6 Brownian Motion

6.1. Basic Definitions

Brownian motion is the most important example of a process with continuous time and continuous paths. Mathematically, it lies in the intersection of three classes of stochastic processes: it is a Gaussian process, a Markov process with continuous paths, and has independent increments. Because of this, it is possible to use techniques from all three fields, so there is a rich and detailed theory of its properties. From a practical view, Brownian motion is often used as a component of models of physical, biological, and economic phenomena.

(1.1) **Definition.** $B(t)$ is a (one-dimensional) Brownian motion with variance σ^2 if it satisfies the following conditions:

(a) $B(0) = 0$, a convenient normalization.

(b) *Independent increments.* Whenever $0 = t_0 < t_1 < \ldots < t_n$

$$B(t_1) - B(t_0), \ldots, B(t_n) - B(t_{n-1}) \quad \text{are independent.}$$

(c) *Stationary increments.* The distribution of $B_t - B_s$ only depends on $t - s$.

(d) $B(t)$ is normal$(0, \sigma^2 t)$.

(e) $t \to B_t$ is continuous.

To explain the definition we begin by discussing the situation that gives the process its name. Historically, Brownian motion originates from the fact observed by Robert Brown (and others) in the 1800s that pollen grains under a microscope perform a "continuous swarming motion." To explain the source of this motion, we imagine a spherical pollen grain that collides with water molecules at times of a Poisson process with a large rate λ. The displacement of the pollen grain between time 0 and time 1 is the sum of a large number of

small independent random effects. To see that this implies the change in the x coordinate, $X(1) - X(0)$, will be approximately normal$(0, \sigma^2)$, we recall:

(1.2) **Central Limit Theorem.** *Let $X_1, X_2, \ldots$ be i.i.d. with $EX_i = 0$ and $\text{var}(X_i) = \sigma^2$, and let $S_n = X_1 + \cdots + X_n$. Then for all x we have*

$$P\left(\frac{S_n}{\sigma\sqrt{n}} \le x\right) \to P(\chi \le x)$$

where χ has a standard normal distribution. That is,

$$P(\chi \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy$$

Formally, (1.2) says that $S_n/\sigma\sqrt{n}$ **converges in distribution** to χ, a conclusion we will write as $S_n/\sigma\sqrt{n} \Rightarrow \chi$.

Since there are a huge number of water molecules in the system ($> 10^{20}$), it is reasonable to think that after a collision the molecule disappears into the crowd never to be seen again, and with a small leap of faith we can assert that $X(t)$ will have independent increments. The movement of the pollen grain is clearly continuous in space, so if we impose the condition $X(0) = 0$ by measuring displacements from the initial position, then the first coordinate of the pollen grain will be a Brownian motion with variance σ^2.

Does Brownian motion exist? Is there a process with the properties given in (1.1)? The answer is yes, of course, or this chapter wouldn't exist. The technical problem is to define the uncountably many random variables B_t for each nonnegative real number t on the same space so that $t \to B_t$ is continuous with probability one. One solution to this problem is provided in Exercise 6.11. However, the details of the proof are not important for understanding how B_t behaves, so we suggest that the reader take the existence of Brownian motion on faith.

Multidimensional Brownian motion. If we are looking at the pollen grain through a microscope then we can observe two coordinates of its motion $(X^1(t), X^2(t))$. Again we can argue that the displacement of the pollen grain between time 0 and time t will be the sum of a large number of small independent random effects, so if we suppose that $(X^1(0), X^2(0)) = (0, 0)$, then the position at time t, $(X^1(t), X^2(t))$, will have a two-dimensional normal distribution.

Multidimensional normal distributions $(Z^1, \ldots, Z^n)$ are characterized by giving the vector of means $\mu_i = EZ^i$ and the covariance matrix:

$$\Gamma_{ij} = E\{(Z^i - \mu_i)(Z^j - \mu_j)\}$$

In the case under consideration $\mu_1 = \mu_2 = 0$. The spherical symmetry of the pollen grain and of the physical system implies that

$$\Gamma_{ij} = \sigma^2 I_{ij}$$

where I is the identity matrix: $I_{ii} = 1$ and $I_{ij} = 0$ for $i \neq j$. Recalling that:

(1.3) **Lemma.** *Two components Z_i and Z_j of a multidimensional normal are independent if and only if their covariance $E(Z_iZ_j) - EZ_iEZ_j = 0$*

we see the two components of the movement of the pollen grain are independent. Generalizing (and specializing), we state:

(1.4) **Definition.** $(B^1(t), \dots B^n(t))$, $t \geq 0$ *is a standard n-dimensional Brownian motion if each of its components are independent one-dimensional Brownian motions with* $\sigma^2 = 1$.

We have removed the parameter σ^2 since it is trivial to reintroduce it. If B_t is a standard Brownian motion, then σB_t gives a Brownian motion with variance σ^2. A slight variation on the last calculation gives the following useful result:

(1.5) **Scaling relation.** *The processes* $\{B_{ct}, t \geq 0\}$ *and* $\{c^{1/2}B_t, t \geq 0\}$ *have the same distribution.*

Proof. Speeding up time by a factor of c does not change properties (a), (b), (c), or (e), but in (d) it makes the variance $c(t-s)$. By the previous observation this can be achieved by multiplying a standard Brownian motion by $c^{1/2}$. □

A second basic property of Brownian motion is:

(1.6) **Covariance formula.** $E(B_sB_t) = s \wedge t$.

Proof. If $s = t$, this says $EB_s^2 = s$, which is correct since B_s is normal$(0, s)$, so suppose now that $s < t$. Writing $B_t = B_s + (B_t - B_s)$ and using the fact that B_s and $B_t - B_s$ are independent with mean 0, we have

$$E(B_sB_t) = E(B_s^2) + E(B_s(B_t - B_s)) = s + 0 = s = s \wedge t \qquad \square$$

It follows easily from the definition given in (1.1) that Brownian motion is a **Gaussian process**. That is, for any $t_1 < t_2 \dots < t_n$ the vector

$$(B(t_1), B(t_2), \dots B(t_n))$$

has a multivariate normal distribution. Since multivariate normal distributions are characterized by giving their means and covariances, it follows that Brownian motion is the only Gaussian process with continuous paths that has $EB_s = 0$ and $E(B_sB_t) = s \wedge t$.

We will not use the Gaussian viewpoint here, except in a couple of the exercises. However, to complete the story we should mention two important Gaussian relatives of Brownian motion:

Example 1.1. Brownian bridge. Let $B_t^0 = B_t - tB_1$. By definition $B_1^0 = 0$. The basic fact about this process is the following:

(1.7) **Theorem.** $\{B_t^0, 0 \leq t \leq 1\}$ *and* $(\{B_t, 0 \leq t \leq 1\} | B_1 = 0)$ *have the same distribution. Each is a Gaussian process with mean 0 and covariance* $s(1-t)$ *if* $s \leq t$.

Why is this true? Since each of the processes is Gaussian and has mean 0, it suffices to show that their covariances are equal to $s(1-t)$. If $s < t$,

$$\begin{aligned} E(B_s^0B_t^0) &= E((B_s - sB_1)(B_t - tB_1)) = EB_sB_t - sEB_1B_t - tEB_sB_1 + stB_1^2 \\ &= s - st - st + st = s(1-t) \end{aligned}$$

With considerably more effort (see Exercise 6.2) one can compute the covariance function for $(B_t|B_1 = 0)$ and show that it is also $s(1-t)$, so the two processes have the same distribution. □

Example 1.2. Ornstein–Uhlenbeck process. Let

$$Z(t) = e^{-t}B(e^{2t}) \qquad \text{for } -\infty < t < \infty$$

The definition may look strange, but note the very special property that for each t, $Z(t)$ has a normal(0,1) distribution. To compute the joint distribution of $Z(t)$ and $Z(t+s)$ we note that

$$\begin{aligned} Z(s+t) &= e^{-(s+t)}B(e^{2(s+t)}) \\ &= e^{-(s+t)}B(e^{2t}) + e^{-(s+t)}(B(e^{2(s+t)}) - B(e^{2t})) \end{aligned}$$

The first term is equal to $e^{-s}Z(t)$. The second is independent of $Z(t)$ and has a normal distribution with variance $e^{-2(s+t)}(e^{2(s+t)} - e^{2t}) = 1 - e^{-2s}$. Thus introducing an independent standard normal χ we can write

$$(\star) \qquad Z(s+t) = e^{-s}Z(t) + \chi\sqrt{1 - e^{-2s}}$$

From this we can see that the effect of $Z(t)$ decays exponentially in time and that as $s \to \infty$, $Z(s+t)$ is almost an independent standard normal. Using $(\star)$ we can also compute the covariance for of the Ornstein–Uhlenbeck process: $E(Z(t)Z(t+s)) = e^{-s}EZ(t)^2 + 0 = e^{-s}$.

