1

Basic notions

In a first course on probability one typically works with a sequence of random variables $X_1, X_2, \ldots$ For stochastic processes, instead of indexing the random variables by the positive integers, we index them by $t \in [0, \infty)$ and we think of X_t as being the value at time t. The random variable could be the location of a particle on the real line, the strength of a signal, the price of a stock, and many other possibilities as well.

We will also work with increasing families of σ-fields $\{\mathcal{F}_t\}$, known as filtrations. The σ-field $\mathcal{F}_t$ is supposed to represent what we know up to time t.

1.1 Processes and σ-fields

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A real-valued *stochastic process* (or simply a process) is a map X from $[0, \infty) \times \Omega$ to the reals. We write $X_t = X_t(\omega) = X(t, \omega)$. We will impose stronger measurability conditions shortly, but for now we require that the random variables X_t be measurable with respect to $\mathcal{F}$ for each $t \geq 0$.

A collection of σ-fields $\mathcal{F}_t$ such that $\mathcal{F}_t \subset \mathcal{F}$ for each t and $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$ is called a *filtration*. Define $\mathcal{F}_{t+} = \cap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$. A filtration is *right continuous* if $\mathcal{F}_{t+} = \mathcal{F}_t$ for all $t \geq 0$. The σ-field $\mathcal{F}_{t+}$ is supposed to represent what one knows if one looks ahead an infinitesimal amount. Most of the filtrations we will come across will be right continuous, but see Exercise 1.1.

A *null set* N is one that has outer probability 0. This means that

$$\inf\{\mathbb{P}(A) : N \subset A, A \in \mathcal{F}\} = 0.$$

A filtration is *complete* if each $\mathcal{F}_t$ contains every null set. A filtration that is right continuous and complete is said to satisfy the *usual conditions*.

Given a filtration $\{\mathcal{F}_t\}$, whether or not it satisfies the usual conditions, we define $\mathcal{F}_\infty$ to be the σ-field generated by $\cup_{t\geq 0} \mathcal{F}_t$, that is, the smallest σ-field containing $\cup_{t\geq 0} \mathcal{F}_t$, and we write

$$\mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t.$$

Recall that the arbitrary intersection of σ-fields is a σ-field, but the union of even two σ-fields need not be a σ-field.

We say that a stochastic process X is *adapted* to a filtration $\{\mathcal{F}_t\}$ if X_t is $\mathcal{F}_t$ measurable for each t. Often one starts with a stochastic process X and wants to define a filtration with respect to which X is adapted.

The simplest way to do this is to let $\mathcal{F}_t$ be the σ-field generated by the random variables $\{X_s, s \leq t\}$. More often one wants to have a slightly larger filtration than the one generated by X.

We define the *minimal augmented filtration* generated by X to be the smallest filtration that is right continuous and complete and with respect to which the process X is adapted. For each t, $\mathcal{F}_t$ is in general strictly larger than the smallest σ-field with respect to which $\{X_s : s \leq t\}$ is measurable because of the inclusion of the null sets. It is important to include the null sets; see Exercise 1.5. There is no widely accepted name for what we call the minimal augmented filtration; I like this nomenclature because it is descriptive and sufficiently different from "filtration generated by X" to avoid confusion.

The minimal augmented filtration generated by the process X_t can be constructed in three steps. First, let $\{\mathcal{F}_t^{00}\}$ be the smallest filtration with respect to which X is adapted, that is,

$$\mathcal{F}_t^{00} = \sigma(X_s; s \leq t). \tag{1.1}$$

Let $\mathbb{P}^*$ be the outer probability corresponding to $\mathbb{P}$: for $A \subset \Omega$,

$$\mathbb{P}^*(A) = \inf\{\mathbb{P}(B) : B \in \mathcal{F}, A \subset B\}.$$

Let $\mathcal{N}$ be the collection of null sets, so that $\mathcal{N} = \{A \subset \Omega : \mathbb{P}^*(A) = 0\}$. The second step is to let $\mathcal{F}_t^0$ be the smallest σ-field containing $\mathcal{F}_t^{00}$ and $\mathcal{N}$, or

$$\mathcal{F}_t^0 = \sigma(\mathcal{F}_t^{00} \cup \mathcal{N}). \tag{1.2}$$

The third step is to let

$$\mathcal{F}_t = \cap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}^0. \tag{1.3}$$

Exercise 1.2 asks you to check that $\{\mathcal{F}_t\}$ is the minimal augmented filtration generated by X. We will refer to $\{\mathcal{F}_t^{00}\}$ as the *filtration generated by* X.

