1

Quadratic Functionals of Brownian Motion

Brownian motion (Bm) has been a central topic, not only in probability theory but also in various applied fields. Throughout this book we deal with Bm from a statistical viewpoint. In particular, we deal with quadratic functionals of Bm or ratios of those functionals, and consider their distributions, where functionals are expressed by the Riemann–Stieltjes double integral with respect to Bm. Since it is hard in general to derive distribution functions of such functionals explicitly, we attempt to derive the associated characteristic function (c.f.). For this purpose we use the theory of integral equations of Fredholm type, among which is the Fredholm determinant. The notion of the resolvent is also necessary if we deal with quadratic plus linear or bilinear functionals of Bm. In this chapter we give an introductory discussion on these. In particular, we indicate by some theorems and simple examples how to relate the c.f. with the Fredholm determinant and the resolvent.

1.1 Brownian Motion and Some Statistical Properties

Brownian motion (Bm) plays a fundamental role in subsequent discussions. We thus start by defining Bm. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space on which Bm is defined.

Definition 1.1 (Brownian motion) The stochastic process $\{W(t)\}$ for $t \in [0, 1]$ is called Bm if the following conditions are satisfied.

(1) $P(W(0) = 0) = 1$.
(2) Increments $W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ are independent for any positive integer n and time points $0 \leq t_0 < t_1 < \cdots < t_n \leq 1$.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

(3) $W(t) - W(s) \sim \mathrm{N}(0,\ t-s)$ for any $0 \le s < t \le 1$, where $\mathrm{N}(\mu,\ \sigma^2)$ stands for a normal distribution with mean μ and variance σ^2.

Note that the parameter t denotes time and extends over the unit interval $[0, 1]$ throughout this book, unless otherwise stated. Note also that we shall always use $\{W(t)\}$ for Bm without mentioning its definition. It follows from (3) above that Bm is a zero-mean Gaussian process, which means that the finite-dimensional distributions of $W(t_1), \ldots, W(t_n)$ for each collection $0 \le t_1 < \cdots < t_n \le 1$ are multivariate normal with means 0 for all n. Thus it suffices to know only the covariance $\mathrm{Cov}(W(s),\ W(t))$ to determine the distributions. We have from (2) above, for $s < t$,

$$\begin{aligned}\mathrm{Cov}(W(s),\ W(t)) &= \mathrm{E}(W(s)\,W(t)) = \mathrm{E}\,[W(s)\,(W(t) - W(s) + W(s))]\\ &= \mathrm{E}\,[W(s)\,(W(t) - W(s))] + \mathrm{E}(W^2(s)) = s,\end{aligned}$$

so that $\mathrm{Cov}(W(s),\ W(t)) = \min(s, t)$. It also follows from (2) that the increment $W(t) - W(s)$ for $t > s$ is independent of the past $W(u)$ for $u \le s$. More generally $W(t)$ is independent of $\mathcal{F}_s$, the σ-field generated by $W(u)$ for $u \le s$, which implies that Bm is a *martingale* satisfying

$$\mathrm{E}(|W(t)|) = \sqrt{\frac{2\,t}{\pi}} < \infty, \quad \mathrm{E}(W(t)|\mathcal{F}_s) = W(s) \quad (s \le t).$$

The mean square (m.s.) continuity of Bm is another consequence from the definition. It holds that

$$\lim_{h \to 0}\ \mathrm{E}\left[(W(t+h) - W(t))^2\right] = \lim_{h \to 0}\ |h| = 0,$$

which says that $W(t)$ is continuous at t in the m.s. sense. This is abbreviated as

$$\underset{h \to 0}{\mathrm{l.i.m.}}\ W(t+h) = W(t),$$

where "l.i.m." stands for limit in mean square. The second moment is necessary for the m.s. limit to exist. It follows that $W(t)$ is not m.s. differentiable in the sense that

$$\underset{h \to 0}{\mathrm{l.i.m.}}\ \frac{W(t+h) - W(t)}{h}$$

does not exist (Exercise 1.1.1 in this section).

Bm is also continuous with probability 1. In fact, it holds that

$$\mathrm{E}\left[(W(t) - W(s))^4\right] = 3\,(t-s)^2,$$

which implies that $W(t)$ is continuous with probability 1 because of Kolmogorov's continuity criterion (Liptser and Shiryaev, 2001a, p.23).

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Some other properties (a) through (d) of Bm follow (Exercises 1.1.2, 1.1.3, 1.1.4).

(a) If $\{X(t)\}$ is a zero-mean Gaussian process with covariance function given by $\mathrm{Cov}(X(s), X(t)) = \min(s,t)$, then $\{X(t)\}$ is Bm.

(b) $W(t)$ is of unbounded variation, that is, for any $0 \le t_0 < t_1 < \cdots < t_n \le 1$, and $\Delta_n = \max_{1\le k\le n}(t_k - t_{k-1})$,

$$\lim_{\Delta_n \to 0} \sum_{k=1}^{n} \mathrm{E}\left(|W(t_k) - W(t_{k-1})|\right) = \infty.$$

(c) The quadratic variation is finite and bounded away from 0. More specifically, for any $0 \le a = t_0 < t_1 < \cdots < t_n = b \le 1$, and $\Delta_n = \max_{1\le k\le n}(t_k - t_{k-1})$,

$$\lim_{\Delta_n \to 0} \sum_{k=1}^{n} \mathrm{E}\left(|W(t_k) - W(t_{k-1})|^2\right) = b - a.$$

(d) The stochastic process $X(t)$ defined in the m.s. sense by

$$X(t) = \sum_{n=1}^{\infty} \frac{\sqrt{2}\,\sin\left(n - \frac{1}{2}\right)\pi\, t}{\left(n - \frac{1}{2}\right)\pi} Z_n \tag{1.1}$$

is Bm, where $\{Z_n\}$ is a sequence of independent N(0, 1) random variables, which is denoted as $\{Z_n\} \sim \mathrm{NID}(0,1)$. The expansion in (1.1) is called the *Karhunen–Loève expansion*, which will be explained in detail in Chapter 2.

If $W(t)$ is conditioned on $P(W(1) = 0) = 1$, the resulting process is called the *Brownian bridge* (Bb), which may be defined as follows:

Definition 1.2 (Brownian bridge) The stochastic process $\{\tilde{W}(t)\}$ for $t \in [0,1]$ is called the Bb if the following condition is satisfied.

(1) $\{\tilde{W}(t)\}$ is a zero-mean Gaussian process with covariance function given by $\mathrm{Cov}(\tilde{W}(s), \tilde{W}(t)) = \min(s,t) - st$.

The increments of the Bb are not independent, unlike Bm. In fact,

$$\mathrm{Cov}\left(\tilde{W}(t_2) - \tilde{W}(t_1), \tilde{W}(t_4) - \tilde{W}(t_3)\right) = -(t_2 - t_1)(t_4 - t_3) < 0$$

for $0 \le t_1 < t_2 < t_3 < t_4 \le 1$. Some other related properties (e), (f), and (g) of the Bb follow (Exercises 1.1.5 and 1.1.6).

(e) The stochastic process $\{W(t) - tW(1)\}$ is the Bb.

(f) $\mathrm{Cov}(W(s), W(t)|W(1) = 0) = \min(s, t) - st$.

(g) The stochastic process $X(t)$ defined in the m.s. sense by

$$X(t) = \sum_{n=1}^{\infty} \frac{\sqrt{2}\,\sin n\pi t}{n\,\pi}\, Z_n \tag{1.2}$$

is the Bb, where $\{Z_n\} \sim \mathrm{NID}(0, 1)$. This is the Karhunen–Loève expansion for the Bb.

We note in passing that Bm is a continuous-time version of the *random walk*. More specifically, suppose that the random walk is defined by

$$y_j = y_{j-1} + \varepsilon_j = \sum_{i=1}^{j} \varepsilon_i \qquad (j = 1, \ldots, n), \tag{1.3}$$

where $y_0 = 0$ and $\{\varepsilon_i\}$ is a sequence of independent and identically distributed random variables with mean 0 and variance σ^2, which is abbreviated as $\{\varepsilon_i\} \sim$ i.i.d.$(0, \sigma^2)$. Define also the *partial sum process*

$$\begin{aligned} Y_n(t) &= \frac{1}{\sqrt{n}\sigma}\, y_{j-1} + n\left(t - \frac{j-1}{n}\right) \frac{1}{\sqrt{n}\sigma}\, \varepsilon_j \\ &= \frac{1}{\sqrt{n}\sigma}\, y_{[nt]} + (nt - [nt])\, \frac{1}{\sqrt{n}\sigma}\, \varepsilon_{[nt]+1}, \end{aligned} \tag{1.4}$$

where $(j-1)/n \le t \le j/n$ $(j = 1, \ldots, n)$, and $[x]$ denotes the largest integer not exceeding x. Note that $Y_n(t)$ is continuous on $[0, 1]$ by connecting $y_{j-1}/(\sqrt{n}\sigma)$ and $y_j/(\sqrt{n}\sigma)$ $(j = 1, \ldots, n)$. Then it holds that, as $n \to \infty$,

$$\{Y_n(t)\} \Rightarrow \{W(t)\}, \tag{1.5}$$

where $\Rightarrow$ stands for the *weak convergence*. This is the result established by Donsker (1951) and is referred to as *Donsker's theorem* or the *functional central limit theorem* (FCLT) because the weak convergence is concerned with the function space $C[0, 1]$, where $C[0, 1]$ is the space of all real-valued continuous functions defined on $[0, 1]$. Billingsley (1999) discusses details on the FCLT and weak convergence.

Our main subject in this book is to deal with functionals of Bm. In particular we consider functionals expressed by integrals. For this purpose we describe the stochastic integrals associated with Bm in the next section.

Exercises for Section 1.1

Exercise 1.1.1 Show that Bm is not m.s. differentiable.

Exercise 1.1.2 Prove that, if $\{X(t)\}$ is a zero-mean Gaussian process with covariance function given by $\mathrm{Cov}(X(s), X(t)) = \min(s, t)$, then $\{X(t)\}$ is Bm.

Exercise 1.1.3 Prove that

$$\lim_{\Delta_n \to 0} \sum_{k=1}^{n} \mathrm{E}\left(|W(t_k) - W(t_{k-1})|\right) = \infty,$$

$$\lim_{\Delta_n \to 0} \sum_{k=1}^{n} \mathrm{E}\left(|W(t_k) - W(t_{k-1})|^2\right) = b - a$$

for $a = t_0 < t_1 < \cdots < t_n = b$, and $\Delta_n = \max_{1 \le k \le n}(t_k - t_{k-1})$.

Exercise 1.1.4 Show that the stochastic process $X(t)$ defined in (1.1) is Bm, using the formula

$$\sum_{n=1}^{\infty} \frac{\cos\left(n - \frac{1}{2}\right)\pi x}{\left(n - \frac{1}{2}\right)^2 \pi^2} = \frac{1}{2}(1 - |x|) \qquad (|x| \le 2).$$

Exercise 1.1.5 Prove that Bm $W(t)$ conditioned on $W(1) = 0$ is the Bb. Prove also that the stochastic process $\{W(t) - tW(1)\}$ is a Bb that is independent of $W(1)$.

Exercise 1.1.6 Show that the stochastic process $X(t)$ defined in (1.2) is the Bb, using the formula

$$\sum_{n=1}^{\infty} \frac{\cos n\pi x}{n^2\pi^2} = \frac{1}{4}(|x| - 1)^2 - \frac{1}{12} \qquad (|x| \le 2).$$

1.2 Mean Square Integral Associated with Brownian Motion

Stochastic integrals associated with Bm play an important role in the development of subsequent discussions. We define such integrals in the m.s. sense. For this purpose we introduce the space L_2, which is the space of random variables defined on a common probability space with finite second moment. It is known that the space L_2 is complete, so that, if $X_n \in L_2$ and $X = \underset{n\to\infty}{\text{l.i.m.}}\, X_n$ exists, X also belongs to L_2 (Loève, 1977, p.163). Moreover, X_n converges in the m.s. sense if and only if $\mathrm{E}(X_m X_n)$ converges as $m, n \to \infty$ in any manner. We also note that "l.i.m." and "expectation" commute in the following sense. If $\underset{n\to\infty}{\text{l.i.m.}}\, X_n = X$ and $\underset{n\to\infty}{\text{l.i.m.}}\, Y_n = Y$, it holds (Exercise 1.2.1) that

$$\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X), \qquad \lim_{m,n\to\infty} \mathrm{E}(X_m Y_n) = \mathrm{E}(XY). \tag{1.6}$$

Using the above m.s. properties, we present stochastic integrals of two types in this section. Integrals of other types will be presented in the subsequent two sections.

1.2.1 Mean Square Riemann Integral

Let a stochastic process $\{X(t)\}$ belong to L_2. Then consider

$$Y = \int_0^1 X(t)\,dt = \underset{\substack{n\to\infty \\ \Delta_n \to 0}}{\text{l.i.m.}}\, Y_n, \quad Y_n = \sum_{i=1}^n X(t_i')(t_i - t_{i-1}), \tag{1.7}$$

where $\Delta_n = \max_{1\le i\le n}(t_i - t_{i-1})$ and $t_i' \in [t_{i-1}, t_i)$ $(i = 1, \ldots, n)$. The integral in (1.7) is called the *m.s. Riemann integral* of $X(t)$, and $X(t)$ is said to be m.s. Riemann integrable if $Y = \underset{n\to\infty}{\text{l.i.m.}}\, Y_n$ exists. It follows that a sufficient condition for the m.s. integral in (1.7) to be well defined is that $X(t)$ is m.s. continuous (Exercise 1.2.2). It can also be shown (Exercise 1.2.3) that, if $X(t)$ is m.s. continuous,

$$\mathrm{E}\left(\int_0^1 X(t)\,dt\right) = \int_0^1 \mathrm{E}(X(t))\,dt,$$

$$\mathrm{E}\left[\left(\int_0^1 X(t)\,dt\right)^2\right] = \int_0^1\int_0^1 \mathrm{E}(X(s)X(t))\,ds\,dt.$$

The following are three simple examples of the m.s. Riemann integral:

$$I_1 = \int_0^1 W(t)\,dt, \quad I_2 = \int_0^1 |W(t)|\,dt, \quad I_3 = \int_0^1 e^{W(t)}\,dt,$$

where I_1 is a linear functional of $W(t)$ so that it is normal because "l.i.m." and "normality" commute (Exercise 1.2.4), whereas I_2 and I_3 are not normal. We have (Exercise 1.2.5)

$$\mathrm{E}(I_1) = 0, \qquad \mathrm{E}(I_2) = \frac{2}{3}\sqrt{\frac{2}{\pi}}, \qquad \mathrm{E}(I_3) = 2(\sqrt{e} - 1),$$

$$\mathrm{E}\big(I_1^2\big) = \frac{1}{3}, \qquad \mathrm{E}\big(I_2^2\big) = \frac{3}{8}, \qquad \mathrm{E}\big(I_3^2\big) = \frac{2}{3}e^2 - \frac{8}{3}\sqrt{e} + 2.$$

The following are three examples of the m.s. Riemann integral of quadratic functions of Bm:

$$I_4 = \int_0^1 W^2(t)\,dt, \quad I_5 = \int_0^1 \tilde{W}^2(t)\,dt, \quad I_6 = \int_0^1 \bar{W}^2(t)\,dt,$$

where $\tilde{W}(t)$ is the Bb, whereas

$$\bar{W}(t) = W(t) - \int_0^1 W(s)\,ds$$

is called *demeaned Bm*. We have (Exercise 1.2.6)

$$\mathrm{E}(I_4) = \frac{1}{2}, \quad \mathrm{E}\big(I_4^2\big) = \frac{7}{12}, \quad \mathrm{E}(I_5) = \mathrm{E}(I_6) = \frac{1}{6}, \quad \mathrm{E}\big(I_5^2\big) = \mathrm{E}\big(I_6^2\big) = \frac{1}{20}.$$

It is noted that I_5 and I_6 have the same mean and variance. It will be shown later that the two have the same distribution.

Some other properties (a), (b), and (c) of the m.s. Riemann integral follow (Soong, 1973; Kuo, 2006; Klebaner, 2012).

(a) If $f(W(t))$ is m.s. continuous, $f(W(t))$ is m.s. Riemann integrable.
(b) The m.s. Riemann integral of $f(W(t))$, if it exists, is unique.
(c) If $f(W(t))$ is m.s. integrable, $X(t) = \int_0^t f(W(s))\,ds$ is m.s. differentiable with derivative given by $\dot{X}(t) = f(W(t))$.

1.2.2 Mean Square Riemann–Stieltjes Integral

Let $g(t)$ be a deterministic function of $t \in [0,1]$. Then we define

$$A = \int_0^1 g(t)\,dW(t) = \underset{\substack{n\to\infty \\ \Delta_n\to 0}}{\text{l.i.m.}} \sum_{i=1}^n g(t_i')(W(t_i) - W(t_{i-1})), \tag{1.8}$$

$$B = \int_0^1 W(t)\,dg(t) = \underset{\substack{n\to\infty \\ \Delta_n\to 0}}{\text{l.i.m.}} \sum_{i=1}^n W(t_i')(g(t_i) - g(t_{i-1})), \tag{1.9}$$

where $t_i' \in [t_{i-1}, t_i)$ $(i = 1, \ldots, n)$. The integrals in (1.8) and (1.9) are called the *m.s. Riemann–Stieltjes integrals.* Normality is retained in these integrals with $\mathrm{E}(A) = \mathrm{E}(B) = 0$ (Exercise 1.2.7). Note that the Riemann–Stieltjes integral in (1.9) reduces to the Riemann integral in (1.7) when $g(t)$ is differentiable with $|g'(t)| < \infty$ on $[0,1]$.

Because of the definition of the m.s. limit, the existence of the above integrals is determined by the existence of appropriate ordinary integrals, which leads us to

Theorem 1.3 *The m.s. Riemann–Stieltjes integrals in (1.8) and (1.9) are well defined if the following limits of the Riemann–Stieltjes double sums exist.*

$$\begin{aligned} E(A^2) &= \lim_{n\to\infty} \sum_{i=1}^n \sum_{j=1}^n g(t_i)g(t_j)E\left(\Delta W(t_i)\Delta W(t_j)\right) \\ &= \lim_{n\to\infty} \sum_{i=1}^n g^2(t_i)(t_i - t_{i-1}) \\ &= \int_0^1 g^2(t)\,dt, \end{aligned} \tag{1.10}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

$$
\begin{aligned}
E(B^2) &= \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left(W(t_i)W(t_j)\right) \Delta g(t_i)\Delta g(t_j) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \min(t_i, t_j)\Delta g(t_i)\Delta g(t_j) \\
&= \int_0^1 \int_0^1 \min(s,t)\, dg(s)\, dg(t), \qquad (1.11)
\end{aligned}
$$

where $\Delta W(t_i) = W(t_i) - W(t_{i-1})$ *and* $\Delta g(t_i) = g(t_i) - g(t_{i-1})$.

