

Chapter 2

Brownian Motion.

2.1 Brownian Motion.

As discussed earlier, the “integrated noise” process should be a process with stationary independent increments.

DEFINITION 2.1

A process $Z = \{Z(t) : t \geq 0\}$ is said to have **stationary independent increments** if:

- (i) For $t_1 < t_2 < \dots < t_n$, the random variables $Z(t_n) - Z(t_{n-1}), \dots, Z(t_1) - Z(0)$ are independent;
- (ii) $Z(t + s) - Z(s) \stackrel{\mathcal{D}}{=} Z(t) - Z(0)$ for $s, t \geq 0$.

EXERCISE 2.1

Prove that if Z is a process with stationary independent increments, with $\sigma^2 = \text{Var}[Z(1)]$, then $\text{Var}[Z(t)] = \sigma^2 t$ for $t \geq 0$.

What makes Brownian motion particularly appropriate as an integrated noise process? The key is that it arises naturally as an approximation to more complex processes, in the same way that the normal distribution naturally arises whenever sums of random variables are considered.

Specifically, consider a random walk process.

DEFINITION 2.2

A process $\{S_n : n \geq 0\}$ is said to be a **random walk** if $S_0 = 0$ and the increments $S_1, S_2 - S_1, \dots, S_n - S_{n-1}, \dots$ are i.i.d. random variables.

Random walks occur in many applied problem settings (e.g. fortune of gambler at time n , total cost of running system to time n). The two most fundamental results for random walks are the **law of large numbers** (LLN) and the **central limit theorem** (CLT). Set $X_i = S_i - S_{i-1}$.

THEOREM 2.1 (STRONG LAW OF LARGE NUMBERS)

Suppose that $\{S_n : n \geq 0\}$ is a random walk with $E|X_i| < \infty$. Then,

$$\frac{1}{n}S_n \rightarrow \mu \triangleq E[X_1] \quad \text{a.s. as } n \rightarrow \infty. \quad (2.1)$$

From an applications viewpoint, the Strong Law of Large Numbers (SLLN) allows one to approximate S_n as

$$S_n \stackrel{\mathcal{D}}{\approx} n\mu \quad (2.2)$$

where $\stackrel{\mathcal{D}}{\approx}$ denotes “has approximately the same distribution as”.

REMARK 2.1 The notation $\overset{\mathcal{D}}{\approx}$ has no rigorous mathematical meaning; it is used only to suggest a certain approximation. The rigorous mathematical statement underlying (2.2) is the limit theorem (2.1).

One difficulty with (2.2) is that it approximates the random variable S_n by the deterministic constant $n\mu$. A better approximation can be obtained from the CLT.

THEOREM 2.2 (CENTRAL LIMIT THEOREM)

Suppose that $\{S_n : n \geq 0\}$ is a random walk with $\sigma^2 = \text{Var}[X_i] < \infty$. Then,

$$\sqrt{n} \left(\frac{S_n}{n} - \mu \right) \Rightarrow \sigma N(0, 1) \quad \text{as } n \rightarrow \infty. \quad (2.3)$$

The weak convergence statement (2.3) suggests the approximation

$$S_n \overset{\mathcal{D}}{\approx} n\mu + \sqrt{n}\sigma N(0, 1) \quad (2.4)$$

for large n . As we all know, (2.4) is widely used to approximate the distribution of a sum of random variables.

Relation (2.4) provides an approximation to the distribution of a random walk at a single (large) time point n . What about if we want an approximation to the process $\{S_n : n \geq 0\}$? Let $Z = \{Z(t) : t \geq 0\}$ be such an approximating process. Since S is a discrete-time process with stationary independent increments, Z should inherit this property. And (2.4) suggests that we should approximate S_n by $Z(n)$, where $Z(n) \overset{\mathcal{D}}{=} N(n\mu, n\sigma^2)$. So let $Z = \{Z(t) : t \geq 0\}$ be a real-valued stochastic process such that:

- (i) Z has stationary independent increments;
- (ii) $Z(t) \overset{\mathcal{D}}{=} N(t\mu, t\sigma^2)$ for $t \geq 0$.

The above two properties completely define the finite-dimensional distributions of Z . In other words, properties (i) and (ii) permit one to compute

$$P((Z(t_1), Z(t_2), \dots, Z(t_n)) \in \cdot)$$

for any finite collection $t_1 < t_2 < \dots < t_n$ of time points.

One might expect that this completely specifies the process Z . The following example shows that this is not the case.

EXAMPLE 2.1

Suppose that $Z(t) \equiv 0$ for $t \geq 0$. Let $\tau \sim \exp(1)$ and set

$$\tilde{Z} = \begin{cases} 0, & t \neq \tau \\ 1, & t = \tau \end{cases} \quad (2.5)$$

Since $P(\tau = s) = 0$ for each $s \geq 0$, it follows that the finite-dimensional distributions of \tilde{Z} match those of Z . But Z and \tilde{Z} are clearly different processes.

To completely specify the process Z that approximates the random walk, we will add on an additional requirement related to Z 's sample path behavior (Note that in Example 2.1, there is a unique continuous process, $Z(t) \equiv 0$, that has the common set of finite-dimensional distributions). What should we add?

We might hope that we can find a “version” of Z that has differentiable sample paths, so that the realizations of Z are smooth almost surely. Then, there would exist a stochastic process $\{Z'(t) : t \geq 0\}$ such that for each $t \geq 0$,

$$Z(t+h) - Z(t) = Z'(t)h + o(h) \quad \text{a.s.} \quad (2.6)$$

as $h \searrow 0$. Taking the variance of both sides of (2.5), we get

$$\text{Var}[Z(t+h) - Z(t)] = \text{Var}[Z'(t)]h^2 + o(h) \quad (2.7)$$

as $h \searrow 0$. But the stationary increments property of Z implies that $\text{Var}[Z(t+h) - Z(t)] = \text{Var}[Z(h)]$. And

$$\text{Var}[Z(h)] = \text{Var}[N(\mu h, \sigma^2 h)] = \sigma^2 h.$$

Of course, this contradicts (2.7). We conclude that demanding differentiability of the sample paths is inconsistent with the two properties for Z stated above.

One very important result in the theory of Brownian motion is the following:

THEOREM 2.3

There exists a probability space supporting a stochastic process $\{Z(t) : t \geq 0\}$ satisfying (i) and (ii) above such that $P(\omega : Z(\cdot, \omega) \text{ is continuous}) = 1$.

In other words, properties (i) and (ii) are consistent with continuous sample paths. This theorem is due to (20). Consequently, Brownian motion is often referred to as a Wiener process.

Given Theorem 2.3, it follows that we can, without any real loss of generality, view the sample space as the space of continuous functions. Specifically, we can (if we wish) set $\Omega = C[0, \infty)$, the space of continuous functions $\omega : [0, \infty) \rightarrow \Re$. Then, Theorem 2.3 guarantees that there exists a probability (measure) P on $C[0, \infty)$ such that the co-ordinate random variables

$$Z(t, \omega) \triangleq \omega(t)$$

satisfy (i) and (ii) (Of course, $Z(\cdot, \omega) = \omega(\cdot)$ will necessarily be continuous). Is P unique?

