# Stochastic Processes: Lecture 2

Andrew J. Holbrook

UCLA Biostatistics 270

Spring 2022

# Brownian motion

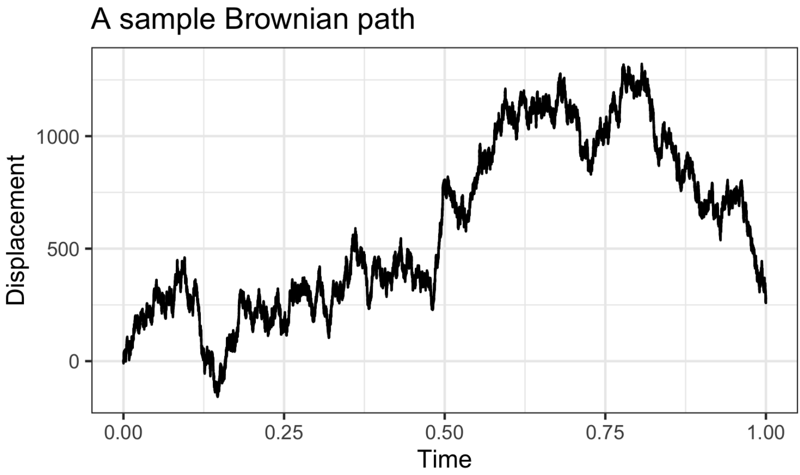
Theorem 1
*There exists a probability distribution over the set of continuous functions* $B : \mathbb{R} \to \mathbb{R}$ *satisfying the following conditions:*

(i) $B(0) = 0$

(ii) *for all* $0 \leq s < t$, $B(t) - B(s) \sim N(0, t - s)$.

(iii) $B(t_i) - B(s_i) \perp B(t_j) - B(s_j)$ *for* $s_i < t_i \leq s_j < t_j$.

Item (ii) is stationarity, where $t - s$ is the variance. Item (iii) is independence over non-overlapping increments.


A sample Brownian path
Displacement
1000
500
0
0.00
0.25
0.50
0.75
1.00
Time

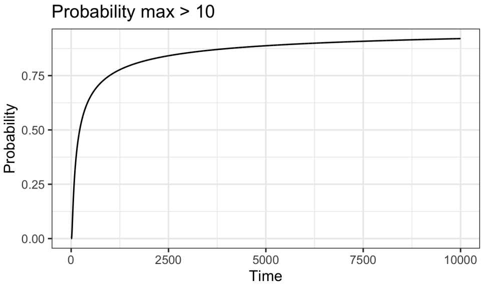
# Brownian motion is not differentiable

Let $M(t) = \max_{0 \leq s \leq t} B(s)$. $M(t)$ is well-defined because $B$ is continuous and $[0, t]$ is compact.

Proposition 1

*The following holds:*

$$Pr(M(t) > a) = 2Pr(B(t) > a) = 2 - 2\Phi\left(\frac{a}{\sqrt{t}}\right).$$



# Brownian motion is not differentiable

Proof.

Define $\tau_a = \min_s\{s : B(s) = a\}$. Note that for all $0 \leq s < t$

$$\Pr(B(t) - B(s) < 0) = \Pr(B(t) - B(s) > 0)\,,$$

and, because the distribution of $B(t) - B(\tau_a)$ is not influenced by conditioning on $\tau_a < t$,

$$\Pr(B(t) - B(\tau_a) < 0|\tau_a < t) = \Pr(B(t) - B(\tau_a) > 0|\tau_a < t)\,,$$

i.e., the "reflection principle" holds:

$$\Pr(B(t) < a|\tau_a < t) = \Pr(B(t) > a|\tau_a < t)\,.$$

Now,

$$\begin{aligned}\Pr(M(t) \geq a) = \Pr(\tau_a < t) &= \Pr(B(t) < a|\tau_a < t) + \Pr(B(t) > a|\tau_a < t)\\ &= 2\Pr(B(t) > a|\tau_a < t) = 2\Pr(B(t) > a)\,.\end{aligned}$$

$\square$

# Brownian motion is not differentiable

Proposition 2

*For $t \geq 0$, the Brownian motion is a.s. not differentiable at $t$.*

Proof.

Assume Brownian motion $B$ is differentiable at a fixed $t_0$. Then there exists constants $A$ and $\epsilon_0$ s.t. for all $0 < \epsilon < \epsilon_0$, $B(t) - B(t_0) < A\epsilon$ holds for all $0 < t - t_0 \leq \epsilon$.

