1 Notes on Brownian Motion

We present an introduction to *Brownian motion*, an important continuous-time stochastic process that serves as a continuous-time analog to the simple symmetric random walk on the one hand, and shares fundamental properties with the Poisson counting process on the other hand.

Throughout, we use the following notation for the real numbers, the non-negative real numbers, the integers, and the non-negative integers respectively:

$$\mathbb{R} \stackrel{\text{def}}{=} (-\infty, \infty) \tag{1}$$

$$\mathbb{R}_+ \stackrel{\text{def}}{=} [0, \infty) \tag{2}$$

$$\mathbb{Z} \stackrel{\text{def}}{=} \{\cdots, -2, -1, 0, 1, 2, \cdots\} \tag{3}$$

$$\mathbb{N} \stackrel{\text{def}}{=} \{0, 1, 2, \cdots\}. \tag{4}$$

Before our study of Brownian motion, we must review the normal distribution, and its importance due to the central limit theorem. We do so next.

1.1 Normal distribution

Of particular importance in our study is the normal distribution, $N(\mu, \sigma^2)$, with mean $-\infty < \mu < \infty$ and variance $0 < \sigma^2 < \infty$; the probability density function and cdf are given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R}, \tag{5}$$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(y-\mu)^2}{2\sigma^2}} dy, \ x \in \mathbb{R}. \tag{6}$$

The normal distribution is also called the *Gaussian* distribution after the famous German mathematician and physicist Carl Friedrich Gauss (1777 - 1855). The description “bell curve” is given to the shape of the density function $y = f(x), x \in \mathbb{R}$ when graphed in the $x - y$ plane: It looks like a bell centered symmetrically about the mean value μ.

When $\mu = 0$ and $\sigma^2 = 1$ we obtain the *standard* (or *unit*) normal distribution, $N(0, 1)$, and the density and cdf reduce to

$$\phi(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \tag{7}$$

$$\Phi(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-y^2}{2}} dy. \tag{8}$$

We usually denote a $N(0, 1)$ rv by Z and write $Z \sim N(0, 1)$; $\Phi(x) = P(Z \leq x), \ x \in \mathbb{R}$.

That $\phi(x)$ really is a density function, that is, that $\int_{-\infty}^{\infty} \phi(x)dx = 1$:

$\phi(x) \geq 0$ (non-negativity holds), so we must only prove that $C = \int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$; that the normalizing factor is $\sqrt{2\pi}$.

To do so, we will show that $C^2 = 2\pi$ by using a change into polar coordinates $x = r\cos(\theta)$, $y = r\sin(\theta)$, with $\theta \in (0, 2\pi]$, and $r^2 = x^2 + y^2$ with $r \in \mathbb{R}_+$, and $dxdy = rdrd\theta$ in what follows:

$$
\begin{aligned}
C^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2/2}dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2/2}dy\right) && (9)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{-x^2+y^2}{2}}dxdy && (10)\\
&= \int_0^{2\pi}\int_0^{\infty} re^{-r^2/2}drd\theta, \text{ (Polar coordinates)} && (11)\\
&= \int_0^{\infty}\int_0^{2\pi} re^{-r^2/2}d\theta dr, \text{ (Fubini's Theorem (Tonelli's version))} && (12)\\
&= 2\pi\int_0^{\infty} re^{-r^2/2}dr && (13)\\
&= 2\pi\int_0^{\infty} e^{-u}du, \text{ } (u = r^2/2 \text{ change of variables}) && (14)\\
&= 2\pi \times 1 && (15)\\
&= 2\pi && (16)
\end{aligned}
$$

As we shall see over and over again in our study of Brownian motion, one of its nice features is that many computations involving it are based on evaluating $\Phi(x)$, and hence are computationally elementary.

It is easily seen that

1. If $Z \sim N(0,1)$, then $X = \sigma Z + \mu$ has the $N(\mu, \sigma^2)$ distribution.

2. Conversely, if $X \sim N(\mu, \sigma^2)$, then $Z = (X - \mu)/\sigma$ has the standard normal distribution.

For example if $X = \sigma Z + \mu$, then $F(x) = P(X \le x) = P(\sigma Z + \mu \le x) = P(Z \le (x - \mu)/\sigma) = \Phi((x - \mu)/\sigma)$, differentiating and using (7), $\Phi'((x-\mu)/\sigma) = \phi((x-\mu)/\sigma)(1/\sigma) = f(x)$. In other words X has the $N(\mu, \sigma^2)$ density given in (5).

Another important (and easy to derive) fact is that

> *if $X \sim N(\mu, \sigma^2)$, then $-X \sim N(-\mu, \sigma^2)$, and in particular $-Z$ remains a unit normal; it has the same distribution as Z. (This is due to symmetry about the origin 0.)*

1.2 Simulating from the normal distribution

As pointed out already, once we have a copy of $Z \sim N(0,1)$ we can transform it into an $X \sim N(\mu, \sigma^2)$ via setting $X = \sigma Z + \mu$. Thus it suffices to have a simulation algorithm for generating iid copies of $Z \sim N(0,1)$. The inverse transform method can't be used because we do not have an explicit form of the CDF $\Phi(x) = P(Z \le x)$ in (8) let alone its inverse. One might first try to approximate $\Phi^{-1}(y)$ by an explicit tractable function so as to use the inverse transform method to obtain approximate copies of Z, and that is an approach sometimes used in practice. However, we can actually exactly simulate copies of Z using a clever different approach called the *polar method*. What is interesting about this method is that it requires the use of 2 iid $Unif(0,1)$ rvs and in return hands you back 2 iid copies of Z, $X = Z_1, Y = Z_2$.

