2

Brownian Motion

2.1 Definition of Brownian Motion

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A stochastic process is a measurable function $X(t, \omega)$ defined on the product space $[0, \infty) \times \Omega$. In particular,

(a) for each t, $X(t, \cdot)$ is a random variable,
(b) for each ω, $X(\cdot, \omega)$ is a measurable function (called a *sample path*).

For convenience, the random variable $X(t, \cdot)$ will be written as $X(t)$ or X_t. Thus a stochastic process $X(t, \omega)$ can also be expressed as $X(t)(\omega)$ or simply as $X(t)$ or X_t.

Definition 2.1.1. *A stochastic process $B(t, \omega)$ is called a* Brownian motion *if it satisfies the following conditions:*

(1) $P\{\omega\,;\, B(0, \omega) = 0\} = 1$.

(2) *For any $0 \le s < t$, the random variable $B(t) - B(s)$ is normally distributed with mean 0 and variance $t - s$, i.e., for any $a < b$,*

$$P\big\{a \le B(t) - B(s) \le b\big\} = \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b e^{-x^2/2(t-s)}\, dx.$$

(3) *$B(t, \omega)$ has independent increments, i.e., for any $0 \le t_1 < t_2 < \cdots < t_n$, the random variables*

$$B(t_1),\ B(t_2) - B(t_1),\ \ldots\ , B(t_n) - B(t_{n-1}),$$

are independent.

(4) *Almost all sample paths of $B(t, \omega)$ are continuous functions, i.e.,*

$$P\big\{\omega\,;\, B(\cdot, \omega) \text{ is continuous}\big\} = 1.$$

In Remark 1.2.3 we mentioned that the limit

$$B(t) = \lim_{\delta \to 0} Y_{\delta, \sqrt{\delta}}(t)$$

is a Brownian motion. However, this fact comes only as a consequence of an intuitive observation. In the next chapter we will give several constructions of Brownian motion. But before these constructions we shall give some simple properties of a Brownian motion and define the Wiener integral.

A Brownian motion is sometimes defined as a stochastic process $B(t, \omega)$ satisfying conditions (1), (2), (3) in Definition 2.1.1. Such a stochastic process always has a continuous realization, i.e., there exists Ω_0 such that $P(\Omega_0) = 1$ and for any $\omega \in \Omega_0$, $B(t, \omega)$ is a continuous function of t. This fact can be easily checked by applying the Kolmogorov continuity theorem in Section 3.3. Thus condition (4) is automatically satisfied.

The Brownian motion $B(t)$ in the above definition starts at 0. Sometimes we will need a Brownian motion starting at x. Such a process is given by $x + B(t)$. If the starting point is not 0, we will explicitly mention the starting point x.

2.2 Simple Properties of Brownian Motion

Let $B(t)$ be a fixed Brownian motion. We give below some simple properties that follow directly from the definition of Brownian motion.

Proposition 2.2.1. *For any $t > 0$, $B(t)$ is normally distributed with mean 0 and variance t. For any $s, t \geq 0$, we have $E[B(s)B(t)] = \min\{s, t\}$.*

Remark 2.2.2. Regarding Definition 2.1.1, it can be proved that condition (2) and $E[B(s)B(t)] = \min\{s, t\}$ imply condition (3).

Proof. By condition (1), we have $B(t) = B(t) - B(0)$ and so the first assertion follows from condition (2). To show that $EB(s)B(t) = \min\{s, t\}$ we may assume that $s < t$. Then by conditions (2) and (3),

$$E\big[B(s)B(t)\big] = E\big[B(s)\big(B(t) - B(s)\big) + B(s)^2\big] = 0 + s = s,$$

which is equal to $\min\{s, t\}$. □

Proposition 2.2.3. (Translation invariance) *For fixed $t_0 \geq 0$, the stochastic process $\widetilde{B}(t) = B(t + t_0) - B(t_0)$ is also a Brownian motion.*

Proof. The stochastic process $\widetilde{B}(t)$ obviously satisfies conditions (1) and (4) of a Brownian motion. For any $s < t$,

$$\widetilde{B}(t) - \widetilde{B}(s) = B(t + t_0) - B(s + t_0). \tag{2.2.1}$$

By condition (2) of $B(t)$, we see that $\widetilde{B}(t)-\widetilde{B}(s)$ is normally distributed with mean 0 and variance $(t+t_0)-(s+t_0)=t-s$. Thus $\widetilde{B}(t)$ satisfies condition (2). To check condition (3) for $\widetilde{B}(t)$, we may assume that $t_0>0$. Then for any $0\le t_1<t_2<\cdots<t_n$, we have $0<t_0\le t_1+t_0<\cdots<t_n+t_0$. Hence by condition (3) of $B(t)$, $B(t_k+t_0)-B(t_{k-1}+t_0)$, $k=1,2,\ldots,n$, are independent random variables. Thus by Equation (2.2.1), the random variables $\widetilde{B}(t_k)-\widetilde{B}(t_{k-1})$, $k=1,2,\ldots,n$, are independent and so $\widetilde{B}(t)$ satisfies condition (3) of a Brownian motion. □

The above translation invariance property says that a Brownian motion starts afresh at any moment as a new Brownian motion.

Proposition 2.2.4. (Scaling invariance) *For any real number* $\lambda>0$, *the stochastic process* $\widetilde{B}(t)=B(\lambda t)/\sqrt{\lambda}$ *is also a Brownian motion.*

Proof. Conditions (1), (3), and (4) of a Brownian motion can be readily checked for the stochastic process $\widetilde{B}(t)$. To check condition (2), note that for any $s<t$,

$$\widetilde{B}(t)-\widetilde{B}(s)=\frac{1}{\sqrt{\lambda}}\big(B(\lambda t)-B(\lambda s)\big),$$

which shows that $\widetilde{B}(t)-\widetilde{B}(s)$ is normally distributed with mean 0 and variance $\frac{1}{\lambda}(\lambda t-\lambda s)=t-s$. Hence $\widetilde{B}(t)$ satisfies condition (2). □

It follows from the scaling invariance property that for any $\lambda>0$ and $0\le t_1<t_2<\cdots<t_n$ the random vectors

$$\big(B(\lambda t_1),B(\lambda t_2),\ldots,B(\lambda t_n)\big),\quad \big(\sqrt{\lambda}B(t_1),\sqrt{\lambda}B(t_2),\ldots,\sqrt{\lambda}B(t_n)\big)$$

have the same distribution.

