9

Game-Theoretic Probability in Finance

In this introductory chapter, we sketch our game-theoretic approach to some basic topics of finance theory. We concentrate on how the Black-Scholes method for pricing and hedging a European option can be made purely game-theoretic, but begin with an examination of the apparently random behavior of stock-market prices, and we conclude with a discussion of informational efficiency.

The usual derivation of the Black-Scholes formula for the price of an option relies on the assumption that the market price $S(t)$ of the underlying security $\mathcal{S}$ follows a diffusion process. As we explain in the first two sections of this chapter, this stochastic assumption is used in two crucial ways:

Taming the Market The diffusion model limits the wildness of fluctuations in $S(t)$. This is the celebrated $\sqrt{dt}$ effect: the change in $S(t)$ over an increment of time of positive length dt has the order of magnitude $(dt)^{1/2}$. This is wild enough, because $(dt)^{1/2}$ is much larger than dt when dt is small, but one can imagine much wilder fluctuations—say fluctuations of order $(dt)^{1/3}$.

Averaging Market Changes The diffusion model authorizes the use of the law of large numbers on a relatively fine time scale. The model says that relative changes in $S(t)$ over nonoverlapping time intervals are independent, no matter how small the intervals, and so by breaking a small interval of time $[t_1, t_2]$ into many even smaller intervals, we can use the law of large numbers to replace certain effects by their theoretical mean values.

Our purely game-theoretic approach does not need the diffusion model for either of these purposes.

A limit on the wildness of price changes can be expressed straightforwardly in our game-theoretic framework: as a constraint on the market, it can be listed among the rules of the game between Investor and Market. Market simply is not allowed to move too wildly. In §9.1, we discuss how such a constraint can be expressed. In a realistic discrete-time framework, it can be expressed in terms of the variation spectrum of the price series $S(t)$. In a theoretical continuous-time framework, where clearer and more elegant theorems can be proven, it can be expressed in terms of the variation exponent or the Hölder exponent of $S(t)$.

The notion that the market is constrained in how it can change prices should, of course, be taken with a grain of salt. The market can do what it wants. But no theory is possible without some regularity assumptions, and the game-theoretic framework can be credited for making clear the nature of the assumptions that the Black-Scholes argument requires.

As we explain in §9.2, the use of the law of large numbers on a fine time scale is more problematic. The problem is that the hedging of options must actually be implemented on a relatively coarse time scale. Transaction costs limit the frequency with which it is practical or desirable to trade in $\mathcal{S}$, and the discreteness of actual trades limits the fineness of the scale at which the price process $S(t)$ is even well defined. In practice, the interval dt between adjustments in one's holdings of $\mathcal{S}$ is more likely to be a day than a millisecond, and this makes the appeal to the law of large numbers in the Black-Scholes argument appear pollyannaish, for this appeal requires the total change in $S(t)$ to remain negligible during enough dts for the law of large numbers to do its work.

In our judgment, the appeal to the law of large numbers is the weak point of the Black-Scholes method and may be partly responsible for the substantial departures from the Black-Scholes formula often observed when option prices are determined by supply and demand. In any case, the appeal is unpersuasive in the game-theoretic framework, and in order to eliminate it, we need a significant change in our understanding of how options should be priced and hedged.

When time is measured in days, acceptance of the Black-Scholes use of the law of large numbers amounts, roughly speaking, to the assumption that a derivative security that pays daily dividends equal to $(dS(t)/S(t))^2$ should decrease in price linearly: its price at time t should be $\sigma^2(T-t)$, where σ^2 is the variance rate of the process $S(t)$ and T is the date the commitment to pay the dividends expires. We propose to drop this assumption and have the market actually price this dividend-paying derivative, the *variance derivative*, as we shall call it. As we show in §9.3, this produces a purely game-theoretic version of the Black-Scholes formula, in which the current market price of the variance derivative replaces the statistical estimate of $\sigma^2(T-t)$ that appears in the standard Black-Scholes formula. A derivative that pays dividends may not be very manageable, and in §12.2 we explain how it might be replaced with a more ordinary derivative. But the variance derivative is well suited to this chapter's explanation of our fundamental idea: cure the shortcomings of the Black-Scholes method and make it purely game-theoretic by asking the market to price $\mathcal{S}$'s volatility as well as $\mathcal{S}$ itself.

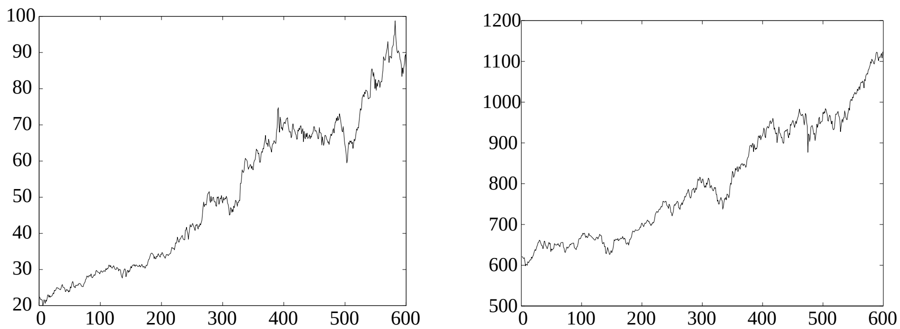

Fig. 9.1 The graph on the left shows daily closing prices, in dollars, for shares of Microsoft for 600 working days starting January 1, 1996. The graph on the right shows daily closing values of the S&P 500 index for the same 600 days.

After discussing stochastic and game-theoretical models for the behavior of stock-market prices and their application to option pricing, we move on, in §9.4, to a more theoretical topic: the efficient-market hypothesis. This section serves as an introduction to Chapter 15.

In an appendix, §9.5, we discuss various ways that the prices given by the Black-Scholes formula are adjusted in practice to bring them into line with the forces of supply and demand. In some cases these adjustments can be thought of as responses to the market's pricing of the volatility of $S(t)$. So our proposal is not radical with respect to practice.

In a second appendix, §9.6, we provide additional information on stochastic option pricing, including a derivation of the stochastic differential equation for the logarithm of a diffusion process, a statement of Itô's lemma, and a sketch of the general theory of risk-neutral valuation.

9.1 THE BEHAVIOR OF STOCK-MARKET PRICES

The erratic and apparently random character of changes in stock-market prices has been recognized ever since the publication of Louis Bachelier's doctoral dissertation, *Théorie de la Spéculation*, in 1900 [9, 60]. To appreciate just how erratic the behavior of these prices is, it suffices to glance at a few time-series plots, as in Figure 9.1.

Bachelier proposed that the price of a stock moves like what is now called *Brownian motion*. This means that changes in the price over nonoverlapping intervals of time are independent and Gaussian, with the variance of each price change proportional to the length of time involved. Prominent among the several arguments Bachelier gave for each change being Gaussian was the claim that it is sum of many smaller changes, resulting from many independent influences; the idea that the Gaussian distribution appears whenever many small independent influences are at play was already widely accepted in 1900 ([3], Chapter 6).

An important shortcoming of Bachelier's model, which he himself recognized, is that it allows the price to become negative. The price of a share of a corporation cannot be negative, because the liability of the owner of the share is limited to what he pays to buy it. But this difficulty is easily eliminated: we may assume that the logarithm of the share price, rather than the price itself, follows a Brownian motion. In this case, we say that the price follows a *geometric Brownian motion*.

Norbert Wiener (1894–1964) at MIT in the 1920s. He was the first to put Brownian motion into the measure-theoretic framework.

In this section, we study how these stochastic models constrain the jaggedness of the path followed by the prices. In the next two chapters, we will use what we learn here to express these constraints directly in game-theoretic terms—as constraints on Market's moves in the game between Investor and Market. As we will see, this has many advantages. One advantage is that we can study the constraints in a realistically finite and discrete setting, instead of relying only on asymptotic theory obtained by making the length of time between successive price measurements or portfolio adjustments infinitely small. A second advantage is that we cannot avoid acknowledging the contingency of the constraints. At best, they are expectations based on the past behavior of Market, or perhaps on our understanding of the strategic interaction among the many players who comprise Market, and Market may well decide to violate them if Investor, perhaps together with some of these other players, puts enough money on the table.

Our main tool for describing constraints on Market's moves in a discrete-time game between Market and Investor is the variation spectrum. We define the variation spectrum in this section, and we explain how it behaves for the usual stochastic models and for typical price series such as those in Figure 9.1. We also discuss the variation and Hölder exponents. These exponents can be defined only asymptotically, but once we understand how they are related to the variation spectrum, which is meaningful in the discrete context, we will be able to relate continuous-time theory based on them to discrete games between Investor and Market.

Brownian Motion

Five years after Bachelier published his dissertation, which was concerned with the motion of stock prices, the physicists Albert Einstein (1879–1955) and Marian von Smoluchowski (1872–1917) proposed the same model for the motion of particles suspended in a liquid (Einstein 1905, Smoluchowski 1906; see also [45, 110, 318]).

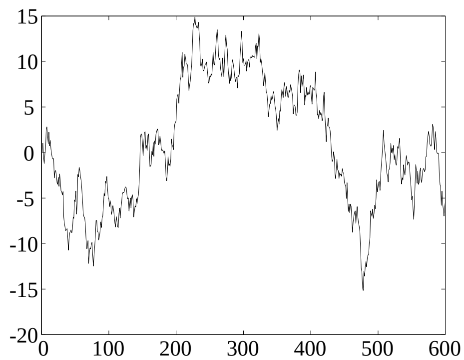

Fig. 9.2 A realization of the Wiener process. Each of the 600 steps in this graph is independent, with mean zero and variance one.

Because experimental study of the motion of such particles had been initiated by the naturalist Robert Brown in 1828, and because Einstein's and Smoluchowski's predictions were promptly verified experimentally, the stochastic process that Bachelier, Einstein, and Smoluchowski studied became known as *Brownian motion*.

Intuitively, Brownian motion is a continuous limit of a random walk. But as Norbert Wiener showed in the early 1920s, it can be described directly in terms of a probability measure over a space of continuous paths [102, 346]. As Wiener showed, it is legitimate to talk about a random real-valued continuous function W on $[0, \infty)$ such that

- $W(0) = 0$,
- for each $t > 0$, $W(t)$ is Gaussian with mean zero and variance t, and
- if the intervals $[t_1, t_2]$ and $[u_1, u_2]$ do not overlap, then the random variables $W(t_2) - W(t_1)$ and $W(u_2) - W(u_1)$ are independent.