Polar Method

Suppose that X and X are iid copies of $N(0,1)$. If we graph the vector (X,Y) in the Cartesian $x-y$ plane and then transform into polar coordinates,
$R^2 = X^2 + Y^2 \in \mathbb{R}_+, \Theta = \arctan(Y/X) \in [0, 2\pi)$ then from classical multi-dimensional calculus (compute the joint density of (R^2, Θ) by using the *Jacobian matrix/determinant* of the transformation $(x,y) \longrightarrow (x^2+y^2, \arctan(y/x)))$, it can be shown that

1. R^2 has an exponential distribution with mean 2 (hence rate 1/2).
2. Θ has a continuous uniform distribution over the interval $[0, 2\pi)$.
3. R^2 and Θ are independent random variables.

In other words, the joint density of (R^2, Θ) denote by $f(u,\theta)$ is given by a product

$$f(u,\theta) = \frac{1}{2}e^{-u/2}\frac{1}{2\pi}, \ u \geq 0, \ \theta \in [0, 2\pi).$$

Using the above facts in reverse we conclude that if R^2 has an exponential distribution with mean 2, and independently Θ has a continuous uniform distribution over the interval $[0, 2\pi)$, then (converting back into Cartesian coordinates), with radius $R = \sqrt{R^2}$, we have that the following 2 rvs X, Y are iid $N(0,1)$:

$$\begin{aligned} X &= R\cos\Theta \\ Y &= R\sin\Theta \end{aligned}$$

Letting U_1, U_2 be iid $Unif(0,1)$, we can generate our exponential via $R^2 = -2\ln(U_1)$ and our uniform via $\Theta = 2\pi U_2$ leading to

Polar Algorithm

1. Generate U_1, U_2.
2. Set $R^2 = -2\ln(U_1)$, $\Theta = 2\pi U_2$ and set $R = \sqrt{R^2}$.
3. Set

$$\begin{aligned} X &= R\cos\Theta \\ Y &= R\sin\Theta. \end{aligned}$$

4. Stop. Output X, Y.

1.3 Moment generating function of a normal distribution

In general, for a random variable X, the moment generating function (MGF) $M_X(s), \ s \in \mathbb{R}$ of X (or of its distribution) is defined by

$$M_X(s) \stackrel{\text{def}}{=} E(e^{sX}), \ s \in \mathbb{R}.$$

It is a function of $s \in \mathbb{R}$. If X has probability density function $g(x)$, then

$$M_X(s) = E(e^{sX}) = \int_{-\infty}^{\infty} e^{sx} g(x)dx.$$

For some rvs, the MGF might be infinite for some or all values of $s \neq 0$. For example, if X has an exponential distribution at rate λ, then (easily derived):

$$M_X(s) = \frac{\lambda}{\lambda - s}, \ s < \lambda,$$

and $M_X(s) = \infty$ for $s \geq \lambda$. Other examples typically yield an interval of the form $(-\epsilon, \epsilon)$ for which $M_X(s) < \infty, \ s \in (-\epsilon, \epsilon)$, for a sufficiently small $\epsilon > 0$, and infinite for other values. Some distributions, however, such as the normal distribution, have a finite MGF for all $s \in \mathbb{R}$; we will explicitly derive its MGF next.

Letting $X \sim N(\mu, \sigma^2)$, the moment generating function (MGF) of the normal distribution can be derived explicitly and the result is

$$\begin{aligned} M_X(s) &= E(e^{sX}) \\ &= \int_{-\infty}^{\infty} e^{sx} f(x) dx \\ &= e^{s\mu + s^2\sigma^2/2}, \ -\infty < s < \infty. \end{aligned} \tag{17}$$

Deriving (17): First we derive $M_Z(s) = e^{s^2/2}$, that is, the case when $X = Z$ is the unit normal.

$$\begin{aligned} M_Z(s) &= E(e^{sZ}) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} e^{\frac{-x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x^2-2sx)}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{s^2/2} e^{\frac{-(x-s)^2}{2}} dx \\ &= e^{s^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-s)^2}{2}} dx \\ &= e^{s^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} du, \ (u = x - s \text{ change of variables}) \\ &= e^{s^2/2} \int_{-\infty}^{\infty} \phi(u) du \\ &= e^{s^2/2} \times 1, \ (\phi(u) \text{ is a density function hence sums to 1}) \\ &= e^{s^2/2}. \end{aligned}$$

To obtain the general form in (17): If $X \sim N(\mu, \sigma^2)$, then it can be expressed as $X = \sigma Z + \mu$, and thus
$M_X(s) = E(e^{sX}) = e^{s\mu} E(e^{\sigma s Z})$
$= e^{s\mu} M_Z(\sigma s)$
$= e^{s\mu} e^{(\sigma s)^2/2}$
$= e^{s\mu + s^2\sigma^2/2}$; we have derived (17).

