Wiener Processes (WP)

7

7.1   Summary

The Wiener process (or the Brownian motion) is the starting point and the basis for all the following chapters.[1] This is why we will consider this process more explicitly. It is a continuous-time process having as prominent a position in stochastic calculus as the Gaussian distribution in statistics. After introducing its defining properties intuitively, we will discuss important characteristics in the third section. Examples derived from the Wiener process will conclude the exposition.

7.2   From Random Walk to Wiener Process

We consider a nonstationary special case of the AR(1) process and thereby try to arrive at the Wiener process which is the most important continuous-time process in the fields of our applications.

Random Walks

The cumulation of white noise is labeled **random walk**,

$$x_t = \sum_{j=1}^{t} \varepsilon_j, \quad t \in \{1, 2, \ldots, n\}.$$

---

[1]Norbert Wiener, 1894–1964, was a US-American mathematician. He succeeded in finding a mathematically solid definition and discussion of the so-called Brownian motion named after the nineteenth century British botanist Brown. With a microscope, Brown initially observed and described erratic paths of molecules.

Obviously, it holds that

$$x_t = x_{t-1} + \varepsilon_t, \quad x_0 = 0.$$

In other words: The random walk results as an AR(1) process for the parameter value $a = 1$ and with the starting value zero[2]

$$x_t = a x_{t-1} + \varepsilon_t, \quad a = 1, \; x_0 = 0.$$

As the process is nonstationary,

$$\mathrm{E}(x_t) = 0, \quad \mathrm{Var}(x_t) = \sigma^2 t,$$

it cannot have an infinitely long past, i.e. the index set is finite, $\mathbb{T} = \{1, 2, \ldots, n\}$.

In a way, the random walk models the way home of a drunk who at a point in time t turns to the left or to the right by chance and uncorrelated with his previous path. Put more formally: The random walk is a martingale. We briefly want to convince ourselves of this fact. By substitution the AR(1) process yields

$$x_{t+s} = a^s x_t + \sum_{j=0}^{s-1} a^j \varepsilon_{t+s-j}.$$

Therefore, for $s > 0$ it holds that

$$\mathrm{E}(x_{t+s} | \mathcal{I}_t) = a^s x_t + 0,$$

where $\mathcal{I}_t$ again denotes the information set of the AR(1) process. Thus, the martingale condition (2.11) for AR(1) processes is fulfilled if and only if $a = 1$. The second martingale condition, $\mathrm{E}(|x_t|) < \infty$, is given as $\sigma^2 < \infty$ and hence $\mathrm{E}(x_t^2) = t\sigma^2 < \infty$.

Example 7.1 (Discrete-Valued Random Walk)   Let the set of outcomes contain only two elements (e.g. coin toss: heads or tails),

$$\Omega = \{\omega_0, \, \omega_1\},$$

with probabilities $\mathrm{P}(\{\omega_1\}) = \frac{1}{2} = \mathrm{P}(\{\omega_0\})$. Let $\{\varepsilon_t\}$ be a white noise process assigning the numerical values 1 and -1 to the events,

$$\varepsilon(t; \omega_1) = 1, \quad \varepsilon(t; \omega_0) = -1, \quad t = 1, 2, \ldots, n.$$

---

[2]This special assumption for the starting value is made out of convenience; it is by no means crucial for the behavior of a random walk.

For every point in time, this induces the probabilities

$$P_\varepsilon(\varepsilon_t = 1) = P(\{\omega_1\}) = P_\varepsilon(\varepsilon_t = -1) = P(\{\omega_0\}) = \frac{1}{2}.$$

Then, for expectation and variance it is immediately obtained:

$$E(\varepsilon_t) = 0, \quad \mathrm{Var}(\varepsilon_t) = 1^2 \frac{1}{2} + (-1)^2 \frac{1}{2} = 1.$$

For $t = 1, \ldots, n$, the corresponding random walk $x_t = \sum_{j=1}^{t} \varepsilon_j$ can only take on the countably many values $\{-n, -n+1, \ldots, n-1, n\}$ and is therefore also called discrete-valued. ∎

Example 7.2 (Continuous-Valued Random Walk) If $\{\varepsilon_t\}$ is a Gaussian random process,

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

then, obviously, the random walk based thereon is also Gaussian, where the variance grows linearly with time:

$$x_t = \sum_{j=1}^{t} \varepsilon_j \sim \mathcal{N}(0, \sigma^2 t).$$

In this case, $\{x_t\}$ is a continuous random variable by assumption and hence this random walk is also called continuous-valued. ∎

Wiener Process

At this point, the continuous-time Wiener process will not yet be defined rigorously, but we will approach it intuitively step by step. In order to do so, we choose the index set $\mathbb{T} = [0, 1]$ with the equidistant, disjoint partition

$$[0, 1) = \bigcup_{i=1}^{n} \left[\frac{i-1}{n}, \frac{i}{n} \right).$$

Now, the random walk is multiplied by the factor $1/\sqrt{n}$ and expanded into a step function $X_n(t)$. Interval by interval, we define as continuous-time process:

$$X_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} \varepsilon_j \quad \text{for } t \in \left[\frac{i-1}{n}, \frac{i}{n} \right), \ i = 1, \ldots, n. \tag{7.1}$$

In addition, we assume $\varepsilon_t \in \{-1, 1\}$ and set for $t = 1$

$$X_n(1) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j .$$

For $t = 0$, i.e. $i = 1$ in (7.1), we follow the convention that a sum equals zero if the upper summation limit is smaller than the lower one, which is why $X_n(0) = 0$ begins in the origin. Apparently, $X_n(t)$ is a constant step function on an interval of the length $1/n$, respectively; if $X_n(t)$ was only observed at the jump discontinuities, a time-discrete random walk would be obtained. Being dependent on the choice of n (i.e. the fineness of the partition), the process $X_n(t)$ is indexed accordingly.

If ε_t is from Example 7.1, i.e. $|\varepsilon_t| = 1$, this means that each individual step of the step function has the height $1/\sqrt{n}$ in absolute value. Hence, $X_n(t)$ only takes on values from

$$\left\{ \frac{-n}{\sqrt{n}} , \frac{-n+1}{\sqrt{n}} , \dots , \frac{n-1}{\sqrt{n}} , \frac{n}{\sqrt{n}} \right\} .$$

Therefore, $X_n(t)$ is a continuous-time but discrete-valued process.

Now, the starting point for the Wiener process is the step function $X_n(t)$ with ε_t from Example 7.1. The number of the steps obviously depends on n which indicates the fineness of the partition of the unit interval. Simultaneously, the step height of the steps with $n^{-0.5}$ becomes flatter, the finer it is partitioned. Note that, due to this fact, the range becomes finer and finer and larger and larger as n grows. Hence, with n growing, $X_n(t)$ becomes "more continuous", in the sense that the step heights $n^{-0.5}$ turn out to be smaller; simultaneously, the jump discontinuities move together more closely (the steps of the width $1/n$ get narrower) such that $X_n(t)$ can take on more and more possible values. In the limit ($n \to \infty$) a process named after Norbert Wiener is obtained which we will always denote by W in the following:

$$X_n(t) \; \Rightarrow \; W(t) \quad \text{for } n \to \infty ,$$

where "$\Rightarrow$" denotes a mode of convergence which will be clarified in Chap. 14. Intuitively speaking, it holds that for each of the uncountably many points in time t the function $X_n(t)$ converges in distribution to $W(t)$ just at this point. The transition of discrete-time and discrete-valued step functions from (7.1) to the Wiener process (for n growing) is illustrated in Fig. 7.1.[3]