We now call such a random function a *Wiener process*. It is nowhere differentiable, and its jaggedness makes it similar to observed price series like those in Figure 9.1. If dt is a small positive number, then the increment $W(t + dt) - W(t)$ is Gaussian with mean zero and variance dt. This means in particular that its order of magnitude is $(dt)^{1/2}$; this is the $\sqrt{dt}$ effect.

One realization of the Wiener process (one continuous path chosen at random according to the probabilities given by the Wiener measure) is shown in Figure 9.2. Actually, this figure shows a random walk with 600 steps up or down, each Gaussian with mean zero and variance one. In theory, this is not the same thing as a realization of the Wiener process, which has the same jagged appearance no matter how fine the scale at which it is viewed. But this finer structure would not be visible at the scale of the figure.

Diffusion Processes

In practice, we may want to measure Brownian motion on an arbitrary scale and allow the mean to change. So given a Wiener process W, we call any process S of the form

$$S(t) = \mu t + \sigma W(t), \tag{9.1}$$

where $\mu \in \mathbb{R}$ and $\sigma \geq 0$, a *Brownian motion*. The random variable $S(t)$ then has mean μt and variance $\sigma^2 t$. We call μ the *drift* of the process, we call σ its *volatility*, and we call σ^2 its *variance*.

Equation (9.1) implies that any positive real number dt, we may write

$$dS(t) = \mu dt + \sigma dW(t), \tag{9.2}$$

where, as usual, $dS(t) := S(t+dt) - S(t)$ and $dW(t) := W(t+dt) - W(t)$. When dt is interpreted as an infinitely small number rather than an ordinary real number, this is called a *stochastic differential equation* and is given a rigorous meaning in terms of a corresponding stochastic integral equation (see §9.6).

We obtain a wider class of stochastic processes when we allow the drift and volatility in the stochastic differential equation to depend on S and t:

$$dS(t) = \mu(S(t), t)dt + \sigma(S(t), t)dW(t). \tag{9.3}$$

Stochastic processes that satisfy stochastic differential equations of this form are called *diffusion processes*, because of the connection with the diffusion equation (also known as the heat equation; see p. 128) and because the random walk represented by W diffuses the probabilities for the position of the path as time goes on.

A diffusion process S that satisfies (9.3) has the Markov property: the probabilities for what S does next depend only on the current state, $S(t)$. We can generalize further, allowing the drift and volatility to depend on the whole preceding path of S rather than merely the current value $S(t)$. This produces a wider class of processes, which are often called *Itô processes* in honor of the Japanese mathematician Kiyosi Itô. But in this book, we consider only the Markov case, (9.3).

Itô developed a calculus for the study of stochastic differential equations, the *stochastic calculus*. The centerpiece of this calculus is Itô's lemma, which allows us write down, from knowledge of the stochastic differential equation governing an Itô process S, the stochastic differential equation governing the Itô process $U(S)$, where U is a well-behaved function. We state Itô's lemma in §9.5, and we prove a game-theoretic variant in §14.2. But in the body of this chapter we avoid Itô's lemma in favor of more direct heuristic arguments, whose robustness is more easily analyzed when we relax the rather strong assumptions that define the Wiener process.

The measure-theoretic account of continuous-time stochastic processes is essentially asymptotic: it shows us only what is left in the limit as the time steps of the underlying random walk (represented by W) are made shorter and shorter. Although it makes for beauty and clarity, the asymptotic character of the theory can create difficulties when we want to gauge the value and validity of applications to phenomena such as finance, which are discrete on a relatively macroscopic level. One of our goals in this book is to develop less asymptotic tools, whose relevance to discrete phenomena can be understood more directly and more clearly.

Osborne's Log-Gaussian Model (Geometric Brownian Motion)

Although Bachelier continued publishing on probability and finance through the 1930s, his fellow probabilists gave his work on probability little prominence and ignored his work on finance [60]. Consequently, the British and American statisticians and economists who began examining stock-market prices empirically and theoretically in the middle of the twentieth century had to rediscover for themselves the relevance of the Wiener process. The principal contribution was made in 1959 by the American astrophysicist M. F. Maury Osborne, who was the first to publish a detailed study of the hypothesis that $S(t)$ follows a geometric Brownian motion. This has been called Osborne's *log-Gaussian model*, because it says that the logarithm of the price $S(t)$, rather than $S(t)$ itself, follows a Brownian motion [237, 337].

If $\ln S(t)$ follows a Brownian motion, then we may write

$$d \ln S(t) = \mu_0 dt + \sigma_0 dW(t).$$

It follows (see p. 231) that the relative increments $dS(t)/S(t)$ satisfy a stochastic differential equation of the same form:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t). \tag{9.4}$$

So S itself is also a diffusion process:

$$dS(t) = \mu S(t) dt + \sigma S(t) dW(t). \tag{9.5}$$

The stochastic differential equation (9.5) will be the starting point for our review in the next section of the Black-Scholes model for option pricing.

Figure 9.3 shows two realized paths of the random motion defined by (9.5), with parameters chosen to match the apparent trend and volatility of the Microsoft prices shown in Figure 9.1. The parameters are the same for the two paths. Both start at 22.4, the initial price of Microsoft in Figure 9.1, and for both the theoretical drift μ is 0.0024 (the average daily return on Microsoft's stock), and the theoretical volatility σ is 0.0197 (the square root of the average daily squared return on Microsoft's stock). The fact that σ is multiplied by $S(t)$ in (9.5) is visible in these graphs: the magnitude of the fluctuations tends to be larger when $S(t)$ is larger. This same tendency towards larger fluctuations when the price is higher is also evident in Figure 9.1. On the other hand, the two paths show very different overall trends; the first ends up about 50% higher than Microsoft does at the end of 600 days, while the second ends up only about $1/3$ as high. This illustrates one aspect of the relatively weak role played by the drift μ in the diffusion model; it can easily be dominated by the random drift produced by the failure of the $dW(t)$ to average exactly to zero.

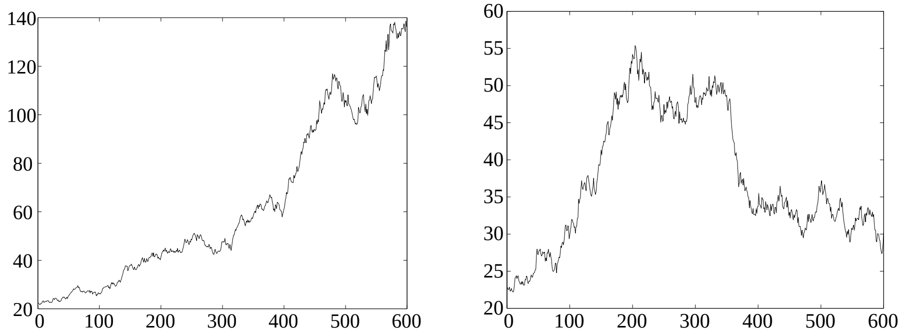

Fig. 9.3 Two realizations of the same geometric Brownian motion, whose parameters are chosen to match the apparent trend and volatility in the Microsoft price series in Figure 9.1: $S(0) = 22.4375$, $\mu = 0.0024$, and $\sigma = 0.0197$.

Benoit Mandelbrot (born 1924) in 1983. Since the 1960s, Mandelbrot has championed alternatives to the log-Gaussian model.

In the four decades since Osborne's formulation of the log-Gaussian model, abundant evidence has been found for deviations from it: dependencies, skewness, non-Gaussian kurtosis, and heteroskedasticity [212, 216, 208]. Some authors, most notably Benoit Mandelbrot, have argued for alternative stochastic models. But most researchers in finance have not found these alternatives attractive or useful. They have continued to take the log-Gaussian model as their starting point, making ad hoc adjustments and elaborations as necessary when empirical deviations become too troublesome.

We agree with Mandelbrot that the log-Gaussian model is a doubtful starting point. But we do not propose replacing it with a different stochastic model. Instead, we propose dropping the assumption that the market is governed by a stochastic model. We now turn to a tool that can help us do this, the variation spectrum.

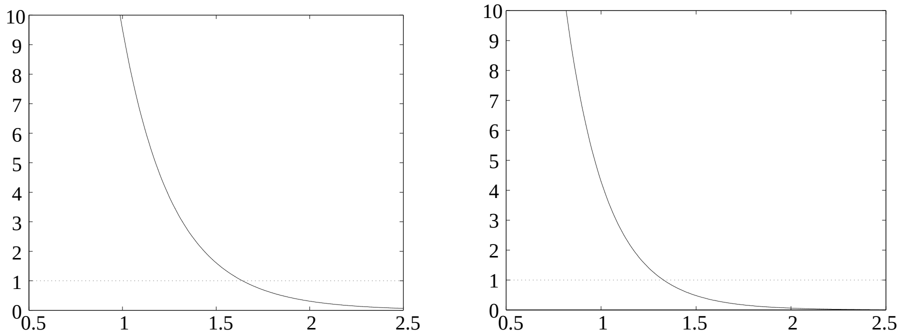

Fig. 9.4 Variation spectra for the two price series in Figure 9.1—Microsoft on the left and the S&P 500 on the right. In both cases, we have rescaled the data in Figure 9.1 by taking the median value as our unit. The median price for Microsoft over the 600 working days in Figure 9.1 is approximately $\$48.91$; the median value of the S&P 500 index is approximately 775. This unit determines the vertical scale in the picture of the variation spectrum. As an example, consider $\mathbf{var}_{600}(1)$, which is the sum of the absolute values of the daily changes. In the case of Microsoft, $\mathbf{var}_{600}(1)$ is approximately 9.5, meaning that the total of the absolute sizes of the daily changes over 600 days is 9.5 times the median price. In the case of the S&P 500, $\mathbf{var}_{600}(1)$ is approximately 4.3, meaning that the total of the absolute sizes of the daily changes is 4.3 times the median value of the index.

The Variation Spectrum

Consider a continuous function $S(t)$ on the finite interval $[0, T]$. Choose an integer N, set

$$x_n = S\left(n\frac{T}{N}\right) - S\left((n-1)\frac{T}{N}\right)$$

for $n = 1, \ldots, N$, and set

$$\mathbf{var}_{S,N}(p) := \sum_{n=1}^{N} |x_n|^p \tag{9.6}$$

for all $p > 0$. We call $\mathbf{var}_{S,N}(p)$ the p*-variation* of S, and we call the function $\mathbf{var}_{S,N}$ the *variation spectrum* for S. We abbreviate $\mathbf{var}_{S,N}$ to $\mathbf{var}_S$ or $\mathbf{var}_N$ or even to $\mathbf{var}$ in contexts that fix the missing parameters. If we set $dt := T/N$, then we can also write (9.6) in the form

$$\mathbf{var}_{S,N}(p) := \sum_{n=0}^{N-1} |dS(ndt)|^p, \tag{9.7}$$

where, as usual, $dS(t) := dS(t + dt) - dS(t)$.