A sufficient condition for the existence of (1.8) is that $g(t)$ is continuous, whereas (1.9) exists if $g(t)$ is of bounded variation. Note that

$$
A \sim \mathrm{N}\left(0, \int_0^1 g^2(t)\, dt\right), \quad B \sim \mathrm{N}\left(0, \int_0^1 \int_0^1 \min(s,t)\, dg(s)\, dg(t)\right). \tag{1.12}
$$

Here are two examples of (1.8).

$$
A_1 = \int_0^1 dW(t), \qquad A_2 = \int_0^1 (1-t)\, dW(t).
$$

Then A_1 and A_2 are well defined with $g(t) = 1,\ 1-t$, respectively, in (1.8). Thus we have, from (1.10),

$$
\begin{aligned}
\mathrm{E}(A_1^2) &= \mathrm{E}\left(\int_0^1 \int_0^1 dW(s)\, dW(t)\right) = \int_0^1 dt = 1, \\
\mathrm{E}(A_2^2) &= \mathrm{E}\left(\int_0^1 \int_0^1 (1-s)(1-t) dW(s)\, dW(t)\right) = \int_0^1 (1-t)^2\, dt = \frac{1}{3}.
\end{aligned}
$$

Here the Riemann–Stieltjes double integral appears, although its definition is deferred until the next section. In any case the above computation entails the following relation:

$$
\mathrm{E}\left(dW(s)\, dW(t)\right) = \delta_{st}\, dt = \begin{cases} dt & (s = t), \\ 0 & (s \neq t). \end{cases}
$$

An important and useful property associated with the two integrals in (1.8) and (1.9) is that the existence of one integral implies the other and vice versa, and they can be combined together in the following theorem for *integration by parts* (Soong, 1973, theorem 4.5.3; Kuo, 2006, p.9; Klebaner, 2012, p.12).

Theorem 1.4 *If either A in (1.8) or B in (1.9) exists, then both integrals exist, and*

$$\int_0^1 g(t)\, dW(t) = \big[g(t)W(t)\big]_0^1 - \int_0^1 W(t)\, dg(t). \tag{1.13}$$

For example, we have

$$A_3 = \int_0^1 dW(t) = [W(t)]_0^1 = W(1),$$

$$A_4 = \int_0^1 (1-t)\, dW(t) = [(1-t)W(t)]_0^1 - \int_0^1 (-1)W(t)\, dt = \int_0^1 W(t)\, dt.$$

As another example, let us consider, for $t \in [0,1]$,

$$F_g(t) = \int_0^t \frac{(t-s)^g}{g!}\, dW(s) \quad (g = 1, 2, \dots), \quad F_0(t) = W(t). \tag{1.14}$$

This integral can be transformed into

$$\begin{aligned} F_g(t) &= \int_0^t \left(\int_s^t \frac{(u-s)^{g-1}}{(g-1)!}\, du \right) dW(s) = \int_0^t \left(\int_0^u \frac{(u-s)^{g-1}}{(g-1)!}\, dW(s) \right) du \\ &= \int_0^t F_{g-1}(u)\, du = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{g-1}} W(t_g)\, dt_g\, dt_{g-1} \cdots dt_1. \end{aligned} \tag{1.15}$$

The stochastic process $\{F_g(t)\}$ is called *g-fold integrated Bm* because of this last expression, and it holds that

$$F_g(t) \sim \mathrm{N}\left(0, \frac{t^{2g+1}}{(2g+1)(g!)^2}\right) \quad (g = 0, 1, \dots), \tag{1.16}$$

which means that the variation becomes smaller quickly as g gets large. Note that, unlike the Bm $F_0(t) = W(t)$, $F_g(t)$ for $g \geq 1$ is m.s. g-times continuously differentiable, and it holds that

$$\frac{dF_g(t)}{dt} = F_{g-1}(t), \qquad \int_0^t F_g(s)\, dF_g(s) = \frac{1}{2} F_g^2(t) \qquad (g \geq 1).$$

Figure 1.1 draws sample paths of $F_0(t)$, $F_1(t)$, and $F_2(t)$ with variations adjusted allowing for (1.16) so that the three sample paths can be compared on the same sheet without the vertical axis. These were generated by using the Karhunen–Loève expansion for Bm given in (1.1). It can be seen that the sample path of $F_2(t)$ is quite smooth and that of $F_1(t)$ is moderately smooth, whereas the Bm exhibits a zigzag fluctuation.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

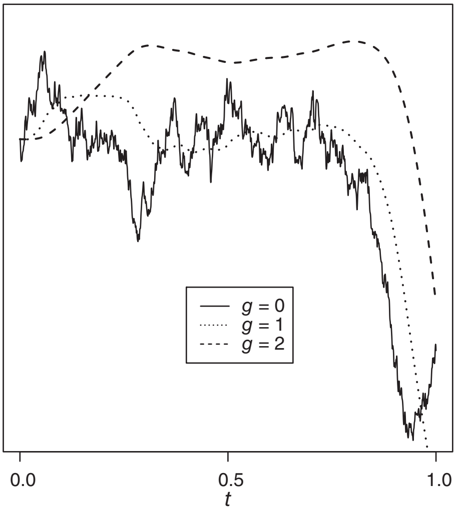

Figure 1.1 Sample paths of $\{F_g(t)\}$ for $g = 0, 1, 2$.

g-fold integrated Bm arises from the discrete-time process $\{y_j^{(g)}\}$ generated by

$$(1 - L)^g \, y_j^{(g)} = \varepsilon_j \qquad (j = 1, \ldots, n), \tag{1.17}$$

where L is the *lag-operator* with the effect of lagging a variable $Ly_j = y_{j-1}, L^2 y_j = y_{j-2}$, and so on, whereas $\{\varepsilon_j\} \sim$ i.i.d.$(0, \sigma^2)$. We assume that $y_{-k}^{(g)} = 0$ for $k = 0, 1, \ldots, g - 1$. The process $\{y_j^{(g)}\}$ is called the I(g) process, where "I" stands for "integrated." Note that the random walk corresponds to the case of $g = 1$. It follows from (1.17) that

$$y_j^{(g)} = y_{j-1}^{(g)} + y_j^{(g-1)} = y_1^{(g-1)} + \cdots + y_j^{(g-1)}, \qquad y_j^{(0)} = \varepsilon_j.$$

Roughly speaking, the process $F_g(t)$ is related to $y_j^{(g)}$ in the following way. For Bm it holds that

$$W(t) = \int_0^t dW(u) \quad \Leftrightarrow \quad y_j^{(1)} = (1 - L)^{-1}\varepsilon_j = \sum_{i=1}^{j} \varepsilon_i,$$

where $(1 - L)^{-1}$ is interpreted as a cumulative sum operator. Similarly it holds that, for $F_1(t)$,

$$\int_0^t W(u)\, du = \int_0^t (t - u)\, dW(u) \;\Leftrightarrow\; y_j^{(2)} = (1 - L)^{-2}\varepsilon_j \approx \sum_{i=1}^{j} (j - i)\varepsilon_i.$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

In general, we have, for $F_g(t)$,

$$\int_0^t \frac{(t-u)^g}{g!}\, dW(u) \ \Leftrightarrow\ y_j^{(g+1)} = (1-L)^{-(g+1)}\varepsilon_j \approx \frac{1}{g!}\sum_{i=1}^{j}(j-i)^g\varepsilon_i.$$

The above relationship between $F_g(t)$ and $y_j^{(g)}$ can be made exact in the following way. Let us put, for $g \geq 1$ and $(j-1)/n \leq t \leq j/n$,

$$\begin{aligned} Y_n^{(g)}(t) &= \frac{1}{n^{g-1/2}\,\sigma}\, y_{j-1}^{(g)} + n\left(t - \frac{j-1}{n}\right)\frac{1}{n^{g-1/2}\sigma}\, y_j^{(g-1)} \\ &= \frac{1}{n}\sum_{i=1}^{j-1} Y_n^{(g-1)}\left(\frac{i}{n}\right) + n\left(t - \frac{j-1}{n}\right)\frac{1}{n^{g-1/2}\sigma}\, y_j^{(g-1)}. \end{aligned}$$

Then it was shown by Chan and Wei (1988) that

$$\left\{Y_n^{(g)}(t)\right\} \Rightarrow \{F_{g-1}(t)\} \quad (g = 1, 2, \ldots), \tag{1.18}$$

which is an extended version of Donsker's FCLT in (1.5).

Exercises for Section 1.2

Exercise 1.2.1 If $\underset{n\to\infty}{\text{l.i.m.}}\, X_n = X$ and $\underset{n\to\infty}{\text{l.i.m.}}\, Y_n = Y$, prove that

$$\underset{n\to\infty}{\text{l.i.m.}}\,(aX_n + bY_n) = aX + bY, \qquad \lim_{m,n\to\infty} \mathrm{E}(X_m Y_n) = \mathrm{E}(XY),$$

where a and b are constants.

Exercise 1.2.2 Show that the m.s. integral $\int_0^1 X(t)\,dt$ is well defined if $X(t)$ is m.s. continuous.

Exercise 1.2.3 Show that, if $X(t)$ is m.s. continuous, it holds that

$$\mathrm{E}\left(\int_0^1 X(t)\,dt\right) = \int_0^1 \mathrm{E}(X(t))\,dt,$$

$$\mathrm{E}\left[\left(\int_0^1 X(t)\,dt\right)^2\right] = \int_0^1\int_0^1 \mathrm{E}(X(s)X(t))\,ds\,dt.$$

Exercise 1.2.4 Prove that, if $\{X_n(t)\}$ is a Gaussian process and $\underset{n\to\infty}{\text{l.i.m.}}\, X_n(t)$ exists, then $\{X(t)\}$ is also a Gaussian process.

Exercise 1.2.5 Compute the means and second moments of

$$I_1 = \int_0^1 W(t)\,dt, \quad I_2 = \int_0^1 |W(t)|\,dt, \quad I_3 = \int_0^1 e^{W(t)}\,dt.$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Exercise 1.2.6 Compute the means and second moments of

$$I_4 = \int_0^1 W^2(t)\,dt, \quad I_5 = \int_0^1 (W(t) - tW(1))^2\,dt,$$
$$I_6 = \int_0^1 \left(W(t) - \int_0^1 W(s)\,ds\right)^2 dt.$$

Exercise 1.2.7 Prove that $\int_0^1 g(t)\,dW(t) \sim \mathrm{N}\left(0, \int_0^1 g^2(t)\,dt\right)$.

1.3 Statistics of Quadratic Functionals of Brownian Motion

Here we extend the m.s. Riemann–Stieltjes integral discussed in the previous section to the corresponding double integral, which will be frequently used in subsequent discussions. Let the partition $p_{m,n}$ of the unit square $[0, 1] \times [0, 1]$ be

$$p_{m,n} \;:\; 0 = s_0 < s_1 < \cdots < s_m = 1;\ 0 = t_0 < t_1 < \cdots < t_n = 1,$$

and put

$$\Delta_{m,n} = \max_{1\le i,\, j\le n} \left(s_i - s_{i-1},\ t_j - t_{j-1}\right).$$

Then we consider, for a symmetric function $K(s,t)$ defined in $[0, 1] \times [0, 1]$,

$$\begin{aligned} X &= \int_0^1 \int_0^1 K(s,t)\,dW(s)\,dW(t) \\ &= \underset{\substack{m,n\to\infty \\ \Delta_{m,n}\to 0}}{\mathrm{l.i.m.}} \sum_{i=1}^m \sum_{j=1}^n K\left(s_i', t_j'\right) \Delta W(s_i)\,\Delta W(t_j), \end{aligned} \tag{1.19}$$

where $s_i' \in [s_{i-1}, s_i)$, $t_j' \in [t_{j-1}, t_j)$, and $\Delta W(s_i) = W(s_i) - W(s_{i-1})$. The integral in (1.19) is called the *m.s. Riemann–Stieltjes double integral*, whereas X is referred to as a *quadratic functional* of Bm. In Section 1.2.1 we dealt with the m.s. Riemann integral of the quadratic functions of Bm. There is a close relationship between the two integrals, which will be explained later.

A sufficient condition for the existence of the integral in (1.19) is that $K(s,t)$ is continuous in $[0, 1] \times [0, 1]$. If $K(s,t)$ is symmetric and continuous, it holds (Exercise 1.3.1) that

$$\begin{aligned} \mathrm{E}(X) &= \int_0^1 K(t,t)\,dt, \\ \mathrm{Var}(X) &= 2\int_0^1 \int_0^1 K^2(s,t)\,ds\,dt = 4\int_0^1 \int_0^t K^2(s,t)\,ds\,dt. \end{aligned} \tag{1.20}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Moreover, the c.f. of X in (1.19) is given by

$$\mathrm{E}\left(e^{i\theta X}\right) = \mathrm{E}\left[\exp\left\{i\theta \int_0^1 \int_0^1 K(s,t)\, dW(s)\, dW(t)\right\}\right] = (D(2i\theta))^{-1/2}, \tag{1.21}$$

where $D(\lambda)$ is called the *Fredholm determinant* (FD) of $K(s,t)$. This result originates from Anderson and Darling (1952), whose proof will be given in Theorem 2.3 under an additional assumption on $K(s,t)$. It holds that the FD of $K(s,t)$ is the same as that of $K(1-s,1-t)$ because of *Mercer's theorem* (Theorem 2.2). Thus we have

$$\int_0^1 \int_0^1 K(s,t)\, dW(s)\, dW(t) \stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 K(1-s,1-t)\, dW(s)\, dW(t), \tag{1.22}$$

where $\stackrel{\mathcal{D}}{=}$ stands for distributional equivalence. The results described in (1.20) will also be derived easily via the properties of the FD (see Theorem 2.4).

Let us deal with some examples of (1.19). For this purpose we put, for $j = 1, 2, \ldots,$

$$X_j = \int_0^1 \int_0^1 K_j(s,t)\, dW(s)\, dW(t) \stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 K_j(1-s,1-t)\, dW(s)\, dW(t). \tag{1.23}$$

Consider first X_1 with

$$K_1(s,t) = \sum_{k=1}^{n} g_k(s) g_k(t),$$

where $g_k(t)$ $(k = 1, \ldots, n)$ are continuous and linearly independent on $[0, 1]$. In this case $K_1(s,t)$ is said to be *degenerate*. To obtain the c.f. of X_1, we have

$$X_1 = \sum_{k=1}^{n} \int_0^1 \int_0^1 g_k(s) g_k(t)\, dW(s)\, dW(t) = \sum_{k=1}^{n} Y_k^2,$$

where

$$Y_k = \int_0^1 g_k(t)\, dW(t) \sim \mathrm{N}\left(0, \int_0^1 g_k^2(t)\, dt\right),$$

$$\mathrm{Cov}(Y_k, Y_\ell) = \int_0^1 g_k(t)\, g_\ell(t)\, dt.$$

Each Y_k^2 is a constant multiple of a $\chi^2(1)$ random variable, where $Y_1, \ldots, Y_n$ are not independent. We can compute the c.f. of X_1 following

$$\mathrm{E}\left(e^{i\theta X_1}\right) = |I_n - 2i\theta\Sigma|^{-1/2} = \prod_{k=1}^{n} (1 - 2i\theta\lambda_k)^{-1/2},$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where I_n is the $n \times n$ identity matrix and Σ is the covariance matrix of $\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$, whereas $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of Σ, which are all positive. It holds that

$$X_1 \stackrel{\mathcal{D}}{=} \sum_{k=1}^{n} \lambda_k Z_k^2, \quad \boldsymbol{Z} = (Z_1, \ldots, Z_n) \sim \mathrm{N}(\boldsymbol{0}, I_n).$$

According to the result in (1.21), the FD of $K_1(s,t)$ is given by

$$D_1(\lambda) = \left| I_n - \lambda \Sigma \right| = \prod_{k=1}^{n} \left(1 - \lambda \lambda_k\right), \tag{1.24}$$

where it is noted that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the solutions to $D_1\left(\lambda^{-1}\right) = 0$.

As a simple example of X_1, let us consider X_2 with

$$K_2(s,t) = 1 + st, \quad g_1(t) = 1, \quad g_2(t) = t. \tag{1.25}$$

Then we obtain (Exercise 1.3.2)

$$\begin{aligned}
\mathrm{E}(X_2) &= \frac{4}{3}, \quad \mathrm{Var}(X_2) = \frac{29}{9}, \\
\mathrm{E}\left(e^{i\theta X_2}\right) &= (D_2(2i\theta))^{-1/2} = \left(1 - \frac{8}{3}\, i\theta - \frac{1}{3}\,\theta^2\right)^{-1/2}, \\
X_2 &\stackrel{\mathcal{D}}{=} \lambda_1 Z_1^2 + \lambda_2 Z_2^2 \quad \left(\lambda_1, \lambda_2 = \frac{2}{3} \pm \frac{\sqrt{13}}{6}\right),
\end{aligned}$$

where $D_2(\lambda) = 1 - 4\lambda/3 + \lambda^2/12$, which is the FD of $K_2(s,t)$, whereas λ_1 and λ_2 are the solutions to $D_2\left(\lambda^{-1}\right) = 0$.

As another example, consider X_3 with

$$K_3(s,t) = 1 + st + s^2 t^2, \quad g_1(t) = 1, \quad g_2(t) = t, \quad g_3(t) = t^2. \tag{1.26}$$

We have

$$\begin{aligned}
\mathrm{E}(X_3) &= \frac{23}{15}, \quad \mathrm{Var}(X_3) = \frac{1199}{300}, \\
\mathrm{E}\left(e^{i\theta X_3}\right) &= (D_3(2i\theta))^{-1/2} = \left(1 - \frac{46}{15}\, i\theta - \frac{127}{180}\theta^2 + \frac{1}{270}\, i\theta^3\right)^{-1/2}, \\
X_3 &\stackrel{\mathcal{D}}{=} \lambda_1 Z_1^2 + \lambda_2 Z_2^2 + \lambda_3 Z_3^2,
\end{aligned}$$

where $D_3(\lambda) = 1 - 23\lambda/15 + 127\lambda^2/720 - \lambda^3/2160$, which is the FD of $K_3(s,t)$, whereas λ_1, λ_2, and λ_3 are the solutions to $D_3\left(\lambda^{-1}\right) = 0$.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

In the above examples, the number of solutions to $D_j(\lambda) = 0$ $(j = 1, 2, 3)$ is finite. For these cases the functions $K_j(s, t)$ $(j = 1, 2, 3)$ are degenerate.

The next example deals with the case where the function $K(s, t)$ is *non-degenerate*, that is, the case where the number of zeros of the FD is infinite. Consider

$$K_4(s, t) = \sum_{n=1}^{\infty} \frac{f_n(s)\, f_n(t)}{\lambda_n}, \tag{1.27}$$

where λ_n is positive for all n and the sum converges absolutely and uniformly in the unit square. It is also assumed that $\{f_n(t)\}$ is an *orthonormal* sequence of functions on [0, 1] that satisfies

$$\int_0^1 f_m(t)\, f_n(t)\, dt = \delta_{mn} = \begin{cases} 1 & (m = n), \\ 0 & (m \neq n). \end{cases}$$

Then we can change the order of the sum and integration to get

$$X_4 = \int_0^1 \int_0^1 \sum_{n=1}^{\infty} \frac{f_n(s)\, f_n(t)}{\lambda_n}\, dW(s)\, dW(t) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} Z_n^2,$$

where

$$Z_n = \int_0^1 f_n(t)\, dW(t) \sim \mathrm{NID}(0, 1), \quad \mathrm{E}(X_4) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n}, \quad \mathrm{Var}(X_4) = \sum_{n=1}^{\infty} \frac{2}{\lambda_n^2}.$$

The c.f. of X_4 is obtained as

$$\mathrm{E}\left(e^{i\theta X_4}\right) = \mathrm{E}\left[\exp\left\{i\theta \sum_{n=1}^{\infty} \frac{1}{\lambda_n} Z_n^2\right\}\right] = \prod_{n=1}^{\infty} \left(1 - \frac{2i\theta}{\lambda_n}\right)^{-1/2}.$$

According to the result in (1.21), the FD of $K_4(s, t)$ is given by

$$D_4(\lambda) = \prod_{n=1}^{\infty} \left(1 - \frac{\lambda}{\lambda_n}\right),$$

where the solutions to $D_4(\lambda) = 0$ are λ_n $(n = 1, 2, \ldots)$.

