Interactive demonstration. A script file "bm.R" contains all the interactive demonstrations accompanying this lecture. You can download `bm.R` and start it by executing the following at the native R (Rgui, not in RStudio).

```
> source("bm.R")
```

Or, you can choose `bm.R` from R via [File]→[Source R code...]. In either way it opens a menu screen, entering into an interactive demonstration immediately. To start each topic, move the cursor on the top of title and click the left mouse button. To return the menu at the end of demonstration, move the cursor to "[Exit]" and click the left mouse button.

Brownian motion. Given a sample space Ω and a fixed real value x_0, a **sample path** $B = (B_s)_{s\geq 0}$ (a collection of real-valued random variables indexed by "time s") is called a **Brownian motion** if

1. $B_0(\omega) = x_0$ for all $\omega \in \Omega$ (that is, it always starts at x_0).
2. For each $\omega \in \Omega$, the "sample path" $B_s(\omega)$ is continuous for $s \in [0, \infty)$.
3. For each $0 \leq s \leq t < u$, the increment $B_u - B_t$ is independent of $B_s - B_0$ in the past, and it follows a normal (or "Gaussian") distribution with mean 0 and variance $(u - t)$.

```
brownian.motion> Click [Run] to generate sample path
brownian.motion> Click bewteen 0 and T above [Probability Density]
```

Gaussian property of Brownian motion. Let $F(B)$ be a function of a sample path B. Suppose that F is a simple function formed by $\prod_{i=1}^{N} I_{E_i}(B(t_i))$ with Borel subsets E_i's and $0 = t_0 < t_1 < \cdots < t_N \leq T$. Then the expectation of $F(B)$ is obtained by

$$\mathbf{E}[F(B)] = \int_{E_0} \cdots \int_{E_N} \prod_{i=1}^{N} b(t_i - t_{i-1}, x_{i-1}, x_i) dx_i$$

where

$$b(t, x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{(x-y)^2}{2t}\right] \tag{6.1}$$

is the **transition probability density**. If F is not a simple function, it can be expressed as the limit of weighted sum of simple functions; thus, the expectation of $F(B)$ is also obtained by the limit.

Canonical form of Brownian motion. Because outcomes are sample paths, we can identify the sample space Ω with the space $C([0, \infty), \mathbb{R})$ of continuous functions.

- We view an outcome ω as a continuous sample path, often write $\omega(s) = B_s(\omega)$ and use ω and B interchangeably.
- The probability measure $\mathbb{P}_{x_0}$ is defined on Ω along with σ-algebra $\mathcal{F}$, and it gives a Gaussian property of Brownian motion $(B_s)_{s\geq 0}$ starting from the state x_0 at time $s = 0$.
- The quadruple $(\Omega, \mathcal{F}, \mathbb{P}_{x_0}, B_s)$ is called a **canonical form**.

Having fixed $T > 0$, it is often of our interest to construct sample paths on the interval $[0, T]$. Here the sample space is identified with $C([0, T], \mathbb{R})$, and denoted by Ω_T. Then the canonical form becomes $(\Omega_T, \mathcal{F}, \mathbb{P}_{x_0}, B_s)$.

Why study Brownian motion? At the beginning of their book "*Diffusions, Markov Processes, and Martingales*," Rogers and Williams answered this question as follows:

1. Every interesting class of processes, (Gaussian process, Markov process, or diffusion) contains Brownian motion.
2. It is sufficiently concrete for explicit calculations which are impossible in general.
3. It can be used as a building block for other processes.
4. It is a rich and beautiful mathematical object in its own right (continuous but nowhere differentiable sample path).

Gradient descent algorithm. Consider an objective function $f(x)$ on $\mathbb{R}^n$, and the optimization problem to minimize $f(x)$ over $\mathbb{R}^n$. Gradient descent algorithm employs a descent direction of the negative gradient

$$-\nabla f(x) = -\left[\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right]^T$$

1. Choose the initial state x_0 at time $t = 0$.
2. Suppose x_t at time t. Update the next position $x_{t+\Delta t}$ at time $t + \Delta t$ by

$$x_{t+\Delta t} = x_t - \left(\frac{\Delta t}{2}\right) \nabla f(x_t) \tag{6.2}$$

   with step size Δt small enough.

Langevin algorithm. We can relax a strict descent criteria and add "noise" to (6.2) by introducing a Brownian motion B_t.

$$X_{t+\Delta t} = X_t - \left(\frac{\Delta t}{2}\right) \nabla f(X_t) + B_{t+\Delta t} - B_t \tag{6.3}$$

By letting $\Delta t \to 0$, the dynamical system (6.3) can be viewed as a solution X_t to the stochastic differential equation (SDE)

$$dX_t = -\frac{1}{2} \nabla f(X_t)\, dt + dB_t \tag{6.4}$$

with initial state $X_0 = x_0$.