Figure 9.4 shows the variation spectra for the Microsoft and S&P 500 data shown in Figure 9.1. These spectra display some typical features: the p-variation has large values for p less than 1 and small values for p greater than 2. This is to be expected; if, e.g., $0 < |x_n| < 1$ for all n, then $\mathbf{var}_{S,N}(p)$ decreases continuously in p, with $\lim_{p\to 0} \mathbf{var}_{S,N}(p) = N$ and $\lim_{p\to\infty} \mathbf{var}_{S,N}(p) = 0$. The fall from large to small values is gradual, however, and the apparent position of the transition depends strongly on the unit of measurement for the original series. This dependence is illustrated by Table 9.1, which gives $\mathbf{var}_{600}(2)$ for the Microsoft and S&P 500 data measured with several different units.

In many of the games between Investor and Market that we will study, the ability of Investor to hedge the sale of an option depends on Market choosing his moves $x_1, \ldots, x_N$ so that $\mathbf{var}_N(2+\epsilon)$ is small for some small positive ϵ. In Chapter 10, for example, we prove that under certain conditions, which include $\mathbf{var}_N(2+\epsilon) \le \delta$, the game-theoretic upper and lower prices of an option are approximated in discrete time by our Black-Scholes formula with an error no greater than $K\delta^{1/4}$ for some constant K. Here it makes sense that $\mathbf{var}_N$ should depend on the unit of measurement for the price of the underlying security, for the error in pricing the option is also measured using some arbitrary unit, and the constant K depends on how these two units are related. But the devil is in the details. It is not enough to say that the Black-Scholes formula will be accurate if $\mathbf{var}_N(2+\epsilon)$ is small. Just how small a bound we can put on $\mathbf{var}_N(2+\epsilon)$ will determine whether the formula will give a reasonable approximation or be an asymptotic irrelevance.

The Variation and Hölder Exponents

In addition to a practical theory, in which we wade through messy calculations to obtain bounds on the accuracy of hedging in terms of bounds on 2-variation, we do also want an asymptotic theory, in which we clear away the clutter of practical detail in order to see the big picture more clearly. This book is far from practical, and asymptotic theory will occupy most of the remaining chapters. In this asymptotic theory, we will emphasize the value of p at which the p-variation drops from large (asymptotically infinite) to small (asymptotically zero). In practice, as we have just seen, this value is scarcely sharply defined. But an asymptotic theory will assume,

Table 9.1 Dependence of $\mathbf{var}_{600}(2)$ and $\mathbf{var}_{600}(2.5)$ on the choice of a unit for the Microsoft price or the S&P 500 index. The p-variations are given to three significant figures or three decimal places, whichever is less.

Microsoft			S&P 500		
unit	$\mathbf{var}(2)$	$\mathbf{var}(2.5)$	unit	$\mathbf{var}(2)$	$\mathbf{var}(2.5)$
dollar	742	$1,100$	index point	$37,700$	$160,000$
median ($48.91)	0.310	0.066	median (775)	0.063	0.010
initial ($22.44)	1.473	0.463	initial (621)	0.098	0.017
final ($86.06)	0.100	0.016	final (1124)	0.030	0.004

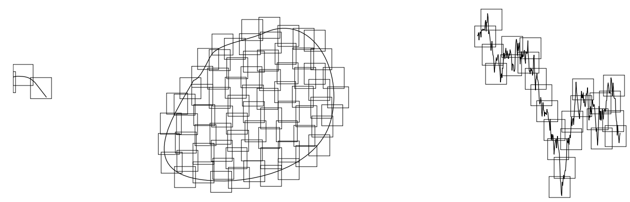

Fig. 9.5 Mandelbrot's concept of box dimension ([216], pp. 172–173). Intuitively, the box dimension of an object in the plane is the power to which we should raise $1/dt$ to get, to the right order of magnitude, the number of $dt \times dt$ boxes required to cover it. In the case of an object of area A, about $A/(dt)^2$ boxes are required, so the box dimension is 2. In the case of a smooth curve of length T, T/dt boxes are required, so the box dimension is 1. In the case of the graph of a function on $[0, T]$ with Hölder exponent H, we must cover a vertical distance $(dt)^H$ above the typical increment dt on the horizontal axis, which requires $(dt)^H/dt$ boxes. So the order of magnitude of the number of boxes needed for all T/dt increments is $(dt)^{H-2}$; the box dimension is $2 - H$.

one way or another, that it is defined. It is called the *variation exponent*, and its inverse is called the *Hölder exponent*, in honor of Ludwig Otto Hölder (1859–1937).

We brush shoulders with the idea of a Hölder exponent whenever we talk informally about the order of magnitude of small numbers. What is the order of magnitude of $x_1, \ldots, x_N$ relative to $1/N$? If N is large and the $|x_n|$ have the same order of magnitude as $(1/N)^H$ on average, where $0 < H \leq 1$, then the order of magnitude of the p-variation will be

$$N\left(\frac{1}{N}\right)^{Hp} = N^{1-Hp}. \tag{9.8}$$

This will be close to zero for $p > 1/H + \epsilon$ and large for $p < 1/H - \epsilon$, where ϵ is some small positive number. Intuitively, $1/H$ is the variation exponent, and H is the Hölder exponent. In other words, the Hölder exponent of $S(t)$ is the number H such that $dS(t)$ has the order of magnitude $(dt)^H$.

Following Mandelbrot ([215], Part II, §2.3; [216], p. 160), we also introduce a name for $2 - H$; this is the *box dimension*. Mandelbrot's rationale for this name is sketched in Figure 9.5. Although this rationale is only heuristic, the concept gives visual meaning to the Hölder exponent. In general, the box dimension of the graph of a time series falls between 1 and 2, because the jaggedness of the graph falls somewhere between that of a line or a smooth curve and that of a graph so dense that it almost completely fills a two-dimensional area.

As Figure 9.5 suggests, an ordinary continuous function should have box dimension 1, which means that its Hölder exponent and variation exponent should also be 1. For a wide class of stochastic processes with independent increments, including all diffusion processes with $\sigma > 0$, we expect the Hölder exponent of a path to be $1/2$; this expresses the idea that the increment dS has the order of magnitude $(dt)^{1/2}$. Functions that fall between an ordinary function and a diffusion process in their jaggedness may be called *substochastic*. These benchmarks are summarized in Table 9.2.

For a greater variety of values for the Hölder exponent, we may consider the fractional Brownian motions. The *fractional Brownian motion* with index $h \in (0, 1)$ is a stochastic process B_h such that $B_h(0) = 0$, values $B_h(t)$ for $t > 0$ are jointly Gaussian, and the variance of an increment $B_h(t) - B_h(s)$, where $0 < s < t$, is $(t - s)^{2h}$. If $h = 0.5$, then the fractional Brownian motion reduces to the usual Brownian motion, but for other values of h even nonoverlapping increments are correlated. We can assume that the sample paths are continuous (cf. [262], Corollary 25.6 on p. 61 of Volume 1). The Hölder exponent should approximate the index h. Figure 9.6 shows some sample paths for different h. See [13, 48, 214, 216].

We use such cautious turns of phrase ("the Hölder exponent is supposed to be such and such") because there is no canonical way to define the Hölder exponent precisely. In practice, we cannot determine a Hölder exponent for an arbitrary continuous function $S(t)$ on $[0, t]$ with any meaningful precision unless we chop the interval into an absolutely huge number of increments, and even then there will be an unattractive dependence on just how we do the chopping. In order to formulate a theoretical definition, we cannot merely look at the behavior in the limit of a division of $[0, T]$ into N equal parts; in general, we must pass to infinity much faster than this, and exactly how we do it matters. In the nonstochastic and nonstandard framework that we use in Chapter 11, this arbitrariness is expressed as the choice of an arbitrary infinite positive integer N.

The practical implication of the relative indeterminacy of the Hölder exponent is clear. Unlike many other theoretical continuous-time concepts, the Hölder exponent is not something that we should try to approximate in a discrete reality. It has a less direct meaning. When a statement in continuous-time theory mentions a Hölder exponent H, we can use that statement in a discrete reality only after translating it into a statement about the p-variation for values of p at a decent distance from $1/H$. For example, the statement that Market's path will have Hölder exponent $1/2$ or less must be translated into a condition about the relative smallness of the p-variation for $p = 2 + \epsilon$, where $\epsilon > 0$ is not excessively small.

Table 9.2 Typical values for our three related measures of the jaggedness of a continuous real-valued function S. The figures given for a substochastic function are typical of those reported for price series by Mandelbrot.

	Hölder exponent H	variation exponent vex	box dimension dim
definition		$\text{vex} := 1/H$	$\text{dim} := 2 - H$
range	$0 \leq H \leq 1$	$1 \leq \text{vex} \leq \infty$	$1 \leq \text{dim} \leq 2$
ordinary function	1	1	1
substochastic	0.57	1.75	1.43
diffusion process	0.5	2	1.5

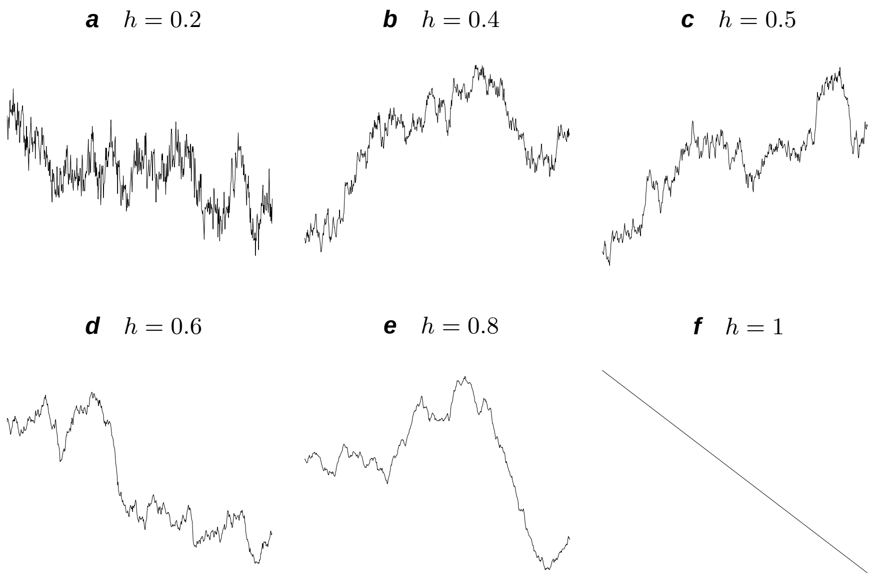

Fig. 9.6 Sample paths for fractional Brownian motions with different values for the index h. The Hölder exponent of a sample path is supposed to approximate the index h. When $h = 0.5$, the fractional Brownian motion reduces to ordinary Brownian motion. When $h = 1$, it reduces to a straight line.