6.2. Markov Property, Reflection Principle

Intuitively, the fact that Brownian motion has stationary independent increments (i.e., (c) and (b) of the definition hold) should immediately imply:

(2.1) **Markov property.** *If $s > 0$, then $B_{s+t} - B_s$, $t \geq 0$ is a Brownian motion independent of B_r, $r \leq s$.*

In words, the future increments $B_{s+t} - B_s$, $t \geq 0$ are independent of "the behavior of the process before time s." The phrase in quotation marks is intuitive but a little slippery to articulate mathematically. To be precise we should say instead of (2.1) that

(2.1′) For any times $0 \leq r_1 \cdots < r_n = s$, $B_{s+t} - B_s$, $t \geq 0$ is a Brownian motion independent of the vector $(B(r_1), \dots, B(r_n))$.

However, we prefer the more informal phrase "independent of $\{B_r, r \leq s\}$" so we will say the short phrase and ask that the reader understand that what we mean precisely is the longer one given in (2.1′).

As in discrete time, the Markov property is also true at **stopping times** T which are defined by the requirement that the decision to stop before time s, $T \leq s$, can be determined from the values of B_r at times $r \leq s$.

(2.2) **Strong Markov property.** *If T is a stopping time, then $B_{T+t} - B_T$, $t \geq 0$ is a Brownian motion independent of B_r, $r \leq T$.*

The proof of (2.2), or even a precise definition of the notion of stopping time, requires considerable sophistication. Historically, it was not given until many years after Brownian motion was first used as a model. Here, we will follow early researchers and use (2.2) without bothering about its proof.

To illustrate the use of (2.2), let $T_a = \min\{t \geq 0 : B_t = a\}$ be the first time that Brownian motion hits a. Our first fact is:

(2.3) T_a, $a \geq 0$ **has stationary independent increments.** *That is, (i) if $a < b$, then the distribution of $T_b - T_a$ is the same as T_{b-a}. (ii) if $a_0 = 0 < a_1 < \cdots < a_n$, then $T_{a_1} - T_{a_0}, \dots, T_{a_n} - T_{a_{n-1}}$ are independent.*

Proof. Using (2.2) with $T = T_a$ we see that $B_{T+t} - B_T$ is a Brownian motion independent of B_r, $r \leq T$. Since $B_T = a$, we see that $T_b - T_a$ is just the hitting time of $b - a$ for $B_{T+t} - B_T$ and hence the distribution of $T_b - T_a$ is the same as T_{b-a}.

To prove (ii) now, we use (2.2) with $T = T_{a_{n-1}}$ to see that $B_{T+t} - B_T$ is a Brownian motion independent of B_r, $r \leq T$. Since $T_{a_n} - T_{a_{n-1}}$ is determined by $B_{T+t} - B_T$ for $t \geq 0$, while $(T_{a_1} - T_{a_0}, \dots, T_{a_{n-1}} - T_{a_{n-2}})$ is determined by B_r, $r \leq T$, we see that the last increment $T_{a_n} - T_{a_{n-1}}$ is independent of the previous $n-1$ increments. This result and an induction argument shows that the increments are independent. □

Distribution of T_a. Suppose $a > 0$. Since $P(B_t = a) = 0$ we have

$$(\alpha) \qquad P(T_a \leq t) = P(T_a \leq t, B_t > a) + P(T_a \leq t, B_t < a)$$

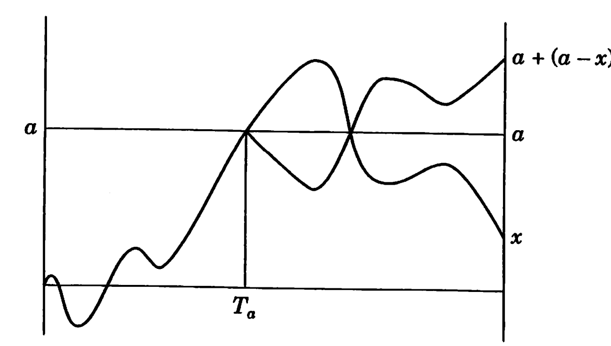
Now any path that starts at 0 and ends above $a > 0$ must cross a, so

$$(\beta) \qquad P(T_a \leq t, B_t > a) = P(B_t > a)$$

To handle the second term in (α), we note that (i) the strong Markov property implies that $B(T_a+t) - B(T_a) = B(T_a+t) - a$ is a Brownian motion independent of B_r, $r \leq T_a$, and (ii) the normal distribution is symmetric about 0, so

$$(\gamma) \qquad P(T_a \leq t, B_t < a) = P(T_a \leq t, B_t > a)$$

The last argument is sometimes called the **reflection principle** since it was originally obtained by arguing that a path that started at 0 and ended above a had the same probability as the one obtained by reflecting the path B_t, $t \geq T_a$ is an mirror located at level a.



Combining (α), (β), and (γ), we have

(2.4) $$P(T_a \le t) = 2P(B_t > a) = 2\int_a^\infty (2\pi t)^{-1/2}\exp(-x^2/2t)\,dx$$

To find the probability density of T_a, we change variables $x = t^{1/2}a/s^{1/2}$, with $dx = -t^{1/2}a/2s^{3/2}ds$ to get

$$\begin{aligned} P_0(T_a \le t) &= 2\int_t^0 (2\pi t)^{-1/2}\exp(-a^2/2s)\left(-t^{1/2}a/2s^{3/2}\right) ds \\ &= \int_0^t (2\pi s^3)^{-1/2} a \exp(-a^2/2s)\,ds \end{aligned}$$

Thus we have shown:

(2.5) *T_a has density function*

$$(2\pi s^3)^{-1/2} a \exp(-a^2/2s)$$

Using (2.5) and integrating, we can show:

(2.6) *The probability that B_t has no zero in the interval (s,t) is*

$$\frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}$$

Proof. If $B_s = a > 0$, then symmetry implies that the probability of at least one zero in (s,t) is $P(T_a \le t-s)$. Breaking things down according to the value of B_s, which has a normal$(0,s)$ distribution, and using the Markov property, it follows that the probability of interest is

$$2\int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-a^2/2s} \int_0^{t-s} (2\pi r^3)^{-1/2} a \exp(-a^2/2r)\,dr\,da$$

Interchanging the order of integration, we have

(*) $$\frac{1}{\pi\sqrt{s}}\int_0^{t-s} r^{-3/2}\int_0^\infty a\exp\left(-a^2/2\left[\frac{1}{s}+\frac{1}{r}\right]\right) da\,dr$$

The inside integral can be evaluated as

$$= \frac{sr}{r+s}\exp\left(-\frac{a^2 rs}{2(r+s)}\right)\Bigg|_0^\infty = \frac{sr}{r+s}$$

so the expression in (*) is equal to

$$\frac{\sqrt{s}}{\pi}\int_0^{t-s}\frac{1}{(s+r)\sqrt{r}}\,dr$$

Changing variables $r = sv^2$, $dr = 2sv\,dv$ produces

$$\frac{2}{\pi}\int_0^{\sqrt{(t-s)/s}}\frac{dv}{1+v^2} = \frac{2}{\pi}\arctan\sqrt{\frac{t-s}{s}}$$

By considering a right triangle with sides $\sqrt{s}$, $\sqrt{t-s}$ and $\sqrt{t}$ one can see that when the tangent is $\sqrt{(t-s)/s}$, the cosine is $\sqrt{s/t}$. At this point we have shown that the probability of at least one zero in (s,t) is $(2/\pi)\arccos(\sqrt{s/t})$. To get the result given, recall $\arcsin(x) = (\pi/2) - \arccos(x)$ when $0 \le x \le 1$. $\square$

Let $L_t = \max\{s < t : B_s = 0\}$. From (2.6) we see that

(2.7) $$P(L_t < s) = \frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}$$

Recalling that the derivative of $\arcsin(x)$ is $1/\sqrt{1-x^2}$, we see that L_t has density function

(2.8) $$P(L_t = s) = \frac{2}{\pi}\frac{1}{\sqrt{1-s/t}}\cdot\frac{1}{2\sqrt{st}} = \frac{1}{\pi\sqrt{s(t-s)}}$$

In view of the definition of L_t as the LAST ZERO before time t, it is somewhat surprising that the density function of L_t

(i) is symmetric about $t/2$, and

(ii) tends to ∞ as $s \to 0$.

