18.3: The Brownian Bridge

Basic Theory

Definition and Constructions

In the most common formulation, the *Brownian bridge process* is obtained by taking a standard Brownian motion process $\boldsymbol{X}$, restricted to the interval $[0, 1]$, and conditioning on the event that $X_1 = 0$. Since $X_0 = 0$ also, the process is “tied down” at both ends, and so the process in between forms a “bridge” (albeit a very jagged one). The Brownian bridge turns out to be an interesting stochastic process with surprising applications, including a very important application to statistics. In terms of a definition, however, we will give a list of characterizing properties as we did for standard Brownian motion and for Brownian motion with drift and scaling.

A *Brownian bridge* is a stochastic process $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ with state space $\mathbb{R}$ that satisfies the following properties:

1. $X_0 = 0$ and $X_1 = 0$ (each with probability 1).
2. $\boldsymbol{X}$ is a Gaussian process.
3. $\mathbb{E}(X_t) = 0$ for $t \in [0, 1]$.
4. $\operatorname{cov}(X_s, X_t) = \min\{s, t\} - st$ for $s,\ t \in [0, 1]$.
5. With probability 1, $t \mapsto X_t$ is continuous on $[0, 1]$.

So, in short, a Brownian bridge $\boldsymbol{X}$ is a continuous Gaussian process with $X_0 = X_1 = 0$, and with mean and covariance functions given in (c) and (d), respectively. Naturally, the first question is whether there exists such a process. The answer is yes, of course, otherwise why would we be here? But in fact, we will see several ways of constructing a Brownian bridge from a standard Brownian motion. To help with the proofs, recall that a standard Brownian motion process $\boldsymbol{Z} = \{Z_t : t \in [0, \infty)\}$ is a continuous Gaussian process with $Z_0 = 0$, $\mathbb{E}(Z_t) = 0$ for $t \in [0, \infty)$ and $\operatorname{cov}(Z_s, Z_t) = \min\{s, t\}$ for $s,\ t \in [0, \infty)$. Here is our first construction:

Suppose that $\boldsymbol{Z} = \{Z_t : t \in [0, \infty)\}$ is a standard Brownian motion, and let $X_t = Z_t - tZ_1$ for $t \in [0, 1]$. Then $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ is a Brownian bridge.

Proof

1. Note that $X_0 = Z_0 = 0$ and $X_1 = Z_1 - Z_1 = 0$.
2. Linear combinations of the variables in $\boldsymbol{X}$ reduce to linear combinations of the variables in $\boldsymbol{Z}$ and hence have normal distributions. Thus $\boldsymbol{X}$ is a Gaussian process.
3. $E(X_t) = \mathbb{E}(Z_t) - t\mathbb{E}(Z_1) = 0$ for $t \in [0, 1]$
4. $\operatorname{cov}(X_s, X_t) = \operatorname{cov}(Z_s - sZ_1, Z_t - tZ_1) = \operatorname{cov}(Z_s, Z_t) - t\operatorname{cov}(Z_s, Z_1) - s\operatorname{cov}(Z_1, Z_t) + st\operatorname{cov}(Z_1, Z_1) = \min\{s, t\} - st - st + st$ for $s,\ t \in [0, 1]$.
5. $t \mapsto X_t$ is continuous on $[0, 1]$ since $t \mapsto Z_t$ is continuous on $[0, 1]$.

Let's see the Brownian bridge in action.

Run the simulation of the Brownian bridge process in single step mode a few times.

For the Brownian bridge $\boldsymbol{X}$, note in particular that X_t is normally distributed with mean 0 and variance $t(1-t)$ for $t \in [0, 1]$. Thus, the variance increases and then decreases on $[0, 1]$ reaching a maximum of $1/4$ at $t = 1/2$. Of course, the variance is 0 at $t = 0$ and $t = 1$, since $X_0 = X_1 = 0$ deterministically.

Open the simulation of the Brownian bridge process. Vary t and note the change in the probability density function and moments. For various values of t, run the simulation 1000 times and compare the empirical density function and moments to the true density function and moments.

Conversely to the construction above, we can build a standard Brownian motion on the time interval $[0, 1]$ from a Brownian bridge.

Suppose that $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ is a Brownian bridge, and suppose that Z is a random variable with a standard normal distribution, independent of $\boldsymbol{X}$. Let $Z_t = X_t + tZ$ for $t \in [0, 1]$. Then $\boldsymbol{Z} = \{Z_t : t \in [0, 1]\}$ is a standard Brownian motion on $[0, 1]$.