2.3 Wiener Integral

In Section 1.1 we raised the question of defining the integral $\int_a^b f(t)\,dg(t)$. We see from Example 1.1.3 that in general this integral cannot be defined as a Riemann–Stieltjes integral.

Now let us consider the following integral:

$$\int_a^b f(t)\,dB(t,\omega),$$

where f is a deterministic function (i.e., it does not depend on ω) and $B(t,\omega)$ is a Brownian motion. Suppose for each $\omega\in\Omega$ we want to use Equation (1.1.2) to define this integral in the Riemann–Stieltjes sense by

$$(RS)\int_a^b f(t)\,dB(t,\omega)=f(t)B(t,\omega)\Big]_a^b-(RS)\int_a^b B(t,\omega)\,df(t). \qquad (2.3.1)$$

Then the class of functions $f(t)$ for which the integral $(RS)\int_a^b f(t)\,dB(t,\omega)$ is defined for each $\omega \in \Omega$ is rather limited, i.e., $f(t)$ needs to be a continuous function of bounded variation. Hence for a continuous function of unbounded variation such as $f(t) = t\sin\frac{1}{t}$, $0 < t \le 1$, and $f(0) = 0$, we cannot use Equation (2.3.1) to define the integral $\int_0^1 f(t)\,dB(t,\omega)$ for each $\omega \in \Omega$.

We need a different idea in order to define the integral $\int_a^b f(t)\,dB(t,\omega)$ for a wider class of functions $f(t)$. This new integral, called the Wiener integral of f, is defined for all functions $f \in L^2[a,b]$. Here $L^2[a,b]$ denotes the Hilbert space of all real-valued square integrable functions on $[a,b]$. For example, $\int_0^1 t\sin\frac{1}{t}\,dB(t)$ is a Wiener integral.

Now we define the Wiener integral in two steps:

Step 1. Suppose f is a step function given by $f = \sum_{i=1}^n a_i\, 1_{[t_{i-1},t_i)}$, where $t_0 = a$ and $t_n = b$. In this case, define

$$I(f) = \sum_{i=1}^n a_i \big(B(t_i) - B(t_{i-1})\big). \tag{2.3.2}$$

Obviously, $I(af + bg) = aI(f) + bI(g)$ for any $a, b \in \mathbb{R}$ and step functions f and g. Moreover, we have the following lemma.

Lemma 2.3.1. *For a step function f, the random variable $I(f)$ is Gaussian with mean 0 and variance*

$$E\big(I(f)^2\big) = \int_a^b f(t)^2\,dt. \tag{2.3.3}$$

Proof. It is well known that a linear combination of independent Gaussian random variables is also a Gaussian random variable. Hence by conditions (2) and (3) of Brownian motion, the random variable $I(f)$ defined by Equation (2.3.2) is Gaussian with mean 0. To check Equation (2.3.3), note that

$$E\big(I(f)^2\big) = E\sum_{i,j=1}^n a_i\, a_j \big(B(t_i) - B(t_{i-1})\big)\big(B(t_j) - B(t_{j-1})\big).$$

By conditions (2) and (3) of Brownian motion,

$$E\big(B(t_i) - B(t_{i-1})\big)^2 = t_i - t_{i-1},$$

and for $i \ne j$,

$$E\big(B(t_i) - B(t_{i-1})\big)\big(B(t_j) - B(t_{j-1})\big) = 0.$$

Therefore,

$$E\big(I(f)^2\big) = \sum_{i=1}^n a_i^2\,(t_i - t_{i-1}) = \int_a^b f(t)^2\,dt. \qquad \square$$

Step 2. We will use $L^2(\Omega)$ to denote the Hilbert space of square integrable real-valued random variables on Ω with inner product $\langle X, Y\rangle = E(XY)$. Let $f \in L^2[a,b]$. Choose a sequence $\{f_n\}_{n=1}^{\infty}$ of step functions such that $f_n \to f$ in $L^2[a,b]$. By Lemma 2.3.1 the sequence $\{I(f_n)\}_{n=1}^{\infty}$ is Cauchy in $L^2(\Omega)$. Hence it converges in $L^2(\Omega)$. Define

$$I(f) = \lim_{n\to\infty} I(f_n), \quad \text{in } L^2(\Omega). \tag{2.3.4}$$

Question 2.3.2. Is $I(f)$ well-defined?

In order for $I(f)$ to be well-defined, we need to show that the limit in Equation (2.3.4) is independent of the choice of the sequence $\{f_n\}$. Suppose $\{g_m\}$ is another such sequence, i.e., the g_m's are step functions and $g_m \to f$ in $L^2[a,b]$. Then by the linearity of the mapping I and Equation (2.3.3),

$$E\big(|I(f_n) - I(g_m)|^2\big) = E\big(|I(f_n - g_m)|^2\big) = \int_a^b \big(f_n(t) - g_m(t)\big)^2 \, dt.$$

Write $f_n(t) - g_m(t) = \big[f_n(t) - f(t)\big] - \big[g_m(t) - f(t)\big]$ and then use the inequality $(x-y)^2 \le 2(x^2 + y^2)$ to get

$$\begin{aligned} \int_a^b \big(f_n(t) - g_m(t)\big)^2 \, dt &\le 2 \int_a^b \Big(\big[f_n(t) - f(t)\big]^2 + \big[g_m(t) - f(t)\big]^2\Big) \, dt \\ &\to 0, \qquad \text{as } n, m \to \infty. \end{aligned}$$

It follows that $\lim_{n\to\infty} I(f_n) = \lim_{m\to\infty} I(g_m)$ in $L^2(\Omega)$. This shows that $I(f)$ is well-defined.

