8 Riemann Integrals

8.1 Summary

In this chapter we deal with stochastic Riemann integrals, i.e. with ordinary Riemann integrals with a stochastic process as the integrand.[1] Mathematically, these constructs are relatively unsophisticated, they can be defined pathwise for continuous functions as in conventional (deterministic) calculus. However, this pathwise definition will not be possible any longer for e.g. Ito integrals in the chapter after next. Hence, at this point we propose a way of defining integrals as a limit (in mean square) which will be useful later on. If the stochastic integrand is in particular a Wiener process, then the Riemann integral follows a Gaussian distribution with zero expectation and the familiar formula for the variance. A number of examples will facilitate the understanding of this chapter.

8.2 Definition and Fubini's Theorem

As one has done in deterministic calculus, we will define the Riemann integral by an adequate partition as the limit of a sum.

Partition

In order to define an integral of a function from 0 to t, we decompose the interval into n adjacent, non-overlapping subintervals which are allowed to intersect at the

[1] Bernhard Riemann (1826–1866) studied with Gauss in Göttingen where he himself became a professor. Already before his day, integration had been used as a technique which reverses differentiation by forming an antiderivative. However, Riemann explained for the first time under which conditions a function possesses an antiderivative at all.

U. Hassler, *Stochastic Processes and Calculus*, Springer Texts in Business and Economics, DOI 10.1007/978-3-319-23428-1_8

endpoints:

$$P_n([0, t]): \quad 0 = s_0 < s_1 < \ldots < s_n = t\,. \tag{8.1}$$

In the following, we always assume that the **partition** $P_n([0, t])$ becomes increasingly fine with n growing ("adequate partition"):

$$\max_{1 \le i \le n} (s_i - s_{i-1}) \;\to\; 0 \quad \text{for } n \to \infty\,. \tag{8.2}$$

By s_i^* we denote an arbitrary point in the i-th interval,

$$s_i^* \in [s_{i-1}, s_i]\,, \quad i = 1, \ldots, n\,.$$

Occasionally, we will sum up the lengths of the subintervals. Obviously, it holds that

$$\sum_{i=1}^{n} (s_i - s_{i-1}) = s_n - s_0 = t\,.$$

In general, for a function φ one obtains:

$$\sum_{i=1}^{n} (\varphi(s_i) - \varphi(s_{i-1})) = \varphi(t) - \varphi(0)\,. \tag{8.3}$$

Sometimes, we will operate with the example of the equidistant partition. It is given by $s_i = it/n$:

$$0 = s_0 < s_1 = \frac{t}{n} < \ldots < s_{n-1} = \frac{n-1}{n} t < s_n = t\,.$$

Due to $s_i - s_{i-1} = 1/n$ the required refinement from (8.2) for $n \to \infty$ is guaranteed.

Definition and Existence

Now, the product of a deterministic function f and a stochastic process X is to be integrated. To this end, the **Riemann sum** is defined by means of the notation introduced:

$$R_n = \sum_{i=1}^{n} f(s_i^*)\, X(s_i^*)\, (s_i - s_{i-1}). \tag{8.4}$$

Here, we have a sum of rectangular areas, each with a width of $(s_i - s_{i-1})$ and the height $f(s_i^*)\, X(s_i^*)$. With n growing, the area beneath $f(s)\, X(s)$ on $[0, t]$ is to be approximated all the better. If the limit of this sum for $n \to \infty$ exists uniquely and independently of the partition and of the choice of s_i^*, then it is defined as a (stochastic) Riemann integral. In this case, the convergence occurs in mean square ("$\xrightarrow{2}$")[2]:

$$R_n = \sum_{i=1}^{n} f(s_i^*)\, X(s_i^*)\, (s_i - s_{i-1}) \;\xrightarrow{2}\; \int_0^t f(s)\, X(s)\, ds\,.$$

One then says that the **Riemann integral** exists. For this existence, there is a sufficient and a necessary condition formulated in the following proposition. The proof is carried out with part (b) from Lemma 8.2 below, see Problem 8.1. Further elaborations on mean square convergence can be found at the end of the chapter.

Proposition 8.1 (Existence of the Riemann Integral) *The Riemann sum from Eq. (8.4) converges in mean square for $n \to \infty$ under (8.2) if and only if the double integral*

$$\int_0^t \int_0^t f(s) f(r)\, E\,(X(s)\, X(r))\; drds$$

exists.

A sufficient condition for the existence of the Riemann integral is that the function f is continuous and that furthermore $\mathrm{E}\,(X(s)\, X(r))$ is continuous in both arguments. In order to find this, we define

$$\varphi(s) := f(s) \int_0^t f(r)\, \mathrm{E}\,(X(s)\, X(r))\; dr\,.$$

Now, if the function $\mathrm{E}\,(X(s)\, X(r))$ is continuous in both arguments, this implies continuity of φ for a continuous f as the integral is a continuous functional, see e.g. Trench (2013, p. 462). Therefore, the ordinary Riemann integral of φ exists,

$$\int_0^t \varphi(s)\, ds = \int_0^t \int_0^t f(s) f(r)\, \mathrm{E}\,(X(s)\, X(r))\; drds\,.$$

Hence, the Riemann sum from (8.4) converges due to Proposition 8.1.