1.3.1 Application of moment generating functions to moments of the lognormal distribution

The moment generating results of the previous section will be used crucially when we study *geometric Brownian motion* later on. To prepare for that, let us introduce the *lognormal* *distribution:* Y is said to have a lognormal distribution if it is of the form $Y = e^X$ where $X \sim N(\mu, \sigma^2)$. The point is that by taking the natural logarithm of Y, we get back the normal distribution; $X = \ln(Y)$. More generally, $Y = ce^X$ for any constant $c > 0$ yields a lognormal distribution; $\ln(ce^X) = \ln(c) + X \sim N(\ln(c) + \mu, \sigma^2)$.

We can now use formula (17) to compute the expected value of a lognormal rv $Y = ce^X$:

$$E(Y) = cE(e^X) = cM_X(1) = ce^{\mu+\sigma^2/2}.$$

The point here is that by letting $s = 1$ in formula (17), we obtain our desired expected value. summarizing:

Proposition 1.1 *If $Y = ce^X$ is lognormal (with $c > 0$ and $X \sim N(\mu, \sigma^2)$), then*

$$E(Y) = ce^{\mu+\sigma^2/2}.$$

Further note that higher moments of Y can easily be derived in the same manner. For example, since $Y^2 = c^2e^{2X}$, we have

$$E(Y^2) = c^2E(e^{2X}) = c^2M_X(2) = c^2e^{2\mu+2\sigma^2};$$

we use value $s = 2$ in formula (17). This, for example, then allows us to compute the variance of Y:

$$Var(Y) = E(Y^2) - E^2(Y) = c^2e^{2\mu+2\sigma^2} - c^2e^{2\mu+\sigma^2} = c^2e^{2\mu+\sigma^2}(e^{\sigma^2} - 1).$$

In general, for the n^{th} moment, $E(Y^n) = c^nM_X(n) = c^ne^{n\mu+n^2\sigma^2/2}$.

1.3.2 Central limit theorem (CLT)

Theorem 1.1 *If $\{X_i : i \geq 1\}$ are iid with finite mean $E(X) = \mu \in \mathbb{R}$ and finite non-zero variance $\sigma^2 = Var(X)$, then*

$$Z_n \stackrel{\text{def}}{=} \frac{1}{\sigma\sqrt{n}}\Big(\sum_{i=1}^{n} X_i - n\mu\Big) \Longrightarrow N(0,1),\ \ n \to \infty,\ \ \textit{in distribution};$$

$\lim_{n\to\infty} P(Z_n \leq x) = \Phi(x),\ \ x \in \mathbb{R}$.

If $\mu = 0$ and $\sigma^2 = 1$, then the CLT becomes

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \Longrightarrow N(0,1).$$

Moreover, for any constant $c \neq 0$, since $cZ \sim N(0, c^2)$, we obtain:
If $\mu = 0$ and $\sigma^2 = 1$, then the CLT becomes for any constant $c \neq 0$

$$c\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \Longrightarrow N(0,c^2). \tag{18}$$

Remark 1.1 By defining the empirical average via

$$\overline{X}(n) = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

the Z_n can be re-written alternatively as

$$Z_n = \frac{\sqrt{n}}{\sigma}\Big(\overline{X}(n) - \mu\Big),$$

and the CLT is commonly written as

$$\frac{\sqrt{n}}{\sigma}\Big(\overline{X}(n) - \mu\Big) \Longrightarrow N(0,1),\ n \to \infty,\ \textit{in distribution.}$$

The Strong Law of Large Numbers (SLLN) ensures that, wp1, $\overline{X}(n) \to \mu$,
and hence $(\overline{X}(n) - \mu) \to 0$. Thus this alternative way of writing the CLT expresses the fact that by scaling $(\overline{X}(n) - \mu)$ by something tending to ∞, we get a proper limit; the correct scaling is $\sqrt{n}$.

1.4 Construction of Brownian motion from the simple symmetric random walk

Recall the simple *symmetric* random walk, $R_0 = 0$,

$$R_n = \Delta_1 + \cdots + \Delta_n = \sum_{i=1}^{n} \Delta_i,\ n \geq 1,$$

where the Δ_i are iid with $P(\Delta = -1) = P(\Delta = 1) = 0.5$. Thus $E(\Delta) = 0$ and $Var(\Delta) = E(\Delta^2) = 1$.

We view time n in minutes, and R_n as the position at time n of a particle, moving on the time line $\mathbb{R}$, which every minute takes a step, of size 1, equally likley to be forwards or backwards. Because $E(\Delta) = 0$ and $Var(\Delta) = 1$, it follows that $E(R_n) = 0$ and $Var(R_n) = n,\ n \geq 0$.

Choosing a large integer $k > 1$, if we instead make the particle still start at the origin but instead take a step every $1/k$ minutes and make the step size $1/\sqrt{k}$, then by time t the particle will have taken a large number, $n = tk$, of steps and its position will be

$$B_k(t) = \frac{1}{\sqrt{k}}\sum_{i=1}^{tk} \Delta_i\ t \geq 0, \tag{19}$$

with $B_k(0) = 0$. (By convention if tk is not an integer then we replace it by the largest integer less than or equal to it; denoted by $[tk]$.) This leads to the particle taking many many iid steps, but each of small magnitude, in any given interval of time. We expect that as $k \to \infty$, these small steps become a continuum and the process $\{B_k(t) : t \geq 0\}$ should converge to a process $\{B(t) : t \geq 0\}$ with continuous sample paths. We call this process Brownian motion (BM) after the Scottish botanist Robert Brown.[1] Its properties will be derived next. First note that since $B_k(0) = 0$ for any $k \geq 1$, we must end up with $B(0) = 0$ too.