Denote this event $E_{\epsilon,A}$ and let $E_A = \cap_\epsilon E_{\epsilon,A}$ . But note that

$$\begin{aligned}
\Pr(E_{\epsilon,A}) &= \Pr\left(B(t) - B(t_0) < A\epsilon, \text{ for all } 0 < t - t_0 \leq \epsilon\right) \\
&= 1 - \Pr(M(\epsilon) > A\epsilon) = 1 - 2\Pr(B(\epsilon) > A\epsilon) \\
&= 1 - 2(1 - \Phi\left(\frac{A\epsilon}{\sqrt{\epsilon}}\right)) = 1 - 2(1 - \Phi(A\sqrt{\epsilon}))
\end{aligned}$$

Taking the RHS to 0 takes the LHS to 0, and thus $P(E_A) = 0$. $\square$

# Quadratic variation: $(dB)^2 = dt$

Theorem 2

*For a partition* $\Pi = \{t_0, t_1, \ldots, t_j\}$ *of the interval* $[0, T]$*, let* $|\Pi| = \max_i(t_{i+1} - t_i)$*. A Brownian motion satisfies the following equation with probability 1:*

$$\lim_{|\Pi|\to 0} \sum_i (B(t_{i+1}) - B(t_i))^2 = T\,.$$

Proof.

For simplicity, assume gaps $t_{i+1} - t_i$ are uniform. Then $t_i = iT/n$ for $i = 0, \ldots, n-1$ and $B(t_{i+1}) - B(t_i) \sim N(0, T/n)$. Then by the LLN, for $n$ large,

$$\frac{1}{n} \sum_{i=0}^{n-1} (B(t_{i+1}) - B(t_i))^2 \approx T/n\,.$$

$\square$

# Quadratic variation: $(dB)^2 = dt$

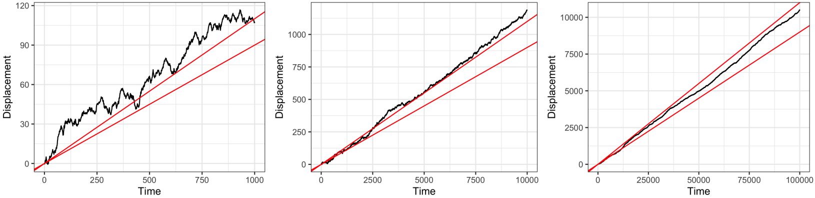
This result suggests that Brownian motion moves around a lot. For reference, assume $f$ is differentiable. Then,

$$\begin{aligned}
\sum_i (f(t_{i+1}) - f(t_i))^2 &\le \sum_i (t_{i+1} - t_i)^2 f'(s_i)^2 \\
&\le \max_{s \in [0,T]} f'(s)^2 \sum_i (t_{i+1} - t_i)^2 \\
&\le \max_{s \in [0,T]} f'(s)^2 \cdot \max_i (t_{i+1} - t_i) \cdot T \,.
\end{aligned}$$

Sending the $\max_i (t_{i+1} - t_i)$ to 0 sends the LHS to 0.

# Brownian motion with drift

We can always add a drift term and consider $X(t) = \mu t + B(t)$. The drift term overpowers diffusion in a certain sense: for any $\epsilon > 0$, as $t$ gets large, $X(t)$ is always within the lines $y = (\mu \pm \epsilon)t$.



# Ito's lemma

We know $\frac{dB_t}{dt}$ does not exist: $B(t)$ is nowhere differentiable with probability 1. But we define the infinitesimal $df$ for a smooth function $f(B(t))$? We know we cannot simply apply the chain rule:

$$df = \left( f'(B_t)\frac{dB_t}{dt} \right) dt\,.$$

But maybe we can do this anyway by using $dB_t$ directly instead? Then the previous equation becomes

$$df = f'(B_t)dB_t\,.$$

But this only works when $\Delta x \cdot f'(x)$ dominates all other terms in the Taylor expansion

$$f(x+\Delta x) - f(x) = \Delta x \cdot f'(x) + \frac{(\Delta x)^2}{2} f''(x) + \cdots\,.$$

# Ito's lemma

Let's plug $\Delta B_t$ into the Taylor expansion:

$$\Delta f = \Delta B_t \cdot f'(B_t) + \frac{(\Delta B_t)^2}{2} f''(B_t) + \cdots \,.$$

But we know that $\mathrm{E}(\Delta B_t)^2 = \Delta t$ (quadratic variation), so

$$\Delta f = \Delta B_t \cdot f'(B_t) + \frac{\Delta t}{2} f''(B_t) + \cdots \,.$$

This gives us the simplest statement of Ito's lemma:

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt \,.$$

# Ito's lemma

More generally, for a smooth function $f(t, x)$, we have

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx \,.$$

In Ito calculus, this becomes:

$$\begin{aligned} df(t, B_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dB_t)^2 \\ &= \left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t \,. \end{aligned}$$

# Ito's lemma

Theorem 3
*Let $f(t,x)$ be a smooth function, and let $X_t$ be a stochastic process satisfying $dX_t = \mu_t dt + \sigma_t dB_t$. Then*

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}\right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t \,. \end{aligned}$$

# Ito calculus

Define integration as the inverse of differentiation, i.e.,

$$F(t, B_t) = \int f(t, B_t)dB_t + \int g(t, B_t)dt$$

$$\iff$$

$$dF(t, B_t) = f(t, B_t)dB_t + g(t, B_t)dt\,.$$

# Fundamental theorem of calculus

If $f(x) = x^2/2$, then

$$df(B_t) = B_t dB_t + \frac{1}{2} dt\,.$$

This means that

$$B_T^2/2 = \int_0^T B_t dB_t + \int_0^T \frac{1}{2} dt = \int_0^T B_t dB_t + T/2$$

and thus

$$\int_0^T B_t dB_t = B_T^2/2 - T/2\,.$$

# Solving an SDE

If $f(t,x) = \exp(\mu t + \sigma x)$, then

$$df(t, B_t) = (\mu + \frac{1}{2}\sigma^2) f(t, B_t) dt + \sigma f(t, B_t) dB_t\,.$$

Question: which stochastic process $X_t(t, B_t)$ satisfies the SDE

$$dX_t = \sigma X_t dB_t\,?$$

Solution: set $\mu = -\sigma^2/2$ to get

$$X(t, B_t) = \exp\left(-\sigma^2 t/2 + \sigma B_t\right)\,.$$

# Ito calculus

Theorem 4

*Let $\Delta(t)$ be a nonrandon function of time. Suppose the stochastic process $I(t)$ satisfies*

$$dI(t) = \Delta_s dB_s\,, \quad i.e., \quad I(t) = \int_0^t \Delta_s dB_s\,,$$

*where $I(0) = 0$. Then for each $t > 0$, $I(t)$ is normally distributed.*

# Ito calculus

Let $X_t$ be a stochastic process. A process $\Delta_t$ is an *adapted* process w.r.t. $X_t$ if for all $t \geq 0$, the random variable $\Delta_t$ depends only on $X_s$ for $s \leq t$.

- The process $\Delta_t = X_t$ is an adapted process.
- The process $\Delta_t = \min(X_t, c)$ for $c$ constant is an adapted process.
- The process $\Delta_t = \max_{0 \leq t \leq T} X_t$ is not an adapted process.
- If $\tau$ is a stopping time, then $X_\tau$ is an adapted process.

Recall that a stochastic process $X_t$ is a *martingale* if $\mathrm{E}|X_t| < \infty$ and $\mathrm{E}(X_t | \{X_\tau, \tau \leq s\}) = X_s$ for all $s \leq t$.

# Ito calculus

Theorem 5

*For all adapted processes $g(t, B_t)$ satisfying the $L^2$ bound*

$$\int \int_0^t g^2(s, B_s) ds \, dB_s < \infty$$

*the integral*

$$\int_0^t g(s, B_s) dB_s$$

*is a martingale.*

# Ito calculus

The process $B_t$ itself is an adapted process. Recall that

$$\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t) \quad \text{and} \quad \mathsf{E} B_t^2 = t\,.$$

Hence

$$\mathsf{E}\left(\int_0^t B_s dB_s\right) = 0\,.$$

More generally,

$$\begin{aligned}\mathsf{E}\left(\int_{t_1}^{t_2} B_s dB_s \,\middle|\, \mathcal{F}_{t_1}\right) &= \mathsf{E}\left(\frac{1}{2}(B_{t_2}^2 - t_2) \,\middle|\, \mathcal{F}_{t_1}\right) - \frac{1}{2}(B_{t_1}^2 - t_1) \\ &= \frac{1}{2}(t_2 - t_1) + \frac{1}{2}B_{t_1}^2 - \frac{t_2}{2} - \frac{1}{2}(B_{t_1}^2 - t_1) = 0\,.\end{aligned}$$

The theorem is confirmed for $g(s, B_s) = B_s$.