[2] A definition and discussion of this mode of convergence can be found in the fourth section of this chapter.

Fubini's Theorem

Frequently, we are interested in the average behavior, i.e. the expected value of Riemann integrals. The expected value, however, is defined as an integral itself such that one is confronted with double integrals. For calculating these, there is a simple rule which is finally based on the fact that the order of integration does not matter for double integrals over continuous functions. In deterministic calculus, this fact is also known as "Fubini's theorem" (see Footnote 6 in Sect. 2.3). Adapted to our problem of the expected value of a Riemann integral, the corresponding circumstances are given in the following proposition, also cf. Billingsley (1986, Theorem 18.3).

Proposition 8.2 (Fubini's Theorem) *If $\int_0^t E(|X(s)|)\,ds$ exists, it holds for a continuous X that:*

$$E\left(\int_0^t X(s)\,ds\right) = \int_0^t E(X(s))\,ds\,.$$

The statement is easy to comprehend if one thinks of the integral as a finite Riemann sum. As is well known, in the discrete case, summation and expectation is interchangeable:

$$\mathrm{E}\left(\sum_{i=1}^{n} X(s_i^*)\,(s_i - s_{i-1})\right) = \sum_{i=1}^{n} \mathrm{E}(X(s_i^*))\,(s_i - s_{i-1})\,.$$

Now, Fubini's theorem just guarantees a continuation of this interchangeability for $n \to \infty$.

Example 8.1 (Expected Value of the Integrated WP) Consider the special case of the integrated WP with $X(s) = W(s)$ and $f(s) = 1$. With the WP being continuous, $|W(t)|$ is a continuous process as well. In (7.8) we have determined the following expression as the expected value:

$$\mathrm{E}(|W(t)|) = \sqrt{\frac{2t}{\pi}}.$$

Before applying Proposition 8.2, we check:

$$\begin{aligned}
\int_0^t \mathrm{E}\left(|W(s)|\right)\,ds &= \int_0^t \sqrt{\frac{2}{\pi}}\,s^{\frac{1}{2}}ds \\
&= \sqrt{\frac{2}{\pi}}\left[\frac{2}{3}\,s^{\frac{3}{2}}\right]_0^t \\
&= \sqrt{\frac{2}{\pi}}\,\frac{2}{3}\,t^{\frac{3}{2}}\,.
\end{aligned}$$

As this quantity is finite, the requirements of Fubini's theorem are fulfilled. Hence, it follows that

$$\mathrm{E}\left(\int_0^t W(s)\,ds\right) = \int_0^t \mathrm{E}(W(s))\,ds = 0\,. \quad \blacksquare$$

General Rules

Note that our definition of the integral seems to be unnecessarily restrictive. However, the restriction on the interval $[0, t]$ is by no means crucial. The usual rules apply and are here symbolically described for an integrand g (without proof):

$$\int_a^b g(x)\,dx = \int_a^c g(x)\,dx + \int_c^b g(x)\,dx \quad \text{for } a \leq c \leq b\,,$$

$$\int (g_1(x) + g_2(x))\,dx = \int g_1(x)\,dx + \int g_2(x)\,dx\,,$$

$$\int c\,g(x)\,dx = c\int g(x)\,dx \quad \text{for } c \in \mathbb{R}\,.$$

8.3 Riemann Integration of Wiener Processes

In this section, we concentrate on Riemann integrals where the stochastic part of the integrand is a WP: $X(t) = W(t)$.

Normal Distribution

Frequently, Gaussian random variables are hidden behind Riemann integrals. In fact, it holds that all of the integrals discussed in this section follow Gaussian distributions with zero expectation. The variances can be determined according to the following proposition (for a proof see Problem 8.3).

Proposition 8.3 (Normality of Riemann Integrals) *Let f be a continuous deterministic function on* $[0, t]$. *Then, it holds*

$$\int_0^t f(s)\,W(s)\,ds \;\sim\; \mathcal{N}\left(0\,,\int_0^t\int_0^t f(r)f(s)\min(r,s)drds\right)\,.$$

The normality follows from the fact that the WP is a Gaussian process. Hence, the Riemann sum R_n from (8.4) follows a Gaussian distribution for finite n. As R_n converges in mean square, it follows from Lemma 8.1 (see below) that the limit is Gaussian as well. Note that the finiteness of the variance expression from Proposition 8.3 is just sufficient and necessary for the existence of the Riemann integral (Proposition 8.1).