As a specific example of $K_4(s, t)$, let us take up

$$K_5(s, t) = \sum_{n=1}^{\infty} \frac{2\cos\left(n - \frac{1}{2}\right)\pi s\, \cos\left(n - \frac{1}{2}\right)\pi t}{\left(n - \frac{1}{2}\right)^2 \pi^2} = 1 - \max(s, t), \tag{1.28}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where $\left\{\sqrt{2}\cos\left(n-\frac{1}{2}\right)\pi t\right\}$ is orthonormal on $[0,1]$, and this last equality is ensured by Mercer's theorem (Theorem 2.2). Then we have, for the corresponding X_5,

$$\mathrm{E}(X_5) = \sum_{n=1}^{\infty} \frac{1}{\left(n-\frac{1}{2}\right)^2 \pi^2} = \int_0^1 (1-t)\,dt = \frac{1}{2},$$

$$\mathrm{Var}(X_5) = \sum_{n=1}^{\infty} \frac{2}{\left(n-\frac{1}{2}\right)^4 \pi^4} = 4\int_0^1 \int_0^t (1-\max(s,t))^2\,ds\,dt = \frac{1}{3}.$$

The c.f. of X_5 is given by

$$\mathrm{E}\left(e^{i\theta X_5}\right) = \prod_{n=1}^{\infty}\left(1-\frac{2i\theta}{\left(n-\frac{1}{2}\right)^2\pi^2}\right)^{-1/2} = \left(\cos\sqrt{2i\theta}\right)^{-1/2}, \tag{1.29}$$

where this last equality comes from the infinite product expansion formula for the cos function:

$$\cos z = \prod_{n=1}^{\infty}\left(1-\frac{z^2}{\left(n-\frac{1}{2}\right)^2\pi^2}\right). \tag{1.30}$$

Thus the FD of $K_5(s,t) = 1-\max(s,t)$ is given by

$$D_5(\lambda) = \cos\sqrt{\lambda} = \prod_{n=1}^{\infty}\left(1-\frac{\lambda}{\left(n-\frac{1}{2}\right)^2\pi^2}\right),$$

where the solutions to $D_5(\lambda)=0$ are $(n-1/2)^2\pi^2$ $(n=1,2,\ldots)$.

We note here that

$$\begin{aligned}
&\int_0^1\int_0^1 K_5(s,t)\,dW(s)\,dW(t)\\
&= \int_0^1\int_0^1 [1-\max(s,t)]\,dW(s)\,dW(t)\\
&= \int_0^1 \left(W(1) - t\int_0^t dW(s) - \int_t^1 s\,dW(s)\right) dW(t)\\
&= \int_0^1 \left(\int_t^1 W(s)\,ds\right) dW(t) = \int_0^1 W^2(t)\,dt.
\end{aligned}$$

This relation may be established in a reversed direction in a general form as

$$\int_0^1 g(t)\, W^2(t)\, dt = \int_0^1 g(t) \left(\int_0^t \int_0^t dW(u)\, dW(v) \right) dt$$
$$= \int_0^1 \int_0^1 \left(\int_{\max(u,v)}^1 g(t)\, dt \right) dW(u)\, dW(v),$$

where it is justified to interchange the order of integration (Exercise 1.3.3). For example, we have

$$\int_0^1 t^{2k}\, W^2(t)\, dt = \int_0^1 \int_0^1 \frac{1}{2k+1} \left[1 - (\max(s,t))^{2k+1} \right] dW(s)\, dW(t),$$

where $k > -1/2$.

It is noted that, because of the distributional equivalence described in (1.22), the distribution of X_6 with

$$K_6(s,t) = K_5(1-s, 1-t) = \sum_{i=1}^{\infty} \frac{2 \sin\left(n - \frac{1}{2}\right) \pi s \, \sin\left(n - \frac{1}{2}\right) \pi t}{\left(n - \frac{1}{2}\right)^2 \pi^2}$$
$$= 1 - \max(1-s, 1-t) = \min(s,t) \tag{1.31}$$

is the same as that of X_5. Moreover, for X_6, it holds (Exercise 1.3.4) that

$$X_6 = \int_0^1 \int_0^1 \min(s,t)\, dW(s)\, dW(t) = \int_0^1 (W(1) - W(t))^2\, dt.$$

Thus it follows that

$$\begin{aligned} \int_0^1 W^2(t)\, dt &= \int_0^1 \int_0^1 [1 - \max(s,t)]\, dW(s)\, dW(t) \\ &\stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 \min(s,t)\, dW(s)\, dW(t) \\ &= \int_0^1 (W(1) - W(t))^2\, dt, \end{aligned} \tag{1.32}$$

which shows a close relationship between the m.s. Riemann integral of squared functions of Bm and the m.s. Riemann–Stieltjes double integral with respect to Bm. More generally, it will be shown in Theorem 2.11 that any zero-mean Gaussian process $\{X(t)\}$ defined on $[0, 1]$ satisfies

$$\int_0^1 X^2(t)\, dt \stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 \mathrm{Cov}\,(X(s), X(t))\; dW(s)\, dW(t). \tag{1.33}$$

Let us consider another example of the nondegenerate case for which

$$K_7(s,t) = \sum_{n=1}^{\infty} \frac{2\sin n\pi s\ \sin n\pi t}{n^2\pi^2} = \min(s,t) - st, \tag{1.34}$$

where this last equality also comes from Mercer's theorem. Note here that $K_7(s,t) = K_7(1-s, 1-t)$, unlike $K_5(s,t) = 1 - \max(s,t)$. Then it is shown (Exercise 1.3.5) that the corresponding X_7 has the following properties:

$$\mathrm{E}(X_7) = \sum_{n=1}^{\infty} \frac{1}{n^2\pi^2} = \frac{1}{6}, \quad \mathrm{Var}(X_7) = \sum_{n=1}^{\infty} \frac{2}{n^4\pi^4} = \frac{1}{45},$$

$$\mathrm{E}\left(e^{i\theta X_7}\right) = \prod_{n=1}^{\infty}\left(1 - \frac{2i\theta}{n^2\pi^2}\right)^{-1/2} = \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-1/2}, \tag{1.35}$$

where this last equality comes from the infinite product expansion formula for the sin function:

$$\frac{\sin z}{z} = \prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2\pi^2}\right). \tag{1.36}$$

Thus the FD of $K_7(s,t) = \min(s,t) - st$ is given by

$$D_7(\lambda) = \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}} = \prod_{n=1}^{\infty}\left(1 - \frac{\lambda}{n^2\pi^2}\right),$$

where the solutions to $D_7(\lambda) = 0$ are given by $n^2\pi^2$ $(n = 1, 2, \ldots)$. It is also shown (Exercise 1.3.6) that

$$\begin{aligned}\int_0^1 \left(W(t) - \int_0^1 W(s)\,ds\right)^2 dt &= \int_0^1 W^2(t)\,dt - \left(\int_0^1 W(s)\,ds\right)^2 \\ &= \int_0^1\int_0^1 [\min(s,t) - st]\,dW(s)\,dW(t),\end{aligned}$$

where the integrand on the left-hand side is the square of demeaned Bm.

The next example gives the same FD as $K_7(s,t)$, which is given by

$$K_8(s,t) = \sum_{n=1}^{\infty} \frac{2\cos n\pi s\ \cos n\pi t}{n^2\pi^2} = \frac{1}{3} - \max(s,t) + \frac{s^2+t^2}{2}, \tag{1.37}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where this last equality comes from Mercer's theorem. Using the relation described in (1.33), we can relate the distribution of X_7 with that of X_8 as follows (Exercise 1.3.7):

$$\int_0^1 \int_0^1 K_7(s,t)\, dW(s)\, dW(t)$$
$$= \int_0^1 \int_0^1 [\min(s,t) - st]\, dW(s)\, dW(t) = \int_0^1 \left(W(t) - \int_0^1 W(s)\, ds \right)^2 dt$$
$$\stackrel{\mathcal{D}}{=} \int_0^1 (W(t) - tW(1))^2\, dt = \int_0^1 \int_0^1 K_8(s,t)\, dW(s)\, dW(t), \qquad (1.38)$$

where $\{W(t) - tW(1)\}$ is the Bb.

The next example is concerned with the quadratic functional of the g-fold integrated Bm $\{F_g(t)\}$ discussed in the last section. Consider here the case of $g = 1$. Then we have

$$X_9 = \int_0^1 F_1^2(t)\, dt = \int_0^1 \left(\int_0^t (t-s)\, dW(s) \right)^2 dt$$
$$= \int_0^1 \int_0^1 K_9(s,t)\, dW(s)\, dW(t) \qquad (1.39)$$
$$\stackrel{\mathcal{D}}{=} X_{10} = \int_0^1 \int_0^1 K_{10}(s,t)\, dW(s)\, dW(t), \qquad (1.40)$$

where, for $s \le t$,

$$K_9(s,t) = \frac{1}{6}(1-t)^2(t+2-3s), \qquad (1.41)$$
$$K_{10}(s,t) = \mathrm{Cov}\,(F_1(s), F_1(t)) = -\frac{1}{6}s^3 + \frac{1}{2}s^2 t. \qquad (1.42)$$

The c.f. of this statistic cannot be obtained by the approach discussed so far. However, it can be obtained by the method we shall use later, which yields

$$\mathrm{E}\left(e^{i\theta X_9}\right) = \mathrm{E}\left(e^{i\theta X_{10}}\right) = \left[\frac{1}{2}\left(1 + \cos(2i\theta)^{1/4} \cosh(2i\theta)^{1/4} \right) \right]^{-1/2}.$$

Thus the FD of $K_9(s,t)$ is equal to that of $K_{10}(s,t)$ and is given by

$$D_9(\lambda) = D_{10}(\lambda) = \frac{1}{2}\left(1 + \cos\lambda^{1/4} \cosh\lambda^{1/4} \right).$$

The computation of the c.f. becomes more complicated as g becomes large, which we discuss in detail in Chapter 2.

The class of quadratic functionals of Bm is quite large. We shall examine many quadratic functionals in later chapters.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Exercises for Section 1.3

Exercise 1.3.1 Suppose that

$$X = \int_0^1 \int_0^1 K(s,t)\, dW(s)\, dW(t),$$

where $K(s,t)$ is symmetric and continuous in $[0,1] \times [0,1]$. Then prove that

$$\begin{aligned}
\mathrm{E}(X) &= \int_0^1 K(t,t)\, dt, \\
\mathrm{Var}(X) &= 2\int_0^1 \int_0^1 K^2(s,t)\, ds\, dt = 4\int_0^1 \int_0^t K^2(s,t)\, ds\, dt.
\end{aligned}$$

Exercise 1.3.2 For X_2 and X_3 defined in (1.23) with $K_2(s,t) = 1 + st$ and $K_3(s,t) = 1 + st + s^2t^2$, show that

$$\begin{aligned}
&\mathrm{E}(X_2) = \frac{4}{3}, \quad \mathrm{Var}(X_2) = \frac{29}{9}, \quad \mathrm{E}\left(e^{i\theta X_2}\right) = \left(1 - \frac{8}{3}\, i\theta - \frac{1}{3}\, \theta^2\right)^{-1/2}, \\
&\mathrm{E}(X_3) = \frac{23}{15}, \quad \mathrm{Var}(X_3) = \frac{1199}{300}, \\
&\mathrm{E}\left(e^{i\theta X_3}\right) = \left(1 - \frac{46}{15}\, i\theta - \frac{127}{180}\, \theta^2 + \frac{1}{270}\, i\theta^3\right)^{-1/2}.
\end{aligned}$$

Exercise 1.3.3 Justify the following computation:

$$\int_0^1 g(t) W^2(t)\, dt = \int_0^1 \int_0^1 \left(\int_{\max(s,t)}^1 g(u)\, du \right) dW(s)\, dW(t),$$

where $g(t)$ is continuous on $[0,1]$.

Exercise 1.3.4 Prove the following equality by showing that (i) the left-hand side leads to the right-hand side and (ii) the right-hand side leads to the left-hand side:

$$\int_0^1 \int_0^1 \min(s,t)\, dW(s)\, dW(t) = \int_0^1 (W(1) - W(t))^2\, dt.$$

Exercise 1.3.5 For X_7 defined in (1.23) with

$$K_7(s,t) = \sum_{n=1}^{\infty} \frac{2 \sin n\pi s\, \sin n\pi t}{n^2 \pi^2},$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

show that

$$\mathrm{E}(X_7) = \sum_{n=1}^{\infty} \frac{1}{n^2\pi^2} = \frac{1}{6}, \quad \mathrm{Var}(X_7) = \sum_{n=1}^{\infty} \frac{2}{n^4\pi^4} = \frac{1}{45},$$

$$\mathrm{E}\left(e^{i\theta X_7}\right) = \prod_{n=1}^{\infty} \left(1 - \frac{2i\theta}{n^2\pi^2}\right)^{-1/2} = \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-1/2}.$$

Exercise 1.3.6 Show that

$$\int_0^1 \left(W(t) - \int_0^1 W(s)\,ds\right)^2 dt = \int_0^1 \int_0^1 [\min(s,t) - st]\,dW(s)\,dW(t).$$

Exercise 1.3.7 Show that

$$\begin{aligned}
&\int_0^1 (W(t) - tW(1))^2\,dt \\
&\quad = \int_0^1 \int_0^1 \left[\frac{1}{3} - \max(s,t) + \frac{s^2+t^2}{2}\right] dW(s)\,dW(t) \\
&\quad \overset{\mathcal{D}}{=} \int_0^1 \int_0^1 [\min(s,t) - st]\,dW(s)\,dW(t).
\end{aligned}$$

1.4 Mean Square Itô Integral

Here we discuss the m.s. *Itô integral* in connection with quadratic functionals of Bm. Let $\{X(t)\}$ $(\in L_2)$ be a stochastic process that is m.s. continuous and independent of $W(v) - W(u)$ for all $0 \le t \le u \le v \le 1$. Let the partition of $[0,1]$ be $p_n : 0 = t_0 < t_1 < \cdots < t_n = 1$, and put $\Delta_n = \max_{1 \le i \le n}(t_i - t_{i-1})$. Then the m.s. Itô integral of $X(t)$ with respect to $W(t)$ is defined by

$$Y = \int_0^1 X(t)\,dW(t) = \underset{\substack{n\to\infty \\ \Delta_n \to 0}}{\mathrm{l.i.m.}} \sum_{i=1}^{n} X(t_{i-1})(W(t_i) - W(t_{i-1})). \tag{1.43}$$

The Itô integral defined in this way exists and is unique. For the proof, see Jazwinski (1970, theorem 4.2), Soong (1973, theorem 5.2.1), Kuo (2006, theorem 4.7.1), and Klebaner (2012, theorem 4.3). It evidently holds that

$$\mathrm{E}(Y) = 0, \quad \mathrm{Var}(Y) = \int_0^1 \mathrm{E}(X^2(t))\,dt.$$

We notice in (1.43) that the value of $X(t)$ taken in the interval $[t_{i-1}, t_i)$ is $X(t_{i-1})$, unlike in the m.s. Riemann or Riemann–Stieltjes integral. If we take t_i' $(\ne t_{i-1})$ in $[t_{i-1}, t_i)$, then the m.s. limit is different, which will be shown later.

According to the definition of the Itô integral, the following relations can be established (Exercise 1.4.1):

$$\int_0^1 X(t)\,(dW(t))^j = \begin{cases} \displaystyle\int_0^1 X(t)\,dt & (j=2), \\ 0 & (j \geq 3). \end{cases} \tag{1.44}$$

In particular, putting $X(t) \equiv 1$, we formally have

$$(dW(t))^j = \begin{cases} dt & (j=2), \\ 0 & (j \geq 3). \end{cases} \tag{1.45}$$

To examine some other properties, consider the stochastic process defined by

$$Y(t) = \int_0^t X(s)\,dW(s), \tag{1.46}$$

which is well defined for each $t \in [0,1]$. It is clear that $\{Y(t)\}$ is m.s. continuous because, for $0 \leq t < t+h \leq 1$,

$$\begin{aligned} \mathrm{E}\left[(Y(t+h)-Y(t))^2\right] &= \mathrm{E}\left[\left(\int_t^{t+h} X(s)\,dW(s)\right)^2\right] \\ &= \int_t^{t+h} \mathrm{E}(X^2(s))\,ds \\ &\leq h \max_{t \leq s \leq t+h} \mathrm{E}(X^2(s)). \end{aligned}$$

The m.s. differentiability of $Y(t)$ is not ensured, but we express (1.46) as

$$Y(t) = \int_0^t X(s)\,dW(s) \qquad \Leftrightarrow \qquad dY(t) = X(t)\,dW(t). \tag{1.47}$$

The differential expression on the right is called the *stochastic differential*, which is always taken to imply the left-hand integral. It also holds that $\{Y(t)\}$ is a zero-mean martingale, that is, $\mathrm{E}(Y(t)) = 0$ and, for $s \leq t$,

$$\mathrm{E}[Y(t)|Y(s)] = Y(s) + \mathrm{E}[Y(t)-Y(s)|Y(s)] = Y(s).$$

As an example, let us consider

$$Y_1(t) = \int_0^t W(s)\,dW(s). \tag{1.48}$$

We cannot apply Theorem 1.4 for integration by parts to $Y_1(t)$. If it is applied, we have $Y_1(t) = W^2(t)/2$, which is evidently incorrect because $\{Y_1(t)\}$ is a zero-mean martingale, whereas $\mathrm{E}(W^2(t)/2) = t/2 \neq 0$. Let the partition

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

of $[0,t]$ be $p_m(t)\ :\ 0 = t_0 < t_1 < \cdots < t_m = t$, and put $\Delta_m(t) = \max_{1\le i\le m}(t_i - t_{i-1})$. Then, as $m\to\infty$ and $\Delta_m(t)\to 0$,

$$\begin{aligned}
Y_{m1}(t) &= \sum_{i=1}^{m} W(t_{i-1})(W(t_i) - W(t_{i-1})) \\
&= -\frac{1}{2}\left[\sum_{i=1}^{m}(W(t_i) - W(t_{i-1}))^2 - \sum_{i=1}^{m} W^2(t_i) + \sum_{i=1}^{m} W^2(t_{i-1})\right] \\
&= \frac{1}{2}\,W^2(t) - \frac{1}{2}\sum_{i=1}^{m}(W(t_i) - W(t_{i-1}))^2 \\
&\Rightarrow \frac{1}{2}\,W^2(t) - \frac{1}{2}\,t,
\end{aligned}$$

where the last result follows from the quadratic variation property of Bm described in Section 1.1. Thus it holds that

$$Y_1(t) = \int_0^t W(s)\,dW(s) = \underset{\substack{m\to\infty \\ \Delta_m(t)\to 0}}{\text{l.i.m.}}\ Y_{m1}(t) = \frac{1}{2}\,(W^2(t) - t), \tag{1.49}$$

and formally we have

$$d\left(\frac{1}{2}(W^2(t) - t)\right) = W(t)\,dW(t) \quad \Leftrightarrow \quad d(W^2(t)) = 2W(t)\,dW(t) + dt. \tag{1.50}$$

This is a simplified version of the Itô calculus discussed below. In connection with the random walk $\{y_j\}$ defined in (1.3), it holds (Exercise 1.4.2) that, as $n\to\infty$,

$$\frac{1}{n\sigma^2}\sum_{j=1}^{n} y_{j-1}\varepsilon_j \Rightarrow \frac{1}{2}(W^2(1) - 1) = \int_0^1 W(t)\,dW(t), \tag{1.51}$$

where $y_j = y_{j-1} + \varepsilon_j$, $y_0 = 0$ and $\{\varepsilon_j\}$ ~i.i.d.$(0,\sigma^2)$. Because of the *continuous mapping theorem* (CMT) and joint weak convergence (Billingsley, 1999, chapter 2; Phillips, 1987a), it also holds that

$$S_n = \frac{\frac{1}{n}\sum_{j=2}^{n} y_{j-1}\varepsilon_j}{\frac{1}{n^2}\sum_{j=2}^{n} y_{j-1}^2} \Rightarrow \frac{\int_0^1 W(t)\,dW(t)}{\int_0^1 W^2(t)\,dt}. \tag{1.52}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

The distribution of this last random variable is called the *unit root distribution* because

$$S_n = n\left(\frac{\sum_{j=2}^n y_{j-1}y_j}{\sum_{j=2}^n y_{j-1}^2} - 1\right) = n(\hat{\rho} - 1),$$

where $\hat{\rho}$ is the least squares estimator (LSE) of ρ for the model $y_j = \rho y_{j-1} + \varepsilon_j$, with the true ρ being unity.