The Wiener process $W(t)$ as a limit of $X_n(t)$ is continuous-valued with range $\mathbb{R} = (-\infty, \infty)$ and of course it is continuous-time with $t \in [0, 1]$. Furthermore, the Wiener process is a Gaussian process (normally distributed) which is not that surprising. As, due to the central limit theorem for $n \to \infty$, it holds for the

---

[3]In order not to violate the concept of functions, strictly speaking, the vertical lines would not be allowed to occur in the graphs of the figure. We ignore this subtlety to enhance clarity.
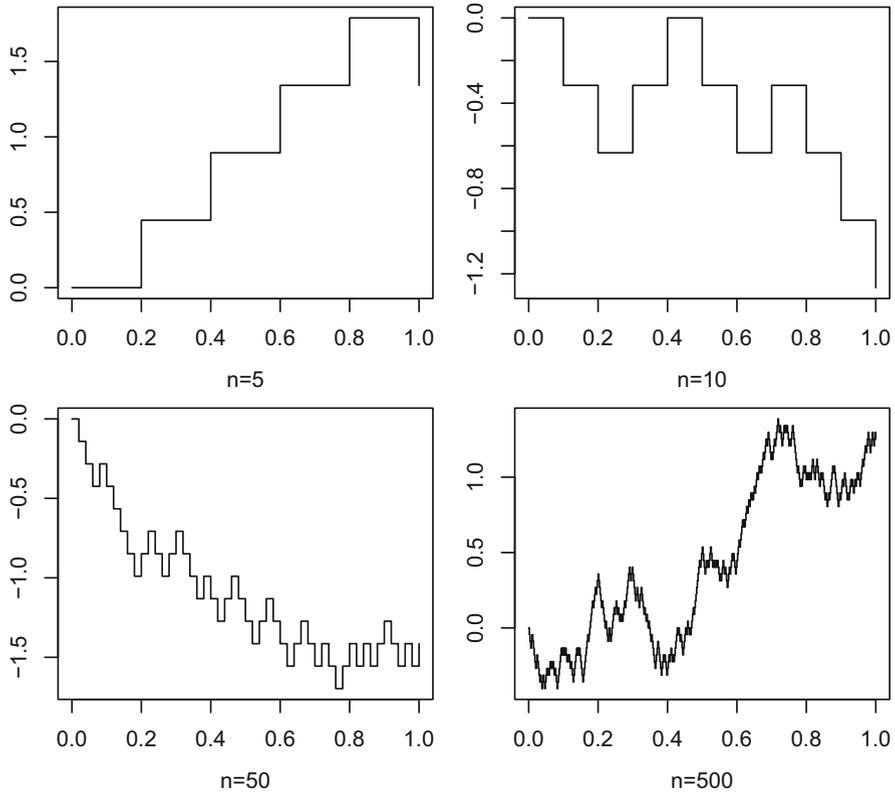
Fig. 7.1 Step function from (7.1) on the interval [0,1]

standardized sum of uncorrelated random variables $\{\varepsilon_j\}$ (whose variance is one and whose expectation is zero) that it tends to a standard normal distribution:

$$X_n(1) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j = \frac{\sum_{j=1}^{n} \varepsilon_j - \mathrm{E}\left(\sum_{j=1}^{n} \varepsilon_j\right)}{\sqrt{\mathrm{Var}\left(\sum_{j=1}^{n} \varepsilon_j\right)}} \xrightarrow{d} \mathcal{N}(0,1) . \qquad (7.2)$$

Here, "$\xrightarrow{d}$" denotes the usual convergence in distribution; cf. Sect. 8.4. As $X_n(1)$ tends to $W(1)$ at the same time, the Wiener process has to be a standard normally distributed random variable at $t = 1$. After giving a formal definition, we will again bridge the gap from the Wiener process to $X_n(t)$.

Formal Definition

The Wiener process (WP) $W(t)$, $t \in [0, T]$, is defined by three assumptions. Put into words, they read: It is a process with starting value zero and independent, normally distributed, stationary increments. These assumptions are to be concretized and specified. Hence, the **Wiener process** is defined by:

(W1) The starting value is zero with probability one, $P(W(0) = 0) = 1$;

(W2) non-overlapping increments $W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1})$, with $0 \leq t_0 \leq t_1 \leq \ldots \leq t_n$, are independent for arbitrary n;

(W3) the increments follow a Gaussian distribution with the variance equalling the difference of the arguments, $W(t) - W(s) \sim \mathcal{N}(0, t - s)$ with $0 \leq s < t$.

Note that the variance of the increments does not depend on the point in time but only on the temporal difference. Furthermore, the covariance of non-overlapping increments is zero due to the independence and the joint distribution results as the product of the marginal distributions. Hence, the joint distribution of non-overlapping increments is multivariate normal. If all increments are measured over equidistant constant time intervals, $t_i - t_{i-1} = const$, then the variances are identical. Therefore, such a series of increments is (strictly) stationary.

Although the WP is defined by its increments, they translate into properties of the level. Obviously, the first and the third property[4] imply

$$W(t) \sim \mathcal{N}(0, t), \tag{7.3}$$

i.e. the Wiener process is clearly a stochastic function being normally distributed at every point in time with linearly growing variance t. More precisely, the WP is even a Gaussian process in the sense of the definition from Chap. 2. The autocovariances being necessary for the complete characterization of the multivariate normal distribution $(W(t_1), \ldots, W(t_n))'$, are determined as follows (see Problem 7.3):

$$\text{Cov}(W(t), W(s)) = \min(s, t). \tag{7.4}$$

The Wiener process, which here will always be denoted by W, is for us a special case of the more general Brownian motion.[5] So to speak, it takes over the role of the *standard* normal distribution, and by multiplication with a constant a general **Brownian motion** is obtained as

$$B(t) = \sigma W(t), \quad \sigma > 0.$$

---

[4]To be completely accurate, this needs to read: $W(t) - W(0) \sim \mathcal{N}(0, t)$. As $W(0)$ is zero with probability one, we set $W(0)$ equal to zero here and in the following; then, the corresponding statements only hold with probability one.

[5]This convention does not hold beyond these pages. Many authors use the terms Wiener process or Brownian motion interchangeably or they apply one of them exclusively.

The assumptions (W1) to (W3) seem very natural if the WP is accepted as a limit of $X_n(t)$ from (7.1). For this process it holds by construction that

- $X_n(t) = 0$ for $t \in [0, 1/n)$,
- e.g. the increments

$$X_n\left(\frac{k+1}{n}\right) - X_n\left(\frac{k}{n}\right) = \frac{\varepsilon_{k+1}}{\sqrt{n}}$$

and

$$X_n\left(\frac{k}{n}\right) - X_n(0) = \frac{1}{\sqrt{n}} \sum_{j=1}^{k} \varepsilon_j$$

are uncorrelated (or even independent if $\{\varepsilon_t\}$ is a pure random process),
- $X_n(1) - X_n(0)$ is approximately normally distributed due to (7.2).

The three properties (W1) to (W3) just reflect the properties of the step function $X_n(t)$.

7.3 Properties

We have already come to know some properties of the WP, for example its autocovariance structure and the Gaussian distribution. For the understanding and handling of Wiener processes further properties are important.

