Applied Stochastic Analysis

Zixun Huang, 2200010794

January 12, 2026

Contents

1 Homework 1 (Sep 8th)

Solution 1.1. *We want to compute*

$$C = AB = \sum_{i=1}^{n} A_{:,i} B_{i,:},$$

that is, a sum of n rank-one outer products. The randomized estimator is constructed by sampling indices $i_m \in \{1, \ldots, n\}$, $m = 1, \ldots, K$, with probabilities $\{p_i\}$, and defining

$$L^{(m)} = \frac{1}{\sqrt{K p_{i_m}}} A_{:,i_m}, \quad R^{(m)} = \frac{1}{\sqrt{K p_{i_m}}} B_{i_m,:}.$$

The approximation is

$$\widehat{C} = \sum_{m=1}^{K} L^{(m)} R^{(m)}.$$

Taking expectation,

$$\mathbb{E}[L^{(m)} R^{(m)}] = \sum_{i=1}^{n} p_i \cdot \frac{1}{K p_i} A_{:,i} B_{i,:} = \frac{1}{K} \sum_{i=1}^{n} A_{:,i} B_{i,:}.$$

Summing over $m = 1, \ldots, K$ gives

$$\mathbb{E}[\widehat{C}] = C,$$

so the estimator is unbiased.

The accuracy can be described in terms of the variance

$$\begin{aligned}
\mathbb{E}\left[\|\widehat{C} - C\|_F^2\right] &= \mathbb{E}\left[\left\|\sum_{m=1}^{K} L^{(m)} R^{(m)} - C\right\|_F^2\right] \\
&= \frac{1}{K^2} \sum_{m=1}^{K} \mathbb{E}\left[\left\|L^{(m)} R^{(m)} - \frac{1}{K} C\right\|_F^2\right] \\
&= \frac{1}{K} \mathbb{E}\left[\left\|L^{(1)} R^{(1)} - \frac{1}{K} C\right\|_F^2\right]
\end{aligned}$$

This shows that the estimator is unbiased and its variance decays as $1/K$, so the accuracy improves with more samples, and the bound highlights that the choice of sampling distribution $\{p_i\}$ is crucial: selecting $p_i \propto \|A_{:,i}\| \|B_{i,:}\|$ minimizes the variance up to constants.

Solution 1.2. *Let $[0,1]$ be partitioned uniformly with step $h = 1/n$ and midpoints $m_i = (i - \frac{1}{2})h$. On each cell $I_i = [(i-1)h, ih]$, Taylor expand f about m_i:*

$$f(x) = f(m_i) + f'(m_i)(x - m_i) + \tfrac{1}{2} f''(\xi_{i,x})(x - m_i)^2, \qquad x \in I_i.$$

Integrate over I_i; the odd term vanishes:

$$\int_{I_i} f(x)\, dx = h\, f(m_i) + \tfrac{1}{2} \int_{-h/2}^{h/2} f''(\xi_{i,m_i+t})\, t^2\, dt$$

for some $\xi_i \in I_i$. *Hence the midpoint rule*

$$Q_h := h \sum_{i=1}^{n} f(m_i)$$

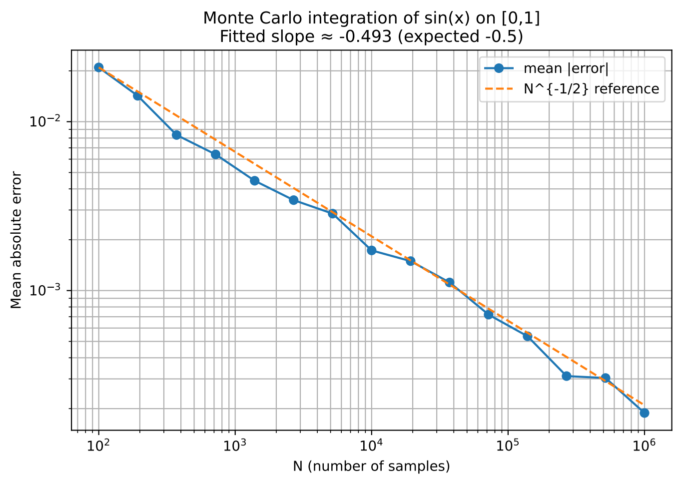
satisfies

$$\int_0^1 f(x)\,dx - Q_h = \sum_{i=1}^{n} \tfrac{1}{2} \int_{-h/2}^{h/2} f''(\xi_{i,m_i+t})\, t^2\, dt.$$

If $f \in C^2[0,1]$ *and* $\|f''\|_\infty \le M$, *then*

$$\left| \int_0^1 f - Q_h \right| \le \frac{n}{2} M \frac{h^3}{12} = \frac{M}{24} h^2.$$

so the midpoint rule has global error $O(h^2)$ *(second-order convergence).*



Solution 1.3.

2 Homework 2 (Sep 10th)

Solution 2.1. *Let* (A_n) *be independent with* $\sum_n \mathbb{P}(A_n) = \infty$. *For* $m \ge 1$, *set* $B_m = \bigcup_{n \ge m} A_n$. *Then* $\{A_n \text{ i.o.}\} = \bigcap_{m \ge 1} B_m$, *so* $\mathbb{P}(A_n \text{ i.o.}) = \lim_{m\to\infty} \mathbb{P}(B_m) = 1 - \lim_{m\to\infty} \mathbb{P}(B_m^c)$. *Now* $B_m^c = \bigcap_{n \ge m} A_n^c$ *and by independence*

$$\mathbb{P}\left(\bigcap_{n=m}^{M} A_n^c \right) = \prod_{n=m}^{M} \big(1 - \mathbb{P}(A_n)\big) \le \exp\left(- \sum_{n=m}^{M} \mathbb{P}(A_n) \right) \xrightarrow[M\to\infty]{} 0,$$

since the series diverges. Hence $\mathbb{P}(B_m^c) = 0$ *for each* m, *so the limit is* 0 *and* $\mathbb{P}(A_n \text{ i.o.}) = 1$.

Solution 2.2. *If* $X \sim \mathrm{Pois}(\lambda)$, $Y \sim \mathrm{Pois}(\mu)$ *are independent, then for* $k \in \mathbb{Z}_{\ge 0}$,

$$\mathbb{P}(X+Y=k) = \sum_{i=0}^{k} \mathbb{P}(X=i)\mathbb{P}(Y=k-i) = e^{-(\lambda+\mu)} \sum_{i=0}^{k} \frac{\lambda^i \mu^{k-i}}{i!(k-i)!} = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!},$$

which is the pmf of $\mathrm{Pois}(\lambda+\mu)$.