[1] Brown himself noticed in 1827, while carrying out some experiments, the unusual "motion" of particles within pollen grains suspended in water, under his microscope. The physical cause of such motion (bombardment of the particles by water molecules undergoing thermal motion) was not formalized via kinetic theory until Einstein in 1905. The rigorous mathematical construction of a stochastic process as a model for such motion is due to the mathematician Norbert Weiner; that is why it is sometimes called a Weiner process.

Notice that for fixed k, any increment

$$B_k(t) - B_k(s) = \frac{1}{\sqrt{k}} \sum_{i=sk+1}^{tk} \Delta_i, \; 0 \le s < t,$$

has a distribution that only depends on the length, $t - s$, of the time interval $(s, t]$ because it only depends on the number, $k(t - s)$, of iid Δ_i making up its construction. Thus we deduce that the limiting process (as $k \to \infty$) will possess *stationary increments*: *The distribution of any increment $B(t) - B(s)$ has a distribution that only depends on the length of the time interval $t - s$. In particular, $B(t) - B(s)$ has the same distribution as does $B(t - s) = B(t - s) - B(0)$ (since $B(0) = 0$).*

Notice further that given two non-overlapping time intervals, $(t_1, t_2]$ and $(t_3, t_4]$, $0 \le t_1 < t_2 < t_3 < t_4$, the corresponding increments

$$B_k(t_4) - B_k(t_3) = \frac{1}{\sqrt{k}} \sum_{i=t_3k+1}^{t_4k} \Delta_i, \tag{20}$$

$$B_k(t_2) - B_k(t_1) = \frac{1}{\sqrt{k}} \sum_{i=t_1k+1}^{t_2k} \Delta_i, \tag{21}$$

are independent because they are constructed from different Δ_i. Thus we deduce that the limiting process (as $k \to \infty$) will also possess *independent increments*: *For any non-overlapping time intervals, $(t_1, t_2]$ and $(t_3, t_4]$, the increment rvs $I_1 = B(t_2) - B(t_1)$ and $I_2 = B(t_4) - B(t_3)$ are independent.*

Observing that $E(B_k(t)) = 0$ and $Var(B_k(t)) = [tk]/k \to t, \; k \to \infty$, we infer that the limiting process will satisfy $E(B(t)) = 0, \; Var(B(t)) = t$ just like the simple symmetric random walk $\{R_n\}$ does in discrete-time n ($E(R_n) = 0, \; Var(R_n) = n$).

Finally, a direct application of the CLT (using (18)) yields (via setting $n = tk, \; \mu = 0, \; \sigma^2 = 1, \; c = \sqrt{t}$)

$$B_k(t) = \sqrt{t}\Big(\frac{1}{\sqrt{kt}} \sum_{i=1}^{tk} \Delta_i\Big) \Longrightarrow N(0, t), \; k \to \infty, \; \textit{in distribution},$$

and we conclude that for each fixed $t > 0$, $B(t)$ has a normal distribution with mean 0 and variance t. Similarly, using the stationary and independent increments property, we conclude that $B(t) - B(s)$ has a normal distribution with mean 0 and variance $t - s$, and more generally:

> *the limiting BM process is a process with continuous sample paths that has both stationary and independent normally distributed (Gaussian) increments: If $t_0 = 0 < t_1 < t_2 < \cdots < t_n$, then the rvs. $B(t_i) - B(t_{i-1}), \; i \in \{1, \ldots n\}$, are independent with $B(t_i) - B(t_{i-1}) \sim N(0, t_i - t_{i-1})$.*

If, for a given fixed $\sigma > 0, \; \mu \in \mathbb{R}$, we define $X(t) = \sigma B(t) + \mu t$, then $X(t) \sim N(\mu t, \sigma^2 t)$, and we obtain, by such scaling and translation, more generally, a process with stationary and independent increments in which $X(t) - X(s)$ has a normal distribution with mean $\mu(t - s)$ and variance $\sigma^2(t - s)$.

When $\sigma^2 = 1$ and $\mu = 0$ (as in our construction) the process is called *standard Brownian motion*, and denoted by $\{B(t) : t \ge 0\}$. Otherwise, it is called Brownian motion with variance term σ^2 and *drift* μ.

Definition 1.1 *A stochastic process* $\mathbf{B} = \{B(t) : t \geq 0\}$ *possessing (wp1) continuous sample paths is called standard Brownian motion if*

1. $B(0) = 0$.
2. $\mathbf{B}$ *has both stationary and independent increments.*
3. $B(t) - B(s)$ *has a normal distribution with mean* 0 *and variance* $t - s,\ 0 \leq s < t$.

For Brownian motion with variance σ^2 and drift μ, $X(t) = \sigma B(t) + \mu t$, the definition is the same except that *3* must be modified to

$3'$. $X(t) - X(s)$ *has a normal distribution with mean* $\mu(t-s)$ *and variance* $\sigma^2(t-s)$.

Remark 1.2 It can in fact be proved that Condition *3* above is redundant: a stochastic process with stationary and independent increments that possesses (wp1) continuous sample paths must be Brownian motion, that is, the increments must be normally distributed. This is analogous to the Poisson counting process which is the unique simple *counting process* that has both stationary and independent increments: the stationary and independent increments property forces the increments to be Poisson distributed. (*Simple* means that the arrival times of the underlying point process are strictly increasing; no batches.)