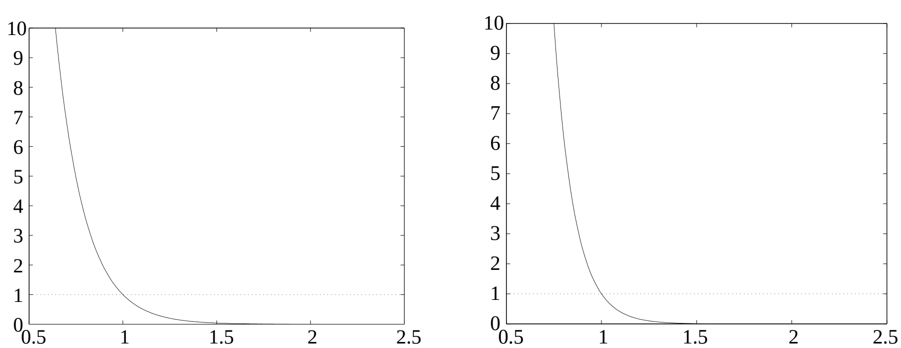

Fig. 9.7 Graphs of variation spectra for the straight line $S(t) = t$ on $0 \leq t \leq 1$. On the left is the variation spectrum based on a division of the interval $[0, 1]$ into 600 steps. On the right is the variation spectrum based on a division into $10,000$ steps. These graphs fall from values greater than one to values less than one when we cross $p = 1$ on the horizontal axis, but the fall is less than abrupt.

In order to be fully persuasive on this point, we should perhaps look at some further examples of the indeterminacy of the Hölder exponent even when N is huge. Consider the best behaved function there is: the linear function $S(t) = t$ for $0 \leq t \leq 1$. We split the interval $[0, 1]$ into N parts, and (9.7) becomes

$$\mathbf{var}_{S,N}(p) := \sum_{n=0}^{N-1} \left(\frac{1}{N}\right)^p = N^{1-p}.$$

Figure 9.7 graphs this function for $N = 600$ and $N = 10,000$. As N tends to infinity, N^{1-p} tends to zero for $p > 1$ and to infinity for $p < 1$, confirming that the Hölder exponent for S is 1. But how large does N have to be in order for the variation spectrum to identify this value to within, say, 10%? For example, how large does N have to be in order for

$$\mathbf{var}_{S,N}(p) \geq 10 \text{ for } p \leq 0.9 \qquad \text{and} \qquad \mathbf{var}_{S,N}(p) \leq 0.1 \text{ for } p \geq 1.1 \tag{9.9}$$

to hold? This is not hard to answer: the conditions $N^{1-0.9} \geq 10$ and $N^{1-1.1} \leq 0.1$ are both equivalent to $N \geq 10^{10}$. And even when N is in the billions, there are still arbitrary choices to be made. If we had chosen our linear function on $[0, 1]$ to be $S(t) = 100t$ instead of $S(t) = t$, $\mathbf{var}_N(1)$ would not be 1, as shown in Figure 9.7 and taken for granted by our question about (9.9). Instead, it would be 100.

The story is similar for other continuous functions, including sample paths for diffusion processes: we do not get even the appearance of a sharply defined value for the Hölder exponent unless N is in the billions. Figure 9.8 illustrates the point with the variation spectra for two realizations of the Wiener process, one with $N = 600$ and one with $N = 10,000$. In this case, the Hölder exponent is 0.5: the p-variation tends to infinity with N for p less than 2 and to zero for p greater than 2. The scale in which the Wiener process is measured has the convenient feature that $\mathbb{E}\,\mathbf{var}_N(2) = 1$ for all N, but the drop at the point $p = 2$ is no more abrupt than for the linear function. We should note that Figure 9.8 is not affected by sampling error; different realizations give curves that cannot be distinguished visually.

The $\mathbf{var}_{600}$ for Microsoft and the $\mathbf{var}_{600}$ for the S&P 500, both of which we displayed in Figure 9.4, are similar to the $\mathbf{var}_{600}$ for the Wiener process in Figure 9.8. But in the Microsoft and S&P 500 cases, we must use empirical scales for the data (in Figure 9.1, our unit was the observed median in each case), and so we cannot create the illusion that we can identify 2 as the crossover point from large to small p-variation.

Why Do Stock-Market Prices Move Like Brownian Motion?

As we have explained, we will use the variation spectrum and the variation exponent in game-theoretic versions of Black-Scholes pricing that do not rely on the assumption that stock-market prices are stochastic. In these versions, the assumption that the stock price $S(t)$ follows a geometric Brownian motion with theoretical volatility σ is replaced by the assumption that the market prices a derivative security $\mathcal{D}$ that

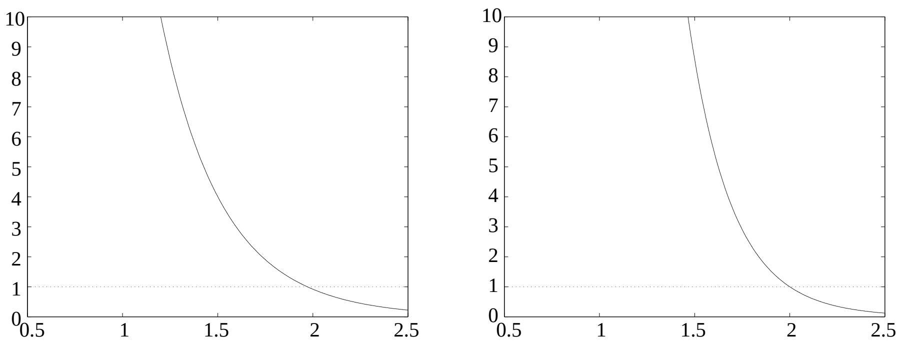

Fig. 9.8 The variation spectrum for a sample path of the Wiener process, for 600 and $10,000$ sample points.

pays dividends that add up to the actual future (relative) variance of $S(t)$. We then assume bounds on the p-variations of $S(t)$ and the price $D(t)$ of the traded derivative. In discrete time, we assume upper bounds on $\mathbf{var}_S(2+\epsilon)$ and $\mathbf{var}_D(2-\epsilon)$ (Proposition 10.3, p. 249). In our continuous-time idealization, bounds on the wildness of $S(t)$ are not even needed if we can assume that the payoff of the option we want to price is always nonnegative and that $\mathcal{S}$ does not become worthless; in this case it is enough that the variation exponent of the variance security $\mathcal{D}$ be greater than 2 (or, equivalently, that its Hölder exponent be less than $1/2$) (Theorem 11.2, p. 280).

The movement of stock-market prices does look roughly like Brownian motion. But in our theory this is a consequence, not an assumption, of our market games for pricing options. This is made clear in our continuous-time theory, where we prove (Proposition 11.1, p. 281) that Market can avoid allowing Investor to become infinitely rich only by choosing his $dS(t)$ so that $S(t)$ has variation exponent exactly equal to 2. This is a way of expressing the truism that it is the operation of the market that makes price changes look like a random walk. The game-theoretic framework clears away the sediment of stochasticism that has obscured this truism, and in its discrete-time form it allows us to explore how far the effect must go in order for Black-Scholes pricing to work.

9.2 THE STOCHASTIC BLACK-SCHOLES FORMULA

The *Black-Scholes formula* was published in 1973, in two celebrated articles, one by Fischer Black and Myron Scholes, the other by Robert C. Merton [30, 229, 28, 16]. This formula, which prices a wide class of derivatives, has facilitated an explosive growth in the markets for options and more complex derivatives. In addition to filling the need for objective pricing (a starting point for valuation by creditors and auditors), it gives those who write (originate) options guidance on how to hedge the risks they are taking, and it permits a great variety of adjustments that can bring the ultimate pricing into line with supply and demand. It was recognized by a Nobel Prize in 1997.

As we explained in the introduction to this chapter, the Black-Scholes formula relies on the assumption that the price of the underlying stock follows a geometric Brownian motion. In this section, we review the derivation, informally but with careful attention to how the stochastic assumption is used. This will set the stage for our presentation, in the next section, of our purely game-theoretic Black-Scholes formula.

European Options

Fischer Black (1938–1995) in 1975. Because of his early death, Black did not share in the 1997 Nobel prize for economics, which was awarded to Myron Scholes and Robert C. Merton.

Recall that a derivative (or, more properly, a derivative security) is a contract whose payoff depends on the future movement of the prices of one or more commodities, securities, or currencies. The derivative's payoff may depend on these prices in a complicated way, but *European options* are relatively simple; their payoffs depend only on the price of a single security at a fixed date of maturity. A European option $\mathcal{U}$ on an underlying security $\mathcal{S}$ is characterized by its maturity date, say T, and its payoff function, say U. Its payoff at time T is, by definition,

$$\mathcal{U}(T) := U(S(T)). \tag{9.10}$$

The problem is to price $\mathcal{U}$ at a time t before T. What price should a bank charge at time 0, say, for a contract that requires it to pay (9.10) at time T?

The most familiar European option is the *European call*, which allows the holder to buy a security at a price set in advance. The holder of a European call on $\mathcal{S}$ with *strike price* c and maturity T will exercise it only if $S(T)$ exceeds c, and he can then immediately realize the profit $S(T) - c$ by reselling the security. So the payoff function U of the call is

$$U(S) := \begin{cases} S - c & \text{if } S > c \\ 0 & \text{if } S \le c. \end{cases}$$

In practice, the bank selling the option and the customer buying it usually do not bother with the underlying security; the bank simply agrees to pay the customer $U(S(T))$.

At first glance, one might expect that buyers of call options would be motivated by the belief that the price of the stock $\mathcal{S}$ will go up. This is often true in practice, especially for buyers who cannot afford to buy the stock outright. But buying the stock is a more straightforward and efficient way to bet on a price increase. The buyer of a call option will generally be charged more for his potential payoff than the interest on the money needed to buy the stock, because he is not risking the loss a buyer of the stock would incur if its price goes down instead of up. If he can afford to buy the stock but buys the stock option instead, the fact that he is paying to avoid this risk reveals that he is not really so confident that the price will go up. He may be counting only on a big change in the price. If it goes up, he will make a lot of money. If it goes down, he will lose only the relatively small price he pays for the option.

