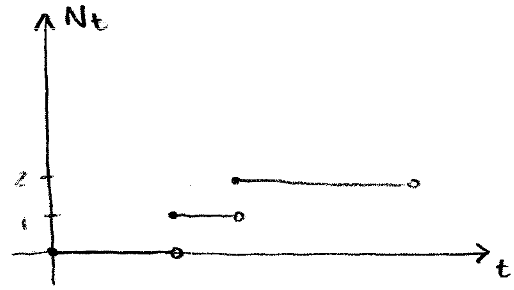
Brownian motion

Read Taylor & Karlin pp. 473 – 475.

Brownian motion is best thought of in terms of 3 qualitative axioms, just like the Poisson process.

Poisson process



. State space $S = \{0, 1, 2, 3, \ldots\}$, $N_0 = 0$.

(i) N increases, by jumps of size 1

(ii) $N_{t+s} - N_t$ is independent of N_r, $r \le t$

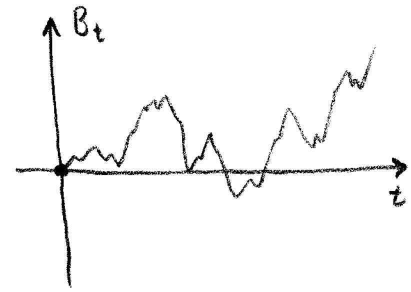
(iii) distribution of $N_{t+s} - N_t$ doesn't depend on t.

Then we can show that N_t has a Poisson distribution; in fact $N_t \sim \text{Poisson}(\lambda t)$ for some λ in $(0, \infty)$.

First we used axiom (ii) to show that N is a Markov process, then axiom (iii) to show that it is time homogeneous.

Brownian motion

Stochastic process B_t, $t \ge 0$, state space $\mathbb{R}$, $B_0 = 0$.



(i) realizations of B are continuous
that is, $t \mapsto B_t$ is always continuous

(ii) $B_{t+s} - B_t$ is independent of B_r, $r \le t$

(iii) distribution of $B_{t+s} - B_t$ doesn't depend on t

As before, it is easy enough to show that B is a Markov process:

$$P(B_{t+s} \in A \mid B_t = x, B_r, r \le t) = \mathbb{P}(B_{t+s} \in A \mid B_t = x)$$

Also, note that for any integer $n > 0$,

$$B_1 = B_{\frac{1}{n}} + \left(B_{\frac{2}{n}} - B_{\frac{1}{n}}\right) + \left(B_{\frac{3}{n}} - B_{\frac{2}{n}}\right) + \cdots + \left(B_1 - B_{\frac{n-1}{n}}\right),$$

so that B_1 is the sum of n iid random variables. As long as these have finite variance, as $n \to \infty$ the distribution of the sum becomes normal.

Thus, B_1 is <u>normally distributed</u>!

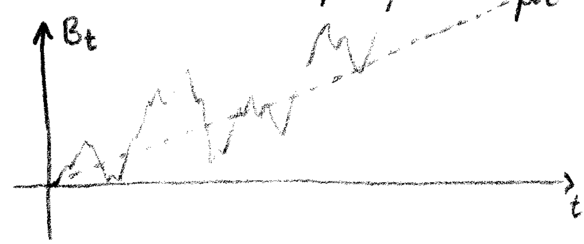
In fact, for some parameters μ and σ^2, $B_1 \sim N(\mu, \sigma^2)$.

And in the same way, $B_t \sim N(\mu t, \sigma^2 t)$; normal by the same argument as before, and linear mean because

$$\begin{aligned} m(t+s) = \mathbb{E}\, B_{t+s} &= \mathbb{E}\, B_{t+s} - B_t + B_t \\ &= \mathbb{E}(B_{t+s} - B_t) + \mathbb{E}(B_t) \\ &= \mathbb{E}\, B_s + \mathbb{E}\, B_t \qquad \text{axiom (iii)} \\ &= m(s) + m(t). \end{aligned}$$

Thus $m(t) = \mu t$ for some μ.
The fact that the variance is linear in t is shown in the same way.

The parameter μ is called the drift, since $\mathbb{E} B_t = \mu t$.



The parameter σ^2 is called the diffusion coefficient.

The name comes from the fact that Brownian motion was originally used to model molecular diffusion, and σ^2 measures how quickly/violently the molecule moves around.