Donsker's theorem

Our construction of Brownian motion as a limit is in fact a rigorous one, but requires more advanced mathematical tools (beyond the scope of these lecture notes) in order to state it precisely and to prove it. We have (due to the CLT) proved that fixed increments of $B_k(t)$ converge to a normal rv, but more generally it can be proved that the stochastic process $\{B_k(t) : t \geq 0\}$ as defined by (19) converges in distribution (weak convergence in path (function) space), as $k \to \infty$, to Brownian motion $\{B(t) : t \geq 0\}$. This is known as *Donsker's theorem* or the *functional central limit theorem.* The point is that it is a generalization of the central limit theorem, because it involves an entire stochastic process (with all its multi-dimensional joint distributions, for example) as opposed to just a one-dimensional limit (such as for fixed $t > 0$, $B_k(t) \to N(0, t)$ in distribution). Donsker's theorem implies, for example, that the vector $(B_k(t_1), \ldots, B_k(t_n))$ converges (jointly) in distribution to the vector $(B(t_1), \ldots, B(t_n))$: for any time points $0 \leq t_1 < t_2 < \cdots < t_n$ and $x_i \in \mathbb{R},\ 1 \leq i \leq n$, it holds that (as $k \to \infty$)

$$P(B_k(t_1) \leq x_1, \ldots, B_k(t_n) \leq x_n) \to P(B(t_1) \leq x_1, \ldots, B(t_n) \leq x_n).$$

1.5 Hitting times for standard BM

Consider two values $a > 0$, $b > 0$ and the two-point set $\{a, -b\}$. Let

$$\tau = \min\{t \geq 0 : B(t) \in \{a, -b)\}\},$$

the first time that BM hits either a or $-b$. (The continuity of sample paths implies that $B(t)$ either hits a or hits $-b$ (e.g, there is no overshoot/jumps).)

Proposition 1.2 *For standard BM, the probability that* a *is first hit before* $-b$ *is given by*

$$p_a = \frac{b}{a+b},\ a > 0, b > 0.$$

We offer a sketch of the proof (a more rigorous proof is provided by *martingale theory*): Recall from the gambler's ruin problem that for the simple *symmetric* random walk $\{R_n\}$, $p_a = \frac{b}{a+b}$, where $a > 0$, $b > 0$ (integers), and p_a denotes the probability that the random walk starting at $R_0 = 0$ first hits a before hitting $-b$.
Thus p_a denotes the probability that R_n goes up a steps before going down b steps.

For the process $\{B_k(t) : t \geq 0\}$ to hit a before $-b$ would require the random walk

$$R_n^{(k)} = \frac{1}{\sqrt{k}} \sum_{i=1}^{n} \Delta_i, \; n \geq 0$$

to hit a before $-b$. This would require (approximately) that the original random walk,

$$R_n = \sum_{i=1}^{n} \Delta_i, \; n \geq 0,$$

hits $a\sqrt{k}$ before $-b\sqrt{k}$; yielding the same answer

$$p_a = \frac{b\sqrt{k}}{a\sqrt{k} + b\sqrt{k}} = \frac{b}{a+b}.$$

We thus deduce (via letting $k \to \infty$) that the same holds for standard BM (where a and b need not be integers now).

One can also derive:

Proposition 1.3 *For standard BM, if $\tau = \min\{t \geq 0 : B(t) \in \{a, -b\} | B(0) = 0\}$, the first time that BM hits either a or $-b$, then*

$$E(\tau) = ab.$$

Note that if a variance term is introduced, $\sigma B(t)$, $\sigma > 0$, then $\sigma B(t) \in \{a, -b\}$ if and only if $B(t) \in \{a/\sigma, -b/\sigma\}$ yielding $E(\tau) = \frac{ab}{\sigma^2}$.

(In a later section, we will learn the corresponding (more complicated) formulas for BM with drift μ and variance parameter σ^2.)

Examples

1. A particle moves on a line according to a standard BM, $B(t)$. What is its expected position at time $t = 6$? What is the variance of its position at time $t = 6$?

   SOLUTION: $B(t)$ has a normal distribution with mean $E(B(t)) = 0$ and variance $Var(B(t)) = t$, hence the answers are 0 and 6.

2. *Continuation:*

   If the particle is at position 1.7 at time $t = 2$, what is its expected position at time $t = 4$?

   SOLUTION: $B(4) = B(2) + B(4) - B(2) = (B(2) - B(0)) + (B(4) - B(2))$, where the two increments are independent; $B(2)$ is independent of $(B(4) - B(2))$.

   $E(B(4)|B(2) = 1.7) = 1.7 + E(B(4) - B(2)|B(2) = 1.7)$
   $= 1.7 + E(B(4) - B(2))$ (independent increments)
   $= 1.7 + 0 = 1.7$, since all increments have mean 0, $E(B(t) - B(s)) = 0$.

3. *Continuation:*

   What is the probability that the particle hits level 10 before level -2? What is the expected length of time until either 10 or -2 are hit?

   SOLUTION: $a = 10$, and $b = 2$ in the formula for $p_a = b/(a+b) = 1/6$, and $E(\tau) = ab = 20$.

4. The price of a commodity moves according to a BM, $X(t) = \sigma B(t) + \mu t$, with variance term $\sigma^2 = 4$ and drift $\mu = -5$. Given that the price is 4 at time $t = 8$, what is the probability that the price is below 1 at time $t = 9$?