Two stochastic processes X and Y are said to be *indistinguishable* if $\mathbb{P}(X_t \neq Y_t$ for some $t \geq 0) = 0$. X and Y are *versions* of each other if for each $t \geq 0$, we have $\mathbb{P}(X_t \neq Y_t) = 0$. An example of two processes that are versions of each other but are not indistinguishable is to let $\Omega = [0, 1]$, $\mathcal{F}$ the Borel σ-field on $[0, 1]$, $\mathbb{P}$ Lebesgue measure on $[0, 1]$, $X(t, \omega) = 0$ for all t and ω, and $Y(t, \omega)$ equal to 1 if $t = \omega$ and 0 otherwise. Note that the functions $t \to X(t, \omega)$ are continuous for each ω, but the functions $t \to Y(t, \omega)$ are not continuous for any ω.

If X is a stochastic process, the functions $t \to X(t, \omega)$ are called the *paths* or *trajectories* of X. There will be one path for each ω. If the paths of X are continuous functions, except for a set of ω's in a null set, then X is called a *continuous process*, or is said to be continuous. We similarly define right continuous process, left continuous process, etc.

A function $f(t)$ is *right continuous with left limits* if $\lim_{h>0, h\downarrow 0} f(t+h) = f(t)$ for all t and $\lim_{h<0, h\uparrow 0} f(t+h)$ exists for all $t > 0$. Almost all our stochastic processes will have the property that except for a null set of ω's the function $t \to X(t, \omega)$ is right continuous and has left limits. One often sees *cadlag* to refer to paths that are right continuous with left limits; this abbreviates the French "continue à droite, limite à gauche."

1.2 Laws and state spaces

Let $\mathcal{S}$ be a topological space. The Borel σ-field on $\mathcal{S}$ is defined to be the σ-field generated by the open sets of $\mathcal{S}$. A function $f\colon \mathcal{S} \to \mathbb{R}$ is Borel measurable if $f^{-1}(G)$ is in the Borel σ-field of $\mathcal{S}$ whenever G is an open subset of $\mathbb{R}$. A random variable $Y\colon \Omega \to \mathcal{S}$ is measurable with respect to a σ-field $\mathcal{F}$ of subsets of Ω if $\{\omega \in \Omega : Y(\omega) \in A\}$ is in $\mathcal{F}$ whenever A is in the Borel σ-field on $\mathcal{S}$.

A stochastic process taking values in a topological space $\mathcal{S}$ is a map $X\colon [0,\infty) \times \Omega \to \mathcal{S}$, where for each t, the random variable X_t is measurable with respect to $\mathcal{F}$.

Recall that if we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $Y\colon \Omega \to \mathbb{R}$ is a random variable, then the *law* of Y is the probability measure $\mathbb{P}_Y$ on the Borel subsets of $\mathbb{R}$ defined by $\mathbb{P}_Y(A) = \mathbb{P}(Y \in A)$. Similarly, if $Y\colon \Omega \to \mathbb{R}^d$ is a d-dimensional random vector, then the law of Y is the probability measure $\mathbb{P}_Y$ on the Borel subsets of $\mathbb{R}^d$ defined by $\mathbb{P}_Y(A) = \mathbb{P}(Y \in A)$. We extend this definition to random variables Y taking values in a topological space $\mathcal{S}$. In this case $\mathbb{P}_Y$ is a probability measure on the Borel subsets of $\mathcal{S}$ with the same definition: $\mathbb{P}_Y(A) = \mathbb{P}(Y \in A)$. In particular, if Y and Z are two random variables with the same state space $\mathcal{S}$, then Y and Z will have the same law if $\mathbb{P}(Y \in A) = \mathbb{P}(Z \in A)$ for all Borel subsets A of $\mathcal{S}$.

The relevance of the preceding paragraph to stochastic processes is this. Suppose X and Y are stochastic processes with continuous paths. Let $\mathcal{S} = C[0,\infty)$ be the collection of real-valued continuous functions on $[0,\infty)$ together with the usual metric defined in terms of the supremum norm:

$$d(f,g) = \sup_{0 \le t} |f(t) - g(t)|.$$

(Strictly speaking, we should write $C([0,\infty))$, but we follow the usual convention and drop the outside parentheses.) Let the random variable $\overline{X}$ taking values in $\mathcal{S}$ be defined by setting $\overline{X}(\omega)$ to be the continuous function $t \to X(t,\omega)$, and define $\overline{Y}$ similarly. More precisely, $\overline{X}\colon \Omega \to \mathcal{S}$ with

$$\overline{X}(\omega)(t) = X(t,\omega), \qquad t \ge 0.$$

Then $\overline{X}$ and $\overline{Y}$ are random variables taking values in the metric space $\mathcal{S}$, and saying that $\overline{X}$ and $\overline{Y}$ have the same law means that $\mathbb{P}(\overline{X} \in A) = \mathbb{P}(\overline{Y} \in A)$ for all Borel subsets A of $\mathcal{S}$. When this happens, we also say that the stochastic processes X and Y have the same law.

