6. Brownian Motion

- definition and invariance principle, properties
- reflected Brownian motion
- Brownian bridge, Ornstein-Uhlenbeck proc.
- diffusion processes

Brownian motion is an important stochastic model used in many applications, e.g. in finance, industry and elsewhere. It also plays important roles in mathematical statistics, e.g. in studyinging goodness-of-fit, nonparametric function estimation and many other procedures. Brownian motion is the ultimate extension of a random walk and is crucial to the definition of a diffusion process.

Def. 6.1 Let $\{B(t)\}_{t\geq 0}$ be a stochastic process with the following assumptions.

(i) B has independent increments.
(ii) $B(t)-B(s) \sim \text{Normal}(0, t-s)$.

$\Big\} \Rightarrow B(t)$ is a Levy process (recall Def. 4.19)

(iii) $B(t)$ is a continuous function, with probability 1.
(iv) $B(0)=0$ w.p. 1.

$B(t)$ is called <u>standard Brownian Motion</u> (or <u>standard Wiener process</u>).
A more general Brownian motion is $W(t) = W_0 + \sigma B(t)$ where $\sigma > 0$ and $W_0 \sim \text{normal}(x, \eta)$, independent of B, $\eta > 0$, $x \in \mathbb{R}$.

That is, $B(t)$ is a <u>Gaussian process</u>.

<u>Thm. 6.2</u> $B(t)$ is standard Brownian Motion iff for every $t_1, \ldots, t_k, k$, $(B(t_1), \ldots, B(t_k)) \sim \text{normal}(0, \Sigma)$ where Σ is the matrix with (i,j) element $\min(t_i, t_j)$ and B has continuous paths w.p. 1.

proof. $\Rightarrow$ By (ii) & (iv), $B(t) \sim \text{normal}(0, t)$ for each t.

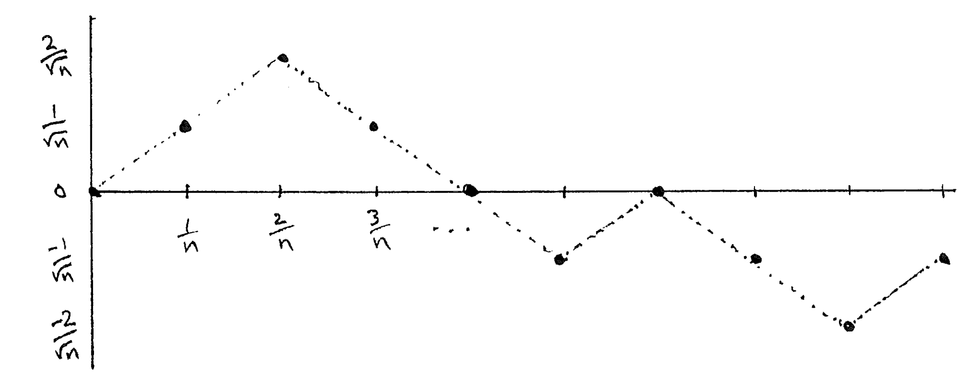
let $s<t$.

$$\begin{aligned} \operatorname{cov}(B(t), B(s)) &= E(B(t)B(s)) \\ &= E(B^2(s)) + E\big((B(t)-B(s))B(s)\big) \\ &= s + 0 \quad \text{(since increments are independent)} \\ &= \min(t,s). \end{aligned}$$

Of course, since $(B(t_1), B(t_2)-B(t_1), \ldots, B(t_k)-B(t_{k-1}))$ is multivariate normal, $(B(t_1), \ldots, B(t_k))$ is also.

$\Leftarrow$ Reversing the argument, the increments are multivariate normal and, by the form of Σ, uncorrelated, hence independent. Also, $B(0) \sim \text{normal}(0,0)$, i.e. $B(0)=0$ w.p. 1. $\square$

Brownian motion is a limiting version of a zero mean random walk with both the time axis and the vertical axis shrunk.



<u>Thm 6.3</u> (Functional CLT - loosely) Let $\{X_n\}$ be iid r.v.'s with mean 0 and variance 1. Let $S_n = X_1 + \cdots + X_n$, $S_0 = 0$ and

$$B_n(t) = \frac{S_{[nt]}}{\sqrt{n}}, \qquad ([s] = \text{greatest integer less than or equal to } s)$$

Then, $(B_n(t_1), \ldots, B_n(t_k)) \Rightarrow (B(t_1), \ldots, B(t_k))$ for all $t_1, \ldots, t_k, k$.

<u>Remark</u> This is finite dimensional convergence. "Functional" CLT actually refers to convergence in a functional metric space, a stronger result.

proof. We may assume $t_1 < t_2 \le \cdots < t_n$.

$$B_n(t_j) - B_n(t_{j-1}) = \frac{1}{\sqrt{n}} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i$$

$$\stackrel{d}{=} \frac{1}{\sqrt{n}} S_{[nt_j]-[nt_{j-1}]}$$

$$\Rightarrow \text{Normal}(0, (t_j - t_{j-1})).$$

(if $\frac{\alpha_n}{n} \to \alpha$ then $\frac{1}{\sqrt{n}} S_{\alpha_n} \Rightarrow$ normal$(0,\alpha)$ (check))

Also, the increments of B_n : $(B_n(t_1), B_n(t_2)-B_n(t_1), \ldots, B_n(t_k)-B_n(t_{k-1}))$ are independent so they converge jointly to independent normal r.v.'s, in dist. The conclusion follows from this. $\square$

We now state the stronger version.

Thm 6.4 (Functional CLT - or Invariance Principle) Let $C[0,\infty)$ be the metric space of continuous functions on $[0,\infty)$ with metric

$$\rho(f_1, f_2) = \sum_{n=1}^{\infty} 2^{-n} \min(1, \sup_{t \le n} |f_1(t) - f_2(t)|).$$

(thus $\rho(f_m, f) \to 0 \iff \sup_{t \le n} |f_m(t) - f(t)| \to 0$, local unif. conv.)

Define $B_n(t) = \frac{S_{[nt]}}{\sqrt{n}} + (nt - [nt]) \frac{X_{[nt]+1}}{\sqrt{n}}$ (which is continuous - "connects the dots").

If $T: C[0,\infty) \to \mathbb{R}$ is continuous (i.e. $\rho(f_m, f) \to 0 \Rightarrow T(f_m) \to T(f)$)
Then $T(B_n) \Rightarrow T(B)$.

Remark The invariance principle is extremely important in statistics because with it we can study the behavior of $\frac{S_{[nt]}}{\sqrt{n}}$, not just of $\frac{S_n}{\sqrt{n}}$. For example,

$T(B_n(t)) = \max_{0 \le t \le 1} |B_n(t)| \Rightarrow \max_{0 \le t \le 1} |B(t)|$, which is useful for sequential testing.

Ex. 6.1 Suppose $Y_1, Y_2, \ldots$ are iid with mean 0 and variance 1. For fixed sample size n, consider the statistic

$$M_n = \max_{k \le n} S_k = \max_{k \le n} \sum_{j=1}^{k} Y_j .$$

Note that $\frac{M_n}{\sqrt{n}} = \max_{0 \le t \le 1} B_n(t)$, where B_n is as in Thm. 6.4.

$T(f) = \max_{0 \le t \le 1} f(t)$ is a cont. functional on $C[0,\infty)$.

Thus, $\frac{M_n}{\sqrt{n}} \Rightarrow \max_{0 \le t \le 1} B(t)$. We show later that this has the same dist. as $|Z|$, $Z \sim$ normal$(0,1$

We now discuss basic properties of Brownian motion.

Thm. 6.5 Let $B(t)$ be standard Brownian motion.

(i) $B(t)$ is a Markov process with homogeneous transition probabilities such that the cond. dist. of $B(t+s)$, given $B(t)=x$ is normal(x,s).

(ii) $-B(t)$ is standard B.M.

(iii) Let $B_s(t) = B(t+s) - B(s)$ with $s>0$ fixed. B_s is standard B.M. and independent of $\{B(u)\}_{0\le u\le s}$.