# Ito isometry

Theorem 6

*For all adapted processes $\Delta_t$ w.r.t. $B_t$*

$$E\left(\left(\int_0^t \Delta_s dB_s\right)^2\right) = E\left(\int_0^t \Delta_s^2 ds\right) .$$

Let $\Delta(t) = 1$. Then

$$\mathrm{E}\left(\left(\int_0^t \Delta_s dB_s\right)^2\right) = \mathrm{E}(B_t^2) = t ,$$

and

$$\mathrm{E}\left(\int_0^t \Delta_s^2 ds\right) = t .$$

# Stochastic differential equations

We wish to solve equations of the form

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t)\,.$$

A function $X$ satisfies this equation if

$$X_T = \int_0^T \mu(t, X_t)dt + \int_0^T \sigma(t, X(t)dB(t)\,.$$

# Stochastic differential equations

## Theorem 7 (Existence and uniqueness)

*If the coefficients of the SDE*

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t)$$
$$X(0) = x_0\,, \quad 0 \leq t \leq T$$

*satisfy the conditions*

$$|\mu(t,x) - \mu(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 \leq K|x-y|^2$$

*and*

$$|\mu(t,x)|^2 + |\sigma(t,x)|^2 \leq K(1 + |x|^2)\,,$$

*then there is an adapted process solution $X(t)$ that satisfies the $L^2$ bound. If $X$ and $Y$ are both continuous solutions satisfying the $L^2$ bound, then*

$$Pr(X(t) = Y(t), \forall t \in [0, T]) = 1\,.$$

# Solving $dX(t) = \mu X(t)dt + \sigma X(t)dB(t)\,, \quad X(0) = x_0 > 0$

Step 1: assume $X(t) = f(t, B(t))$, then

$$dX(t) = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \frac{\partial f}{\partial x} dB(t)\,.$$

Step 2: equate

$$\mu X(t) = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right) \quad \text{and} \quad \sigma X(t) = \frac{\partial f}{\partial x}\,.$$

Step 3: solve the second equation with

$$f(t, x) = x_0 \exp\left(\sigma x + g(t)\right)\,.$$

Step 4: plug this into first equation

$$\mu f = g'(t) f + \frac{\sigma^2}{2} f \quad \text{to get} \quad g'(t) = \mu - \sigma^2/2\,.$$

Step 5: recognize that

$$f(t, x) = x_0 \exp(\sigma x + (\mu - \sigma^2/2)t) \quad \text{or} \quad X(t) = x_0 \exp(\sigma B(t) + (\mu - \sigma^2/2)t)\,.$$

# Solving $dX(t) = -\alpha X(t)dt + \sigma dB(t)\,, \quad X(0) = x_0$

Try the test function

$$X(t) = a(t)\left(x_0 + \int_0^t b(s)dB(s)\right), \quad a(0) = 1\,, \quad a(t) > 0\,, \forall t\,.$$

Differentiating gives

$$\begin{aligned} dX(t) &= a'(t)dt\left(x_0 + \int_0^t b(s)dB(s)\right) + a(t)b(t)dB(t) \\ &= \frac{a'(t)}{a(t)}X(t)dt + a(t)b(t)dB(t)\,. \end{aligned}$$

Matching this to the original SDE gives

$$-\alpha = \frac{a'(t)}{a(t)}\,, \quad \sigma = a(t)b(t)\,.$$

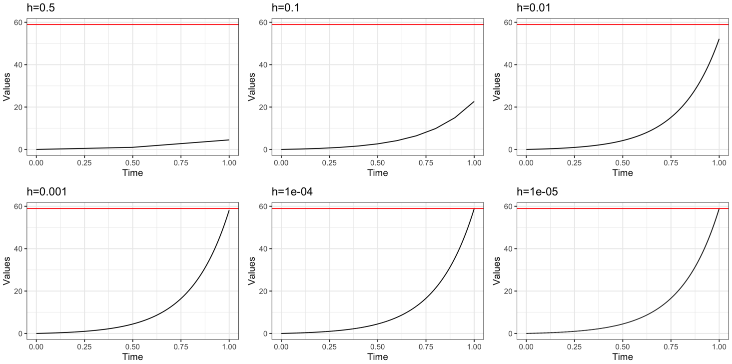
Thus $a(t) = \exp(-\alpha t)$, $b(t) = \sigma\exp(\alpha t)$ and

$$X(t) = x_0\exp(-\alpha t) + \sigma\int_0^t \exp(\alpha(s-t))dB(s)\,.$$

# Euler's method for ODEs

Problem: obtain $u(1)$ for ODE $u'(x) = 5u(x) + 2$ with $u(0) = 0$.
Solution: select small number $h > 0$ and use Taylor approximation at each step for times $t = 0, 1h, 2h, \ldots, (1/h - 1)/h, 1$.