Two stochastic processes X and Y have the same *finite-dimensional distributions* if for every $n \ge 1$ and every $t_1 < \cdots < t_n$, the laws of $(X_{t_1}, \ldots, X_{t_n})$ and $(Y_{t_1}, \ldots, Y_{t_n})$ are equal.

Most often the topological spaces we will consider will also be metric spaces, but there will be a few occasions when we want to consider topological spaces that are not metric spaces. Suppose $\mathcal{S} = \mathbb{R}^{[0,\infty)}$. We furnish $\mathcal{S}$ with the product topology. $\mathcal{S}$ can be identified with the collection of real-valued functions on $[0,\infty)$, but the topology is not given by the supremum norm nor by any other metric. We use f for elements of $\mathcal{S}$, where $f(t)$ is the tth coordinate of f. We call a subset A of $\mathcal{S}$ a *cylindrical set* if there exist $n \ge 1$, non-negative reals $t_1, t_2, \ldots, t_n$, and a Borel subset B of $\mathbb{R}^n$ such that

$$A = \{f \in \mathcal{S} : (f(t_1), \ldots, f(t_n)) \in B\}.$$

The appropriate σ-field to use on $\mathcal{S}$ is the one generated by the collection of cylindrical sets.

We want to generalize this notion slightly by allowing more general index sets and by allowing for the possibility of considering only a subset of the product space.

Definition 1.1 Let $\mathcal{U}$ be a topological space, T an arbitrary index set, and B a subset of $\mathcal{U}^T$, the collection of functions from T into $\mathcal{U}$. We say a set C is a *cylindrical subset* of B if there exist $n \geq 1, t_1, \ldots, t_n \in T$, and a Borel subset A of $\mathbb{R}^n$ such that

$$C = \{f \in B : (f(t_1), \ldots, f(t_n)) \in A\}.$$

Exercises

1.1 This exercise gives an example where $\{\mathcal{F}_t^{00}\}$ defined by (1.1) is not right continuous. Let $\Omega = \{a, b\}$, let $\mathcal{F}$ be the collection of all subsets of Ω, and let $\mathbb{P}(\{a\}) = \mathbb{P}(\{b\}) = \frac{1}{2}$. Define

$$X_t(\omega) = \begin{cases} 0, & t \leq 1; \\ 0, & t > 1 \text{ and } \omega = a; \\ t-1, & t > 1 \text{ and } \omega = b. \end{cases}$$

Calculate $\mathcal{F}_t^{00} = \sigma(X_s; s \leq t)$ and show $\{\mathcal{F}_t^{00}\}$ is not right continuous.

1.2 If X is a stochastic process, let $\mathcal{F}_t^{00}$, $\mathcal{F}_t^0$, and $\mathcal{F}_t$ be defined by (1.1), (1.2), and (1.3), respectively. Show that $\{\mathcal{F}_t\}$ is the minimal augmented filtration generated by X.

1.3 Let $\{\mathcal{F}_t\}$ be a filtration satisfying the usual conditions and let $\mathcal{B}[0, t]$ be the Borel σ-field on $[0, t]$. A real-valued stochastic process X is *progressively measurable* if for each $t \geq 0$, the map $(s, \omega) \to X(s, \omega)$ from $[0, t] \times \Omega$ to $\mathbb{R}$ is measurable with respect to the product σ-field $\mathcal{B}[0, t] \times \mathcal{F}_t$.

(1) If X is adapted to $\{\mathcal{F}_t\}$ and we define

$$X_t^{(n)}(\omega) = \sum_{k=0}^{\infty} X_{k/2^n}(\omega) 1_{[k/2^n, (k+1)/2^n)}(t),$$

show that $X^{(n)}$ is progressively measurable for each $n \geq 1$.

(2) Use (1) to show that if X is adapted to $\{\mathcal{F}_t\}$ and has left continuous paths, then X is progressively measurable.