Proof

1. Note that $Z_0 = X_0 = 0$.
2. Linear combinations of the variables in $\boldsymbol{Z}$ reduce to linear combinations of the variables in $\boldsymbol{X}$ and hence have normal distributions. Thus $\boldsymbol{Z}$ is a Gaussian process.
3. $\mathbb{E}(Z_t) = \mathbb{E}(X_t) + t\mathbb{E}(Z) = 0$ for $t \in [0, 1]$.
4. $\operatorname{cov}(Z_s, Z_t) = \operatorname{cov}(X_s + sZ, X_t + tZ) = \operatorname{cov}(X_s, X_t) + t\operatorname{cov}(X_s, Z) + s\operatorname{cov}(X_t, Z) + st\operatorname{var}(Z) = \min\{s, t\} - st + 0 + 0 + st = \min\{s, t\}$ for $s,\ t \in [0, 1]$.
5. $t \mapsto Z_t$ is continuous on $[0, 1]$ since $t \mapsto X_t$ is continuous on $[0, 1]$.

Here's another way to construct a Brownian bridge from a standard Brownian motion.

Suppose that $\boldsymbol{Z} = \{Z_t : t \in [0, \infty)\}$ is a standard Brownian motion. Define $X_1 = 0$ and

$$X_t = (1-t) Z\left(\frac{t}{1-t}\right), \quad t \in [0, 1) \tag{18.3.1}$$

Then $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ is a Brownian bridge.

Proof

1. Note that $X_0 = Z_0 = 0$ and by definition, $X_1 = 0$.
2. Linear combinations of variables in $\boldsymbol{X}$ reduce to linear combinations of variables in $\boldsymbol{Z}$ and hence have normal distributions. Thus $\boldsymbol{X}$ is a Gaussian process.
3. For $t \in [0, 1]$,

$$\mathbb{E}(X_t) = (1-t)\mathbb{E}\left[Z\left(\frac{t}{1-t}\right)\right] = 0 \tag{18.3.2}$$

4. If $s, t \in [0, 1)$ with $s < t$ then $s/(1-s) < t/(1-t)$ so

$$\operatorname{cov}(X_s, X_t) = \operatorname{cov}\left[(1-s)Z\left(\frac{s}{1-s}\right), (1-t)Z\left(\frac{t}{1-t}\right)\right] = (1-s)(1-t)\frac{s}{1-s} = s(1-t) \tag{18.3.3}$$

5. Finally, $t \mapsto X_t$ is continuous with probability 1 on $[0, 1)$, and with probability 1, $X_t = (1-t)Z\left[t/(1-t)\right] \to 0$ as $t \uparrow 1$.

Conversely, we can construct a standard Brownian motion from a Brownian bridge.

Suppose that $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ is a Brownian bridge. Define

$$Z_t = (1+t) X\left(\frac{t}{1+t}\right), \quad t \in [0, \infty) \tag{18.3.4}$$

Then $\boldsymbol{Z} = \{Z_t : t \in [0, \infty)\}$ is a standard Brownian motion process.

Proof

1. Note that $Z_0 = X_0 = 0$
2. Linear combinations of the variables in $\boldsymbol{Z}$ reduce to linear combinations of the variables in X, and hence have normal distributions. Thus $\boldsymbol{Z}$ is a Gaussian process.
3. For $t \in [0, \infty)$,

$$\mathbb{E}(Z_t) = (1+t)\mathbb{E}\left[X\left(\frac{t}{1+t}\right)\right] = 0 \tag{18.3.5}$$

4. If $s, t \in [0, 1]$ with $s < t$ Then $s/(1+s) < t/(1+t)$ so

$$\operatorname{cov}(Z_s, Z_t) = \operatorname{cov}\left[(1+s)X\left(\frac{s}{1+s}\right), (1+t)X\left(\frac{t}{1+t}\right)\right] = (1+s)(1+t)\left[\frac{s}{1+s} - \frac{s}{1+s}\frac{t}{1+t}\right] = s \tag{18.3.6}$$

5. Since $t \mapsto X_t$ is continuous, $t \mapsto Z_t$ is continuous

We return to the comments at the beginning of this section, on conditioning a standard Brownian motion to be 0 at time 1. Unlike the previous two constructions, note that we are not transforming the random variables, rather we are changing the underlying *probability measure*.

Suppose that $\boldsymbol{X} = \{X_t : t \in [0, \infty)\}$ is a standard Brownian motion. Then conditioned on $X_1 = 0$, the process $\{X_t : t \in [0, 1]\}$ is a Brownian bridge process.