In a different direction, we also find it remarkable that $P(L_t > 0) = 1$. To see why this is surprising note that if we change our previous definition of the hitting time of 0 to $T_0^+ = \min\{s > 0 : B_s = 0\}$ then this implies that for all $t > 0$ we have $P(T_0^+ \le t) = P(L_t > 0) = 1$, and hence

(2.9) $$P(T_0^+ = 0) = 1$$

Since T_0^+ is defined to be the minimum over all strictly positive times, (2.9) implies that there is a sequence of times $s_n \to 0$ at which $B(s_n) = 0$. At first this may sound innocent, but using this observation with the strong Markov property, we see that every time Brownian motion returns to 0 it will immediately have infinitely many 0's!

6.3. Martingales, Hitting Times

As in our earlier study of random walks in Section 2.4, martingales are useful for proving results about Brownian motion. The first and simplest is the:

a. Linear martingale.

(3.1) **Theorem.** *B_t is a martingale. That is, if $s < t$, then*

$$E(B_t|B_r, r \le s) = B_s$$

What does this formula mean? In words, it says that the conditional expectation of B_t given the entire history of the process before time s is equal to the value at time s. As was the case with the Markov property, it is difficult to define formally the conditional expectation given the entire past, so as in the discussion of the Markov property in Section 6.2, what this really means is that

(3.1′) For any times $0 \le r_1 \cdots < r_n < s$,

$$E(B_t|B_s, B_{r_n}, \dots, B_{r_1}) = B_s$$

Again, we prefer the more informal formulation of the conditional expectation as "given B_r, $r \le s$" rather than the technically correct "given $B_s, B_{r_n}, \dots, B_{r_1}$ for any times $0 \le r_1 \cdots < r_n < s$." Again, we will say the short phrase and ask the reader to remember that what we mean precisely is the long one.

Proof of (3.1). Given B_r, $r \le s$, the value of B_s is known while $B_t - B_s$ is independent and has mean 0.

$$\begin{aligned} E(B_t|B_r, r \le s) &= E(B_s + B_t - B_s|B_r, r \le s) \\ &= B_s + E(B_t - B_s|B_r, r \le s) = B_s \end{aligned} \qquad \square$$

As in Chapter 2, all of our computations will be based on the:

(3.2) **Stopping theorem for bounded martingales.** *Let M_t, $t \ge 0$ be a martingale with continuous paths. Suppose T is a stopping time with $P(T < \infty) = 1$, and there is a constant K so that $|M_{T\wedge t}| \le K$ for all t. Then*

$$EM_T = EM_0$$

As in Section 2.4, we can use the linear martingale to compute the:

Example 3.1. Exit distribution from an interval. Define the exit time by $\tau = \min\{t : B_t \notin (a,b)\}$.

$$P(B_\tau = b) = \frac{-a}{b-a} \quad \text{and} \quad P(B_\tau = a) = \frac{b}{b-a} \tag{3.3}$$

Proof. To check τ is a stopping time, note that

$$\{\tau \le s\}^c = \{\tau > s\} = \{B_r \in (a,b) \text{ for all } r \le s\}$$

which can be determined by looking at B_r for $r \le s$. The fact that $P(\tau < \infty) = 1$ can be shown by noting that $\tau \le T_a$ while $P(T_a < \infty) = 1$ by the formula for the density in (2.5). It is clear that $|B(\tau \wedge t)| \le |a| + b$, so using (3.2) now we see that

$$0 = EB_\tau = aP(B_\tau = a) + bP(B_\tau = b)$$

Since $P(B_\tau = a) = 1 - P(B_\tau = b)$, we have $0 = a + (b-a)P(B_\tau = b)$. Solving gives the first formula. The second follows from that fact that $P(B_\tau = a) = 1 - P(B_\tau = b)$. $\square$

Using the hitting times $T_x = \min\{t \ge 0 : B_t = x\}$, (3.3) can be written as

$$P(T_a < T_b) = \frac{-a}{b-a} \quad \text{and} \quad P(T_b < T_a) = \frac{b}{b-a}$$

Fixing the value of $a < 0$ and letting $b \to \infty$, it follows that $P(T_a < \infty) = 1$ for $a < 0$. Since this is clearly also true for $a > 0$ and trivial for $a = 0$, we have:

(3.4) **Theorem.** *For all a, $P(T_a < \infty) = 1$.*

So Brownian motion will certainly visit every point on the real line. A little thought reveals that it can only do this if and only if

(3.4′) *With probability one,*

$$\limsup_{t\to\infty} B_t = \infty \quad \text{and} \quad \liminf_{t\to\infty} B_t = -\infty$$

In words, B_t oscillates between ∞ and $-\infty$, and hence visits each value infinitely many times.

For a gambling application of (3.3) consider:

Example 3.2. Optimal doubling in Backgammon. In our idealization, B_t is the probability you will win given the events in the game up to time t. Thus, backgammon is a Brownian motion starting at 1/2 run until it hits 1 (win) or 0 (loss). Initially the "doubling cube" sits in the middle of the board and either player can "double" = tell the other player to play on for twice the stakes or give up and pay the current wager. If a player accepts the double (i.e., decides to play on), she gets possession of the doubling cube and is the only one who can offer the next double.

A doubling strategy is given by two numbers $b < 1/2 < a$, i.e., offer a double when $B_t \geq a$ and give up if the other player doubles and $B_t < b$. It is not hard to see that for the optimal strategy $b^* = 1 - a^*$ and that when $B_t = b^*$ accepting and giving up must have the same payoff. If you accept when your probability of winning is b^* then you lose 2 dollars when your probability hits 0 but you win 2 dollars when your probability of winning hits a^*, since at that moment you can double and the other player gets the same payoff if they give up or play on. If giving up or playing on at b^* is to have the same payoff we must have

$$-1 = \frac{b^*}{a^*} \cdot 2 + \frac{a^* - b^*}{a^*} \cdot (-2)$$

Writing $b^* = c$ and $a^* = 1 - c$ and solving, we have

$$-(1-c) = 2c - 2(1-2c) \qquad 1 = 5c$$

so $b^* = 1/5$ and $a^* = 4/5$. In words, you should offer a double when your chances of winning are at least 80% and refuse if your chances of winning are less than 20%. □

b. Quadratic martingale

(3.5) **Theorem.** *$B_t^2 - t$ is a martingale. That is, if $s < t$, then*

$$E(B_t^2 - t | B_r, r \leq s) = B_s^2 - s$$

Proof. Again, given B_r, $r \leq s$, the value of B_s is known, while $B_t - B_s$ is independent with mean 0 and variance $t - s$

$$\begin{aligned} E(B_t^2 | B_r, r \leq s) &= E(\{B_s + B_t - B_s\}^2 | B_r, r \leq s) \\ &= B_s^2 + 2B_s E(B_t - B_s | B_r, r \leq s) + E((B_t - B_s)^2 | B_r, r \leq s) \\ &= B_s^2 + 0 + t - s \end{aligned}$$

Subtracting t from each side gives the desired conclusion. □

(3.5) allows us to obtain information about the:

Example 3.3. Mean exit time from an interval. Define the exit time by $\tau = \min\{t : B_t \notin (a,b)\}$.

(3.6) $$E\tau = -ab$$

Why is this true? The martingale $M_t = B_t^2 - t$ does not have the property that $|M_{\tau\wedge t}| \leq K$. However, if we ignore that and apply the stopping theorem, (3.2), and the formula for the distribution of B_τ given in (3.3), then we get

$$0 = E(B_\tau^2 - \tau) = b^2 \cdot \frac{-a}{b-a} + a^2 \cdot \frac{b}{b-a} - E\tau$$

Rearranging now gives

$$E\tau = \frac{-b^2 a + ba^2}{b-a} = -ba \qquad \square$$

Proof. To justify the last computation we let n be a positive integer and consider the stopping time $T = \tau \wedge n$, which has $|M_{T\wedge t}| \leq a^2 + b^2 + n$ for all t. Using (3.2) now as before and rearranging, we have

$$E(\tau \wedge n) = EB_{\tau\wedge n}^2$$

Letting $n \to \infty$ now gives (3.6). □

To see what (3.6) says we will now derive two consequences. First and simplest is to set $a = -r$, $b = r$ to see that if τ_r is the exit time from $(-r, r)$, then

$$E\tau_r = r^2$$

To see why this should be true note that the scaling relation (1.5) implies that τ_r has the same distribution as $r^2\tau_1$.