(3) If X is adapted to $\{\mathcal{F}_t\}$ and we define

$$Y_t^{(n)}(\omega) = \sum_{k=0}^{\infty} X_{(k+1)/2^n}(\omega) 1_{[k/2^n, (k+1)/2^n)}(t),$$

show that for each $t \geq 0$, the map $(s, \omega) \to Y^{(n)}(s, \omega)$ from $[0, t] \times \Omega$ to $\mathbb{R}$ is measurable with respect to $\mathcal{B}[0, t] \times \mathcal{F}_{t+2^{-n}}$.

(4) Show that if X is adapted to $\{\mathcal{F}_t\}$ and has right continuous paths, then X is progressively measurable.

1.4 Let $\mathcal{S} = \mathbb{R}^{[0,1]}$, the set of functions from $[0, 1]$ to $\mathbb{R}$, and let $\mathcal{F}$ be the σ-field generated by the cylindrical sets. The purpose of this exercise is to show that the elements of $\mathcal{F}$ depend on only countably many coordinates.

Let $\mathcal{S}_0 = \{(x_1, x_2, \dots)\}$, the set of sequences taking values in $\mathbb{R}$. Let $\mathcal{F}_0$ be the σ-field generated by the cylindrical subsets of $\mathbb{R}^{\mathbb{N}}$, where $\mathbb{N} = \{1, 2, \dots\}$.

Show that $B \in \mathcal{F}$ if and only if there exist $t_1, t_2, \dots$ in $[0, 1]$ and a set $C \in \mathcal{F}_0$ such that

$$B = \{f \in \mathcal{S} : (f(t_1), f(t_2), \dots) \in C\}.$$

1.5 Null sets are sometimes important! Let $\mathcal{S}$ and $\mathcal{F}$ be as in Exercise 1.4. Show that $D \notin \mathcal{F}$, where

$$D = \{f \in \mathcal{S} : f \text{ is a continuous function on } [0, 1]\}.$$

1.6 Suppose X is a stochastic process, $\{\mathcal{F}_t\}$ its minimal augmented filtration, and $\mathcal{F}_\infty = \vee_{t \ge 0} \mathcal{F}_t$. Suppose with probability one, the paths of X are right continuous with left limits. Let $X_{t-} = \lim_{s<t, s\to t} X_s$, the left-hand limit at time t, and $\Delta X_t = X_t - X_{t-}$, the size of the jump at time t. If

$$A = \{\exists t \ge 0 : \Delta X_t > 1\},$$

prove $A \in \mathcal{F}_\infty$.

1.7 Suppose X is a stochastic process, $\{\mathcal{F}_t\}$ is the minimal augmented filtration for X, and $\mathcal{F}_\infty = \vee_{t \ge 0} \mathcal{F}_t$. If the paths of X are right continuous with left limits with probability one, show that the event

$$A = \{X \text{ has continuous paths}\}$$

is in $\mathcal{F}_\infty$.

Notes

The older literature sometimes uses the notion of a separable stochastic process, but this is rarely seen nowadays. For much more on measurability, see Chapter 16. For the complete story on the foundations of stochastic processes, see Dellacherie and Meyer (1978).

2

Brownian motion

Brownian motion is by far the most important stochastic process. It is the archetype of Gaussian processes, of continuous time martingales, and of Markov processes. It is basic to the study of stochastic differential equations, financial mathematics, and filtering, to name only a few of its applications.

In this chapter we define Brownian motion and consider some of its elementary aspects. Later chapters will take up the construction of Brownian motion and properties of Brownian motion paths.

2.1 Definition and basic properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{\mathcal{F}_t\}$ be a filtration, not necessarily satisfying the usual conditions.

Definition 2.1 $W_t = W_t(\omega)$ is a one-dimensional *Brownian motion* with respect to $\{\mathcal{F}_t\}$ and the probability measure $\mathbb{P}$, started at 0, if

(1) W_t is $\mathcal{F}_t$ measurable for each $t \geq 0$.
(2) $W_0 = 0$, a.s.
(3) $W_t - W_s$ is a normal random variable with mean 0 and variance $t - s$ whenever $s < t$.
(4) $W_t - W_s$ is independent of $\mathcal{F}_s$ whenever $s < t$.
(5) W_t has continuous paths.