Pathwise Properties

Loosely speaking, it holds that the Brownian motion is everywhere (i.e. for all t) continuous in terms of conventional calculus[6]; however, it is nowhere differentiable. These are pathwise properties, i.e. for a given ω_0, $W(t) = W(t; \omega_0)$ can be regarded as a function which is continuous in t but which is nowhere differentiable. This is a matter of properties being mathematically rather deep but which can be made

---

[6]Occasionally, the pathwise continuity is claimed to be the fourth defining property. This is to be understood as follows. Billingsley (1986, Theorem 37.1) proves more or less the following: If one has a WP W with (W1) to (W3) at hand, then a process W^* can be constructed which is a WP in the sense of (W1) to (W3), as well, which has the same distribution as W *and* which is pathwise continuous. As W^* and W are equal in distribution, they cannot be distinguished and therefore, w.l.o.g. it can safely be assumed that one may work with the continuous W^*.
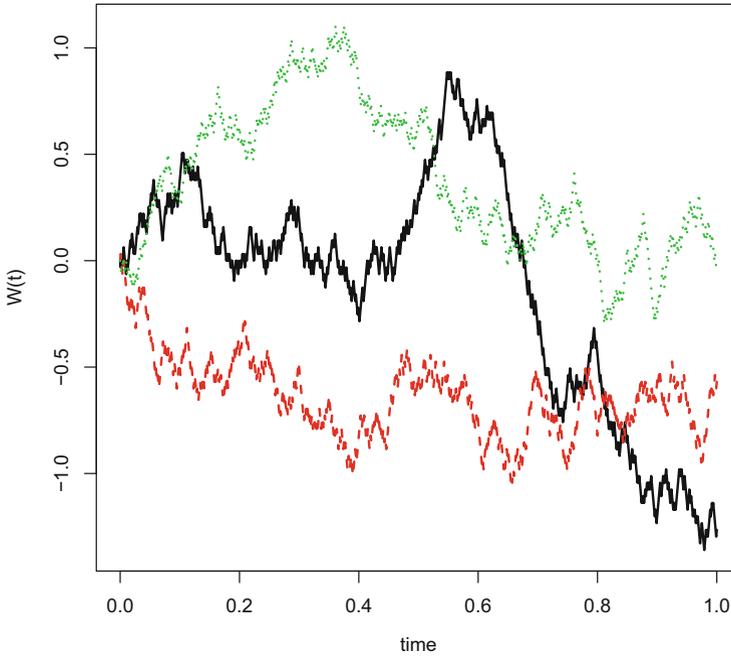
Fig. 7.2 Simulated paths of the WP on the interval [0,1]

heuristically plausible at least. Concerning this, let us consider

$$W(t + h) - W(t) \sim \mathcal{N}(0, h), \quad h > 0.$$

For $h \to 0$ the given Gaussian distribution degenerates to zero suggesting continuity: $W(t + h) - W(t) \approx 0$ for $h \approx 0$. Analogously, we observe a difference quotient whose variance tends to infinity for $h \to 0$,

$$\frac{W(t + h) - W(t)}{h} \sim \mathcal{N}\left(0, \frac{1}{h}\right),$$

which suggests that a usual derivative does not exist. Related to contents, this means that it is not possible to add a tangent line to $W(t)$ which would allow for approximating $W(t + h)$ for an ever so small h. Three simulated paths of the WP in Fig. 7.2 illustrate these properties.

Markov and Martingale Property

In the previous section, we have learned that the random walk is a martingale. For the WP as a continuous-time counterpart, a corresponding result can be obtained

(where $\mathcal{I}_t = \sigma\left(W(r),\ r \leq t\right)$ contains all the information about the past up to t):

$$E\left(|W(t)|\right) < \infty\,,$$

$$E\left(W(t+s)\,|\,\mathcal{I}_t\right) = W(t)\,.$$

The WP satisfies the Markov property (2.9) as well. In order to show this, we use the fact that the increment $W(t+s) - W(t)$ for $s > 0$ is independent of the information set $\mathcal{I}_t$ due to (W2). Hence, for $W(t) = v$ it holds that:

$$P\left(W(t+s) \leq w\,|\,\mathcal{I}_t\right) = P\left(W(t+s) - W(t) \leq w - v\,|\,\mathcal{I}_t\right)$$
$$= P\left(W(t+s) - W(t) \leq w - v\right)\,.$$

At the same time it holds again due to independence that:

$$P\left(W(t+s) \leq w\,|\,W(t) = v\right) = P\left(W(t+s) - W(t) \leq w - v\,|\,W(t) = v\right)$$
$$= P\left(W(t+s) - W(t) \leq w - v\right)\,,$$

which just verifies the Markov property.

Scale Invariance

The Wiener process is a function being Gaussian for every point in time t with expectation zero and variance t. However, time can be measured in minutes, hours or other units. If the time scale is blown up or squeezed by the factor $\sigma > 0$, then it holds that

$$W(\sigma\,t) \ \sim \ \mathcal{N}(0, \sigma\,t)\,.$$

The same distribution is obtained for the $\sqrt{\sigma}$-fold of the Wiener process:

$$\sqrt{\sigma}\,W(t) \ \sim \ \mathcal{N}(0, \sigma\,t)\,.$$

That is why the Wiener process is called scale-invariant (or self-similar). Hence, $W(\sigma\,t)$ and $\sqrt{\sigma}\,W(t)$ are equal in distribution, which we formulate as

$$\sqrt{\sigma}\,W(t) \ \sim \ W(\sigma\,t) \tag{7.5}$$

as well. Such an equality in distribution is to be handled with care and by no means to be confused with ordinary equality. Naturally, it does not hold that, e.g. the double of $W(t)$ is equal to the value at the point in time $4t$:

$$\sqrt{\sigma}\,W(t) \ \neq \ W(\sigma\,t)\,.$$

In other words: Scale invariance is a distributional and not a pathwise property.

Up to this point, it has not been emphasized that the Wiener process is nonstationary. This can already be noted in (7.4) as for $s > 0$ the covariance $\text{Cov}(W(t), W(t + s)) = t$ does not result as dependent on the temporal distance s, but as dependent on the point in time t itself. As we have said, the increments of the WP from (W2), however, are multivariate normal with expectations and covariances of zero and variances which are not affected by a shift on the time axis. The trending behavior of the nonstationary Wiener process will now be clarified by two propositions.

Hitting Time

Let T_b be the point in time at which the WP attains (or hits) a given value $b > 0$ for the first time.[7] By variable transformation it is shown that this random variable has the distribution function (see Eq. (7.14) in Problem 7.5)

$$F_b(t) := \mathrm{P}(T_b \le t) = \frac{2}{\sqrt{2\pi}} \int_{b/\sqrt{t}}^{\infty} e^{-y^2/2} dy.$$

Thereby statement (a) from the following proposition is proved; statement (b) is obtained by means of the corresponding density function (see Problem 7.5).

Proposition 7.1 (Hitting Time) *For the hitting time T_b, where the WP hits $b > 0$ for the first time, it holds that:*

(a) $\mathrm{P}(T_b > t) \to 0$ for $t \to \infty$;
(b) $\mathrm{E}(T_b)$ does not exist.

The result $T_b > t$ is tantamount to the fact that $W(s)$ has not attained the value b up to t:

$$\mathrm{P}(T_b > t) = \mathrm{P}\left(\max_{0 \le s \le t} W(s) < b \right).$$

Laxly formulated this proposition implies that, paradoxically, (a) sooner or later, the WP exceeds every value with certainty; (b) on average, this takes infinitely long: $\mathrm{E}(T_b) = \infty$.

---

[7] The random variable T_b is a so-called "stopping time". This is a term from the theory of stochastic processes which we will not elaborate on here.