Example 8.2 (Variance of the Integrated WP) Consider an integrated WP with $f(s) = 1$ as in Example 8.1. We look for a closed expression for the variance of $\int_0^t W(s)ds$. Due to Proposition 8.3, the starting point is:

$$\mathrm{Var}\left(\int_0^t W(s)ds\right) = \int_0^t \int_0^t \min(r,s)\, drds\,.$$

Now, we employ a useful trick for many applications. The integral with respect to r is decomposed into the sum of two integrals with s as the integration limit such that the minimum function can be specified explicitly:

$$\begin{aligned}\int_0^t \int_0^t \min(r,s)\, drds &= \int_0^t \left[\int_0^s \min(r,s)\, dr + \int_s^t \min(r,s)\, dr\right] ds \\ &= \int_0^t \left[\int_0^s r\, dr + \int_s^t s\, dr\right] ds\,.\end{aligned}$$

Now, the integration of the power functions yields the requested variance:

$$\begin{aligned}\mathrm{Var}\left(\int_0^t W(s)ds\right) &= \int_0^t \left[\int_0^s r\, dr + \int_s^t s\, dr\right] ds \\ &= \int_0^t \left[\frac{s^2}{2} + s\,(t-s)\right] ds \\ &= \left[\frac{s^2\, t}{2} - \frac{s^3}{6}\right]_0^t \\ &= \frac{t^3}{3}\,. \quad \blacksquare\end{aligned}$$

Autocovariance Function

With the time-dependent integration limit, $\int_0^t f(s)\, W(s)\, ds$ itself is a stochastic process. Therefore, it suggests itself to not only determine the variance as in Proposition 8.3, but the covariance function as well. The general result is given in the following proposition, which will be verified in Problem 8.7.

Proposition 8.4 (Autocovariance Function) *For a continuous function f with integrable antiderivative F and $Y(t) = \int_0^t f(s)\, W(s)ds$ it holds that:*

$$E(Y(t)\, Y(t+h)) = \int_0^t f(s) \left[s\,F(s) - \int_0^s F(r)dr + s(F(t+h) - F(s)) \right] ds\,,$$

where $h \geq 0$.

Therefore, with $h = 0$ an alternative expression for the variance from Proposition 8.3 is obtained. For concrete functions f, the formula can be simplified considerably. This is to be shown by the following example.

Example 8.3 (Autocovariance of the Integrated WP) Once again, we examine the integrated WP with $f(s) = 1$ and $F(s) = s$ as in Examples 8.1 and 8.2. Then, plugging in yields:

$$\begin{aligned} \mathrm{E}\left(\int_0^t W(s)ds \int_0^{t+h} W(r)dr \right) &= \int_0^t \left[s^2 - \frac{1}{2}s^2 + s((t+h) - s) \right] ds \\ &= \int_0^t \left[s(t+h) - \frac{1}{2}s^2 \right] ds \\ &= \frac{t^2(t+h)}{2} - \frac{t^3}{6}. \end{aligned}$$

Hence, for $h = 0$ the variance of the integrated Wiener process reads:

$$\mathrm{Var}\left(\int_0^t W(s)ds \right) = \frac{t^3}{2} - \frac{t^3}{6} = \frac{t^3}{3}.$$

Of course, we already know this from Example 8.2. ■

Examples

For three special Gaussian integrals, which we will be confronted with over and over, the variances are to be calculated. We put the results in front.

Corollary 8.1 *It holds that*

$$
\begin{aligned}
&(a) \quad \int_0^1 W(s)\,ds \quad \sim \mathcal{N}(0,1/3)\,,\\
&(b) \; W(1) - \int_0^1 W(s)\,ds \sim \mathcal{N}(0,1/3)\,,\\
&(c) \; \int_0^1 (s-c)\,W(s)\,ds \;\sim\; \mathcal{N}(0,\sigma_R^2)\,,
\end{aligned}
$$

where $c \in \mathbb{R}$ *and* $\sigma_R^2 = \frac{8-25c+20c^2}{60} > 0$.

The normality in (a) and (c) is clear due to Proposition 8.3. In (b) we have the sum of two Gaussian random variables which does not necessarily have to be Gaussian again unless a multivariate Gaussian distribution is present. Thus, the normality of (b) can only be proven in connection with Stieltjes integrals (see Problem 9.2).

The result from (a) is a special case of Example 8.2 with $t = 1$. We show in Problem 8.4 that the variance in (b) is just $1/3$. The proof of (c) for $c = 0$ is given in Problem 8.5; for an arbitrary c, the proof is basically similar, however, it gets computationally more involved. Note that the variance σ_R^2 cannot be zero or negative for any c (Problem 8.6).