If instead of (2) we have $W_0 = x$, we say we have a Brownian motion started at x. Definition 2.1(4) is referred to as the *independent increments* property of Brownian motion. The fact that $W_t - W_s$ has the same law as W_{t-s}, which follows from Definition 2.1(3), is called the *stationary increments* property. When no filtration is specified, we assume the filtration is the filtration generated by W, i.e., $\mathcal{F}_t = \sigma(W_s; s \leq t)$. Sometimes a one-dimensional Brownian motion started at 0 is called a *standard Brownian motion*.

Figure 2.1 is a simulation of a typical Brownian motion path.

We define d-dimensional Brownian motion with respect to a filtration $\{\mathcal{F}_t\}$ and started at $x = (x_1, \ldots, x_d)$ to be $(W_t^{(1)}, \ldots, W_t^{(d)})$, where the $W^{(i)}$ are each one-dimensional Brownian motions with respect to $\{\mathcal{F}_t\}$ started at x_i, respectively, and $W^{(1)}, \ldots, W^{(n)}$ are all independent.

The law of a Brownian motion is called *Wiener measure*. More precisely, given a Brownian motion W, we can view it as a random variable taking values in $C[0, \infty)$, the space of real-valued continuous functions on $[0, \infty)$. The law of W is the measure $\mathbb{P}_W$ on

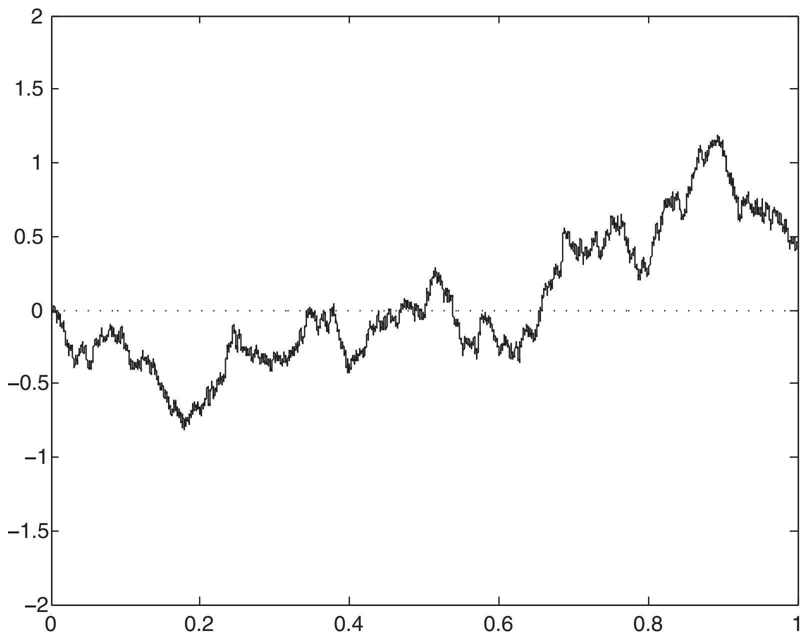

Figure 2.1 Simulation of a typical Brownian motion path.

$C[0, \infty)$ defined by $\mathbb{P}_W(A) = \mathbb{P}(W \in A)$ for all Borel subsets A of $C[0, \infty)$. The measure $\mathbb{P}_W$ is Wiener measure.

There are a number of transformations one can perform on a Brownian motion that yield a new Brownian motion. The first one is called the *scaling property of Brownian motion*, or simply *scaling*.

Proposition 2.2 *If W is a Brownian motion started at 0, $a > 0$, and $Y_t = aW_{t/a^2}$, then Y_t is a Brownian motion started at 0.*

Proof We use $\mathcal{G}_t = \mathcal{F}_{t/a^2}$ for the filtration for Y. Clearly Y_t has continuous paths, $Y_0 = 0$, a.s., and Y_t is $\mathcal{G}_t$ measurable. If $s < t$,

$$Y_t - Y_s = a(W_{t/a^2} - W_{s/a^2})$$

is independent of $\mathcal{F}_{s/a^2}$, hence is independent of $\mathcal{G}_s$. Finally, if $s < t$, and if $s < t$, then $Y_t - Y_s$ will be a normal random variable with mean zero and

$$\text{Var}\,(Y_t - Y_s) = a^2 \text{Var}\,(W_{t/a^2} - W_{s/a^2}) = a^2\left(\frac{t}{a^2} - \frac{s}{a^2}\right) = t - s.$$

This suffices to give our result. □

For some other transformations, see Exercises 2.3 and 2.5.

Recall what it means for a finite collection of random variables to be jointly normal; see (A.29). A stochastic process X is *Gaussian* or *jointly normal* if all its finite-dimensional distributions are jointly normal, that is, if for each $n \geq 1$ and $t_1 < \cdots < t_n$, the collection of random variables $X_{t_1}, \ldots, X_{t_n}$ is a jointly normal collection.