Zero Crossing

Next, let $p(t_1, t_2)$ with $0 < t_1 < t_2$ be the probability of the WP hitting the zero line at least once between t_1 and t_2 (even if not necessarily crossing it). We then talk about a zero crossing. The following proposition states how to calculate it. For a proof see e.g. Klebaner (2005, Theorem 3.25).

Proposition 7.2 (Arcus Law)
The probability of a zero crossing between t_1 and t_2, $0 < t_1 < t_2$, equals

$$p(t_1, t_2) = \frac{2}{\pi} \arctan \sqrt{\frac{t_2 - t_1}{t_1}} \,,$$

where arctan *denotes the inverse of the tangent function* $\tan = \frac{\sin}{\cos}$.

It is interesting to fathom out the limiting cases of Proposition 7.2. From the shape of the inverse function of the tangent function it results that

$$\lim_{x \to \infty} \arctan x = \frac{\pi}{2} \quad \text{and} \quad \lim_{x \to 0} \arctan x = 0 \,.$$

Hence, substantially, for $t_2 \to \infty$ it follows that the probability of attaining the zero line tends to one; for $t_2 \to t_1$, however, it naturally converges to zero.

In the literature, an equivalent formulation of the Arcus Law is found:

$$p(t_1, t_2) = \frac{2}{\pi} \arccos \sqrt{\frac{t_1}{t_2}} \,.$$

The equivalence is based on the formula

$$\arctan x = \arccos \frac{1}{\sqrt{1 + x^2}} \,,$$

see e.g. Gradshteyn and Ryzhik (2000, 1.624-8), where "arccos" stands for the inverse of the cosine function.

7.4   Functions of Wiener Processes

When applying stochastic calculus, one is often concerned with processes derived from the Brownian motion. In this section, some of these will be covered and illustrated graphically. We simulate processes on the interval $[0, 1]$; for this purpose, the theoretically continuous processes are calculated at 1000 sampling points and plotted. The resulting graphs are based on pseudo-random variables. Details
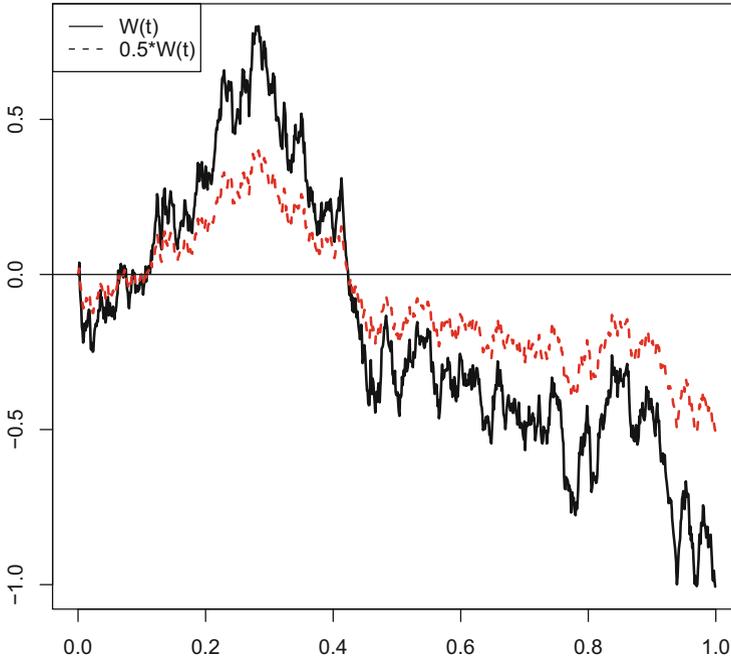
Fig. 7.3 WP and Brownian motion with $\sigma = 0.5$

regarding the simulation of stochastic processes are treated in Chap. 12 on stochastic differential equations.

Brownian Motion $B(t)$

In Fig. 7.3 a path of a WP and a Brownian motion based thereon with only half the standard deviation,

$$W(t) \text{ and } B(t) = 0.5\,W(t)\,,$$

are depicted. Obviously, the one graph is just half of the other.

Brownian Motion with Drift $X(t) = \mu\,t + \sigma\,W(t)$

Here it holds that both the expectation and the variance grow linearly with t:

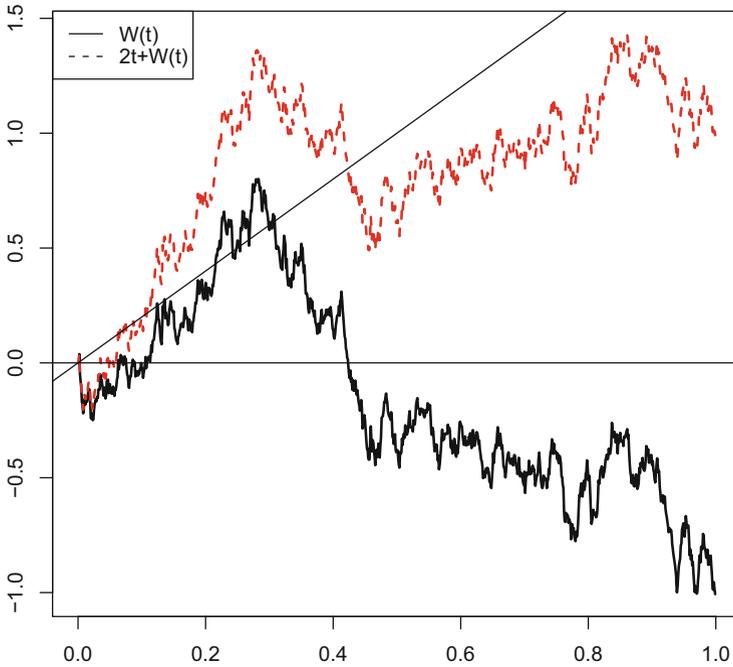$$X(t) \sim \mathcal{N}(\mu t, \sigma^2 t)\,.$$

Fig. 7.4 WP and Brownian motion with drift, where $\sigma = 1$

In Fig. 7.4 the WP from Fig. 7.3 and the process based thereon with drift are depicted. The drift parameter is $\mu = 2$ for $\sigma = 1$, and the expectation function $2\,t$ is also displayed.

Brownian Bridge $X(t) = B(t) - t\,B(1)$

This process is based on the Brownian motion, $B(t) = \sigma\,W(t)$, and fundamentally, it is only defined for $t \in [0, 1]$. The name comes from the fact that the starting and the final value are equal with probability one by construction: $X(0) = X(1) = 0$. One can verify easily that (see Problem 7.6):