Solution 2.3. *With $X \sim \text{Pois}(\lambda)$, $Y \sim \text{Pois}(\mu)$ independent and $N = X + Y$ fixed, for $x = 0, 1, \ldots, N$,*

$$\mathbb{P}(X = x \mid X{+}Y = N) = \frac{\mathbb{P}(X = x, Y = N - x)}{\mathbb{P}(X + Y = N)} = \frac{\left(e^{-\lambda} \frac{\lambda^x}{x!}\right)\left(e^{-\mu} \frac{\mu^{N-x}}{(N-x)!}\right)}{e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^N}{N!}} = \binom{N}{x}\left(\frac{\lambda}{\lambda+\mu}\right)^x \left(\frac{\mu}{\lambda+\mu}\right)^{N-x}.$$

Hence $X \mid (X + Y = N) \sim \text{Bin}(N, \lambda/(\lambda+\mu))$ (and symmetrically for Y).

Solution 2.4. *(1) Let $X \sim \text{Exp}(\lambda)$ with tail $\bar{F}(x) = \mathbb{P}(X > x) = e^{-\lambda x}$ for $x \geq 0$. For $s, t > 0$,*

$$\mathbb{P}(X > s + t \mid X > s) = \frac{\bar{F}(s+t)}{\bar{F}(s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X > t).$$

This is the memoryless property.

(2) Assume $\mathbb{P}(X > s + t) = \mathbb{P}(X > s)\mathbb{P}(X > t)$ for all $s, t > 0$. Let $\phi(t) = \mathbb{P}(X > t)$ for $t \geq 0$. Then $\phi(0) = 1$, ϕ is nonincreasing and right–continuous, and

$$\phi(s + t) = \phi(s)\phi(t) \qquad (s, t \geq 0).$$

Put $g(t) = -\log \phi(t)$ (well-defined since $\phi(t) \in (0, 1]$). Then $g(0) = 0$, g is measurable and $g(s{+}t) = g(s){+}g(t)$, so by Cauchy's functional equation in the measurable/monotone case, $g(t) = \lambda t$ for some $\lambda \geq 0$. Because $\phi(t) \to 0$ as $t \to \infty$ for a proper r.v., we must have $\lambda > 0$. Hence $\phi(t) = e^{-\lambda t}$ and therefore $X \sim \text{Exp}(\lambda)$.

Solution 2.5. *Let $(X_1, \ldots, X_n)$ be centered jointly Gaussian with covariance matrix $\Sigma = (\Sigma_{ij})$. The mgf of $T = \sum_{i=1}^n t_i X_i$ is*

$$M_T(u) = \mathbb{E}e^{uT} = \exp\left(\tfrac{1}{2} u^2\, t^\top \Sigma t\right).$$

By multilinearity,

$$\mathbb{E}(X_{i_1} \cdots X_{i_k}) = \frac{\partial^k}{\partial t_{i_1} \cdots \partial t_{i_k}} \exp\left(\tfrac{1}{2} t^\top \Sigma t\right)\Big|_{t=0}.$$

If k is odd, all derivatives vanish at $t = 0$ and the moment is 0. For even $k = 2m$, differentiating the quadratic exponent generates a sum over all pairings π of $\{1, \ldots, 2m\}$; each pairing contributes $\prod_{(a,b)\in\pi} \Sigma_{ab}$. Thus

$$\mathbb{E}(X_1 \cdots X_k) = \begin{cases} \displaystyle\sum_{\textit{pairings } \pi} \prod_{(a,b)\in\pi} \mathbb{E}(X_a X_b), & k \textit{ even}, \\ 0, & k \textit{ odd}. \end{cases}$$

This is Wick's (Isserlis') theorem; e.g. for (X, Y, Z) one obtains the expansion illustrated in the prompt by listing all pairings with their multiplicities.

Solution 2.6. *Suppose (A_n) are mutually independent, $\mathbb{P}(\cup_{n\geq 1} A_n) = 1$, and $\mathbb{P}(A_n) < 1$ for each n. Then by independence,*

$$\mathbb{P}\left(\bigcap_{n=1}^N A_n^c\right) = \prod_{n=1}^N \left(1 - \mathbb{P}(A_n)\right) \xrightarrow[N\to\infty]{} 0$$

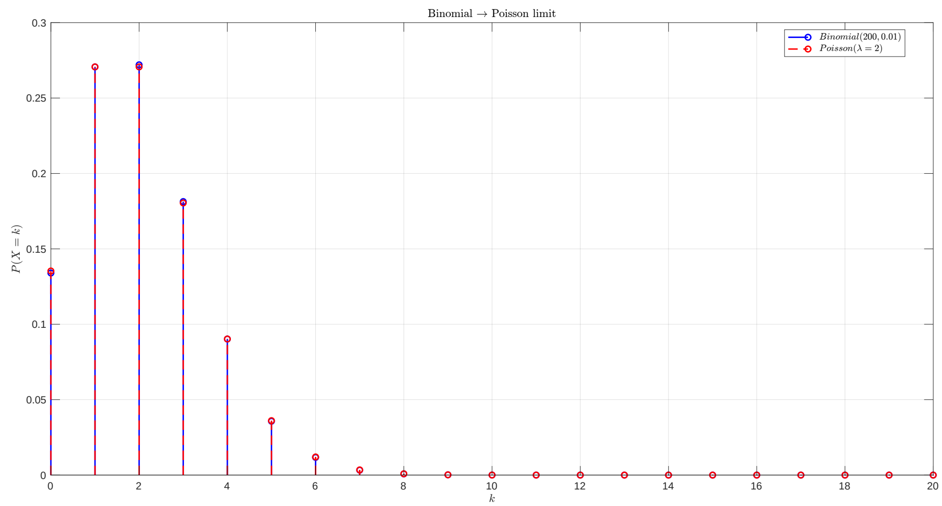
because $\mathbb{P}(\cap_{n\geq 1} A_n^c) = 1 - \mathbb{P}(\cup_{n\geq 1} A_n) = 0$. Fix m. For $M \geq m$,