(iv) $B_s^*(t) = \begin{cases} B(t), & t\le s, \\ 2B(s) - B(t), & t \ge s \end{cases}$ is stand. B.M. (Reflection at $t=s$)

(v) For any $c>0$, $\{\sqrt{c}\, B(t/c)\}$ is stand. B.M. (self-similarity)

(vi) $\{t B(1/t)\}$ is stand. B.M. (time reversal)

proof. (i) This follows from the fact that B has independent, stationary normal increments:

$$\begin{aligned} P(B(t_{n+1}) \le y \mid B(t_i)=x_i, i\le n) &= P(B(t_{n+1}) - B(t_n) \le y - x_n \mid B(t_i) = x_i, i \le n) \\ &= P(B(t_{n+1}) - B(t_n) \le y - x_n) \\ &= \Phi\left(\frac{y-x_n}{\sqrt{t_{n+1}-t_n}}\right). \end{aligned}$$

(ii) Easy to check.

(iii) Like (i), check the finite-dim. distributions.

(iv) Let $t_1 < t_2 < \cdots < t_k$. It suffices to assume $s = t_j$ (o.w. include s - we still will characterize all finite dim. dist.'s)

The increments of B_s^* are

$$B(t_1),\ B(t_2) - B(t_1), \ldots, B(s) - B(t_{j-1}),$$
$$B(s) - B(t_{j+1}),\ B(t_{j+1}) - B(t_{j+2}), \ldots, B(t_{k-1}) - B(t_k)$$

which are independent, normal and have the right variances. Note that B_s^* is cont. also. So this ensures it is B.M.

(Sketch)

(v) $\sqrt{c}\,B(0/c) = 0$, $\sqrt{c}B(t/c)$ is cont., has indep. increments and $\sqrt{c}\,B(t/c) - \sqrt{c}\,B(s/c) \sim \text{Normal}(t-s)$.

(vi) Define $\tilde{B}(t) = \begin{cases} t\,B(1/t), & \text{if } t > 0 \\ 0, & \text{if } t = 0 \end{cases}$

Obviously, $(\tilde{B}(t_1), \dots, \tilde{B}(t_k))$ is multivariate normal with mean 0. So we need only to check it has the correct covariances (by Thm. 6.2). Let $s \geq 0, t \geq 0$.

$$\begin{aligned} \operatorname{cov}(\tilde{B}(s), \tilde{B}(t)) &= E(st\,B(1/s)B(1/t)) \\ &= st\min(\tfrac{1}{s}, \tfrac{1}{t}) = \min(s,t). \end{aligned}$$

$\tilde{B}$ is continuous, except possibly at 0, which we need to check. (Continuity at 0 means $\frac{1}{s}B(s) \to 0$ as $s \to \infty$.)

Let $Z_n = \max\limits_{n \leq t \leq n+1} |B(t) - B(n)|$. The Z_n's are iid.

Then, for $n \leq s \leq n+1$:

$$\left| \frac{B(s)}{s} - \frac{B(n)}{n} \right| \leq \frac{|B(s) - B(n)|}{s} + |B(n)| \left(\frac{1}{n} - \frac{1}{s} \right)$$

$$\leq \frac{Z_n}{n} + \frac{|B(n)|}{n^2}.$$

$B(n) = \sum_{i=1}^{n} (B(i) - B(i-1))$ is a sum of iid r.v.'s w/ mean 0.

So $B(n)/n \to 0$ w.p.1 and $|B(n)|/n^2 \to 0$ w.p.1.

Likewise, if $E(Z_1) < \infty$, then $\frac{1}{n}\sum_{i=1}^{n} Z_i \to E(Z_1)$ w.p.1 (strong law of large numbers) and hence $Z_n/n \to 0$ w.p.1. (Z_n's are nonneg.) $\square$

We will show later that $Z_1 = \sup\limits_{t \leq 1} |B(t)|$ has a finite mean (Cor. 6.10).

In Thm. 6.5(iv) the B.M. was reflected at a fixed time s. It is also useful to reflect it at the <u>random</u> time it reaches a fixed boundary a.

<u>Def. 6.6</u> Let $\{X(t)\}$ be a stochastic process for $t \in [0,\infty)$. A nonnegative random variable T is a <u>stopping time</u> for $\{X(t)\}$ iff the event $\{T \le t\}$ is determined to have occurred or not by the process up to time t.

(More precisely, $\{T \le t\}$ is in the σ-field of events generated by events $\{X(s) \in A\}$, $0 \le s \le t$, Borel sets A. Also, some stronger assumptions are required, including $X(t)$ is right-cont, w/ left-hand limits

As in the discrete time case, if T is a stopping time for a process $\{X(t)\}$ which you have been following, then at any fixed time t you know either (a) the value of T and that $T \le t$, or (b) that $T > t$.

<u>Thm. 6.7</u> (Strong Markov Property for B.M.) Suppose $\{B(t)\}$ is Brownian motion and T is a stopping time for $B(t)$. Then, conditional on the event $\{T < \infty\}$, the process

$$\{\tilde{B}(t)\} = \{B(T+t) - B(T)\}_{t \ge 0}$$

is Brownian motion and is independent of T & $\{B(s)\}_{0 \le s \le T}$.

proof. (see book - Sec. 6.6. Also, recall Thm. 2.16.) □

<u>Thm. 6.8</u> (Reflection property of B.M.) Define $T_a = \inf\{t : B(t) = a\}$, $a > 0$. (T_a is the first time $B(\cdot)$ hits level a.) Define

$$B^*(t) = \begin{cases} B(t), & t \le T_a, \\ 2a - B(t), & t > T_a. \end{cases}$$

Then $B^*(t)$ is Brownian motion.

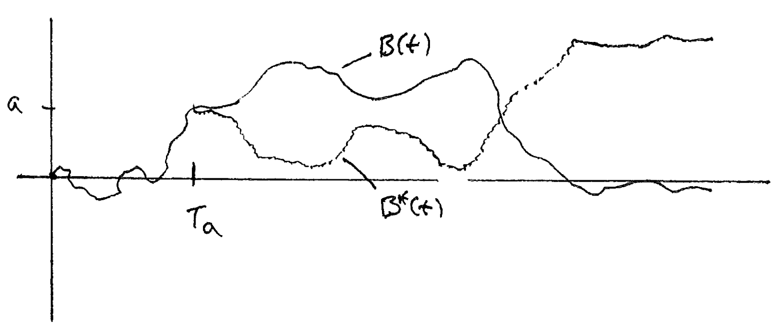
proof. Given $T_a < \infty$, $\tilde{B}(t) = B(T_a + t) - a = B(T_a + t) - B(T_a)$ is B.M. by Thm 6.7 and independent of $\{B(s)\}_{s \le T_a}$. Since T_a is a stopping time, it may be shown it is a function solely of $\{B(s)\}_{s \le T_a}$ and is therefore independent of $\tilde{B}(\cdot)$. (this can be made precise in measure theoretic terms.)

First, we compute

$$
\begin{aligned}
P(B(t) \ge a) &= P(B(t) \ge a, T_a \le t) \qquad \text{since } B(t) \ge a \Rightarrow T_a \le t \\
&= P(B(t) \ge a, T_a \le t \mid T_a < \infty)\, P(T_a < \infty) \\
&= P(T_a < \infty) \int_0^t P(\tilde{B}(t-s) \ge 0 \mid T_a = s)\, \underset{T_a}{F}(ds) \qquad \uparrow \text{cond. dist. given } T_a < \infty \\
\overset{\text{indep.}}{=} & \ \tfrac{1}{2} P(T_a < \infty) \int_0^t F_{T_a}(ds) \\
&= \tfrac{1}{2} P(T_a \le t)
\end{aligned}
$$

Since $P(B(t) \ge a) = 1 - \Phi(\frac{a}{\sqrt{t}}) \to \frac{1}{2}$ as $t \to \infty$, we have

$$P(T_a < \infty) = \lim_{t \to \infty} P(T_a \le t) = 1.$$



Now we observe that the process (not just at each fixed t)

$$B^*(t) = \begin{cases} B(t), & t \le T_a \\ B(T_a) - \tilde{B}(t - T_a), & t > T_a \end{cases} \overset{D}{=} \begin{cases} B(t), & t \le T_a \\ B(T_a) + \tilde{B}(t - T_a), & t > T_a \end{cases} = B(t),$$

since $\tilde{B} \overset{d}{=} -\tilde{B}$ and $\tilde{B}$ is indep. of $\{B(u)\}_{u \le T_a}$ (This can be made precise by using finite-dim distributions or by using functionals.)

That is, B^* is Brownian motion. $\square$

The reflection property is idiosyncratic of Brownian motion but it is also very useful for discovering a lot about the extremal nature of Brownian motion.