Definition 2.3.3. *Let $f \in L^2[a,b]$. The limit $I(f)$ defined in Equation* (2.3.4) *is called the* Wiener integral *of f.*

The Wiener integral $I(f)$ of f will be denoted by

$$I(f)(\omega) = \Big(\int_a^b f(t)\, dB(t)\Big)(\omega), \quad \omega \in \Omega, \ \text{almost surely}.$$

For simplicity, it will be denoted by $\int_a^b f(t)\, dB(t)$ or $\int_a^b f(t)\, dB(t,\omega)$. Note that the mapping I is linear on $L^2[a,b]$.

Theorem 2.3.4. *For each $f \in L^2[a,b]$, the Wiener integral $\int_a^b f(t)\, dB(t)$ is a Gaussian random variable with mean* 0 *and variance $\|f\|^2 = \int_a^b f(t)^2\, dt$.*

Proof. By Lemma 2.3.1, the assertion is true when f is a step function. For a general $f \in L^2[a,b]$, the assertion follows from the following well-known fact: If X_n is Gaussian with mean μ_n and variance σ_n^2 and X_n converges to X in $L^2(\Omega)$, then X is Gaussian with mean $\mu = \lim_{n\to\infty} \mu_n$ and variance $\sigma^2 = \lim_{n\to\infty} \sigma_n^2$. □

Thus the Wiener integral $I : L^2[a,b] \to L^2(\Omega)$ is an isometry. In fact, it preserves the inner product, as shown by the next corollary.

Corollary 2.3.5. *If $f, g \in L^2[a,b]$, then*

$$E\big(I(f)\, I(g)\big) = \int_a^b f(t)g(t)\, dt. \tag{2.3.5}$$

In particular, if f and g are orthogonal, then the Gaussian random variables $I(f)$ and $I(g)$ are independent.

Proof. By the linearity of I and Theorem 2.3.4 we have

$$\begin{aligned} E\big[(I(f)+I(g))^2\big] &= E\big[(I(f+g))^2\big] \\ &= \int_a^b \big(f(t)+g(t)\big)^2\, dt \\ &= \int_a^b f(t)^2\, dt + 2\int_a^b f(t)g(t)\, dt + \int_a^b g(t)^2\, dt. \end{aligned} \tag{2.3.6}$$

On the other hand, we can also use Theorem 2.3.4 to obtain

$$\begin{aligned} &E\big[(I(f)+I(g))^2\big] \\ &\quad = E\big[I(f)^2 + 2I(f)I(g) + I(g)^2\big] \\ &\quad = \int_a^b f(t)^2\, dt + 2E\big[I(f)I(g)\big] + \int_a^b g(t)^2\, dt. \end{aligned} \tag{2.3.7}$$

Obviously, Equation (2.3.5) follows from Equations (2.3.6) and (2.3.7). □

Example 2.3.6. The Wiener integral $\int_0^1 s\, dB(s)$ is a Gaussian random variable with mean 0 and variance $\int_0^1 s^2\, ds = \frac{1}{3}$.

Theorem 2.3.7. *Let f be a continuous function of bounded variation. Then for almost all $\omega \in \Omega$,*

$$\left(\int_a^b f(t)\, dB(t)\right)(\omega) = (RS)\int_a^b f(t)\, dB(t,\omega),$$

where the left-hand side is the Wiener integral of f and the right-hand side is the Riemann–Stieltjes integral of f defined by Equation (2.3.1).

Proof. For each partition $\Delta_n = \{t_0, t_1, \ldots, t_{n-1}, t_n\}$ of $[a,b]$, we define a step function f_n by

$$f_n = \sum_{i=1}^n f(t_{i-1}) 1_{[t_{i-1}, t_i)}.$$

Note that f_n converges to f in $L^2[a,b]$ as $n \to \infty$, i.e., as $\|\Delta_n\| \to 0$. Hence by the definition of the Wiener integral in Equation (2.3.4),

$$\int_a^b f(t)\,dB(t) = \lim_{n\to\infty} \sum_{i=1}^n f(t_{i-1})\big(B(t_i) - B(t_{i-1})\big), \quad \text{in } L^2(\Omega). \tag{2.3.8}$$

On the other hand, by Equation (2.3.1), the following limit holds for each $\omega \in \Omega_0$ for some Ω_0 with $P(\Omega_0) = 1$,

$$\begin{aligned}
&(RS)\int_a^b f(t)\,dB(t,\omega) \\
&\quad = f(b)B(b,\omega) - f(a)B(a,\omega) - \lim_{n\to\infty}\sum_{i=1}^n B(t_i,\omega)\big(f(t_i) - f(t_{i-1})\big) \\
&\quad = \lim_{n\to\infty}\Big(f(b)B(b,\omega) - f(a)B(a,\omega) - \sum_{i=1}^n B(t_i,\omega)\big(f(t_i) - f(t_{i-1})\big)\Big),
\end{aligned}$$

which, after regrouping the terms, yields the following equality for each ω in Ω_0:

$$(RS)\int_a^b f(t)\,dB(t,\omega) = \lim_{n\to\infty}\sum_{i=1}^n f(t_{i-1})\big(B(t_i) - B(t_{i-1})\big). \tag{2.3.9}$$

Since $L^2(\Omega)$-convergence implies the existence of a subsequence converging almost surely, we can pick such a subsequence of $\{f_n\}$ to get the conclusion of the theorem from Equations (2.3.8) and (2.3.9). □

Example 2.3.8. Consider the Riemann integral $\int_0^1 B(t,\omega)\,dt$ defined for each $\omega \in \Omega_0$ for some Ω_0 with $P(\Omega_0) = 1$. Let us find the distribution of this random variable. Use the integration by parts formula to get

$$\begin{aligned}
\int_0^1 B(t,\omega)\,dt &= B(t,\omega)(t-1)\Big]_0^1 - \int_0^1 (t-1)\,dB(t,\omega) \\
&= (RS)\int_0^1 (1-t)\,dB(t,\omega).
\end{aligned}$$