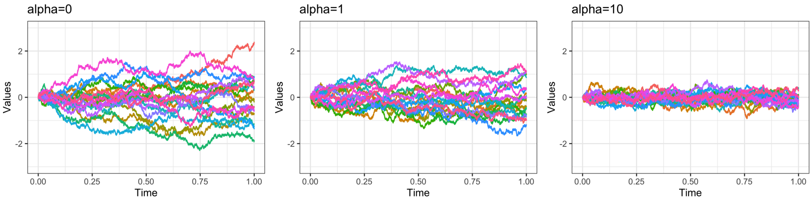
$$u(t+h) \approx u(t) + h \cdot u'(t) = u(t) + h \cdot (5u(x) + 2)\,.$$



# Euler-Maruyama

Problem: obtain distribution of $X(1)$ for OU equation $dX(t) = -\alpha X(t)dt + \sigma dB(t)$ with $X(0) = 0$.
Solution: select small number $h > 0$ and use Taylor approximation at each step for times $t = 0, 1h, 2h, \ldots, (1/h - 1)/h, 1$.

$$X(t+h) \approx X(t) + dX(t) = X(t) - h\alpha X(t) + \sigma\sqrt{h}Z_{t+h}\,.$$



# Langevin Monte Carlo

We are interested in generating samples from a target distribution

$$\pi(\theta) \propto \exp(-U(\theta)),$$

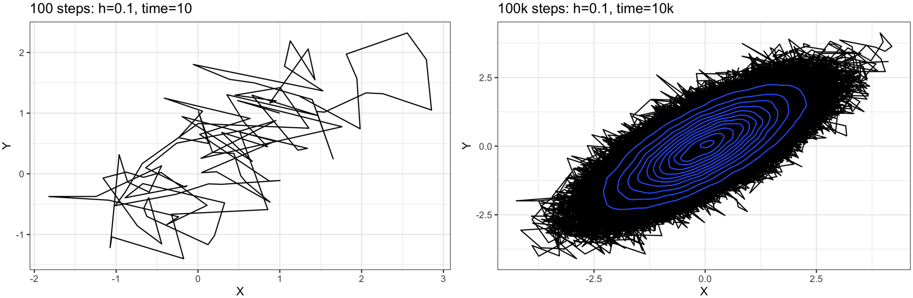
so we simulate the diffusion that solves the SDE

$$\begin{aligned} d\theta(t) &= -\nabla U(\theta(t))dt + \sqrt{2}dB(t) \\ &= \nabla \log \pi(\theta(t))dt + \sqrt{2}dB(t) \end{aligned}$$

using the Euler-Maruyama method, e.g.,

$$\theta(t+h) = \theta(t) + h\nabla \log \pi(\theta(t)) + \sqrt{2h}Z_{t+h}.$$

# Langevin Monte Carlo



# Justifying LMC

The stochastic process $\theta(t)$ that satisfies

$$d\theta(t) = \nabla \log \pi(\theta(t))dt + \sqrt{2}dB(t)$$

leaves $\pi(\theta)$ invariant. To see this, we use a PDE that describes the evolution of the probability density function of $X(t)$ with time for the general Ito diffusion

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t)\,.$$

In 1D, this PDE is the Fokker-Plank equation:

$$\frac{\partial}{\partial t}p(t,x) = -\frac{\partial}{\partial x}\left(\mu(t,x)p(t,x)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(t,x)p(t,x)\right)\,.$$

For us, this becomes

$$\frac{\partial}{\partial t}p(t,\theta) = -\frac{\partial}{\partial \theta}\left(\frac{\partial}{\partial \theta}\log \pi(\theta)p(t,\theta)\right) + \frac{\partial^2}{\partial \theta^2}p(t,\theta)\,.$$

# Justifying LMC

For us, this becomes:

$$\frac{\partial}{\partial t}p(t,\theta) = -\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}\log\pi(\theta)p(t,\theta)\right) + \frac{\partial^2}{\partial\theta^2}p(t,\theta)\,.$$

Want to show: if $p(t,\theta) = \pi(\theta)$, then $\frac{\partial}{\partial t}p(t,\theta) = 0$. Plug it in:

$$\begin{aligned}\frac{\partial}{\partial t}p(t,\theta) &= -\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}\log\pi(\theta)\pi(\theta)\right) + \frac{\partial^2}{\partial\theta^2}\pi(\theta)\\ &= \frac{\partial}{\partial\theta}\left(-\frac{\partial}{\partial\theta}\pi(\theta) + \frac{\partial}{\partial\theta}\pi(\theta)\right) = 0\,.\end{aligned}$$