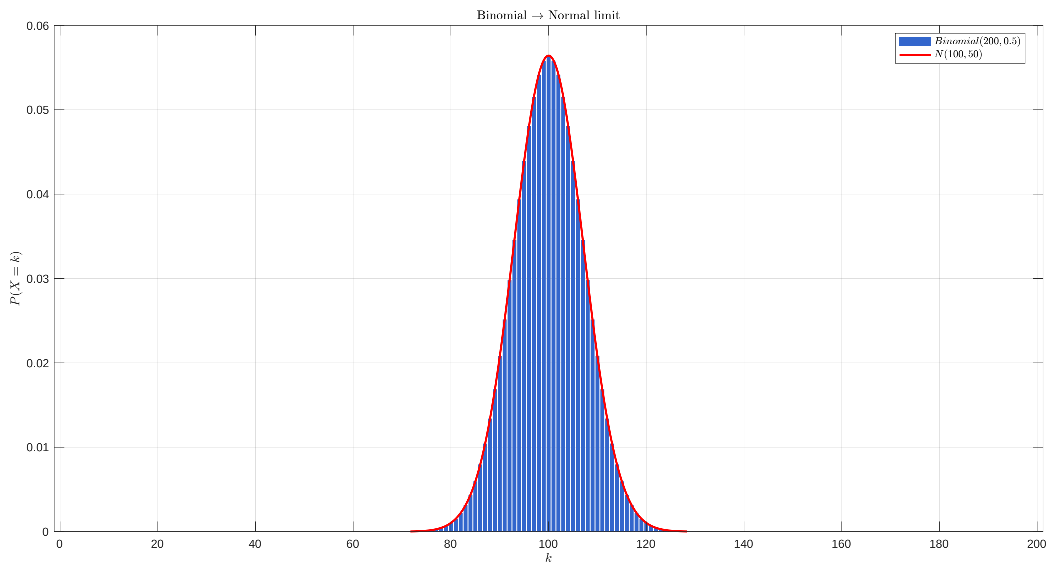
$$\mathbb{P}\left(\bigcap_{n=m}^M A_n^c\right) = \frac{\prod_{n=1}^M (1 - \mathbb{P}(A_n))}{\prod_{n=1}^{m-1} (1 - \mathbb{P}(A_n))} \xrightarrow[M\to\infty]{} 0,$$

since the denominator is strictly positive by $\mathbb{P}(A_n) < 1$. *Hence* $\mathbb{P}(\cup_{n \geq m} A_n) = 1$ *for every* m, *and therefore*

$$\mathbb{P}(A_n \text{ } i.o.) = \mathbb{P}\left(\bigcap_{m \geq 1} \bigcup_{n \geq m} A_n \right) = 1.$$

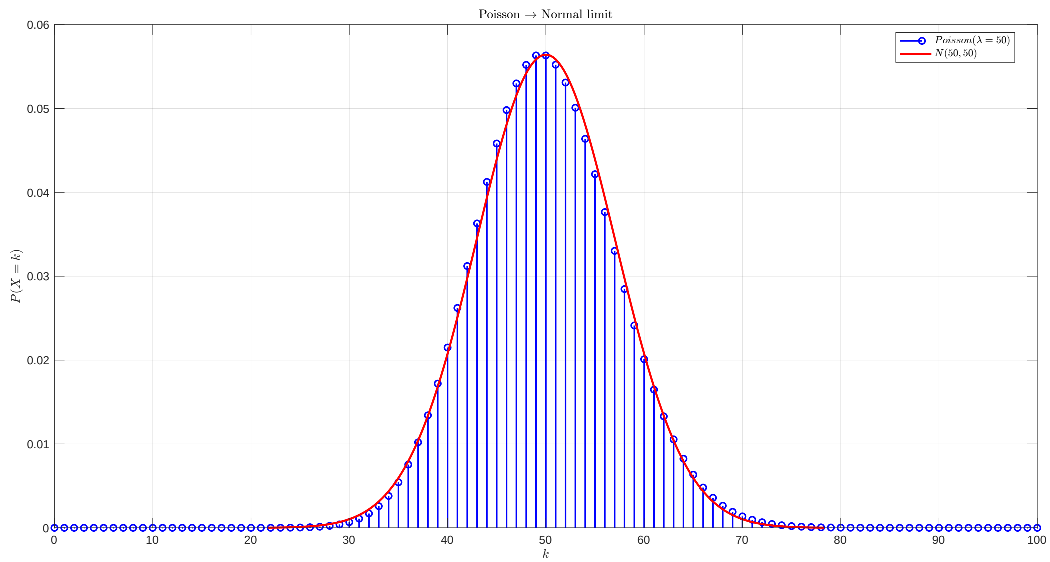
Solution 2.7. *We choose* $n = 200$ *as sample size,* $p_1 = 0.01$ *for the approximation of Poisson distribution,* $p_2 = 0.5$ *for the approximation of Normal distribution.*





3 Homework 3 (Sep 15th)

Solution 3.1. *A complete implementation is provided in the accompanying MATLAB code.*



Solution 3.2. *There are several intuitive methods to generate a random point uniformly on the sphere surface S^2. We describe two classical approaches:*

Approach 1. Normalized Gaussian vectors.

1. *Generate $Z_1, Z_2, Z_3 \sim \mathcal{N}(0, 1)$ independently.*
2. *Form the vector $Z = (Z_1, Z_2, Z_3)$.*
3. *Normalize: $X = Z/\|Z\|$.*

Because the multivariate Gaussian distribution is rotationally invariant, the direction $Z/\|Z\|$ is uniformly distributed on S^2.

Approach 2. Direct spherical coordinates.

1. *Generate $\phi \sim \mathrm{Unif}[0, 2\pi]$.*
2. *Generate $u \sim \mathrm{Unif}[-1, 1]$ and set $\cos\theta = u$.*
3. *Convert to Cartesian coordinates:*

$$x = \sqrt{1-u^2}\cos\phi, \qquad y = \sqrt{1-u^2}\sin\phi, \qquad z = u.$$

This uses the fact that the surface element on S^2 is proportional to $\sin\theta\, d\theta\, d\phi$, so $\cos\theta$ is uniform on $[-1, 1]$.

Comparison. *The Gaussian normalization method is very general and works in higher dimensions. The spherical coordinate method is more geometric and provides intuition about the distribution of latitude and longitude.*

Solution 3.3. *In Algorithm 2.6 we draw X from the density $g(x) = f(x)/A$ by the inverse transform $X = F^{-1}(AZ)$ (here $A = \int f(x)\, dx$ and F is the primitive of f), then draw $Y \mid X \sim \mathrm{Unif}[0, f(X)]$ and accept iff $Y < p(X)$.*

Conditioning on $X = x$, the acceptance probability is

$$\mathbb{P}(\mathit{accept} \mid X = x) = \frac{p(x)}{f(x)} \qquad (\mathit{since}\ Y \sim \mathrm{Unif}[0, f(x)]).$$

Hence the unconditional acceptance probability is

$$\mathbb{P}(accept) = \mathbb{E}\left[\frac{p(X)}{f(X)}\right] = \int \frac{p(x)}{f(x)} \frac{f(x)}{A}\, dx = \frac{1}{A}\int p(x)\, dx.$$

If p is a normalized pdf, $\int p(x)\, dx = 1$ and thus $\mathbb{P}(accept) = 1/A$. Therefore the rejection probability *of one trial is*

$$\mathbb{P}(reject) = 1 - \frac{\int p(x)\, dx}{A} = 1 - \frac{1}{A}.$$

Solution 3.4. ***Correctness.*** *Because $g_\ell(X) \le p(X)$, the rule above is equivalent to*

$$accept \iff U \le \frac{p(X)}{M g_m(X)}.$$

Therefore

$$\mathbb{P}(X \in dx,\ accept) = g_m(x)\, dx \cdot \frac{p(x)}{M g_m(x)} = \frac{p(x)}{M}\, dx.$$

Let $Z_p := \int p(x)\, dx$. Then $\mathbb{P}(accept) = \int \frac{p(x)}{M}\, dx = \frac{Z_p}{M}$ and, conditioning on acceptance,

$$\mathbb{P}(X \in dx \mid accept) = \frac{p(x)dx/M}{Z_p/M} = \frac{p(x)}{Z_p}\, dx,$$

i.e., the accepted X has density proportional to p, hence exactly p when p is normalized.