It was mentioned above that the m.s. limit of the sum that defines the Itô integral crucially depends on the choice of the value of the integrand on each interval. For example, we have (Exercise 1.4.3)

$$\underset{\substack{m\to\infty \\ \Delta_m(t)\to 0}}{\text{l.i.m.}} \sum_{i=1}^m W(\tau_{i-1})(W(t_i) - W(t_{i-1})) = \frac{1}{2}(W^2(t) - t) + \lambda t,$$

where $\tau_{i-1} = (1-\lambda)t_{i-1} + \lambda t_i$ for $0 \le \lambda \le 1$. The case of $\lambda = 0$ corresponds to the Itô integral, whereas the integral with $\lambda = 1/2$ is called the *Stratonovich integral*. When $\lambda = 0$, the above quantity becomes a zero-mean martingale.

We now discuss the *Itô calculus*. Let $\{X(t)\}$ be a stochastic process on [0, 1] that belongs to L_2, and consider the integral equation of the form

$$X(t) = X(0) + \int_0^t \mu(X(s), s)\, ds + \int_0^t \sigma(X(s), s)\, dW(s). \tag{1.53}$$

Notice that there are two types of m.s. integrals. One is the Riemann integral and the other the Itô integral. Whether these integrals can be defined in the m.s. sense, and, more importantly, whether this integral equation has a unique m.s. solution $X(t)$, can be answered in the affirmative by the following theorem (Jazwinski, 1970, theorem 4.5; Kuo, 2006, theorem 10.3.5).

Theorem 1.5 *Suppose that*

(i) $X(0)$ *is any random variable that belongs to* L_2 *and is independent of* $W(t) - W(s)$ *for all* $0 \le s \le t \le 1$.

(ii) There is a positive constant K *such that*

$$|\mu(x,t) - \mu(y,t)| \le K|x-y|, \quad |\sigma(x,t) - \sigma(y,t)| \le K|x-y|,$$

$$|\mu(x,s) - \mu(x,t)| \le K|s-t|, \quad |\sigma(x,s) - \sigma(x,t)| \le K|s-t|,$$

$$|\mu(x,t)| \le K(1+x^2)^{1/2}, \quad |\sigma(x,t)| \le K(1+x^2)^{1/2}.$$

Then (1.53) has a unique m.s. continuous solution $X(t)$ *on* $[0,1]$ *such that* $X(t) - X(0)$ *is independent of* $W(v) - W(u)$ *for all* $0 \le t \le u \le v \le 1$.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

This theorem ensures that the two integrals appearing in (1.53) are well defined in the m.s. sense. Note also that the solution process $\{X(t)\}$ is not m.s. differentiable in general. Nonetheless we write (1.53) as

$$dX(t) = \mu(X(t),t)\,dt + \sigma(X(t),t)\,dW(t). \tag{1.54}$$

Here $dX(t)$ is called the *stochastic differential* of $X(t)$, and we call this equation the Itô *stochastic differential equation* (SDE), which is always understood in terms of the integral equation (1.53).

The idea of stochastic differentials can further be developed for functions of $X(t)$ and t. Namely, we can consider another SDE that $f(X(t),t)$ satisfies on the basis of the following theorem (Jazwinski, 1970, lemma 4.2; Kuo, 2006, theorem 7.4.3; Klebaner, 2012, theorem 4.16).

Theorem 1.6 (Itô's formula) *Suppose that $X(t)$ has the stochastic differential (1.54), where the conditions (i) and (ii) in Theorem 1.5 are satisfied. Let $f(x,t)$ denote a continuous function in $(-\infty,\infty)\times[0,1]$ with continuous partial derivatives $f_x = \partial f(x,t)/\partial x$, $f_{xx} = \partial^2 f(x,t)/\partial x^2$, and $f_t = \partial f(x,t)/\partial t$. Assume further that $f_t(X(t),t)$ and $f_{xx}(X(t),t)\,\sigma^2(X(t),t)$ are m.s. Riemann integrable. Then $f(X(t),t)$ has the stochastic differential*

$$df(X(t),t) = f_x(X(t),t)\,dX(t) + \left(f_t(X(t),t) + \frac{1}{2}\,f_{xx}(X(t),t)\sigma^2(X(t),t)\right)dt. \tag{1.55}$$

Some important implications of Itô's formula follow.

(a) If we expand $df(X(t),t)$ formally as

$$df \sim f_x\,dX(t) + f_t\,dt + \frac{1}{2}\,(f_{xx}\,d^2X(t) + 2f_{xt}\,dX(t)\,dt + f_{tt}\,d^2t),$$

then (1.55) implies that we may put $d^2X(t) = \sigma^2(X(t),t)\,dt$, $dX(t)\,dt = 0$ and $d^2t = 0$.

(b) The integral equation corresponding to (1.55) may be written as

$$\begin{aligned}\int_a^b f_x(X(t),t)\,dX(t) &= f(X(b),b) - f(X(a),a)\\ &\quad - \int_a^b \left(f_t(X(t),t) + \frac{1}{2}\,f_{xx}(X(t),t)\,\sigma^2(X(t),t)\right)dt,\end{aligned}$$

where $0 \le a \le t \le b \le 1$. The integral on the left-hand side has never appeared before, and may be called the extended Itô integral. The above relation tells us that the extended Itô integral can be computed via the

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Riemann integral on the right-hand side. In particular, for $f(x,t) = x^2$, we have

$$\int_0^t X(s)\,dX(s) = \frac{1}{2}\left(X^2(t) - X^2(0)\right) - \frac{1}{2}\int_0^t \sigma^2(X(s),s)\,ds. \quad (1.56)$$

(c) For any integer $n \geq 2$, the following stochastic differential can be derived from (1.55) as

$$d(W^n(t)) = nW^{n-1}(t)\,dW(t) + \frac{n(n-1)}{2}W^{n-2}(t)\,dt. \quad (1.57)$$

In particular, when $n = 2$, we have

$$d(W^2(t)) = 2W(t)\,dW(t) + dt,$$

which was earlier presented in (1.50) as a simplified version of the Itô calculus. The relation (1.57) is equivalent to

$$\int_0^t W^n(s)\,dW(s) = \frac{1}{n+1}W^{n+1}(t) - \frac{n}{2}\int_0^t W^{n-1}(s)\,ds, \quad (1.58)$$

which reduces to (1.49) when $n = 1$. We see that the formulas (1.56) and (1.58) enable us to convert (extended) Itô integrals into the usual Riemann integrals.

(d) Itô's formula is useful for solving SDEs in conjunction with the uniqueness property of the solution described in Theorem 1.5, which will be exemplified shortly.

Two examples follow in connection with quadratic functionals associated with the Itô integral.

Example 1.7 Let us consider the SDE

$$d\left(g(t)W^2(t)\right) = 2g(t)W(t)\,dW(t) + (g'(t)W^2(t) + g(t))\,dt,$$

which yields (Exercise 1.4.4)

$$\int_0^1 g(t)W(t)\,dW(t) = \frac{1}{2}\int_0^1\int_0^1 g(\max(s,t))\,dW(s)\,dW(t) - \frac{1}{2}\int_0^1 g(t)\,dt.$$

In particular, we have

$$\int_0^1 t^k\,W(t)\,dW(t) = \int_0^1\int_0^1 \frac{1}{2}(\max(s,t))^k\,dW(s)\,dW(t) - \frac{1}{2(k+1)}, \quad (1.59)$$

where $k \geq 0$.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Example 1.8 Consider the SDE

$$dX(t) = \gamma\, X(t)\, dt + dW(t) \quad \Leftrightarrow \quad X(t) = e^{\gamma t} X(0) + e^{\gamma t} \int_0^t e^{-\gamma s}\, dW(s), \tag{1.60}$$

which is called the *Ornstein–Uhlenbeck (O–U) process* defined on [0, 1], where $X(0)$ is independent of $\{W(t)\}$. Then it follows from (1.56) that

$$\int_0^t X(s)\, dX(s) = \frac{1}{2}\left(X^2(t) - X^2(0) - t\right), \tag{1.61}$$

which is of the nature of the Itô integral. The O–U process has the following moments:

$$\begin{aligned} \mathrm{E}(X(t)) &= e^{\gamma t} \mathrm{E}(X(0)), \\ \mathrm{Cov}(X(s),\, X(t)) &= e^{\gamma(s+t)} \left[\mathrm{Var}(X(0)) + \frac{1 - e^{-2\gamma \min(s,t)}}{2\gamma} \right]. \end{aligned}$$

It turns out that, if γ is positive, $\{X(t)\}$ is explosive and $\mathrm{Var}(X(t))$ increases with t. On the other hand, if γ is negative and $X(0) \sim \mathrm{N}(0, -1/(2\gamma))$, then $\{X(t)\}$ is stationary with $\mathrm{E}(X(t)) = 0$ and $\mathrm{Cov}(X(s),\, X(t)) = -e^{\gamma|s-t|}/(2\gamma)$ so that $\mathrm{Var}(X(t)) = -1/(2\gamma)$ (Exercise 1.4.5).

The O–U process $\{X(t)\}$ in (1.60) arises from the *near-random walk* defined by

$$y_j = \rho\, y_{j-1} + \varepsilon_j \qquad (j = 1, \ldots, n), \quad y_0 = 0, \tag{1.62}$$

where $\rho = 1 + \gamma/n$ and $\{\varepsilon_j\} \sim \mathrm{i.i.d.}(0, \sigma^2)$. In fact, if we construct the partial sum process

$$X_n(t) = \frac{1}{\sqrt{n}\sigma}\, y_{[nt]} + (nt - [nt]) \frac{y_{[nt]+1} - y_{[nt]}}{\sqrt{n}\sigma}, \tag{1.63}$$

it holds (Bobkoski, 1983; Phillips, 1987b) that

$$\{X_n(t)\} \;\Rightarrow\; \{X(t)\}, \quad X(t) = e^{\gamma t} \int_0^t e^{-\gamma s}\, dW(s). \tag{1.64}$$

In particular, the CMT gives

$$\frac{1}{n\sigma^2}\, y_n^2 = X_n^2(1) \;\Rightarrow\; X^2(1) = \left(\int_0^1 e^{\gamma(1-s)}\, dW(s) \right)^2. \tag{1.65}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Let us consider the asymptotic distribution of the LSE $\hat{\rho}$ of ρ in (1.62), where

$$n(\hat{\rho}-1) = n\left(\frac{\sum_{j=2}^{n} y_{j-1}y_j}{\sum_{j=2}^{n} y_{j-1}^2} - 1\right) = \frac{\frac{1}{n}\sum_{j=2}^{n} y_{j-1}(y_j - y_{j-1})}{\frac{1}{n^2}\sum_{j=2}^{n} y_{j-1}^2}.$$

The FCLT and CMT lead us to

$$\begin{aligned}
\frac{1}{n\sigma^2}\sum_{j=2}^{n} y_{j-1}(y_j - y_{j-1}) &= -\frac{1}{2n\sigma^2}\left(\sum_{j=2}^{n}(y_j - y_{j-1})^2 - \sum_{j=2}^{n} y_j^2 + \sum_{j=2}^{n} y_{j-1}^2\right) \\
&= -\frac{1}{2n\sigma^2}\left(\sum_{j=2}^{n}\left(\frac{\gamma}{n}\, y_{j-1} + \varepsilon_j\right)^2 - y_n^2 + y_1^2\right) \\
&\Rightarrow \frac{1}{2}\,(X^2(1) - 1) = \int_0^1 X(t)\, dX(t),
\end{aligned}$$

$$\frac{1}{n^2\sigma^2}\sum_{j=2}^{n} y_j^2 = \frac{1}{n}\sum_{j=2}^{n} X_n^2\left(\frac{j}{n}\right) \Rightarrow \int_0^1 X^2(t)\, dt.$$

Thus, because of joint weak convergence (Phillips, 1987b), we have

$$n(\hat{\rho}-1) \;\Rightarrow\; R = \frac{\int_0^1 X(t)\, dX(t)}{\int_0^1 X^2(t)\, dt}. \tag{1.66}$$

The distribution of R in (1.66) is called the *near-unit root distribution*, which will be examined in detail in Chapter 7.

To describe an extended version of Itô's formula, we introduce

Definition 1.9 (Vector Brownian motion) A multiple stochastic process $\boldsymbol{W}(t) = (W_1(t), \ldots, W_q(t))'$ is q-dimensional Bm that satisfies

(1) $P(\boldsymbol{W}(0) = \boldsymbol{0}) = 1$.
(2) $\boldsymbol{W}(t_1) - \boldsymbol{W}(t_0),\ \boldsymbol{W}(t_2) - \boldsymbol{W}(t_1),\ \ldots,\ \boldsymbol{W}(t_n) - \boldsymbol{W}(t_{n-1})$ are independent for any positive integer n and time points $t_0 < t_1 < \cdots < t_n$.
(3) $\boldsymbol{W}(t) - \boldsymbol{W}(s) \sim \mathrm{N}(\boldsymbol{0},\ (t-s)I_q)$ for any $0 \le s < t \le 1$.

We now extend Itô's formula to the vector case (Jazwinski, 1970, lemma 4.2; Kuo, 2006, p.107; Klebaner, 2012, p.119).

Theorem 1.10 (An extended version of Itô's formula) *Let the p-dimensional stochastic process $\{\boldsymbol{Y}(t)\}$ be the unique solution of the Itô SDE :*

$$d\boldsymbol{Y}(t) = \boldsymbol{\mu}(\boldsymbol{Y}(t),t)dt + G(\boldsymbol{Y}(t),t)\,d\boldsymbol{W}(t), \tag{1.67}$$

where $\boldsymbol{\mu}$ is $p\times 1$ and G is $p\times q$. Let $g(\boldsymbol{y},t)$ be a real-valued function of $\boldsymbol{y}:\ p\times 1$ and $t\in[0,1]$ with continuous partial derivatives $g_y = \partial g/\partial \boldsymbol{y}:\ p\times 1$, $g_{yy} = \partial^2 g/(\partial\boldsymbol{y}\partial\boldsymbol{y}'):\ p\times p$, and $g_t = \partial g/\partial t$. Assume further that g_t and $GG'g_{yy}$ are m.s. Riemann integrable. Then $g(\boldsymbol{Y}(t),t)$ has the stochastic differential

$$dg = g_y' d\boldsymbol{Y}(t) + \left(g_t + \frac{1}{2}\,\mathrm{tr}(GG'g_{yy})\right)dt. \tag{1.68}$$

Some applications of Theorem 1.10 follow. Putting $g(\boldsymbol{Y}(t),t) = \boldsymbol{W}'(t)H\boldsymbol{W}(t)$ with H being a $q\times q$ symmetric matrix, it follows from (1.68) that

$$d(\boldsymbol{W}'(t)H\boldsymbol{W}(t)) = 2\boldsymbol{W}'(t)Hd\boldsymbol{W}(t) + \mathrm{tr}\,(H)\,dt,$$

which yields

$$S_1 = \int_0^1 \boldsymbol{W}'(t)H\,d\boldsymbol{W}(t) = \frac{1}{2}\left(\boldsymbol{W}'(1)H\boldsymbol{W}(1) - \mathrm{tr}\,(H)\right). \tag{1.69}$$

Thus it is easy to derive the c.f. of S_1 (Exercise 1.4.6). Note that, if $H = I_q$, we have

$$\begin{aligned}\int_0^1 \boldsymbol{W}'(t)\,d\boldsymbol{W}(t) &= \sum_{k=1}^q \int_0^1 W_k(t)\,dW_k(t)\\ &= \sum_{k=1}^q \frac{1}{2}\left(W_k^2(1) - 1\right) = \frac{1}{2}\left(\boldsymbol{W}'(1)\boldsymbol{W}(1) - q\right),\end{aligned}$$

which may be established from the scalar Itô integral in (1.49).

As another simple case of (1.69), consider the case of $q=2$, where

$$S_2 = \int_0^1 \boldsymbol{W}'(t)\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix} d\boldsymbol{W}(t) = \int_0^1 \left(W_1(t)\,dW_2(t) + W_2(t)\,dW_1(t)\right). \tag{1.70}$$

Then we have (Exercise 1.4.7)

$$S_2 = \frac{1}{2}\boldsymbol{W}'(1)\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}\boldsymbol{W}(1) = W_1(1)W_2(1) \overset{\mathcal{D}}{=} \frac{1}{2}\left(W_1^2(1) - W_2^2(1)\right).$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

If the matrix H is not symmetric, the formula (1.69) is invalid. As an example, consider the case of $q = 2$, where

$$\begin{aligned} S_3 &= \int_0^1 \boldsymbol{W}'(t) \begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix} d\boldsymbol{W}(t) \\ &= \frac{1}{2} \int_0^1 \Big(W_1(t)\, dW_2(t) - W_2(t)\, dW_1(t) \Big), \end{aligned} \tag{1.71}$$

which is called *Lévy's stochastic area*. If we apply the right-hand side of (1.69) to this integral, it reduces to 0, which is evidently incorrect.

To derive the distribution of S_3, we note that the conditional distribution of S_3 given $\{W_1(t)\}$ is normal with mean 0. Since it holds that

$$\mathrm{E}(W_2(s)dW_2(t)) = \begin{cases} 0 & (s < t), \\ dt & (s \geq t), \end{cases}$$

we have

$$\begin{aligned} &\mathrm{Var}\Big(S_3 |\, \{W_1(t)\}\Big) \\ &= \frac{1}{4} \mathrm{E}\Big[\int_0^1 \int_0^1 \Big(W_1(s)\, dW_2(s) - W_2(s)\, dW_1(s) \Big) \\ &\qquad \times \Big(W_1(t)\, dW_2(t) - W_2(t)\, dW_1(t) \Big) \Big| \{W_1(t)\} \Big] \\ &= \frac{1}{4} \Big[\int_0^1 W_1^2(t)\, dt - 2 \int\!\!\int_{0 \leq s \leq t \leq 1} W_1(s)\, ds\, dW_1(t) \\ &\qquad + \int_0^1 \int_0^1 \min(s,t)\, dW_1(s)\, dW_1(t) \Big] \\ &= \frac{1}{4} \int_0^1 \Big[W_1^2(t) - 2(W_1(1) - W_1(t))W_1(t) + (W_1(1) - W_1(t))^2 \Big] dt \\ &= \int_0^1 \Big(W_1(t) - \frac{1}{2} W_1(1) \Big)^2 dt \\ &\stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 \frac{1}{4} \Big[1 - 2|s - t| \Big] dW_1(s) dW_1(t). \end{aligned} \tag{1.72}$$

Thus it is possible to obtain, using the conditional argument and the result in (1.21),

$$
\begin{aligned}
\mathrm{E}\left(e^{i\theta S_3}\right) &= \mathrm{E}\left[\mathrm{E}\left(e^{i\theta S_3}\Big|\,\{W_1(t)\}\right)\right] \;=\; \mathrm{E}\left[\exp\left\{-\frac{\theta^2}{2}\,\mathrm{Var}\left(S_3|\,\{W_1(t)\}\right)\right\}\right]\\
&= \mathrm{E}\left[\exp\left\{-\frac{\theta^2}{2}\int_0^1\left(W_1(t)-\frac{1}{2}W_1(1)\right)^2 dt\right\}\right]\\
&= \mathrm{E}\left[\exp\left\{-\frac{\theta^2}{2}\int_0^1\int_0^1\frac{1}{4}\Big[1-2|s-t|\Big]\,dW_1(s)\,dW_1(t)\right\}\right]\\
&= \left(D(-\theta^2)\right)^{-1/2} = \left(\cosh\frac{\theta}{2}\right)^{-1},
\end{aligned}
$$

where $D(\lambda)=\left(\cos(\sqrt{\lambda}/2)\right)^2$ is the FD of $K(s,t)=(1-2|s-t|)/4$, and this last result will be proved in Section 2.4. Assuming this result, it can be shown that

$$
S_3 \stackrel{\mathcal{D}}{=} \frac{1}{4}\sum_{n=1}^{\infty}\frac{Z_{n1}^2+Z_{n2}^2-Z_{n3}^2-Z_{n4}^2}{\left(n-\frac{1}{2}\right)\pi},
$$

where $\left(Z_{n1},\ Z_{n2},\ Z_{n3},\ Z_{n4}\right)'\sim \mathrm{NID}(\mathbf{0},\ I_4)$. It is interesting to note that S_2 and S_3 resemble each other closely, but the two distributions are totally different. The distribution of S_2 is that of the weighted sum of a finite number of $\chi^2(1)$ random variables, whereas that of S_3 is the weighted sum of an infinite number of $\chi^2(2)$ random variables.