(extremal process)

<u>Thm. 6.9</u> Let $B(t)$ be stand. B.M. and define $M(t) = \max_{0 \le s \le t} B(s)$.

(i) For all $y \geq 0$, $a > 0$,

$$P(B(t) \le a-y,\ M(t) \geq a) = P(B(t) > a+y).$$

(ii) For each fixed $t \geq 0$, $M(t) \overset{d}{=} |B(t)|$, with mean $\sqrt{\frac{2t}{\pi}}$ and variance $(1-\frac{2}{\pi})t$.

proof. (i) $M(t) \geq a \iff T_a \le t \iff T_a^* \le t$ $\quad (T_a = T_a^*)$

where $T_a^* = \inf\{t : B^*(t) = a\}$ and B^* is as in Thm. 6.8.

So $$\begin{aligned} P(B(t) \le a-y,\ M(t) \geq a) &= P(B(t) \le a-y,\ T_a \le t) \\ &= P(B^*(t) \le a-y,\ T_a^* \le t) \quad \text{by Thm 6.8} \\ &= P(2a - B(t) \le a-y,\ T_a \le t) \quad \text{by def. of } B^* \\ &= P(B(t) \geq a+y). \quad (B(t) \geq a+y \Rightarrow T_a \le t) \end{aligned}$$

(ii) $$\begin{aligned} P(M(t) \geq a) &= P(M(t) \geq a,\ B(t) \le a) + P(M(t) \geq a,\ B(t) > a) \\ &= 2P(B(t) > a) \quad \text{by (i)} \\ &= P(|B(t)| > a) \quad \text{by symmetry of normal dist.} \end{aligned}$$

Thus $M(t) \overset{d}{=} |B(t)|$, for each fixed t. That is, $\frac{M(t)}{\sqrt{t}} \overset{D}{=} |Z|$, $Z \sim \text{normal}(0,1)$. Also, this tells us $E(M(t)) = \sqrt{\frac{2t}{\pi}}$ and $E(M^2(t)) = t$. □

<u>Cor 6.10</u> $Z_1 = \max_{0 \le t \le 1} |B(t)|$ has a finite mean.

proof. $Z_1 \le \max_{0 \le t \le 1} B(t) + \max_{0 \le t \le 1} (-B(t))$. So

$E(Z_1) \le 2E(M(1)) < \infty$ by Thm. 6.9(ii). □

Cor. 6.11 For fixed $t>0$, the joint density of $(B(t), M(t))$ is
$$f_{B(t),M(t)}(x,w) = \sqrt{\frac{2}{\pi}}\ \frac{2w-x}{t^{3/2}}\ e^{-(2w-x)^2/2t}, \qquad x\in\mathbb{R},\ w>0,\ x<w.$$

proof. The density is
$$-\frac{d^2}{dx\,dw}P(B(t)\le x, M(t)\ge w) = -\frac{d^2}{dx\,dw}P(B(t)>2w-x)$$
by Thm 6.9(i).
$$= \frac{d^2}{dx\,dw}\Phi\left(\frac{2w-x}{\sqrt{t}}\right), \text{ where } \Phi(z) = \text{std. normal cdf.} \ \square$$

Cor. 6.12 Let $T_a = \min\{t : B(t)=a\}$, $a\ge 0$, as before. Then
(i) $P(T_a\le t) = 2\left(1-\Phi\left(\frac{a}{\sqrt{t}}\right)\right)$ and $E(e^{-\lambda T_a}) = e^{-\sqrt{2\lambda}\,a}$.
 (This is the dist. of $\frac{a^2}{Z^2}$, where $Z\sim$ std. normal.)
(ii) $T_a \stackrel{D}{=} a^2 T_1$ for all $a>0$.
(iii) As a process, $\{T_a\}_{a\ge 0}$ has stationary & independent increments. (That is, it is another Lévy process.)
(iv) $\{T_a\}$ is a pure jump process increasing on every interval.
(v) $\{T_a\} \stackrel{D}{=} \{c^2 T_{a/c}\}$ (self-similarity of order $\frac{1}{2}$).

proof (i) This follows directly from $T_a\le t \Leftrightarrow M(t)\ge a$ and Thm. 6.9(ii). The Laplace transform can be found simply – but by use of a clever "trick". We do not show the calculation here.
(ii) This is immediate from (i).
(Note the difference: $B(t) \stackrel{D}{=} \sqrt{t}\,B(1)$.)
(iii) By use of induction, it suffices to show $T_b - T_a$ is indp. of $\{T_u\}_{u\le a}$ and has the same dist. as T_{b-a}.

Let $\tilde{B}(t) = B(T_a + t) - a$, $t \geq 0$, which is a B.M., by Thm. 6.7. Let $\tilde{T}_{b-a} = \min\{t : \tilde{B}(t) = b-a\}$. Then $T_b = \tilde{T}_{b-a} + T_a$, which says

$$T_b - T_a = \tilde{T}_{b-a} = T_{b-a}, \text{ since } \tilde{B}(t) \text{ is a std B.M.}$$

like $B(t)$. Moreover, $\tilde{T}_{b-a}$ is independent of $\{B_t\}_{t \leq T_a}$ (since $\tilde{B}$ is) & thus $T_b - T_a$ is independent of $\{B_t\}_{t \leq T_a}$. However, T_u, $u \leq a$, is completely characterized by B_t, $t \leq T_a$. So $T_b - T_a$ is independent of $\{T_u\}_{u \leq a}$.

(iv) Here is a heuristic explanation. Let $T_{a-} = \sup\{T_b : b < a\}$. Consider that the extremal process $M(t)$ can have flat stretches w/ pos prob. (since it can take a while for $B(t)$ to return to its previous maximum value). This means $P(T_a - T_{a-} > 0) > 0$. By using the self-similarity property of $B(t)$ (Thm. 6.5(v)) it can be shown that in fact

$P(T_a - T_{a-} > 0) = 1$.

(v) This also follows from the self-similarity of $\{B(t)\}$ (exercise). ☐

Caution: it is not generally true that $|X_t|$ is a Markov process when X_t is Markov.

<u>Thm. 6.13</u> $\{|B(t)|\}$ is a homogeneous Markov process.
This is B.M. repeatedly reflected at 0.

proof. The joint densities for $\{|B(t)|\}$ are $(t_1 < t_2 < \cdots < t_k)$

$$f_{|B(t_1)|, \ldots, |B(t_k)|}(y_1, \ldots, y_k) = \sum_{\substack{\delta_i = \pm 1 \\ i \leq k}} f_{B(t_1), \ldots, B(t_k)}(\delta_1 y_1, \ldots, \delta_k y_k)$$

$$= \sum_{\substack{\delta_i = \pm 1 \\ i \leq k}} f_{B(t_1), \ldots, B(t_{k-1})}(\delta_1 y_1, \ldots, \delta_{k-1} y_{k-1}) \, \varphi\!\left(\frac{\delta_k y_k - \delta_{k-1} y_{k-1}}{\sqrt{t_k - t_{k-1}}}\right) \frac{1}{\sqrt{t_k - t_{k-1}}}$$

$$= \sum_{\substack{\delta_i = \pm 1 \\ i \leq k-1}} f_{B(t_1), \ldots, B(t_{k-1})}(\delta_1 \ldots, \delta_{k-1} y_{k-1}) \frac{1}{\sqrt{t_k - t_{k-1}}} \left(\varphi\!\left(\frac{y_k - y_{k-1}}{\sqrt{t_k - t_{k-1}}}\right) + \varphi\!\left(\frac{y_k + y_{k-1}}{\sqrt{t_k - t_{k-1}}}\right) \right)$$

$$= f_{|B(t_1)|,\dots,|B(t_{k-1})|}(y_1,\dots,y_{k-1}) \underbrace{\left(\varphi\left(\frac{y_k - y_{k-1}}{\sqrt{t_k - t_{k-1}}}\right) + \varphi\left(\frac{y_k + y_{k-1}}{\sqrt{t_k - t_{k-1}}}\right)\right)\frac{1}{\sqrt{t_k - t_{k-1}}}}$$

This must be the cond. density of $|B(t_k)|$, given $|B(t_i)| = y_i$, $i < k$. Since it depends only on y_{k-1}, $\{|B(t)|\}$ is Markov and its transition density is

$$p_s(y|x) = \frac{d}{dy} P(X(s) \le y \,|\, X(0) = x)$$
$$= \frac{1}{\sqrt{t_k - t_{k-1}}}\left(\varphi\left(\frac{y-x}{\sqrt{s}}\right) + \varphi\left(\frac{y+x}{\sqrt{s}}\right)\right). \quad \square$$

We now turn to the Brownian bridge process.