Advantage over Algorithm 2.6. *The overall acceptance probability remains Z_p/M (the same as the standard envelope Mg_m). However, Step 2 provides a* squeeze*: whenever $U \le g_\ell(X)/(Mg_m(X))$ we accept* without *evaluating the potentially expensive $p(x)$. Only in the remaining cases do we compute $p(x)$ to decide. Thus the algorithm:*

- *reduces the expected number of evaluations of $p(\cdot)$ per accepted sample;*
- *yields faster simulation when g_ℓ is easy to compute and is a good lower bound of p.*

4 Homework 4 (Sep 22th)

Solution 4.1. *Let $X \sim U[0,1]$ and let $f : [0,1] \to \mathbb{R}$ be monotone (hence measurable and integrable). Note that $1 - X \sim U[0,1]$ and $\mathbb{E}[f(1-X)] = \mathbb{E}[f(X)]$. Therefore*

$$\mathrm{Cov}(f(X), f(1-X)) = \mathbb{E}[f(X)f(1-X)] - \big(\mathbb{E}[f(X)]\big)^2 = \int_0^1 f(x)f(1-x)\, dx - \Big(\int_0^1 f(x)\, dx\Big)^2.$$

Consider the double integral

$$I \;=\; \int_0^1\int_0^1 \big(f(x) - f(u)\big)\big(f(1-x) - f(1-u)\big)\, du\, dx.$$

Expanding and integrating term-by-term gives

$$I = 2\int_0^1 f(x)f(1-x)\, dx - 2\Big(\int_0^1 f(x)\, dx\Big)^2,$$

hence

$$\int_0^1 f(x)f(1-x)\, dx - \Big(\int_0^1 f(x)\, dx\Big)^2 = \frac{I}{2}.$$

If f is nondecreasing, then for any $x, u \in [0,1]$,

$$x \geq u \ \Rightarrow\ f(x) - f(u) \geq 0, \qquad 1 - x \leq 1 - u \ \Rightarrow\ f(1-x) - f(1-u) \leq 0,$$

so the product $\big(f(x)-f(u)\big)\big(f(1-x)-f(1-u)\big) \leq 0$. The same conclusion holds if f is nonincreasing (the two factors switch signs). Therefore $I \leq 0$ and thus

$$\mathrm{Cov}(f(X), f(1-X)) = \frac{I}{2} \leq 0.$$

Solution 4.2. *Let $Y := f(\mathbf{X})$ and let $\mathscr{G} := \sigma(\mathbf{X}^{(2)})$. Then $\mathbb{E}[Y \mid \mathscr{G}]$ is $\mathscr{G}$–measurable and $\mathbb{E}\big[Y - \mathbb{E}(Y \mid \mathscr{G}) \mid \mathscr{G}\big] = 0$. Write the orthogonal decomposition*

$$Y - \mathbb{E}Y = \big(Y - \mathbb{E}(Y \mid \mathscr{G})\big) + \big(\mathbb{E}(Y \mid \mathscr{G}) - \mathbb{E}Y\big).$$

Squaring and taking expectations gives

$$\mathrm{Var}(Y) = \mathbb{E}\big[(Y - \mathbb{E}(Y \mid \mathscr{G}))^2\big] + \mathbb{E}\big[(\mathbb{E}(Y \mid \mathscr{G}) - \mathbb{E}Y)^2\big] + 2\,\mathbb{E}[(Y - \mathbb{E}(Y \mid \mathscr{G}))(\mathbb{E}(Y \mid \mathscr{G}) - \mathbb{E}Y)]\,.$$

The cross term is zero by conditional expectation:

$$\mathbb{E}[(Y - \mathbb{E}(Y \mid \mathscr{G}))(\mathbb{E}(Y \mid \mathscr{G}) - \mathbb{E}Y)] = \mathbb{E}[\mathbb{E}[Y - \mathbb{E}(Y \mid \mathscr{G}) \mid \mathscr{G}]\,(\mathbb{E}(Y \mid \mathscr{G}) - \mathbb{E}Y)] = 0.$$

Hence

$$\mathrm{Var}(Y) = \mathbb{E}\big[(Y - \mathbb{E}(Y \mid \mathscr{G}))^2\big] + \mathbb{E}\big[(\mathbb{E}(Y \mid \mathscr{G}) - \mathbb{E}Y)^2\big] = \mathbb{E}[\mathrm{Var}(Y \mid \mathscr{G})] + \mathrm{Var}(\mathbb{E}(Y \mid \mathscr{G}))\,.$$

Restoring $Y = f(\mathbf{X})$ and $\mathscr{G} = \sigma(\mathbf{X}^{(2)})$ yields

$$\mathrm{Var}(f(\mathbf{X})) = \mathrm{Var}\big(\mathbb{E}[f(\mathbf{X}) \mid \mathbf{X}^{(2)}]\big) + \mathbb{E}\big[\,\mathrm{Var}(f(\mathbf{X}) \mid \mathbf{X}^{(2)})\big],$$

which is the desired identity.

Solution 4.3. *Let f, g be probability densities on a common space $(\mathscr{X}, \mathscr{A}, \mu)$ with $g > 0$ a.e. and set*

$$D(f\|g) \;=\; \int_{\mathscr{X}} f(x)\, \log \frac{f(x)}{g(x)}\, d\mu(x).$$

Write $\phi(t) = t \log t$, which is strictly convex on $(0, \infty)$. Then

$$D(f\|g) = \int_{\mathscr{X}} g(x)\, \phi\!\left(\frac{f(x)}{g(x)}\right)\, d\mu(x).$$

By Jensen's inequality applied to the probability measure $g\, d\mu$ and the convex function ϕ,

$$\int g\, \phi\!\left(\frac{f}{g}\right)\, d\mu \;\geq\; \phi\!\left(\int g \cdot \frac{f}{g}\, d\mu\right) = \phi\!\left(\int f\, d\mu\right) = \phi(1) = 0,$$

whence $D(f\|g) \geq 0$.