   SOLUTION:

   $P(X(9) < 1|X(8) = 4) = P(X(9) - X(8) < -3|X(8) = 4) = P(X(9) - X(8) < -3)$
   (independent increments, $X(9) - X(8)$ is independent of $X(8) - X(0) = X(8)$)
   $= P(X(1) < -3)$ (stationary increments)
   $= P(2Z - 5 < -3)$ (since $X(1) \sim N(-5, 4)$ can be represented in terms of a unit normal, Z, as $2Z - 5$)
   $= P(Z < 1) = \Theta(1) = 0.8413$ (via a Table for the standard normal distribution, as found in any statistics textbook for example).

5. A stock price per share moves according to *geometric BM*,

   $$S(t) = S_0 e^{B(t)},\ t \geq 0.$$

   Suppose that $S_0 = 4$, $S(t) = 4e^{B(t)}$. What is the probability that the stock price will reach a high of 7 before a low of 2?

   SOLUTION:

   Taking natural logarithms, we can convert the problem into *What is the probability that* $\ln(4) + B(t)$ *hits a high of* $\ln(7)$ *before a low of* $\ln(2)$*?*

   Because $\ln 7 - \ln 4 = \ln(7/4)$, this is equivalent to

   What is the probability that $B(t)$ *hits a high of* $\ln(7/4)$ *before a low of* $\ln(1/2)$*?*

   Noting that $\ln(1/2) = -\ln(2)$ we can set $a = \ln(7/4)$ and $b = \ln(2)$ in the formula $p_a = b/(a+b) = \ln(2)/(\ln(7/2))$.

1.6 BM as a Gaussian process

We observe that the vector $(B(t_1), \ldots, B(t_n))$ has a multivariate normal distribution because the event

$$\{B(t_1) = x_1, \ldots, B(t_n) = x_n\}$$

can be re-written in terms of independent increment events

$$\{B(t_1) = x_1, B(t_2) - B(t_1) = x_2 - x_1, \ldots, B(t_n) - B(t_{n-1}) = x_n - x_{n-1}\},$$

yielding the joint density of $(B(t_1), \ldots, B(t_n))$ as

$$f(x_1, \ldots, x_n) = f_{t_1}(x_1) f_{t_2 - t_1}(x_2 - x_1) \cdots f_{t_n - t_{n-1}}(x_n - x_{n-1}),$$

where

$$f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}}$$

is the density for the $N(0,t)$ distribution.

The finite dimensional distributions of BM are thus multivariate normal, or Gaussian, and BM is an example of a *Gaussian process*, that is, a process with continuous sample paths in which the finite dimensional distributions are multivariate normal, that is, for any fixed choice of n time points $0 \leq t_1 < t_2 < \cdots < t_n,\ n \geq 1$, the joint distribution of the vector $(X(t_1),\ldots,X(t_n))$ is multivariate normal.

Since a multivariate normal distribution is completely determined by its mean and covariance parameters, we conclude that a Gaussian process is completely determined by its mean and covariance function $m(t) \stackrel{\text{def}}{=} E(X(t)),\ a(s,t) \stackrel{\text{def}}{=} cov(X(s),X(t)),\ 0 \leq s \leq t.$

For standard BM, $m(t) = 0$ and $a(s,t) = s$:

$$\begin{aligned} cov(B(s),B(t)) &= cov(B(s),B(s)+B(t)-B(s)) \\ &= cov(B(s),B(s)) + cov(B(s),B(t)-B(s)) \\ &= var(B(s)) + 0 \textit{ (via independent increments)} \\ &= s. \end{aligned}$$

Thus standard BM is the unique Gaussian process with $m(t) = 0$ and $a(s,t) = \min\{s,t\}$. Similarly, BM with variance σ^2 and drift μ, $X(t) = \sigma B(t) + \mu t$, is the unique Gaussian process with $m(t) = \mu t$ and $a(s,t) = \sigma^2 \min\{s,t\}$.

1.7 BM as a Markov Processes

If B is standard BM, then the independent increments property implies that $B(s+t) = B(s) + (B(s+t) - B(s))$, in which $B(s)$ and $(B(s+t) - B(s))$ are independent. The independent increments property implies further that $(B(s+t) - B(s))$ is also independent of the past before time s, $\{B(u) : 0 \leq u < s\}$.

Thus the future, $B(s+t)$, given the present state, $B(s)$, only depends on a rv, $B(s+t) - B(s)$, that is independent of the past. Thus we conclude that BM satisfies the Markov property. Since the increments are also stationary, we conclude that BM is a time-homogenous Markov process.

Letting $p(x,t,y)$ denote the probability density function for $B(s+t) = y$ given $B(s) = x$, we see, from $B(s+t) = x + (B(s+t) - B(s))$, that $p(x,t,y)$ is the density for $x + B(s+t) - B(s)$. But $x + B(s+t) - B(s) = y$ if and only if $(B(s+t) - B(s)) = y - x$, yielding

$$p(x,t,y) = f_t(y-x) = \frac{1}{\sqrt{2\pi t}} e^{\frac{-(y-x)^2}{2t}}. \tag{22}$$

More generally, $X(t) = \sigma B(t) + \mu t$ is a Markov process with

$$p(x,t,y) = \frac{1}{\sigma\sqrt{2\pi t}} e^{\frac{-((y-x-\mu t)^2}{2\sigma^2 t}}. \tag{23}$$