$$\mathrm{Var}(X(t)) = t\,(1-t)\,\sigma^2 < t\,\sigma^2. \tag{7.6}$$

Hence, for $t \in (0, 1]$ it holds that $\mathrm{Var}(X(t)) < \mathrm{Var}(B(t))$. This is intuitively clear: With being forced back to zero, the Brownian bridge has to exhibit less variability than the Brownian motion. This is also shown in Fig. 7.5 for $\sigma = 1$.
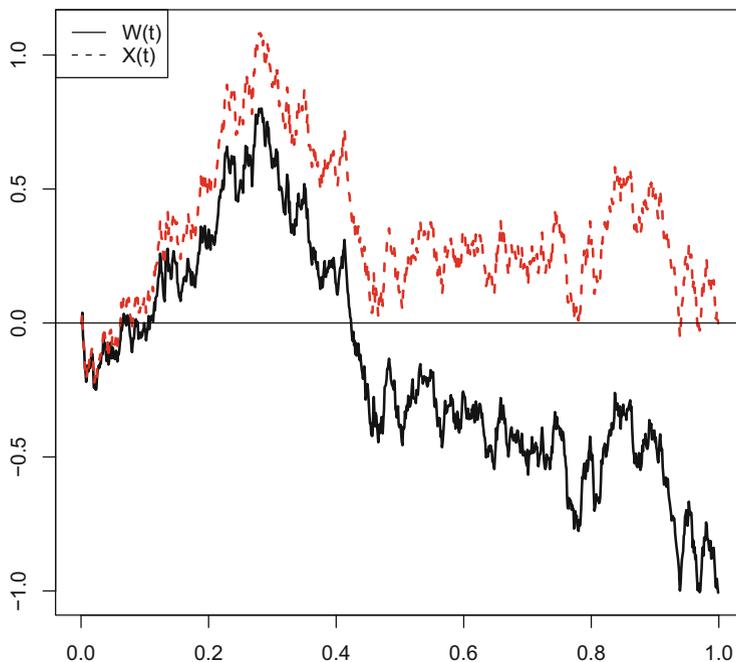
Fig. 7.5 WP and Brownian bridge ($\sigma = 1$)

Reflected Wiener Process $X(t) = |W(t)|$

Due to $W(t) \sim \mathcal{N}(0, t)$, for the distribution function it is elementary to obtain (see Problem 7.7):

$$P(X(t) \leq x) = \frac{2}{\sqrt{2\pi t}} \int_{-\infty}^{x} \exp\left(\frac{-y^2}{2t}\right) dy - 1 .$$

Note that one integrates over twice the density of a Gaussian random variable with expectation zero. Therefore it immediately holds that

$$P(X(t) \leq x) = \frac{2}{\sqrt{2\pi t}} \int_{0}^{x} \exp\left(\frac{-y^2}{2t}\right) dy . \tag{7.7}$$

Expectation and variance of the reflected Wiener process can be determined from the corresponding density function. They read (see Problem 7.7):

$$E(X(t)) = \sqrt{\frac{2t}{\pi}} , \quad \text{Var}(X(t)) = t \left(1 - \frac{2}{\pi}\right) < t = \text{Var}(W(t)) . \tag{7.8}$$
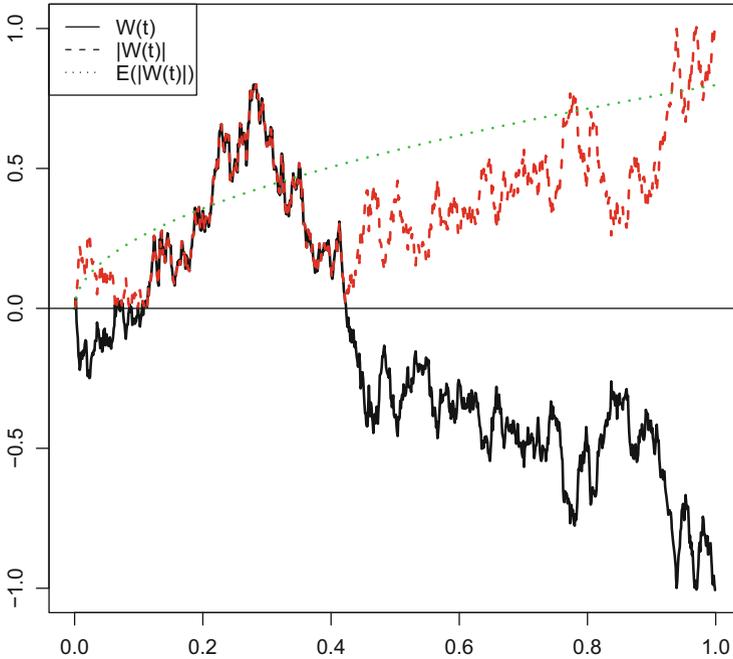
Fig. 7.6 WP and reflected WP along with expectation

As the reflected Wiener process cannot become negative, it has a positive expected value growing with t. For the same reason its variance is smaller than the one of the unrestricted Wiener process, see Fig. 7.6.

Geometric Brownian Motion $X(t) = e^{\mu\, t + \sigma\, W(t)}$

By definition, it holds in this case that the logarithm[8] of the process is a Brownian motion with drift and therefore Gaussian,

$$\log X(t) = \mu\, t + \sigma\, W(t) \sim \mathcal{N}(\mu t, \sigma^2 t)\,.$$

A random variable Y whose logarithm is Gaussian is called – as would seem natural – **log-normal** (logarithmically normally distributed). If it holds that

$$\log Y \sim \mathcal{N}(\mu_y, \sigma_y^2),$$

---

[8] By "log" we denote the natural logarithm and not the common logarithm.

then we know how the two first moments of Y look like, cf. e.g. Sydsæter, Strøm, and Berck (1999, p. 189) or Johnson, Kotz, and Balakrishnan (1994, Ch. 14):

$$E(Y) = e^{\mu_y + \sigma_y^2/2}, \quad Var(Y) = e^{2\mu_y + \sigma_y^2}\left(e^{\sigma_y^2} - 1\right).$$

Hence, by plugging in we obtain for the geometric Brownian motion

$$E(X(t)) = e^{(\mu + \sigma^2/2)\,t} \quad \text{and} \quad Var(X(t)) = e^{(2\mu + \sigma^2)t}\,(e^{\sigma^2 t} - 1). \tag{7.9}$$

While $\log X(t)$ is Gaussian with a linear trend, μt, as expectation, $X(t)$ exhibits an exponentially growing expectation function. Particularly for $\mu = 0$ and $\sigma = 1$ the results

$$E(X(t)) = e^{t/2} \quad \text{and} \quad Var(X(t)) = e^t\,(e^t - 1) \tag{7.10}$$

are obtained. The on average exponential growth in the case of $\mu > -\sigma^2/2$ is illustrated in Fig. 7.7. In Fig. 7.8 we find graphs of the WP and a geometric Brownian motion with expectation one, namely with $\mu = -0.5$ and $\sigma = 1$. Generally, for $\mu = -\sigma^2/2$ an expectation function of one is obtained. Then, one also says that the process does not exhibit a trend (or drift).
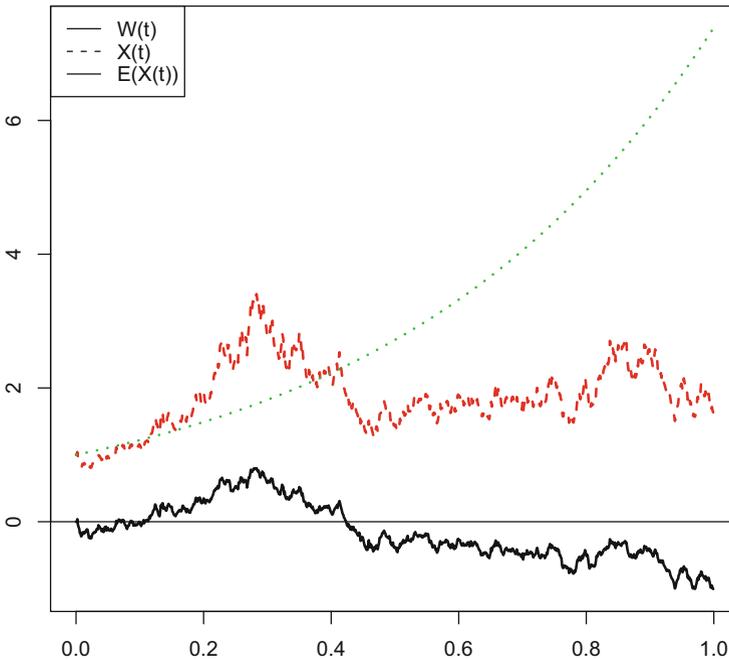