Moreover, since ϕ is strictly convex, equality in Jensen holds iff $\frac{f(x)}{g(x)}$ is constant g-a.e. Because $\int f\, d\mu = 1 = \int g\, d\mu$, this constant must be 1, i.e., $f(x) = g(x)$ g-a.e. (and hence μ-a.e.). Conversely, if $f = g$ a.e., then integrand is 0 a.e. and $D(f\|g) = 0$.

With the convention $0 \log 0 = 0$ (and taking $f \ll g$ so the ratio is well-defined g-a.e.), the result follows.

5 Homework 5 (Sep 24th)

Solution 5.1. *Let* $(X_j)_{j\geq 1}$ *be i.i.d. with* $X_j \sim U[0,1]$.

(i) Harmonic mean. *We claim*

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{X_j} \xrightarrow[n\to\infty]{a.s.} \infty \qquad \Longrightarrow \qquad \frac{n}{X_1^{-1}+\cdots+X_n^{-1}} \xrightarrow{a.s.} 0.$$

Define the truncations $Y_j^{(k)} := \min\{X_j^{-1}, k\}$ *for* $k \in \mathbb{N}$*. For each fixed* k, $Y_j^{(k)}$ *are i.i.d. with*

$$\mathbb{E}\,Y_1^{(k)} = \int_0^1 \min\left(\frac{1}{x}, k\right)\,dx = \int_0^{1/k} k\,dx + \int_{1/k}^1 \frac{1}{x}\,dx = 1 + \log k \uparrow \infty \quad (k\to\infty).$$

By the strong law of large numbers (SLLN),

$$\frac{1}{n}\sum_{j=1}^{n} Y_j^{(k)} \xrightarrow{a.s.} \mathbb{E}\,Y_1^{(k)} = 1+\log k.$$

Fix $M>0$*. Choose* k *so large that* $1+\log k > M$*. Then, almost surely, for all sufficiently large* n*,*

$$\frac{1}{n}\sum_{j=1}^{n}\frac{1}{X_j} \;\geq\; \frac{1}{n}\sum_{j=1}^{n} Y_j^{(k)} \;>\; M/2.$$

Since M *was arbitrary,* $\frac{1}{n}\sum_{j=1}^{n} X_j^{-1} \to \infty$ *a.s., hence*

$$\frac{n}{X_1^{-1}+\cdots+X_n^{-1}} \xrightarrow{a.s.} 0.$$

(ii) Geometric mean. *Because* $\mathbb{E}|\log X_1| = \int_0^1 |\log x|\,dx = 1 < \infty$*, the SLLN gives*

$$\frac{1}{n}\sum_{j=1}^{n}\log X_j \xrightarrow{a.s.} \mathbb{E}[\log X_1] = \int_0^1 \log x\,dx = -1.$$

By continuity of the exponential,

$$\sqrt[n]{X_1X_2\cdots X_n} = \exp\left(\frac{1}{n}\sum_{j=1}^{n}\log X_j\right) \xrightarrow{a.s.} e^{-1}.$$

(iii) Quadratic mean. *Since* $\mathbb{E}[X_1^2] = \int_0^1 x^2\,dx = \frac{1}{3}$*, the SLLN yields*

$$\frac{1}{n}\sum_{j=1}^{n} X_j^2 \xrightarrow{a.s.} \frac{1}{3}.$$

Applying the continuous mapping theorem with the square root,

$$\sqrt{\frac{X_1^2+\cdots+X_n^2}{n}} \xrightarrow{a.s.} \sqrt{\frac{1}{3}} = \frac{1}{\sqrt{3}}.$$

Combining the three parts,

$$\boxed{\lim_{n\to\infty}\frac{n}{X_1^{-1}+\cdots+X_n^{-1}} = 0, \quad \lim_{n\to\infty}\sqrt[n]{X_1\cdots X_n} = e^{-1}, \quad \lim_{n\to\infty}\sqrt{\frac{X_1^2+\cdots+X_n^2}{n}} = \frac{1}{\sqrt{3}}}$$

a.s.

Solution 5.2. *Let $X_1, X_2, \dots$ be i.i.d. random variables with $\mathbb{E}[X_i] = 0$. Assume that*

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}} \stackrel{d}{\Rightarrow} X, \qquad Z_{2n} = \frac{X_1 + \cdots + X_{2n}}{\sqrt{2n}} \stackrel{d}{\Rightarrow} X$$

and denote the characteristic function of X by $f(\xi) = \mathbb{E}[e^{i\xi X}]$.

***(a) Show that** $f(\xi) = f^2(\xi/\sqrt{2})$.*

Let

$$S_n^{(1)} = \sum_{j=1}^{n} X_j, \qquad S_n^{(2)} = \sum_{j=n+1}^{2n} X_j.$$

Then $S_n^{(1)}$ and $S_n^{(2)}$ are independent and have the same distribution as S_n. We can write

$$Z_{2n} = \frac{S_n^{(1)} + S_n^{(2)}}{\sqrt{2n}} = \frac{1}{\sqrt{2}}\Big(\frac{S_n^{(1)}}{\sqrt{n}} + \frac{S_n^{(2)}}{\sqrt{n}}\Big) = \frac{1}{\sqrt{2}}\big(Z_n^{(1)} + Z_n^{(2)}\big),$$

where $Z_n^{(1)}, Z_n^{(2)}$ are i.i.d. copies of Z_n.

Let φ_n be the characteristic function of Z_n. Then

$$\varphi_{2n}(\xi) = \big(\varphi_n(\xi/\sqrt{2})\big)^2.$$

By Lévy's continuity theorem, $\varphi_n(\xi) \to f(\xi)$ and $\varphi_{2n}(\xi) \to f(\xi)$, hence

$$f(\xi) = f^2(\xi/\sqrt{2}).$$

(b) If** $f \in C^2(\mathbb{R})$**, then** f **is Gaussian.

Let $\phi(\xi) = \log f(\xi)$ (the continuous branch near 0, noting $f(0) = 1$). From (a),

$$\phi(\xi) = 2\phi(\xi/\sqrt{2}).$$

Define $h(\xi) = \phi(\xi)/\xi^2$ for $\xi \neq 0$. Then

$$h(\xi) = h(\xi/\sqrt{2}) \quad \Rightarrow \quad h(\xi) = \lim_{k\to\infty} h(\xi/2^{k/2}) = h(0).$$

Since $f \in C^2$, $\phi''(0) = f''(0) - [f'(0)]^2$ exists, and $\phi(0) = \phi'(0) = 0$, so $h(0) = \frac{1}{2}\phi''(0) = -\frac{1}{2}\sigma^2$ for some $\sigma^2 \geq 0$. Hence

$$\phi(\xi) = -\tfrac{1}{2}\sigma^2\xi^2, \qquad f(\xi) = e^{-\frac{1}{2}\sigma^2\xi^2},$$

which is the characteristic function of $N(0, \sigma^2)$.