Hence by Theorem 2.3.7 we see that for almost all $\omega \in \Omega$,

$$\int_0^1 B(t,\omega)\,dt = \Big(\int_0^1 (1-t)\,dB(t)\Big)(\omega),$$

where the right-hand side is a Wiener integral. Thus $\int_0^1 B(t)\,dt$ and the Wiener integral $\int_0^1 (1-t)\,dB(t)$ have the same distribution, which is easily seen to be Gaussian with mean 0 and variance

$$E\bigg(\int_0^1 (1-t)\,dB(t)\bigg)^2 = \int_0^1 (1-t)^2\,dt = \frac{1}{3}.$$

2.4 Conditional Expectation

In this section we explain the concept of conditional expectation, which will be needed in the next section and other places. Let $(\Omega, \mathcal{F}, P)$ be a fixed probability space. For $1 \leq p < \infty$, we will use $L^p(\Omega)$ to denote the space of all random variables X with $E(|X|^p) < \infty$. It is a Banach space with norm

$$\|X\|_p = \Big(E\big(|X|^p\big)\Big)^{1/p}.$$

In particular, $L^2(\Omega)$ is the Hilbert space used in Section 2.3. In this section we use the space $L^1(\Omega)$ with norm given by $\|X\|_1 = E|X|$. Sometimes we will write $L^1(\Omega, \mathcal{F})$ when we want to emphasize the σ-field $\mathcal{F}$.

Suppose we have another σ-field $\mathcal{G} \subset \mathcal{F}$. Let X be a random variable with $E|X| < \infty$, i.e., $X \in L^1(\Omega)$. Define a real-valued function μ on $\mathcal{G}$ by

$$\mu(A) = \int_A X(\omega)\, dP(\omega), \quad A \in \mathcal{G}. \tag{2.4.1}$$

Note that $|\mu(A)| \leq \int_A |X|\, dP \leq \int_\Omega |X|\, dP = E|X|$ for all $A \in \mathcal{G}$. Moreover, the function μ satisfies the following conditions:

(a) $\mu(\emptyset) = 0$;
(b) $\mu\big(\cup_{n\geq 1} A_n\big) = \sum_{n\geq 1} \mu(A_n)$ for any disjoint sets $A_n \in \mathcal{G}$, $n = 1, 2, \ldots$;
(c) If $P(A) = 0$ and $A \in \mathcal{G}$, then $\mu(A) = 0$.

A function $\mu\colon \mathcal{G} \to \mathbb{R}$ satisfying conditions (a) and (b) is called a *signed measure* on $(\Omega, \mathcal{G})$. A signed measure μ is said to be *absolutely continuous* with respect to P if it satisfies condition (c). Therefore, the function μ defined in Equation (2.4.1) is a signed measure on $(\Omega, \mathcal{G})$ and is absolutely continuous with respect to P.

Apply the Radon–Nikodym theorem (see, e.g., the book by Royden [73]) to the signed measure μ defined in Equation (2.4.1) to get a $\mathcal{G}$-measurable random variable Y with $E|Y| < \infty$ such that

$$\mu(A) = \int_A Y(\omega)\, dP(\omega), \quad \forall A \in \mathcal{G}. \tag{2.4.2}$$

Suppose $\widetilde{Y}$ is another such random variable, namely, it is $\mathcal{G}$-measurable with $E|\widetilde{Y}| < \infty$ and satisfies

$$\mu(A) = \int_A \widetilde{Y}(\omega)\, dP(\omega), \quad \forall A \in \mathcal{G}. \tag{2.4.3}$$

Then by Equations (2.4.2) and (2.4.3), we have $\int_A (Y - \widetilde{Y})\, dP = 0$ for all $A \in \mathcal{G}$. This implies that $Y = \widetilde{Y}$ almost surely.

The above discussion shows the existence and uniqueness of the conditional expectation in the next definition.

Definition 2.4.1. *Let $X \in L^1(\Omega, \mathcal{F})$. Suppose $\mathcal{G}$ is a σ-field and $\mathcal{G} \subset \mathcal{F}$. The* conditional expectation of X given $\mathcal{G}$ *is defined to be the unique random variable Y (up to P-measure 1) satisfying the following conditions:*

(1) *Y is $\mathcal{G}$-measurable;*

(2) *$\int_A X\,dP = \int_A Y\,dP$ for all $A \in \mathcal{G}$.*

We will freely use $E[X|\mathcal{G}]$, $E(X|\mathcal{G})$, or $E\{X|\mathcal{G}\}$ to denote the conditional expectation of X given $\mathcal{G}$. Notice that the $\mathcal{G}$-measurability in condition (1) is a crucial requirement. Otherwise, we could take $Y = X$ to satisfy condition (2), and the above definition would not be so meaningful. The conditional expectation $E[X|\mathcal{G}]$ can be interpreted as the best guess of the value of X based on the information provided by $\mathcal{G}$.

Example 2.4.2. Suppose $\mathcal{G} = \{\emptyset, \Omega\}$. Let X be a random variable in $L^1(\Omega)$ and let $Y = E[X|\mathcal{G}]$. Since Y is $\mathcal{G}$-measurable, it must be a constant, say $Y = c$. Then use condition (2) in Definition 2.4.1 with $A = \Omega$ to get

$$\int_\Omega X\,dP = \int_\Omega Y\,dP = c.$$

Hence $c = EX$ and we have $E[X|\mathcal{G}] = EX$. This conclusion is intuitively obvious. Since the σ-field $\mathcal{G} = \{\emptyset, \Omega\}$ provides no information, the best guess of the value of X is its expectation.