Kolmogorov backward equation. Corresponding to (6.4) we can introduce a diffusion operator $\mathcal{A}_x$ by

$$\mathcal{A}_x = \frac{1}{2} \sum_{i=1}^{n} \left[\frac{\partial^2}{\partial x_i^2} - \frac{\partial f}{\partial x_i}(x) \frac{\partial}{\partial x_i}\right]$$

It uniquely determines a transition density function $p(t,x,y)$ as a fundamental solution to

$$\frac{\partial}{\partial t} p(t,x,y) = \mathcal{A}_x p(t,x,y) \tag{6.5}$$

which is referred as **Kolmogorov backward equation**.

Time-reversibility of Langevin algorithm. The fundamental solution $p(t,x,y)$ of (6.5) implies the symmetry

$$\pi(x)p(t,x,y) = \pi(y)p(t,y,x) \tag{6.6}$$

with nonnegative **invariant function**

$$\pi(x) \propto \exp(-f(x)) \tag{6.7}$$

In general $\pi(x)$ could be improper, that is, $\int \pi(x)\,dx = \infty$. If (6.7) is a probability density function then the corresponding SDE

$$dX_t = \frac{1}{2}\nabla \ln \pi(X_t)\,dt + dB_t \tag{6.8}$$

converges to the stationary distribution π, and (6.8) is appropriately called **Langevin algorithm**.

Time-reversed sample path. The time-reversibility (6.6) of Langevin algorithm leads to the following lemma (an exercise by Rogers and Williams).

Lemma 1. *Assuming that a function F on $C([0,T],\mathbb{R})$ is integrable in either side of (6.9), we have*

$$\int \mathbf{E}(F(X)|X_0 = x)\pi(x)dx = \int \mathbf{E}(F(X_{T-\cdot})|X_0 = y)\pi(y)dy \tag{6.9}$$

where $X_{T-\cdot}$ is the time-reversed sample path $\hat{X}_t = X_{T-t}$, $0 \le t \le T$.

Forward and backward stochastic flow. Consider the Langevin algorithm (6.8) over $\mathbb{R}$, and assume that $\pi(x)$ is symmetric, that is, that $\pi(x) = \pi(-x)$. Then we can construct a **stochastic flow** $\Phi_{t,s}$ for $0 \le t \le s < \infty$. The map $\Phi_{t,s}(\cdot, B)$ is homeomorphic, and the inverse map $\Phi_{t,s}^{-1}(\cdot, B)$ exists. Having fixed t and x,

$$\Phi_{t,s}(x,B) = X_s, \quad t \le s,$$

is a solution to (6.8) starting from $X_t = x$ and moving forward. Similarly having fixed s and y, $Y_t = \Phi_{t,s}^{-1}(y,B)$, $0 \le t \le s$, is a solution to (6.8) starting from $Y_s = y$ and moving backward.

```
forward.backward> Construct [Forward]/[Backward] flow
forward.backward> [Run] to generate a new sample
```

Skorohod equation. Let $\kappa(s)$ be a real-valued continuous function. Assuming $\kappa(0) \ge 0$, we call

$$\eta(s) = \kappa(s) + \ell(s) \tag{6.10}$$

a **Skorohod equation** if $\eta(s)$ is a nonnegative continuous function and $\ell(s)$ is a nondecreasing continuous function with $\ell(0) = 0$, satisfying

$$\ell(s) = \int_0^s I_{\{\eta(u)=0\}} d\ell(u),$$

where $I_{\{\eta(u)=0\}}$ is the indicator function of a statement $\{\eta(u) = 0\}$, taking values 1 or 0 accordingly as the statement is true or not. Given $\kappa(s)$, a pair (κ, ℓ) of functions forms the Skorohod equation, and the nonnegative function η of (6.10) is uniquely determined by

$$\ell(s) = -\min_{0\le u\le s}[\kappa(u) \wedge 0].$$

Coalescing flow. (i) Consider a sample path $Y_s = \Phi_{s,T}^{-1}(y, -B)$ starting at $s = T$ from y and moving **backward** until $s = t$. (ii) Construct a sample path X_s, $t \le s \le T$, starting at $s = t$ from x and moving **forward** by

$$X_s = \begin{cases} \Phi_{t,s}(x, -B) & \text{if } x > Y_t; \\ \Phi_{t,s}(x, B) - \ell(s) & \text{if } x \le Y_t, \end{cases} \tag{6.11}$$

and Skorohod equation for $x \le Y_t$ with

$$\ell(s) = \int_t^s I_{\{X_v = Y_v\}} d\ell(v).$$

We construct (6.11) as a map, and denote it by $\psi_{t,s,y,T}(x, B) = X_s$, $t \le s \le T$.

```
coalescing.flow> Construct [Forward]/[Backward] flow
coalescing.flow> [Run] to generate a new sample
```