Fig. 7.7 Geometric Brownian motion with $\mu = 1.5$ and $\sigma = 1$ along with expectation
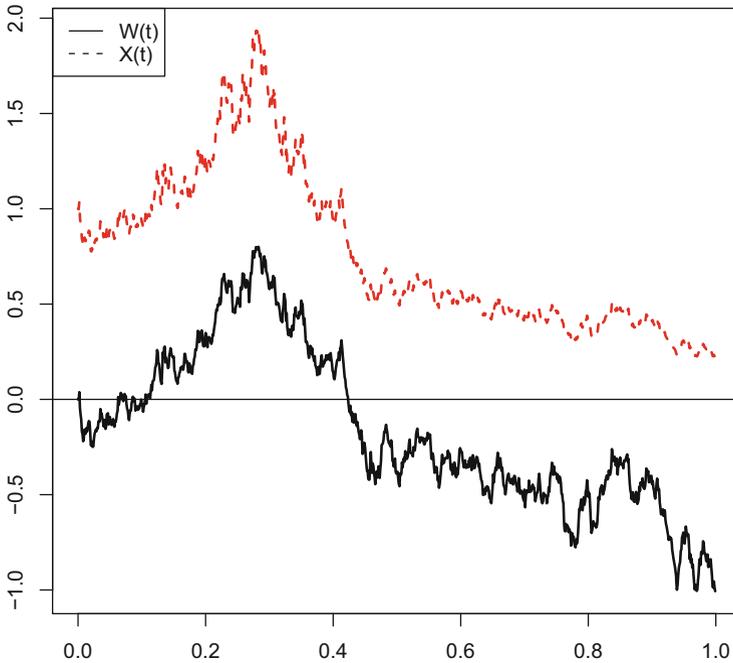
Fig. 7.8 WP and geometric Brownian motion with $\mu = -0.5$ and $\sigma = 1$

In comparison to the expectation, the median of a geometric Brownian motion does not depend on σ. Rather, it holds that (see Problem 7.8):

$$P\left(e^{\mu t + \sigma W(t)} \leq e^{\mu t}\right) = 0.5,$$

such that the median results as $e^{\mu t}$.

Maximum of a WP $X(t) = \max_{0 \leq s \leq t} W(s)$

At t, the maximum process is assigned the maximal value which the WP has taken on up to this point in time. Therefore, in periods of a decreasing Wiener process path, $X(t)$ is constant on the historic maximum until a new relative maximum is attained. However, this process has a distribution function that we have already come to know. By applying the distribution function of the stopping time which is given above Proposition 7.1, one shows (see Problem 7.9) that the maximum process and the reflected WP are equal in distribution:

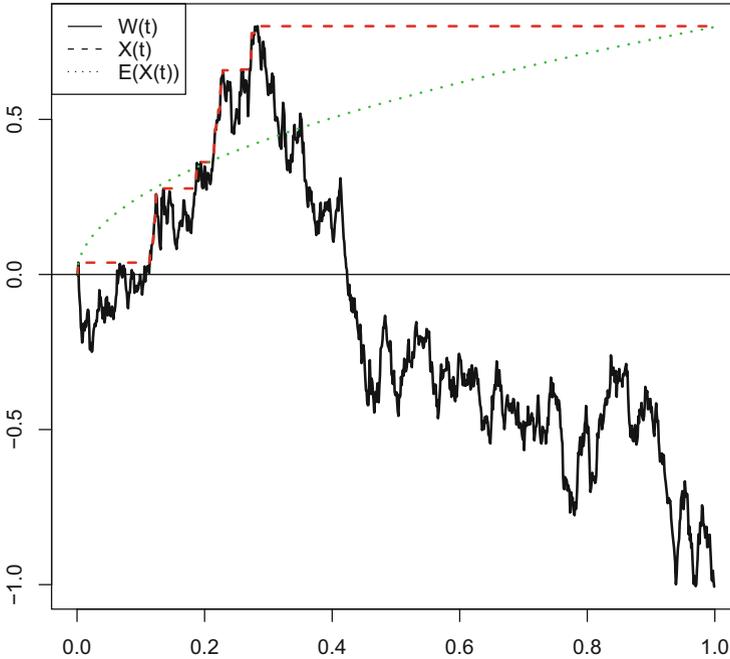$$P(X(t) \leq b) = P(|W(t)| \leq b).$$

Fig. 7.9 WP and maximum process along with expectation

Therefore, expectation and variance of the maximum process of $|W(t)|$ can naturally be copied:

$$\mathrm{E}(X(t)) = \sqrt{\frac{2t}{\pi}}, \quad \mathrm{Var}(X(t)) = t\left(1 - \frac{2}{\pi}\right) < t = \mathrm{Var}(W(t)). \tag{7.11}$$

The expected value is positive and grows with time as the WP again will replace a relative positive maximum by a new relative maximum. Due to the process being again and again constant over times, it is not surprising that its variance is smaller than the one of the underlying WP, cf. Fig. 7.9.

Integrated Wiener Process $X(t) = \int_0^t W(s)\, ds$

As the Brownian motion is a continuous function, the Riemann integral can be defined pathwise. Hence, e.g. the following random variable is obtained:

$$\int_0^1 B(t)\, dt = \sigma \int_0^1 W(t)\, dt.$$

Behind this random variable hides a normal distribution. The latter can be proved by using the definition of the Riemann integral or as a simple conclusion of the Proposition 8.3 below:

$$\int_0^1 W(t)\, dt \sim \mathcal{N}(0, 1/3)\,. \tag{7.12}$$

Basically, by using the integral of a WP, a new stochastic process can also be generated by making the upper limit of integration time-dependent:

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

This idea forms the starting point for the subsequent chapter. In Fig. 7.10 the relation is shown between the WP and the integral $X(t)$ as the area beneath the graph.
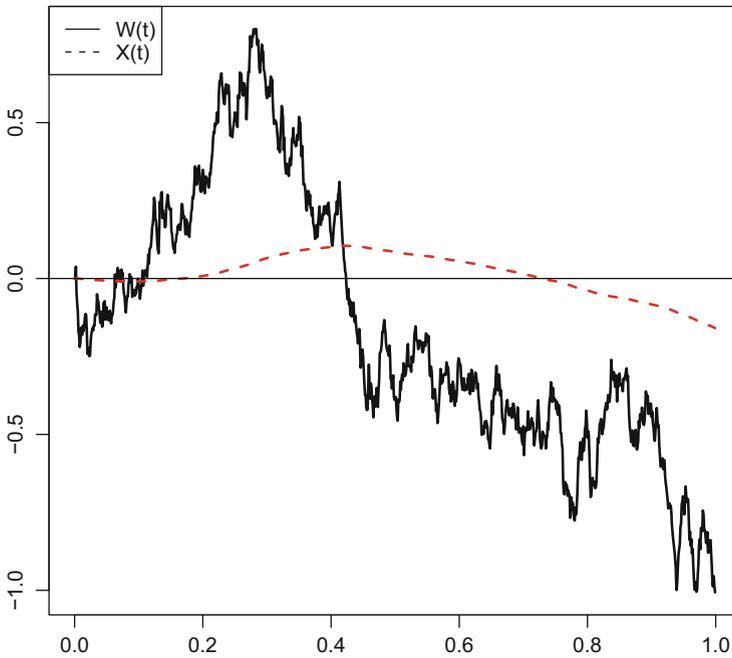

Fig. 7.10 WP and integrated WP

7.5   Problems and Solutions

Problems

7.1 Consider with $\sigma > 0$

$$X(t) = W(1) - \sigma W(1 - t) \quad \text{for } 0 \leq t \leq 1 .$$

Determine the mean and variance of $X(t)$.