Example 2.4.3. Suppose $\Omega = \cup_n A_n$ is a disjoint union (finite or countable) with $P(A_n) > 0$ for each n. Let $\mathcal{G} = \sigma\{A_1, A_2, \ldots\}$, the σ-field generated by the A_n's. Let $X \in L^1(\Omega)$ and $Y = E[X|\mathcal{G}]$. Since Y is $\mathcal{G}$-measurable, it must be constant, say c_n, on A_n for each n. Use condition (2) in Definition 2.4.1 with $A = A_n$ to show that $c_n = P(A_n)^{-1} \int_{A_n} X\,dP$. Therefore, $E[X|\mathcal{G}]$ is given by

$$E[X|\mathcal{G}] = \sum_n \left(\frac{1}{P(A_n)} \int_{A_n} X\,dP \right) 1_{A_n},$$

where 1_{A_n} denotes the characteristic function of A_n.

Example 2.4.4. Let Z be a discrete random variable taking values $a_1, a_2, \ldots$ (finite or countable). Let $\sigma\{Z\}$ be the σ-field generated by Z. Then

$$\sigma\{Z\} = \sigma\{A_1, A_2, \ldots\},$$

where $A_n = \{Z = a_n\}$. Let $X \in L^1(\Omega)$. We can use Example 2.4.3 to obtain

$$E\big[X|\sigma\{Z\}\big] = \sum_n \left(\frac{1}{P(A_n)} \int_{A_n} X\,dP \right) 1_{A_n},$$

which can be rewritten as $E\big[X|\sigma\{Z\}\big] = \theta(Z)$ with the function θ defined by

$$\theta(x) = \begin{cases} \dfrac{1}{P(Z = a_n)} \displaystyle\int_{Z=a_n} X\,dP, & \text{if } x = a_n,\ n \geq 1; \\ 0, & \text{if } x \notin \{a_1, a_2, \ldots\}. \end{cases}$$

Note that the conditional expectation $E[X|\mathcal{G}]$ is a random variable, while the expectation EX is a real number. Below we list several properties of conditional expectation and leave most of the proofs as exercises at the end of this chapter.

Recall that $(\Omega, \mathcal{F}, P)$ is a fixed probability space. The random variable X below is assumed to be in $L^1(\Omega, \mathcal{F})$ and $\mathcal{G}$ is a sub-σ-field of $\mathcal{F}$, namely, $\mathcal{G}$ is a σ-field and $\mathcal{G} \subset \mathcal{F}$. All equalities and inequalities below hold almost surely.

1. $E\big(E[X|\mathcal{G}]\big) = EX$.

   Remark: Hence the conditional expectation $E[X|\mathcal{G}]$ and X have the same expectation. When written in the form $EX = E\big(E[X|\mathcal{G}]\big)$, the equality is often referred to as *computing expectation by conditioning*. To prove this equality, simply put $A = \Omega$ in condition (2) of Definition 2.4.1.
2. *If X is $\mathcal{G}$-measurable, then $E[X|\mathcal{G}] = X$.*
3. *If X and $\mathcal{G}$ are independent, then $E[X|\mathcal{G}] = EX$.*

   Remark: Here X and $\mathcal{G}$ being independent means that $\{X \in U\}$ and A are independent events for any Borel subset U of $\mathbb{R}$ and $A \in \mathcal{G}$, or equivalently, the events $\{X \leq x\}$ and A are independent for any $x \in \mathbb{R}$ and $A \in \mathcal{G}$.
4. *If Y is $\mathcal{G}$-measurable and $E|XY| < \infty$, then $E[XY|\mathcal{G}] = YE[X|\mathcal{G}]$.*
5. *If $\mathcal{H}$ is a sub-σ-field of $\mathcal{G}$, then $E[X|\mathcal{H}] = E\big[E[X|\mathcal{G}] \,|\, \mathcal{H}\big]$.*

   Remark: This property is useful when X is a product of random variables. In that case, in order to find $E[X|\mathcal{H}]$, we can use some factors in X to choose a suitable σ-field $\mathcal{G}$ between $\mathcal{H}$ and $\mathcal{F}$ and then apply this property.
6. *If $X, Y \in L^1(\Omega)$ and $X \leq Y$, then $E[X|\mathcal{G}] \leq E[Y|\mathcal{G}]$.*
7. $\big|E[X|\mathcal{G}]\big| \leq E[|X| \,|\, \mathcal{G}]$.

   Remark: For the proof, let $X^+ = \max\{X, 0\}$ and $X^- = -\min\{X, 0\}$ be the positive and negative parts of X, respectively. Then apply Property 6 to X^+ and X^-.
8. $E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}]$, $\forall a, b \in \mathbb{R}$ *and* $X, Y \in L^1(\Omega)$.

   Remark: By Properties 7 and 8, the conditional expectation $E[\,\cdot\,|\mathcal{G}]$ is a bounded linear operator from $L^1(\Omega, \mathcal{F})$ into $L^1(\Omega, \mathcal{G})$
9. (Conditional Fatou's lemma) *Let $X_n \geq 0$, $X_n \in L^1(\Omega)$, $n = 1, 2, \ldots,$ and assume that $\liminf_{n\to\infty} X_n \in L^1(\Omega)$. Then*
$$E\Big[\liminf_{n\to\infty} X_n \,\Big|\, \mathcal{G}\Big] \leq \liminf_{n\to\infty} E[X_n|\mathcal{G}].$$
10. (Conditional monotone convergence theorem) *Let $0 \leq X_1 \leq X_2 \leq \cdots \leq X_n \leq \cdots$ and assume that $X = \lim_{n\to\infty} X_n \in L^1(\Omega)$. Then*
$$E[X|\mathcal{G}] = \lim_{n\to\infty} E[X_n|\mathcal{G}].$$

11. (Conditional Lebesgue dominated convergence theorem) *Assume that* $|X_n| \leq Y$, $Y \in L^1(\Omega)$, *and* $X = \lim_{n\to\infty} X_n$ *exists almost surely. Then*

$$E[X|\mathcal{G}] = \lim_{n\to\infty} E[X_n|\mathcal{G}].$$

12. (Conditional Jensen's inequality) *Let* $X \in L^1(\Omega)$. *Suppose* ϕ *is a convex function on* $\mathbb{R}$ *and* $\phi(X) \in L^1(\Omega)$. *Then*

$$\phi\big(E[X|\mathcal{G}]\big) \leq E[\phi(X)|\mathcal{G}].$$

2.5 Martingales

Let $f \in L^2[a, b]$ and consider the stochastic process defined by

$$M_t = \int_a^t f(s)\, dB(s), \quad a \leq t \leq b. \tag{2.5.1}$$

We will show that M_t is a martingale. But first we review the concept of the martingale. Let T be either an interval in $\mathbb{R}$ or the set of positive integers.