Another important European option is the European put, which entitles its holder to sell stock at a given price at a given time. It can be analyzed in the same way as we have analyzed the European call. European options are less popular, however, than American options, which can be exercised at a time of the holder's choosing. We analyze American options in Chapter 13.

The Market Game

In the 1950s and early 1960s, when American economists first attacked the problem of pricing derivatives, they considered a number of factors that might influence the price an investor would be willing to pay, including the investor's attitude toward risk and the prospects for the underlying security [56, 114, 264]. But the formula derived by Black and Scholes involved, to their own surprise, only the time to maturity T, the option's payoff function U, the current price $S(t)$ of the underlying security $\mathcal{S}$, and the volatility of $\mathcal{S}$'s price. Moreover, as they came to realize in discussions with Merton, the formula requires very little economic theory for its justification. If the price of $\mathcal{S}$ follows a geometric Brownian motion, then the price the formula gives for $\mathcal{U}$ is its game-theoretic price, in the sense of Part I, in a game between Investor and Market in which Investor is allowed to continuously adjust the amount of the stock $\mathcal{S}$ he holds. The protocol for this game looks like this:

THE BLACK-SCHOLES PROTOCOL
Parameters: $T > 0$ and $N \in \mathbb{N}$; $dt := T/N$
Players: Investor, Market
Protocol:
$\mathcal{I}(0) := 0$.
Market announces $S(0) > 0$.
FOR $t = 0, dt, 2dt, \ldots, T - dt$:
Investor announces $\delta(t) \in \mathbb{R}$.
Market announces $dS(t) \in \mathbb{R}$.
$S(t + dt) := S(t) + dS(t)$.
$\mathcal{I}(t + dt) := \mathcal{I}(t) + \delta(t)dS(t)$. (9.11)

Investor's move $\delta(t)$ is the number of shares of the stock he holds during the period t to $t+dt$, Market's move $dS(t)$ is the change in the price per share over this period, and hence $\delta(t)dS(t)$ is Investor's gain (or loss). We write $\mathcal{I}$ for Investor's capital process.

Investor starts with zero capital, but at each step he can borrow money to buy stock or borrow stock to sell, in amounts as large as he wants. So his move $\delta(t)$ may be any number, positive, zero, or negative. Market may also choose his move positive, zero, or negative, but he cannot allow $S(t)$ to become negative. If $S(t)$ ever becomes zero, the firm issuing the stock is bankrupt, and $S(t)$ must remain zero.

For simplicity, we assume that the interest rate is zero. Thus we do not need to adjust Investor's gain, $\delta(t)dS(t)$, to account for his payment or receipt of interest. We also assume that transaction costs are zero; Investor does not incur any fees when he buys or sells stock in order to change the number of shares he is holding from $\delta(t)$ to $\delta(t+dt)$.

The Stochastic Assumption

To the protocol we have just described, Black and Scholes added the assumption that $S(t)$ follows a geometric Brownian motion. In other words, Market's moves must obey the stochastic differential equation for geometric Brownian motion,

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t). \tag{9.12}$$

But the derivation of the Black-Scholes formula does not actually use the full force of this assumption. What it does use can be boiled down to three subsidiary assumptions:

1. **The $\sqrt{dt}$ effect.** The variation exponent of the $S(t)$ is 2. In other words, the order of magnitude of the $dS(t)$ is $(dt)^{1/2}$. This is a constraint on the wildness of Market's moves. They cannot take too jagged a path.

2. **Standard deviation proportional to price.** The expected value of $(dS(t))^2$ just before Market makes the move $dS(t)$ is approximately $\sigma^2 S^2(t)dt$.

3. **Authorization to use the law of large numbers.** In a calculation where the squared increments $(dS(t))^2$ are added, they can be replaced by their expected values, $\sigma^2 S^2(t)dt$. This follows from the assumption that the $dW(t)$ are independent, together with the law of large numbers, which is applicable if the time increment dt is sufficiently small.

In our judgment, the most troublesome of these three assumptions is the third. The first assumption, as we explained in the preceding section, can be re-expressed in game-theoretic terms. Adjustments (more or less convincing and more or less cumbersome) can be made to correct for deviations from the second. But the third is risky, simply because the number of terms being averaged may fail to be large enough to justify the use of the law of large numbers. The new Black-Scholes method that we introduce in the next section is motivated in part by our dissatisfaction with this risky use of the law of large numbers.

Assumption 1, that the $dS(t)$ have order of magnitude $(dt)^{1/2}$, follows, of course, from (9.12) and the fact that the $dW(t)$, as increments in a Wiener process, have this order of magnitude. The first term on the right-hand side of (9.12), $\mu S(t)dt$, can be neglected, because dt is much smaller than $(dt)^{1/2}$.

Assumption 2 follows similarly when we square both sides of (9.12):

$$(dS(t))^2 = S^2(t)\left(\mu^2(dt)^2 + 2\mu\sigma dt dW(t) + \sigma^2(dW(t))^2\right). \tag{9.13}$$

Because $dW(t)$ is of order $(dt)^{1/2}$, $(dW(t))^2$ is of order dt and dominates the other two terms, and hence the approximate expected value of $(dS(t))^2$ is obtained by dropping them and replacing $(dW(t))^2$ by its expected value, dt.

In order to explain Assumption 3 carefully, we first note that as the square of a Gaussian variable with mean zero and variance dt, $(dW(t))^2$ has mean dt and variance $2(dt)^2$. (The Gaussian assumption is not crucial here; the fact that the coefficient of $(dt)^2$ in the variance is exactly 2 depends on it, but this coefficient is of no importance.) This can also be expressed by writing

$$(dW(t))^2 = dt + z, \tag{9.14}$$

where the z has mean zero and variance $2(dt)^2$. In words: $(dW(t))^2$ is equal to dt plus a fluctuation of order dt. Summing (9.14) over all N increments

$$dW(0), dW(dt), dW(2dt), \dots, dW(T-dt),$$

we obtain

$$\sum_{n=0}^{N-1}(dW(ndt))^2 = T + \sum_{n=0}^{N-1} z_n.$$

Because $\sum_{n=0}^{N-1} z_n$ has a total variance of only $2Tdt$, we may neglect it and say that the $(dW(t))^2$ add to the total time T; the z_n cancel each other out. More generally, if the squared increments $(dW(t))^2$ are added only after being multiplied by slowly varying coefficients, such as $\sigma^2(S(t))^2$, we can still expect the z_n to cancel each other out, and so we can simply replace each $(dW(t))^2$ in the sum with dt. Here it is crucial that the time step dt be sufficiently small; before there is any substantial change in the coefficient $S^2(t)$, there must be enough time increments to average the effects of the z_n to zero.

The Derivation

Now to our problem. We want to find $\mathcal{U}(t)$, the price at time t of the European option $\mathcal{U}$ that pays $U(S(T))$ at its maturity T. We begin by optimistically supposing that there is such a price and that it depends only on t and on the current price of the stock, $S(t)$. This means that there is a function of two variables, $\overline{U}(s,t)$, such that $\mathcal{U}(t) = \overline{U}(S(t),t)$. In order to find $\overline{U}$, we investigate the behavior of its increments by means of a Taylor's series.

Considering only terms of order $(dt)^{1/2}$ or dt (i.e., omitting terms of order $(dt)^{3/2}$ and higher, which are much smaller), we obtain

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \frac{\partial \overline{U}}{\partial t} dt + \frac{1}{2}\frac{\partial^2 \overline{U}}{\partial s^2}(dS(t))^2. \tag{9.15}$$

Here there is one term of order $(dt)^{1/2}$—the term in $dS(t)$. There are two terms of order dt—the term in dt and the term in $(dS(t))^2$. The terms of order dt must be included because their coefficients are always positive and hence their cumulative effect (there are T/dt of them) will be nonnegligible. Individually, the $dS(t)$ are much larger, but because they oscillate between positive and negative values while their coefficient varies slowly, their total effect may be comparable to that of the dt terms. (We made this same argument in our heuristic proof of De Moivre's theorem in §6.2.)

Substituting the right-hand side of (9.13) for $(dS(t))^2$ in (9.15) and again retaining only terms of order $(dt)^{1/2}$ and dt, we obtain

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \frac{\partial \overline{U}}{\partial t} dt + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 \overline{U}}{\partial s^2}(dW(t))^2. \tag{9.16}$$

We still have one term of order $(dt)^{1/2}$ and two terms of order dt.

Because its coefficient in (9.16) varies slowly, we replace $(dW(t))^2$ with dt, obtaining

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \left(\frac{\partial \overline{U}}{\partial t} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 \overline{U}}{\partial s^2}\right) dt. \tag{9.17}$$

This is the risky use of the law of large numbers. It is valid only if the coefficient $S^2(t)\partial^2\overline{U}/\partial s^2$ holds steady during enough dt for the $(dW(t))^2$ to average out. Notice that we simplified in our preliminary discussion of this point. The variability in $(dW(t))^2$'s coefficient comes from $\partial^2\overline{U}/\partial s^2$ in addition to $S^2(t)$.

Now we look again at the Black-Scholes protocol. According to (9.11),

$$d\mathcal{I}(t) = \delta(t)dS(t),$$

where $\delta(t)$ is the amount of stock Investor holds from t to $t+dt$. Comparing this with (9.17), we see that we can achieve our goal by setting

$$\delta(t) := \frac{\partial \overline{U}}{\partial s}(S(t),t) \tag{9.18}$$

if we are lucky enough to have

$$\frac{\partial \overline{U}}{\partial t}(S(t),t) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 \overline{U}}{\partial s^2}(S(t),t) = 0$$

for all t, no matter what value $S(t)$ happens to take.

Our problem is thus reduced to a purely mathematical one: we need to find a function $\overline{U}(s,t)$, for $0 < t < T$ and $0 < s < \infty$, that satisfies the partial differential equation

$$\frac{\partial \overline{U}}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 \overline{U}}{\partial s^2} = 0 \tag{9.19}$$

(this is the *Black-Scholes equation*) and the final condition

$$\overline{U}(s,t) \to U(s) \quad (t \to T).$$

As it turns out (see Chapter 11), there is a solution,

$$\overline{U}(s,t) = \int_{-\infty}^{\infty} U\left(se^z\right)\, \mathcal{N}_{-\sigma^2(T-t)/2,\,\sigma^2(T-t)}(dz). \tag{9.20}$$

As the reader will have noted, (9.19) differs only slightly from the heat equation, which we used in our proof of De Moivre's theorem, and (9.20) is similar to the solution of that equation. Both equations are parabolic equations, a class of partial differential equations that have been thoroughly studied and are easily solved numerically; see §6.3 and [68, 131, 352].