Brownian motion has many good things going for it:

postpone

1. continuous paths
2. B_t has a normal distribution
3. independent increments
4. B is a Markov process
5. B is a martingale.

A Brownian motion with drift μ and diffusivity coefficient σ^2 can always be written as

$$B_t = \mu t + \sigma W_t.$$

$\rightarrow$ HW.

Special case

When $\mu = 0$ and $\sigma^2 = 1$, B is called the Wiener process, W, in honor of Norbert Wiener, who first showed that Brownian motion exists. It is as important as the standard normal r.v.

Recap

$B_t,\ t \ge 0$ is a Brownian motion if

(i) • realizations are continuous

(ii) • $B_{t+r} - B_t \perp\!\!\!\perp B_r,\ r \le t$

(iii) • dist'n of $B_{t+r} - B_t$ does not dep. on t.

$N_t,\ t \ge 0$ is a Poisson process if

(i) • N increases by jumps of size 1 only

(ii) •
(iii) • } same

Claim Then B_t is normally dist'd for each t

N_t has the Poisson dist'n for each t.

Proof of normality

Reads like a proof of the Central Limit Theorem, using characteristic functions.

Write $B_t = B_{\frac{t}{n}} + \left(B_{\frac{2t}{n}} - B_{\frac{t}{n}}\right) + \cdots + \left(B_t - B_{\frac{(n-1)t}{n}}\right)$.

Then B_t is the sum of n iid r. variables.

☆ Key: Because B is continuous, all but a few of these increments will be small, less than some fixed number ε.

The sum of the ones smaller than ε is a good approximation to B_t, and good enough as $n \to \infty$.

This is quite different from the Poisson case, in which most increments are 0 and a few are 1 or 2 or 3 (large numbers).

$B_t \sim N(\mu t, \sigma^2 t)$ μ = drift, σ^2 = diff. coeff.

When $\mu = 0$, $\sigma^2 = 1$, B is written W, Wiener process. Enough to study this

12-10-2000

Recap

B_t, $t \ge 0$ is a Brownian motion if

(i) • realizations are continuous

(ii) • $B_{t+r} - B_t$ $\perp\!\!\!\perp$ B_r, $r \le t$

(iii) • dist'n of $B_{t+r} - B_t$ does not dep. on t.

N_t, $t \ge 0$ is a Poisson process if

(i) • N increases by jumps of size 1 only

(ii) •
(iii) • } same

Claim Then B_t is normally dist'd for each t

N_t has the Poisson dist'n for each t.

Proof of normality

Reads like a proof of the Central Limit Theorem, using characteristic functions.

Write $B_t = B_{\frac{t}{n}} + \left(B_{\frac{2t}{n}} - B_{\frac{t}{n}}\right) + \cdots + \left(B_t - B_{\frac{(n-1)t}{n}}\right)$.

Then B_t is the sum of n iid r. variables.

★ Key: Because B is continuous, all but a few of these increments will be small, less than some fixed number ε.

The sum of the ones smaller than ε is a good approximation to B_t, and good enough as $n \to \infty$.

This is quite different from the Poisson case, in which most increments are 0 and a few are 1 or 2 or 3 (large numbers).

$B_t \sim N(\mu t, \sigma^2 t)$ μ = drift, σ^2 = diff. coeff.
When $\mu = 0$, $\sigma^2 = 1$, B is written W, Wiener process. Enough to study this

Pf. that $B_1 \sim N(\mu, \sigma^2)$. (Çinlar).

$$B_1 = B_{\frac{1}{n}} + (\quad) + \cdots + (B_1 - B_{\frac{n-1}{n}})$$

Most of there are small by continuity of B.
Y_m = same sum, set to 0 if term is too large.

Write $f(\alpha) = \mathbb{E}\, e^{i\alpha B_1}$

$$= \lim \mathbb{E}\, e^{i\alpha Y_m}$$

where Y_m is the sum of the ones that are small.

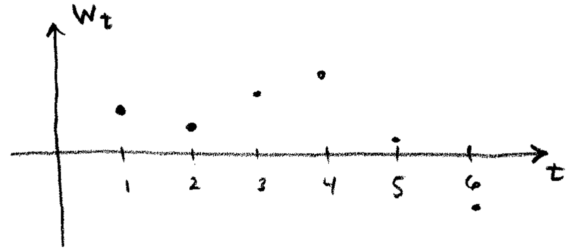
Show that

Characterstic function of B_1 ---- must be that of normal.