Proof

Part of the argument is based on properties of the multivariate normal distribution. The conditioned process is still continuous and is still a Gaussian process. In particular, suppose that $s, t \in [0, 1]$ with $s < t$. Then (X_t, X_1) has a joint normal distribution with parameters specified by the mean and covariance functions of $\boldsymbol{X}$. By standard computations, the conditional distribution of X_t given $X_1 = 0$ is normal with mean 0 and variance $t(1-t)$. Similarly, the joint distribution of (X_s, X_t, X_1) is normal with parameters specified by the mean and covariance functions of $\boldsymbol{X}$. Again, by standard computations, the conditional distribution of (X_s, X_t) given $X_1 = 0$ is bivariate normal with 0 means and with $\operatorname{cov}(X_s, X_t \mid X_1 = 0) = s(1-t)$.

Finally, the Brownian bridge can be defined in terms a stochastic integral

Suppose that $\boldsymbol{Z} = \{Z_t : t \in [0, \infty)\}$ is standard Brownian motions. Define $X_1 = 1$ and

$$X_t = (1-t)\int_0^t \frac{1}{1-s}\, dZ_s, \quad t \in [0, 1) \tag{18.3.7}$$

Then $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ is a Brownian bridge process.

Proof

1. Note that $X_0 = 0$ and by definition, $X_1 = 0$.
2. Since the integrand in the stochastic integral is deterministic, $\boldsymbol{X}$ is a Gaussian process.
3. $\boldsymbol{X}$ is continuous on $[0, 1)$ with probability 1, as a basic property of stochastic integrals. Moreover, $X_t \to 0$ as $t \uparrow 1$ as a consequence of the martingale inequality.
4. $\mathbb{E}(X_t) = 0$ since the stochastic integral has mean 0.
5. Suppose that $s, t \in [0, 1]$ with $s \le t$. Then

$$\operatorname{cov}(X_s, X_t) = \operatorname{cov}\left[(1-s)\int_0^s \frac{1}{1-u}\,dZ_u, (1-t)\left(\int_0^s \frac{1}{1-u}\,dZ_u + \int_s^t \frac{1}{1-u}\,dZ_u\right)\right] \tag{18.3.8}$$

But $\int_0^s \frac{1}{1-u}\,dZ_u$ and $\int_s^t \frac{1}{1-u}\,dZ_u$ are independent,

$$\operatorname{cov}(X_s, X_t) = (1-s)(1-t)\operatorname{var}\left(\int_0^s \frac{1}{1-u}\,dZ_u\right) \tag{18.3.9}$$

But then by the Ito isometry,

$$\operatorname{cov}(X_s, X_t) = (1-s)(1-t)\int_0^s \frac{1}{(1-u)^2}\,du = (1-s)(1-t)\left(\frac{1}{1-s} - 1\right) = (1-t)s \tag{18.3.10}$$

In differential form, the process above can be written as

$$dX_t = \frac{X_t}{1-t}\,dt + dZ_t,\ X_0 = 0 \tag{18.3.11}$$

The General Brownian Bridge

The processes constructed above (in several ways!) is the *standard* Brownian bridge. it's a simple matter to generalize the process so that it starts at a and ends at b, for arbitrary $a, b \in \mathbb{R}$.

Suppose that $\boldsymbol{Z} = \{Z_t : t \in [0, 1]\}$ is a standard Brownian bridge process. Let $a, b \in \mathbb{R}$ and define $X_t = (1-t)a + tb + Z_t$ for $t \in [0, 1]$. Then $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ is a *Brownian bridge process from a to b*.

Of course, any of the constructions above for standard Brownian bridge can be modified to produce a general Brownian bridge. Here are the characterizing properties.

The Brownian bridge process $\boldsymbol{X} = \{X_t : t \in [0, 1]\}$ from a to b is characterized by the following properties:

1. $X_0 = a$ and $X_1 = b$ (each with probability 1).
2. $\boldsymbol{X}$ is a Gaussian process.
3. $\mathbb{E}(X_t) = (1-t)a + tb$ for $t \in [0, 1]$.
4. $\operatorname{cov}(X_s, X_t) = \min\{s, t\} - st$ for $s, t \in [0, 1]$.
5. With probability 1, $t \mapsto X_t$ is continuous on $[0, 1]$.