THEOREM 2.4

Let P_1, P_2 be two probability (measures) on $C[0, \infty)$ that share the same finite-dimensional distributions. Then $P_1 = P_2$.

PROOF. See p. 30 of (2). \diamond

According to Theorem 2.4, we can now assert that there is a unique stochastic process $Z = \{Z(t) : t \geq 0\}$ exhibiting the following properties:

- (i) Z has stationary independent increments;
- (ii) $Z(t) \stackrel{\mathcal{D}}{=} N(\mu t, \sigma^2 t)$ for $t \geq 0$;
- (iii) Z has continuous paths.

DEFINITION 2.3

A process Z having the above three properties is a (μ, σ^2) Brownian motion.

We conclude that if S is a random walk, then it is reasonable to try approximating S as a stochastic process as follows:

$$S_n \stackrel{\mathcal{D}}{\approx} Z(n), \quad n \geq 0, \quad (2.8)$$

where Z is a (μ, σ^2) Brownian motion with $\mu = E[X_1]$ and $\sigma^2 = \text{Var}[X_1]$.

For normally distributed random variables (Gaussian variables), one can always express the random variable in terms of a “standard normal random variable” (i.e. a $N(0, 1)$ random variable). Specifically,

$$N(\mu, \sigma^2) \stackrel{\mathcal{D}}{=} \mu + \sigma N(0, 1).$$

Something similar can be done for Brownian motions.

DEFINITION 2.4

A $(0,1)$ Brownian motion is called a standard Brownian motion. We denote standard Brownian motion by $B = \{B(t) : t \geq 0\}$.

EXERCISE 2.2

Prove that if Z is a (μ, σ^2) Brownian motion, then $Z = \mu e + \sigma B$, where $e(t) = t$ for $t \geq 0$.

With the above exercise, the approximation (2.8) can be re-written as

$$S_n \stackrel{\mathcal{D}}{=} \mu n + \sigma B(n) \quad (2.9)$$

for $n \geq 0$. Thus, Brownian motion is the continuous analog to the random walk. Of course, as we indicated earlier, the notation has no rigorous meaning. In the next section, we provide a rigorous sense in which the random walk is approximated by Brownian motion.

REMARK 2.2 For an accessible proof of Theorem 2.3, see pp. 371-378 of (11).

2.2 Donsker's Theorem.

We now discuss a theorem that clarifies the rigorous connection between the random walk and Brownian motion.

THEOREM 2.5 (STRONG APPROXIMATION THEOREM)

Let $\{X_n : n \geq 1\}$ be a sequence of i.i.d. random variables with $\sigma^2 = \text{Var}[X_i] < \infty$. Set $S_0 = 0$ and $S_n = X_1 + \cdots + X_n$. Then, there exists a probability space supporting a sequence $\{S'_n : n \geq 0\}$ and a standard Brownian motion B such that

- (i) $S' \stackrel{\mathcal{D}}{=} S$;
- (ii) $S'_n = n\mu + \sigma B(n) + o(\sqrt{n})$ a.s. as $n \rightarrow \infty$.

where $\mu = EX_1$ and $\sigma^2 = \text{Var}[X_1]$.

This theorem is due to Komlos, Major and Tusnady, and it is known as a “strong approximation theorem.” It should be noted that (ii) asserts that if $\varepsilon_n(\omega) = S'_n(\omega) - n\mu - \sigma B(n, \omega)$, then

$$P \left\{ \omega : \frac{\varepsilon_n(\omega)}{\sqrt{n}} \rightarrow 0 \text{ as } n \rightarrow \infty \right\} = 1.$$

With Theorem 2.5 in hand, it is relatively straightforward to make rigorous statements in which the random walk is approximated by Brownian motion. As an example, consider the CLT. Note that

$$\frac{S_n - n\mu}{\sqrt{n}} \stackrel{\mathcal{D}}{=} \frac{S'_n - n\mu}{\sqrt{n}}. \quad (2.10)$$

But

$$\frac{S'_n - n\mu}{\sqrt{n}} = \frac{\sigma B(n)}{\sqrt{n}} + \frac{o(\sqrt{n})}{\sqrt{n}} \quad \text{a.s.} \quad (2.11)$$

Also,

$$\frac{B(n)}{\sqrt{n}} \stackrel{\mathcal{D}}{=} N(0, 1) \quad \text{for } n \geq 1. \quad (2.12)$$

It follows from (2.10), (2.11), (2.12) and the “converging together” principle that $(S_n - n\mu)/\sqrt{n} \Rightarrow \sigma N(0, 1)$ as $n \rightarrow \infty$. But much stronger results that connect the dynamical behavior of S to that of Brownian motion are also possible.

To establish such a connection, recall that the CLT studies the random walk S over a long time scale n , examining spatial fluctuations of order \sqrt{n} . This suggests considering the stochastic process $\chi_n = \{\chi_n(t) : t \geq 0\}$ defined by

$$\chi_n(t) = \frac{S_{[nt]} - \mu[nt]}{\sqrt{n}}.$$

The process χ_n packs n units of time in S into a unit time interval of χ_n , and scales space in units of \sqrt{n} . Set

$$\chi'_n(t) = \frac{S'_{[nt]} - \mu[nt]}{\sqrt{n}}$$

and note that $\chi'_n \stackrel{\mathcal{D}}{=} \chi_n$ for $n \geq 0$. We wish to show that χ_n converges weakly to Brownian motion as $n \rightarrow \infty$. We start by observing that χ_n is a random element of $D[0, \infty)$, the space of right-continuous functions on $[0, \infty)$ having left limits. Set $B'_n(t) = B(nt)/\sqrt{n}$.

EXERCISE 2.3

Prove that for $n \geq 1$, $B'_n \stackrel{\mathcal{D}}{=} B$, i.e. $B(n\cdot)/\sqrt{n} \stackrel{\mathcal{D}}{=} B(\cdot)$.

To prove weak convergence on $D[0, \infty)$, let $f : D[0, \infty) \rightarrow \mathfrak{R}$ be bounded and uniformly continuous in the Skorohod topology (see Appendix A.4). Note that for $u \geq 0$,

$$\begin{aligned} \sup_{0 \leq t \leq u} |\chi'_n(t) - \sigma B'_n(t)| &\leq \sup_{0 \leq t \leq u} \frac{o(\sqrt{nt})}{\sqrt{n}} + \sup_{0 \leq t \leq u} \frac{\sigma |B([nt]) - B(nt)|}{\sqrt{n}} \quad \text{a.s.} \\ &\leq o(1) + \sigma \sup_{0 \leq k \leq [nu]} \sup_{0 \leq t \leq 1} \frac{|B(k+t) - B(k)|}{\sqrt{n}} \quad \text{a.s.} \\ &= o(1) + \sigma \sup_{0 \leq k \leq [nu]} \frac{\Gamma_k}{\sqrt{n}} \quad \text{a.s.,} \end{aligned} \quad (2.13)$$

where $\Gamma_k = \sup_{0 \leq t \leq 1} |B(k+t) - B(k)|$. Because B has stationary independent increments, the random variables $(\Gamma_k : k \geq 0)$ are i.i.d.