---

* Variance of each increment goes to 0 quickly enough, like $(\frac{1}{n})^2$,
whereas in Poisson case, var. of each increment goes like $\frac{1}{n}$.

Numerical simulation of the Wiener process



This is not difficult.

$W_0 = 0$

$W_1 \sim N(0,1)$, so generate a standard normal for it.

$W_2 - W_1$ is independent of W_1, and also $N(0,1)$, so generate a standard normal and add it to W_1 to get W_2.

$W_3 = W_2 + (W_3 - W_2)$ add another independent st. normal.

continue. to time N.

To make it look like a continuous process, connect the dots.

As long as N is large enough, you won't notice that you just connected the dots.

(To get a zoomed-in picture, $W_{\Delta t} = N(0, \Delta t) = \sqrt{\Delta t}\ N(0,1)$

$W_{2\Delta t} = W_{\Delta t} + N(0, \Delta t) = W_{\Delta t} + \sqrt{\Delta t}\ N(0,1)$

$\vdots$

this is just more indexing trouble.)

Contrast with symmetric random walk

Let X_n, $n=0,1,2,\ldots$ be a Markov chain on $\mathbb{Z}$ with

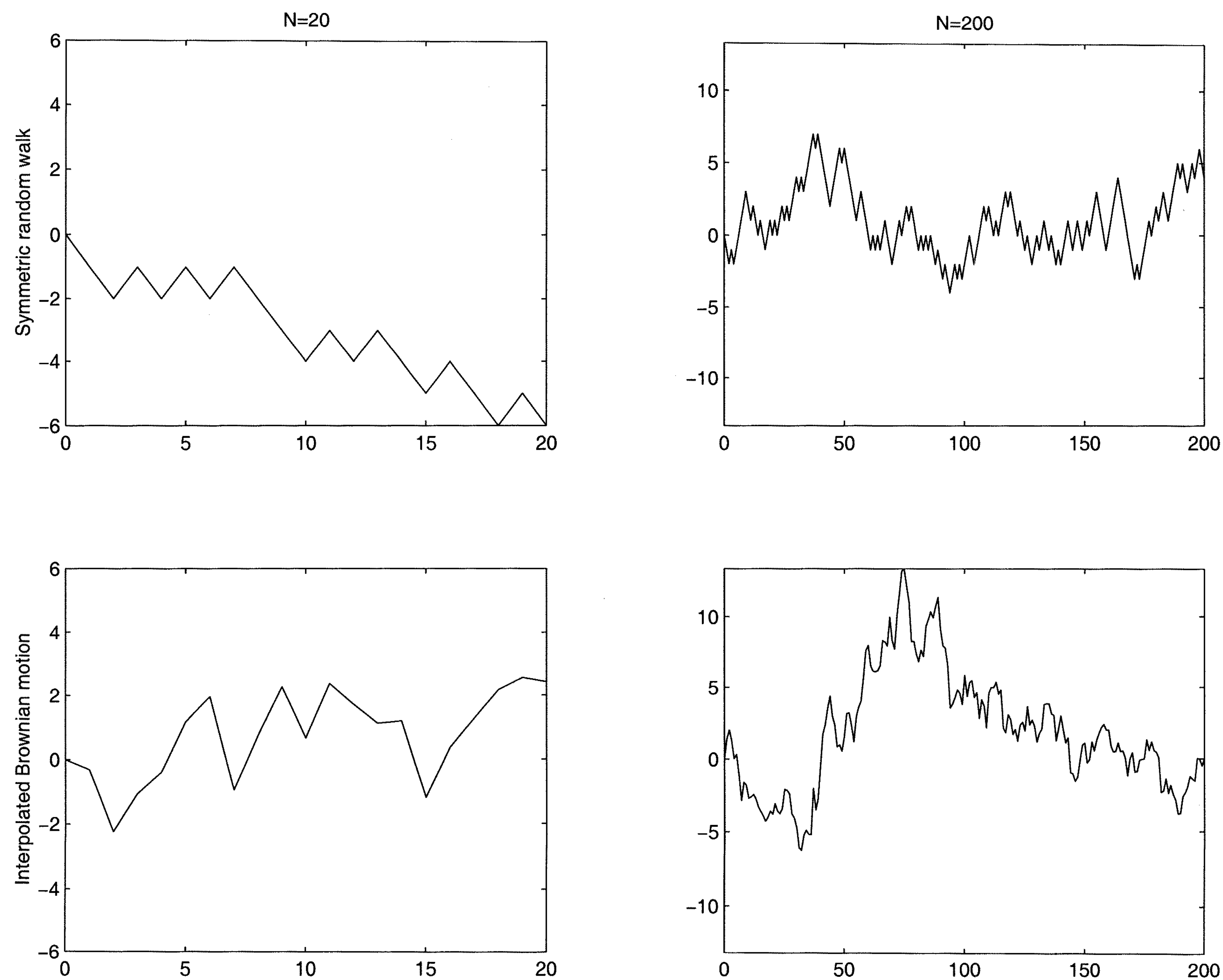
$\mathbb{P}(X_{n+1} = i \pm 1 \mid X_n = i) = \frac{1}{2}$,