If we consider a simplified version of S_3 defined by

$$
S_4=\int_0^1 \boldsymbol{W}'(t)\begin{pmatrix}0 & 1/2\\ 0 & 0\end{pmatrix}d\boldsymbol{W}(t)=\frac{1}{2}\int_0^1 W_1(t)\,dW_2(t), \tag{1.73}
$$

it can be shown (Exercise 1.4.8) that, for $\left(Z_{n1},\ Z_{n2}\right)'\sim \mathrm{NID}(\mathbf{0},\ I_2)$,

$$
S_4 \stackrel{\mathcal{D}}{=} \frac{1}{4}\sum_{n=1}^{\infty}\frac{Z_{n1}^2-Z_{n2}^2}{\left(n-\frac{1}{2}\right)\pi}.
$$

Thus we conclude that S_3 is a convolution of S_4, which gives

$$
S_3 \stackrel{\mathcal{D}}{=} \frac{1}{2}\int_0^1\Big(W_1(t)\,dW_2(t)+W_3(t)\,dW_4(t)\Big),
$$

where $\left(W_1(t),W_2(t),W_3(t),W_4(t)\right)'$ is four-dimensional Bm. The c.f. of S_4 is given by

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

$$\mathrm{E}\left(e^{i\theta S_4}\right) = \prod_{n=1}^{\infty}\left\{\left(1 - \frac{i\theta/2}{\left(n-\frac{1}{2}\right)\pi}\right)\left(1 + \frac{i\theta/2}{\left(n-\frac{1}{2}\right)\pi}\right)\right\}^{-1/2}$$

$$= \prod_{n=1}^{\infty}\left(1 + \frac{\theta^2/4}{\left(n-\frac{1}{2}\right)^2\pi^2}\right)^{-1/2} = \left(\cosh\frac{\theta}{2}\right)^{-1/2},$$

so that the c.f. of S_3 is given by

$$\mathrm{E}\left(e^{i\theta S_3}\right) = \left[\mathrm{E}\left(e^{i\theta S_4}\right)\right]^2 = \left(\cosh\frac{\theta}{2}\right)^{-1}. \tag{1.74}$$

Exercises for Section 1.4

Exercise 1.4.1 Prove that

$$\int_0^1 X(t)\,(dW(t))^j = \begin{cases} \int_0^1 X(t)\,dt & (j=2), \\ 0 & (j \geq 3). \end{cases}$$

Exercise 1.4.2 Prove that

$$\frac{1}{n\sigma^2}\sum_{j=1}^{n} y_{j-1}\varepsilon_j \Rightarrow \frac{1}{2}(W^2(1)-1) = \int_0^1 W(t)\,dW(t),$$

where $y_j = y_{j-1} + \varepsilon_j$, $y_0 = 0$ and $\{\varepsilon_j\} \sim$ i.i.d.$(0, \sigma^2)$.

Exercise 1.4.3 Show that

$$\underset{\substack{m\to\infty \\ \Delta_m(t)\to 0}}{\text{l.i.m.}} \sum_{i=1}^{m} W(t_i')(W(t_i) - W(t_{i-1})) = \frac{1}{2}(W^2(t) - t) + \lambda t,$$

where $0 = t_0 < t_1 < \cdots < t_m = t$, $\Delta_m(t) = \max_{1\leq i\leq m}(t_i - t_{i-1})$ and $t_i' = (1-\lambda)t_{i-1} + \lambda t_i$ for $0 \leq \lambda \leq 1$.

Exercise 1.4.4 Show that

$$\int_0^1 g(t)W(t)\,dW(t)$$
$$= \frac{1}{2}\int_0^1\int_0^1 g(\max(s,t))\,dW(s)\,dW(t) - \frac{1}{2}\int_0^1 g(t)\,dt,$$

where $g(t)$ is an ordinary Riemann integrable function.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Exercise 1.4.5 For the O–U process $dX(t) = \gamma X(t)\,dt + dW(t)$ with $X(0) = 0$, prove the following relations:

$$\int_0^1 X^2(t)\,dt = \int_0^1 \int_0^1 \frac{e^{\gamma(2-s-t)} - e^{\gamma|s-t|}}{2\gamma}\,dW(s)\,dW(t)$$
$$\stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 \frac{e^{\gamma(s+t)} - e^{\gamma|s-t|}}{2\gamma}\,dW(s)\,dW(t).$$

Exercise 1.4.6 Derive the c.f. of

$$S_1 = \int_0^1 \boldsymbol{W}'(t) H\,d\boldsymbol{W}(t),$$

where $\boldsymbol{W}(t)$ is q-dimensional Bm, and H is a $q \times q$ symmetric matrix.

Exercise 1.4.7 Show that

$$S_2 = \int_0^1 \Big(W_2(t)\,dW_1(t) + W_1(t)\,dW_2(t)\Big) \stackrel{\mathcal{D}}{=} \frac{1}{2}\Big(W_1^2(1) - W_2^2(1)\Big),$$

where $\boldsymbol{W}(t) = \big(W_1(t), W_2(t)\big)'$ is two-dimensional Bm.

Exercise 1.4.8 Using (1.29) and the conditional argument, show that

$$\int_0^1 W_1(t)\,dW_2(t) \stackrel{\mathcal{D}}{=} \frac{1}{2}\sum_{n=1}^{\infty} \frac{Z_{n1}^2 - Z_{n2}^2}{\left(n - \frac{1}{2}\right)\pi},$$

where $\big(Z_{n1},\ Z_{n2}\big)' \sim \mathrm{NID}(\boldsymbol{0},\ I_2)$.

1.5 Statistical Examples and the Fredholm Determinant

In this section we explore various statistical examples where quadratic functionals of Bm are used, among which are (1) goodness-of-fit tests and Darling's formula, (2) LBI tests for parameter constancy, (3) LBIU tests for the moving average unit root, and (4) LBI and LBIU tests for the autoregressive unit root.

1.5.1 Goodness-of-Fit Tests and Darling's Formula

Let $X_1, \ldots, X_n$ be a random sample from the common distribution $G(x)$ assumed to be absolutely continuous. The testing problem considered here is to test if $G(x) = F(x;\xi)$, where $F(x;\xi)$ is a distribution function parameterized by ξ. For illustrative purposes we concentrate here on the case where ξ is a single parameter. The multi-parameter case will be examined in Chapters 3 and 4. Depending on whether ξ is specified or not, we deal with two cases.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Case 1.1 The testing problem considered here is

$$H_0: \ G(x) = F(x;\xi),$$

where ξ is a nuisance parameter to be estimated. This case was discussed in Darling (1955), where the test statistic is a modified version of the *Cramér–von Mises–Smirnov statistic* and is given by

$$W_n^2(\hat{\xi}) = n \int_{-\infty}^{\infty} \left(F_n(x) - F(x;\hat{\xi}) \right)^2 dF(x;\hat{\xi}), \tag{1.75}$$

where $\hat{\xi}$ is a suitable estimator of ξ, and $F_n(x)$ is the empirical distribution function defined by

$$F_n(x) = \frac{1}{n} \sum_{j=1}^{n} \epsilon\left(x - X_j\right), \qquad \epsilon(x) = \begin{cases} 1 & (x \geq 0), \\ 0 & (x < 0). \end{cases}$$

For practical computation of $W_n^2(\hat{\xi})$, we rearrange the sample $X_1, \ldots, X_n$ to get $X_1' < \cdots < X_n'$, so that (Exercise 1.5.1)

$$W_n^2(\hat{\xi}) = \sum_{j=1}^{n} \left(F(X_j';\hat{\xi}) - \frac{2j-1}{2n} \right)^2 + \frac{1}{12n}.$$

Let us consider the asymptotic distribution of $W_n^2(\hat{\xi})$ under H_0. Suppose that $\hat{\xi}$ is the maximum likelihood estimator (MLE) of ξ that satisfies some regularity conditions. Here we exclude the super-efficient case such that $n\mathrm{E}[(\hat{\xi} - \xi)^2] \to 0$ as $n \to \infty$. Putting

$$g_\xi(t) = \left. \frac{\partial F(x;\xi)}{\partial \xi} \right|_{x = x(t;\xi)}, \qquad x(t;\xi) = \inf\{x : F(x;\xi) = t\},$$

Darling (1955) considered

$$\begin{aligned} W_n^2(\hat{\xi}) &= \int_{-\infty}^{\infty} \left(\sqrt{n}\,(F_n(x) - F(x;\xi)) - \sqrt{n}\,(\hat{\xi} - \xi) \frac{\partial F(x;\xi^*)}{\partial \xi} \right)^2 dF(x;\hat{\xi}) \\ &= \int_0^1 \left(Y_n(t) - \sqrt{n}\,(\hat{\xi} - \xi)\, g_\xi(t) \right)^2 dt + \delta_n \\ &= \int_0^1 Z_n^2(t)\, dt + \delta_n, \qquad \mathrm{plim}_{n \to \infty} \delta_n = 0, \end{aligned}$$

where ξ^* lies between ξ and $\hat{\xi}$ and plim stands for convergence in probability, whereas

$$Z_n(t) = Y_n(t) - \sqrt{n}\,(\hat{\xi} - \xi)\, g_\xi(t),$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

and

$$Y_n(t) = \sqrt{n}\,(H_n(t) - t), \qquad H_n(t) = \frac{1}{n}\sum_{j=1}^{n} \epsilon\big(t - F(X_j;\xi)\big).$$

Since $F(X_1;\xi), \ldots, F(X_n;\xi)$ are independent and uniformly distributed over [0, 1], it follows (Exercise 1.5.2) that

$$\mathrm{E}(H_n(t)) = t, \qquad \mathrm{Cov}(H_n(s), H_n(t)) = \frac{1}{n}\,(\min(s,t) - st),$$

so that

$$\mathrm{E}(Y_n(t)) = 0, \qquad \mathrm{Cov}(Y_n(s),\, Y_n(t)) = \min(s,t) - st.$$

Then it was shown by Darling (1955) that

$$\begin{aligned} W_n^2(\hat{\xi}) &\Rightarrow \int_0^1 Z^2(t)\,dt \;=\; W^2(\xi), \\ W^2(\xi) &\overset{\mathcal{D}}{=} \int_0^1\int_0^1 K(s,t)\,dW(s)\,dW(t), \end{aligned} \tag{1.76}$$

where $\{Z(t)\}$ is a zero-mean Gaussian process and

$$K(s,t) = \mathrm{Cov}(Z(s),\, Z(t)) = \min(s,t) - st - \sigma^2(\xi)\,g_\xi(s)\,g_\xi(t), \tag{1.77}$$

$$\sigma^{-2}(\xi) = \mathrm{E}\left\{\left(\frac{\partial \log f(X_1;\xi)}{\partial\xi}\right)^2\right\} = \int_{-\infty}^{\infty}\left(\frac{\partial \log f(x;\xi)}{\partial\xi}\right)^2 f(x;\xi)\,dx,$$

with $f(x;\xi) = \partial F(x;\xi)/\partial x$.

It must be ensured that $\sigma(\xi)g_\xi(t)$ does not depend on ξ. If it does, the present test is useless because we cannot compute significance points. It was shown in Darling (1955) that, if ξ is a location or scale parameter, $\sigma(\xi)g_\xi(t)$ does not depend on ξ, so that the $W_n^2(\hat{\xi})$ test is parameter free. In fact, it holds (Exercise 1.5.3) that, when $F(x;\xi) = R(x-\xi)$ and $r(x) = R'(x)$,

$$\sigma(\xi)\,g_\xi(t) = -\left(\int_{-\infty}^{\infty}\frac{(r'(x))^2}{r(x)}\,dx\right)^{-1/2} r(R^{-1}(t)),$$

and, when $F(x;\xi) = R(x/\xi)$,

$$\sigma(\xi)\,g_\xi(t) = -\left(\int_{-\infty}^{\infty}\frac{(x\,r'(x))^2}{r(x)}\,dx - 1\right)^{-1/2} R^{-1}(t)\,r(R^{-1}(t)).$$

It also holds (Exercise 1.5.4) that, when $F(x;\xi) = (R(x))^{\xi}$ $(\xi > 0)$,

$$\sigma(\xi)\,g_\xi(t) = t\,\log t.$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

As a specific example, let us consider

$$F(x;\xi) = \frac{1}{1+e^{-(x-\xi)}},$$

which is the logistic distribution function with location parameter ξ. We have

$$\frac{\partial F(x;\xi)}{\partial \xi} = \frac{-e^{-(x-\xi)}}{\left(1+e^{-(x-\xi)}\right)^2}, \qquad F(x;\xi) = t \ \Leftrightarrow\ e^{-(x-\xi)} = \frac{1-t}{t},$$

which yields

$$g_\xi(t) = \left.\frac{\partial F(x;\xi)}{\partial \xi}\right|_{x=x(t;\xi)} = -t(1-t).$$

Since

$$f(x;\xi) = \frac{e^{-(x-\xi)}}{\left(1+e^{-(x-\xi)}\right)^2}, \quad \frac{\partial \log f(x;\xi)}{\partial \xi} = 1 - \frac{2e^{-(x-\xi)}}{1+e^{-(x-\xi)}},$$

we have

$$\begin{aligned}\sigma^{-2}(\xi) &= \int_{-\infty}^{\infty}\left(1-\frac{2e^{-(x-\xi)}}{1+e^{-(x-\xi)}}\right)^2 \frac{e^{-(x-\xi)}}{\left(1+e^{-(x-\xi)}\right)^2}\,dx\\ &= \int_0^1 (2t-1)^2\,dt = \frac{1}{3}.\end{aligned}$$

Thus the distribution of $W^2(\xi)$ in the present case is given by (1.76), and it follows from (1.77) that

$$\mathrm{Cov}(Z(s),\ Z(t)) = \min(s,t) - st - 3st(1-s)(1-t). \tag{1.78}$$

The computation of the c.f. of $W^2(\xi)$ can be carried out using Darling's formula (see Darling (1955, theorem 6.5) and (2.83)), which gives the FD of $\mathrm{Cov}(Z(s), Z(t))$ in (1.78) as

$$\begin{aligned}D(\lambda) &= \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}} - 6\int_0^1\left(\int_0^t g_\xi'(s)\cos\sqrt{\lambda}s\,ds\right) g_\xi'(t)\cos\sqrt{\lambda}(1-t)\,dt\\ &= \frac{12}{\lambda}\left(\frac{2\left(1-\cos\sqrt{\lambda}\right)}{\lambda} - \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}}\right).\end{aligned} \tag{1.79}$$

Thus we obtain

$$\mathrm{E}\left(e^{i\theta W^2(\xi)}\right) = (D(2i\theta))^{-1/2} = \left[\frac{6}{i\theta}\left(\frac{1-\cos\sqrt{2i\theta}}{i\theta} - \frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)\right]^{-1/2}. \tag{1.80}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

As another example, consider

$$F(x;\xi) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\frac{x}{\xi} \quad (\xi > 0),$$

which is the Cauchy distribution with scale parameter ξ. We have

$$\begin{aligned}\frac{\partial F(x;\xi)}{\partial \xi} &= -\frac{x}{\xi^2\pi}\frac{1}{1+x^2/\xi^2},\\ F(x;\xi) = t \;\Leftrightarrow\; \frac{x}{\xi} &= \tan\pi(t-1/2) = -\cot\pi t,\end{aligned}$$

which yields

$$g_\xi(t) = \left.\frac{\partial F(x;\xi)}{\partial \xi}\right|_{x=x(t;\xi)} = \frac{\cot\pi t}{\xi\pi}\frac{1}{1+\cot^2\pi t} = \frac{1}{2\xi\pi}\sin 2\pi t.$$

Since

$$f(x;\xi) = \frac{1}{\xi\pi}\frac{1}{1+x^2/\xi^2}, \qquad \frac{\partial \log f(x;\xi)}{\partial \xi} = -\frac{1}{\xi} + \frac{2x^2/\xi^3}{1+x^2/\xi^2},$$

we have

$$\begin{aligned}\sigma^{-2}(\xi) &= \int_{-\infty}^{\infty}\left(-\frac{1}{\xi} + \frac{2x^2/\xi^3}{1+x^2/\xi^2}\right)^2\frac{1}{\xi\pi}\frac{1}{1+x^2/\xi^2}\,dx\\ &= \frac{1}{\xi^2} + \frac{8}{\xi^2\pi}\int_0^{\infty}\left(\frac{-u^2}{(1+u^2)^2} + \frac{u^4}{(1+u^2)^3}\right)du\\ &= \frac{1}{2\xi^2},\end{aligned}$$

where we have used the fact that

$$\int_0^{\infty}\frac{x^{a-1}}{(1+x^c)^b}\,dx = \frac{1}{c}B(b-a/c, a/c) \quad (a>0,\ b>a/c,\ c>0).$$

Here $B(p,q)$ is the beta function defined by

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,dt \quad (p>0,\ q>0).$$

Thus the distribution of $W^2(\xi)$ in the present case is given by (1.76) with

$$\mathrm{Cov}(Z(s),\ Z(t)) = \min(s,t) - st - \frac{1}{2\pi^2}\sin 2\pi s\ \sin 2\pi t. \tag{1.81}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Then, using Darling's formula in (2.83), the FD of $\mathrm{Cov}(Z(s), Z(t))$ in (1.81) is given by

$$\begin{aligned} D(\lambda) &= \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}} - 4\int_0^1 \left(\int_0^t \cos 2\pi s\ \cos\sqrt{\lambda}s\, ds\right) \cos 2\pi t\ \cos\sqrt{\lambda}(1-t)\, dt \\ &= \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}} \Bigg/ \left(1 - \frac{\lambda}{4\pi^2}\right). \end{aligned} \tag{1.82}$$

Thus we now obtain

$$\mathrm{E}\left(e^{i\theta W^2(\xi)}\right) = (D(2i\theta))^{-1/2} = \left[\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}} \Bigg/ \left(1 - \frac{i\theta}{2\pi^2}\right)\right]^{-1/2}. \tag{1.83}$$

Case 1.2 Here we deal with the testing problem

$$H_0:\ G(x) = F(x;\xi_0) \quad \text{versus} \quad H_1:\ G(x) = F(x;\xi), \quad \xi = \xi_0 + \frac{\gamma}{\sqrt{n}},$$

where ξ_0 is a specified value and γ is a constant. This case was discussed in Durbin (1973). Using the Cramér–von Mises–Smirnov statistic

$$W_n^2(\xi_0) = n\int_{-\infty}^{\infty} \Big(F_n(x) - F(x;\xi_0)\Big)^2 dF(x;\xi_0), \tag{1.84}$$

the null hypothesis H_0 is rejected if $W_n^2(\xi_0)$ is sufficiently large.