Markov process with coffin state. We can construct a sample path Y_t^*, $s \le t \le T$, starting from $Y_s^* = y > 0$ by

$$Y_t^* = \begin{cases} \Phi_{T-t,T-s}^{-1}(y, -B_{T-\cdot}) & \text{if } s \le t < \tau; \\ 0 & \text{if } \tau \le t, \end{cases}$$

and stopping at $\tau = \inf\{s \le t \le T : Y_t^* = 0\}$. Since

$$Y_t^* = \Phi_{T-t,T-s}^{-1}(y, -B_{T-\cdot})$$

does not depend on T, it determines the stochastic flow

$$\Xi_{s,t}^*(y, B) = Y_t^*$$

and $\Xi_{s,t}^*(\cdot, B)$ maps from $D^* = (0, \infty) \cup \{0\}$ to itself, where 0 is viewed as a "coffin state."

Subprobability transition function. The stochastic flow $\Xi_{s,t}^*(y, B) = Y_t^*$, $s \le t < \infty$, is viewed as a Markov process absorbed at 0, and later it is called a **Liggett dual**. It has the subprobability transition function $q(t, x, y)$ for $0 < x, y < \infty$, where $\int_0^\infty q(t, x, y)\, dy < 1$.

```
liggett.dual> Click [Run] to generate sample path
liggett.dual> Click bewteen 0 and T above [Probability Density]
```

Liggett duality. Observe that $Y_T^* = \Xi_{0,T}^*(y, B)$ starts from $Y_0^* = y$. Define

$$\Gamma(y,x) = \begin{cases} 1 & \text{if } x \in (-y, y]; \\ 0 & \text{otherwise.} \end{cases}$$

Then we obtain

$$\Gamma(Y_T^*, x) = \Gamma(y, X_T) \tag{6.12}$$

where X_T is constructed by the coalescing flow (6.11) starting from $X_0 = x$. We call Y_T^* a **Liggett dual** of X_T.

Imputing a Liggett dual. Suppose that the Brownian motion B_t is missing. Provided a sample path $\xi(t)$ moving backward for $s \le t \le T$, we can impute a sample path B_t so that

$$\xi(t) = \Phi_{0,T-t}(\xi(T), B_{T-\cdot})$$

for $s \le t \le T$ as if $X_u = \xi(T-u)$ and $Y_u = \eta(T-u)$ satisfy (6.11) for $0 \le u \le T-s$. Then $\eta(t)$ moves forward by setting

$$\eta(t) = \begin{cases} \Phi_{T-t,T-s}^{-1}(y, B_{T-\cdot}) & \text{if } \xi(s) > y; \\ \Phi_{T-t,T-s}^{-1}(y, -B_{T-\cdot}) + \ell(t) & \text{if } \xi(s) \le y, \end{cases} \tag{6.13}$$

where

$$\ell(t) = \int_s^t I_{\{\xi(v)=\eta(v)\}} d\ell(v).$$

Imputing a Liggett dual, continued. Thus, we can map from ξ to the imputed B, and define $\Theta_{y,s}(\xi) = B$. For $\xi \in C([0,T], \mathbb{R})$ and $y \ge 0$ the map

$$\Psi_{s,t}^*(y, \xi) = \Xi_{s,t}^*(y, \Theta_{y,s}(\xi))$$

allows us to impute a Liggett dual provided a sample path ξ. Then $\Psi_{s,t}^*(y,\xi)$ is a stochastic flow, and by Liggett duality it satisfies

$$\Gamma(y, \xi(s)) = \Gamma(\Psi_{s,T}^*(y, \xi), \xi(T)), \quad 0 \le s \le T. \tag{6.14}$$

```
imputing.dual> Construct [Forward]/[Backward] flow
imputing.dual> [Run] to generate a new sample
```

Λ-linked coupling. Recall the time-reversibility of X_s. For any function g on $[0,\infty) \times \mathbb{R}$ we obtain

$$\begin{aligned} &\int_{-\infty}^{\infty} \mathbf{E}(g(\Psi_{0,T}^*(y,X), X_T) | X_0 = x) \Gamma(y,x) \pi(x) dx \\ &= \int_{-\infty}^{\infty} \mathbf{E}(g(\Psi_{0,T}^*(y,X), X_T) \Gamma(\Psi_{0,T}^*(y,X), X_T) | X_0 = x) \pi(x) dx \\ &= \int_{-\infty}^{\infty} \mathbf{E}(g(\Psi_{0,T}^*(y, X_{T-\cdot}), z) \Gamma(\Psi_{0,T}^*(y, X_{T-\cdot}), z) | X_0 = z) \pi(z) dz \end{aligned} \tag{6.15}$$

For the reason explained later the resulting construction of $(\Psi^*_{0,T}(y, X), X_T)$ is called **Λ-linked coupling**.