Then $X_{n+1} = X_n + A_n$, $n = 0,1,2,\ldots$

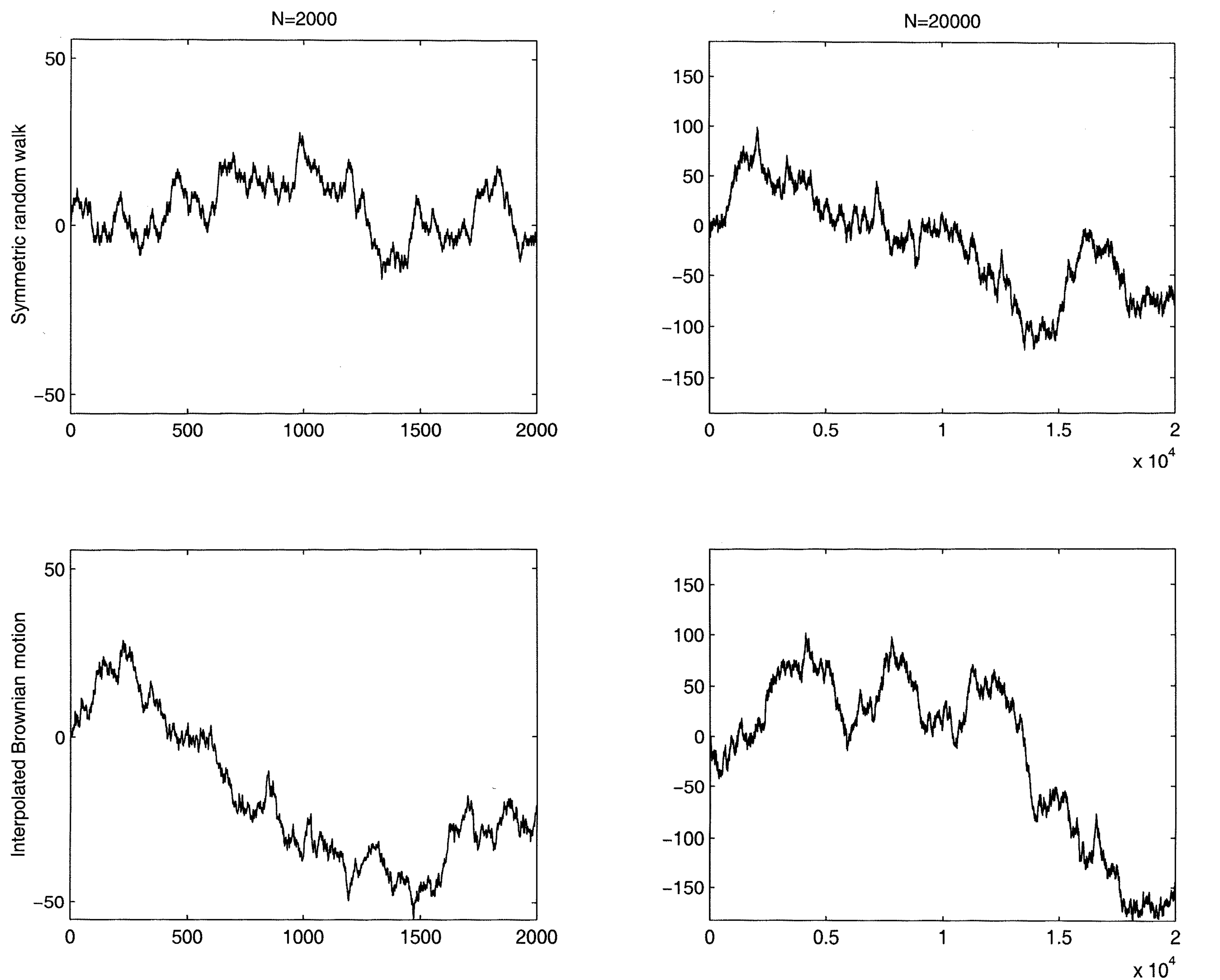
where $A_0, A_1, A_2, \ldots$ are iid r.v. taking values -1 and 1 w/p $\frac{1}{2}$.