Taking $b = N$ in (3.6) and letting $\tau_{a,N} = \min\{t : B_t \notin (a, N)\}$, we see that if $a < 0$, then

$$ET_a \geq E\tau_{a,N} = -aN \to \infty$$

as $N \to \infty$. The same conclusion holds for $a > 0$ by symmetry, so we have

(3.7) *For all $a \neq 0$, $ET_a = \infty$.*

This result does not hold for $a = 0$ since $T_0 = \min\{t \geq 0 : X_t = 0\} = 0$ by definition.

c. Exponential martingale

Before introducing our third and final martingale, we need some algebra,

$$-\frac{x^2}{2u} + \theta x = -\frac{(x-u\theta)^2}{2u} + \frac{u\theta^2}{2}$$

and some calculus.

$$\begin{aligned} E(\exp(\theta B_u)) &= \int e^{\theta x} \frac{1}{\sqrt{2\pi u}} e^{-x^2/2u}\,dx \\ &= e^{u\theta^2/2} \int \frac{1}{\sqrt{2\pi u}} \exp\left(-\frac{(x-u\theta)^2}{2u}\right) dx = e^{u\theta^2/2} \end{aligned}$$

since the integrand is the density of a normal$(u\theta, u)$ distribution.

(3.8) **Theorem.** $\exp(\theta B_t - t\theta^2/2)$ *is a martingale. That is, if* $s < t$, *then*

$$E(\exp(\theta B_t - t\theta^2/2)|B_r, r \le s) = \exp(\theta B_s - s\theta^2/2)$$

Proof. Again, given B_r, $r \le s$, the value of B_s is known, while $B_t - B_s$ is an independent normal$(0, t-s)$.

$$\begin{aligned} E(\exp(\theta B_t)|B_r, r \le s) &= \exp(\theta B_s) E(\exp(\theta(B_t - B_s))|B_r, r \le s) \\ &= \exp(\theta B_s + (t-s)\theta^2/2) \end{aligned}$$

by our calculus fact with $u = t - s$. Multiplying each side by $\exp(-t\theta^2/2)$ gives the desired result. □

Example 3.4. Probability of ruin when playing a favorable game. Let $X_t = \sigma B_t + \mu t$ be a Brownian motion with drift μ and variance σ^2. Let

$$R_a = \min\{t : X_t = a\} = \min\{t : B_t = (a - \mu t)/\sigma\}$$

(3.9) **Theorem.** *If* $\mu > 0$ *and* $a < 0$, *then* $P(R_a < \infty) = e^{2\mu a/\sigma^2}$.

To see that this answer is reasonable, note that if a and b are negative, then in order for X_t to reach $a + b$ it must first reach a, an event of probability $P(R_a < \infty)$, and then go from a to $a + b$, an event of probability $P(R_b < \infty)$, so the strong Markov property implies

$$P(R_{a+b} < \infty) = P(R_a < \infty) P(R_b < \infty)$$

This tells us that $P(R_a < \infty) = e^{ca}$ for some c, but cannot tell us that the value of $c = 2\mu/\sigma^2$. To see that the answer should only depend on μ/σ^2 note that

$$X(t/\sigma^2) = \sigma B(t/\sigma^2) + (\mu/\sigma^2)t$$

which by the scaling relation (1.5) has the same distribution as $B_t + (\mu/\sigma^2)t$.

Why is (3.9) true? If we stop the martingale $M_t^\theta = \exp(\theta B_t - t\theta^2/2)$ at time R_a, then $B(R_a) = (a - \mu R_a)/\sigma$, so

$$M^\theta(R_a) = \exp\{\theta(a - \mu R_a)/\sigma - R_a\theta^2/2\} = e^{\theta a/\sigma} \exp\{-R_a(\mu\theta/\sigma + \theta^2/2)\}$$

This suggests choosing θ so that $\mu\theta/\sigma + \theta^2/2 = 0$; i.e., $\theta = -2\mu\sigma$. Taking $T = R_a$ in the optional stopping theorem, (3.2), and supposing there is no contribution from the set $\{R_a = \infty\}$, we have

$$1 = EM^\theta(R_a) = e^{-2\mu a/\sigma^2} P(R_a < \infty)$$

and rearranging gives the desired result. □

Proof of (3.9). Applying the stopping theorem (3.2) at time $R_a \wedge t$ we get

$$\begin{aligned} & 1 = E\exp\{-(2\mu/\sigma)B_{R_a\wedge t} - (2\mu/\sigma)^2(R_a \wedge t)/2\} \\ (\star) \qquad & \quad = E\left(\exp\{-(2\mu/\sigma)B_{R_a} - (2\mu^2/\sigma^2)R_a\}\,; R_a \le t\right) \\ & \quad + E\left(\exp\{-(2\mu/\sigma)B_t - (2\mu^2/\sigma^2)t\}\,; R_a > t\right) \end{aligned}$$

Since $B(R_a) = (a - \mu R_a)/\sigma$, the second line is

$$E\left(e^{-2\mu a/\sigma^2}; R_a \le t\right) = e^{-2\mu a/\sigma^2} P(R_a \le t) \to e^{-2\mu a/\sigma^2} P(R_a < \infty)$$

as $t \to \infty$. To deal with the third line in $(\star)$, we note that the strong law of large numbers in Exercise 1.6 implies that with probability one $B_t/t \to 0$. Thus on the event $\{R_a > t\}$ we have

$$\exp(-(2\mu/\sigma)B_t - (2\mu^2/\sigma^2)t) \to 0$$

and the contribution to the expected value from $R_a > t$ tends to 0. Combining the last three displays gives the desired result. □

d. Higher-order martingales

The result in (3.8) says

$$E(\exp(\theta B_t - \theta^2 t/2)|B_r, r \le s) = \exp(\theta B_s - \theta^2 s/2)$$

Differentiating with respect to θ (and not worrying about the details) gives

$$(*) \quad E((B_t - \theta t)\exp(\theta B_t - \theta^2 t/2)|B_r, r \le s) = (B_s - \theta s)\exp(\theta B_s - \theta^2 s/2)$$

Setting $\theta = 0$, we get the martingale in (3.1):

$$E(B_t|B_r, r \le s) = B_s$$

Differentiating $(*)$ again gives

$$(**) \quad \begin{aligned} &E(\{(B_t - \theta t)^2 - t\}\exp(\theta B_t - \theta^2 t/2)|B_r, r \le s) \\ &= \{(B_s - \theta s)^2 - s\}\exp(\theta B_s - \theta^2 s/2) \end{aligned}$$

Setting $\theta = 0$, we get the martingale in (3.5):

$$E(B_t^2 - t|B_r, r \le s) = B_s^2 - s$$

Having reinvented our linear and quadratic martingales, it is natural to differentiate more to find a new martingales. To get organized for doing this, we let

$$f(\theta, x, t) = \exp(\theta x - \theta^2 t/2)$$

let $f_k(\theta, x, t)$ be the k the derivative w.r.t. θ, and let $h_k(x,t) = f_k(0, x, t)$. Repeating the arguments for $(*)$ and $(**)$ we see that:

(3.10) **Theorem.** *$h_k(B_t, t)$ is a martingale.*

As noted above, we have already seen $h_1(B_t, t)$ and $h_2(B_t, t)$, but the other ones are new:

k	$f_k(\theta, x, t)$	$h_k(B_t, t)$
1	$(x - \theta t)f(\theta, x, t)$	B_t
2	$\{(x - \theta t)^2 - t\}f(\theta, x, t)$	$B_t^2 - t$
3	$\{(x - \theta t)^3 - 3t(x - \theta t)\}f(\theta, x, t)$	$B_t^3 - 3tB_t$
4	$\{(x - \theta t)^4 - 6t(x - \theta t)^2 + 3t^2\}f(\theta, x, t)$	$B_t^4 - 6tB_t^2 + 3t^2$

The next result illustrates the use of the martingale $h_4(B_t, t)$:

(3.11) *If $\tau_a = \inf\{t : B_t \notin (-a, a)\}$, then*

$$E\tau_a^2 = 5a^4/3$$

To see why the answer is a multiple of a^4, note that the scaling relationship tells us that τ_a has the same distribution as $a^2\tau_1$, so $E\tau_a^2 = a^4 E\tau_1^2$.

Proof. Using (3.2) at time $\tau_a \wedge t$, we have

$$E(B(\tau_a \wedge t)^4 - 6(\tau_a \wedge t)B(\tau_a \wedge t)^2) = -3E(\tau_a \wedge t)^2$$

From (3.6) we know that $E\tau_a = a^2 < \infty$. Letting $t \to \infty$, gives

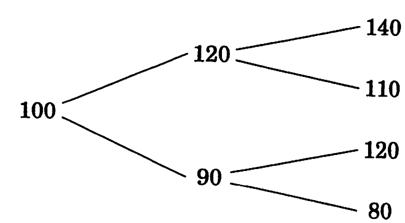
$$a^4 - 6a^2 E\tau_a = -3E(\tau_a^2)$$

Plugging in $E\tau_a = a^2$ and doing a little arithmetic gives the desired result. □

6.4. Option Pricing in Discrete Time

In the next section we will use our knowledge of Brownian motion to derive the Black–Scholes formula. To prepare for that computation, we will consider the much simpler problem of pricing options when there are a finite number of time periods and two possible outcomes at each stage. The restriction to two outcomes is not as bad as one might think. One justification for this is that we are looking at the process on a very slow time scale, so at most one interesting event happens (or not) per time period. We begin by considering a very simple special case.