7.2 Consider

$$X(t) = t\, W(t^{-1}) \quad \text{for } t > 0 .$$

Determine the covariance of $X(t)$ and $W(t)$, $\text{Cov}(X(t), W(t))$.

7.3 Derive the autocovariance function of a WP, (7.4). Find a simple expression in t and s only for the autocorrelations

$$\rho(s, t) = \frac{\text{Cov}(W(t), W(s))}{\sqrt{\text{Var}(W(t))\, \text{Var}(W(s))}} .$$

7.4 Choose $d \in \mathbb{R}$ such that $T^{d-0.5}\, W(t)$ and $W(T\, t)$ are equal in distribution.

7.5 Prove Proposition 7.1 using the hints given in the text.

7.6 Derive the autocovariance function of a Brownian bridge, and hence show (7.6).

7.7 Determine the distribution function, (7.7), and the moments, (7.8), of a reflected Wiener process.

7.8 Show that in the general case of a geometric Brownian motion, $e^{\mu t + \sigma W(t)}$, the median is given by $e^{\mu t}$.

7.9 Show by means of the hints in the text that the maximum process of a WP and the corresponding reflected WP are equal in distribution:

$$P\left(\max_{0 \leq s \leq t} W(s) \leq b \right) = P(|W(t)| \leq b) .$$

Solutions

7.1 As the Wiener process is on average zero at every point in time, this obviously holds for $X(t)$ as well. Therefore, the variance is calculated as follows:

$$
\begin{aligned}
\mathrm{Var}(X(t)) &= \mathrm{E}(X^2(t)) \\
&= \mathrm{E}[W^2(1) - 2\sigma W(1)W(1-t) + \sigma^2 W^2(1-t)] \\
&= \mathrm{Var}(W(1)) - 2\sigma \mathrm{Cov}(W(1), W(1-t)) + \sigma^2 \mathrm{Var}(W(1-t)) \\
&= 1 - 2\sigma \min(1,\ 1-t) + \sigma^2(1-t) \\
&= 1 - 2\sigma(1-t) + \sigma^2(1-t) \\
&= t + (1-t)(1-\sigma)^2\,.
\end{aligned}
$$

7.2 Due to $\mathrm{E}(W(t)) = \mathrm{E}(X(t)) = 0$ one obtains:

$$
\begin{aligned}
\mathrm{Cov}(X(t), W(t)) &= \mathrm{E}(X(t)W(t)) \\
&= t\,\mathrm{E}(W(t^{-1})W(t)) \\
&= t\,\min(t^{-1}, t).
\end{aligned}
$$

Because of

$$
\min(t^{-1}, t) = \begin{cases} t, & 0 < t \le 1 \\ t^{-1}, & t \ge 1 \end{cases}
$$

it follows that

$$
\mathrm{Cov}(X(t), W(t)) = \begin{cases} t^2, & 0 < t \le 1 \\ 1, & t \ge 1 \end{cases}.
$$

7.3   We simply apply the defining properties (W1), (W2) and (W3) or put differently (7.3). Due to (7.3) the WP has an expectation of zero such that

$$
\mathrm{Cov}(W(t), W(s)) = \mathrm{E}(W(t)W(s))\,.
$$

W.l.o.g. let $s \le t$. By using (W1) and (W2) and after adding zero, we then write:

$$
\begin{aligned}
\mathrm{E}(W(t)W(s)) &= \mathrm{E}([W(s) + W(t) - W(s)]\,[W(s) - W(0)]) \\
&= \mathrm{E}([W(s)]^2) + \mathrm{E}([W(t) - W(s)]\,[W(s) - W(0)]) \\
&= s + 0,
\end{aligned}
$$

where the last equality uses $\mathrm{Var}(W(s)) = s$ and the independence of non-overlapping increments. As one could also assume $t \leq s$ w.l.o.g., (7.4) is verified.

With the autocovariance one obtains

$$\rho(s, t) = \frac{\min(s, t)}{\sqrt{t\,s}} = \frac{\min(s, t)}{\sqrt{\max(s, t)\,\min(s, t)}} = \sqrt{\frac{\min(s, t)}{\max(s, t)}}\,.$$

7.4 This is a problem on self-similarity or scale invariance. Due to (7.3) it obviously holds that:

$$T^{d-0.5}W(t) \sim \mathcal{N}(0, T^{2d-1}t)$$

and

$$W(Tt) \sim \mathcal{N}(0, Tt)\,.$$

Therefore, the corresponding variances are equal for $d = 1$. The corresponding result is obtained from (7.5) as well:

$$T^{0.5}W(t) \sim W(Tt)\,.$$

7.5 Proof of Proposition 7.1(a): Our proof consists of three steps. At first, we establish the equation

$$\mathrm{P}(T_b < t) = 2\,\mathrm{P}(W(t) > b)\,. \tag{7.13}$$

Secondly, by using this we show:

$$F_b(t) := \mathrm{P}(T_b \leq t) = \frac{2}{\sqrt{2\pi}} \int_{b/\sqrt{t}}^{\infty} e^{-y^2/2}\,dy\,. \tag{7.14}$$

Note that in (7.14) the integrand amounts to the density function of the standard normal distribution. Hence, thirdly for $t \to \infty$ the claim immediately follows from (7.14) due to $\mathrm{P}(T_b > t) = 1 - \mathrm{P}(T_b \leq t)$.

In order to accept (7.13), we remember (7.3). Accordingly, for $T_b < t$ it holds that

$$W(t) - W(T_b) \sim \mathcal{N}(0, t - T_b)\,,$$

which is why from the symmetry of the Gaussian distribution with $W(T_b) = b$ it follows for the conditional probability that:

$$\mathrm{P}\left(W(t) > b \mid T_b < t\right) = \mathrm{P}\left(W(t) - W(T_b) > 0 \mid T_b < t\right) = \frac{1}{2}\,.$$

Hence, it results that:

$$\frac{1}{2} = \frac{P(W(t) > b \text{ and } T_b < t)}{P(T_b < t)} = \frac{P(W(t) > b)}{P(T_b < t)},$$

where the last equality is caused by the fact that $T_b < t$ is implied by $W(t) > b$. Thus, we obtain Eq. (7.13) which will now be applied for deriving (7.14).

Due to (7.3) it holds by definition that

$$P(W(t) > b) = \int_b^\infty \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dx.$$

By variable transformation, $y = \frac{x}{\sqrt{t}}$, it follows that

$$P(W(t) > b) = \int_{b/\sqrt{t}}^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy,$$

whereby (7.14) and hence claim (a) is proved due to $P(T_b < t) = P(T_b \le t)$.