<u>Thm 6.14</u> Let $\{B(t)\}$ be stand. B.M. and define the <u>Brownian bridge process</u> $B^{(0)}(t) = B(t) - tB(1)$, $0 \le t \le 1$. Then $\{B^0(t)\}$ is a Gaussian process (i.e. has multivariate normal finite dimensional distributions) with mean 0 and covariance function

$$\operatorname{cov}(B^{(0)}(t_1), B^{(0)}(t_2)) = t_1(1-t_2) \quad \text{for } 0 \le t_1 \le t_2 \le 1.$$

proof. $(B^{(0)}(t_1), \dots, B^{(0)}(t_k))$ is clearly multivariate normal for any $t_1, \dots, t_k$ and has mean 0.

Let $0 \le t_1 \le t_2 \le 1$. Then

$$\operatorname{cov}(B^{(0)}(t_1), B^{(0)}(t_2)) = E\left((B(t_1) - t_1 B(1))(B(t_2) - t_2 B(1))\right)$$
$$= \min(t_1, t_2) - t_1 \min(1, t_2) - t_2 \min(1, t_1) + t_1 t_2$$
$$= t_1(1-t_2). \quad \square$$

<u>Remark</u> Let $\{X(t)\}$ be a Gaussian process and let $g(t)$ be nondecreasing. Then $\{X(g(t))\}$ is a Gaussian process also. In particular, $\{B^{(0)}(g(t))\}$ has covariance function

$$\operatorname{cov}(B^{(0)}(g(t_1)), B^{(0)}(g(t_2))) = g(t_1)(1-g(t_2)), \quad \text{if } t_1 \le t_2.$$

<u>Example 6.2</u> Suppose $X_1, X_2, \ldots \sim$ iid F. To estimate F we can use the <u>empirical distribution function</u>

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(X_i) \qquad (= \text{rel. freq. of data not greater than } x).$$

Since $1_{(-\infty,x]}(X_i) \sim$ iid Bernoulli $(p = F(x))$, we know that

$$\sqrt{n}\,(F_n(x) - F(x)) \Rightarrow \text{normal}\,(0, F(x)(1-F(x)))$$

by the central limit theorem.

Now suppose that $x_1 \le x_2$. Then

$$\text{cov}\left(1_{(-\infty,x_1]}(X_1),\ 1_{(-\infty,x_2]}(X_1)\right) = F(x_1)(1-F(x_2)) \quad (\text{check})$$

Thus $\sqrt{n}\left[\begin{pmatrix} F_n(x_1) \\ F_n(x_2)\end{pmatrix} - \begin{pmatrix} F(x_1) \\ F(x_2)\end{pmatrix}\right] \Rightarrow$ bivariate normal w/ the appropriate covariance matrix.

In fact we have the following.

<u>Thm. 6.15</u> Let $X_1, X_2, \ldots \sim$ iid F and let $F_n(\cdot)$ be the empirical dist. function. Then

$$\sqrt{n}\,(F_n(\cdot) - F(\cdot)) \Rightarrow B^{(0)}(F(\cdot))$$

in the finite dimensional distribution sense.

(Actually, convergence can be stated in a stronger sense as well.)

proof. Extending Ex. 6.2 above, if $x_1 < x_2 < \cdots < x_k$,

$$\sqrt{n}\left[\begin{pmatrix} F_n(x_1) \\ \vdots \\ F_n(x_k)\end{pmatrix} - \begin{pmatrix} F(x_1) \\ \vdots \\ F(x_k)\end{pmatrix}\right] \Rightarrow \text{mult. normal}\,(0, \Sigma)$$

where, for $i \le j$, the $(i,j)^{\text{th}}$ element of Σ is $F(x_i)(1-F(x_j))$.

That is, the finite dim. dist's of $\sqrt{n}\,(F_n(\cdot) - F(\cdot))$ converge to the finite dim. dists. of $B^{(0)}(F(\cdot))$. $\square$

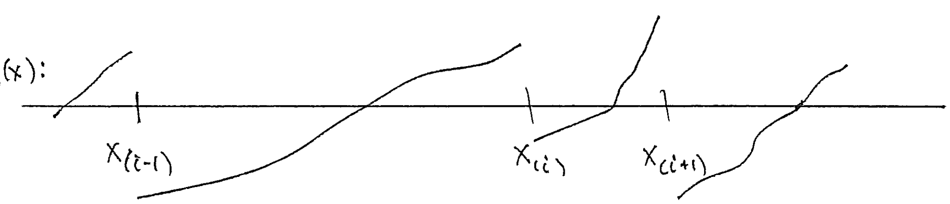
Def. 6.16 Let $X_1, \ldots, X_n \sim$ iid F. The Kolmogorov-Smirnov statistic (for testing $H_0 : F = F_0$) is

$$D_n = \sqrt{n} \sup_x |F_n(x) - F_0(x)|$$

$$= \sqrt{n} \max_{1 \le i \le n} \left(\max \left(\left| F_0(X_{(i)}) - \frac{i}{n} \right|, \left| F_0(X_{(i)}) - \frac{i-1}{n} \right| \right) \right),$$

where $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ are the order statistics of the sample. The latter formula is correct because $F_n(x)$ is a step function with jumps at $X_{(i)}$'s. So for $x \in (X_{(i)}, X_{(i-1)})$, $F_0(x) - F_n(x)$ is a nondecreasing function with maximum absolute value at one of the endpoints ($X_{(i)}$ or $X_{(i-1)}$).

$F_0(x) - F_n(x)$:



Theorem 6.17 Assume $X_1, X_2, \ldots \sim$ iid F where F is continuous and let D_n be the Kolmogorov-Smirnov statistic. If $F = F_0$, then

$$D_n \Rightarrow D \quad \text{where } P(D > y) = 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 y^2}.$$

proof. First we verify a very useful result about $F(X_1)$. Let $F^{\leftarrow}$ be the inverse of F (see Def. 4.9). Then, if $0 < u < 1$,

$$F(x) \ge u \iff x \ge F^{\leftarrow}(u).$$

Therefore, $P(F(X_1) \ge u) = P(X_1 \ge F^{\leftarrow}(u))$

$$= 1 - F(F^{\leftarrow}(u)) = 1 - u \quad \text{(by Thm. 4.10)}.$$

That is, $F(X_1) \sim \text{Uniform}(0,1)$.

It follows that $F(X_{(1)}), \ldots, F(X_{(n)})$ are distributed like the order statistics for a sample from the Uniform$(0,1)$ dist.

Hence, if $F_0 = F$, the distribution of D_n is independent of F. (That is, D_n is a "pivotal" statistic.) It suffices, therefore, to continue as if $F(x) = x$, $0 \le x \le 1$, and $X_i \sim$ iid Uniform(0,1).

Thus, we have

$$\sqrt{n}\,(F_n(\cdot) - F(\cdot)) \Rightarrow B^{(0)}(F(\cdot)) \text{ in the fin. dim. dist. sense.}$$

If we knew we had a stronger, functional convergence (as in Thm. 6.4) we could deduce that

$$D_n = \max_{0 \le t \le 1} |\sqrt{n}\,(F_n(t) - t)| \Rightarrow \max_{0 \le t \le 1} |B^{(0)}(t)|.$$

(Of course, such a result is well-known. See the book for a proof which determines the convergence above directly by applying Thm. 6.4.)

Calculating the distribution of $\max_{0 \le t \le 1} |B^{(0)}(t)|$ is tedious but it may be shown it has the probability tail given in the theorem statement. $\square$

We now turn to Brownian motion with drift.

Def. 6.18 Let $B(t)$ be standard B.M. and $B_\mu(t) = B(t) + \mu t$, $t \ge 0$. The process $\{B_\mu(t)\}$ is called Brownian motion with drift μ. (This also is a Lévy process – check. Hence, it is Markov.)

Obviously, $B_\mu(t) \sim$ normal $(\mu t, t)$ for each t.