***(c) Replace** $1/\sqrt{n}$ **by** $1/n$.*

Now

$$\widetilde{Z}_n = \frac{X_1 + \cdots + X_n}{n} \stackrel{d}{\Rightarrow} X, \qquad \widetilde{Z}_{2n} \stackrel{d}{\Rightarrow} X.$$

Analogously,

$$f(\xi) = f^2(\xi/2).$$

Let $\phi = \log f$. Then $\phi(\xi) = 2\phi(\xi/2)$. If f is even (symmetric) or constant, define $h(\xi) = \phi(\xi)/|\xi|$. Then $h(\xi) = h(\xi/2) \Rightarrow h(\xi) = h(0) = -\gamma$ with $\gamma \geq 0$. Thus

$$f(\xi) = e^{-\gamma|\xi|},$$

which is the characteristic function of a (centered) Cauchy–Lorentz distribution (or degenerate at 0 if $\gamma = 0$).

(d) Replace $1/\sqrt{n}$ ***by*** $1/n^\alpha$***.***

The same reasoning gives

$$f(\xi) = f^2(\xi/2^\alpha), \qquad \phi(\xi) = 2\phi(\xi/2^\alpha).$$

Assume $\phi(\xi) = -c|\xi|^p$ *for some* $c > 0$*. Then*

$$-c|\xi|^p = 2\big(-c|\xi|^p/2^{\alpha p}\big) \quad \Rightarrow \quad 2^{1-\alpha p} = 1 \quad \Rightarrow \quad p = \frac{1}{\alpha}.$$

Hence

$$f(\xi) = \exp(-c|\xi|^{1/\alpha}),$$

the characteristic function of a symmetric α*-stable law.*

For $\exp(-c|\xi|^p)$ *to be a characteristic function,* $0 < p \le 2$*. Therefore*

$$\boxed{\alpha \ge \tfrac{1}{2},}$$

where $\alpha = \frac{1}{2}$ *gives the Gaussian,* $\alpha = 1$ *gives the Cauchy case, and larger* α *yield heavier-tailed stable distributions.*

Remark 5.1. *Solve function equations.*

Solution 5.3. *This example shows that the law of large numbers fails when its assumptions are not satisfied. Let* $\{X_j\}_{j=1}^\infty$ *be i.i.d. random variables following the Cauchy distribution with probability density function*

$$f(x) = \frac{1}{\pi(1+x^2)}, \qquad x \in \mathbb{R}.$$

We check three key properties.

(1) Expectation. *Although the Cauchy distribution is symmetric about* 0*, the mean* $\mathbb{E}[X_j]$ *does* not exist*. Indeed, the integral*

$$\int_{\mathbb{R}} |x| f(x)\,dx = \frac{2}{\pi}\int_0^\infty \frac{x}{1+x^2}\,dx = \frac{1}{\pi}\Big[\log(1+x^2)\Big]_0^\infty = \infty.$$

Hence $\mathbb{E}|X_j| = \infty$*, and* $\mathbb{E}[X_j]$ *is undefined (it is not absolutely integrable).*

(2) Divergent moments. *Similarly,*

$$\mathbb{E}[X_j^2] = \int_{\mathbb{R}} \frac{x^2}{\pi(1+x^2)}\,dx = \infty,$$

so the variance does not exist either.

(3) Distribution of the sample mean. *Let* $S_n = X_1 + \cdots + X_n$*. The characteristic function of* X_1 *is*

$$\phi(t) = \mathbb{E}[e^{itX_1}] = e^{-|t|}, \qquad t \in \mathbb{R}.$$

By independence,

$$\phi_{S_n/n}(t) = \mathbb{E}[e^{itS_n/n}] = \prod_{j=1}^n \mathbb{E}[e^{itX_j/n}] = \big(\phi(t/n)\big)^n = \big(e^{-|t|/n}\big)^n = e^{-|t|}.$$

Thus, $\phi_{S_n/n}(t) = \phi(t)$*, meaning*

$$\frac{S_n}{n} \stackrel{d}{=} X_1 \quad \text{for all } n.$$

Hence the distribution of the sample mean is the same as that of each X_j*, and the sequence* $\{S_n/n\}$ *does* not *converge to a constant. Both the weak and strong laws of large numbers therefore fail.*

Solution 5.4. *Let* $\psi(x) := h(0) - h(x) \geq 0$*. Then* $\psi(0) = 0$*,* $\psi(x) > 0$ *for* $x > 0$*,* $\psi'(x) = -h'(x)$ *on* $(0, \infty)$*, and* $\psi'(0) = -h'(0) =: a > 0$*. Write*

$$I(t) := \int_0^\infty e^{th(x)}\, dx = e^{th(0)} \int_0^\infty e^{-t\psi(x)}\, dx.$$

Fix $\varepsilon \in (0, a/2)$*. By continuity of* ψ' *at* 0*, there exists* $\delta > 0$ *such that*

$$a - \varepsilon \leq \psi'(x) \leq a + \varepsilon \qquad (0 \leq x \leq \delta). \tag{1}$$

Split the integral:

$$\int_0^\infty e^{-t\psi(x)}\, dx = \int_0^\delta e^{-t\psi(x)}\, dx + \int_\delta^\infty e^{-t\psi(x)}\, dx =: I_1(t) + I_2(t).$$

Tail estimate *Since* ψ *is increasing and* $\psi(\delta) > 0$*,*

$$0 \leq I_2(t) \leq= e^{-(t-1)\psi(\delta)} \int_\delta^\infty e^{-\psi(x)}\, dx \leq C\, e^{-ct}$$

for some $C, c > 0$*.*

Main part *On* $[0, \delta]$ *the function* ψ *is strictly increasing, hence a* C^1 *bijection onto* $[0, \psi(\delta)]$*. Make the change of variables*

$$y = t\,\psi(x), \qquad x = x_t(y) := \psi^{-1}(y/t),$$

to get

$$I_1(t) = \frac{1}{t} \int_0^{t\psi(\delta)} e^{-y}\, \frac{1}{\psi'(x_t(y))}\, dy.$$

By (1)*,*

$$\frac{1}{a + \varepsilon} \leq \frac{1}{\psi'(x_t(y))} \leq \frac{1}{a - \varepsilon} \qquad (0 \leq y \leq t\psi(\delta)).$$

Moreover, $x_t(y) = \psi^{-1}(y/t) \to 0$ *for each fixed* y *as* $t \to \infty$*, so* $\frac{1}{\psi'(x_t(y))} \to \frac{1}{a}$*. By the dominated convergence theorem,*

$$\lim_{t\to\infty} t\, I_1(t) = \int_0^\infty e^{-y}\, \frac{1}{a}\, dy = \frac{1}{a}.$$

Combining with the tail estimate,

$$\int_0^\infty e^{-t\psi(x)}\, dx = \frac{1}{t}\,\frac{1}{a} + o\left(\frac{1}{t}\right).$$

Therefore

$$I(t) = e^{th(0)} \left(\frac{1}{at} + o\left(\frac{1}{t}\right)\right) = \frac{e^{th(0)}}{-t\, h'(0)}\, (1 + o(1)),$$

which proves the claimed asymptotic.