Definition 2.5.1. *A* filtration *on* T *is an increasing family* $\{\mathcal{F}_t |\, t \in T\}$ *of* σ*-fields. A stochastic process* $X_t, t \in T$, *is said to be* adapted *to* $\{\mathcal{F}_t |\, t \in T\}$ *if for each* t, *the random variable* X_t *is* $\mathcal{F}_t$*-measurable.*

Remark 2.5.2. A σ-field $\mathcal{F}$ is called *complete* if $A \in \mathcal{F}$ and $P(A) = 0$ imply that $B \in \mathcal{F}$ for any subset B of A. We will always assume that all σ-fields $\mathcal{F}_t$ are complete.

Definition 2.5.3. *Let* X_t *be a stochastic process adapted to a filtration* $\{\mathcal{F}_t\}$ *and* $E|X_t| < \infty$ *for all* $t \in T$. *Then* X_t *is called a* martingale *with respect to* $\{\mathcal{F}_t\}$ *if for any* $s \leq t$ *in* T,

$$E\{X_t |\, \mathcal{F}_s\} = X_s, \quad \text{a.s. } (\textit{almost surely}). \tag{2.5.2}$$

In case the filtration is not explicitly specified, then the filtration $\{\mathcal{F}_t\}$ is understood to be the one given by $\mathcal{F}_t = \sigma\{X_s\, ;\, s \leq t\}$.

The concept of the martingale is a generalization of the sequence of partial sums arising from a sequence $\{X_n\}$ of independent and identically distributed random variables with mean 0. Let $S_n = X_1 + \cdots + X_n$. Then the sequence $\{S_n\}$ is a martingale.

Submartingale and supermartingale are defined by replacing the equality in Equation (2.5.2) with $\geq$ and $\leq$, respectively, i.e., for any $s \leq t$ in T,

$$E\{X_t |\, \mathcal{F}_s\} \geq X_s, \quad \text{a.s.} \quad \text{(submartingale)},$$
$$E\{X_t |\, \mathcal{F}_s\} \leq X_s, \quad \text{a.s.} \quad \text{(supermartingale)}.$$

Let $\{X_n\}$ be a sequence of independent and identically distributed random variables with finite expectation and let $S_n = X_1 + \cdots + X_n$. Then $\{S_n\}$ is a submartingale if $EX_1 \geq 0$ and a supermartingale if $EX_1 \leq 0$.

A Brownian motion $B(t)$ is a martingale. To see this fact, let

$$\mathcal{F}_t = \sigma\{B(s)\,;\, s \leq t\}.$$

Then for any $s \leq t$,

$$E\{B(t) | \mathcal{F}_s\} = E\{B(t) - B(s) | \mathcal{F}_s\} + E\{B(s) | \mathcal{F}_s\}.$$

Since $B(t) - B(s)$ is independent of $\mathcal{F}_s$, we have $E\{B(t) - B(s) | \mathcal{F}_s\} = E\{B(t) - B(s)\}$. But $EB(t) = 0$ for any t. Hence $E\{B(t) - B(s) | \mathcal{F}_s\} = 0$. On the other hand, $E\{B(s) | \mathcal{F}_s\} = B(s)$ because $B(s)$ is $\mathcal{F}_s$-measurable. Thus $E\{B(t) | \mathcal{F}_s\} = B(s)$ for any $s \leq t$ and this shows that $B(t)$ is a martingale. In fact, it is the most basic martingale stochastic process with time parameter in an interval.

Now we return to the stochastic process M_t defined in Equation (2.5.1) and show that it is a martingale in the next theorem.

Theorem 2.5.4. *Let $f \in L^2[a, b]$. Then the stochastic process*

$$M_t = \int_a^t f(s)\, dB(s), \quad a \leq t \leq b,$$

is a martingale with respect to $\mathcal{F}_t = \sigma\{B(s)\,;\, s \leq t\}$.

Proof. First we need to show that $E|M_t| < \infty$ for all $t \in [a, b]$ in order to take the conditional expectation of M_t. Apply Theorem 2.3.4 to get

$$E\big(|M_t|^2\big) = \int_a^t |f(s)|^2\, ds \leq \int_a^b |f(s)|^2\, ds.$$

Hence $E|M_t| \leq \big\{E\big(|M_t|^2\big)\big\}^{1/2} < \infty$. Next we need to prove that $E\{M_t | \mathcal{F}_s\} = M_s$ a.s. for any $s \leq t$. But

$$M_t = M_s + \int_s^t f(u)\, dB(u)$$

and M_s is $\mathcal{F}_s$-measurable. Hence

$$E\{M_t | \mathcal{F}_s\} = M_s + E\bigg\{\int_s^t f(u)\, dB(u) \bigg| \mathcal{F}_s\bigg\}.$$

Thus it suffices to show that for any $s \leq t$,

$$E\bigg\{\int_s^t f(u)\, dB(u) \bigg| \mathcal{F}_s\bigg\} = 0. \tag{2.5.3}$$

First suppose f is a step function $f = \sum_{i=1}^n a_i \, 1_{[t_{i-1}, t_i)}$, where $t_0 = s$ and $t_n = t$. In this case, we have

$$\int_s^t f(u)\, dB(u) = \sum_{i=1}^n a_i \big(B(t_i) - B(t_{i-1})\big).$$

But $B(t_i) - B(t_{i-1})$, $i = 1, \ldots, n$, are all independent of the σ-field $\mathcal{F}_s$. Hence $E\{B(t_i) - B(t_{i-1}) \,|\, \mathcal{F}_s\} = 0$ for all i and so Equation (2.5.3) holds.