More generally, set $B_\mu(t) = B(t) + \mu t + X_0$, X_0 indep. of $B(\cdot)$. So this is B.M. with drift μ and initial state X_0. It follows that $B_\mu(t)$ is a homogeneous Markov process with

$$f_{B_\mu(t_2) | B_\mu(t_1) = x}(y) = \frac{1}{\sqrt{t_2 - t_1}}\, \varphi\left(\frac{y - \mu(t_2 - t_1) - x}{\sqrt{t_2 - t_1}}\right) \text{ (exercise).}$$

For simplicity, let the cond. pdf of $B_\mu(t)$, given $B_\mu(0)=x$, be written

$$f_t(y;x) = \frac{1}{\sqrt{t}}\varphi\left(\frac{y-\mu t-x}{\sqrt{t}}\right) = \frac{1}{\sqrt{2\pi t}} e^{-(y-\mu t-x)^2/2t}.$$

<u>Thm.6.19</u> (i) $f_t(y;x)$ solves the Fokker-Planck equations

$$\frac{\partial}{\partial t} f_t(y;x) = \mu\frac{\partial}{\partial x} f_t(y;x) + \frac{1}{2}\frac{\partial^2}{\partial x^2} f_t(y;x)$$

$$\text{and also} \quad = -\mu\frac{\partial}{\partial y} f_t(y;x) + \frac{1}{2}\frac{\partial^2}{\partial y^2} f_t(y;x).$$

(ii) In fact, it is the unique cond. pdf to solve the equations such that $f_t(y;x)\to 0$ and $\frac{\partial}{\partial t} f_t(y;x)\to 0$ as $t\to 0$.

proof (i) (exercise)

(ii) This is a familiar example from partial differential equations, for which there are several different methods of proof. We do not give a demostration here. □

Diffusion properties

$$\frac{1}{t}E\left(B_\mu(t)-B_\mu(0)\,\middle|\,B_\mu(0)=x\right)=\mu \quad \text{for all } x$$

$$\frac{1}{t}Var\left(B_\mu(t)-B_\mu(0)\,\middle|\,B_\mu(0)=x\right)=1 \quad \text{for all } x.$$

Specifically, over a small time interval t, B_μ is moved, on average, by an amount μt. But it is also perturbed by a random amount with variance t. The Fokker-Planch equations, desribing the perturbation of the pdf of B_μ, given $B_\mu(0)=x$, are reflective of these diffusion properties.

We define hitting times the same as we did for B.M.

$$T_a = \inf\{t : B_\mu(t) = a\}, \text{ for any } a \in \mathbb{R}.$$

These are stopping times, of course.

Again we would like to compute hitting time probabilities.

<u>Thm 6.20</u> Suppose $a \le x \le b$. Then

$$P(T_a < T_b \mid B_\mu(0) = x) = \begin{cases} \dfrac{e^{-2\mu b} - e^{-2\mu x}}{e^{-2\mu b} - e^{-2\mu a}} & \text{if } \mu \neq 0. \\ \dfrac{b}{a+b} & \text{if } \mu = 0 \end{cases}$$

<u>Proof.</u> (heuristic) Define $\tilde{B}_\mu(t) = B_\mu(t+h)$.
So $\tilde{B}_\mu(t)$ is B.M. w/ drift μ & initial state $B_\mu(h)$.
Furthermore, it is independent of $B_\mu(s), s \le h$.

Suppose $a < x < b$ and define $g(x) = P(T_a < T_b \mid B_\mu(0) = x)$.
We will find $g(x)$ by showing it satisfies a differential equation.

$\{T_a < T_b\}$ is the event that a is visited before b.
Consider first the possibility that either a or b are visited quickly. For small h,

$$P(\min(T_a, T_b) \le h) = P\left(\max_{s \le h}(B(s) + \mu s) \ge b - x \text{ or } \min_{s \le h}(B(s) + \mu s) \le a - x\right)$$

$$\le P(M(h) \ge b - x - |\mu| h) + P(M(h) \ge x - a - |\mu| h) \quad \left(M(t) = \max_{s \le t} B(s)\right)$$

$= o(h)$, using the dist. for $M(h)$ as given in Thm. 6.9.

This says we can safely compute, for very small h, as if neither a nor b have been visited in $[0,h]$.

That is, conditioning on $B_\mu(h)=y$,

$$g(x) = \int_{-\infty}^{\infty} P(B_\mu \text{ visits a before b} \mid B_\mu(h)=y)\, f_{B_\mu(h)}(y)\, dy$$

$$= \int_{-\infty}^{\infty} P(\tilde{B}_\mu \text{ visits a before b} \mid B_\mu(h)=y)\, f_{B_\mu(h)}(y)\, dy + o(h)$$

$$= \int_{-\infty}^{\infty} g(y) f_{B_\mu(h)}(y)\, dy + o(h), \text{ since } \tilde{B}_\mu \text{ has initial state } y.$$

$$= \int_{-\infty}^{\infty} \left(g(x) + (y-x)g'(x) + \frac{(y-x)^2}{2} g''(x) \right) f_{B_\mu(h)}(y)\, dy + o(h)$$

using Taylor's expansion

$$= g(x) + g'(x)\, E(B_\mu(h)-x) + \tfrac{1}{2} g''(x)\, E((B_\mu(h)-x)^2) + o(h)$$
$$= g(x) + \mu h g'(x) + \frac{h}{2} g''(x) + \frac{(\mu h)^2}{2} g''(x) + o(h)$$

Subtracting the left-hand side and dividing by h, we get

$$0 = \mu g'(x) + \tfrac{1}{2} g''(x) + \tfrac{1}{h} o(h), \text{ where } \tfrac{1}{h} o(h) \to 0.$$

So $g''(x) = -2\mu g'(x)$, which implies $g'(x) = c_1 e^{-2\mu x}$ and thus (for $\mu \neq 0$)

$$g(x) = \frac{-c_1}{2\mu} e^{-2\mu x} + c_2, \text{ for some } c_1 \,\&\, c_2.$$

Plugging in the obvious initial conditions

$$g(a)=1 \text{ and } g(b)=0$$

gives the values for $c_1 \,\&\, c_2$, and hence the result.

The case $\mu=0$ is similar, or can be obtained by letting $\mu \to 0$. □

Cor. 6.21 Again assume $a<x<b$. Then

$$P(T_a < \infty \mid B_\mu(0)=x) = P(\sup_t B_\mu(t) \geq b) = \begin{cases} 1 & \text{if } \mu \leq 0 \\ e^{-2\mu(x-a)} & \text{if } \mu > 0. \end{cases}$$

proof. use Thm 6.20 and take $\lim_{b\to\infty} P(T_a < T_b \mid B_\mu(0)=x)$. □

The Fokker-Planck equation relates the time movement of the process to the value movement of the process "instantaneously". This is characteristic of diffusion processes.

<u>Def. 6.22</u> A Markov process $\{X(t)\}$ with values in an interval of $\mathbb{R}$ is a <u>diffusion</u> with <u>instantaneous drift</u> $\mu(x)$ and <u>instantaneous diffusion coefficient</u> $\sigma^2(x)$ if

(i) it has continuous sample paths w.p. 1

and (ii) $P(|X(t)-x|>\varepsilon \mid X(0)=x) = o(t)$ for each $\varepsilon>0$,

$E((X(t)-x) \mid X(0)=x) = \mu(x)t + o(t)$,

$E((X(t)-x)^2 \mid X(0)=x) = \sigma^2(x)t + o(t)$, as $t \downarrow 0$.

$1_{|X(t)-x|\le\varepsilon}$ can be inserted to ensure finite expectations.

<u>Ex. 6.4</u> B.M. w/ drift μ. Let $Z \sim$ normal$(0,1)$.

So $B_\mu(t) \mid B_\mu(0)=x \overset{d}{=} x + \mu t + \sqrt{t}\,Z$.

then

check, using $Z \sim$ normal$(0,1)$

$$P(|B_\mu(t)-x|>\varepsilon \mid B_\mu(0)=x) = P\left(|Z+\sqrt{t}\mu| > \frac{\varepsilon}{\sqrt{t}}\right) = o(t),\ t\downarrow 0.$$

$$E((B_\mu(t)-x) \mid B_\mu(0)=x) = \mu t,$$ so $\mu(x)=\mu$ for all x.

$$E((B_\mu(t)-x)^2 \mid B_\mu(0)=x) = \mu^2t^2 + t = t + o(t).$$

So $\sigma^2(x)=1$ for all x.