Let us consider the limiting distribution of $W_n^2(\xi_0)$ as $n \to \infty$. We first deal with the distribution under H_0. Putting $G(x) = F(x;\xi_0) = t$, we have

$$\begin{aligned} W_n^2(\xi_0) &= n\int_{-\infty}^{\infty} \left(\frac{1}{n}\sum_{j=1}^n \epsilon(x - X_j) - F(x;\xi_0)\right)^2 dF(x;\xi_0) \\ &= n\int_0^1 \left(\frac{1}{n}\sum_{j=1}^n \epsilon\Big(t - F(X_j;\xi_0)\Big) - t\right)^2 dt \\ &= \int_0^1 Y_n^2(t)\, dt, \end{aligned}$$

where

$$Y_n(t) = \sqrt{n}\,(H_n(t) - t), \qquad H_n(t) = \frac{1}{n}\sum_{j=1}^n \epsilon\Big(t - F(X_j;\xi_0)\Big).$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

It follows that $\{Y_n(t)\} \Rightarrow \{Y(t)\}$, where $\{Y(t)\}$ is the Bb. Thus it holds that

$$\begin{aligned} W_n^2(\xi_0) \Rightarrow W^2(\xi_0) &= \int_0^1 Y^2(t)\,dt \\ &\stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 [\min(s,t) - st]\,dW(s)\,dW(t), \end{aligned} \tag{1.85}$$

and the c.f. of $W^2(\xi_0)$ is given from (1.35) by

$$\mathrm{E}\left(e^{i\theta W^2(\xi_0)}\right) = (D(2i\theta))^{-1/2} = \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-1/2},$$

where $D(\lambda)$ is the FD of $K(s,t) = \min(s,t) - st$.

We next consider the limiting distribution of $W_n^2(\xi_0)$ under $H_1 : \xi = \xi_0 + \gamma/\sqrt{n}$. Expanding $F(x;\xi_0)$ around $\xi = \xi_0 + \gamma/\sqrt{n}$, we consider, in place of $W_n^2(\xi_0)$,

$$\begin{aligned} R_n^2(\xi_0) &= n\int_{-\infty}^{\infty}\left(F_n(x) - F(x;\xi) + \frac{\gamma}{\sqrt{n}}\left.\frac{\partial F(x;\xi)}{\partial \xi}\right|_{\xi=\xi_0}\right)^2 dF(x;\xi) \\ &= \int_0^1 \left(Y_n(t) + \gamma g_{\xi_0}(t)\right)^2\,dt, \end{aligned} \tag{1.86}$$

where

$$Y_n(t) = \sqrt{n}(H_n(t) - t), \qquad H_n(t) = \frac{1}{n}\sum_{j=1}^{n} \epsilon\big(t - F(X_j;\xi)\big),$$

$$g_{\xi_0}(t) = \left.\frac{\partial F(x;\xi)}{\partial \xi}\right|_{x=x(t;\xi_0)}, \qquad x(t;\xi_0) = \inf\{x : F(x;\xi_0) = t\}.$$

It follows from Durbin (1973) that, under H_1 and some regularity conditions, the limiting distributions of $W_n^2(\xi_0)$ and $R_n^2(\xi_0)$ are the same and

$$R_n^2(\xi_0) \;\Rightarrow\; R^2(\xi_0) = \int_0^1 \left(Y(t) + \gamma g_{\xi_0}(t)\right)^2\,dt, \tag{1.87}$$

where $\{Y(t)\}$ is the Bb. Unlike the previous statistics, the statistic $R^2(\xi_0)$ is a quadratic functional of a nonzero-mean Gaussian process $\{Y(t) + \gamma g_{\xi_0}(t)\}$. To derive the c.f. of such a functional, it is not sufficient to compute the FD of $K(s,t) = \mathrm{Cov}(Y(s), Y(t))$; we also need to deal with the *resolvent* associated with $K(s,t)$ and the function $m(t) = \gamma g_{\xi_0}(t)$.

Discussions on the resolvent are deferred until Chapter 3. Here we take up two examples only for demonstration purposes. Consider first the logistic

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

distribution function $F(x;\xi) = 1/(1+e^{-(x-\xi)})$ dealt with in Case 1.1. Then we have

$$g_{\xi_0}(t) = \left.\frac{\partial\left(1+e^{-(x-\xi)}\right)^{-1}}{\partial\xi}\right|_{x=x(t;\xi_0)} = -t(1-t).$$

Thus (1.87) holds under H_1 with $g_{\xi_0}(t) = -t(1-t)$. The corresponding c.f. of $R^2(\xi_0)$, the derivation of which involves the resolvent and will be explained in Section 3.2, is given by

$$\begin{aligned}
&\mathrm{E}\left(e^{i\theta R^2(\xi_0)}\right)\\
&\quad = \mathrm{E}\left[\exp\left\{i\theta\int_0^1 (Y(t)-\gamma t(1-t))^2\,dt\right\}\right]\\
&\quad = \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-1/2}\exp\left[\frac{\gamma^2}{\theta^2}\left(\frac{\cos\sqrt{2i\theta}-1}{\sin\sqrt{2i\theta}/\sqrt{2i\theta}}+i\theta-\frac{\theta^2}{6}\right)\right].
\end{aligned}\tag{1.88}$$

It can be seen that $R^2(\xi_0)$ does not depend on ξ_0 and tends to normality as $|\gamma| \to \infty$.

As the second example, consider the Cauchy distribution $F(x;\xi) = 1/2 + \tan^{-1}(x/\xi)/\pi$ also dealt with in Case 1.1, where $\xi > 0$. We have

$$g_{\xi_0}(t) = \left.\frac{\partial}{\partial\xi}\left(\frac{1}{2}+\frac{1}{\pi}\tan^{-1}\frac{x}{\xi}\right)\right|_{x=x(t;\xi_0)} = \frac{1}{2\xi_0\pi}\sin 2\pi t,$$

and

$$\begin{aligned}
\mathrm{E}\left(e^{i\theta R^2(\xi_0)}\right) &= \mathrm{E}\left(\exp\left\{i\theta\int_0^1\left(Y(t)+\frac{\gamma}{2\xi_0\pi}\sin 2\pi t\right)^2 dt\right\}\right)\\
&= \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-1/2}\exp\left[\frac{\gamma^2}{\xi_0^2}\frac{i\theta}{4(2\pi^2-i\theta)}\right].
\end{aligned}\tag{1.89}$$

The derivation of this c.f. also involves the resolvent and will be demonstrated in Section 3.2. Unlike $R^2(\xi_0)$ in (1.88), the present statistic $R^2(\xi_0)$ depends on ξ_0, although it also tends to normality as $|\gamma| \to \infty$.

Goodness-of-fit tests will be discussed in more detail in Chapter 4, dealing with the multi-parameter case.

1.5.2 LBI Tests for Parameter Constancy

Let us consider the following *state space model*:

$$y_j = \beta_j + \varepsilon_j,\quad \beta_j = \beta_{j-1} + \eta_j \qquad (j = 1, \ldots, n), \tag{1.90}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where

(i) $\{y_j\}$ is an observable sequence, whereas $\{\beta_j\}$ is an unobservable sequence starting from an unknown constant β_0;
(ii) $\{\varepsilon_j\} \sim \mathrm{NID}(0, \sigma_\varepsilon^2)$, $\{\eta_j\} \sim \mathrm{NID}(0, \sigma_\eta^2)$, and these are independent of each other.

The state space model is composed of two equations. One is a measurement equation that describes the relation between the observed variable y_j and the unobserved state variable β_j. The other is a state equation that reflects the dynamics of the state variable β_j.

Our concern is to test if β_j, the level of the process, is constant, that is, $\beta_j = \beta_0$ for all j. This is equivalent to testing

$$H_0 : \rho = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} = 0 \qquad \text{versus} \qquad H_1 : \rho > 0. \tag{1.91}$$

The present testing problem was first discussed in Nyblom and Mäkeläinen (1983) and Tanaka (1983), and was generalized by Nabeya and Tanaka (1988) and Nabeya (1989) to various cases, which we shall deal with in subsequent chapters.

Since $y_j = \beta_0 + \varepsilon_j + \eta_1 + \cdots + \eta_j$, the observation vector $\boldsymbol{y} = (y_1, \ldots, y_n)'$ follows

$$\boldsymbol{y} = \beta_0 \boldsymbol{e} + \boldsymbol{\varepsilon} + C\boldsymbol{\eta} \sim \mathrm{N}\big(\beta_0 \boldsymbol{e}, \sigma_\varepsilon^2 \Omega(\rho)\big), \tag{1.92}$$

where $\boldsymbol{e} = (1, \ldots, 1)' : n \times 1$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)'$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)'$, and

$$\Omega(\rho) = I_n + \rho\, CC', \ C = \begin{pmatrix} 1 & & & & \\ \cdot & \cdot & & 0 & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ 1 & \cdot & \cdot & \cdot & 1 \end{pmatrix}, \ C^{-1} = \begin{pmatrix} 1 & & & & \\ -1 & \cdot & & 0 & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ 0 & & & -1 & 1 \end{pmatrix}. \tag{1.93}$$

Defining the parameter vector $\boldsymbol{\delta} = (\rho, \beta_0, \sigma_\varepsilon^2)'$, the log-likelihood for $\boldsymbol{y}$ is given by

$$L(\boldsymbol{\delta}) = -\frac{n}{2} \log 2\pi\sigma_\varepsilon^2 - \frac{1}{2}\log|\Omega(\rho)| - \frac{1}{2\sigma_\varepsilon^2}(\boldsymbol{y} - \beta_0\boldsymbol{e})'\Omega^{-1}(\rho)(\boldsymbol{y} - \beta_0\boldsymbol{e}). \tag{1.94}$$

The matrix C necessarily arises from the random walk process $\beta_j = \beta_{j-1} + \eta_j$ and may be called the *random walk-generating matrix*. Note that the (j,k)th element of CC' is $\min(j,k)$, and the eigenvalues of CC' were obtained by

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Rutherford (1946) as

$$\lambda_j = \frac{1}{4}\left(\sin\frac{j-\frac{1}{2}}{2n+1}\,\pi\right)^{-2} \qquad (j = 1, \ldots, n).$$

For the present problem there exists no uniformly best test. Here we consider the *locally best invariant* (LBI) test (Ferguson, 1967). It follows from (1.92) that the testing problem (1.91) is invariant under the group of transformations: $\boldsymbol{y} \to a\boldsymbol{y} + b\boldsymbol{e}$ and $(\rho, \beta_0, \sigma_\varepsilon^2) \to (\rho, a\beta_0 + b, a^2\sigma_\varepsilon^2)$, where $a > 0$. Following Kariya (1980), choose an $n \times (n-1)$ matrix H such that $H'H = I_{n-1}$ and $HH' = I_n - \boldsymbol{ee}'/n = M = M' = M^2$. Let us put $\boldsymbol{w} = H'\boldsymbol{y} = H'(\beta_0\boldsymbol{e} + \boldsymbol{\varepsilon} + C\boldsymbol{\eta})$. Noting that $M\boldsymbol{e} = HH'\boldsymbol{e} = \boldsymbol{0}$ and thus $H'\boldsymbol{e} = \boldsymbol{0}$, we have

$$\boldsymbol{w} \sim \mathrm{N}\left(\boldsymbol{0}, \sigma_\varepsilon^2 H'\Omega(\rho)H\right), \tag{1.95}$$

and the statistic $s(\boldsymbol{w}) = \boldsymbol{w}/\sqrt{\boldsymbol{w}'\boldsymbol{w}}$ is a maximal invariant under the above group. Let $P_\rho(\cdot)$ be the distribution of $s(\boldsymbol{w})$, and put

$$f_n(s(\boldsymbol{w})|\rho) = |H'\Omega(\rho)H|^{-1/2}\left[\frac{\boldsymbol{w}'(H'\Omega(\rho)H)^{-1}\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{w}}\right]^{-(n-1)/2}, \tag{1.96}$$

which is the probability density of $P_\rho(\cdot)$ with respect to $P_0(\cdot)$. The rejection region of the LBI test is now obtained (Ferguson, 1967, p.235) as

$$\left.\frac{\partial \log f_n(s(\boldsymbol{w})|\rho)}{\partial\rho}\right|_{\rho=0} > \text{constant}, \tag{1.97}$$

which leads to the rejection region given by

$$R_{n1} = \frac{\boldsymbol{w}'H'CC'H\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{w}} = \frac{\boldsymbol{y}'MCC'M\boldsymbol{y}}{\boldsymbol{y}'M\boldsymbol{y}} > \text{constant}. \tag{1.98}$$

The derivation of the LBI test is somewhat complicated. It can be shown (Tanaka, 1996, chapter 9) that the LBI test for the present problem is equivalent to the *Lagrange multiplier* (LM) *test*, which is easier to derive. In fact, the rejection region of the LM test is given by

$$\left.\frac{\partial L(\boldsymbol{\delta})}{\partial\rho}\right|_{H_0} > \text{constant}, \tag{1.99}$$

where the partial derivative is evaluated at the MLE $\tilde{\boldsymbol{\delta}} = (\tilde{\rho}, \tilde{\beta}_0, \tilde{\sigma}_\varepsilon^2)'$ of $\boldsymbol{\delta}$ under H_0. Note that $\tilde{\rho} = 0$. We also have

$$\tilde{\beta}_0 = \frac{1}{n}\sum_{j=1}^n y_j = \frac{1}{n}\boldsymbol{e}'\boldsymbol{y}, \quad \tilde{\sigma}_\varepsilon^2 = \frac{1}{n}\sum_{j=1}^n (y_j - \bar{y})^2 = \frac{1}{n}\boldsymbol{y}'M\boldsymbol{y}.$$

Then, using

$$\frac{d\log|\Omega(\rho)|}{d\rho}=\mathrm{tr}\left(\Omega^{-1}(\rho)\frac{d\Omega(\rho)}{d\rho}\right),\quad \frac{d\Omega^{-1}(\rho)}{d\rho}=-\Omega^{-1}(\rho)\frac{d\Omega(\rho)}{d\rho}\Omega^{-1}(\rho),$$

we have

$$\begin{aligned}\left.\frac{\partial L(\boldsymbol{\delta})}{\partial\rho}\right|_{H_0}&=-\frac{1}{2}\mathrm{tr}(CC')+\frac{1}{2\tilde{\sigma}_\varepsilon^2}\,\boldsymbol{y}'MCC'M\boldsymbol{y}\\&=-\frac{n(n+1)}{4}+\frac{n}{2}\frac{\boldsymbol{y}'MCC'M\boldsymbol{y}}{\boldsymbol{y}'M\boldsymbol{y}}.\end{aligned}\tag{1.100}$$

Thus it is seen that the LM test is equivalent to the LBI test. An interesting feature of the present test is that $\partial L(\boldsymbol{\delta})/\partial\rho|_{H_0}$ does not tend to normality, which will be seen below. This contrasts with the standard situation.

Let us derive the asymptotic distribution of the LBI statistic R_{n1} in (1.98) under $H_0:\ \rho=0$. For this purpose we can assume that $\sigma_\varepsilon^2=1$ without any loss of generality and consider

$$S_{n1}=\frac{1}{n}R_{n1}=\frac{1}{n}\frac{\boldsymbol{y}'MCC'M\boldsymbol{y}}{\boldsymbol{y}'M\boldsymbol{y}}.\tag{1.101}$$

Under H_0, it holds that $\boldsymbol{y}=\beta_0\boldsymbol{e}+\boldsymbol{\varepsilon}$ and

$$\frac{1}{n}\,\boldsymbol{y}'M\boldsymbol{y}=\frac{1}{n}\,\boldsymbol{\varepsilon}'M\boldsymbol{\varepsilon}\ \to\ 1\quad\text{in probability,}\tag{1.102}$$

$$\begin{aligned}\frac{1}{n^2}\,\boldsymbol{y}'MCC'M\boldsymbol{y}&=\frac{1}{n^2}\,\boldsymbol{\varepsilon}'MCC'M\boldsymbol{\varepsilon}=\frac{1}{n}\,\boldsymbol{\varepsilon}'B_{n1}\boldsymbol{\varepsilon}\\&=\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}B_{n1}(j,k)\varepsilon_j\varepsilon_k,\end{aligned}\tag{1.103}$$

where $B_{n1}=MCC'M/n$, and $B_{n1}(j,k)$ is the (j,k)th element of B_{n1} given by

$$\begin{aligned}B_{n1}(j,k)\ =\ &\frac{1}{n}\min(j,k)-\frac{1}{2n^2}\,j(2n-j+1)-\frac{1}{2n^2}\,k(2n-k+1)\\&+\frac{1}{6n^3}\,n(n+1)(2n+1).\end{aligned}\tag{1.104}$$

To derive the limiting distribution of the quadratic form in (1.103), we can use the following theorem. Its extended version is given in Nabeya and Tanaka (1988) and will be proved in Chapter 5 (see Theorem 5.1).

Theorem 1.11 *Consider*

$$S_n=\frac{1}{n}\boldsymbol{\varepsilon}'B_n\boldsymbol{\varepsilon}=\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}B_n(j,k)\varepsilon_j\varepsilon_k,$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where $\{\varepsilon_j\} \sim i.i.d.(0,1)$ *and* B_n *is an* $n \times n$ *symmetric and nonnegative definite matrix. Assume that, for the* (j,k)*th element* $B_n(j,k)$ *of* B_n*, there exists a symmetric and continuous function* $K(s,t)$ *satisfying*

$$\lim_{n\to\infty} \max_{1\le j,k\le n} \left| B_n(j,k) - K\left(\frac{j}{n}, \frac{k}{n}\right) \right| = 0. \tag{1.105}$$

Then it holds that, as $n \to \infty$*,*

$$S_n \;\Rightarrow\; S = \int_0^1 \int_0^1 K(s,t)\, dW(s)\, dW(t). \tag{1.106}$$

Note that, because B_n is nonnegative definite and the condition (1.105) is imposed, the statistic S is nonnegative. This can be weakened to the extent that the matrix B_n is indefinite so that S is not necessarily nonnegative to deal with more general cases, which will be discussed in Chapter 5.

Let us return to (1.103) and (1.104). It holds that the matrix $B_{n1} = MCC'M/n$ with $B_{n1}(j,k)$ in the (j,k)th place is nonnegative definite and the function

$$K_1(s,t) = \min(s,t) - s - t + \frac{s^2+t^2}{2} + \frac{1}{3} = \frac{1}{3} - \max(s,t) + \frac{s^2+t^2}{2}$$

is the uniform limit of $B_{n1}(j,k)$ in the sense of (1.105). Thus Theorem 1.11 leads us to establish that

$$\begin{aligned} \frac{1}{n^2}\, \mathbf{y}'MCC'M\mathbf{y} \;&=\; \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} B_{n1}(j,k)\varepsilon_j\varepsilon_k \\ &\Rightarrow \int_0^1 \int_0^1 K_1(s,t)\, dW(s)\, dW(t). \end{aligned} \tag{1.107}$$

Therefore it follows that

$$S_{n1} = \frac{1}{n}\, \frac{\mathbf{y}'MCC'M\mathbf{y}}{\mathbf{y}'M\mathbf{y}} \;\Rightarrow\; S_1 = \int_0^1 \int_0^1 K_1(s,t)\, dW(s)\, dW(t). \tag{1.108}$$

The c.f. of S_1 is given from (1.35) and (1.38) by $(D_1(2i\theta))^{-1/2}$, where $D_1(\lambda) = \sin\sqrt{\lambda}/\sqrt{\lambda}$ is the FD of $K_1(s,t)$.