So an approximate price at time t for the European option $\mathcal{U}$ with maturity T and payoff $U(S(T))$ is given by

$$\mathcal{U}(t) = \int_{-\infty}^{\infty} U\left(S(t)e^z\right)\, \mathcal{N}_{-\sigma^2(T-t)/2,\,\sigma^2(T-t)}(dz). \tag{9.21}$$

This is the *Black-Scholes formula* for an arbitrary European option. (The formula is more often stated in a form that applies only to calls and puts.) We can replicate the option $\mathcal{U}$ at the price (9.21) by holding a continuously adjusted amount of the underlying security $\mathcal{S}$, the amount $\delta(t)$ held at time t being given by (9.18). Financial institutions that write options often do just this; it is called *delta-hedging*.

Only one of the parameters in (9.12), the volatility σ, plays a role in the derivation we have just outlined. The other parameter, the drift μ, does not appear in the Black-Scholes equation or in the Black-Scholes formula.

Most expositions simplify the argument by using Itô's lemma (p. 232). We have avoided this simplification, because Itô's lemma itself is based on an asymptotic application of the law of large numbers, and so using it would obscure just where such asymptotic approximation comes into play. As we have explained, we are uncomfortable with the application of the law of large numbers that takes us from (9.16) to (9.17), because in practice the length of time dt may be equal to a day or longer, and it may be unreasonable to expect $S^2\partial^2\overline{U}/\partial s^2$ to hold steady for a large number of days.

9.3 A PURELY GAME-THEORETIC BLACK-SCHOLES FORMULA

We now turn to our game-theoretic version of the Black-Scholes formula. We have already explained the main ideas: (1) Market is asked to price both $\mathcal{S}$ and a derivative security $\mathcal{D}$ that pays dividends $(dS(t)/S(t))^2$, and (2) constraints on the wildness of price changes are adopted directly as constraints on Market's moves. We now explain informally how these ideas produce a Black-Scholes formula. The argument is made rigorous in the next two chapters, in discrete time in Chapter 10 and in continuous time in Chapter 11.

Another Look at the Stochastic Derivation

Consider again the derivation of the stochastic Black-Scholes formula. It begins with a Taylor's series:

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \frac{\partial \overline{U}}{\partial t} dt + \frac{1}{2}\frac{\partial^2 \overline{U}}{\partial s^2}(dS(t))^2. \qquad (9.22)$$

The right-hand side of this approximation is the increment in the capital process of an investor who holds shares of two securities during the period from t to $t + dt$:

- $\partial\overline{U}/\partial s$ shares of $\mathcal{S}$, and
- $-\sigma^{-2}\partial\overline{U}/\partial t$ shares of a security $\mathcal{D}$ whose price per share at time t is $\sigma^2(T-t)$ (the remaining variance of $\mathcal{S}$), and which pays a continuous dividend per share amounting, over the period from t to $t + dt$, to

$$-\frac{\sigma^2}{\partial\overline{U}/\partial t}\frac{1}{2}\frac{\partial^2 \overline{U}}{\partial s^2}(dS(t))^2. \qquad (9.23)$$

  The second term on the right-hand side of (9.22) is the capital gain from holding the $-\sigma^{-2}\partial\overline{U}/\partial t$ shares of $\mathcal{D}$, and the third term is the total dividend.

The Black-Scholes equation tells us to choose the function $\overline{U}$ so that the dividend per share, (9.23), reduces to

$$\left(\frac{dS(t)}{S(t)}\right)^2,$$

and the increment in the capital process, (9.22), becomes

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \frac{\partial \overline{U}}{\partial t} dt - \frac{\partial \overline{U}}{\partial t}\frac{(dS(t))^2}{\sigma^2 S^2(t)}. \qquad (9.24)$$

Only at this point do we need the assumption that $S(t)$ follows a geometric Brownian motion. It tells us that $(dS(t))^2 \approx \sigma^2 S^2(t) dt$, so that (9.24) reduces to

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \frac{\partial \overline{U}}{\partial t} dt - \frac{\partial \overline{U}}{\partial t} dt,$$

which will be easier to interpret if we write it in the form

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t) - \sigma^{-2}\frac{\partial \overline{U}}{\partial t}(-\sigma^2 dt) - \sigma^{-2}\frac{\partial \overline{U}}{\partial t}(\sigma^2 dt).$$

The capital gain on each share of $\mathcal{D}$, $-\sigma^2 dt$, is cancelled by the dividend, $\sigma^2 dt$. So there is no point in holding $-\sigma^{-2}\partial\overline{U}/\partial t$ shares, or any number of shares, of $\mathcal{D}$. The increment in the capital process is simply

$$d\overline{U}(S(t),t) \approx \frac{\partial \overline{U}}{\partial s} dS(t),$$

which we achieve just by holding $\partial\overline{U}/\partial s$ shares of $\mathcal{S}$.

This way of organizing the Black-Scholes argument points the way to the elimination of the stochastic assumption. We can do without the assumption if the market really does price a security $\mathcal{D}$ whose dividend accounts for the $(dS(t))^2$ term in the Taylor's series.

The Purely Game-Theoretic Derivation

Assume now that between 0 and T, Investor trades in two securities: (1) a security $\mathcal{S}$ that pays no dividends and (2) a security $\mathcal{D}$, each share of which periodically pays the dividend $(dS(t)/S(t))^2$. This produces the following protocol:

THE NEW BLACK-SCHOLES PROTOCOL
Parameters: $T > 0$ and $N \in \mathbb{N}$; $dt := T/N$
Players: Investor, Market
Protocol:
  Market announces $S(0) > 0$ and $D(0) > 0$.
  $\mathcal{I}(0) := 0$.
  FOR $t = 0, dt, 2dt, \ldots, T - dt$:
    Investor announces $\delta(t) \in \mathbb{R}$ and $\lambda(t) \in \mathbb{R}$.
    Market announces $dS(t) \in \mathbb{R}$ and $dD(t) \in \mathbb{R}$.
    $S(t + dt) := S(t) + dS(t)$.
    $D(t + dt) := D(t) + dD(t)$.
    $\mathcal{I}(t + dt) := \mathcal{I}(t) + \delta(t) dS(t) + \lambda(t) \left(dD(t) + (dS(t)/S(t))^2 \right)$. (9.25)
Additional Constraints on Market: (1) $D(t) > 0$ for $0 < t < T$ and $D(T) = 0$, (2) $S(t) \geq 0$ for all t, and (3) the wildness of Market's moves is constrained.

Once D pays its last dividend, at time T, it is worthless: $D(T) = 0$. So Market is constrained to make his $dD(t)$ add to $-D(0)$. We also assume, as we did in the previous section, that the interest rate is zero. We do not spell out the constraints on the wildness of Market's moves, which will take different forms in the different versions of game-theoretic Black-Scholes pricing that we will study in the next two chapters. Here we simply assume that these constraints are sufficient to justify our usual approximation by a Taylor's series.

Consider a European option $\mathcal{U}$ with maturity date T and payoff function U. We begin by optimistically assuming that the price of $\mathcal{U}$ before T is given in terms of the current prices of $\mathcal{D}$ and $\mathcal{S}$ by

$$\mathcal{U}(t) = \overline{U}(S(t), D(t)),$$

where the function $\overline{U}(s, D)$ satisfies the initial condition

$$\overline{U}(s, 0) = U(s). \tag{9.26}$$

We approximate the increment in $\mathcal{U}$'s price from t to $t + dt$ by a Taylor's series:

$$d\overline{U}(S(t), D(t)) \approx \frac{\partial \overline{U}}{\partial s} dS(t) + \frac{\partial \overline{U}}{\partial D} dD(t) + \frac{1}{2} \frac{\partial^2 \overline{U}}{\partial s^2} (dS(t))^2. \tag{9.27}$$

We assume that the rules of the game constrain Market's moves $dS(t)$ and $dD(t)$ so that higher order terms in the Taylor expansion are negligible.

Comparing Equations (9.25) and (9.27), we see that we need

$$\delta(t) = \frac{\partial \overline{U}}{\partial s}, \quad \lambda(t) = \frac{\partial \overline{U}}{\partial D}, \quad \text{and} \quad \frac{\lambda(t)}{S^2(t)} = \frac{1}{2} \frac{\partial^2 \overline{U}}{\partial s^2}.$$

The two equations involving $\lambda(t)$ require that the function $\overline{U}$ satisfy the partial differential equation

$$-\frac{\partial \overline{U}}{\partial D} + \frac{1}{2} s^2 \frac{\partial^2 \overline{U}}{\partial s^2} = 0$$

for all s and all $D > 0$. This is the Black-Scholes equation, adapted to the market in which both $\mathcal{S}$ and $\mathcal{D}$ are traded. Its solution, with the initial condition (9.26), is

$$\overline{U}(s, D) = \int_{-\infty}^{\infty} U\left(se^z\right) \mathcal{N}_{-D/2,D}(dz).$$

This is the Black-Scholes formula for this market.

To summarize, the price for the European option $\mathcal{U}$ in a market where both the underlying security $\mathcal{S}$ and a volatility security $\mathcal{D}$ with dividend $(dS(t)/S(t))^2$ are traded is

$$\mathcal{U}(t) = \int_{-\infty}^{\infty} U\left(S(t)e^z\right) \mathcal{N}_{-D(t)/2,D(t)}(dz). \tag{9.28}$$

To hedge this price, we hold a continuously changing portfolio, containing

$$\frac{\partial \overline{U}}{\partial s}(S(t), D(t)) \text{ shares of } \mathcal{S}$$

and

$$\frac{\partial \overline{U}}{\partial D}(S(t), D(t)) \text{ shares of } \mathcal{D}$$

at time t.

By the argument of the preceding subsection, the derivative $\mathcal{D}$ is redundant if $S(t)$ follows a geometric Brownian motion. In this case, $\mathcal{D}$'s dividends are independent nonnegative random variables with expected value $\sigma^2 dt + (\mu dt)^2 \approx \sigma^2 dt$. By the law of large numbers, the remaining dividends at time t will add to almost exactly $\sigma^2(T - t)$, and hence this will be the market price, known in advance.