Example 4.1. Two-period binary tree. Suppose that a stock price starts at 100 at time 0. At time 1 (one day or one month or one year later) it will either be worth 120 or 90. If the stock is worth 120 at time 1, then it might be worth 130 or 110 at time 2. If the price is 90 at time 1, then the possibilities at time 2 are 120 and 80. The last three sentences can be simply summarized by the following tree.



Suppose now that you are offered a **European call option** with **strike price** 100 and **expiry** 2. This means you have an option to buy the stock (but not an obligation to do so) for 100 at time 2, i.e., after seeing the outcome of the first and second stages. If the stock price is 80, you will not exercise the option to purchase the stock and your profit will be 0. In the other cases you will choose to buy the stock at 100 and then immediately sell it at X_2 to get a payoff of $X_2 - 100$ where X_2 is the stock price at time 2. Combining the two cases we can write the payoff in general as $(X_2 - 100)^+$, where $z^+ = \max\{z, 0\}$ denotes the positive part of z.

Our problem is to figure out what is the right price for this option. At first glance this may seem impossible since we have not assigned probabilities to the various events. However, it is a miracle of "pricing by the absence of arbitrage" that in this case we do not have to assign probabilities to the events to compute the price. To explain this we start by considering a small piece of the tree. When $X_1 = 90$, X_2 will be 120 ("up") or 80 ("down") for a profit of 30 or a loss of 10, respectively. If we pay c for the option, then when X_2 is up we make a profit of $20 - c$, but when it is down we make $-c$. The last two sentences are summarized in the following table

	stock	option
up	30	$20 - c$
down	-10	$-c$

Suppose we buy x units of the stock and y units of the option, where negative numbers indicate that we sold instead of bought. One possible strategy is to choose x and y so that the outcome is the same if the stock goes up or down:

$$30x + (20 - c)y = -10x + (-c)y$$

Solving, we have $40x + 20y = 0$ or $y = -2x$. Plugging this choice of y into the last equation shows that our profit will be $(-10 + 2c)x$. If $c > 5$, then we can make a large profit with no risk by buying large amounts of the stock and selling twice as many options. Of course, if $c < 5$, we can make a large profit by doing the reverse. Thus, in this case the only sensible price for the option is 5.

A scheme that makes money without any possibility of a loss is called an **arbitrage opportunity**. It is reasonable to think that these will not exist in financial markets (or at least be short-lived) since if and when they exist people take advantage of them and the opportunity goes away. Using our new terminology we can say that the only price for the option which is consistent with absence of arbitrage is $c = 5$, so that must be the price of the option (at time 1 when $X_1 = 90$).

Before we try to tackle the whole tree to figure out the price of the option at time 0, it is useful to look at things in a different way. Generalizing our example, let $a_{i,j}$ be the profit for the ith security when the jth outcome occurs.

(4.1) **Theorem.** *Exactly one of the following holds:*

(i) There is a betting scheme x_i so that $\sum_{i=1}^m x_i a_{i,j} \geq 0$ for each j and $\sum_{i=1}^m x_i a_{i,k} > 0$ for some k.

(ii) There is a probability vector $p_j > 0$ so that $\sum_{j=1}^n a_{i,j} p_j = 0$ for all i.

Here an x satisfying (i) is an arbitrage opportunity. We never lose any money but for at least one outcome we gain a positive amount. Turning to (ii), the vector p_j is called a martingale measure since if the probability of the jth outcome is p_j, then the expected change in the price of the ith stock is equal to 0. Combining the two interpretations we can restate (4.1) as:

(4.2) **Theorem.** *There is no arbitrage if and only if there is a strictly positive probability vector so that all the stock prices are martingale.*

Why is (4.1) true? One direction is easy. If (i) is true, then for any strictly positive probability vector $\sum_{i=1}^m \sum_{j=1}^n x_i a_{i,j} p_j > 0$, so (ii) is false.

Suppose now that (i) is false. The linear combinations $\sum_{i=1}^m x_i a_{i,j}$ when viewed as vectors indexed by j form a linear subspace of n-dimensional Euclidean space. Call it $\mathcal{L}$. If (i) is false, this subspace intersects the positive orthant $\mathcal{O} = \{y : y_j \geq 0 \text{ for all } j\}$ only at the origin. By linear algebra we know that $\mathcal{L}$ can be extended to an $n - 1$ dimensional subspace $\mathcal{H}$ that only intersects $\mathcal{O}$ at the origin.

Since $\mathcal{H}$ has dimension $n-1$, it can be written as $\mathcal{H} = \{y : \sum_{j=1}^n y_j p_j = 0\}$. Since for each fixed i the vector $a_{i,j}$ is in $\mathcal{L} \subset \mathcal{H}$, (ii) holds. To see that all the $p_j > 0$ we leave it to the reader to check that if not, there would be a non-zero vector in $\mathcal{O}$ that would be in $\mathcal{H}$. □

To apply (4.1) to our simplified example we begin by noting that in this case $a_{i,j}$ is given by

		$j = 1$	$j = 2$
stock	$i = 1$	30	-10
option	$i = 2$	$20 - c$	$-c$

By (4.2) if there is no arbitrage, then there must be an assignment of probabilities p_j so that

$$30p_1 - 10p_2 = 0 \qquad (20 - c)p_1 + (-c)p_2 = 0$$

From the first equation we conclude that $p_1 = 1/4$ and $p_2 = 3/4$. Rewriting the second we have

$$c = 20p_1 = 20 \cdot (1/4) = 5$$

To generalize from the last calculation to finish our example we note that the equation $30p_1 - 10p_2 = 0$ says that under p_j the stock price is a martingale (i.e., the average value of the change in price is 0), while $c = 20p_1 + 0p_2$ says that the price of the option is then the expected value under the martingale probabilities. Using these ideas we can quickly complete the computations in our example. When $X_1 = 120$ the two possible scenarios lead to a change of $+20$ or -5, so the relative probabilities of these two events should be $1/5$ and $4/5$. When $X_0 = 0$ the possible price changes on the first step are $+20$ and -10, so their relative probabilities are $1/3$ and $2/3$. Making a table of the possibilities, we have

X_1	X_2	probability	$(X_2 - 100)^+$
120	140	$(1/3)\cdot(1/5)$	40
120	115	$(1/3)\cdot(4/5)$	15
90	120	$(2/3)\cdot(1/4)$	20
90	80	$(2/3)\cdot(3/4)$	0

so the value of the option is

$$\frac{1}{15}\cdot 40 + \frac{4}{15}\cdot 15 + \frac{1}{6}\cdot 20 = \frac{80+120+100}{30} = 10 \qquad \square$$

The last derivation may seem a little devious, so we will now give a second derivation of the price of the option. In the scenario described above, our investor has four possible actions:

A_0. Put \$1 in the bank and end up with \$1 in all possible scenarios.

A_1. Buy one share of stock at time 0 and sell it at time 1.

A_2. Buy one share at time 1 if the stock is at 120, and sell it at time 2.

A_3. Buy one share at time 1 if the stock is at 90, and sell it at time 2.

These actions produce the following payoffs in the indicated outcomes

X_1	X_2	A_0	A_1	A_2	A_3	option
120	140	1	20	20	0	40
120	115	1	20	-5	0	15
90	120	1	-10	0	30	20
90	80	1	-10	0	-10	0

Noting that the payoffs from the four actions are themselves vectors in four-dimensional space, it is natural to think that by using a linear combination of these actions we can reproduce the option exactly. To find the coefficients we write four equations in four unknowns,

$$\begin{aligned} z_0 + 20z_1 + 20z_2 &= 40 \\ z_0 + 20z_1 - 5z_2 &= 15 \\ z_0 - 10z_1 + 30z_3 &= 20 \\ z_0 - 10z_1 - 10z_3 &= 0 \end{aligned}$$

Subtracting the second equation from the first and the fourth from the third gives $25z_2 = 25$ and $40z_3 = 20$ so $z_2 = 1$ and $z_3 = 1/2$. Pugging in these values, we have two equations in two unknowns:

$$z_0 + 20z_1 = 20 \qquad z_0 - 10z_1 = 5$$

Taking differences, we conclude $30z_1 = 15$, so $z_1 = 1/2$ and $z_0 = 10$.