<u>Ex. 6.5</u> Ornstein-Uhlenbeck process. Let $\alpha>0, \beta>0$.

$X(t) = e^{-\alpha t/2} B(\beta e^{\alpha t})$, where $B(\cdot)$ is stand. B.M.

This is a Gaussian process (fin. dim. dist. are mult. normal) with mean 0 and

Note: $X(0)=B(\beta) \sim$ normal$(0,\beta)$

$$\mathrm{cov}(X(s),X(t)) = \beta e^{-\alpha s/2-\alpha t/2} \min(e^{\alpha s}, e^{\alpha t})$$
$$= \beta\, e^{-\alpha|s-t|/2}.$$

So the process is stationary. $(X(\cdot+t) \overset{d}{=} X(\cdot))$

For $t>0$, $X(t) \mid X(0)=x \overset{D}{=} \frac{\mathrm{cov}(X(t),X(0))}{\mathrm{var}(X(0))} x + \left(\mathrm{var}(X(t)) - \frac{\mathrm{cov}^2(X(t),X(0))}{\mathrm{Var}(X(0))}\right)^{1/2} Z$

(regression equation for bivariate normal rvs)

$$= e^{-\alpha t/2} x + \sqrt{\beta\,(1-e^{-\alpha t})}\, Z \quad \text{(Check)}$$

so $X(t) - x \overset{D}{\doteq} \underbrace{-\frac{\alpha x}{2}}_{\mu(x)} t + \underbrace{\sqrt{\beta\,\alpha}}_{\sigma(x)\text{, which is constant}} \sqrt{t}\, Z$ as $t \downarrow 0$.

↑ note neg. drift when $x>0$, pos. drift when $x<0$ } this keeps the process from wandering away from 0

The Ornstein-Uhlenbeck process has a number of applications

(1) velocity of large solute molecule in a liquid.
due to friction the expected (instant.) change in velocity is proportional to the velocity & in the opposite direction: $E(\Delta X(t)) \doteq -\frac{\alpha}{2} X(t)$
due to randomness (friction is result of many small particles), the change in velocity is variable: $\mathrm{var}(\Delta X(t)) \doteq \sigma \Delta t$.

(2) as a limit of the Ehrenfest model for heat exchange (in same way that B.M. w/ drift is limit of S.R.W.)
finite state M.C. with

$$P_{ij} = \begin{cases} i/2m, & j=i-1 \\ 1-i/2m, & j=i+1 \\ 0, & \text{o.w.} \end{cases} \quad \text{for } i=1,\dots,m-1$$

and $P_{01} = P_{m,m-1} = 1$ (reflecting boundaries)

(3) Cont. time analog of a AR(1) time series model.

$$Y_n = a Y_{n-1} + \varepsilon_n, \quad 0<a<1, \quad \varepsilon_1, \varepsilon_2, \dots \text{ iid w/ variance } \gamma$$

if $1-a \doteq \frac{\alpha}{2m}$, $\gamma \doteq \frac{\alpha}{m}$, $\beta = 1$,

then $X_m(t) = Y_{[mt]} \Longrightarrow$ Ornstein-Uhlenbeck process, as $m \to \infty$.

check: $\mathrm{var}(Y_0) = \mathrm{var}(Y_1) = \alpha^2 \mathrm{Var}(Y_0) + \gamma \Rightarrow \mathrm{Var}(Y_0) = \frac{\gamma}{1-\alpha^2} \to 1$ as $m \to \infty$

$\mathrm{cov}(X_m(t), X_m(0)) = \mathrm{cov}(Y_0, Y_{[mt]}) = a^{[mt]} \mathrm{Var}(Y_0)$ (by induction)

$\to e^{-\alpha t/2}$.

Ex. 6.6 Suppose $X(t) = X_0 + \sigma B(t) + \mu t$ where X_0 & $B(\cdot)$ are independent, $B(\cdot)$ is B.M. Define

$$Y(t) = e^{X(t)}.$$

This is called geometric Brownian motion. Since $x \to e^x$ is a 1-1 function, $Y(t)$ is a Markov process. It has continuous sample paths. Therefore $Y(t)$ is a diffusion.

Using the moment general function for the normal distribution: $W \sim \text{normal}(\mu, \sigma^2)$ then $E(e^{\lambda W}) = e^{\lambda\mu + \lambda^2 \frac{\sigma^2}{2}}$, we can compute

$$E(Y(t) - Y(0) \mid Y(0) = y) = E(y e^{\mu t + \sigma B(t)} - y)$$

$$= y(e^{\mu t + \sigma^2 t/2} - 1). \quad (\text{since } E(\mu t + \sigma B(t)) = \mu t \text{ \& } \operatorname{var}(\mu t + \sigma B(t)) = \sigma^2 t)$$

Thus

$$\frac{1}{t} E(Y(t) - Y(0) \mid Y(0) = y) = y \, \frac{e^{(\mu + \sigma^2/2)t} - 1}{t}$$

$$\to y(\mu + \sigma^2/2) \quad \text{as } t \to 0.$$

This is the drift coefficient function for $Y(t)$. Note that it depends on σ^2 as well as on μ.

Getting the diffusion coefficient is a little more involved. First, analogous to the above,

$$\frac{1}{t} E(Y^2(t) - Y^2(0) \mid Y(0) = y) = \frac{y^2}{t} E(e^{2\mu t + 2\sigma B(t)} - 1)$$

$$= \frac{y^2}{t}(e^{2\mu t + 2\sigma^2 t} - 1) \to 2y^2(\mu + \sigma^2).$$

Second,

$$\frac{1}{t} E((Y(t) - Y(0))^2 \mid Y(0) = y)$$

$$= \frac{1}{t} E(Y^2(t) - Y^2(0) \mid Y_0 = y) - \frac{2}{t} E(Y(0)(Y(t) - Y(0)) \mid Y(0) = y)$$

$$\to 2y^2(\mu + \sigma^2) - 2y^2(\mu + \sigma^2/2)$$

$$= \sigma^2 y^2.$$

The key to understanding a diffusion process is the relationship between time changes and state changes, the so-called infinitesimal behavior of the process.

<u>Thm 6.23</u> Suppose $X(t)$ is a diffusion w/ drift function $\mu(x)$ and diffusion function $\sigma^2(x)$. We assume these are continuous with "nice" derivatives at almost all x. Suppose $f(x)$ is smooth - it has two continuous derivatives and $f''(x)$ is bounded.

$$\frac{\partial}{\partial t} E(f(X(t))-f(X(0)) \mid X(0)=x)\Big|_{t=0}$$
$$= \mu(x) f'(x) + \frac{\sigma^2(x)}{2} f''(x).$$

The operator A that maps $f(x)$ to this limit is called the <u>infinitesimal generator</u> of the process.

Under some regularity conditions, A uniquely characterizes the distribution of the diffusion process.

proof (very heuristic) This is similar to a couple examples above. By a second order Taylor's expansion,

$$\frac{1}{t} E\left(f(X(t))-f(X(0)) \mid X(0)=x\right)$$
$$\doteq \frac{1}{t} E\left(f'(x)(X(t)-X(0)) \mid X(0)=x\right) + \frac{1}{t} E\left(f''(x)\tfrac{1}{2}(X(t)-X(0))^2 \middle| X(0)\right.$$
$$\xrightarrow{\text{as } t\downarrow 0} f'(x)\mu(x) + f''(x)\tfrac{1}{2}\sigma^2(x)$$ by the definition of diffusion process. $\square$

The derivative (limit) in Thm. 6.23 can hold for more general functions, depending on the particular process.

Ex. 6.5 (cont.) Ornstein-Uhlenbeck process $X(t)$ with $\mu(x) = -\frac{\alpha x}{2}$ and $\sigma^2(x) = \alpha\beta$. Let $\lambda = \frac{\alpha}{2}$, $\sigma^2 = \alpha\beta$. Let $f(x) = e^{rx}$. Although the derivatives are unbounded, the O-U process is sufficiently nice for the result to hold.

$$Af(x) = \mu(x) f'(x) + \frac{\sigma^2(x)}{2} f''(x)$$
$$= -\lambda x f'(x) + \frac{\sigma^2}{2} f''(x)$$

$$= \frac{1}{2}(-\alpha r x + \alpha\beta r^2) e^{rx}.$$

So

$$\lim_{t\to 0} \frac{1}{t} E\left(e^{rX(t)} - e^{rx} \middle| X(0) = x\right) =$$

In this case, since $X(t)$ has a normal distribution, one can also check this directly using the mgf for normal rvs.