1.8 BM as a martingale

Standard BM is a *martingale*:

$$E(B(t+s)|B(u) : 0 \leq u \leq s) = B(s),\ t \geq 0,\ s \geq 0,$$

which asserts that the conditional expectation of BM at any time in the future after time s equals the original value at time s. This of course follows from the independent increments property and using $B(s+t) = B(s) + (B(s+t) - B(s))$:

$$\begin{aligned} E(B(t+s)|B(u):0 \le u \le s) &= E(B(t+s)|B(s)), \textit{ via the Markov property} \\ &= E(B(s) + (B(s+t) - B(s))|B(s)) \\ &= B(s) + E(B(s+t) - B(s)|B(s)) \\ &= B(s) + E(B(s+t) - B(s)), \textit{ via independent increments} \\ &= B(s) + 0 \\ &= B(s). \end{aligned}$$

A martingale captures the notion of a fair game, in that regardless of your current and past fortunes, your expected fortune at any time in the future is the same as your current fortune: on average, you neither win nor lose any money.

The simple symetric random walk is a martingale (and a Markov chain) in discrete time;

$$E(R_{n+k}|R_n, \ldots, R_0) = E(R_{n+k}|R_n) = R_n, \; k \ge 0, \; n \ge 0,$$

because

$$R_{n+k} = R_n + \sum_{i=1}^{k} \Delta_{n+i},$$

and $\sum_{i=1}^{k} \Delta_{n+i}$ is independent of R_n (and the past before time n) and has mean 0.

1.9 Further results on hitting times

Let

$$T_x = \min\{t \ge 0 : B(t) = x \mid B(0) = 0\},$$

the hitting time to $x > 0$. From our study of the simple symmetric random walk, we expect $P(T_x < \infty) = 1$, but $E(T_x) = \infty$: although any level x will be hit with certainty, the mean length of time required is infinite. We will prove this directly and derive the cdf $P(T_x \le t)$, $t \ge 0$ along the way.

The key to our analysis is based on a simple observation involving the symmetry of standard BM: If $T_x < t$, then $B(s) = x$ for some $s < t$. Thus the value of $B(t)$ is determined by where the BM went in the remaining $t - s$ units of time after hitting x. But BM, having stationary and independent Gaussian increments, will continue having them after hitting x (strong Markov property). So by symmetry (about x), the path of BM during the time interval $(s, t]$ with $B(s) = x$ is just as likely to lead to $B(t) > x$ as to $B(t) < x$. So the events $\{B(t) > x | \; T_x \le t\}$ and $\{B(t) < x | \; T_x \le t\}$ are equally likely; both have probability 1/2. ($P(B(t) = x) = 0$ since $B(t)$ has a continuous distribution.) To be precise, if $T_x = s < t$, then $B(t) = x + B(t) - B(s)$ which has the $N(x, t-s)$ distribution (which is symmetric about x). Thus $P(B(t) > x \mid T_x \le t) = 1/2$. On the other hand $P(B(t) > x \mid T_x > t) = 0$ because BM (having continuous sample paths) can not be above x at time t if it never hit x prior to t. Summarizing yields

$$\begin{aligned} P(B(t) > x) &= P(B(t) > x \mid T_x \le t)P(T_x \le t) + P(B(t) > x \mid T_x > t)P(T_x > t) \\ &= P(B(t) > x \mid T_x \le t)P(T_x \le t) + 0 \\ &= \frac{1}{2}P(T_x \le t), \end{aligned}$$

or

$$P(T_x \le t) = 2P(B(t) > x) = \frac{2}{\sqrt{2\pi t}} \int_x^\infty e^{\frac{-y^2}{2t}} dy,$$

because $B(t) \sim N(0,t)$. Changing variables $u = y/\sqrt{t}$ then yields

Proposition 1.4 *For standard BM, for any fixed $x \neq 0$*

$$P(T_x \le t) = 2P(B(t) > x) = \frac{2}{\sqrt{2\pi}} \int_{\frac{x}{\sqrt{t}}}^\infty e^{\frac{-y^2}{2}} dy = 2(1 - \Phi(x/\sqrt{t})), \ t \ge 0.$$

In particular T_x is a proper random variable; $P(T_x < \infty) = 1$.

$P(T_x < \infty) = 1$ because taking the limit as $t \to \infty$ yields $P(T_x < \infty) = 2(1 - \Phi(0)) = 2(1 - 0.5) = 1$.

Corollary 1.1 *For standard BM, for any fixed $x \neq 0$, $E(T_x) = \infty$.*

Proof : We shall proceed by computing $E(T_x) = \infty$ by integrating the tail $P(T_x > t)$;

$$E(T_x) = \int_0^\infty P(T_x > t) dt.$$

To this end, $P(T_x > t) = 1 - P(T_x \le t) = \frac{2}{\sqrt{2\pi}} \int_0^{\frac{x}{\sqrt{t}}} e^{\frac{-y^2}{2}} dy$. Since the constant factor $\frac{2}{\sqrt{2\pi}}$ plays no role in wether the integrated tail is infinite or finite, we leave it out for simplicity. It thus suffices to show that

$$\int_0^\infty \int_0^{\frac{x}{\sqrt{t}}} e^{\frac{-y^2}{2}} dy dt = \infty.$$

Changing the order of integration, we re-write as

$$\begin{aligned}
\int_0^\infty \int_0^{\frac{x^2}{y^2}} e^{\frac{-y^2}{2}} dt dy &= x^2 \int_0^\infty \frac{1}{y^2} e^{\frac{-y^2}{2}} dy \\
&\ge x^2 \int_0^1 \frac{1}{y^2} e^{\frac{-y^2}{2}} dy \\
&\ge x^2 e^{-1/2} \int_0^1 \frac{1}{y^2} dy \\
&= \infty.
\end{aligned}$$

The second inequality is due to the fact that the decreasing function $e^{\frac{-y^2}{2}}$ is minimized over the interval $(0, 1]$ at the end point $y = 1$.