EXERCISE 2.4

Let V_0, V_1, \dots be a sequence of i.i.d. random variables with $E|V_i|^p < \infty$.

- (a) Prove that $\frac{1}{n} \sum_{j=0}^{n-1} |V_j|^p \rightarrow E|V_0|^p$ a.s. as $n \rightarrow \infty$.
- (b) Conclude that $|V_n|^p/n \rightarrow 0$ a.s. as $n \rightarrow \infty$.
- (c) Establish that $\max_{1 \leq k \leq n} |V_k|/(n^{1/p}) \rightarrow 0$ a.s. as $n \rightarrow \infty$.

EXERCISE 2.5

Prove that $E\Gamma_0^2 < \infty$ (in fact, $E\Gamma_0^p < \infty$ for all $p > 0$).

It follows from (2.13) and the above exercises that for $u \leq 0$,

$$\sup_{0 \leq t \leq u} |\chi'_n(t) - \sigma B'_n(t)| \rightarrow 0 \quad \text{a.s. } n \rightarrow \infty$$

Since f is uniformly continuous, $f(\chi'_n) - f(\sigma B'_n) \rightarrow 0$ as $n \rightarrow \infty$. Consequently, the bounded convergence theorem implies that $Ef(\chi'_n) - Ef(\sigma B'_n) \rightarrow 0$ a.s. as $n \rightarrow \infty$. But $Ef(\chi'_n) = Ef(\chi_n)$ and $Ef(\sigma B'_n) = Ef(\sigma B_n)$. Thus, we've proved the following result, known as **Donsker's Theorem** (see Proposition A.2).

THEOREM 2.6 (DONSKER'S THEOREM)

Let $\{X_n : n \geq 1\}$ be a sequence of i.i.d. random variables with $\sigma^2 = \text{Var}[X_i] < \infty$. Set $S_0 = 0$ and $S_n = X_1 + \dots + X_n$. Then,

$$\chi_n \Rightarrow \sigma B \quad \text{as } n \rightarrow \infty$$

in $D[0, \infty)$, equipped with the Skohorod topology.

Theorem 2.6 is a rigorous statement of the sense in which the random walk can be weakly approximated by Brownian motion. It asserts that if we look at large time scales, then spatial fluctuations of the order of the square root of time can be approximated by Brownian motion.

Now consider the following problem, which exemplifies how one can use such an approximation. Suppose that $S_n = X_1 + \cdots + X_n$ with $\mu = EX_i$, $\sigma^2 = \text{Var}[X_i] < \infty$, and X_1, X_2, \dots are i.i.d. To develop an approximation to $M_n = \max_{1 \leq k \leq n} S_k$, Theorem 2.5 suggests approximating S via Brownian motion:

$$S_{\lfloor t \rfloor} \stackrel{\mathcal{D}}{\approx} \mu t + \sigma B(t) \quad \text{for } t \geq 0.$$

Hence,

$$M_n \stackrel{\mathcal{D}}{\approx} \max_{0 \leq s \leq n} [\mu s + \sigma B(s)]. \quad (2.14)$$

To make a rigorous mathematical statement out of (2.14), let us assume that $\mu = 0$ (if $\mu \neq 0$, $M_n \approx \max(0, \mu n)$ is explained primarily by the drift of the random walk). Then, observe that

$$\max_{0 \leq s \leq n} B(s) = \max_{0 \leq s \leq 1} B(ns) \stackrel{\mathcal{D}}{=} \sqrt{n} \max_{0 \leq s \leq 1} B(s).$$

Also, $M_n/\sqrt{n} = \max_{0 \leq s \leq 1} \chi_n(s) = g(\chi_n)$, where $g : D[0, \infty) \rightarrow \mathfrak{R}$ is given by $g(x) = \max_{0 \leq t \leq 1} x(t)$. It is clear that $g(\chi_n) \rightarrow g(\chi)$ whenever

$$\sup_{0 \leq t \leq u} |\chi_n(t) - \chi(t)| \rightarrow 0$$

as $n \rightarrow \infty$ for each $u \geq 0$, so evidently $P(\sigma B \in D_g) = 0$ (see Appendix D). Hence, Proposition A.4 applies, yielding the following proposition:

PROPOSITION 2.1

If S satisfies the hypothesis of Theorem 2.6 and $\mu = 0$, then

$$\frac{1}{\sqrt{n}} \max_{0 \leq k \leq n} S_k \Rightarrow \sigma \max_{0 \leq k \leq n} B(s) \quad \text{as } n \rightarrow \infty \quad (2.15)$$

There are many other such approximations for the random walk that can be obtained by a combination of Theorem 2.6 and the continuous mapping principle. Note that the approximation (2.14) depends on the random walk only through its mean and variance, and is invariant to all the other characteristics of the increments. As a consequence, Theorem 2.6 is often referred to as an **invariance principle**.

EXERCISE 2.6

Suppose that $g : D[0, \infty) \rightarrow \mathfrak{R}$ is given by

$$g(x) = \sup_{t \geq 0} |x(t) - x(t^-)|.$$

- (a) Show that if x_n converges uniformly to x on compact sets, $g(x_n)$ need not converge to $g(x)$.
- (b) Show, by example, that $g(\chi_n)$ need not converge weakly to $g(\sigma B)$.
- (c) Suppose that we modify g as follows:

$$g(x) = \sup_{0 \leq t \leq 1} |x(t) - x(t^-)|.$$

Show that under the assumptions of Theorem 2.6, $g(\chi_n) \Rightarrow 0$ as $n \rightarrow \infty$.

2.3 The Single-Server Queue in Heavy Traffic.

Systems in which resource contention takes place are often modelled as queues. The most fundamental non-trivial queue is the G/G/1 queue, in which the interarrival times U_1, U_2, \dots are assumed i.i.d. and independent of the processing time sequence V_0, V_1, \dots , which is also assumed to be i.i.d. Typically, one requires that the waiting room have infinite capacity, and that the queue uses a “first come, first serve” queue discipline; we shall do both here.

One process of great interest is the waiting time sequence $W = \{W_n : n \geq 0\}$. Specifically, W_n is the time that customer n spends in the system waiting (exclusive of service).

The dynamics of W are governed by a recursion. In particular, note that if

$$\begin{aligned} A_n &= \text{arrival time of customer } n \\ D_n &= \text{departure time of customer } n, \end{aligned}$$

then $W_{n+1} = D_n - A_{n+1}$ if $D_n \geq A_{n+1}$, and zero otherwise. In other words,

$$W_{n+1} = [D_n - A_{n+1}]^+,$$

where $[x]^+ \triangleq \max(x, 0)$. But clearly, $D_n = A_n + W_n + V_n$, so

$$W_{n+1} = [W_n + V_n - U_{n+1}]^+,$$

which is the desired recursion. Set $X_n = V_{n-1} - U_n$, and note that the X_j 's are i.i.d. under our assumptions.