You can easily see that
$\mathbb{E}\, X_n = 0$ (like W)
$Var(X_n) = n$ (like W)
$X_n \sim 2 \cdot Bin(n, \frac{1}{2}) - n$ (close to normal!)

Simulate X in the same way as W, up to time N.
They look different for small N, but the same for large N.
In every way, Brownian motion is the continuous-time version of the symmetric random walk. $\rightarrow$ This satisfies the same 3 axioms, essentially, just in discrete time.



N=20
N=200
Symmetric random walk
Interpolated Brownian motion


N=2000
N=20000
Symmetric random walk
Interpolated Brownian motion
x 10^4

Rescaling the Wiener process

Let W be a Wiener process and fix $c \in \mathbb{R}$.

Let $\hat{W}_t = \frac{1}{c} W_{c^2 t}$, $t \geq 0$.

Then $\hat{W}$ is also a Wiener process.

Check the axioms and normalization:

(i) realizations of $\hat{W}$ are continuous: $t \longmapsto \hat{W}_t$ is continuous.

(ii) $\hat{W}_{t+s} - \hat{W}_t$ is independent of $\hat{W}_r$, $r \leq t$

$\frac{1}{c} W_{c^2(t+s)} - \frac{1}{c} W_{c^2 t}$ is independent of $\frac{1}{c} W_{c^2 r}$, $r \leq t$.

(iii) distribution of $\hat{W}_{t+s} - \hat{W}_t = \frac{1}{c}\left(W_{c^2(t+s)} - W_{c^2(t)}\right)$

does not depend on t

Thus, $\hat{W}$ is a Brownian motion.

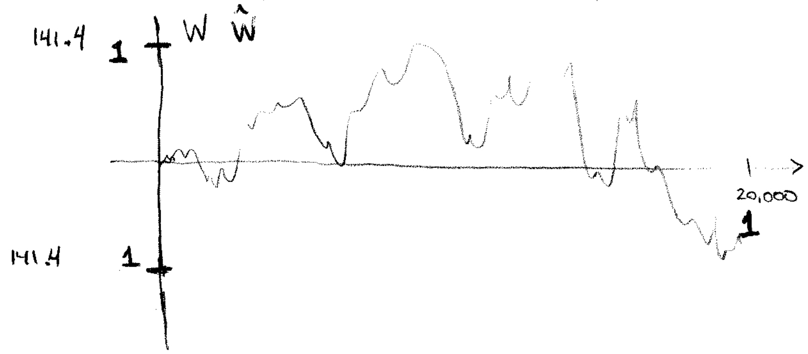
And $$E\,\hat{W}_t = \mathbb{E}\,\frac{1}{c} W_{c^2 t} = 0$$

$$\mathrm{Var}(\hat{W}_t) = \mathrm{Var}\left(\frac{1}{c} \cdot W_{c^2 t}\right) = \frac{1}{c^2} \cdot \mathrm{Var}\left(W_{c^2 t}\right) = \frac{1}{c^2} \cdot c^2 t = t.$$

Thus, $\hat{W}$ is a Brownian motion.

Example Let $c^2 = 20{,}000$.

Then W_t, graphed over $[0, 20{,}000]$, looks like



And $\hat{W}_t = \frac{1}{\sqrt{20{,}000}} W_{20{,}000\,t}$ graphed over $[0,1]$

looks exactly the same, but with the axes relabeled.

To be more precise, a Wiener process graphed over t in $[0, c^2]$ with $-c$ to c on the vertical axis looks the same (probabilistically) as a Wiener process graphed over t in $[0,1]$ with -1 to 1 on the vertical axis!

Gaussian processes

A stochastic process X is said to be a Gaussian process if, for all n, all $t_1 \dots t_n$, the random vector $(X_{t_1}, \dots, X_{t_n})$ is multivariate Gaussian (normal).