Proposition 2.3 *If W is a Brownian motion, then W is a Gaussian process.*

Proof Suppose W is a Brownian motion and let $0 = t_0 < t_1 < \cdots < t_n$. Define

$$Z_i = \frac{W_{t_i} - W_{t_{i-1}}}{\sqrt{t_i - t_{i-1}}}, \qquad i = 1, 2, \ldots, n.$$

By Definition 2.1(4), Z_i is independent of $\mathcal{F}_{t_{i-1}}$, and hence independent of $Z_1, \ldots, Z_{j-1}$. By Definition 2.1(3), Z_i is a mean-zero random variable with variance one. We can write

$$W_{t_j} = \sum_{i=1}^{j} (t_i - t_{i-1})^{1/2} Z_i, \qquad j = 1, \ldots, n,$$

and so $(W_{t_1}, \ldots, W_{t_n})$ is jointly normal. It follows that Brownian motion is a Gaussian process. □

Since the law of a finite collection of jointly normal random variables is determined by their means and covariances, let's calculate the covariance of W_s and W_t when W is a Brownian motion. If $s \le t$, then

$$\begin{aligned} t - s = \mathrm{Var}\,(W_t - W_s) &= \mathrm{Var}\, W_t + \mathrm{Var}\, W_s - 2\,\mathrm{Cov}\,(W_s, W_t) \\ &= t + s - 2\,\mathrm{Cov}\,(W_s, W_t) \end{aligned}$$

from Definition 2.1(2) and (3). Hence $\mathrm{Cov}\,(W_s, W_t) = s$ if $s \le t$. This is frequently written as

$$\mathrm{Cov}\,(W_s, W_t) = s \wedge t. \tag{2.1}$$

We have the following converse.

Theorem 2.4 *If W is a process such that all the finite-dimensional distributions are jointly normal, $\mathbb{E}\, W_s = 0$ for all s, $\mathrm{Cov}\,(W_s, W_t) = s$ when $s \le t$, and the paths of W_t are continuous, then W is a Brownian motion.*

Proof For $\mathcal{F}_t$ we take the filtration generated by W. If we take $s = t$, then $\mathrm{Var}\, W_t = \mathrm{Cov}\,(W_t, W_t) = t$. In particular, $\mathrm{Var}\, W_0 = 0$, and since $\mathbb{E}\, W_0 = 0$, then $W_0 = 0$, a.s. We have

$$\begin{aligned} \mathrm{Var}\,(W_t - W_s) &= \mathrm{Var}\, W_t - 2\,\mathrm{Cov}\,(W_s, W_t) + \mathrm{Var}\, W_t \\ &= t - 2s + s = t - s. \end{aligned}$$

We have thus established all the parts of Definition 2.1 except for the independence of $W_t - W_s$ from $\mathcal{F}_s$.

If $r \le s < t$, then

$$\mathrm{Cov}\,(W_t - W_s, W_r) = \mathrm{Cov}\,(W_t, W_r) - \mathrm{Cov}\,(W_s, W_r) = r - r = 0,$$

and so $W_t - W_s$ is independent of W_r by Proposition A.55. This shows that $W_t - W_s$ is independent of $\mathcal{F}_s$. □

We now look at two results that are more technical. These should only be skimmed on the first reading of the book: read the statements, but not the proofs. The first result says that if W is a Brownian motion with respect to the filtration generated by W, then it is also a Brownian motion with respect to the minimal augmented filtration.

Proposition 2.5 *Let W_t be a Brownian motion with respect to $\{\mathcal{F}_t^{00}\}$, where $\mathcal{F}_t^{00} = \sigma(W_s; s \le t)$. Let $\mathcal{N}$ be the collection of null sets, $\mathcal{F}_t^0 = \sigma(\mathcal{F}_t^{00} \cup \mathcal{N})$, and $\mathcal{F}_t = \cap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}^0$.*
(1) W is a Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$.
(2) $\mathcal{F}_t = \mathcal{F}_t^0$ for each t.