Again, there should be a word of warning concerning equality in distribution. From (b) it follows that:

$$\int_0^1 W(s)\,ds - W(1) \;\sim\; \mathcal{N}(0,1/3)\,.$$

Therefore, the following random variables are equal in distribution,

$$\int_0^1 W(s)\,ds - W(1) \;\sim\; \int_0^1 W(s)\,ds\,,$$

although, pathwise it obviously holds that:

$$\int_0^1 W(s)\,ds - W(1) \;\neq\; \int_0^1 W(s)\,ds\,.$$

8.4 Convergence in Mean Square

Now, we hand in some basics which are not necessary for the understanding of Riemann integrals; however, they are helpful for some technical properties. In particular, for the elaboration on the Ito integral following below, the knowledge of convergence in mean square is advantageous for a complete understanding. For a brief introduction to the basics of asymptotic theory, Pötscher and Prucha (2001) can be recommended.

Definition and Properties

Let $\{X_n\}$, $n \in \mathbb{N}$, be a sequence of real random variables with

$$\mathrm{E}(X_n^2) < \infty . \tag{8.5}$$

For a sequence $\{X_n\}$ and a random variable X, we define the **mean squared error** as distance or norm:

$$\mathrm{MSE}(X_n, X) := \mathrm{E}\left[(X_n - X)^2\right] ,$$

One says, $\{X_n\}$ converges in mean square to X for n tending to infinity if

$$\mathrm{MSE}(X_n, X) \;\to\; 0 , \quad n \to \infty .$$

Abbreviating, we write for this as well

$$X_n \;\stackrel{2}{\to}\; X .$$

This limit is unique with probability one. Of course, it can be a random variable itself or a constant. In any case, due to (8.5) it holds that: $\mathrm{E}(X^2) < \infty$. In fact, expected value and variance of X can be determined from the moments of X_n. In particular, the limit of Gaussian random variables is again Gaussian. More precisely, the following lemma holds (see Problem 8.8 for a proof).

Lemma 8.1 (Properties of the Limit in Mean Square) *Let $\{X_n\}$ with (8.5) converge in mean square to X. Then it holds for $n \to \infty$:*

(a) $E(X_n) \to E(X)$;
(b) $E(X_n^2) \to E(X^2)$;
(c) if $\{X_n\}$ is Gaussian, then X follows a Gaussian distribution as well.

Naturally, the parameters of the Gaussian distribution X from (c) follow according to (a) and (b).

Convergence to a Constant

If the limit is a constant, then it is particularly easy to establish convergence in mean square. For this purpose, we consider the following derivation. By zero addition and the binomial formula, the following expression is obtained:

$$\begin{aligned}[X_n - X]^2 &= [(X_n - \mathrm{E}(X_n)) - (X - \mathrm{E}(X_n))]^2 \\ &= (X_n - \mathrm{E}(X_n))^2 - 2(X_n - \mathrm{E}(X_n))(X - \mathrm{E}(X_n)) + (X - \mathrm{E}(X_n))^2.\end{aligned}$$

The expectation operator yields:

$$\begin{aligned}\mathrm{MSE}(X_n, X) = &\ \mathrm{Var}(X_n) \\ &-2\,\mathrm{E}\left[(X_n - \mathrm{E}(X_n))(X - \mathrm{E}(X_n))\right] + \mathrm{E}\left[(X - \mathrm{E}(X_n))^2\right].\end{aligned}$$

If X is a constant (a "degenerate random variable"), $X = c$, then the second term becomes zero and the third term is the expected value of a constant. In other words, this yields:

$$\mathrm{MSE}(X_n, c) = \mathrm{Var}(X_n) + [c - \mathrm{E}(X_n)]^2 .$$

Hence, $\{X_n\}$ converges in mean square to a constant c if and only if it holds that

$$\mathrm{Var}(X_n) \to 0 \quad \text{and} \quad \mathrm{E}(X_n) \to c\,, \quad n \to \infty\,.$$

As is well known, this implies that $\{X_n\}$ converges to c in probability as well (see Lemma 8.3 below). Next, we cover criteria of convergence.

Test of Convergence

Now, we still need a convenient criterion in order to decide whether a series is convergent in mean square. In fact, we have two equivalent criteria. For the proof see Problem 8.9. The name goes back to the famous French mathematician Augustin Louis Cauchy (1789–1857).

Lemma 8.2 (Cauchy Criterion) *A series $\{X_n\}$ with (8.5) converges in mean square . . .*

(a) . . . if and only if it holds for arbitrary n and m that

$$E\left[(X_m - X_n)^2\right] \to 0\,, \quad m,\, n \to \infty\,;$$

(b) . . . or put equivalently, if and only if it holds for arbitrary n and m that

$$E\,(X_m X_n) \to c < \infty\,, \quad m,\, n \to \infty\,,$$

where $c \in \mathbb{R}$ is a constant.

Note that the convergence of the Cauchy criterion holds independently of how m and n tend to infinity. As well, the constant c results independently of the choice of m and n. As the criteria are sufficient and necessary, the proof of existence for mean square convergence can be supplied without determining the limit explicitly.