Remark 5.2. *Use the change of variables.*

Solution 5.5. *Recall that for a random variable* X *with moment generating function*

$$M(\lambda) = \mathbb{E}[e^{\lambda X}], \qquad \Lambda(\lambda) = \log M(\lambda),$$

the rate function (Legendre–Fenchel transform) is defined as

$$I(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\}.$$

(1) Normal distribution *$X \sim N(\mu, \sigma^2)$.*

$$M(\lambda) = \exp\Big(\mu\lambda + \tfrac{1}{2}\sigma^2\lambda^2\Big), \quad \Lambda(\lambda) = \mu\lambda + \tfrac{1}{2}\sigma^2\lambda^2.$$

Thus

$$I(x) = \sup_{\lambda\in\mathbb{R}}\left\{\lambda(x-\mu) - \tfrac{1}{2}\sigma^2\lambda^2\right\}.$$

Maximizing in λ gives $\lambda^* = (x-\mu)/\sigma^2$, and

$$\boxed{I(x) = \frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}.$$

(2) Exponential distribution *$X \sim \mathrm{Exp}(\lambda)$ with pdf $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$.*

$$M(\theta) = \mathbb{E}[e^{\theta X}] = \frac{\lambda}{\lambda-\theta}, \quad \theta < \lambda, \qquad \Lambda(\theta) = -\log\Big(1 - \frac{\theta}{\lambda}\Big).$$

Then

$$I(x) = \sup_{\theta<\lambda}\left\{\theta x + \log\Big(1-\frac{\theta}{\lambda}\Big)\right\}.$$

Setting the derivative to zero gives

$$x - \Lambda'(\theta) = 0 \quad \Longrightarrow \quad \Lambda'(\theta) = \frac{1}{\lambda-\theta} = x,$$

so $\theta^* = \lambda - \frac{1}{x}$ (valid only for $x > 0$). Then

$$1 - \frac{\theta^*}{\lambda} = \frac{1}{\lambda x}, \qquad \Lambda(\theta^*) = \log(\lambda x).$$

Therefore

$$I(x) = \theta^* x - \Lambda(\theta^*) = (\lambda x - 1) - \log(\lambda x), \qquad x > 0.$$

Since the exponential distribution is supported on $[0,\infty)$,

$$\boxed{I(x) = \begin{cases} \lambda x - 1 - \log(\lambda x), & x > 0, \\ +\infty, & x \leq 0. \end{cases}}$$

Note that $I(x)$ attains its minimum 0 at $x = 1/\lambda$, the mean of the exponential distribution.

6 Homework 6 (Sep 29th)

Solution 6.1. ***Ehrenfest's model.*** *Consider the classical Ehrenfest urn model with N identical particles (or balls) distributed between two boxes (labeled A and B). At each discrete time step, one of the N particles is chosen uniformly at random and moved to the other box.*

Let X_t denote the number of particles in box A at time t. Then $\{X_t\}_{t\geq 0}$ is a Markov chain with state space $\{0,1,2,\ldots,N\}$ and transition probabilities

$$P(i,i+1)=\frac{N-i}{N},\qquad P(i,i-1)=\frac{i}{N},\qquad i=0,1,\ldots,N.$$

These correspond respectively to moving a ball from B to A and from A to B.

Invariant (stationary) distribution. *We seek $\pi=(\pi_0,\pi_1,\ldots,\pi_N)$ such that $\pi P=\pi$. The detailed balance equations (reversibility) are*

$$\pi_i\,P(i,i+1)=\pi_{i+1}\,P(i+1,i),$$

which gives

$$\pi_{i+1}=\pi_i\,\frac{P(i,i+1)}{P(i+1,i)}=\pi_i\,\frac{N-i}{i+1}.$$

Iterating from $i=0$ yields

$$\pi_i=\pi_0\,\frac{N!}{i!\,(N-i)!}=\pi_0\binom{N}{i}.$$

Normalizing so that $\sum_{i=0}^{N}\pi_i=1$ gives

$$\pi_i=2^{-N}\binom{N}{i},\qquad i=0,1,\ldots,N.$$

Interpretation. *The invariant distribution is the binomial distribution* $\mathrm{Bin}(N,1/2)$. *Intuitively, in equilibrium, each particle independently occupies either box A or B with equal probability* $1/2$. *Thus the probability of having i particles in box A is $\pi_i=\binom{N}{i}2^{-N}$.*

Solution 6.2. *Let $\{N(t)\}_{t\geq 0}$ be a (simple) Poisson process with rate $\lambda>0$, characterized by*

(i) $N(0)=0$, *(ii) stationary independent increments,*
(iii) $\mathbb{P}\{N(h)=1\}=\lambda h+o(h)$, $\mathbb{P}\{N(h)\geq 2\}=o(h)$ $(h\downarrow 0)$.

For fixed $t>0$ define the characteristic function of $N(t)$:

$$\phi_t(u):=\mathbb{E}\big[e^{iuN(t)}\big],\qquad u\in\mathbb{R}.$$

Using independent, stationary increments and the small–time behavior (iii), for $h>0$ small we write, conditioning on the increment $N(t+h)-N(t)$,

$$\phi_{t+h}(u)=\mathbb{E}\Big[e^{iuN(t)}\mathbb{E}\Big(e^{iu(N(t+h)-N(t))}\mid N(t)\Big)\Big]=\phi_t(u)\,\mathbb{E}\Big(e^{iuN(h)}\Big).$$

By (iii),

$$\mathbb{E}\Big(e^{iuN(h)}\Big)=1\cdot\mathbb{P}\{N(h)=0\}+e^{iu}\mathbb{P}\{N(h)=1\}+\mathbb{E}\Big(e^{iuN(h)};N(h)\geq 2\Big)=1+\lambda h\,(e^{iu}-1)+o(h).$$

Hence

$$\phi_{t+h}(u)-\phi_t(u)=\phi_t(u)\,\big(\lambda h\,(e^{iu}-1)+o(h)\big),$$

so, dividing by h and letting $h\downarrow 0$,

$$\frac{d}{dt}\phi_t(u)=\lambda\,(e^{iu}-1)\,\phi_t(u),\qquad \phi_0(u)=1.$$

Solving this linear ODE gives

$$\boxed{\phi_t(u) = \exp\{\lambda t\,(e^{iu}-1)\}}.$$

To extract the distribution of $N(t)$*, expand:*

$$\phi_t(u) = e^{-\lambda t}\exp\big(\lambda t\, e^{iu}\big) = e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}\,e^{iun}.$$