Proof (1) The only property we need to check is Definition 2.1(4). If f is a continuous bounded function on $\mathbb{R}$, $A \in \mathcal{F}_s^{00}$, and $s < t$, then because W is a Brownian motion with respect to $\{\mathcal{F}_t^{00}\}$, the independent increments property shows that

$$\mathbb{E}[f(W_t - W_s); A] = \mathbb{E}[f(W_t - W_s)]\,\mathbb{P}(A). \tag{2.2}$$

If A is such that $A \backslash B$ and $B \backslash A$ are null sets for some $B \in \mathcal{F}_s^{00}$, it is easy to see that (2.2) continues to hold. By linearity, it also holds if A is a finite disjoint union of such sets. If $\mathcal{C}_1$ is the collection of subsets of $\mathcal{F}_s^0$ that are finite disjoint unions of such sets, then $\mathcal{C}_1$ is an algebra of subsets of $\mathcal{F}_s^0$. Let $\mathcal{M}_1$ be the collection of subsets of $\mathcal{F}_s^0$ for which (2.2) holds. It is readily checked that $\mathcal{M}_1$ is a monotone class. By the monotone class theorem (Theorem B.2), $\mathcal{M}_1$ is equal to the smallest σ-field containing $\mathcal{C}_1$, which is $\mathcal{F}_s^0$. Therefore (2.2) holds for all $A \in \mathcal{F}_s^0$.

Now suppose $A \in \mathcal{F}_s = \mathcal{F}_{s+}^0$. Then for each $\varepsilon > 0$, $A \in \mathcal{F}_{s+\varepsilon}^0$, and so using (2.2) with s replaced by $s + \varepsilon$ and t replaced by $t + \varepsilon$, we have

$$\mathbb{E}[f(W_{t+\varepsilon} - W_{s+\varepsilon}); A] = \mathbb{E}[f(W_{t+\varepsilon} - W_{s+\varepsilon})]\,\mathbb{P}(A). \tag{2.3}$$

Letting $\varepsilon \to 0$ and using the facts that f is bounded and continuous and W has continuous paths, the dominated convergence theorem implies that

$$\mathbb{E}[f(W_t - W_s); A] = \mathbb{E}[f(W_t - W_s)]\,\mathbb{P}(A). \tag{2.4}$$

This equation holds whenever f is continuous and $A \in \mathcal{F}_s$. By a limit argument, (2.4) holds whenever f is the indicator of a Borel subset of $\mathbb{R}$. That says that $W_t - W_s$ and $\mathcal{F}_s$ are independent.

(2) Fix t and choose $t_0 > t$. Let $\mathcal{M}_2$ be the collection of subsets of $\mathcal{F}_{t_0}^{00}$ whose conditional expectation with respect to $\mathcal{F}_t$ is $\mathcal{F}_t^0$ measurable, that is, $A \in \mathcal{M}_2$ if $A \in \mathcal{F}_{t_0}^{00}$ and $\mathbb{E}[1_A \mid \mathcal{F}_t]$ is $\mathcal{F}_t^0$ measurable. Let $\mathcal{C}_2$ be the collection of events A for which there exist $n \ge 1$, $0 \le s_0 < s_1 < \cdots < s_n \le t_0$ with t equal to one of the s_i, and Borel subsets $B_1, \ldots, B_n$ of $\mathbb{R}$ such that

$$A = (W_{s_1} - W_{s_0} \in B_1, \ldots, W_{s_n} - W_{s_{n-1}} \in B_n).$$

Suppose A is of this form, and suppose $t = s_i$. Then by the independence result that we proved in (1),

$$\begin{aligned}\mathbb{E}[1_A \mid \mathcal{F}_t] = 1_{(W_{s_1} - W_{s_0} \in B_1, \ldots, W_{s_i} - W_{s_{i-1}} \in B_i)}\\ \times \mathbb{P}(W_{s_{i+1}} - W_{s_i} \in B_{i+1}, \ldots, W_{s_n} - W_{s_{n-1}} \in B_n),\end{aligned}$$

which is $\mathcal{F}_t^0$ measurable. Thus $\mathcal{C}_2 \subset \mathcal{M}_2$. Finite unions of sets in $\mathcal{C}_2$ form an algebra of subsets of $\mathcal{F}_{t_0}^{00}$ that generate $\mathcal{F}_t^{00}$. It is easy to check that $\mathcal{M}_2$ is a monotone class, so by the monotone class theorem, $\mathcal{M}_2$ equals $\mathcal{F}_t^{00}$. By linearity and taking monotone limits, if Y is non-negative and $\mathcal{F}_t^{00}$ measurable, then $\mathbb{E}[Y \mid \mathcal{F}_t]$ is $\mathcal{F}_t^0$ measurable.