Example 8.4 (Law of Large Numbers) Let $\{\varepsilon_t\}$ be a white noise process, $\varepsilon_t \sim \mathrm{WN}(0, \sigma^2)$. Then, it can be shown that the arithmetic mean,

$$X_n := \overline{\varepsilon}_n = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t \,,$$

converges in mean square without specifying the limit. It namely holds that

$$\begin{aligned}
\mathrm{E}\left(\overline{\varepsilon}_n\, \overline{\varepsilon}_m\right) &= \frac{1}{mn}\mathrm{E}\left[\sum_{t=1}^{n}\varepsilon_t \sum_{t=1}^{m}\varepsilon_t\right] = \frac{1}{mn}\sum_{t=1}^{\min(n,m)} \mathrm{E}\left(\varepsilon_t^2\right) \\
&= \sigma^2\,\frac{\min(n,m)}{m\,n} \;\to\; 0
\end{aligned}$$

for $m\,, n \to \infty$. Due to Lemma 8.2(b) we conclude that $\overline{\varepsilon}_n$ has a limit in mean square.

Let the limit of $\overline{\varepsilon}_n$ simply be called ε. Naturally, it can be determined immediately. Due to

$$\mathrm{E}\left(\overline{\varepsilon}_n\right) = 0 \quad \text{and} \quad \mathrm{Var}\left(\overline{\varepsilon}_n\right) = \frac{\sigma^2}{n}$$

it follows from Lemma 8.1(a) and (b) for the limit that

$$\mathrm{E}\left(\varepsilon\right) = 0 \quad \text{and} \quad \mathrm{Var}\left(\varepsilon\right) = 0\,.$$

Hence, the limit is equal to zero (with probability one). From this, it follows for $x_t = \mu + \varepsilon_t$ that the arithmetic mean of x_t converges in mean square to the true expected value, μ. In the literature, this fact is also known as the "law of large numbers". ■

Further Modes of Convergence

Two weaker concepts of convergence can be defined via probability statements. First, we say $\{X_n\}$ converges **in probability** to X if it holds for arbitrary $\varepsilon > 0$ that:

$$\lim_{n\to\infty} \mathrm{P}(|X_n - X| \geq \varepsilon) = 0\,.$$

Symbolically, this is denoted as[3]

$$X_n \xrightarrow{p} X \quad \text{for } n \to \infty\,.$$

Secondly, one talks about **convergence in distribution** of $\{X_n\}$ to X if it holds for all points $x \in \mathbb{R}$ at which the distribution function $F_n(x)$ of X_n is continuous, that $F_n(x)$ tends to the distribution function $F(x)$ of X:

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \mathrm{P}(X_n \le x) = \mathrm{P}(X \le x) = F(x)\,.$$

The word "distribution" suggests the symbolic notation:

$$X_n \xrightarrow{d} X \quad \text{for } n \to \infty\,.$$

From Grimmett and Stirzaker (2001, p. 310) or Pötscher and Prucha (2001, Theorem 5 and 9) we adopt the following results.

Lemma 8.3 (Implications of Convergence) *The following implications hold, $n \to \infty$.*

(a) Convergence in mean square implies convergence in probability:

$$(X_n \xrightarrow{2} X) \quad \Longrightarrow \quad (X_n \xrightarrow{p} X)\,.$$

(b) Convergence in probability implies convergence in distribution:

$$(X_n \xrightarrow{p} X) \quad \Longrightarrow \quad (X_n \xrightarrow{d} X)\,.$$

In general, the converse of Lemma 8.3(a) or (b) does not hold. However, if $X = c$ is a constant, then convergence in probability and convergence in distribution are equivalent, see Grimmett and Stirzaker (2001, p. 310) or Pötscher and Prucha (2001, Theorem 10).

8.5 Problems and Solutions

Problems

8.1 Prove Proposition 8.1.
Hint: Use Lemma 8.2.

[3] In particular in econometrics, one often writes alternatively $\operatorname{plim} X_n = X$ as $n \to \infty$.

8.2 Determine the expected value from Proposition 8.3.

8.3 Determine the variance from Proposition 8.3.

8.4 Calculate the variance from Corollary 8.1(b).

8.5 Calculate the variance from Corollary 8.1(c) for the special case of $c = 0$.

8.6 Show that the variance from Corollary 8.1(c) is positive.

8.7 Prove Proposition 8.4.

8.8 Prove Lemma 8.1.

8.9 Prove Lemma 8.2.

Solutions

8.1 Analogously to the partition (8.1) and the Riemann sum R_n from (8.4), we define for arbitrary m with $m \to \infty$:

$$P_m\left([0,\, t]\right): \quad 0 = r_0 < \ldots < r_m = t\,, \; \max_{1 \leq j \leq m} \left(r_j - r_{j-1}\right) \;\to\; 0\,,$$

$$R_m = \sum_{j=1}^{m} f(r_j^*)\, X(r_j^*)\, (r_j - r_{j-1})\,, \qquad r_j^* \in [r_{j-1}, r_j]\,.$$