The reader may have already noticed that $z_0 = 10$ is the option price. This is no accident. What we have shown is that with \$10 cash we can buy and sell shares of stock to produce the outcome of the option in all cases. In the terminology of Wall Street, $z_1 = 1/2$, $z_2 = 1$, $z_3 = 1/2$ is a **hedging strategy** that allows us to **replicate the option**. Once we can do this it follows that the fair price must be \$10. To do this note that if we could sell it for \$12 then we can take \$10 of the cash to replicate the option and have a sure profit of \$2.

6.5. The Black–Scholes Formula

In this section we will find the fair price of a European call option $(X_t - K)^+$ with strike price K and expiry t based on a stock price that follows

$$X_t = X_0 \cdot \exp(\mu t + \sigma B_t) \tag{5.1}$$

Here B_t is a standard Brownian motion, μ is the exponential growth rate of the stock, and σ is its volatility.

In writing the model we have assumed that the growth rate and volatility of the stock are constant. If we also assume that the interest rate r is constant, then the discounted stock price is

$$e^{-rt}X_t = X_0 \cdot \exp((\mu - r)t + \sigma B_t)$$

Here we have to multiply by e^{-rt}, since \$1 at time t has the same value as e^{-rt} dollars today.

Extrapolating wildly from (4.1), we can say that any consistent set of prices must come from a martingale measure. Combining this with the fact proved in (3.8) that $\exp(\theta B_t - (t\theta^2/2))$ is martingale, we see that if

$$\mu = r - \sigma^2/2 \tag{5.2}$$

then the discounted stock price, $e^{-rt}X_t$ is a martingale. To compute the value of the call option, we need to compute its value in the model in (5.1) for this special value of μ. Using the fact that $\log(X_t/X_0)$ has a normal$(\mu t, \sigma^2 t)$ distribution, we see that

$$E(e^{-rt}(X_t - K)^+) = e^{-rt}\int_{\log(K/X_0)}^{\infty} (X_0 e^y - K)\frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-(y-\mu t)^2/2\sigma^2 t}\, dy$$

Changing variables $y = \mu t + w\sigma\sqrt{t}$, $dy = \sigma\sqrt{t}\, dw$ the integral is equal to

$$(\star) \qquad = e^{-rt}X_0 e^{\mu t}\frac{1}{\sqrt{2\pi}}\int_\alpha^\infty e^{w\sigma\sqrt{t}}e^{-w^2/2}\, dw - e^{-rt}K\frac{1}{\sqrt{2\pi}}\int_\alpha^\infty e^{-w^2/2}\, dw$$

where $\alpha = (\log(K/X_0) - \mu t)/\sigma\sqrt{t}$. The first integral

$$\frac{1}{\sqrt{2\pi}}\int_\alpha^\infty e^{w\sigma\sqrt{t}}e^{-w^2/2}\, dw = e^{t\sigma^2/2}\int_\alpha^\infty \frac{1}{\sqrt{2\pi}}e^{-(w-\sigma\sqrt{t})^2/2}\, dw$$
$$= e^{t\sigma^2/2}\, P(\text{normal}(\sigma\sqrt{t}, 1) > \alpha)$$

The last probability can be written in terms of the distribution function Φ of a standard normal χ, i.e., $\Phi(t) = P(\chi \le t)$, by noting

$$P(\text{normal}(\sigma\sqrt{t}, 1) > \alpha) = P(\chi > \alpha - \sigma\sqrt{t})$$
$$= P(\chi \le \sigma\sqrt{t} - \alpha) = \Phi(\sigma\sqrt{t} - \alpha)$$

where in the middle equality we have used the fact that χ and $-\chi$ have the same distribution. Using the last two computations in $(\star)$ converts it to

$$e^{-rt}X_0 e^{\mu t}e^{\sigma^2 t/2}\Phi(\sigma\sqrt{t} - \alpha) - e^{-rt}K\Phi(-\alpha)$$

Using (5.2) now the expression simplifies to the

(5.3) **Black–Scholes formula.** *The price of the European call option $(X_T - K)^+$ is given by*

$$X_0\Phi(\sigma\sqrt{t} - \alpha) - e^{-rt}K\Phi(-\alpha)$$

where $\alpha = \{\log(K/X_0 e^{\mu t})\}/\sigma\sqrt{t}$ and $\mu = r - \sigma^2/2$.

To try to come to grips with this ugly formula note that $K/X_0 e^{\mu t}$ is the ratio of the strike price to the expected value of the stock at time t under the martingale probabilities, while $\sigma\sqrt{t}$ is the standard deviation of $\log(X_t/X_0)$.

Example 5.1. Microsoft call options. The February 23, 1998, *Wall Street Journal* listed the following prices for July call options on Microsoft stock.

strike	75	80	85
price	11	8 1/8	5 1/2

On this date Microsoft stock was trading at 81 5/8, while the annual interest rate was about 4% per year. Should you buy the call option with strike 80?

Solution. The answer to this question will depend on your opinion of the volatility of the market over the period. Suppose that we follow a traditional rule of thumb and decide that $\sigma = 0.3$; i.e., over a one-year period a stock's price might change by about 30% of its current value. In this case the drift rate for the martingale measure is

$$\mu = r - \sigma^2/2 = .04 - (.09)/2 = .04 - .045 = -.005$$

and so the log ratio is

$$\log(K/X_0 e^{\mu t}) = \log(80/(81.625e^{-.005(5/12)}) = \log(80/81.455) = -.018026$$

Five months corresponds to $t = 5/12$, so the standard deviation

$$\sigma\sqrt{t} = .3\sqrt{5/12} = .19364$$

and $\alpha = -.018026/.19364 = -.09309$. Plugging in now, we have a price of

$$81.625\Phi(.19365 + .09309) - e^{-.04(5/12)}80\Phi(.09309)$$
$$= 81.625\Phi(.28674) - 78.678\Phi(.09309)$$
$$= 81.625(.6128) - 78.678(.5371) = 50.02 - 42.25 = 7.76$$

This is somewhat lower than the price quoted in the paper. There are two reasons for this. First, the options listed in the *Wall Street Journal* are **American call options.** The holder has the right to exercise at any time during the life of the option. Since one can ignore the additional freedom to exercise early, American options are at least as valuable as their European counterparts.

Second, and perhaps more importantly, we have not spent much effort on our estimate of r and σ. In connection with the last point it is interesting to note that:

(5.4) **Lemma.** *If we fix the value of r, then the Black–Scholes price is an increasing function of the volatility.*

Thus for any given interest rate, we can, by varying the volatility, match the Black–Scholes price exactly. The magic value of the volatility that gives us the option price is called the **implied volatility**. A consequence of this observation is that we cannot test the adequacy of the Black–Scholes model by looking only at one price.

Example 5.2. Intel call options. Again consulting the *Wall Street Journal* for February 23, 1998, we find the following prices listed for July call options on Intel stock, which was trading at 94 3/16.

strike	70	75	80	85	90	95	100	105
price	26	22	18	$14\frac{1}{2}$	$11\frac{3}{8}$	$8\frac{3}{4}$	$6\frac{1}{2}$	$4\frac{3}{8}$

We leave it as an exercise for the reader to compute the volatilities implied these prices. I have not done this, but folk wisdom suggests that the volatilities will not be constant but "smile" in some convex sort of way. My personal intuition is that large movements in the market are more likely than what is predicted by the Brownian model, so options that are well "out of the money" are overpriced. This situation is also reasonable from an economic point of view as well. Such financial instruments have low payoff relative to their uncertainty and hence are not attractive to sell.

Replication. By analogy with the discrete case, we might ask: Can the option be replicated by using cash and trading in the underlying security? The answer to this questions is yes. Its proof is beyond the scope of this book. However, it is possible to describe the hedging strategy. Let

$$v(s,x) = E((X_t - K)^+ | X_s = x)$$

be the value of the option at time s when the stock price is x where the expected value is computed with respect to the martingale measure. A result known as Ito's formula allows us to conclude

$$v(t, X_t) = v(0, X_0) + \int_0^t \frac{\partial v}{\partial x}(s, X_s)\, dX_s$$

The term $v(0, X_0)$ is the cash we put in to start the replication process. The second term is a "stochastic integral." Intuitively, at time s we hold $(\partial v/\partial x)(s, X_s)$ of the stock, and the second term computes our profit (or loss from using this trading strategy. At time t, $v(t, X_t) = (X_t - K)^+$ is the outcome of the option, so this strategy reproduces the option exactly. As in Section 6.4, the assumption of no arbitrage implies that the price of the option must be $v(0, X_0)$

6.6. Exercises

6.1. Show that if $s < t$, the conditional distribution $(B_s | B_t = z)$ is normal with mean zs/t and variance $s(t-s)/t$.