Proof of Proposition 7.1(b): With the density function $f_b(t) = F_b'(t)$ the approach for the expected value reads as follows:

$$E(T_b) = \int_0^\infty t f_b(t) \, dt.$$

Note that the distribution function derived in (7.14) has the following structure with the antiderivative H, $H' = h$:

$$F_b(t) = \int_{g(t)}^\infty h(y) \, dy = \lim_{c \to \infty} H(c) - H(g(t)).$$

Therefore, due to the chain rule it holds for the density that

$$F_b'(t) = -h(g(t)) \, g'(t) = \frac{b \, e^{-\frac{b^2}{2t}} t^{-\frac{3}{2}}}{\sqrt{2\pi}}.$$

Hence, the variable transformation results in $t = b^2 u^{-2}$ with $dt = -2b^2 u^{-3} du$:

$$E(T_b) = \int_0^\infty t F_b'(t) \, dt$$

$$= \frac{b}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{b^2}{2t}} t^{-\frac{1}{2}} dt = \frac{-2b^2}{\sqrt{2\pi}} \int_\infty^0 e^{-\frac{u^2}{2}} u^{-2} \, du$$

$$= \frac{2b^2}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{u^2}{2}} u^{-2} du$$

$$\geq \frac{2b^2}{\sqrt{2\pi}} \int_0^1 e^{-\frac{u^2}{2}} u^{-2} du$$

$$\geq \frac{2b^2}{\sqrt{2\pi}} e^{-\frac{1}{2}} \int_0^1 u^{-2} \, du.$$

However, this last integral written down symbolically does not exist since the antiderivative of u^{-2} is $-u^{-1}$, and

$$\int_\varepsilon^1 u^{-2} \, du = \varepsilon^{-1} - 1$$

diverges as $\varepsilon \to 0$. This completes the proof.

7.6 First, we determine that

$$X(t) = W(t) - t\, W(1)\,, \quad t \in [0, 1]\,,$$

has zero expectation:

$$\mathrm{E}(X(t)) = \mathrm{E}(W(t)) - t\,\mathrm{E}(W(1)) = 0 - 0\,.$$

By multiplying out and application of (7.4) one can show that

$$\begin{aligned}
\mathrm{Cov}(X(t),\, X(s)) &= \mathrm{E}(X(t)\,X(s))\\
&= \mathrm{E}\left(W(t)W(s) - tW(1)W(s) - sW(1)W(t) + st W^2(1)\right)\\
&= \min(s, t) - t\min(s, 1) - s\min(t, 1) + st\\
&= \min(s, t) - st - st + st\\
&= \min(s, t) - st\,.
\end{aligned}$$

In particular, for $s = t$ the variance formula (7.6) is obtained.

7.7 At first we determine the distribution function (7.7) for $X(t) = |W(t)|$:

$$\begin{aligned}
F_x(x) &= \mathrm{P}(X(t) \leq x)\,, \quad x \geq 0\,,\\
&= \mathrm{P}(W(t) \leq x) - \mathrm{P}(W(t) \leq -x)\\
&= \mathrm{P}(W(t) \leq x) - (1 - \mathrm{P}(W(t) \leq x))\\
&= 2\,\mathrm{P}(W(t) \leq x) - 1\,,
\end{aligned}$$

where the symmetry of the Gaussian distribution was used. With $W(t) \sim \mathcal{N}(0, t)$ we therefore have

$$F_x(x) = \frac{2}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{y^2}{2t}} \, dy - 1 \,,$$

or for the density

$$f_x(x) = F_x'(x) = \frac{2}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \,.$$

In order to calculate expectation and variance, we determine the r-th moment in general:

$$E(X^r(t)) = \int_0^\infty x^r f_x(x) \, dx = \frac{2}{\sqrt{2\pi t}} \int_0^\infty x^r e^{-\frac{x^2}{2t}} \, dx \,.$$

By substitution, these moments can be reduced to the Gamma function, which was introduced in Problem 5.3, see also below Eq. (5.20). For $a > 0$ and with $a x^2 = u$ and $du = 2ax \, dx$, we obtain:

$$
\begin{aligned}
\int_0^\infty x^r e^{-ax^2} dx &= \int_0^\infty \left(\frac{u}{a}\right)^{\frac{r-1}{2}} e^{-u} \frac{du}{2a} \\
&= \frac{1}{2} a^{-\frac{r+1}{2}} \int_0^\infty u^{\frac{r+1}{2}-1} e^{-u} du \\
&= \frac{1}{2} a^{-\frac{r+1}{2}} \Gamma\left(\frac{r+1}{2}\right) \,.
\end{aligned}
$$

The Gamma function possesses a number of nice properties and special values, remember in particular e.g.

$$\Gamma(1) = 1 \,, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \,, \quad \Gamma(n+1) = n\,\Gamma(n) \,.$$

With $a = \frac{1}{2t}$, for the moments it therefore follows that:

$$E(X(t)) = \sqrt{\frac{2t}{\pi}} \,, \quad E(X^2(t)) = \frac{2}{\sqrt{2\pi t}} \frac{1}{2} (2t)^{\frac{3}{2}} \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = t \,.$$

The variance formula is obtained by the usual variance decomposition:

$$\mathrm{Var}(X(t)) = E\left(X^2(t)\right) - (E(X(t)))^2 \,.$$

7.8 The random variable $\sigma\, W(t)$ follows for fixed t a Gaussian distribution with expectation and median equal to zero. Hence, it follows that

$$P(\sigma\, W(t) \le 0) = P\left(e^{\sigma\, W(t)} \le 1\right) = 0.5\,.$$

By multiplying the inequality by $e^{\mu t}$, we obtain

$$P\left(e^{\mu t}\, e^{\sigma\, W(t)} \le e^{\mu t}\right) = P\left(e^{\mu t + \sigma\, W(t)} \le e^{\mu t}\right) = 0.5\,.$$

Therefore, the median of $X(t) = e^{\mu t + \sigma\, W(t)}$ is determined independently of σ as $e^{\mu t}$ as claimed.

7.9 As $X(t) = \max_{0 \le s \le t} W(s)$ is a continuous random variable for given t, it holds that

$$F_x(b) := P(X(t) \le b) = P\left(\max_{0 \le s \le t} W(s) < b\right)\,.$$

Remember the random variable T_b from Proposition 7.1 specifying the point in time at which $W(t)$ hits the value b for the first time. The event $\max_{0 \le s \le t} W(s) < b$ is equivalent to the fact that the hitting time of b is larger than t. Therefore, when using the distribution function from Proposition 7.1, it holds that

$$P(X(t) \le b) = 1 - P(T_b \le t) = 1 - \frac{2}{\sqrt{2\pi}} \int_{b/\sqrt{t}}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz\,.$$

Naturally, the number 1 can be written as an integral over the density of the standard normal distribution:

$$P(X(t) \le b) = \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz - \frac{2}{\sqrt{2\pi}} \int_{b/\sqrt{t}}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz$$

$$= \frac{2}{\sqrt{2\pi}} \int_{0}^{b/\sqrt{t}} \exp\left(-\frac{z^2}{2}\right) dz\,.$$

By substitution,

$$z = \frac{y}{\sqrt{t}} \quad \text{and} \quad dz = \frac{dy}{\sqrt{t}}\,,$$

and due to (7.7) the desired result is immediately obtained:

$$P(X(t) \le b) = \frac{2}{\sqrt{2\pi t}} \int_{0}^{b} \exp\left(-\frac{y^2}{2t}\right) dy$$

$$= P\left(|W(t)| \le b\right)\,.$$

References

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

Gradshteyn, I. S., & Ryzhik, I. M. (2000). *Table of integrals, series, and products* (6th ed.). London/San Diego: Academic Press.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions, Volume 1* (2nd ed.). New York: Wiley.

Klebaner, F. C. (2005). *Introduction to stochastic calculus with applications* (2nd ed.). London: Imperical College Press.

Sydsæter, K., Strøm, A., & Berck, P. (1999). *Economists' mathematical manual* (3rd ed.). Berlin/New York: Springer.