To approximate $W = \{W_n : n \geq 0\}$ by some functional of Brownian motion, we need to express it in terms of a random walk. With this in mind, suppose that $W_0 = 0$. Then

$$\begin{aligned} W_1 &= [X_1]^+ = \max(0, X_1) \\ W_2 &= [W_1 + X_2]^+ \\ &= \max(0, W_1 + X_2) \\ &= \max(\max(0, X_1) + X_2, 0) \\ &= \max(\max(X_2, X_1 + X_2), 0) \\ &= \max(X_1 + X_2, X_2, 0) \\ &= \max(S_2, S_2 - S_1, S_2 - S_2). \end{aligned}$$

An easy induction shows that $W_n = \max_{0 \leq k \leq n} (S_n - S_k)$.

Based on Theorem 2.5, the obvious approximation to W is

$$W_{[t]} \stackrel{\mathcal{D}}{\approx} \max_{0 \leq s \leq t} [\mu t + \sigma B(t) - \mu s - \sigma B(s)] = \mu t + \sigma B(t) - \min_{0 \leq s \leq t} [\mu s + \sigma B(s)] \triangleq Z(t). \quad (2.16)$$

The process $Z = \{Z(t) : t \geq 0\}$ is called a **reflected Brownian motion** (or regulated Brownian motion). The process $L(t) = -\min_{0 \leq s \leq t} [\mu s + \sigma B(s)]$ is a non-decreasing process which “pushes” Z back into $[0, \infty)$ whenever $\mu e + \sigma B$ attempts to go negative. Consequently, (2.16) provides an approximation to W via reflected Brownian motion. This approximation is widely used in practice.

When would one expect (2.16) to be a good approximation? This has to do with the conditions under which one can prove a rigorous limit theorem for W . Intuitively, reflected Brownian motion should be a good approximation when the queue spends relatively little idle time (for then it is undergoing a lot of large excursions that are explained well by Brownian motion). This occurs when the work arriving to the system is close to equilibrium with the capacity, i.e. $EV \approx EU$.

Let ρ , the utilization factor, be given by the ratio of the arrival rate to the service rate. In order that the work-in-system not “blow up”, it is necessary that $\rho < 1$. Then, the reflected Brownian motion (RBM) approximation should be good when $\rho \nearrow 1$. This asymptotic regime is called “heavy traffic”.

It is known that for the M/M/1 queue (the special case of G/G/1 in which the arrival process is Poisson and the processing times are exponential), the work-in-system grows as $(1 - \rho)^{-1}$ as $\rho \nearrow 1$. Because of the temporal/spatial scaling we have already seen in Donsker's theorem, this suggests that our limit theorem should involve a corresponding time scale of order $(1 - \rho)^{-2}$. It remains to prove this rigorously.

To accomplish this, we need to consider a family of queues, indexed by ρ . In the ρ th system, the inter-arrival times U_1, U_2, \dots are i.i.d., whereas the processing times are $V_0(\rho), V_1(\rho), \dots$ (also i.i.d.). We assume that the $V_i(\rho)$'s are obtained by scaling (this is not necessary, but it simplifies the math). So,

$$V_i(\rho) = \rho V_i.$$

We require that $\sigma_U^2 = \text{Var}[U_1] < \infty$, $\sigma_V^2 = \text{Var}[V_1] < \infty$, and $EU = EV$ (so that the system will be in heavy traffic as $\rho \nearrow 1$). Let $X_i(\rho) = V_{i-1}(\rho) - U_i$, $S_n(\rho) = X_1(\rho) + \dots + X_n(\rho)$, and

$$W_n(\rho) = S_n(\rho) - \min_{0 \leq k \leq n} S_k(\rho).$$

Because of the finite variance assumptions, we can invoke Theorem 2.5 twice (once for the U_i 's, once for the V_i 's), yielding

$$\begin{aligned} \sum_{i=1}^n U_i &= nEU + \sigma_U B_1(n) + o(\sqrt{n}) \quad \text{a.s. and} \\ \sum_{i=1}^n V_i &= nEV + \sigma_V B_2(n) + o(\sqrt{n}) \quad \text{a.s.} \end{aligned}$$

as $n \rightarrow \infty$, where B_1 and B_2 are independent standard Brownian motions. So,

$$\begin{aligned} W_n(\rho) &= \rho\sigma_V B_2(n) - \sigma_U B_1(n) - n(1 - \rho)EU \\ &\quad - \min_{0 \leq k \leq n} [\rho\sigma_V B_2(k) - \sigma_U B_1(k) - k(1 - \rho)EU] + o(\sqrt{n}) \quad \text{a.s.} \\ &= \rho\sigma_V B_2(n) - \sigma_U B_1(n) - n(1 - \rho)EU \\ &\quad - \min_{0 \leq s \leq n} [\rho\sigma_V B_2(s) - \sigma_U B_1(s) - s(1 - \rho)EU] + o(\sqrt{n}) \quad \text{a.s.} \end{aligned}$$

(See Exercise 2.4 and 2.5 to see how to replace the “discrete minimum” by a “continuous minimum”.) Then, for $u \geq 0$,

$$\begin{aligned} \sup_{0 \leq t \leq u} \left| (1 - \rho)W_{\lfloor t/(1-\rho)^2 \rfloor} - (1 - \rho) \left\{ \sigma_U B_1(t/(1 - \rho)^2) - \rho\sigma_V B_2(t/(1 - \rho)^2) - t(1 - \rho)EU \right. \right. \\ \left. \left. - \min_{0 \leq s \leq t/(1-\rho)^2} [\sigma_U B_1(s) - \rho\sigma_V B_2(s) - s(1 - \rho)EU] \right\} \right| \rightarrow 0 \quad \text{a.s.} \quad (2.17) \end{aligned}$$

EXERCISE 2.7

Let B_1 and B_2 be two independent standard Brownian motions. Prove that

$$a_1 B_1 + a_2 B_2 \stackrel{\mathcal{D}}{=} \sqrt{a_1^2 + a_2^2} B.$$

From the above exercise, it follows that

$$(1 - \rho)\sigma_U B_1(\cdot/(1 - \rho)^2) - (1 - \rho)\rho\sigma_V B_2(\cdot/(1 - \rho)^2) \stackrel{\mathcal{D}}{=} \sqrt{\sigma_U^2 + \rho^2 \sigma_V^2} B(\cdot).$$

Then, for f uniformly continuous,

$$Ef((1 - \rho)W_{\lfloor t/(1-\rho)^2 \rfloor}) = Ef(\sqrt{\sigma_U^2 + \rho^2 \sigma_V^2} B(t) - tEU - \min_{0 \leq s \leq t} [\sqrt{\sigma_U^2 + \rho^2 \sigma_V^2} B(s) - sEU]) \rightarrow 0$$

as $\rho \nearrow 1$. Since $\sqrt{\sigma_U^2 + \rho^2 \sigma_V^2} B(\cdot) \rightarrow \sqrt{\sigma_U^2 + \sigma_V^2} B(\cdot)$ a.s. in $D[0, \infty)$, we arrive at the following theorem.