Applications

The Empirical Distribution Function

We start with a problem that is one of the most basic in statistics. Suppose that T is a real-valued random variable with an unknown distribution. Let F denote the distribution function of T, so that $F(t) = \mathbb{P}(T \le t)$ for $t \in \mathbb{R}$. Our goal is to construct an estimator of F, so naturally our first step is to *sample* from the distribution of T. This generates a sequence $\boldsymbol{T} = (T_1, T_2, \ldots)$ of independent variables, each with the distribution of T (and so with distribution function F). Think of $\boldsymbol{T}$ as a sequence of independent *copies* of T. For $n \in \mathbb{N}_+$ and $t \in \mathbb{R}$, the natural estimator of $F(t)$ based on the first n sample values is

$$F_n(t) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(T_i \le t) \tag{18.3.12}$$

which is simply the proportion of the first n sample values that fall in the interval $(-\infty, t]$. Appropriately enough, F_n is known as the *empirical distribution function* corresponding to the sample of size n. Note that $(\mathbf{1}(T_1 \le t), \mathbf{1}(T_2 \le t), \ldots)$ is a sequence of independent, identically distributed indicator variables (and hence is a sequence of Bernoulli trials), and corresponds to sampling from the distribution of $\mathbf{1}(T \le t)$. The estimator $F_n(t)$ is simply the sample mean of the first n of these variables. The numerator, the number of the original sample variables with values in $(-\infty, t]$, has the binomial distribution with parameters n and $F(t)$. Like all sample means from independent, identically distributed samples, $F_n(t)$ satisfies some basic and important properties. A summary is given below, but to make sense of some of these facts, you need to recall the mean and variance of the indicator variable that we are sampling from: $\mathbb{E}[\mathbf{1}(T \le t)] = F(t)$, $\operatorname{var}[\mathbf{1}(T \le t)] = F(t)[1 - F(t)]$

For fixed $t \in \mathbb{R}$,

1. $\mathbb{E}[F_n(t)] = F(t)$ so $F_n(t)$ is an unbiased estimator of $F(t)$
2. $\operatorname{var}[F_n(t)] = F(t)[1 - F(t)]/n$ so $F_n(t)$ is a consistent estimator of $F(t)$

3. $F_n(t) \to F(t)$ as $n \to \infty$ with probability 1, the strong law of large numbers.
4. $\sqrt{n}\,[F_n(t) - F(t)]$ has mean 0 and variance $F(t)\,[1 - F(t)]$ and converges to the normal distribution with these parameters as $n \to \infty$, the central limit theorem.

The theorem above gives us a great deal of information about $F_n(t)$ for *fixed* t, but now we want to let t vary and consider the expression in (d), namely $t \mapsto \sqrt{n}\,[F_n(t) - F(t)]$, as a random process for each $n \in \mathbb{N}_+$. The key is to consider a very special distribution first.

Suppose that T has the standard uniform distribution, that is, the continuous uniform distribution on the interval $[0, 1]$. In this case the distribution function is simply $F(t) = t$ for $t \in [0, 1]$, so we have the sequence of stochastic processes $\boldsymbol{X}_n = \{X_n(t) : t \in [0, 1]\}$ for $n \in \mathbb{N}_+$, where

$$X_n(t) = \sqrt{n}\,[F_n(t) - t] \tag{18.3.13}$$

Of course, the previous results apply, so the process $\boldsymbol{X}_n$ has mean function 0, variance function $t \mapsto t(1-t)$, and for fixed $t \in [0, 1]$, the distribution $X_n(t)$ converges to the corresponding normal distribution as $n \to \infty$. Here is the new bit of information, the covariance function of $\boldsymbol{X}_n$ is the same as that of the Brownian bridge!

$\operatorname{cov}[X_n(s), X_n(t)] = \min\{s, t\} - st$ for $s,\, t \in [0, 1]$.

Proof

Suppose that $s \le t$. From basic properties of covariance,

$$\operatorname{cov}[X_n(s), X_n(t)] = n \operatorname{cov}[F_n(s), F_n(t)] = \frac{1}{n}\operatorname{cov}\left(\sum_{i=1}^n \mathbf{1}(T_i \le s), \sum_{j=1}^n \mathbf{1}(T_j \le t)\right) = \frac{1}{n}\sum_{i=1}^n \sum_{j=1}^n \operatorname{cov}[\mathbf{1}(T_i \le s)\mathbf{1}(T_j \le t)] \tag{18.3.14}$$

But if $i \ne j$, the variables $\mathbf{1}(T_i \le s)$ and $\mathbf{1}(T_j \le t)$ are independent, and hence have covariance 0. On the other hand,

$$\operatorname{cov}[\mathbf{1}(T_i \le s), \mathbf{1}(T_i \le t)] = \mathbb{P}(T_i \le s, T_i \le t) - \mathbb{P}(T_i \le s)\mathbb{P}(T_i \le t) = \mathbb{P}(T_i \le s) - \mathbb{P}(T_i \le s)\mathbb{P}(T_i \le t) = s - st \tag{18.3.15}$$

hence

$$\operatorname{cov}[X_n(s), X_n(t)] = \frac{1}{n}\sum_{i=1}^n \operatorname{cov}[\mathbf{1}(T_i \le s), \mathbf{1}(T_i \le t)] = s - st \tag{18.3.16}$$