In the following chapters, we refine and elaborate the derivation of (9.28) in various ways. In Chapter 10, we derive (9.28) as an approximate price in a discrete-time game in which Market is constrained to keep $\mathbf{var}_S(2+\epsilon)$ and $\mathbf{var}_D(2-\epsilon)$ small. In Chapter 11, we derive it as an exact price in a continuous-time game in which Market is constrained to make $\operatorname{vex} D < 2$; it turns out that in this game Market is further obliged to make $\operatorname{vex} S = 2$ in order to avoid allowing Investor to become infinitely rich. As we have already noted, this gives some insight into why stock-market prices resemble diffusion processes as much as they do: the game itself pushes them in this direction. In Chapter 12, we extend the argument to the case of a known interest rate and show that we can replace the dividend-paying security $\mathcal{D}$ with a derivative that pays at time T a strictly convex function of $S(T)$.

Other Choices for the Dividend-Paying Security

The core idea of the preceding argument is to have the market price by supply and demand a derivative security $\mathcal{D}$ that pays a continuous dividend locally proportional to $\mathcal{S}$'s incremental variance, $(dS(t))^2$. We chose for $\mathcal{D}$'s dividend to be $(dS(t))^2/S^2(t)$, but this is not the only possible choice. If we take $\mathcal{D}$'s dividend to be

$$\frac{(dS(t))^2}{g(S(t))}, \tag{9.29}$$

then we obtain the partial differential equation

$$-\frac{\partial \overline{U}}{\partial D} + \frac{1}{2} g(s) \frac{\partial^2 \overline{U}}{\partial s^2} = 0$$

for the price $\overline{U}$, and there are many different functions $g(s)$ for which this equation has solutions. The choice $g(s) := s^2$, which we have just studied and will study further in the next two chapters, leads to the Black-Scholes formula. The choice $g(s) := 1$, which we will also study in the next two chapters, leads to Bachelier's formula. Bachelier's formula makes sense only if $S(t)$ can be negative, which is impossible for the price of stock in a limited-liability corporation. Powers of s intermediate between 0 and 2 (as in the Constant Elasticity of Variance model of Cox and Ross 1976) also have this defect, but there are many choices for $g(s)$ that avoid it; it is sufficient that $g(s)$ go to 0 fast enough as s goes to 0 ([107], p. 294).

In general, the game in which Investor can buy a derivative that pays the dividend (9.29) has as its stochastic counterpart the diffusion model

$$dS(t) = \mu(S(t), t)dt + \sqrt{g(S(t))}dW(t). \tag{9.30}$$

As we explain in an appendix (p. 231), the stochastic theory of risk-neutral valuation, which generalizes the Black-Scholes theory, tells us that if $S(t)$ follows the diffusion model (9.30), then all well-behaved derivatives are exactly priced, and the dividend (9.29), will total to exactly $T - t$ over the period from t to T. So this diffusion model makes it redundant to trade in a derivative that pays the dividend (9.29), just as geometric Brownian motion makes it redundant to trade in a derivative that pays the dividend $(dS(t))^2/S^2(t)$.

Since the Taylor's series on which our reasoning is based is only an approximation, the local proportionality of the dividend to $(dS(t))^2$ does not need to be exact, and this suggests another possibility: $(dS(t))^2$ might be smoothed to limit the dependence on extreme values and the susceptibility of the market to manipulation by major investors. But as a practical matter, it might be more promising to take the more conventional approach we have already discussed: we ask the market to price a derivative that pays a strictly convex function of $S(T)$ at T, and we calculate from its price an implied price for our theoretical derivative $\mathcal{D}$ (§12.2).

9.4 INFORMATIONAL EFFICIENCY

The hypothesis that capital markets are informationally efficient emerged from efforts in the 1960s to give an economic explanation for the apparent randomness of prices in markets for stocks and markets for commodity futures, and it is formulated in the context of a stochastic assumption. According to this stochastic assumption, each price p in such a market is based on a probability distribution for the ultimate value x of the contract being priced—the discounted value of the future stream of dividends in the case of a stock, the value at delivery of the commodity in the case of a futures contract. Neglecting interest, transactions costs, and so on, the assumption is that p is the expected value of x conditional on certain current information. What information? Different answers are possible. The hypothesis of informational efficiency says that p is the expected value of x conditional on *all* information available to investors, including all the information in past prices, so that an investor cannot expect, on average, to profit from buying at current prices and selling later.

Our rejection of stochasticity obviously undercuts this whole discussion. If there is no probability distribution for x, then there is no point to arguing about how the market uses such a probability distribution. But as we pointed out in §1.1, our game-theoretic framework permits a much simpler interpretation of the hypothesis of informational efficiency: it is simply the hypothesis of the impossibility of a gambling strategy in a game where the imaginary player Skeptic is allowed to buy and sell securities at current prices. It says that Skeptic does not, in this setting, have a strategy that will make him very rich without risk of bankruptcy. No assumptions of stochasticity are made, and yet there are many ways of testing the hypothesis: any strategy that does not risk bankruptcy can be the basis for such a test. As we will see in Chapter 15, under certain conditions there are strategies that allow Skeptic to become rich without risk of bankruptcy if returns do not average to zero in the long-run. So tests of the stochastic hypothesis of market efficiency that check whether returns do approximately average to zero can be made into tests of our hypothesis of market efficiency.

In addition to allowing us to test market efficiency, this understanding of the market efficiency also opens the possibility of using game-theoretic probability in various contexts where established finance theory uses stochastic ideas. We explore a couple of examples in Chapter 15: the trade-off between risk and return, and the measurement of value at risk.

Why Should Prices Be Stochastic?

Why should prices in markets for stocks and commodity futures be stochastic? In 1972, Paul Samuelson summarized the answer that comes to an economist's mind as follows ([264], p. 17):

> Expected future price must be closely equal to present price, or else present price will be different from what it is. If there were a bargain, which all could recognize, that fact would be "discounted" in advance and acted upon, thereby raising or lowering present price until the expected discrepancy with the future price were sensibly zero. It is true that people in the marketplace differ in their guesses about the future: and that is a principal reason why there are transactions in which one man is buying and another is selling. But at all times there is said to be as many bulls as bears, and in some versions there is held to be a wisdom in the resultant of the mob that transcends any of its members and perhaps transcends that of any outside jury of scientific observers. The opinions of those who make up the whole market are not given equal weights: those who are richer, more confident, perhaps more volatile, command greater voting power; but since better-informed, more-perceptive speculators tend to be more successful, and since the unsuccessful tend both to lose their wealth and voting potential and also to lose their interest and participation, the verdict of the marketplace as recorded in the record of auction prices is alleged to be as accurate ex ante and ex post as one can hope for and may perhaps be regarded as more accurate and trustworthy than would be the opinions formed by governmental planning agencies.

Samuelson did not represent this argument as his own opinion, and his tone suggests some misgivings. But he did represent it as "a faithful reproduction of similar ideas to be found repeatedly in the literature of economics and of practical finance". He cited a collection of articles edited by Cootner [56], which included a translation of Louis Bachelier's dissertation.

As Samuelson had observed in 1965, the assumption that the current price of a stock (or a futures contract) is the expected value of its price at some future time has a simple consequence: the successive prices of the stock will form a martingale [263]. This means that if p_t is the price of a stock at time t, then

$$p_t = \mathbb{E}_t(p_{t+1}),$$

or

$$\mathbb{E}_t(p_{t+1} - p_t) = 0, \tag{9.31}$$

where $\mathbb{E}_t$ represents the expected value conditional on information available at time t. Before Samuelson's observation, economists had been investigating the hypothesis that prices follow a random walk—that is, have statistically independent increments [116]. The increments of an arbitrary martingale have only the weaker property (9.31); each has expected value zero just before it is determined. Subsequent to Samuelson's calling attention to the martingale property, economists shifted from testing for a random walk to testing (9.31), and they began saying that they are testing market efficiency.

Eugene Fama (born 1939) in 1999. His work on efficient markets has helped make him the most frequently cited professor of finance.

Tests of the stochastic efficiency of markets have spawned an immense literature, chronicled in successive reviews by Eugene Fama [117, 118, 119]. Many authors contend that the empirical results in this literature confirm that financial markets generally are efficient; as Fama put it in 1998, "the expected value of abnormal returns is zero, but chance generates deviations from zero (anomalies) in both directions" ([119], p. 284). Other authors see deviations from efficiency everywhere [288] and conclude that stock-market prices are the result of "indifferent thinking by millions of people" ([286], p. 203) that can hardly identify correct probabilities for what will happen in the future. Yet other authors have suggested that the financial markets can be considered efficient even if they do not conform exactly to a stochastic model or eliminate entirely the possibility for abnormal returns [48, 140, 207, 208].

The diversity of interpretation of the empirical results can be explained in part by the fact, acknowledged by everyone in the debate, that the efficient-markets hypothesis cannot really be tested by itself. By itself, it says only that prices are expected values with respect to some stochastic model. An effective test requires that we specify the stochastic model, substantially if not completely, and then we will be testing not merely the efficient-markets hypothesis but also specific model. This is the *joint hypothesis problem* ([48], p. 24; [118], pp. 1575–1576).

Game-Theoretic Efficiency

Our game-theoretic efficient-market hypothesis is in the spirit of Samuelson's argument but draws a weaker conclusion. We do not suppose that there is some mysteriously correct probability distribution for future prices, and therefore we reject the words with which Samuelson's argument begins: "expected future price". But we accept the notion that an efficient market is one in which bargains have already been discounted in advance and acted upon. We hypothesize that our Skeptic cannot become rich without risking bankruptcy because any bargains providing Skeptic security against large loss would have already been snapped up, so much so that prices would have adjusted to eliminate them. By the principles of §1.3 and §8.3, this is enough to determine game-theoretic upper and lower probabilities for other events in the market being considered.

The purely game-theoretic approach obviously avoids the joint-hypothesis problem. We do not assume a stochastic model, and so we do not need to specify one in order to test our efficient-market hypothesis. We must specify, however, just what market we are talking about. Are we asserting that Skeptic cannot get rich without risking bankruptcy by trading in stocks on the New York Stock Exchange? By trading in options on the Chicago Board Options Exchange? Or by trading just in stocks in the S&P 500 index? These are all well-defined markets, and the hypothesis that Skeptic cannot get rich is a different hypothesis for each one of them, requiring different tests and perhaps leading to different practical conclusions. Ours is an efficient-market hypothesis, not an efficient-markets hypothesis.

We must also specify a unit of measurement for Skeptic's gains—a *numéraire*. We may hypothesize that Skeptic cannot get rich relative to the total value of the market (if this is well-defined for the particular market we are considering). Or we may hypothesize that he cannot get rich in terms of some monetary unit, such as the dollar or the yen. Or we may hypothesize that he cannot get rich relative to the value of a risk-free bond. And so on. These are all different hypotheses, subject to different tests and possibly having different implications concerning what we should expect in the future.