Ex. 6.6 (cont) Geom. B.M. $Y(t)$ with

$$\mu(y) = (\mu + \tfrac{\sigma^2}{2}) y, \quad \sigma^2(y) = \sigma^2 y^2.$$

Suppose $f(y) = \log(1+y)$. Then

$$\frac{1}{t} E\left(\log(1+Y(t)) - \log(1+y) \middle| Y(0) = y\right)$$

$$\to (\mu + \tfrac{\sigma^2}{2}) y \frac{1}{1+y} - \frac{\sigma^2 y^2}{2} \frac{1}{(1+y)^2}, \text{ as } t \to 0.$$

Thm 6.24 (Kolmogorov's Backward & Forward Equations)
Assume $X(t)$ & $f(x)$ are as in Thm 6.23. Let

$$g_t(x) = E(f(X(t)) \mid X(0)=x) \quad (\text{often denoted } (P^t f)(x))$$

(i) $$\frac{\partial}{\partial t} g_t(x) = \mu(x) \frac{\partial}{\partial x} g_t(x) + \frac{\sigma^2(x)}{2} \frac{\partial^2}{\partial x^2} g_t(x).$$

(ii) $$\frac{\partial}{\partial t} g_t(x) = E\left(-\mu(X(t)) f'(X(t)) + \frac{\sigma^2(X(t))}{2} f''(X(t)) \,\middle|\, X(0)=x\right)$$

The 1^{st} is called the "backward" equation as it references perturbations of the initial state x. The 2^{nd}, which refers to $X(t)$ is called the forward equation.

proof. These can be argued heuristically in a manner similar to the proof of Thm. 6.23.

Thm 6.24 is used to prove many results about diffusions. The concept actually extends to (most) continuous time Markov processes, including those with countable state space - if the infinitesimal generator A is defined appropriately.
The next result is one of the principle applications.

Thm 6.25 Let $T_x = \min\{t : X(t) = x\}$. Assume as before and that $\sigma^2(x) > 0$ for all $x \in (c,d)$. For $c < a \le x \le b < d$,

$$P(T_a < T_b \mid X(0)=x) = \frac{\int_x^b e^{-I(y)}\,dy}{\int_a^b e^{-I(y)}\,dy},$$

by convention, $\int_x^y g(z)\,dz = -\int_y^x g(z)\,dz$

where $I(y) = \int_{x_0}^{y} \frac{2\mu(z)}{\sigma^2(z)}\,dz$ (x_0 can be any fixed value in the state space.)

Note: $s'(x) = e^{-I(x)}$ is called the <u>scale density</u> and $s(x) = \int_{x_0}^{x} s'(y)\,dy$ is called the <u>scale function</u>.

Ex. 6.4 (cont.) O-U process. For simplicity, let $\lambda = \frac{\alpha}{2}$ and $\sigma^2 = \alpha\beta$. Then $\mu(x) = -\lambda x$ and $\sigma^2(x) = \sigma^2$ (constant)

We can choose $x_0 = 0$. Then

$$I(y) = \int_0^y \frac{-2\lambda z}{\sigma^2}\,dz = -\frac{\lambda y^2}{\sigma^2} \quad \text{and}$$

$$s(y) = \int_0^y e^{\lambda z^2/\sigma^2}\,dz.$$

So $P(T_a < T_b \mid X(0)=x) = \dfrac{\int_x^b e^{\lambda y^2/\sigma^2}\,dy}{\int_a^b e^{\lambda y^2/\sigma^2}\,dy}$, $a \le x \le b$.

Letting $b \to \infty$, we can see that $P(T_a < \infty \mid X_0 = x) = 1$.
Likewise, $P(T_b < \infty \mid X(0) = x) = 1$.

proof of Thm 6.25. (again, heuristic)
Let $g(x) = P(T_a < T_b \mid X(0) = x)$.

1^{st} step. Similar to our earlier proof (Thm. 6.20) for B.M. with drift, let $X_s(t) = X(s+t)$ for small s

$$P(\min(T_a, T_b) \le s \mid X(0)=x) = o(s) \quad \text{as } s \to 0$$

So

$$\begin{aligned} g(x) &= P(T_a < T_b, \min(T_a,T_b) > s \mid X(0)=x) + o(s) \\ &= E\big(P(X(s+t) \text{ visits } a \text{ before } b \mid X(s)) \mid X(0)=x\big) + o(s) \\ &= E(g(X(s)) \mid X(0)=x) + o(s) \end{aligned}$$

Letting $s \to 0$,

$$\frac{\partial}{\partial s} E\big(g(X(s)) - g(X(0)) \mid X(0)=x\big) = 0.$$

2^{nd} step. By Thm 6.23, we obtain

$$0 = \mu(x)\,g'(x) + \frac{\sigma^2(x)}{2}\,g''(x).$$

(We are assuming we know that $g''(x)$ is cont & bounded on the interval $[a,b]$.)

That is, $\quad g''(x) = -\frac{2\mu(x)}{\sigma^2(x)}\,g'(x).$

This implies $g'(x) = c_1 s'(x)$ and $g(x) = c_0 + c_1 s(x)$ for some constants c_0 & c_1.

3^{rd} step. Using the initial conditions $g(a)=1$ and $g(b)=0$, we can solve for c_0 & c_1. Plugging those values in leads to the desired conclusion. $\square$

Def. 6.26. Let $\mathbb{X}$ be the state space for $X(t)$ (an interval subset of $\mathbb{R}$).

(i) $X(t)$ is irreducible if $P(\tilde{T}_y < \infty \mid X(0)=x) > 0$ all $x,y \in \mathbb{X}$.
(ii) $X(t)$ is recurrent if $P(T_y < \infty \mid X(0)=x) = 1$ all $x,y \in \mathbb{X}$.
(iii) $X(t)$ is positive recurrent if $E(T_y \mid X(0)=x) < \infty$ all $x,y \in \mathbb{X}$.

Note the distinction w/ definitions for countable state Markov processes (for which only $y=x$ was required). Most diffusions do not have first return times because $X(0)=x \Rightarrow$ infinitely many visits to x in $[0,s]$, for any $s>0$.

Thm. 6.27 Assume $\mu(x)$ & $\sigma^2(x)$ are continuous and $\sigma^2(x) > 0$ for all $x \in \mathbb{X}$. Then $X(t)$ is irreducible.

Henceforth we assume $\mathbb{X} = (x_L, x_u)$ is an open interval. (If $\mathbb{X}$ is closed at either endpoint, or both, there is the issue of what happens when the endpoint is reached. The process could either reflect or be absorbed at that state. Dealing with these cases is a little more complicated.)

Thm. 6.28 Assume as in Thm. 6.25, with $\mathbb{X} = (x_L, x_u)$ and $\sigma^2(x) > 0$ for all x. Suppose $x_L < a < x < b < x_u$.

(i)
$$P(T_a < \infty \mid X_0 = x) = \begin{cases} \dfrac{\int_x^{x_u} s'(y)\,dy}{\int_a^{x_u} s'(y)\,dy} & \text{if the integrals are finite} \\ 1 & \text{o.w.} \end{cases}$$

$$P(T_b < \infty \mid X_0 = x) = \begin{cases} \dfrac{\int_{x_L}^{x} s'(y)\,dy}{\int_{x_L}^{b} s'(y)\,dy} & \text{if the integrals are finite} \\ 1 & \text{o.w.} \end{cases}$$

(ii) $X(t)$ is recurrent iff both
$$\int_{x_0}^{x_u} s'(y)\,dy = \infty \quad \text{and} \quad \int_{x_L}^{x_0} s'(y)\,dy = \infty.$$

proof. Just use Thm 6.25 and let $b \to x_u$ (for T_a) and $a \to x_L$ (for T_b). □

Ex. 6.5 (cont.) O-U process. We saw earlier that $s'(y) = e^{\lambda y^2/\sigma^2}$ and that $P(T_y \mid X(0) = x) = 1$ for all x, y.