Doob h-transform. Let $q(t, x, y)$ be the subprobability transition function for the Liggett dual Y^*_t absorbed at 0. By setting $g(y, x) \equiv 1$ and $Y^*_T = \Psi^*_{0,T}(y, X_{T-\cdot})$ in (6.15) we find the harmonic function

$$\begin{aligned} h(y) &= \int_{-\infty}^{\infty} \Gamma(y, x)\pi(x)dx = \int_{-\infty}^{\infty} \mathbf{E}[\Gamma(Y^*_T, z)]\pi(z)dz \\ &= \int_0^{\infty} q(T, y, x)dx \int_{-\infty}^{\infty} \Gamma(x, z)\pi(z)dz \\ &= \int_0^{\infty} q(T, y, x)h(x)dx \end{aligned} \tag{6.16}$$

We define a probability transition function q^* by

$$q^*(t, x, y) = q(t, x, y)h(y)/h(x), \quad 0 < x, y < \infty,$$

and call the corresponding stochastic process **Doob h-transform**.

Intertwining dual. We introduce an operator Λ by

$$\Lambda f(y) = \int_{-\infty}^{\infty} \lambda(y, x)f(x)dx \tag{6.17}$$

with $\lambda(y, x) = \Gamma(y, x)\pi(x)/h(y)$. Dividing (6.15) by $h(y)$ we obtain

$$\begin{aligned} &\int_{-\infty}^{\infty} \mathbf{E}(g(\Psi^*_{0,T}(y, X), X_T)|X_0 = x)\lambda(y, x)dx \\ &= \int_0^{\infty} q^*(T, y, x)dx \int_{-\infty}^{\infty} g(x, z)\lambda(x, z)dz \end{aligned} \tag{6.18}$$

Observe that $\lambda(y, \cdot)$ in (6.17) has the stationary distribution $\pi(x)$ on $(-y, y]$. By choosing an initial state X_0 from $\lambda(y, \cdot)$ the map $\Psi^*_{0,T}(y, X)$ constructs the Doob h-transform.

By setting $g(y, x) = f(x)$ in (6.18) we find

$$\begin{aligned} &\int_{-\infty}^{\infty} \lambda(y, x)dx \int_{-\infty}^{\infty} p(T, x, \eta)f(\eta)d\eta \\ &= \int_0^{\infty} q^*(T, y, x)dx \int_{-\infty}^{\infty} \lambda(x, z)f(z)dz \end{aligned} \tag{6.19}$$

Thus, p and q^* are intertwined with the operator Λ, and the Doob h-transform q^* is called a **intertwining dual**.

```
intertwining.dual> Construct [Forward] flow
intertwining.dual> [Run] to generate a new sample
```

Λ-linked coupling algorithm. Let $y > 0$ be fixed. We can generate an **intertwining dual** Y^*_t, and couple it with X_t as follows.

1. Generate $X_0 = x$ from the stationary distribution π on $(-y, y]$ [or, distributed as $\lambda(y, \cdot)$];
2. generate $X_t = \Phi_{0,t}(x, B)$;
3. generate a process $Y_t^* = \Psi_{0,t}^*(y, X)$.

By (6.18) we can observe that

1. Y_t^* is a Doob h-transform starting from $Y_0^* = y$;
2. given $Y_t^* = \xi$, X_t is distributed as π on $(-\xi, \xi]$ [or, distributed as $\lambda(\xi, \cdot)$].

Pitman-type theorem. In the construction of intertwining dual Y_t^* we can let $y \to 0$ so that we can choose $X_0 = 0$. Pitman (1975) first obtained

Theorem 2 (Pitman $2M-B$ theorem)**.** *Let $B_t(\omega)$ be a Brownian motion starting from $B_0(\omega) = 0$. Then Y_t^* is a three-dimensional Bessel process starting from $Y_0^*(\omega) = 0$. Furthermore, B_t is uniformly distributed on $(-\xi, \xi]$ given $Y_t^*(\omega) = \xi$ at time t.*

```
pitman.theorem> [Run] to generate a new sample path
```