References

1. Peter A. Abken and Saikat Nandi. Options and volatility. *Economic Review*, 81:21–35, December 1996. Federal Reserve Bank of Atlanta. A survey of competing proposals, deterministic and stochastic, for improving the performance of the Black-Scholes model by modeling changes in volatility.

3. William J. Adams. *The Life and Times of the Central Limit Theorem*. Kaedmon, New York, 1974. A pleasantly written history of the central limit theorem from De Moivre to Lyapunov.

9. Louis Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure, 3*^{e} *série*, 17:21–86, 1900. This was Bachelier's doctoral dissertation. It was reprinted in facsimile in 1995 by Éditions Jacques Gabay, Paris. An English translation, by A. James Boness, appears on pp. 17–78 of [56].

13. Jan Beran. *Statistics for Long-Memory Processes*. Chapman and Hall, New York, 1994.

16. Peter L. Bernstein. *Capital Ideas: The Improbable Origins of Modern Wall Street*. Free Press, New York, 1992. A lively account of the development and intertwining of financial practice and theory, beginning with the stories of Louis Bachelier in Paris and Charles Dow and Eddie Jones on Wall Street a century ago and concluding with the adventures of the Wall Street professors of the 1970s and 1980s. Bernstein, a financial analyst and founding editor of the *Journal of Portfolio Management*, is personally acquainted with many of the individuals who play starring roles in his story.

26. Nicholas H. Bingham and Rüdiger Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer, London, 1998. A textbook on stochastic processes and option pricing. The authors present both topics first for discrete and then for continuous time.

27. Fischer Black. How to use the holes in Black-Scholes. *Journal of Applied Corporate Finance*, 1(4):67–73, Winter 1989. Reprinted in [174].

28. Fischer Black. How we came up with the option formula. *Journal of Portfolio Management*, 15:4–8, Winter 1989.

30. Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973. This is the article in which the Black-Scholes equation and the Black-Scholes formula were first derived. It is concerned as much with the pricing of warrants and other corporate liabilities as with options and derivatives.

45. Stephen G. Brush. A history of random sequences. I: Brownian movement from Brown to Perrin. *Archive for History of Exact Sciences*, 5:1–36, 1968.

47. Laurent Calvet, Adlai Fisher, and Benoit B. Mandelbrot. Large deviations and the distribution of price changes. Cowles Foundation Discussion Paper No. 1166, September 1997. Yale University. Paper available from the SSRN database `http://www.ssrn.com`.

48. John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ, 1997.

52. Neil Chriss and William Morokoff. Volatility and variance swaps. *Risk*, 12, October 1999.

56. Paul H. Cootner, editor. *The Random Character of Stock Market Prices*. MIT Press, Cambridge, MA, 1964.

60. Jean-Michel Courtault et al. Louis Bachelier: On the centenary of *Théorie de la Spéculation*. *Mathematical Finance*, 10:339–353, 2000.

66. John C. Cox and Stephen A. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166, 1976.

68. John Crank. *The Mathematics of Diffusion*. Oxford University Press, Oxford, second edition, 1975.

75. Robert C. Dalang, Andrew Morton, and Walter Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market model. *Stochastics and Stochastics Reports*, 29:185–201, 1990.

102. Joseph L. Doob. Wiener's work in probability theory. *Bulletin of the American Mathematical Society*, 72(1, Part II):69–72, 1966. This is a special obituary issue for Wiener, and its pages are numbered separately from the rest of Volume 72.

107. Darrell Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, NJ, second edition, 1996. In addition to a relatively rigorous treatment of continuous-time arbitrage pricing, this book includes discussions of equilibrium pricing and optimal portfolio selection. The first edition appeared in 1992.

109. Albert Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 17:549–560, 1905.

110. Albert Einstein. *Investigations on the Theory of the Brownian Movement*. Dover, New York, 1956. First published by Dutton, New York, in 1926, this is a translation, by A. D. Cowper, of *Untersuchungen über die Theorie der Brownschen Bewegung*. It includes [109] together with Einstein's other papers on Brownian motion.

113. Robert J. Elliott and P. Ekkehard Kopp. *Mathematics of Financial Markets*. Springer, New York, 1999.

114. Edwin J. Elton and Martin J. Gruber, editors. *Security Evaluation and Portfolio Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1972.

116. Eugene F. Fama. The behavior of stock-market prices. *Journal of Business*, 38:34–105, 1965.

117. Eugene F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

118. Eugene F. Fama. Efficient capital markets, II. *Journal of Finance*, 46:1575–1617, 1991.

119. Eugene F. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49:283–306, 1998.

124. Adlai Fisher, Laurent Calvet, and Benoit B. Mandelbrot. Multifractality of Deutschmark / US Dollar exchange rates. Cowles Foundation Discussion Paper No. 1166, September 1997. Yale University. Available from the SSRN database `http://www.ssrn.com`. Among other things, the paper studies the variation spectrum for the DM/USD market.

131. Avner Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, NJ, 1964.

140. Sanford Grossman and Joseph Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70:393–408, 1980.

147. J. Michael Harrison and Stanley R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981. This article is credited with the first rigorous and general proof that the absence of arbitrage and the ability to hedge depend only on the existence and uniqueness of an equivalent martingale measure.

154. John C. Hull. *Options, Futures, and Other Derivatives*. Prentice-Hall, Upper Saddle River, NJ, fourth edition, 2000. A popular introduction to the study of derivatives at the MBA level. It discusses both the theory and practice of derivatives and includes numerical procedures for implementing the most common pricing algorithms.

156. Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy of Tokyo*, 20:519–524, 1944. Reprinted in [157], pp. 85–90.

157. Kiyosi Itô. *Selected Papers*. Springer, New York, 1987. Edited by Daniel W. Stroock and S. R. Sriniva Varadhan.

174. Robert W. Kolb, editor. *Practical Readings in Financial Derivatives*. Blackwell, Malden, MA, 1998.

184. David M. Kreps. Multiperiod securities and the efficient allocation of risk: A comment on the Black-Scholes option pricing model. In John J. McCall, editor, *The Economics of Information and Uncertainty*, pp. 203–232. Chicago University Press, Chicago, 1982.

207. Andrew W. Lo, editor. *Market Efficiency: Stock Market Behaviour in Theory and Practice*. Elgar, Cheltenham, England, 1997. Two volumes. A comprehensive and extremely expensive collection of articles on the efficient market hypothesis over three decades.

208. Andrew W. Lo and A. Craig MacKinlay, editors. *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton, NJ, 1999.

212. Burton G. Malkiel. *A Random Walk Down Wall Street*. Norton, New York, 1996.

214. Benoit B. Mandelbrot. Une classe de processus stochastiques homothétiques à soi; application á la loi climatologique de H. E. Hurst. *Comptes rendus des Séances de l'Académie des Sciences*, 260:3274–3277, 1965.

215. Benoit B. Mandelbrot. Self-affine fractal sets. In L. Pietronero and E. Tosatti, editors, *Fractals in Physics*, pp. 3–28. North-Holland, Amsterdam, 1986.

216. Benoit B. Mandelbrot. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*. Springer, New York, 1997. This volume includes contributions Mandelbrot made to finance theory in the 1960s, with retrospective comments, together with reports on new developments, including work with Fisher and Calvet on the multifractal model of asset returns; see also [47], [124], and [217].

217. Benoit B. Mandelbrot, Adlai Fisher, and Laurent Calvet. A Multifractal Model of Asset Returns. Cowles Foundation Discussion Paper No. 1164, September 1997. Yale University. Available from the SSRN database `http://www.ssrn.com`.

225. Nina Mehta. Equity vol swaps grow up. *Derivatives Strategy*, 4:11–12, July 1999.

229. Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.

237. M. F. Maury Osborne. Brownian motion in the stock market. *Operations Research*, 7:145–173, 1959.

262. L. Chris G. Rogers and David Williams. *Diffusions, Markov Processes, and Martingales*, volume 1: Foundations. Wiley, Chichester, second edition, 1994. Reissued by Cambridge University Press in *Cambridge Mathematical Library*, 2000.

263. Paul A. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6:41–50, Spring 1965.

264. Paul A. Samuelson. Mathematics of speculative price. *SIAM Review*, 15:1–42, 1973.

286. Robert J. Shiller. *Irrational Exuberance*. Princeton University Press, Princeton, NJ, 2000.

288. Andrei Shleifer. *Inefficient Markets*. Oxford University Press, Oxford, 2000.

291. Marian von Smoluchowski. Zarys kinetycznej teoriji ruchów browna i roztworów metnych. *Rozprawy i Sprawozdania z Posiedzeń Wydzialu Matematyczno-Przyrodniczego Akademii Umiejetności*, 3:257–282, 1906. A German translation appeared in the *Annalen der Physik*, 21, pp. 756–780, in 1906.

318. Jan von Plato. *Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective*. Cambridge University Press, Cambridge, 1994. A study of the emergence of measure-theoretic probability in the first third of the 20th century. It emphasizes statistical physics and includes a detailed look at the foundational work of Kolmogorov and de Finetti.

337. Christian Walter. Une histoire du concept d'efficience sur les marchés financiers. *Annales: Histoire, Sciences Sociales*, number 4:873–905, July–August 1996.

339. Robert E. Whaley. Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives*, 1:71–84, 1993.

340. Robert E. Whaley. The investor fear gauge. *Journal of Derivatives*, 26:12–17, 2000.

341. Norbert Wiener. The mean of a functional of arbitrary elements. *Annals of Mathematics*, 22:66, 1920.

342. Norbert Wiener. The average of an analytical functional. *Proceedings of the National Academy of Science*, 7:253, 1921.

343. Norbert Wiener. The average of an analytical functional and the Brownian movement. *Proceedings of the National Academy of Science*, 7:294, 1921.

344. Norbert Wiener. Differential space. *Journal of Mathematics and Physics*, 2:131, 1923.

345. Norbert Wiener. The average value of a functional. *Proceedings of the London Mathematical Society*, 22:454, 1924.

346. Norbert Wiener. *Collected Works*. MIT Press, Cambridge, MA, 1976–1985. Four volumes. Edited by P. Pisani. Volume 1 includes Wiener's original papers on Brownian motion [341, 342, 343, 344, 345], with a commentary by Kiyosi Itô.

351. Paul Wilmott. *Derivatives: The Theory and Practice of Financial Engineering*. Wiley, Chichester, 1998. An engaging and leisurely (739 pages) introduction to the pricing of derivatives.

352. Paul Wilmott, Jeff Dewynne, and Sam Howison. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, Oxford, 1993. A readable book aimed at the applied mathematician.