In order to apply the existence criterion from Lemma 8.2(b), we formulate the product of the two Riemann sums as follows:

$$R_n\, R_m = \sum_{i=1}^{n} \sum_{j=1}^{m} f(s_i^*) f(r_j^*)\, X(s_i^*) X(r_j^*)\, (r_j - r_{j-1})(s_i - s_{i-1})\,.$$

Hence, the Riemann integral as limit of R_n exists if and only if $\mathrm{E}(R_n\, R_m)$ converges. Further,

$$\mathrm{E}(R_n\, R_m) = \sum_{i=1}^{n} \sum_{j=1}^{m} f(s_i^*) f(r_j^*)\, \mathrm{E}\left(X(s_i^*)\, X(r_j^*)\right)\; (r_j - r_{j-1})(s_i - s_{i-1})\,,$$

converges if and only if the ordinary Riemann double integral

$$\int_0^t \int_0^t f(s) f(r)\, \mathrm{E}\left(X(s)\, X(r)\right)\, drds < \infty$$

exists. The Cauchy criterion from Lemma 8.2 therefore amounts to a proof of Proposition 8.1.

8.2 For the WP it holds that

$$\mathrm{E}(W(t)) = 0 \quad \text{and} \quad \mathrm{E}\left(W(s)\, W(r)\right) = \min(s, r)\,.$$

The minimum function is continuous in both the arguments. If f is a continuous function as well, then we know, with the considerations following Proposition 8.1, that the stochastic Riemann integral

$$\int_0^t f(s)\, W(s)\, ds$$

exists. In order to calculate the expected value, Fubini's theorem will be applied. For this purpose, we check that

$$\begin{aligned}\int_0^t \mathrm{E}\left(|f(s)\, W(s)|\right)\, ds &= \int_0^t |f(s)|\, \mathrm{E}\left(|W(s)|\right)\, ds \\ &\leq \max_{0\leq s\leq t} |f(s)| \int_0^t \mathrm{E}\left(|W(s)|\right)\, ds\end{aligned}$$

is finite. The bound is based on the continuity and hence the finiteness of f. The integral

$$\int_0^t \mathrm{E}\left(|W(s)|\right)\, ds$$

was determined in Example 8.1 for t fixed to be finite. As the WP is continuous, Proposition 8.2 can be applied. According to this, it holds that:

$$\mathrm{E}\left(\int_0^t f(s)\, W(s)\, ds\right) = \int_0^t f(s)\, \mathrm{E}(W(s))\, ds = 0\,.$$

8.3 Let us denote the Riemann integral by $Y(t)$:

$$Y(t) = \int_0^t f(s)\, W(s) ds.$$

We have already shown that $\mathrm{E}(Y(t)) = 0$. Hence, it follows that

$$\begin{aligned}\mathrm{Var}(Y(t)) &= \mathrm{E}\left[Y^2(t)\right] \\ &= \mathrm{E}\left[\int_0^t f(s)\, W(s) ds \int_0^t f(r)\, W(r) dr\right]\end{aligned}$$

$$= \mathrm{E}\left[\int_0^t \left(\int_0^t f(r)\, W(r)dr\right) f(s)\, W(s)ds\right]$$
$$= \mathrm{E}\left[\int_0^t \left(\int_0^t f(r)f(s)\, W(r)W(s)dr\right) ds\right].$$

By applying Fubini's theorem twice, we obtain:

$$\begin{aligned}
\mathrm{E}\left[Y^2(t)\right] &= \int_0^t \mathrm{E}\left(\int_0^t f(r)f(s)\, W(r)W(s)dr\right) ds \\
&= \int_0^t \int_0^t \mathrm{E}(f(r)f(s)\, W(r)W(s))dr\, ds \\
&= \int_0^t \int_0^t f(r)f(s)\, \mathrm{E}(W(r)W(s))dr\, ds \\
&= \int_0^t \int_0^t f(r)f(s)\, \min(r,s)dr\, ds\,,
\end{aligned}$$

which is the requested result.