Equivalently, for all n, all $t_1 \dots t_n$, and all $\alpha_1 \dots \alpha_n$,

$$\alpha_1 X_{t_1} + \cdots + \alpha_n X_{t_n} \text{ is normally distributed.}$$

Gaussian processes are nice because it is enough to know the means, variances, and covariances of X_s and X_t for all s and t in order to know the whole probabilistic description of the process.

It follows from the proof that B_t is normally distributed that Brownian motion is a Gaussian process. That says a lot. The only thing left to say is that

$$\mathbb{E}\, B_t = \mu t$$

$$\mathrm{Var}(B_t) = \sigma^2 t$$

$$\mathrm{Cov}(B_s, B_t) = \sigma^2 \min(s, t).$$

Non-differentiability

When is a function differentiable at a?

If you zoom in enough, the function looks linear,
you can write it as $f(x) = f(a) + m(x-a) + o(x-a)$.

With the Wiener process, when you zoom in, by the rescaling property, what you see looks just like the Wiener process on $[0,1]$ (but random, of course).

No matter how much you zoom in around a, the graph will not look linear, so W is not differentiable at a.

In fact, the non-differentiability is even stronger than that.

You might think that, given a realization of a Wiener process, you could choose a place, a, at which that realization is differentiable.

That is, maybe W is differentiable at certain randomly-chosen points.

But in fact, with probability 1, each realization of W is nowhere differentiable!

This property is shared by Brownian motion with general drift and diffusivity coefficient, and by the Ornstein-Uhlenbeck process.

Transition density of Brownian motion

We have seen that $B_t,\ t \geq 0$ is a Markov process:

$$\mathbb{P}(B_{t+s} \in A \mid B_t = x,\ B_r,\ r < t)$$
$$= \mathbb{P}(B_{t+s} \in A \mid B_t = x)$$
$$= \mathbb{P}(B_s \in A \mid B_0 = x) \qquad \text{time homogeneity.}$$

(This conditional probability is a bit strange because we always have $B_0 = 0$. But that is a matter of what initial distribution μ_0 we use; we always have μ_0 concentrated at 0. But that does not prevent us from considering this transition probability.)

We know that $B_s - B_0 \sim N(\mu s, \sigma^2 s)$, so $B_s \sim N(\mu s + x, \sigma^2 s)$,

$$\mathbb{P}(B_s \in A \mid B_0 = x)$$
$$= \int_A \frac{1}{\sqrt{2\pi \sigma^2 s}} \exp\left(- \frac{(y - (x + \mu s))^2}{2\sigma^2 s} \right) dy$$
$$= \int_A p_s(x,y)\, dy$$

The function p is called the transition density of B.

You can verify directly that p satisfies, for each fixed x,

$$\frac{\partial p_t(x,y)}{\partial t} = -\mu \frac{\partial p_t(x,y)}{\partial y} + \frac{1}{2}\sigma^2 \frac{\partial^2 p_t(x,y)}{\partial y^2}$$
$$= (A^* p_t)(x,y), \qquad p_0(x,\cdot) = \delta\text{-function at } x$$

where A^* is the adjoint of the generator A of B.

The generator itself is a differential operator,

$$(Af)(x) = \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2}.$$

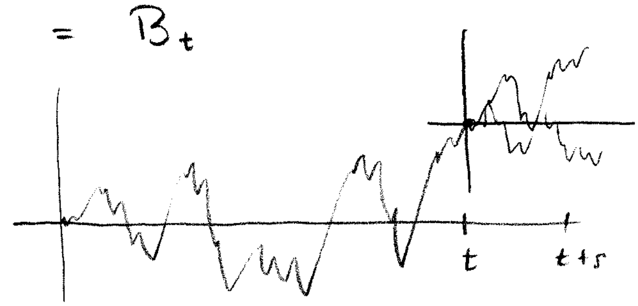
This generalizes the notion of A being a matrix.
Matrices are often used to approximate differential operators.

 12-12-2000

Martingale property

When $\mu = 0$, no drift, Brownian motion has what is called the martingale property:

$$\mathbb{E}\left[B_{t+s} \mid B_t, B_r, r < t \right] = B_t$$



Even given all the past information, there is no future trend to predict.

Ornstein – Uhlenbeck process

Brownian motion starts at 0, and its variance increases linearly from 0. The distribution of B_t is more and more spread out as t increases. There is no sense in which B_t approaches a limiting distribution as $t \to \infty$.

