2 p}{\partial y^2} + a\frac{\partial p}{\partial y} + \frac{\partial p}{\partial \tau} = 0$$

and

$$-\frac{1}{2}\sigma^2\frac{\partial^2 p}{\partial x^2} + a\frac{\partial p}{\partial x} + \frac{\partial p}{\partial t} = 0.$$

**Problem 2.4** Consider the function $\delta_\varepsilon(x)$ defined as follows:

$$\delta_\varepsilon(x) = \begin{cases} 0, & x < -\varepsilon, \\ \frac{3}{4\varepsilon^3}(\varepsilon^2 - x^2), & -\varepsilon \le x \le \varepsilon, \\ 0, & x > \varepsilon. \end{cases}$$

(a) Sketch this function and show that

$$\int_{-\infty}^{\infty} \delta_\varepsilon(x)dx = 1$$

for all $\varepsilon$.

(b) Consider a function $f$, defined on $\mathbb{R}$, which is continuous. Use the Mean Value Theorem of integral calculus to show that

$$\int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(\theta),$$

where $-\varepsilon \le \theta \le \varepsilon$. Hence show that

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(0).$$

(c) Explain how to use the foregoing result to establish the result

$$\lim_{\varepsilon\to 0}\delta_\varepsilon(x) = \delta(x),$$

where $\delta(x)$ is the Dirac delta function.

http://www.springer.com/978-3-662-45905-8

Derivative Security Pricing
Techniques, Methods and Applications
Chiarella, C.; He, X.-Z.; Sklibosios Nikitopoulos, C.
2015, XVI, 616 p. 154 illus., 38 illus. in color., Hardcover
ISBN: 978-3-662-45905-8