THEOREM 2.7

Under the conditions given above,

$$(1 - \rho)W_{\lfloor \cdot / (1 - \rho)^2 \rfloor} \Rightarrow Z(\cdot)$$

in $D[0, \infty)$ as $\rho \nearrow 1$, where

$$Z(t) = \sqrt{\sigma_U^2 + \rho^2 \sigma_V^2} B(t) - tEU - \min_{0 \leq s \leq t} \left[\sqrt{\sigma_U^2 + \rho^2 \sigma_V^2} B(s) - sEU \right]$$

An alternative proof is to use Donsker's theorem and to use the representation (2.16) of W in terms of a random walk to invoke Proposition A.4.

Theorem 2.7 asserts that RBM gives good approximations to W so long as $\rho = EV/EU$ is close to 1, the time horizon under consideration is of order $(1 - \rho)^{-2}$, and spatial variations of order $(1 - \rho)^{-1}$ are of interest.

EXERCISE 2.8

Let $\{X_n : n \geq 1\}$ and X be random elements of $D[0, \infty)$, with the property that there exist steady-state limits $\{X_n(\infty) : n \geq 1\}$ and $X(\infty)$ for which $X_n(t) \Rightarrow X_n(\infty)$ and $X(t) \Rightarrow X(\infty)$ as $t \rightarrow \infty$. Suppose that $X_n \Rightarrow X$ in $D[0, \infty)$. Show, by example, that $X_n(\infty)$ need not converge weakly to $X(\infty)$ as $n \rightarrow \infty$. In other words, the steady-state of X need not be an approximation to the steady-state of X_n . (However, it turns out that for queues in “heavy traffic”, the steady-state approximation is typically valid.)

The RBM approximation considered here arises in many other contexts as well. For example, it arises in the insurance claims setting when the claims payout rate is close to the rate at which premiums are being collected. It is also relevant to the analysis of the CUSUM statistical test that arises in “changepoint” problems (also known as “quickest detection” problems).

EXERCISE 2.9

Let W be the waiting time sequence corresponding to a G/G/1 queue in which the inter-arrival times $\{U_n : n \geq 1\}$ are i.i.d. and independent of the i.i.d. processing times $\{V_n : n \geq 0\}$. Suppose that $\text{Var } U < \infty$, $\text{Var } V < \infty$, and $EU = EV$, so that the system is precisely balanced. Set $W_0 = 0$.

- (a) Prove that $\sqrt{n}W_n \Rightarrow \sigma \max_{0 \leq t \leq 1} B(t)$, where $\sigma^2 = \text{Var } U + \text{Var } V$.
- (b) Provide an explicit approximation for $P(W_n > x)$, valid for large n (and large x too).
- (c) Prove that $\sqrt{n}W_{\lfloor n \cdot \rfloor} \Rightarrow \sigma B(\cdot) - \min_{0 \leq u \leq \cdot} \sigma B(u)$ in $D[0, \infty)$ as $n \rightarrow \infty$.

EXERCISE 2.10

Prove that

$$\sigma B(\cdot) - \min_{0 \leq u \leq \cdot} \sigma B(u) \stackrel{\mathcal{D}}{=} |\sigma B(\cdot)|.$$

(Hint: Prove that both processes are Markov with the same transition density.)

2.4 A Price Process.

Let $\{\beta_n : n \geq 1\}$ be a discrete-time stochastic process in which β_n equals the price of some security on day n . It is natural to write β_n in the form

$$\beta_n = \beta_0 R_1 R_2 \cdots R_n, \tag{2.18}$$

where R_i is the relative change in price on day i , $i \geq 1$. The simplest possible stochastic model is to assume that the R_i 's are i.i.d. positive random variables. Taking logarithms in (2.18), we get

$$\log \beta_n = \log \beta_0 + \sum_{i=1}^n \log R_i. \tag{2.19}$$

Relation (2.19) asserts that $\{\log \beta_n : n \geq 1\}$ evolves as a random walk. Let

$$\begin{aligned}\mu &= E[\log R_i] \\ \sigma^2 &= \text{Var}[\log R_i] < \infty.\end{aligned}$$

Then, Theorem 2.5 suggests approximating the logarithm of β_n via a Brownian motion:

$$\log \beta_{[t]} \stackrel{\mathcal{D}}{\approx} \log \beta_0 + \mu t + \sigma B(t),$$

which in turn yields the approximation

$$\beta_{[t]} \stackrel{\mathcal{D}}{\approx} \beta_0 e^{\mu t + \sigma B(t)}, \quad (2.20)$$

Set $Y(t) = Y(0) \exp(\mu t + \sigma B(t))$, where $Y(0) = \beta_0$. The process $Y = \{Y(t) : t \geq 0\}$ is called a **geometric Brownian motion** (or, alternatively, a log-normal process). The process Y has the following properties:

- (i) For $t_1 < t_2 < \dots < t_n$, the random variables $Y(t_1)/Y(0), \dots, Y(t_n)/Y(t_{n-1})$ are independent.
- (ii) For $0 < s < t$, $Y(t)/Y(s) \stackrel{\mathcal{D}}{=} Y(t-s)/Y(0)$.
- (iii) $Y(t)/Y(0) \stackrel{\mathcal{D}}{=} \text{lognormal}(\mu t, \sigma^2 t)$, i.e. $\log(Y(t)/Y(0)) \stackrel{\mathcal{D}}{=} N(\mu t, \sigma^2 t)$
- (iv) $Y = \{Y(t) : t \geq 0\}$ has continuous paths.

From an approximation viewpoint, it is important to recognize that the exponentiation in (2.20) also exponentiates the “error” in the approximation of the random walk by Brownian motion. Note, for example, that while the Brownian approximation appearing in Donsker’s theorem matches the mean and variance of the random walk, this is not true for (2.20):

$$E\beta_{[t]} = \beta_0 (E[R_1])^{[t]},$$

whereas $EY(t) = \beta_0 \exp(\mu t + \sigma^2 t/2)$.

EXERCISE 2.11

Show, by example, that typically there exists no random variable Z for which

$$\frac{\beta_{[nt]} - E\beta_{[nt]}}{\sqrt{\text{Var} \beta_{[nt]}}} \Rightarrow Z \quad \text{as } n \rightarrow \infty.$$

EXERCISE 2.12

Show how deceptive the approximation (2.20) is. From a rigorous viewpoint, no approximation of the price process itself by Brownian motion (or a functional thereof) is typically possible. It is only the logarithmic price process that can be well-approximated by Brownian motion.

EXERCISE 2.13

Simulate $\{\beta_n : n \geq 1\}$ for some i.i.d. sequence $\{R_n : n \geq 1\}$. Compare the results with what geometric Brownian motion predicts based on (2.20).