Next suppose $f \in L^2[a, b]$. Choose a sequence $\{f_n\}_{n=1}^\infty$ of step functions converging to f in $L^2[a, b]$. Then by the conditional Jensen's inequality with $\phi(x) = x^2$ in Section 2.4 we have the inequality

$$|E\{X \,|\, \mathcal{F}\}|^2 \le E\{X^2 \,|\, \mathcal{F}\},$$

which implies that

$$\begin{aligned}
&\left| E\left\{ \int_s^t \big(f_n(u) - f(u)\big)\, dB(u) \,\Big|\, \mathcal{F}_s \right\} \right|^2 \\
&\qquad \le E\left\{ \left(\int_s^t \big(f_n(u) - f(u)\big)\, dB(u) \right)^2 \,\Big|\, \mathcal{F}_s \right\}.
\end{aligned}$$

Next we use the property $E\big(E\{X \,|\, \mathcal{F}\}\big) = EX$ of conditional expectation and then apply Theorem 2.3.4 to get

$$\begin{aligned}
E\left| E\left\{ \int_s^t \big(f_n(u) - f(u)\big)\, dB(u) \,\Big|\, \mathcal{F}_s \right\} \right|^2 &\le \int_s^t \big(f_n(u) - f(u)\big)^2\, du \\
&\le \int_a^b \big(f_n(u) - f(u)\big)^2\, du \\
&\to 0,
\end{aligned}$$

as $n \to \infty$. Hence the sequence $E\{\int_s^t f_n(u)\, dB(u) \,|\, \mathcal{F}_s\}$ of random variables converges to $E\{\int_s^t f(u)\, dB(u) \,|\, \mathcal{F}_s\}$ in $L^2(\Omega)$. Note that the convergence of a sequence in $L^2(\Omega)$ implies convergence in probability, which implies the existence of a subsequence converging almost surely. Hence by choosing a subsequence if necessary, we can conclude that with probability 1,

$$\lim_{n\to\infty} E\left\{ \int_s^t f_n(u)\, dB(u) \,\Big|\, \mathcal{F}_s \right\} = E\left\{ \int_s^t f(u)\, dB(u) \,\Big|\, \mathcal{F}_s \right\}. \tag{2.5.4}$$

Now $E\big\{ \int_s^t f_n(u)\, dB(u) \,\big|\, \mathcal{F}_s \big\} = 0$ since we have already shown that Equation (2.5.3) holds for step functions. Hence by Equation (2.5.4),

$$E\left\{ \int_s^t f(u)\, dB(u) \,\Big|\, \mathcal{F}_s \right\} = 0,$$

and so Equation (2.5.3) holds for any $f \in L^2[a, b]$. □

2.6 Series Expansion of Wiener Integrals

Let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal basis for the Hilbert space $L^2[a,b]$. Each $f \in L^2[a,b]$ has the following expansion:

$$f = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \phi_n, \tag{2.6.1}$$

where $\langle \cdot, \cdot \rangle$ is the inner product on $L^2[a,b]$ given by $\langle f, g \rangle = \int_a^b f(t)g(t)\,dt$. Moreover, we have the Parseval identity

$$\|f\|^2 = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle^2. \tag{2.6.2}$$

Take the Wiener integral in both sides of Equation (2.6.1) and informally interchange the order of integration and summation to get

$$\int_a^b f(t)\,dB(t) = \sum_{n=1}^{\infty} \langle f, \phi_n \rangle \int_a^b \phi_n(t)\,dB(t). \tag{2.6.3}$$

Question 2.6.1. Does the random series in the right-hand side converge to the left-hand side and in what sense?

First observe that by Theorem 2.3.4 and the remark following Equation (2.3.5), the random variables $\int_a^b \phi_n(t)\,dB(t), n \geq 1$, are independent and have the Gaussian distribution with mean 0 and variance 1. Thus the right-hand side of Equation (2.6.3) is a random series of independent and identically distributed random variables. By the Lévy equivalence theorem [10] [37] this random series converges almost surely if and only if it converges in probability and, in turn, if and only if it converges in distribution. On the other hand, we can easily check the $L^2(\Omega)$ convergence of this random series as follows. Apply Equations (2.3.5) and (2.6.2) to show that

$$\begin{aligned} E\Big(\int_a^b f(t)\,dB(t) &- \sum_{n=1}^{N} \langle f, \phi_n \rangle \int_a^b \phi_n(t)\,dB(t)\Big)^2 \\ &= \int_a^b f(t)^2\,dt - 2\sum_{n=1}^{N} \langle f, \phi_n \rangle^2 + \sum_{n=1}^{N} \langle f, \phi_n \rangle^2 \\ &= \int_a^b f(t)^2\,dt - \sum_{n=1}^{N} \langle f, \phi_n \rangle^2 \\ &\to 0, \end{aligned}$$

as $N \to \infty$. Hence the random series in Equation (2.6.3) converges in $L^2(\Omega)$ to the random variable in the left-hand side of Equation (2.6.3). But the $L^2(\Omega)$ convergence implies convergence in probability. Therefore we have proved the next theorem for the series expansion of the Wiener integral.