8.4 Let us define

$$Y(t) = W(1) - \int_0^1 W(s)\, ds$$

with $\mathrm{E}(Y(t)) = 0$. Then, it holds that

$$\begin{aligned}
\mathrm{Var}(Y(t)) &= \mathrm{E}(Y^2(t)) \\
&= \mathrm{Var}(W(1)) + \mathrm{Var}\left[\int_0^1 W(s)\, ds\right] - 2\mathrm{E}\left[W(1)\int_0^1 W(s)ds\right] \\
&= 1 + \frac{1}{3} - 2\int_0^1 \mathrm{E}(W(1)W(s))ds,
\end{aligned}$$

where the variance from Corollary 8.1(a) and Fubini's theorem were used. On [0, 1] it holds that:

$$\mathrm{E}(W(1)\, W(s)) = \min(1, s) = s.$$

Hence, the variance results as claimed:

$$\mathrm{Var}(Y(t)) = 1 + \frac{1}{3} - 2\left[\frac{1}{2}s^2\right]_0^1 = 1 + \frac{1}{3} - 1 = \frac{1}{3}.$$

8.5 For $c = 0$ the claim reads

$$\mathrm{Var}\left(\int_0^1 s\, W(s)\, \mathrm{d}s\right) = \sigma_R^2 = \frac{2}{15}\,.$$

According to Proposition 8.3, the variance results as a double integral for $f(s) = s$,

$$\sigma_R^2 = \int_0^1 \int_0^1 rs \min(r,s)\, dr\, ds\,,$$

where the inner integral is appropriately decomposed to facilitate the calculation:

$$\begin{aligned}
\int_0^1 \int_0^1 rs \min(r,s)\, dr\, ds &= \int_0^1 s \left[\int_0^s r \min(r,s)\, dr + \int_s^1 r \min(r,s)\, dr\right] ds \\
&= \int_0^1 s \left[\int_0^s r^2\, dr + \int_s^1 rs\, dr\right] ds \\
&= \int_0^1 s \left[\frac{s^3}{3} + \frac{s}{2}(1 - s^2)\right] ds \\
&= \int_0^1 \left(\frac{s^4}{3} + \frac{s^2}{2} - \frac{s^4}{2}\right) ds \\
&= \int_0^1 \left(\frac{s^2}{2} - \frac{s^4}{6}\right) ds \\
&= \left[\frac{s^3}{6} - \frac{s^5}{30}\right]_0^1 \\
&= \frac{1}{6} - \frac{1}{30} = \frac{4}{30}.
\end{aligned}$$

This corresponds to the claimed result.

8.6 We consider the numerator of σ_R^2,

$$n(c) = 8 - 25c + 20c^2,$$

and show that it does not have any real zeros. Setting $n(c) = 0$ yields:

$$\begin{aligned}
c_{1,2} &= \frac{25 \pm \sqrt{25^2 - 4 \cdot 20 \cdot 8}}{2 \cdot 20} \\
&= \frac{25 \pm \sqrt{-15}}{40} \\
&= \frac{25 \pm i\sqrt{15}}{40}\,, \quad i^2 = -1\,.
\end{aligned}$$

Thus, no real zeros exist. Consequently, since $n(0) = 8 > 0$ the in c continuous function $n(c)$ cannot be zero or negative; the same holds true for σ_R^2, which proves the claim.

Of course, an alternative proof consists in determining the real extrema of $n(c)$. There exists an absolute minimum and this turns out to be positive.

8.7 If the Riemann integral is denoted by $Y(t)$, then it holds, as for the derivation of the variance, that:

$$\begin{aligned}\mathrm{E}(Y(t)\,Y(t+h)) &= \int_0^t f(s)\left[\int_0^s f(r)\min(r,s)dr + \int_s^{t+h} f(r)\min(r,s)dr\right]ds\\ &= \int_0^t f(s)\left[\int_0^s f(r)rdr + s\int_s^{t+h} f(r)dr\right]ds.\end{aligned}$$

Partial integration yields the following relation:

$$\int_0^s f(r)rdr = F(s)s - \int_0^s F(r)dr.$$

By plugging in we obtain the claim.

8.8 Proof of (a): By bounding the difference of the two expected values by means of the Cauchy-Schwarz inequality (2.5) one can immediately tell that this difference tends to zero in the case of convergence in mean square:

$$\begin{aligned}|\mathrm{E}(X_n) - \mathrm{E}(X)| &= |\mathrm{E}(X_n - X)|\\ &\leq \sqrt{\mathrm{E}\,[(X_n - X)^2]} = \sqrt{\mathrm{MSE}(X_n, X)}\\ &\to 0\,, \quad n \to \infty\,.\end{aligned}$$

Proof of (b): The simple trick

$$X_n^2 - X^2 = (X_n - X)^2 + 2\,(X_n - X)\,X$$

yields upon expectation:

$$\begin{aligned}\mathrm{E}(X_n^2) - \mathrm{E}(X^2) &= \mathrm{E}\left[(X_n - X)^2\right] + 2\,\mathrm{E}\,[(X_n - X)\,X]\\ &\leq \mathrm{E}\left[(X_n - X)^2\right] + 2\,|\mathrm{E}\,[(X_n - X)X]|\\ &\leq \mathrm{E}\left[(X_n - X)^2\right] + 2\sqrt{\mathrm{E}\,[(X_n - X)^2]}\,\sqrt{\mathrm{E}(X^2)}\\ &= \mathrm{MSE}(X_n, X) + 2\sqrt{\mathrm{MSE}(X_n, X)}\,\sqrt{\mathrm{E}(X^2)}\\ &\to 0\,, \quad n \to \infty\,,\end{aligned}$$

where again the Cauchy-Schwarz inequality (2.5) was used.