Ex. 6.6 (cont.) Geom BM. $\mu(x) = (\mu + \sigma^2/2)x$, $\sigma^2(x) = \sigma^2 x^2$.
So (using $x_0 = 1$) Note: $x_L = 0$, $x_u = \infty$.

$$I(x) = \int_1^x \frac{(2\mu+\sigma^2)y}{\sigma^2 y^2} = \left(1 + \frac{2\mu}{\sigma^2}\right)\log x$$

$$s'(x) = e^{-I(x)} = x^{-(1+2\mu/\sigma^2)}.$$

Then

$$\int_1^\infty s'(x)dx = \infty \iff \mu \le 0 \quad (P(T_y<\infty|X(0)=x)=1 \text{ for } y<x)$$

$$\int_0^1 s'(x)dx = \infty \iff \mu \ge 0. \quad (P(T_y<\infty|X(0)=x)=1 \text{ for } y>x)$$

So Geom BM is recurrent $\iff \mu = 0$.
(Actually, this is already apparent from the fact that Geom BM $= g(B_\mu(t))$ for a 1-1 function $g(x)$.)

Next we look at positive recurrence and stationary distributions. For this we need another function.

Def. 6.29. The speed density is $m'(x) = \frac{2}{\sigma^2(x)}e^{I(x)}$. Use the same x_0 as in defining $I(x)$ and $s(x)$ previously
The speed function is $m(x) = \int_{x_0}^x m'(y)dy$.

Note: $\sigma^2(x)s'(x)m'(x) = 2$ for all x.

Thm. 6.30 Assume as before, and that $X(t)$ is recurrent.
(i) For $y<x$, $E(T_y|X(0)=x)<\infty \iff \int_{x_0}^{x_u} m'(y)dy < \infty$.
For $y>x$, $E(T_y|X(0)=x)<\infty \iff \int_{x_L}^{x_0} m'(y)dy < \infty$.

(ii) If $C = \int_{x_L}^{x_u} m'(y)dy < \infty$ (and $X(t)$ is recurrent) then $X(t)$ is positive recurrent with stationary density $\frac{m'(x)}{C}$, Ex. 6.5 (cont.) O-U process. We saw above that this process is recurrent. Since $\mu(x) = -\lambda x$, $\sigma^2(x) = \sigma^2$ we find $m'(x) = \frac{2}{\sigma^2} e^{-\lambda x^2/\sigma^2}$, which is integrable on $\mathbb{R}$. Therefore, $X(t)$ is positive recurrent with normal$(0, \frac{\sigma^2}{2\lambda})$ stationary distribution.

Ex. 6.6 (cont.) Geom BM w/ $\mu = 0$ is recurrent. However,
$$m'(x) = \frac{2}{\sigma^2 x^2} x = \frac{2}{\sigma^2 x}$$ is not integrable on either $(1, \infty)$ or $(0,1)$. So the process is null recurrent.

Ex. 6.7. (Feller's square root process) This model has been used in finance. We let $\mathbb{X} = (0, \infty)$ and
$$\mu(x) = \lambda(\alpha - x),\ \sigma^2(x) = \beta x\ \left(\sigma(x) = \sqrt{\beta x}\right),\ \lambda > 0, \alpha > 0, \beta > 0$$
Use $x_0 = 1$.
$$I(x) = \int_1^x \frac{2\lambda(\alpha - y)}{\beta y}\, dy = \frac{2\lambda\alpha}{\beta} \log x - \frac{2\lambda}{\beta}(x-1).$$

$$s'(x) = x^{-2\lambda\alpha/\beta} e^{2\lambda/\beta (x-1)}.$$
Clearly, $\int_1^\infty s'(x)\,dx = \infty$. Also, $\int_0^1 s'(x)\,dx = \infty \Leftrightarrow 2\lambda\alpha \geq \beta$.

So $X(t)$ is recurrent iff $2\lambda\alpha \geq \beta$.
Assume now that $2\lambda\alpha \geq \beta$.
$$m'(x) = \frac{2}{\beta} x^{2\lambda\alpha/\beta - 1} e^{-2\lambda/\beta (x-1)},$$ which is integrable on $(0, \infty)$. So $X(t)$ is positive recurrent. The stationary density (proportional to $m'(x)$) is gamma$\left(\frac{2\lambda\alpha}{\beta}, \frac{2\lambda}{\beta}\right)$.

this can be relaxed in some cases

<u>Thm. 6.31</u> Suppose $X(t)$ is a diffusion with (nice) drift $\mu(x)$ and diffusion function $\sigma^2(x)$. Let $g(x)$ be a 1-1 function that is twice continuously differentiable and define $Y(t) = g(X(t))$. Then $Y(t)$ is a diffusion with

$$\text{drift } \tilde{\mu}(y) = \left(\mu(x) g'(x) + \frac{\sigma^2(x)}{2} g''(x)\right)\Big|_{x=g^{-1}(y)}$$

$$\text{and diffusion } \tilde{\sigma}^2(y) = \sigma^2(x)(g'(x))^2\Big|_{x=g^{-1}(y)}.$$

Proof. A 1-1 function of a Markov process is always Markov. Formally, let $t_1 < \cdots < t_k$. Then (assume $g(x)\uparrow$ here)

$$\begin{aligned} &P(Y(t_k) \le y_k \mid Y(t_1) = y_1, \ldots, Y(t_{k-1}) = y_{k-1}) \\ &= P(X(t_k) \le g^{-1}(y_k) \mid X(t_1) = g^{-1}(y_1), \ldots, X(t_{k-1}) = g^{-1}(y_{k-1})) \\ &= P(X(t_k) \le g^{-1}(y_k) \mid X(t_{k-1}) = g^{-1}(y_{k-1})) \\ &= P(Y(t_k) \le y_k \mid Y(t_{k-1}) = y_{k-1}). \end{aligned}$$

Since g is continuous and $X(t)$ has cont. sample paths, $Y(t)$ also has cont. sample paths. Therefore $Y(t)$ is a diffusion.

We already know (Thm 6.23) that

$$\frac{1}{t} E(g(X(t)) - g(X(0)) \mid X(0) = x) \xrightarrow{t \to 0} \mu(x) g'(x) + \frac{\sigma^2(x)}{2} g''(x).$$

Substituting $x = g^{-1}(y)$ therefore gives the drift function for $Y(t) = g(X(t))$.

Likewise,

$$\frac{1}{t} E(g^2(X(t)) - g^2(X(0)) \mid X(0) = x)$$
$$\xrightarrow{t \to 0} 2\mu(x) g'(x) g(x) + \sigma^2(x)\left((g'(x))^2 + g''(x) g(x)\right),$$

by applying Thm. 6.23 to the function $g^2(x)$.

Putting the last two points together,

$$\frac{1}{t}E\left((g(X(t)) - g(X(0)))^2 \mid X(0)=x \right)$$
$$= \frac{1}{t}E\left(g^2(X(t)) - g^2(X(0)) \mid X(0)=x \right)$$
$$- \frac{1}{t} 2g(x)\, E\left(g(X(t)) - g(X(0)) \mid X_0 = x \right)$$
$$\to \sigma^2(x)(g'(x))^2 \qquad \text{(check the algebra)}.$$

Again substituting $x = g^{-1}(y)$, we have deduced the diffusion function for $Y(t) = g(X(t))$. □

Ex. 6.6 (cont.) Geometric BM.

Here $Y(t) = e^{B_\mu(t)}$ where $B_\mu(t)$ is a B.M. w/ drift. Recall the derivations on page 6.21.

Ex. 6.7 (cont.) Consider a special case of Feller's square root process with $\mu(x) = 1-x$ and $\sigma^2(x) = x$. This is a positive process. Let

$$Y(t) = X^2(t).$$

Then $Y(t)$ is a diffusion with positive values,

drift $\tilde{\mu}(y) = \left((1-x)2x + \frac{x}{2}2\right)\Big|_{x=\sqrt{y}} = 2(1-\sqrt{y})\sqrt{y} + \sqrt{y}$
$= 3\sqrt{y} - 2y,$

and diffusion

$$\tilde{\sigma}^2(y) = x(2x)^2\Big|_{x=\sqrt{y}} = 4y^{3/2}.$$

What is the process $W(t) = \log(X(t))$ like? (exercise)

Ex. 6.8 $Y(t) = B^2(t)$ where $B(t)$ is std. B.M. Here, $g(x) = x^2$ is not 1-1. However, $Y(t)$ is Markov (Thm. 6.13). Check that the $\tilde{\mu}(y)$ & $\tilde{\sigma}^2(y)$ in Thm. 6.31 are meaningful.