6.2. (a) Compute the joint distribution $((B_r, B_s) | B_t = 0)$ to show that the covariance is $r(t-s)/t$. (b) Conclude that $B_s^{0,t} = B_s - (s/t)B_t$, $0 \le s \le t$ has the same distribution as $(\{B_s, 0 \le s \le t\} | B_t = 0)$.

6.3. (a) Show that the joint distribution $((B_r - rz/t, B_s - sz/t) | B_t = z)$ is the same as that of $((B_r, B_s) | B_0 = 0)$ (b) Use the result of the previous exercise to conclude that $((B_r, B_s) | B_t = z)$ is a bivariate normal in which the components have means $\mu_r = zr/t$, $\mu_s = zs/t$, variances $\sigma_r^2 = r(t-r)/t$, $\sigma_s^2 = s(t-s)/t$, and covariance $E((B_r, B_s) | B_t = z) = r(t-s)/t$. (b) Show that $B_s^{z,t} = B_s + (s/t)(z - B_t)$, $0 \le s \le t\}$ has the same distribution as $(\{B_s, 0 \le s \le t\} | B_t = z)$.

6.4. Let $Y_t = \int_0^t B_s\, ds$. Find (a) EY_t, (b) EY_t^2, (c) $E(Y_s, Y_t)$.

6.5. Find the conditional distribution of $Y_t = \int_0^t B_s\, ds$ given $B_t = x$.

6.6. *Strong law.* (a) Use the strong law of large numbers to show that as $n \to \infty$ through the integers, we have $B(n)/n \to 0$. (b) Use (ii) of Exercise 2.1 in Chapter 5 to extend the previous conclusion to show that as $u \to \infty$ through the real numbers, we have $B(u)/u \to 0$.

6.7. *Inversion.* Let B_t be a standard Brownian motion. Define a process X by setting $X_0 = 0$ and $X_t = tB(1/t)$ for $t > 0$. Show that X_t is a Gaussian process with $EX_t = 0$ and $E(X_s X_t) = s \wedge t$, so X_t is a standard Brownian motion. (b) Let $u = 1/t$ in the previous exercise to conclude that $\lim_{t \to 0} X_t = 0$.

6.8. Let $B_t^0 = B_t - tB_1$ be Brownian bridge and let $X_t = (1+t)B^0(t/(1+t))$. Show that X_t is a standard Brownian motion.

6.9. Check that $(1 - 3y^{-4})e^{-y^2/2} \le e^{-y^2/2} \le e^{-x^2/2}e^{-x(y-x)}$ for $y \ge x$. Then integrate to conclude

$$(x^{-1} - x^{-3})\exp(-x^2/2) \le \int_x^\infty \exp(-y^2/2)dy \le x^{-1}\exp(-x^2/2)$$

6.10. Show that if X and Y are independent normal$(0,\sigma^2)$, then $U = (X+Y)/2$ and $V = (X-Y)/2$ are independent normal$(0,\sigma^2/2)$.

6.11. *Levy's interpolation construction of Brownian motion.* We construct the Brownian motion only for $0 \le t \le 1$. For each $n \ge 0$ let $Y_{n,1}, \ldots, Y_{n,2^n}$ be independent normal(0,1). Start by setting $B(0) = 0$ and $B(1) = Y_{0,1}$. Assume now that $B(m/2^n)$ has been defined for $0 \le m \le 2^n$ and let

$$B((2m+1)/2^{n+1}) = \frac{\{B((m+1)/2^n) - B(m/2^n)\} + 2^{-n/2}Y_{n,m+1}}{2}$$

(a) Use the previous exercise to show that the increments

$$B(k/2^{n+1}) - B((k-1)/2^{n+1}) \quad 1 \le k \le 2^{n+1}$$

are independent normal$(0,2^{-(n+1)})$. (b) Let $B_n(m/2^n) = B(m/2^n)$ and let $B_n(t)$ be linear on each interval $[m/2^n, (m+1)/2^n]$. Show that

$$\max_t |B_n(t) - B_{n+1}(t)| = \frac{1}{2} \cdot 2^{-n/2} \max_m |Y_{n+1,m}|$$

(c) Use Exercise 6.9 to conclude that

$$P\left(\max_{1\le m\le 2^n} |Y_{n+1,m}| \ge \sqrt{2n}\right) \le 2 \cdot 2^n \cdot e^{-n}$$

(d) Use (b) and (c) to conclude that there is a limit $B(t)$ so that for large n:

$$\max_t |B_n(t) - B(t)| \le \sum_{k=n}^{\infty} 2^{-k/2}\sqrt{2k}$$

This shows that $B(t)$, $0 \le t \le 1$ is a uniform limit of continuous functions. It then follows from a result in real analysis that $B(t)$ is continuous.

6.12. *Nondifferentiability of Browian paths.* The point of this exercise is to show that with probability one, Brownian paths are not Lipschitz continuous (and hence not differentiable) at any point. Fix a constant $C < \infty$ and let $A_n = \{\omega : \text{there is an } s \in [0,1] \text{ so that } |B_t - B_s| \le C|t-s| \text{ when } |t-s| \le 3/n\}$. For $1 \le k \le n-2$ let

$$Y_{k,n} = \max\left\{\left|B\left(\frac{k+j}{n}\right) - B\left(\frac{k+j-1}{n}\right)\right| : j = 0,1,2\right\}$$
$$G_n = \{ \text{ at least one } Y_{k,n} \le 5C/n\}$$

(a) Show that $A_n \subset G_n$. (b) Show that $P(G_n) \to 0$. (c) Conclude that $P(A_m) = 0$ for all m, which gives the desired conclusion.

6.13. *Quadratic variation.* Fix t and let $\Delta_{m,n} = B(tm2^{n-1}) - B(t(m-1)2^{n-1})$. Compute

$$E\left(\sum_{m\le 2^n} \Delta^2_{m,n} - t\right)^2$$

and conclude that $\sum_{m\le 2^n} \Delta^2_{m,n} \to t$ as $n \to \infty$.

6.14. Let $\varphi_a(\lambda) = E\exp(-\lambda T_a)$ be the Laplace transform of T_a. (a) Use (2.3) to conclude that $\varphi_a(\lambda)\varphi_b(\lambda) = \varphi_{a+b}(\lambda)$. (b) Use the scaling relationship to conclude that T_a has the same distribution as a^2T_1. (c) Combine the first two parts to conclude that there is a constant κ so that $\varphi_a(\lambda) = \exp(-a\kappa\sqrt{\lambda})$.

6.15. Let (B^1_t, B^2_t) be a two dimensional Brownian motion starting from 0. Let $T^1_a = \min\{t : B^1_t = a\}$ and let $Y_a = B^2(T^1_a)$. (a) Use the strong Markov property to conclude that if $a < b$ then Y_a and $Y_b - Y_a$ are independent. (b) Use the scaling relationship (1.5) to conclude that Y_a has the same distribution as aY_1. (c) Use (b) the fact that Y_a and $-Y_a$ have the same distribution to show that the Fourier transform $E\exp(i\theta Y_a) = \exp(-a\kappa|\theta|)$ for some constant κ.

6.16. Continuing with the set-up of the previous problem, use the density function of T_a given in (2.5) to show that Y_a has a Cauchy distribution.

6.17. Let $x > 0$. Compute $P(B_t > y, \min_{0\le u\le t} B_t > 0 | B_0 = x)$ and then differentiate to find the density function.

6.18. Let $M_t = \max_{0\le u\le t} B_u$. Find $P(M_t \ge x, B_t \le y)$ for $x > y$ then differentiate to find the joint density of (M_t, B_t).

6.19. Use the fact that $A_s = B_t - B_{t-s}$ is a Brownian motion and the previous exercise to conclude

$$P(B_t > x | B_u \ge 0 \text{ for all } 0 \le u \le t) = P(M_t > x | B_t = M_t) = e^{-x^2/2t}$$

6.20. Let $r < s < t$. Find the conditional probability that a standard Brownian motion is not 0 in the interval (s,t) given that it was not 0 in the interval (r,s).

6.21. Let $R_t = \min\{s > t : B_s = 0\}$. Find the density function $P_0(R_t = u)$.

6.22. Let $T_a = \inf\{t : B_t = a\}$. Compute $P(T_1 < T_{-2} < T_3)$.

6.23. Show that for each $a > 1$ there is a unique b so that $P(T_1 < T_{-a} < T_b) = 1/2$. Find the value of b for $a = 2$.

6.24. Let $(B^1_t, \ldots B^d_t)$ be a d-dimensional Brownian motion and let

$$R_t = \sqrt{(B^1_t)^2 + \cdots + (B^d_t)^2}$$

be its distance from the origin, and $S_r = \min\{t : R_t = r\}$ be the first time the distance from the origin is equal to r. (a) Show that $R^2_t - td$ is a martingale. (b) Imitate the proof of (3.6) to conclude that $ES_r = r^2/d$.