To finish, suppose $A \in \mathcal{F}_t$. Then since $t < t_0$, we see that $A \in \mathcal{F}^0_{t_0}$. By Exercise 2.7, there exists $Y \in \mathcal{F}^{00}_{t_0}$ such that $1_A = Y$, a.s. Then $\mathbb{E}[Y \mid \mathcal{F}_t]$ is $\mathcal{F}^0_t$ measurable. Since $\mathcal{F}^0_t$ contains all the null sets, $1_A = \mathbb{E}[1_A \mid \mathcal{F}_t]$ is also $\mathcal{F}^0_t$ measurable, or $A \in \mathcal{F}^0_t$. This proves (2). □

The final item we consider in this chapter is a subtle one. The question is this: if W and W' are both Brownian motions, do they have all the same properties? To illustrate this issue, let's revisit the example of Chapter 1 where $\Omega = [0, 1]$, $\mathcal{F}$ is the Borel σ-field on $[0, 1]$, $\mathbb{P}$ is Lebesgue measure on $[0, 1]$, $X(t, \omega) = 0$ for all t and ω, and $Y(t, \omega)$ is 1 if $t = \omega$ and 0 otherwise. For each t, $\mathbb{P}(X_t = Y_t) = 1$, so X and Y have the same finite-dimensional distributions. However, if

$$A = \{f : f \text{ is not a continuous function on } [0, 1]\},$$

then $(X \in A)$ is a null set but $(Y \in A)$ is not. Even though X and Y have the same finite-dimensional distributions, X has continuous paths but Y does not.

To rephrase our question, is it true that $\mathbb{P}(W \in A) = \mathbb{P}(W' \in A)$ for every Borel subset A of $C[0, \infty)$? We know W and W' have the same finite-dimensional distributions because each is jointly normal with zero means and $\mathrm{Cov}\,(W_s, W_t) = s \wedge t = \mathrm{Cov}\,(W'_s, W'_t)$. The fact that the answer to our question is yes then comes from the following theorem. We look at $C[0, t_0]$ instead of $C[0, \infty)$ for the sake of simplicity.

Theorem 2.6 *Let $t_0 > 0$ and let X, Y be random variables taking values in $C[0, t_0]$ which have the same finite-dimensional distributions. Then the laws of X and Y are equal.*

Proof Let $\mathcal{M}$ be the collection of Borel subsets A of $C[0, t_0]$ for which $\mathbb{P}(X \in A)$ equals $\mathbb{P}(Y \in A)$. We will show that $\mathcal{M}$ is a monotone class and then use the monotone class theorem to show that $\mathcal{M}$ is equal to the Borel σ-field on $C[0, t_0]$.

First, let $\mathcal{C}$ be the collection of all cylindrical subsets of $C[0, t_0]$ (defined by Definition 1.1). Since the finite-dimensional distributions of X and Y are equal, then $\mathcal{M}$ contains $\mathcal{C}$. It is easy to check that $\mathcal{C}$ is an algebra of subsets of $C[0, t_0]$. If $A_1 \supset A_2 \supset \cdots$ are elements of $\mathcal{M}$, then

$$\mathbb{P}(X \in \cap_n A_n) = \lim_n \mathbb{P}(X \in A_n) = \lim_n \mathbb{P}(Y \in A_n) = \mathbb{P}(Y \in \cap_n A_n)$$

since $\mathbb{P}$ is a finite measure. Therefore $\cap_n A_n \in \mathcal{M}$. A very similar argument shows that if $A_1 \subset A_2 \subset \cdots$ are elements of $\mathcal{M}$, then $\cup_n A_n \in \mathcal{M}$. Therefore $\mathcal{M}$ is a monotone class. By the monotone class theorem, $\mathcal{M}$ contains the smallest σ-field containing $\mathcal{C}$. We will show that $\mathcal{M}$ contains all the open sets; then $\mathcal{M}$ will contain the smallest σ-field containing the open sets, and we will be done.

Since $C[0, t_0]$ is separable, every open set is the countable union of open balls. Because $\mathcal{M}$ is a σ-field, it suffices to show that $\mathcal{M}$ contains the open balls in $C[0, t_0]$, that is, all sets of the form

$$B(f_0, r) = \{f \in C[0, t_0] : \sup_{0 \le t \le t_0} |f(t) - f_0(t)| < r\}$$

where $r > 0$ and $f_0 \in C[0, t_0]$. For each m and n,

$$\{f \in C[0, t_0] : \sup_{0 \le k \le 2^n t_0} |f(k/2^n) - f_0(k/2^n)| \le r - (1/m)\}$$