In the present case there is a simpler way of deriving the limiting distribution of S_{n1}. Noting that

$$\begin{aligned} \frac{1}{n^2}\, \mathbf{y}'MCC'M\mathbf{y} &= \frac{1}{n^2}\, \boldsymbol{\varepsilon}'MCC'M\boldsymbol{\varepsilon} \;\overset{\mathcal{D}}{=}\; \frac{1}{n^2}\, \boldsymbol{\varepsilon}C'MC\boldsymbol{\varepsilon} \\ &= \frac{1}{n}\boldsymbol{\varepsilon}' B_{n1}^{\star}\boldsymbol{\varepsilon} = \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} B_{n1}^{\star}(j,k), \end{aligned}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where $B_{n1}^{\star} = C'MC/n$ and $B_{n1}^{\star}(j,k)$ is the (j,k)th element of $B_{n1}^{\star}$ defined by

$$B_{n1}^{\star}(j,k) = \frac{1}{n}\,\min(n+1-j, n+1-k) - \frac{1}{n^2}\,(n-j+1)(n-k+1), \qquad (1.109)$$

we can find the uniform limit $K_1^{\star}(s,t)$ of $B_{n1}^{\star}(j,k)$ as

$$K_1^{\star}(s,t) = \min(1-s, 1-t) - (1-s)(1-t) = \min(s,t) - st.$$

Thus it follows that

$$S_{n1} \;\Rightarrow\; S_1 \;\stackrel{\mathcal{D}}{=}\; \int_0^1 \int_0^1 [\min(s,t) - st]\, dW(s)\, dW(t). \qquad (1.110)$$

In the present model (1.92), the initial value β_0 was unknown. If β_0 is assumed to be known, in which case we can put $\beta_0 = 0$, then the above results hold true by replacing $K_1(s,t)$ by $\min(s,t)$ (Exercise 1.5.5).

Nabeya and Tanaka (1988) extended the present model to

$$y_j = x_j \beta_j + z_j' \boldsymbol{\gamma} + \varepsilon_j, \quad \beta_j = \beta_{j-1} + \eta_j \quad (j = 1, \ldots, n),$$

where x_j and z_j are scalar and vector deterministic regressors, respectively. We shall deal with this extended case in Chapter 5.

It is also interesting to examine the power of the LBI test. For this purpose we consider the distribution of the LBI test statistic under the local alternative

$$H_1(c,\kappa):\; \rho = \frac{c}{n^{\kappa}},$$

where c and κ are positive constants. We shall discuss the local alternative in Chapter 5.

1.5.3 LBIU Tests for the Moving Average Unit Root

In this section we deal with the first-order moving average (MA(1)) model given by

$$y_j - \mu = \varepsilon_j - \alpha\, \varepsilon_{j-1} \qquad (j = 1, \ldots, n), \qquad (1.111)$$

where μ is an unknown mean and α $(0 < \alpha \le 1)$ is the coefficient parameter, whereas $\{\varepsilon_j\}$ $(j = 0, 1, \ldots, n)$ is an error sequence following $\mathrm{NID}(0, \sigma_\varepsilon^2)$.

Our concern here is to consider testing

$$H_0:\; \alpha = 1 \qquad \text{versus} \qquad H_1:\; \alpha < 1, \qquad (1.112)$$

which was initially discussed in Tanaka (1990b) when $\mu = 0$. The MA(1) model (1.111) is said to be noninvertible under H_0. The log-likelihood for $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)'$ is

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

$$L(\boldsymbol{\delta}) = -\frac{n}{2}\log 2\pi\sigma_\varepsilon^2 - \frac{1}{2}\log|\Omega(\alpha)| - \frac{1}{2\sigma_\varepsilon^2}(\boldsymbol{y}-\mu\boldsymbol{e})'\Omega^{-1}(\alpha)(\boldsymbol{y}-\mu\boldsymbol{e}), \tag{1.113}$$

where $\boldsymbol{\delta} = (\alpha, \mu, \sigma_\varepsilon^2)'$, $\boldsymbol{e} = (1, \ldots, 1)' : n \times 1$, and

$$\Omega(\alpha) = \begin{pmatrix} 1+\alpha^2 & -\alpha & & & 0 \\ -\alpha & 1+\alpha^2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & -\alpha \\ 0 & & & -\alpha & 1+\alpha^2 \end{pmatrix}.$$

Proceeding in the same way as in the last section, the testing problem (1.112) is shown to be invariant under the group of transformations: $\boldsymbol{y} \to a\boldsymbol{y} + b\boldsymbol{e}$ and $(\alpha, \mu, \sigma_\varepsilon^2) \to (\alpha, a\mu+b, a^2\sigma_\varepsilon^2)$, where $a > 0$. Let us put $\boldsymbol{w} = H'\boldsymbol{y}$, where H is the $n \times (n-1)$ matrix such that $H'H = I_{n-1}$ and $HH' = I_n - \boldsymbol{e}\boldsymbol{e}'/n = M$. Then we have

$$\boldsymbol{w} \sim \mathrm{N}\left(\boldsymbol{0}, \sigma_\varepsilon^2 H'\Omega(\alpha)H\right), \tag{1.114}$$

and the probability density of the maximal invariant $s(\boldsymbol{w}) = \boldsymbol{w}/\sqrt{\boldsymbol{w}'\boldsymbol{w}}$ is given by

$$f_n(s(\boldsymbol{w})|\alpha) = |H'\Omega(\alpha)H|^{-1/2}\left[\frac{\boldsymbol{w}'(H'\Omega(\alpha)H)^{-1}\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{w}}\right]^{-(n-1)/2}. \tag{1.115}$$

Using

$$\left.\frac{d\Omega(\alpha)}{d\alpha}\right|_{\alpha=1} = \Omega(1) = \Omega, \qquad \left.\frac{d^2\Omega(\alpha)}{d\alpha^2}\right|_{\alpha=1} = 2I_n, \tag{1.116}$$

we obtain

$$\left.\frac{\partial \log f_n(s(\boldsymbol{w})|\alpha)}{\partial\alpha}\right|_{\alpha=1} = 0,$$
$$\left.\frac{\partial^2 \log f_n(s(\boldsymbol{w})|\alpha)}{\partial\alpha^2}\right|_{\alpha=1} = c_1 + \frac{n-1}{2}\frac{\boldsymbol{y}'H(H'\Omega H)^{-2}H'\boldsymbol{y}}{\boldsymbol{y}'H(H'\Omega H)^{-1}H'\boldsymbol{y}},$$

where c_1 is a constant. Then we can conduct the *unbiased locally best invariant* (LBIU) *test* (Ferguson, 1967, p.237), which rejects H_0 when

$$\left.\frac{\partial^2 \log f_n(s(\boldsymbol{w})|\alpha)}{\partial\alpha^2}\right|_{\alpha=1} > c_2 \quad \Leftrightarrow \quad R_{n2} = \frac{\boldsymbol{y}'H(H'\Omega H)^{-2}H'\boldsymbol{y}}{\boldsymbol{y}'H(H'\Omega H)^{-1}H'\boldsymbol{y}} > c_3, \tag{1.117}$$

where c_2 and c_3 are constants. We can rewrite the LBIU statistic R_{n2} so that it does not depend on H. For this purpose we have (Rao, 1973, p.77)

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Lemma 1.12 *Let $R = (P \ \ Q)$ be an $n \times n$ orthogonal matrix, where P and Q are $n \times (n-p)$ and $n \times p$ matrices, respectively. Then, for any $n \times n$ nonsingular symmetric matrix A, it holds that*

$$P(P'AP)^{-1}P' = A^{-1}N = N'A^{-1} = N'A^{-1}N, \tag{1.118}$$
$$P(P'AP)^{-2}P' = N'A^{-2}N, \tag{1.119}$$

where $N = I_n - Q(Q'A^{-1}Q)^{-1}Q'A^{-1}$.

Proof Let us consider

$$G = \begin{pmatrix} P'A \\ Q' \end{pmatrix}, \quad GR = \begin{pmatrix} P'A \\ Q' \end{pmatrix}(P \quad Q) = \begin{pmatrix} P'AP & P'AQ \\ 0 & I_p \end{pmatrix}.$$

It follows from the assumption that the $n \times n$ matrix G is nonsingular. Then, noting that $P'Q = 0$ and $Q'P = 0$, the first equality in (1.118) is established because

$$\begin{aligned} &G\left[P(P'AP)^{-1}P' - \left(A^{-1} - A^{-1}Q(Q'A^{-1}Q)^{-1}Q'A^{-1}\right)\right] \\ &\quad = G\left[P(P'AP)^{-1}P' - A^{-1}N\right] = 0. \end{aligned}$$

It can be checked that $A^{-1}N = N'A^{-1} = N'A^{-1}N$, which proves (1.118). Since

$$P(P'AP)^{-2}P' = P(P'AP)^{-1}P'P(P'AP)^{-1}P' = N'A^{-1}A^{-1}N = N'A^{-2}N,$$

the relation (1.119) is also established. □

Note that the above lemma holds if P and Q are replaced by PF_1 and QF_2, respectively, where F_1 and F_2 are any nonsigular matrices of size $(n-p) \times (n-p)$ and $p \times p$, respectively. Using this fact, the LBIU statistic R_{n2} in (1.117) can be expressed (Exercise 1.5.6) as

$$R_{n2} = \frac{\boldsymbol{y}'N'\Omega^{-2}N\boldsymbol{y}}{\boldsymbol{y}'N'\Omega^{-1}N\boldsymbol{y}}, \tag{1.120}$$

where $N = I_n - \boldsymbol{e}(\boldsymbol{e}'\Omega^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}'\Omega^{-1}$, $N'\Omega^{-1}N = \Omega^{-1}N = N'\Omega^{-1}$, and

$$\begin{aligned} \Omega^{-1} &= \left[(CC')^{-1} + \boldsymbol{e}_n\boldsymbol{e}_n'\right]^{-1} = CC' - \frac{1}{n+1}C\boldsymbol{e}\boldsymbol{e}'C' \\ &= \left[(C'C)^{-1} + \boldsymbol{e}_1\boldsymbol{e}_1'\right]^{-1} = C'C - \frac{1}{n+1}C'\boldsymbol{e}\boldsymbol{e}'C. \end{aligned} \tag{1.121}$$

Here $\boldsymbol{e}_n = (0, \ldots, 0, 1)' : n \times 1$, $\boldsymbol{e}_1 = (1, 0, \ldots, 0)' : n \times 1$, and C is the random walk-generating matrix defined in (1.93). Note that the (j, k)th element of Ω^{-1} is $\min(j,k) - jk/(n+1)$, and the eigenvalues of Ω^{-1} are given (Anderson, 1971, theorem 6.5.5) by

$$\lambda_j = \frac{1}{4}\left(\sin\frac{j\pi}{2(n+1)}\right)^{-2} \quad (j = 1, \ldots, n).$$

The present LBIU test can also be derived on the basis of the log-likelihood $L(\boldsymbol{\delta})$ defined in (1.113). Using the relations in (1.116) and $\mathrm{tr}(\Omega^{-1}) = n(n+2)/6$, we obtain

$$\left.\frac{\partial L(\boldsymbol{\delta})}{\partial \alpha}\right|_{H_0} = 0, \quad \left.\frac{\partial^2 L(\boldsymbol{\delta})}{\partial \alpha^2}\right|_{H_0} = -\frac{n(n+5)}{6} + n\frac{\boldsymbol{y}'N'\Omega^{-2}N\boldsymbol{y}}{\boldsymbol{y}'N'\Omega^{-1}N\boldsymbol{y}}, \tag{1.122}$$

where the derivatives are evaluated at the MLE of $\boldsymbol{\delta}$ under H_0. Then it is seen that the resulting statistic is the same as the LBIU statistic R_{n2} in (1.117) and (1.120). It is also recognized that $\alpha = 1$ gives a local maximum of $L(\boldsymbol{\delta})$ when $\partial^2 L(\boldsymbol{\delta})/\partial\alpha^2|_{H_0} < 0$, which will be discussed in Chapter 6.

Let us consider the asymptotic distribution of R_{n2} as $n \to \infty$ under H_0. We can assume that $\sigma_\varepsilon^2 = 1$. Putting $\boldsymbol{\zeta} \sim \mathrm{N}(\mathbf{0},\ I_n)$, we have

$$\begin{aligned}
\frac{1}{n}\,\boldsymbol{y}'N'\Omega^{-1}N\boldsymbol{y} &\stackrel{\mathcal{D}}{=} \frac{1}{n}\,\boldsymbol{\zeta}'\Omega^{1/2}N'\Omega^{-1}N\Omega^{1/2}\boldsymbol{\zeta} \stackrel{\mathcal{D}}{=} \frac{1}{n}\,\boldsymbol{\zeta}'\Omega^{-1/2}N\Omega N'\Omega^{-1/2}\boldsymbol{\zeta}\\
&= \frac{1}{n}\,\boldsymbol{\zeta}'\tilde{M}\boldsymbol{\zeta} \;\to\; 1 \quad \text{in probability},\\
\frac{1}{n^2}\,\boldsymbol{y}'N'\Omega^{-2}N\boldsymbol{y} &\stackrel{\mathcal{D}}{=} \frac{1}{n^2}\,\boldsymbol{\zeta}'\Omega^{1/2}N'\Omega^{-2}N\Omega^{1/2}\boldsymbol{\zeta} \stackrel{\mathcal{D}}{=} \frac{1}{n^2}\,\boldsymbol{\zeta}'\Omega^{-1}N\Omega N'\Omega^{-1}\boldsymbol{\zeta}\\
&= \frac{1}{n}\,\boldsymbol{\zeta}'\,B_{n2}\,\boldsymbol{\zeta},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{M} &= I_n - \Omega^{-1/2}\boldsymbol{e}(\boldsymbol{e}'\Omega^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}'\Omega^{-1/2} = \tilde{M}' = \tilde{M}^2,\\
B_{n2} &= \frac{1}{n}\,\Omega^{-1}N = \frac{1}{n}\left(\Omega^{-1} - \Omega^{-1}\boldsymbol{e}(\boldsymbol{e}'\Omega^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}'\Omega^{-1}\right).
\end{aligned}$$

Here $\mathrm{tr}(\tilde{M}) = n - 1$ and Ω^{-1} is given in (1.121), whereas

$$\Omega^{-1}\boldsymbol{e} = C'\boldsymbol{d} - \frac{n}{2}C'\boldsymbol{e}, \quad \boldsymbol{e}'\Omega^{-1}\boldsymbol{e} = \frac{n(n+1)(n+2)}{12},$$

where $\boldsymbol{d} = (1, 2, \ldots, n)'$. Then we have

$$B_{n2}(j,k) = \frac{1}{n}\left(\min(j,k) - \frac{jk}{n+1} - \frac{3jk(n-j+1)(n-k+1)}{n(n+1)(n+2)}\right),$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

which has the uniform limit

$$K_2(s,t) = \min(s,t) - st - 3st(1-s)(1-t). \tag{1.123}$$

It follows from Theorem 1.11 that

$$S_{n2} = \frac{1}{n} R_{n2} = \frac{1}{n} \frac{\boldsymbol{y}' N' \Omega^{-2} N \boldsymbol{y}}{\boldsymbol{y}' N' \Omega^{-1} N \boldsymbol{y}} \Rightarrow S_2 = \int_0^1 \int_0^1 K_2(s,t)\, dW(s)\, dW(t). \tag{1.124}$$

It is interesting to note that the function $K_2(s,t)$ in (1.123) was earlier presented in connection with goodness-of-fit tests in Section 1.5.1, and its FD is given in (1.79) together with the c.f. of S_2 in (1.80).

In this section we have dealt with the MA(1) model with unknown mean. If the mean is known and assumed to be 0, the analysis is easier (Exercise 1.5.7).

In Chapter 6 we shall deal with more general models and derive the asymptotic distributions of the LBIU statistics under the local alternative

$$H_1(c,\kappa) : \ \alpha = 1 - \frac{c}{n^\kappa},$$

where c and κ are positive constants. We shall also deal with the conditional model where the initial value ε_0 of the error process $\{\varepsilon_j\}$ is set at 0. The LBI test is derived for that model, and the power performance of the LBI test is compared with that of the LBIU test discussed in this section.

1.5.4 LBI and LBIU Tests for the Autoregressive Unit Root

Let us consider the following model:

$$y_j = \boldsymbol{\beta}' \boldsymbol{x}_j + u_j, \quad u_j = \rho u_{j-1} + \varepsilon_j \qquad (j = 1, \ldots, n), \tag{1.125}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{x}_j$ are $p \times 1$ coefficient and regression vectors, respectively, and $\{u_j\}$ is an error term that follows the first-order autoregressive (AR(1)) process with $\{\varepsilon_j\} \sim$ i.i.d.$(0, \sigma_\varepsilon^2)$. The AR(1) coefficient parameter ρ is assumed to take any value, and the initial value u_0 of $\{u_j\}$ is assumed to be $u_0 = 0$.

Our concern here is to conduct the test

$$H_0 : \ \rho = 1 \quad \text{versus} \quad H_1 : \ \rho < 1 \quad \text{or} \quad H_1 : \ \rho \neq 1. \tag{1.126}$$

Dickey and Fuller (1979) first suggested the unit root test using the following three models:

$$y_j = \begin{cases} \rho\, y_{j-1} + \varepsilon_j, \\ \\ \rho\, y_{j-1} + \alpha + \varepsilon_j, \\ \\ \rho\, y_{j-1} + \beta + \gamma\, j, \end{cases} \tag{1.127}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where it is assumed that ρ is close to unity and the true value of (α, γ) is $(0, 0)$, while β is arbitrary. This seemingly curious restriction may be understood from the model in (1.125). In fact, when $\boldsymbol{x}_j = (1, j)'$ in (1.125), we have

$$y_j = \beta_1 + \beta_2\, j + u_j = \rho\, y_{j-1} + \beta_1(1-\rho) + \beta_2\rho + \beta_2(1-\rho)\, j + \varepsilon_j,$$

and it is recognized that the above restriction is imposed on this model when $\rho = 1$.