Proof of (c): As is well known, convergence in mean square implies convergence in distribution, see Lemma 8.3. The latter is equivalent to the fact that the characteristic function $\phi_n(u)$ of X_n tends to the characteristic function $\phi(u)$ of X.[4] Now we show: If $\phi_n(u)$ belongs to a Gaussian distribution, then this holds for $\phi(u)$ as well. Hence, (a) and (b) in combination with the premise of a Gaussian sequence $\{X_n\}$ imply:

$$\begin{aligned}\phi_n(u) &= \exp\left\{i\,u\mathrm{E}(X_n) - \frac{u^2\mathrm{Var}(X_n)}{2}\right\}\\ &\to \exp\left\{i\,u\mathrm{E}(X) - \frac{u^2\mathrm{Var}(X)}{2}\right\} = \phi(u)\end{aligned}$$

for $n \to \infty$. Thus, the characteristic function of X as $n \to \infty$ is that of a Gaussian distribution as well.

8.9 Proof of (a): Elementarily, it can be shown that the Cauchy criterion follows from convergence in mean square:

$$\begin{aligned}\mathrm{E}\left[(X_m - X_n)^2\right] &= \mathrm{E}\left[((X_m - X) + (X - X_n))^2\right]\\ &= \mathrm{E}\left[(X_m - X)^2\right] + \mathrm{E}\left[(X - X_n)^2\right]\\ &\quad + 2\,\mathrm{E}\left[(X_m - X)(X - X_n)\right]\\ &\le \mathrm{E}\left[(X_m - X)^2\right] + \mathrm{E}\left[(X - X_n)^2\right]\\ &\quad +2\,\sqrt{\mathrm{E}\left[(X_m - X)^2\right]}\,\sqrt{\mathrm{E}\left[(X - X_n)^2\right]}\\ &= \mathrm{MSE}\,(X_m, X) + \mathrm{MSE}\,(X_n, X)\\ &\quad +2\,\sqrt{\mathrm{MSE}\,(X_m, X)}\,\sqrt{\mathrm{MSE}\,(X_n, X)}\,,\end{aligned}$$

where the bounding is again based on the Cauchy-Schwarz inequality (2.5). It is somewhat more involved that, inversely, the condition from (a) implies convergence in mean square as well. For the proof, we refer e.g. to the exposition on Hilbert spaces in Brockwell and Davis (1991, Ch. 2).

[4] See e.g. sections 5.7 through 5.10 in Grimmett and Stirzaker (2001) for an introduction to the theory and application of characteristic functions. In particular, it holds for the characteristic function of a random variable with a Gaussian distribution, $Y \sim \mathcal{N}(\mu, \sigma^2)$, that:

$$\phi_y(u) = \exp\left\{i\,u\,\mu - \frac{u^2\sigma^2}{2}\right\}\,, \quad i^2 = -1\,,\; u \in \mathbb{R}\,.$$

Proof of (b): If we take (a) for granted, the proof is simple. Due to

$$\mathrm{E}\left[(X_m - X_n)^2\right] = \mathrm{E}(X_m^2) + \mathrm{E}(X_n^2) - 2\,\mathrm{E}(X_m X_n),$$

one can immediately tell that the condition from (a) implies:

$$\mathrm{E}(X_m X_n) \to \frac{\mathrm{E}(X^2) + \mathrm{E}(X^2)}{2} = \mathrm{E}(X^2)\,.$$

Inversely, from the condition from (b) it naturally follows that

$$\begin{aligned}\mathrm{E}\left[(X_m - X_n)^2\right] &= \mathrm{E}(X_m^2) + \mathrm{E}(X_n^2) - 2\,\mathrm{E}(X_m X_n)\\ &\to c + c - 2\,c = 0\,.\end{aligned}$$

This completes the proof.

References

Billingsley, P. (1986). *Probability and measure* (2nd ed.). New York: Wiley.

Brockwell, P. J., & Davis, R. A. (1991). *Time series: Theory and methods* (2nd ed.). New York: Springer.

Grimmett, G. R., & Stirzaker, D. R. (2001). *Probability and random processes* (3rd ed.). Oxford: Oxford University Press.

Pötscher, B. M., & Prucha, I. R. (2001). Basic elements of asymptotic theory. In B. H. Baltagi (Ed.), *A companion to theoretical econometrics* (pp. 201–229). Malden: Blackwell.

Trench, W. F. (2013). *Introduction to real analysis*. Free Hyperlinked Edition 2.04 December 2013. Downloaded on 10th May 2014 from http://digitalcommons.trinity.edu/mono/7.