EXERCISE 2.14

Even the Brownian approximation to random walks can be poor out in the tails (more than order \sqrt{n} from the mean of the walk at time n). To see this, simulate a binomial(n, p) random walk and compare $P(B(n, p) > np + k\sqrt{p(1-p)n})$ to that predicted by Brownian motion for $k = 1, 5, 10$ and 50 .

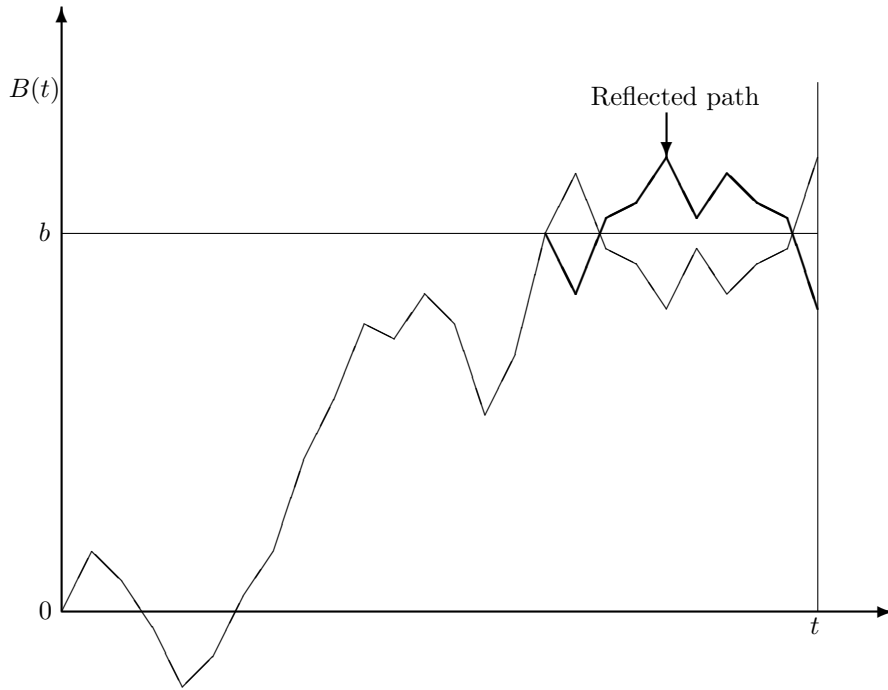


Figure 2.1: The Reflection Principle of Brownian Motion.

2.5 The Reflection Principle.

The approximation theorems for random walk and for queues provide important qualitative insight:

- (1) **Donsker's theorem:** Over large time scales, the random fluctuations occur continuously over spatial scales of order the square root of time. Furthermore, these random fluctuations depend on the underlying random walk only through its mean and variance.
- (2) **Heavy-traffic limit theorem:** If $\rho \approx 1$, the random fluctuations occur over time scales of order $(1 - \rho)^{-2}$ and involve spatial variations of magnitude $(1 - \rho)^{-1}$. Furthermore, these fluctuations depend on the underlying interarrival and processing times only through the mean and variance.

But an equally important reason for the use of Brownian approximations is that one expects them to be computationally and analytically more tractable. Let's illustrate this in computing the distribution of

$$T(b) = \inf\{t \geq 0 : B(t) \geq b\} \quad \text{for } b > 0.$$

Note that

$$P(T(b) \leq t) = P\left(\max_{0 \leq u \leq t} B(u) \geq b\right),$$

so this computation will also provide a means of computing the distribution of

$$\max_{0 \leq u \leq t} B(u).$$

Consider figure 2.1: For every path that hits b and ends up above b , there is a corresponding path that hits b and ends up below b . So, the likelihood of a path that hits b (before t) is twice the likelihood of a path that hits b and ends up above b . However, every path that ends up above b must necessarily pass through b at some prior time (by path continuity). We conclude, at least intuitively, that

$$P(T(b) \leq t) = 2P(B(t) \geq b).$$

This is the **reflection principle**, and relies heavily on the path continuity and symmetry of Brownian motion.

To prove this property more rigorously, we note that B enjoys the **Markov property**: For $t_1 < t_2 < \dots < t_n$,

$$P(B(t_n + \cdot) \in A | B(t_1), \dots, B(t_n)) = P_{B(t_n)}(A), \quad (2.21)$$

where $P_x(A) = P(B(\cdot) \in A | B(0) = x)$. It turns out that $T(b)$ is a special type of random time (known as a stopping time) for which (2.21) extends from deterministic horizon histories to random time horizon histories:

$$P(B(T(b) + \cdot) \in A | B(u) : 0 \leq u \leq T(b)) = P_{B(T(b))}(A) (= P_b(A)). \quad (2.22)$$

(More about this later; (2.22) involves an invocation of the **strong Markov property**.) Then, we note that

$$\begin{aligned} P(T(b) \leq t) &= P(T(b) \leq t, B(t) \geq b) + P(T(b) \leq t, B(t) < b) \\ &= P(B(t) \geq b) + P(T(b) \leq t, B(t) < b). \end{aligned} \quad (2.23)$$

But

$$\begin{aligned} &P(T(b) \leq t, B(t) < b) \\ &= E[P(B(t) < b | B(u) : 0 \leq u \leq T(b)) \cdot I(T(b) \leq t)] \\ &= \int_0^\infty P(B(t) < b | T(b) = s) P(T(b) \in ds) \\ &= \int_0^\infty P(B(T(b) + t - s) < b | T(b) = s) P(T(b) \in ds) \\ &= \int_0^\infty E[P(B(T(b) + t - s) < b | B(u) : 0 \leq u \leq T(b)) | T(b) = s] P(T(b) \in ds) \\ &= \int_0^\infty P_b(B(t - s) < b) P(T(b) \in ds) \\ &= \frac{1}{2} P(T(b) \leq t). \end{aligned}$$

Our conclusions are summarized in Proposition 2.2 and Corollary 2.3:

PROPOSITION 2.2

$$P(T(b) \leq t) = 2P(B(t) > b).$$

COROLLARY 2.3

$$P\left(\max_{0 \leq u \leq t} B(u) \geq b\right) = 2P(B(t) > b).$$

EXERCISE 2.15

Prove that $E\Gamma_1^p < \infty$ for all $p > 0$, where $\Gamma_1 = \max_{0 \leq u \leq t} |B(u)|$. (This provides one avenue to establishing Exercise 2.5)

EXERCISE 2.16

(a) Prove that $B(n)/n \rightarrow 0$ a.s. as $n \rightarrow \infty$ using the strong law of large numbers.

(b) Prove that $\max_{1 \leq k \leq n} B(k) \rightarrow 0$ a.s. as $n \rightarrow \infty$ (see Exercise 2.15).

(c) Use (a) and (b) to prove that $B(t)/t \rightarrow 0$ a.s. as $t \rightarrow \infty$.

EXERCISE 2.17

Let $\tilde{B}(t) = tB(1/t)$ for $t > 0$ and set $\tilde{B}(0) = 0$. Use Exercise 2.16 to prove that $\tilde{B} \stackrel{\mathcal{D}}{=} B$.