6.25. (a) Integrate by parts to show that

$$\int x^m e^{-x^2/2}\,dx = (m-1)\int x^{m-2}e^{-x^2/2}\,dx$$

(b) Use this result to conclude that if χ has a standard normal distribution then for any integer $n \geq 1$

$$E\chi^{2n} = (2n-1)(2n-3)\cdots 3\cdot 1 \quad \text{and} \quad E\chi^{2n-1} = 0$$

That is, $E\chi^4 = 3$, $E\chi^6 = 15$, etc., while all odd moments are 0.

6.26. Verify directly that

$$E(B_t^4 - 6tB_t^2 + 3t^2 | B_r, r \leq s) = B_s^4 - 6sB_s^2 + 3s^2$$

6.27. Let $T_a = \min\{t : B_t = a\}$. Use the exponential martingale in part (c) of Section 6.3 to conclude that $E\exp(-\lambda T_a) = e^{-a\sqrt{2\lambda}}$.

6.28. Let $\tau = \inf\{t : B_t \notin (a,b)\}$ and let $\lambda > 0$. Use the strong Markov property to show

$$E\exp(-\lambda T_a) = E(e^{-\lambda\tau}; T_a < T_b) + E(e^{-\lambda\tau}; T_b < T_a)\, E\exp(-\lambda T_{a-b})$$

(ii) Interchange the roles of a and b to get a second equation, use the previous exercise, and solve to get

$$E_x(e^{-\lambda T}; T_a < T_b) = \sinh(\sqrt{2\lambda} b)/\sinh(\sqrt{2\lambda}(b-a))$$
$$E_x(e^{-\lambda T}; T_b < T_a) = \sinh(-\sqrt{2\lambda} a)/\sinh(\sqrt{2\lambda}(b-a))$$

6.29. As in Example 3.4, let $X_t = \sigma B_t + \mu t$ with $\mu > 0$ and $R_a = \min\{t : X_t = a\}$ where $a < 0$. (i) Set $\mu\theta/\sigma + \theta^2/2 = \lambda$ and solve to get a negative solution $\theta = -\nu - (\nu^2 + 2\lambda)^{1/2}$ where $\nu = \mu/\sigma$. (ii) Use the reasoning for (3.9) to show

$$E\exp(-\lambda R_a) = \exp(-\theta a/\sigma) = \exp\left(\frac{a}{\sigma}\{\nu + (\nu^2 + 2\lambda)^{1/2}\}\right)$$

Taking $\lambda = 0$, we get (3.9): $P(R_a < \infty) = \exp(2\mu a/\sigma^2)$.

6.30. Again let $X_t = \sigma B_t + \mu t$ where $\mu > 0$ but now let $R_b = \min\{t : X_t = b\}$ where $b > 0$. Show that $ER_b = b/\mu$.

6.31. As in the previous exercise, let $X_t = \sigma B_t + \mu t$ with $\mu > 0$ and $R_b = \min\{t : X_t = b\}$ where $b > 0$. (i) Set $\mu\theta/\sigma + \theta^2/2 = \lambda$ and solve to get a positive solution $\theta = -\nu + (\nu^2 + 2\lambda)^{1/2}$ where $\nu = \mu/\sigma$. (ii) Use the reasoning for (3.9) to show

$$E\exp(-\lambda R_b) = \exp(-\theta b/\sigma) = \exp\left(\frac{b}{\sigma}\{\nu - (\nu^2 + 2\lambda)^{1/2}\}\right)$$

Taking $\lambda = 0$ this time we have $P(R_b < \infty) = 1$.

6.32. Let $\tau_a = \inf\{t : B_t \notin (-a,a)\}$ and recall $\cosh(x) = (e^x + e^{-x})/2$. Show that $\exp(-\theta^2 t/2)\cosh(\theta B_t)$ is a martingale and use this to conclude that

$$E\exp(-\lambda\tau_a) = 1/\cosh(a\sqrt{2\lambda})$$

6.33. Let $p_t(x,y) = (2\pi t)^{-1/2} e^{-(y-x)^2/2t}$. Check that

$$\frac{\partial}{\partial t} p_t(x,y) = \frac{1}{2}\frac{\partial^2}{\partial y^2} p_t(x,y)$$

6.34. Let $u(t,x)$ be a function that satisfies

$$(*) \qquad \frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} = 0 \quad \text{and} \quad \left|\frac{\partial^2 u}{\partial x^2}(t,x)\right| \leq C_T \exp(x^2/(t+\epsilon)) \quad \text{for } t \leq T$$

Show that $u(t, B_t)$ is a martingale by using the previous exercise to conclude

$$\frac{\partial}{\partial t} Eu(t, B_t) = \int \frac{\partial}{\partial t}(p_t(0,y)u(t,y))\, dy = 0$$

Examples of functions that satisfy $(*)$ are $\exp(\theta x - \theta^2 t/2)$, x, $x^2 - t$, ...

6.35. Find a martingale of the form $B_t^6 - atB_t^4 + bt^2B_t^2 - ct^3$ and use it to compute the third moment of $\tau_a = \inf\{t : B_t \notin (-r,r)\}$. Note that scaling implies $E\tau_a = ca^6$.

6.36. (a) Use Exercise 6.34 to show that $(1+t)^{-1/2}\exp(B_t^2/2(1+t))$ is a martingale. (b) Use (a) to conclude that for large integers n, we have

$$|B_n| < \sqrt{(3.002)n\log n}$$

6.37. *Law of the iterated logarithm.* By working much harder than in the previous exercise one can show that

$$\limsup_{t\to\infty} B_t/\sqrt{2t\log\log t} = 1$$

Use the inversion $X(t) = tB(1/t)$ to conclude that

$$\limsup_{t\to 0} B_t/\sqrt{2t\log\log(1/t)} = 1$$

6.38. The Cornell hockey team is playing a game against Harvard: it will either win lose or draw. A gambler offers you the following three payoffs, each for a \$1 bet.

	win	lose	draw
Bet 1	0	1	1.5
Bet 2	2	2	0
Bet 3	.5	1.5	0

Assume you are able to buy any amounts (even negative) of these bets. Is there an arbitrage opportunity?

6.39. Suppose Microsoft stock sells for 100 while Netscape sells for 50. Three possible outcomes of a court case will have the following impact on the two stocks.

	Microsoft	Netscape
1 (win)	120	30
2 (draw)	110	55
3 (lose)	84	60

What should we be willing to pay for an option to buy Netscape for 50 after the court case is over?

6.40. *Put-call parity.* $(K - X_T)^+$ is a European put with strike price K and expiry T. Show that the put price can be obtained from the call price by

$$Ee^{-rt}(K - X_T)^+ = Ee^{-rt}(X_T - K)^+ - X_0$$

References

Consult the following sources to learn more about the subject.

Athreya, K.B. and Ney, P.E. (1972) *Branching Processes.* Springer-Verlag, New York

Bailey, N.T.J. (1964) *The Elements of Stochastic Processes: With Applications to the Natural Sciences.* John Wiley and Sons.

Barbour, A.D., Holst, L., and Janson, S. (1992) *Poisson Approximation.* Oxford U. Press

Bhattacharya, R.N. and Waymire, C. *Stochastic Processes with Applications.* John Wiley and Sons, New York

Chung, K.L. (1967) *Markov Chains with Stationary Transition Probabilities.* Second edition, Springer-Verlag, New York

Cox, D.R. and Miller, H.D. (1965) *The Theory of Stochastic Processes.* Methuen & Co, Ltd., London

Doob, J.L. (1953) *Stochastic Processes.* John Wiley and Sons, New York

Durrett, R. (1993) *Essentials of Probability.* Duxbury Press, Belmont, CA

Durrett, R. (1995) *Probability: Theory and Examples.* Duxbury Press, Belmont, CA

Durrett, R. (1996) *Stochastic Calculus: A Practical Introduction.* CRC Press, Boca Raton, FL

Feller, W. (1968) *An Introduction to Probability Theory and its Applications.* Third edition, John Wiley and Sons, New York

Hoel, P.G., Port, S.C. and Stone, C.J. (1972) *Introduction to Stochastic Processes.* Houghton-Mifflin, Boston, MA

Hull, J.C. (1997) *Options, Futures, and Other Derivatives.* Prentice Hall, Upper Saddle River, NJ

Jagers, P. (1975) *Branching Processes with Biological Applications.* John Wiley and Sons, New York

Johnson, N.L. and Kotz, S. (1977) *Urn Models and Their Applications.* John Wiley and Sons, New York