Can we find a process which:

- is Markov
- has continuous paths
- is Gaussian
- is stationary ?

The answer is yes.

Let W be a Wiener process and let V_0 be a standard normal, independent of W.

Set $V_t = e^{-t} W_{e^{2t}}$, for all $t \geq 0$.

This is called the Ornstein – Uhlenbeck process. (simpler than homework!)

Continuity

For each realization of W, the corresponding realization of V is obtained by composition with e^{2t} and multiplication by e^{-t}. Then V has continuous realizations as well.

Gaussian

Fix n, $t_1 \dots t_n$.

Then $(V_{t_1}, \dots, V_{t_n}) = (e^{-t_1} W_{e^{2t_1}}, \dots, e^{-t_n} W_{e^{2t_n}})$,

which is multivariate normal.

Thus, V is a Gaussian process.

Mean: $\mathbb{E} V_t = \mathbb{E}\, e^{-t} W_{e^{2t}} = 0$

Covariance:
$$\begin{aligned}
\mathbb{E} V_s V_t &= \mathbb{E}\, e^{-s} W_{e^{2s}}\, e^{-t} W_{e^{2t}} \\
&= e^{-s-t}\, \mathbb{E}\, W_{e^{2s}} W_{e^{2t}} \\
&= e^{-s-t} \min(e^{2s}, e^{2t}) \\
&= e^{-s-t} e^{2\min(s,t)} \\
&= e^{-|t-s|} \quad \text{by considering cases.}
\end{aligned}$$

In particular, $\mathrm{Var}(V_t) = e^{-|t-t|} = 1.$

Markov property

$$\mathbb{P}(V_{t+s} \le a \mid V_t = x,\ V_r,\ r<t)$$

$$= \mathbb{P}(e^{-t-s} W_{e^{2(t+s)}} \le a \mid e^{-t} W_{e^{2t}} = x,\ e^{-r} W_{e^{2r}},\ r<t)$$

$$= \mathbb{P}(W_{e^{2(t+s)}} \le a e^{t+s} \mid W_{e^{2t}} = x e^{t},\ e^{-r} W_{e^{2r}},\ r<t)$$

$$= \mathbb{P}(W_{e^{2(t+s)}} \le a e^{t+s} \mid W_{e^{2t}} = x e^{t}) \quad \text{since } W \text{ is Markov}$$

(so V is Markov)

$$= \mathbb{P}(e^{-(t+s)} W_{e^{2(t+s)}} \le a \mid e^{-t} W_{e^{2t}} = x) \quad \text{rewrite}$$

$$= \mathbb{P}(e^{-s} \cdot \tfrac{1}{c} W_{c^2 e^{2s}} \le a \mid \tfrac{1}{c} W_{c^2} = x) \quad c = e^t$$

$$= \mathbb{P}(e^{-s} W_{e^{2s}} \le a \mid W_1 = x) \quad \text{rescaling property of the Wiener process.}$$

$$= \mathbb{P}(V_s \le a \mid V_0 = x)$$

This does not depend on t, so V is time homogeneous.

Stationarity

V_0 is $N(0,1)$ and V_t is $N(0,1)$ for all t, so V is a stationary Markov process.

Simulations

$V_t = e^{-t}(V_0 + W_{e^{2t}-1})$. — easier to start with this representation. V_0 is independent of W.

Suppose $V_0 = x$.

Then $\mathbb{E} V_t = e^{-t} \cdot x$

$$\mathrm{Var}(V_t) = \mathrm{Var}\left(e^{-t}(x + W_{e^{2t}-1})\right)$$
$$= e^{-2t} \cdot (e^{2t} - 1)$$
$$= 1 - e^{-2t}.$$

Then $V_t = e^{-t} \cdot x + \sqrt{1-e^{-2t}} \cdot Z$,

where Z is $N(0,1)$.

To make this a simulation scheme,

$$V_t = e^{-t} V_0 + \sqrt{1-e^{-2t}} \cdot Z$$

Fix a timestep Δt; make it small.

Fix V_0.

$$V_{\Delta t} = e^{-\Delta t} V_0 + \sqrt{1-e^{-2\Delta t}}\, Z$$
$$\approx (1-\Delta t) V_0 + \sqrt{1-(1-2\Delta t)}\, Z$$
$$= V_0 - \Delta t\, V_0 + \sqrt{2\Delta t} \cdot Z$$

randomness

drift toward 0, size proportional to V_0

original state