EXERCISE 2.18

Let $L(b) = \sup\{t \geq 0 : B(t)/t \geq b\}$. Use Exercise 2.17 to prove that $L(b)$ is well-defined and finite, and compute the distribution of $L(b)$.

We can also use the reflection principle to compute the joint distribution of the maximum of B over $[0, t]$ and $B(t)$ itself. Note that if $x > y > 0$, then

$$\begin{aligned} P(B(t) < x, \max_{0 \leq u \leq t} B(u) > y) &= P(\max_{0 \leq u \leq t} B(u) > y) - P(B(t) > x, \max_{0 \leq u \leq t} B(u) > y) \\ &= 2P(B(t) > y) - P(B(t) > x). \end{aligned}$$

On the other hand, if $x < y$ with $y > 0$, then

$$\begin{aligned} P(B(t) < x, \max_{0 \leq u \leq t} B(u) > y) &= P(B(t) < x, T(y) \leq t) \\ &= \int_0^t P(B(t) < x | T(y) = s) P(T(y) \in ds) \\ &= \int_0^t P_y(B(t-s) < x) P(T(y) \in ds) \\ &= \int_0^t P_y(B(t-s) > y + (y-x)) P(T(y) \in ds) \\ &= P(B(t) > (2y-x), T(y) \leq t) \\ &= P(B(t) > (2y-x)). \end{aligned}$$

(The reflection principle, based on the symmetry of B , was used in the third-to-last equality.)

PROPOSITION 2.4

If $x > y > 0$, then

$$P(B(t) < x, \max_{0 \leq u \leq t} B(u) > y) = 2P(B(t) > y) - P(B(t) > x),$$

whereas if $x < y$ with $y > 0$, then

$$P(B(t) < x, \max_{0 \leq u \leq t} B(u) > y) = P(B(t) > (2y-x)).$$

EXERCISE 2.19

Use Proposition 2.4 to compute $P_x(\max_{0 \leq u \leq t} B(u) > z | B(t) = y)$ for $z > (x \vee y)$.

EXERCISE 2.20

Use Exercise 2.19 to prove that when $xy > 0$,

$$P_x(B \text{ has no zero in } [0, t] | B(t) = y) = 1 - e^{-2xy/t}$$

EXERCISE 2.21

Prove that

$$P\{(B(u) : 0 \leq u \leq t) \in \cdot | B(t) = y\} = P\{(B(u) + uy/t : 0 \leq u \leq t) \in \cdot | B(t) = 0\}.$$

EXERCISE 2.22

(a) Use Exercise 2.21 to prove that

$$P(\max_{0 \leq u \leq t} B(u) > z | B(t) = y) = P(\max_{0 \leq u \leq t} B(u) > z | B(t) = 0).$$

- (b) Use Exercise 2.19 and part (a) to compute $P(\max_{0 \leq u \leq t} (B(u) - \mu u) > z | B(t) = 0)$ for arbitrary μ .
- (c) Prove that

$$P(B(t) - \mu t - \max_{0 \leq s \leq t} (B(s) - \mu s) > z) = P(\max_{0 \leq u \leq t} (B(u) - \mu u) > z).$$

The “reflection principle” can be used to further derive the joint distribution of the minimum, maximum, and Brownian motion itself. For $y \leq 0 \leq z$ and $u < v$, it can be shown that

$$\begin{aligned} & P\{y < \min_{0 \leq s \leq t} B(s) < \max_{0 \leq s \leq t} B(s) < z, u < B(t) < v\} \\ &= \sum_{k=-\infty}^{\infty} P\{u + 2k(z - y) < B(t) < v + 2k(z - y)\} \\ &= \sum_{k=-\infty}^{\infty} P\{2z - v + 2k(z - y) < B(t) < 2z - u + 2k(z - y)\} \end{aligned}$$

See pp. 78-79 of (2).

EXERCISE 2.23

Prove that

$$P(\max_{0 \leq s \leq t} |B(s)| \leq z) = \sum_{k=-\infty}^{\infty} (-1)^k P((2k - 1)z < B(t) < (2k + 1)z)$$

2.6 Additional Properties of Brownian Motion.

Here, we record some additional properties of Brownian motion that are useful.

DEFINITION 2.5

A **Gaussian random vector** is an \mathbb{R}^d -valued random element having a joint (multivariate) normal distribution.

DEFINITION 2.6

A stochastic process $X = \{X(t) : t \geq 0\}$ is said to be **Gaussian** if for each $t_1 < t_2 < \cdots < t_n$, the vector $(X(t_1), X(t_2), \dots, X(t_n))$ is a Gaussian random vector.

Brownian motion is a Gaussian process. Since Gaussian random vectors are characterized by their means and covariance matrices, it follows that the finite-dimensional distributions of a Gaussian process X are characterized by its mean function $m(t) = EX(t)$ and its covariance function $c(s, t) = \text{Cov}(X(s), X(t))$. For Brownian motion with drift μ and variance σ^2 ,

$$\begin{aligned} m(t) &= \mu t \\ c(s, t) &= \sigma^2(s \wedge t). \end{aligned}$$

Since Brownian motion is Gaussian, we expect any “linear transformation” of B to be Gaussian. For example, consider $I = \int_0^1 B(t)dt$. Since B is continuous, we may define the integral as a Riemann integral. To prove that I is Gaussian, note that

$$I = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} B\left(\frac{j}{n}\right) \quad \text{a.s.}$$

Of course, $I_n = n^{-1} \sum_{j=0}^{n-1} B(j/n)$ is a Gaussian random variable with mean zero and variance

$$\sigma_n^2 = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \text{Cov}\left(B\left(\frac{j}{n}\right), B\left(\frac{k}{n}\right)\right) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{j=0}^{n-1} \frac{j}{n} + \lim_{n \rightarrow \infty} \frac{2}{n^2} \sum_{j=0}^{n-1} \sum_{k=j+1}^{n-1} \frac{j}{n}.$$

It is easily seen that $\sigma_n^2 \rightarrow 1/3$ as $n \rightarrow \infty$. Hence, the Bounded Convergence theorem ensures that

$$Ee^{i\theta I} = E \lim_{n \rightarrow \infty} e^{i\theta I_n} = \lim_{n \rightarrow \infty} Ee^{i\theta I_n} = \lim_{n \rightarrow \infty} Ee^{-\theta^2 \sigma_n^2/2} = e^{-\theta^2/6} = Ee^{-i\theta N(0,1/3)}.$$

It follows that $I \stackrel{\mathcal{D}}{=} N(0, 1/3)$ (using the fact that characteristic functions uniquely determine the distribution of a random variable).

PROPOSITION 2.5

$$\int_0^t B(t)dt \stackrel{\mathcal{D}}{=} N(0, 1/3).$$

EXERCISE 2.24

Prove that if S is a random walk that has mean zero with i.i.d. increments having finite variance σ^2 , then

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} S_j \Rightarrow \sigma N(0, 1/3) \quad \text{as } n \rightarrow \infty.$$

A “tied down” Brownian motion is a Brownian motion that is conditioned to be at the origin at time t . By viewing Brownian motion as a Gaussian process, we can find an alternative representation of “tied down” Brownian motion that removes the conditioning.