Theorem 2.6.2. *Let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal basis for $L^2[a,b]$. Then for each $f \in L^2[a,b]$, the Wiener integral of f has the series expansion*

$$\int_a^b f(t)\,dB(t) = \sum_{n=1}^{\infty} \langle f, \phi_n\rangle \int_a^b \phi_n(t)\,dB(t),$$

with probability 1*, where the random series converges almost surely.*

In particular, apply the theorem to $a=0,\ b=1$, and $f = 1_{[0,t)},\ 0 \le t \le 1$. Then $\int_0^1 f(s)\,dB(s) = B(t)$ and we have the random series expansion,

$$B(t,\omega) = \sum_{n=1}^{\infty} \Big(\int_0^t \phi_n(s)\,ds \Big) \Big(\int_0^1 \phi_n(s)\,dB(s,\omega) \Big).$$

Note that the variables t and ω are separated in the right-hand side. In view of this expansion, we expect that $B(t)$ can be represented by

$$B(t,\omega) = \sum_{n=1}^{\infty} \xi_n(\omega) \int_0^t \phi_n(s)\,ds,$$

where $\{\xi_n\}_{n=1}^{\infty}$ is a sequence of independent random variables having the same Gaussian distribution with mean 0 and variance 1. This method of defining a Brownian motion has been studied in [29] [41] [67].

Exercises

1. Let $B(t)$ be a Brownian motion. Show that $E|B(s)-B(t)|^4 = 3|s-t|^2$.
2. Show that the marginal distribution of a Brownian motion $B(t)$ at times $0 < t_1 < t_2 < \cdots t_n$ is given by

$$\begin{aligned} &P\{B(t_1) \le a_1, B(t_2) \le a_2, \ldots, B(t_n) \le a_n\} \\ &\quad = \frac{1}{\sqrt{(2\pi)^n t_1 (t_2 - t_1) \cdots (t_n - t_{n-1})}} \int_{-\infty}^{a_n} \cdots \int_{-\infty}^{a_1} \\ &\qquad \exp\Big[-\frac{1}{2}\Big(\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \cdots + \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}} \Big) \Big]\, dx_1 dx_2 \cdots dx_n. \end{aligned}$$

3. Let $B(t)$ be a Brownian motion. For fixed t and s, find the distribution function of the random variable $X = B(t) + B(s)$.
4. Let $B(t)$ be a Brownian motion and let $0 < s \le t \le u \le v$. Show that the random variables $\frac{1}{t}B(t) - \frac{1}{s}B(s)$ and $aB(u) + bB(v)$ are independent for any $a, b \in \mathbb{R}$.

5. Let $B(t)$ be a Brownian motion and let $0 < s \leq t \leq u \leq v$. Show that the random variables $aB(s) + bB(t)$ and $\frac{1}{v}B(v) - \frac{1}{u}B(u)$ are independent for any $a, b \in \mathbb{R}$ satisfying the condition $as + bt = 0$.
6. Let $B(t)$ be a Brownian motion. Show that $\lim_{t \to 0^+} tB(1/t) = 0$ almost surely. Define $W(0) = 0$ and $W(t) = tB(1/t)$ for $t > 0$. Prove that $W(t)$ is a Brownian motion.
7. Let $B(t)$ be a Brownian motion. Find all constants a and b such that $X(t) = \int_0^t \left(a + b\frac{u}{t}\right) dB(u)$ is also a Brownian motion.
8. Let $B(t)$ be a Brownian motion. Find all constants a, b, and c such that $X(t) = \int_0^t \left(a + b\frac{u}{t} + c\frac{u^2}{t^2}\right) dB(u)$ is also a Brownian motion.
9. Let $B(t)$ be a Brownian motion. Show that for any integer $n \geq 1$, there exist nonzero constants $a_0, a_1, \ldots, a_n$ such that $X(t) = \int_0^t \left(a_0 + a_1\frac{u}{t} + a_2\frac{u^2}{t^2} + \cdots + a_n\frac{u^n}{t^n}\right) dB(u)$ is also a Brownian motion.
10. Let $B(t)$ be a Brownian motion. Show that both $X(t) = \int_0^t (2t - u)\, dB(u)$ and $Y(t) = \int_0^t (3t - 4u)\, dB(u)$ are Gaussian processes with mean function 0 and the same covariance function $3s^2t - \frac{2}{3}s^3$ for $s \leq t$.
11. Let $B(t) = (B_1(t), \ldots, B_n(t))$ be an $\mathbb{R}^n$-valued Brownian motion. Find the density functions of $R(t) = |B(t)|$ and $S(t) = |B(t)|^2$.
12. For each $n \geq 1$, let X_n be a Gaussian random variable with mean μ_n and variance σ_n^2. Suppose the sequence X_n converges to X in $L^2(\Omega)$. Show that the limits $\mu = \lim_{n\to\infty} \mu_n$ and $\sigma^2 = \lim_{n\to\infty} \sigma_n^2$ exist and that X is a Gaussian random variable with mean μ and variance σ^2.
13. Let $f(x, y)$ be the joint density function of random variables X and Y. The *marginal density function* of Y is given by $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$. The *conditional density function* of X given $Y = y$ is defined by $f_{X|Y}(x|y) = f(x, y)/f_Y(y)$. The *conditional expectation* of X given $Y = y$ is defined by $E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx$. Let $\sigma(Y)$ be the σ-field generated by Y. Prove that
$$E[X|\sigma(Y)] = \theta(Y),$$
where θ is the function $\theta(y) = E[X|Y = y]$.
14. Prove the properties of conditional expectation listed in Section 2.4.
15. Let $B(t)$ be a Brownian motion. Find the distribution of $\int_0^t e^{t-s}\, dB(s)$. Check whether $X_t = \int_0^t e^{t-s}\, dB(s)$ is a martingale.
16. Let $B(t)$ be a Brownian motion. Find the distribution of $\int_0^t B(s)\, ds$. Check whether $Y_t = \int_0^t B(s)\, ds$ is a martingale.
17. Let $B(t)$ be a Brownian motion. Find the distribution of the integral $\int_0^t B(s) \cos(t - s)\, ds$.
18. Let $B(t)$ be a Brownian motion. Show that $X_t = \frac{1}{3}B(t)^3 - \int_0^t B(s)\, ds$ is a martingale.

http://www.springer.com/978-0-387-28720-1

Introduction to Stochastic Integration
Kuo, H.-H.
2006, XIII, 279 p. 2 illus., Softcover
ISBN: 978-0-387-28720-1