The unit root test applied to the model (1.125) is based on $n(\hat{\rho} - 1)$, where $\hat{\rho}$ is the LSE of ρ computed from

$$\hat{\rho} = \frac{\sum_{j=2}^{n} \hat{u}_{j-1}\hat{u}_j}{\sum_{j=2}^{n} \hat{u}_{j-1}^2}, \quad \hat{u}_j = y_j - \hat{\boldsymbol{\beta}}'\boldsymbol{x}_j. \tag{1.128}$$

Here $\hat{\boldsymbol{\beta}}$ is the LSE of $\boldsymbol{\beta}$ obtained from the regression of $\{y_j\}$ on $\{\boldsymbol{x}_j\}$, where we consider three cases for $\boldsymbol{x}_j$:

$$\text{Case 1: } \boldsymbol{x}_j = 0, \quad \text{Case 2: } \boldsymbol{x}_j = 1, \quad \text{Case 3: } \boldsymbol{x}_j = (1,\ j)'.$$

Note that $\hat{\rho}$ computed for the three cases is asymptotically equivalent to the LSE computed from the three models in (1.127). Then it can be shown (Phillips, 1987a; Nabeya and Tanaka, 1990a) that, as $n \to \infty$ under H_0,

$$n(\hat{\rho} - 1) \ \Rightarrow \ \frac{\displaystyle\int_0^1 G(t)\, dW(t)}{\displaystyle\int_0^1 G^2(t)\, dt}, \tag{1.129}$$

where

$$G(t) = \begin{cases} W(t), & \text{(Case 1)} \\[2ex] W(t) - \displaystyle\int_0^1 W(s)\, ds, & \text{(Case 2)} \\[2ex] W(t) + (6t-4)\displaystyle\int_0^1 W(s)\, ds - (12t-6)\displaystyle\int_0^1 sW(s)\, ds. & \text{(Case 3)} \end{cases}$$

We already presented an expression for the limit in distribution of $n(\hat{\rho} - 1)$ for Case 1, which is the unit root distribution given in (1.52). Note that $G(t)$ in Case 2 is referred to as *demeaned Bm*, which comes from

$$\frac{1}{\sqrt{n}\sigma_\varepsilon}\hat{u}_j = \frac{1}{\sqrt{n}\sigma_\varepsilon}\left(u_j - \frac{1}{n}\sum_{k=1}^{n} u_k\right) \quad \Leftrightarrow \quad W(t) - \int_0^1 W(s)\, ds,$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

whereas $G(t)$ in Case 3 is *detrended Bm*, which comes from (Exercise 1.5.8)

$$\begin{aligned}\frac{1}{\sqrt{n}\sigma_\varepsilon}\,\hat{u}_j &= \frac{1}{\sqrt{n}\sigma_\varepsilon}\,(u_j - \hat{\beta}_1 - \hat{\beta}_2\, j)\\ &\Leftrightarrow W(t) + (6t-4)\int_0^1 W(s)\,ds - (12t-6)\int_0^1 sW(s)\,ds.\end{aligned}$$

The limiting distributions of $n(\hat{\rho}-1)$ under H_0 can be expressed as

$$F(x) = \lim_{n\to\infty} P\Big(n(\hat{\rho}-1) \le x\Big) = P(H(x) \ge 0),$$

where

$$H(x) = x\int_0^1 G^2(W)\,dt - \int_0^1 G(t)\,dW(t).$$

The distribution function $F(x)$ can be computed by deriving the FD associated with $H(x)$. The limiting local powers of the unit root tests can also be computed by deriving the associated FD, which was discussed in Nabeya and Tanaka (1990a, 1990b). This topic will be discussed in Chapter 7.

In subsequent discussions, as a continuation of the last two sections, we consider the LBI or LBIU type tests, assuming that $\{\varepsilon_j\} \sim \mathrm{NID}(0, \sigma_\varepsilon^2)$. The model (1.125) can be expressed as

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u}, \qquad \boldsymbol{u} = C(\rho)\boldsymbol{\varepsilon}, \tag{1.130}$$

where

$$\boldsymbol{y} = \begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix},\quad X = \begin{pmatrix} \boldsymbol{x}_1'\\ \boldsymbol{x}_2'\\ \vdots\\ \boldsymbol{x}_n'\end{pmatrix},\quad \boldsymbol{u} = \begin{pmatrix} u_1\\ u_2\\ \vdots\\ u_n\end{pmatrix},\quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{pmatrix},$$

whereas

$$C(\rho) = \begin{pmatrix} 1 & & & & & \\ \rho & 1 & & & 0 & \\ \cdot & & \cdot & \cdot & & \\ \cdot & & & \cdot & \cdot & \\ \cdot & & & & \cdot & \cdot \\ \rho^{n-1} & \cdot & \cdot & \cdot & \rho & 1 \end{pmatrix}. \tag{1.131}$$

Here $C(1) = C$ is the random walk-generating matrix defined in (1.93).

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Putting $\Phi(\rho) = C(\rho)C'(\rho)$, $\Phi = \Phi(1) = CC'$, and noting (Exercise 1.5.9) that

$$\left.\frac{d\Phi(\rho)}{d\rho}\right|_{\rho=1} = \boldsymbol{d}\boldsymbol{d}' - \Phi, \quad \boldsymbol{d} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix}, \tag{1.132}$$

it follows from (1.97) that the LBI test rejects H_0 when

$$R_{n3} = \frac{\boldsymbol{y}'H\left(H'\Phi H\right)^{-1}H'\boldsymbol{d}\boldsymbol{d}'H\left(H'\Phi H\right)^{-1}H'\boldsymbol{y}}{\boldsymbol{y}'H\left(H'\Phi H\right)^{-1}H'\boldsymbol{y}} \tag{1.133}$$

takes small values, where H is an $n \times (n-p)$ matrix such that $H'H = I_{n-p}$, and $HH' = I_n - X(X'X)^{-1}X' = M$. It follows from Lemma 1.12 that

$$H(H'\Phi H)^{-1}H' = \Phi^{-1} - \Phi^{-1}X(X'\Phi^{-1}X)^{-1}X'\Phi^{-1} = \tilde{M},$$

which gives

$$R_{n3} = \frac{\boldsymbol{y}'\tilde{M}\boldsymbol{d}\boldsymbol{d}'\tilde{M}\boldsymbol{y}}{\boldsymbol{y}'\tilde{M}\boldsymbol{y}}. \tag{1.134}$$

It is recognized that R_{n3} in (1.134) reduces to 0 if $\tilde{M}\boldsymbol{d} = \boldsymbol{0}$. This occurs if $\boldsymbol{d}$ belongs to the column space of X because it holds that $\tilde{M}X = 0$. Avoiding this case, let us consider the case where $X = (1, \ldots, 1)' = \boldsymbol{e}$. Then we have

$$\boldsymbol{y}'\tilde{M}\boldsymbol{y} = \sum_{j=2}^{n}(y_j - y_{j-1})^2 = \sum_{j=2}^{n}(u_j - u_{j-1})^2,$$
$$\boldsymbol{y}'\tilde{M}\boldsymbol{d} = y_n - y_1 = u_n - u_1,$$

so that

$$R_{n3} = \frac{\boldsymbol{y}'\tilde{M}\boldsymbol{d}\boldsymbol{d}'\tilde{M}\boldsymbol{y}}{\boldsymbol{y}'\tilde{M}\boldsymbol{y}} = \frac{(u_n - u_1)^2}{\sum_{j=2}^{n}(u_j - u_{j-1})^2} = \frac{U_n}{V_n},$$

where $U_n = (u_n - u_1)^2$ and $V_n = \sum_{j=2}^{n}(u_j - u_{j-1})^2$. It follows that, as $n \to \infty$ under H_0, $R_{n3} \Rightarrow \chi^2(1)$. The power performance of the present LBI test is found to be quite poor and is not recommended; this will be clarified in Chapter 7.

We next consider the LBIU test for the case where $\boldsymbol{d} = (1, 2, \ldots, n)'$ belongs to the column space of X. To derive the test, it is more convenient

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

to use the log-likelihood for $\boldsymbol{y} \sim \mathrm{N}(X\boldsymbol{\beta}, \sigma_\varepsilon^2 \Phi(\rho))$. Noting that $|\Phi(\rho)| = 1$, the log-likelihood is given by

$$L(\boldsymbol{\delta}) = -\frac{n}{2} \log 2\pi\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}(\boldsymbol{y} - X\boldsymbol{\beta})'\Phi^{-1}(\rho)(\boldsymbol{y} - X\boldsymbol{\beta}), \tag{1.135}$$

where $\boldsymbol{\delta} = (\rho, \boldsymbol{\beta}', \sigma_\varepsilon^2)'$. Then the LBIU test rejects H_0 when

$$\begin{aligned} \left.\frac{\partial^2 L(\boldsymbol{\delta})}{\partial\rho^2}\right|_{H_0} &= -\frac{1}{2\sigma_\varepsilon^2}(\boldsymbol{y} - X\boldsymbol{\beta})'\frac{d^2\Phi^{-1}(\rho)}{d^2\rho}(\boldsymbol{y} - X\boldsymbol{\beta})\Big|_{H_0} \\ &= -\frac{1}{2\tilde{\sigma}_\varepsilon^2}\boldsymbol{y}'N'\left.\frac{d^2\Phi^{-1}(\rho)}{d^2\rho}\right|_{\rho=1} N\boldsymbol{y} \end{aligned}$$

becomes large, where $N = I_n - X(X'\Phi^{-1}X)^{-1}X'\Phi^{-1}$ and $\tilde{\sigma}_\varepsilon^2 = \boldsymbol{y}'N'\Phi^{-1}$ $N\boldsymbol{y}/n$. We also obtain (Exercise 1.5.10)

$$\left.\frac{d^2\Phi^{-1}(\rho)}{d\rho^2}\right|_{\rho=1} = 2(I_n - \boldsymbol{e}_n\boldsymbol{e}_n'), \quad \boldsymbol{e}_n = (0, \ldots, 0, 1)' : n \times 1,$$

which yields

$$\left.\frac{\partial^2 L(\boldsymbol{\delta})}{\partial\rho^2}\right|_{H_0} = -n\frac{\boldsymbol{y}'N'(I_n - \boldsymbol{e}_n\boldsymbol{e}_n')N\boldsymbol{y}}{\boldsymbol{y}'N'\Phi^{-1}N\boldsymbol{y}}. \tag{1.136}$$

Since it holds that $N'\Phi^{-1}X = 0$, $\boldsymbol{y}'N'\boldsymbol{e}_n$ reduces to 0 if $\boldsymbol{e}_n$ belongs to the column space of $\Phi^{-1}X$ or if $\Phi\boldsymbol{e}_n = CC'\boldsymbol{e}_n = \boldsymbol{d}$ belongs to the column space of X. In that case we have the LBIU statistic

$$R_{n4} = \frac{\boldsymbol{y}'N'N\boldsymbol{y}}{\boldsymbol{y}'N'\Phi^{-1}N\boldsymbol{y}} = \frac{\boldsymbol{y}'N'N\boldsymbol{y}}{\boldsymbol{y}'\tilde{M}\boldsymbol{y}}, \tag{1.137}$$

and H_0 is rejected when R_{n4} takes small values.

Consider $X = (\boldsymbol{e}, \boldsymbol{d})$, which yields (Exercise 1.5.11)

$$\begin{aligned} \boldsymbol{y}'N'N\boldsymbol{y} &= \sum_{j=2}^{n}\left(y_j - y_1 - \frac{j-1}{n-1}(y_n - y_1)\right)^2 \\ &= \sum_{j=2}^{n}\left(u_j - u_1 - \frac{j-1}{n-1}(u_n - u_1)\right)^2, \\ \boldsymbol{y}'\tilde{M}\boldsymbol{y} &= \sum_{j=2}^{n}\left(y_j - y_{j-1} - \frac{y_n - y_1}{n-1}\right)^2 = \sum_{j=2}^{n}\left(u_j - u_{j-1} - \frac{u_n - u_1}{n-1}\right)^2, \end{aligned}$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

where it holds (Exercise 1.5.12) that, as $n \to \infty$ under H_0, $\mathbf{y}'\tilde{M}\mathbf{y}/(n\sigma_\varepsilon^2) \to 1$ in probability and

$$S_{n4} = \frac{1}{n}\, R_{n4} = \frac{\mathbf{y}'N'N\mathbf{y}/\left(n^2\sigma_\varepsilon^2\right)}{\mathbf{y}'\tilde{M}\mathbf{y}/\left(n\sigma_\varepsilon^2\right)} \;\Rightarrow\; S_4,$$

$$S_4 = \int_0^1 (W(t) - tW(1))^2\, dt \overset{\mathcal{D}}{=} \int_0^1\int_0^1 [\min(s,t) - st]\, dW(s)\, dW(t).$$

The c.f. of S_4 is $(D(2i\theta))^{-1/2}$, where $D(\lambda) = \sin\sqrt{\lambda}/\sqrt{\lambda}$.

The performance of the LBIU test is found to be quite good, unlike the LBI test; this will be clarified in Chapter 7.

Exercises for Section 1.5

Exercise 1.5.1 Show that the modified Cramér–von Mises–Smirnov statistic $W_n^2(\hat{\xi})$ defined in (1.75) can be rewritten as

$$W_n^2(\hat{\xi}) = \sum_{j=1}^n \left(F(X_j';\hat{\xi}) - \frac{2j-1}{2n}\right)^2 + \frac{1}{12n},$$

where $X_1' \le X_2' \le \cdots \le X_n'$ are ordered observations of $X_1, X_2, \ldots, X_n$.

Exercise 1.5.2 Let $X_1, \ldots, X_n$ be a random sample from the continuous distribution $F(x)$. Prove that

$$\mathrm{E}(H_n(t)) = t, \qquad \mathrm{Cov}(H_n(s), H_n(t)) = \frac{1}{n}\,(\min(s,t) - st),$$

where $H_n(t)$ is the empirical distribution function defined by

$$H_n(t) = \frac{1}{n}\sum_{j=1}^n \epsilon\big(t - F(X_j)\big), \quad \epsilon(x) = \begin{cases} 1 & (x \ge 0), \\ 0 & (x < 0). \end{cases}$$

Exercise 1.5.3 Let $F(x;\xi)$ be the continuous distribution function of X with a parameter ξ, and $f(x;\xi)$ be its probability density. Put

$$\sigma^{-2}(\xi) = \int_{-\infty}^{\infty} \left(\frac{\partial \log f(x;\xi)}{\partial\xi}\right)^2 f(x;\xi)\, dx,$$

$$g_\xi(t) = \left.\frac{\partial F(x;\xi)}{\partial \xi}\right|_{x=x(t;\xi)}, \quad x(t;\xi) = \inf\{x : F(x;\xi) = t\}.$$

Then prove that, when $F(x;\xi) = R(x-\xi)$ and $r(x) = R'(x)$,

$$\sigma(\xi)\, g_\xi(t) = -\left(\int_{-\infty}^{\infty} \frac{(r'(x))^2}{r(x)}\, dx\right)^{-1/2} r(R^{-1}(t)).$$

Prove also that, when $F(x;\xi) = R(x/\xi)$,

$$\sigma(\xi)\, g_\xi(t) = -\left(\int_{-\infty}^{\infty} \frac{(x\, r'(x))^2}{r(x)}\, dx - 1\right)^{-1/2} R^{-1}(t)\, r(R^{-1}(t)).$$

Exercise 1.5.4 For the distribution function $F(x;\xi) = (R(x))^\xi$ $(\xi > 0)$, prove that

$$\sigma(\xi)\, g_\xi(t) = t\, \log t,$$

where $\sigma(\xi)$ and $g_\xi(t)$ are as defined in Exercise 1.5.3.

Exercise 1.5.5 Consider the model

$$y_j = \beta_j + \varepsilon_j, \quad \beta_j = \beta_{j-1} + \eta_j \quad (j = 1, \ldots, n),$$

where $\beta_0 = 0$, $\{\varepsilon_j\} \sim \mathrm{NID}(0, \sigma_\varepsilon^2)$ and $\{\eta_j\} \sim \mathrm{NID}(0, \sigma_\eta^2)$, with $\{\varepsilon_j\}$ and $\{\eta_j\}$ being independent of each other. Putting $\rho = \sigma_\eta^2/\sigma_\varepsilon^2$, show that $\partial L(\rho, \sigma_\varepsilon^2)/\partial\rho$ does not tend to normality when $\rho = 0$, where $L(\rho, \sigma_\varepsilon^2)$ is the log-likelihood for $\boldsymbol{y} = (y_1, \ldots, y_n)'$.

Exercise 1.5.6 Prove that

$$H(H'\Omega H)^{-1}H' = N'\Omega^{-1}N, \quad H(H'\Omega H)^{-2}H' = N'\Omega^{-2}N,$$

where Ω is any $n \times n$ nonsingular symmetric matrix, H is an $n \times (n-p)$ matrix with rank $n - p$, $N = I_n - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}$, and X is an $n \times p$ matrix with rank p such that $H'X = 0$.

Exercise 1.5.7 Consider the testing problem $H_0 : \alpha = 1$ for the MA(1) model

$$y_j = \varepsilon_j - \alpha\varepsilon_{j-1} \quad (j = 1, \ldots, n),$$

where $\{\varepsilon_j\} \sim \mathrm{NID}(0, \sigma_\varepsilon^2)$. Derive the LBIU test statistic and its asymptotic distribution under H_0.

Exercise 1.5.8 Consider the model

$$y_j = \beta_1 + \beta_2\, j + u_j, \quad u_j = u_{j-1} + \varepsilon_j \qquad (j = 1, \ldots, n),$$

where $\{\varepsilon_j\} \sim \mathrm{i.i.d.}(0, \sigma_\varepsilon^2)$. Denote by $\hat{u}_j$ the residual obtained from the least squares regression of $\{y_j\}$ on $(1, j)$ $(j = 1, \ldots, n)$. Prove that

$$\frac{1}{n^2\sigma_\varepsilon^2} \sum_{j=1}^{n} \hat{u}_j^2 \Rightarrow \int_0^1 \Bigg(W(t) + (6t - 4)\int_0^1 W(s)\, ds - (12t - 6)\int_0^1 sW(s)\, ds\Bigg)^2 dt.$$

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press

Exercise 1.5.9 Define the $n \times n$ matrix $C(\rho)$ by (1.131), where its (j,k)th element is ρ^{j-k} for $j \geq k$ and 0 for $j < k$. Show that

$$\left.\frac{dC(\rho)C'(\rho)}{d\rho}\right|_{\rho=1} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} (1,\, 2,\, \ldots,\, n) - C(1)C'(1).$$

Exercise 1.5.10 For the matrix $C(\rho)$ given in Exercise 1.5.9, show that

$$\left.\frac{d^2\left\{\left(C(\rho)C'(\rho)\right)^{-1}\right\}}{d\rho^2}\right|_{\rho=1} = 2\Big(I_n - \boldsymbol{e}_n\boldsymbol{e}_n'\Big), \quad \boldsymbol{e}_n = (0,\, \ldots,\, 0, 1)' : n \times 1.$$

Exercise 1.5.11 Consider the model

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u}, \quad \boldsymbol{u} = C(1)\boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathrm{N}\big(\boldsymbol{0}, \sigma_\varepsilon^2 I_n\big),$$

where $X = (\boldsymbol{e}, \boldsymbol{d})$, $\boldsymbol{e} = (1,\, \ldots,\, 1)' : n \times 1$ and $\boldsymbol{d} = (1, 2,\, \ldots,\, n)'$, whereas $C(\rho)$ is as given in Exercise 1.5.9 with $C = C(1)$. Then show that

$$\boldsymbol{y}'N'N\boldsymbol{y} = \sum_{j=2}^{n} \left(u_j - u_1 - \frac{j-1}{n-1}(u_n - u_1)\right)^2,$$

$$\boldsymbol{y}'\tilde{M}\boldsymbol{y} = \sum_{j=2}^{n} \left(u_j - u_{j-1} - \frac{u_n - u_1}{n-1}\right)^2,$$

where $N = I_n - X(X'(CC')^{-1}X)^{-1}X'(CC')^{-1}$ and $\tilde{M} = N'(CC')^{-1}N$.

Exercise 1.5.12 Prove that

$$\frac{1}{n^2}\sum_{j=2}^{n} \left(u_j - u_1 - \frac{j-1}{n-1}(u_n - u_1)\right)^2 \Rightarrow \int_0^1 (W(t) - tW(1))^2\, dt,$$

where $u_j = u_{j-1} + \varepsilon_j$, $u_0 = 0$ and $\{\varepsilon_j\} \sim \mathrm{NID}(0, 1)$.

https://doi.org/10.1017/9781009567008.003 Published online by Cambridge University Press