EXERCISE 2.25

Let $\tilde{B}(t) = B(t) - tB(1)$.

(a) Prove that for $t_1 < t_2 < \dots < t_n < t$,

$$P\{(B(t_1), \dots, B(t_n)) | B(t) = 0\} = P\{(\tilde{B}(t_1), \dots, \tilde{B}(t_n))\}.$$

(b) Prove that $P\{(B(u) : 0 \leq u \leq t) \in \cdot | B(t) = 0\} = P\{(\tilde{B}(u) : 0 \leq u \leq t) \in \cdot\}$.

(c) Compute $P(\int_0^1 B(t)dt \in \cdot | B(1) = 0)$.

DEFINITION 2.7

The process $\tilde{B} = \{\tilde{B}(u) : 0 \leq u \leq t\}$ is called a **Brownian bridge**, where

$$\tilde{B}(u) = B(u) - uB(1).$$

There is also an interesting representation of the **Ornstein-Uhlenbeck** process in terms of Brownian motion.

DEFINITION 2.8

The process $V = \{V(t) : t \geq 0\}$ having continuous sample paths and Gaussian finite-dimensional distributions is said to be a (zero mean) stationary Ornstein-Uhlenbeck process if V is stationary with covariance function

$$c(s, t) = \sigma^2 \exp(-a|t - s|) \quad \text{for } \sigma^2, a > 0.$$

EXERCISE 2.26

Set $B^*(t) = re^{-ut}B(e^{2ut})$ for $t \geq 0$. Show that $B^* \stackrel{\mathcal{D}}{=} V$, where V is Ornstein-Uhlenbeck.

2.7 Generalizations of Donsker's Theorem.

In the pricing and queueing contexts, it may be that it is unreasonable to assume that the increments of the underlying random walk are i.i.d. with finite variance. To be precise, let $S = \{S_n : n \geq 0\}$ be a stochastic sequence (called a “generalized random walk”) with $S_0 = 0$ and stationary increments $\{X_n : n \geq 1\}$.

DEFINITION 2.9

A sequence $\{X_n : n \geq 1\}$ is said to be **stationary** if for $m \geq 0$,

$$(X_0, X_1, \dots) \stackrel{\mathcal{D}}{=} (X_m, X_{m+1}, \dots).$$

A process $\{X(t) : t \geq 0\}$ is said to be **stationary** if for $u \geq 0$,

$$\{X(t) : t \geq 0\} \stackrel{\mathcal{D}}{=} \{X(u+t) : t \geq 0\}.$$

Such an increments process can exhibit no deterministic seasonal effects or growth effects.

To start, suppose that the increments have finite variance. Then, the autocorrelation function $(\rho(n) : n \geq 0)$ is well-defined, where $\rho(n)$ is the coefficient of correlation between X_k and X_{k+n} :

$$\rho(n) = \text{Cov}(X_k, X_{k+n}) / \text{Var}[X_k].$$

Let $\chi_n(t) = (S_{\lfloor nt \rfloor} - \mu \lfloor nt \rfloor) / \sqrt{n}$, where $S_n = X_1 + \dots + X_n$, $\mu = EX_1$, and $\sigma^2 = \text{Var}[X_1]$. Typically, it turns out that Donsker's theorem continues to hold when

$$\sum_{j=0}^{\infty} |\rho(j)| < \infty. \quad (2.24)$$

A process X satisfying (2.24) is said to exhibit **short-range dependence**. When (2.24) holds, it usually follows that $\chi_n \Rightarrow \eta B$ in $D[0, \infty)$ as $n \rightarrow \infty$, where $\eta^2 = \sigma^2(1 + 2 \sum_{j=0}^{\infty} |\rho(j)|)$ (see p. 174 of (2), for a typical such result).

So, what happens if the increments exhibit **long-range dependence**? This is the case in which $\rho(n) \rightarrow 0$ as $n \rightarrow \infty$, but $\sum_{j=0}^{\infty} |\rho(j)| = +\infty$. For example, suppose that

$$\rho(n) \sim cn^{2h-2} \quad \text{as } n \rightarrow \infty \text{ with } c > 0 \text{ and } 0.5 < H < 1 \quad (2.25)$$

(under (2.25), the autocorrelations go to zero slowly; the parameter H appearing in (2.25) is called the **Hurst parameter**).

In this setting, the random walk is not well-approximated by Brownian motion. The autocorrelations are so slowly decaying that S can not be approximated by an independent increments process. Set

$$\tilde{\chi}_n(t) = \frac{S_{\lfloor nt \rfloor} - \mu \lfloor nt \rfloor}{n^H}$$

Then, under certain (technical) conditions, $\tilde{\chi}_n \Rightarrow \eta B_H$ as $n \rightarrow \infty$, where $B_H = \{B_H(t) : t \geq 0\}$ is a so-called **fractional Brownian motion**. Here,

$$\eta^2 = \sigma^2 c H^{-1} (2H^{-1} - 1)^{-1}$$

(see p. 337 of (18), for a supporting result). The process B_H is a zero-mean Gaussian process with covariance function

$$c(s, t) = 0.5\{|s|^{2H} + |t|^{2H} - |t - s|^{2H}\}.$$

EXERCISE 2.27

Prove that B_H is not a Markov process.

EXERCISE 2.28

Prove that B_H is a self-similar (fractal) process in the sense that $B_H(a \cdot) \stackrel{\mathcal{D}}{=} a^H B_H(\cdot)$ for $a > 0$.

The fact that B_H is non-Markov makes it exceedingly hard to work with (even simulating B_H is hard!)

Now, what happens if the finite variance assumption is violated? Specifically, suppose that S is a random walk with i.i.d. increments having infinite variance. Clearly, any approximation to S here must have stationary independent increments (it follows that any limit process here must be Markov). Random variables having infinite variance are called heavy-tailed random variables.

For example, suppose that

$$\begin{aligned} P(X_1 > x) &\sim ax^{-\alpha} \\ P(X_1 < -x) &\sim bx^{-\alpha} \end{aligned}$$

as $x \rightarrow \infty$, where $a, b > 0$ and $1 < \alpha < 2$. Then, set

$$\tilde{\chi}_n^*(t) = \frac{S_{\lfloor nt \rfloor} - \mu \lfloor nt \rfloor}{n^{1/\alpha}}$$

Under certain (technical) conditions, $\tilde{\chi}_n^*(t) \Rightarrow L$ as $n \rightarrow \infty$, where $L = \{L(t) : t \geq 0\}$ is a stationary independent increments process (also known as a **Levy process**) that typically does not have continuous sample paths. The loss of continuity is due to the extraordinarily large jumps that occur occasionally, due to the heavy tails.

REMARK 2.3 Some statisticians claim to have found evidence of long-range dependence and heavy tails in both financial data and various traces of communications network traffic.