11 Brownian Motion

The main motivation for studying random walks on integer grids is the physically much more relevant but mathematically much more difficult notion of the Brownian motion in continuous space.

History. In 1827, the Scottish botanist ROBERT BROWN observed the erratic movement of pollen particles within a waterdrop under the microscope; hence the term Brownian motion. At the end of the 19th century, THORVALD NICOLAI THIELE in Denmark and LOUIS BACHELIER in France described mathematical models of this process, primarily to analyze economic data. Both contributions had little influence on the later developments, perhaps because financial mathematics was in its infancy. In 1905, the German physicist ALBERT EINSTEIN formulated the process in today’s form in order to explain the motion of a particle in terms of the molecular structure of water. The existence of this stochastic process was not established until 1923 when the American mathematician NORBERT WIENER based it on tools from measure theory. Today, the mathematical formalism is called the Wiener process. Another major contribution was the development of a stochastic calculus by ITÔ KIYOSHI, which was instrumental in the adaptation of the Wiener process to applications in various areas. Today, the Brownian motion and related stochastic processes are major probabilistic tools in the natural and social sciences.

Continuous construction. Brownian motion is an example of a continuous stochastic process, perhaps the most important example. It consists of an uncountable collection of random variables $W_t$, for $t \in \mathbb{R}$. We think of $t$ as time and the Brownian motion as a function $t \mapsto W_t$. To get a first approximation, we can understand the function as the limit of a random walk in which the step-size goes to zero. We begin by scaling time and space separately. For the random walk, we have $\Delta t = 1$ and $\Delta x = 1$. We now set $\Delta t = \delta = 1/N$ for a step in time. For large $N$, the process should look like

$$W_{t\delta} \approx \Delta x S_t.$$  

We need to determine $\Delta x$ in terms of $\delta$ so that the process scales correctly. To do this, we normalize such that $E(W_t^2) = 1$. Since

$$E((\Delta x S_N)^2) = (\Delta x)^2 E(S_N^2) = (\Delta x)^2 N,$$

we conclude that $\Delta x = \sqrt{\delta} = 1/\sqrt{N}$. Setting $t = j\delta = j/N$, we write

$$W_t = W_{j/N} \approx S_j/\sqrt{N} = \frac{S_j}{\sqrt{j/N}} = \frac{S_j}{\sqrt{j}}.$$  

As $j$ goes to infinity, the Central Limit Theorem tells us that $S_j/\sqrt{j}$ approaches the normal distribution, with mean 0 and variance 1. This leads us to the first requirement in our construction of the mathematical formalism of Brownian motion, namely that each $W_t$ is normally distributed with mean 0 and variance $t$. More generally, we ask the same for each difference.

I. For $0 \leq s < t < \infty$, the random variable $W_t - W_s$ has a normal distribution with mean 0 and variance $t - s$.

The $W_t - W_s$ are sometimes called the identically distributed normal increments. We would like to have independent increments, like in a Markov process:

II. For all $s < t$, the random variable $W_t - W_s$ is independent of all random variables $W_r$ with $r \leq s$.

Note that this is not saying that $W_s$ and $W_t$ are independent; in fact they are not. Instead $W_s$ and $W_t - W_s$ are independent. The technically most difficult is the third requirement since it needs a significant amount of analysis:

III. The function $t \mapsto W_t$ be continuous.

It is not trivial at all to prove that such a process exists. This is the achievement of NORBERT WIENER, which is why the mathematical formulation is usually referred to as the Wiener process. To generate it, we may proceed as in the random walk, except that the step-size be chosen from a normal distribution with variance $(\Delta x)^2$. This process has a number of interesting properties, that it is nowhere differentiable with probability 1 being one.