■

Let $M_t \stackrel{\text{def}}{=} \max_{0 \le s \le t} B(t)$ denote the maximum value of BM up to time t. Noting that $M_t \ge x$ if and only if $T_x \le t$, we conclude that $P(M_t \ge x) = P(T_x \le t)$ yielding (from Proposition 1.4) a formula for the distribution of M_t:

Corollary 1.2 *For standard BM, for any fixed $t \ge 0$,*

$$P(M_t > x) = \frac{2}{\sqrt{2\pi}} \int_{\frac{x}{\sqrt{t}}}^\infty e^{\frac{-y^2}{2}} dy = 2(1 - \Phi(x/\sqrt{t})), \ x \ge 0.$$

1.10 Hitting times for BM with drift

For $X(t) = \sigma B(t) + \mu t$, let's assume that $\mu < 0$ so that the BM has *negative drift*. This is analogous to the simple random walk with negative drift, that is, $\{R_n\}$ when the increments have distribution $P(\Delta = 1) = p,\ P(\Delta = -1) = q = 1 - p$ and $q > p$. Recall from the gambler's ruin problem that in this case

$$p_a = \frac{1 - (p/q)^b}{(p/q)^{-a} - (p/q)^b},$$

and thus by letting $b \to \infty$ we obtain the probability that the random walk will ever reach at least as high as level a;

$$P(\max_{n \geq 0} R_n \geq a) = \lim_{b \to \infty} p_b = (p/q)^a.$$

We conclude that the maximum of the random walk has a geometric distribution with "success" probability $1 - p/q$. The point is that the negative drift random walk will eventually drift off to $-\infty$, but before it does there is a positive probability, $(p/q)^a$, that it will first reach a (finite) level $\geq a > 0$.

$X(t)$ is similar. We let $M = \max_{t \geq 0} X(t)$ denote the maximum of the BM:

Proposition 1.5 *For BM with negative drift, $X(t) = \sigma B(t) + \mu t,\ \mu < 0$,*

$$p_a = \frac{1 - e^{-\alpha b}}{e^{\alpha a} - e^{-\alpha b}},$$

where $\alpha = 2|\mu|/\sigma^2$; thus (letting $b \to \infty$)

$$P(M > a) = e^{-\alpha a},\ a \geq 0,$$

and we conclude that M has an exponential distribution with mean $\alpha^{-1} = \sigma^2/2|\mu|$.

In general, $\mu > 0$ or $\mu < 0$, the formula for p_a is

$$p_a = \frac{1 - e^{(2\mu/\sigma^2)b}}{e^{(-2\mu/\sigma^2)a} - e^{(2\mu/\sigma^2)b}}.$$

Proof :

Here we use an exponential martingale of the form

$$e^{\lambda X(t) - (\lambda\mu + \frac{1}{2}\lambda^2\sigma^2)t}.$$

This is a MG for any value of λ. Choosing $\lambda = \alpha = -2\mu/\sigma^2$, so that the second term in the exponent vanishes, we have the MG

$$U(t) = e^{\alpha X(t)}.$$

Then for $\tau = \min\{t \geq 0 : X(t) \in \{a, -b\} | X(0) = 0\}$, we use optional sampling to obtain $E(Y(\tau)) = 1$ or $e^{\alpha a}p_a + e^{-\alpha b}(1 - p_a)$; solving for p_a yields the result. ($U(t \wedge \tau)$ is bounded hence UI.)

■

We also have (proof left out):

Proposition 1.6 *For BM with drift, $X(t) = \sigma B(t) + \mu t$, $\mu \neq 0$, if $\tau = \min\{t \geq 0 : X(t) \in \{a, -b\} | X(0) = 0\}$, then*

$$E(\tau) = \frac{a(1 - e^{\frac{2\mu b}{\sigma^2}}) + b(1 - e^{\frac{-2\mu a}{\sigma^2}})}{\mu(e^{\frac{-2\mu a}{\sigma^2}} - e^{\frac{2\mu b}{\sigma^2}})}.$$

What about T_x? If the drift is negative, then we already know that for $x > 0$, the BM might not ever hit x; $P(T_x = \infty) = P(M < x) > 0$. But if the drift is positive, x will be hit with certainty (because this is so even when $\mu = 0$; Proposition 1.4). In this case the mean is finite (proof not given here) :

Proposition 1.7 *For BM with positive drift,$X(t) = \sigma B(t) + \mu t$, $\mu > 0$, if $T_x = \min\{t \geq 0 : X(t) = x | X(0) = 0\}$, then*

$$E(T_x) = \frac{x}{\mu},\ x > 0.$$

Note how, as $\mu \to 0$, $E(T_x) \to \infty$, and this agrees with our previous calculation (Corollary 1.1) that $E(T_x) = \infty$ when $\mu = 0$ (even though $P(T_x < \infty) = 1$).