Gaussian normal distributions and kernels. Since the normal distribution figures so prominently in the construction of the Wiener process, we study some of its properties in detail. In 1 dimension, the Gaussian normal distribution with zero mean and variance $\sigma^2$ is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}.$$  

It has the characteristic shape of a bell, with inflection points at $\pm\sigma$; see Figure 15. In the interval from $-\sigma$ to
The connection to probability is the sum of two random variables. If \( X, Y \) are random variables with distribution functions \( f, g : \mathbb{R} \rightarrow \mathbb{R} \), then \( X + Y \) is a random variable with distribution function \( f \ast g \) defined as

\[
G(x) = \int_y g(x - y) f(y) \, dy.
\]

We can think of replacing \( g(x) \) by a weighted average of the values of \( g \) in a neighborhood of \( x \); the weights are furnished by \( f \). For example, if

\[
f(y) = \begin{cases} 
1 & \text{for } -\frac{1}{2} \leq y \leq \frac{1}{2}, \\
0 & \text{otherwise},
\end{cases}
\]

then \( g(x) \) gets replaced by the average of \( g \) in the interval \([x - \frac{1}{2}, x + \frac{1}{2}]\). If \( g \) is defined the same way, we get \( f \ast g \) as a hat function:

\[
(f \ast g)(x) = \begin{cases} 
1 - \|x\| & \text{for } -1 \leq x \leq 1, \\
0 & \text{otherwise},
\end{cases}
\]

To cast additional light onto this construction, we consider the joint distribution of \( X \) and \( Y \), which is \( F : \mathbb{R}^2 \rightarrow \mathbb{R} \) defined by \( F(x, y) = g(x) \cdot f(y); \) see Figure 16. Assuming both \( f \) and \( g \) are probability distributions,

The property expressed by (19) is referred to as the separability of the Gaussian kernel. Each factor in the product is a 1-dimensional normal distribution, whose integral is equal to 1. We can now prove \( \int f_d(x) \, dx = 1 \) by induction over the dimension.

**Convolution.** This is an operation between two functions; it maps functions \( f \) and \( g \) to the function \( f \ast g \). The graph of the Gaussian normal distribution and its first two derivatives. For better visibility, we show the second derivative multiplied with a quarter.
that the mean (the expectations) are additive in general, and that the variances are additive if $X$ and $Y$ are independent. We still need to show that $X + Y$ is normally distributed. We do this by computing the convolution of the two distributions, $f$ of $X$ and $g$ of $Y$:

$$G(x) = \int_y g(x - y)f(y) \, dy$$

$$= \int_y \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-y-\nu)^2}{2v}} \frac{1}{\sqrt{2\pi u}} e^{-\frac{(y-\mu)^2}{2u}} \, dy.$$ 

The product of the two coefficients can be written as

$$\frac{1}{\sqrt{2\pi v}} \cdot \frac{1}{\sqrt{2\pi u}} = \frac{1}{\sqrt{2\pi (u + v)}}.$$ 

The sum of the two exponents can be written as

$$-\frac{(x - y - \nu)^2}{2v} - \frac{(y - \mu)^2}{2u} \quad \Rightarrow \quad -\frac{(x - \mu - \nu)^2}{2u + v} \cdot \left[ \frac{y - \frac{u(x - \nu) + \nu}{u + v}}{2 \sqrt{\frac{uv}{u + v}}} \right]^2.$$ 

Computing the integral for the second coefficient on the right hand side and the second exponent on the right hand side, we get

$$\int_y \frac{1}{\sqrt{2\pi \sqrt{\frac{uv}{u + v}}}} \cdot e^{-\frac{(y - \frac{u(x - \nu) + \nu}{u + v})^2}{2 \frac{uv}{u + v}}} \, dy.$$ 

But this is the integral of a normal distribution, which evaluates to 1. What remains are the first coefficient and the first exponent, which gives

$$G(x) = \frac{1}{\sqrt{2\pi (u + v)}} \cdot e^{-\frac{(x - (\mu + \nu))^2}{2(u + v)}}. \quad \text{(20)}$$

The same is true in two and higher dimensions:

**Theorem.** Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be Gaussian kernels with mean values $\mu, \nu \in \mathbb{R}^d$ and variances $u, v > 0$. Then the convolution $f * g$ is the Gaussian kernel with mean $\mu + \nu$ and variance $u + v$.

While the proof using convolution is technically involved, the result is not surprising. For example, the sum of two binomially distributed random variables is again a binomially distributed random variable. Indeed, flipping a coin $m$ times and then $n$ times is like flipping it $m + n$ times. As $m$ and $n$ go to infinity, the distributions of the three random variables each approach a normal distribution.