Comparing with the general form $\phi_t(u) = \sum_{n\ge 0} e^{iun}\mathbb{P}\{N(t)=n\}$*, we read off*

$$\boxed{\mathbb{P}\{N(t)=n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \qquad n=0,1,2,\dots}.$$

Thus $N(t)\sim \text{Poisson}(\lambda t)$*.*

Solution 6.3. *Let* $X(t)$ *be a CTMC on a countable state space* S *with generator* $Q=(q_{ij})$*, where* $q_{ij}\ge 0$ *for* $j\ne i$ *and* $q_{ii} = -\sum_{j\ne i} q_{ij} =: -q_i$*. For a bounded (or suitable) function* $f: S\to\mathbb{R}$ *define*

$$h_i(t) := \mathbb{E}^i[f(X(t))], \qquad i\in S,\ t\ge 0.$$

Equivalently, letting $(P_t)_{t\ge 0}$ *be the transition semigroup,* $h(t)=P_t f$*.*

Fix i *and condition on what happens in the first small interval* $[0,dt]$*: with probability* $1-q_i\,dt+o(dt)$ *the chain stays in* i*; with probability* $q_{ij}\,dt+o(dt)$ *it jumps to* $j\ne i$*. Using the Markov property at time* dt*,*

$$\begin{aligned} h_i(t+dt) = \mathbb{E}^i[f(X(t+dt))] &= \big(1-q_i\,dt\big)\mathbb{E}^i[f(X(t))] + \sum_{j\ne i} q_{ij}\,dt\,\mathbb{E}^j[f(X(t))] + o(dt) \\ &= \big(1-q_i\,dt\big)h_i(t) + \sum_{j\ne i} q_{ij}\,dt\,h_j(t) + o(dt). \end{aligned}$$

Hence

$$\frac{h_i(t+dt)-h_i(t)}{dt} = \sum_{j\ne i} q_{ij}\big(h_j(t)-h_i(t)\big) + o(1) = \sum_{j\in S} q_{ij}\,h_j(t) + o(1),$$

because $q_{ii} = -\sum_{j\ne i} q_{ij}$*. Letting* $dt\downarrow 0$ *we obtain the* backward Kolmogorov equation

$$\boxed{\frac{d}{dt}h_i(t) = \sum_{j\in S} q_{ij}\,h_j(t), \qquad h_i(0)=f(i).}$$

In vector form, with $h(t)=(h_i(t))_{i\in S}$ *and* $f=(f(i))_{i\in S}$*,*

$$\boxed{\frac{d}{dt}h(t) = Q\,h(t), \qquad h(0)=f.}$$

Solution 6.4. *Let each trial (coin toss) occur every* τ *units of time with success probability* p*. Let* $N(t)$ *be the number of successes by time* t*. Then*

$$N(t)\sim \text{Binomial}\Big(n=\frac{t}{\tau},\,p\Big), \qquad \mathbb{P}\{N(t)=k\} = \binom{t/\tau}{k}p^k(1-p)^{t/\tau-k}.$$

Now take the limit $p\to 0$*,* $\tau\to 0$ *with* $\dfrac{p}{\tau}\to\lambda>0$*. Let* $n=t/\tau\to\infty$*, so* $np=\lambda t$*. Then*

$$\binom{n}{k}p^k(1-p)^{n-k} \approx \frac{(np)^k}{k!}e^{-np} = e^{-\lambda t}\frac{(\lambda t)^k}{k!}.$$

Hence

$$\mathbb{P}\{N(t) = k\} \to e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

which is the Poisson(λt) *law.*

Conclusion: *Under $p, \tau \to 0$ with $p/\tau \to \lambda$, the binomial counting process converges to a Poisson process with rate λ.*

Solution 6.5. *Let $\{N(t)\}_{t \geq 0}$ be a (nonhomogeneous) Poisson process with time–varying rate $\lambda(t) > 0$ in the sense that for $h \downarrow 0$,*

$$\mathbb{P}\{N(t+h) - N(t) = 1\} = \lambda(t)\, h + o(h), \qquad \mathbb{P}\{N(t+h) - N(t) \geq 2\} = o(h),$$

and the increments over disjoint intervals are independent.

State probabilities. *Let $p_m(t) := \mathbb{P}\{N(t) = m\}$ and set the cumulative rate $\Lambda(t) := \int_0^t \lambda(u)\, du$. The forward equations are*

$$p_m'(t) = -\lambda(t)\, p_m(t) + \lambda(t)\, p_{m-1}(t), \qquad p_m(0) = \mathbf{1}_{\{m=0\}}.$$

Solving (e.g. by induction or generating functions) yields the Poisson law with mean $\Lambda(t)$:

$$\boxed{p_m(t) = e^{-\Lambda(t)} \frac{\Lambda(t)^m}{m!}, \qquad m = 0, 1, 2, \ldots}$$

7 Homework 7 (Oct 27th)

Solution 7.1. ***Estimate*** *$R = \mathbb{E}_{\pi_b}[w(X)]$ with $w(x) = \exp(x^2(1/b - 1/a))$, $\pi_b(x) \propto e^{-x^2/b}\mathbf{1}_{[L,U]}(x)$, $[L, U] = [-10, 10]$.*

Algorithm 1 Metropolis–Hastings targeting π_b with uniform independence proposal

Require: $a > 0, b > 0$, total steps N; $L \leftarrow -10$, $U \leftarrow 10$

draw $x_0 \sim \text{Unif}[L, U]$

for $t = 1 \ldots N$ **do**

  propose $x' \sim \text{Unif}[L, U]$ ▷ $q(x'|x) = \text{Unif}[L, U]$

  $\alpha \leftarrow \min\big(1, \exp\big(-(x'^2 - x_{t-1}^2)/b\big)\big)$

  draw $u \sim \text{Unif}(0, 1)$; $x_t \leftarrow \begin{cases} x', & u < \alpha \\ x_{t-1}, & \text{else} \end{cases}$

Output: $\widehat{R} = \frac{1}{N} \sum_{t=1}^{N} \exp\big(x_t^2(1/b - 1/a)\big)$

Choose sample size $N = 200000$, and the results are as follows.

```
Estimated  R  = 0.9127813562
True       R  = 0.9129045361
Abs. error    = 1.232e-04
```

Solution 7.2. *Check the detailed balance condition for Metropolis and Glauber dynamics.*

For Gibbs distribution $\pi(\sigma) = Z^{-1}e^{-\beta H(\sigma)}$, *with symmetric proposal* $Q(\sigma \to \sigma')$, *the detailed balance condition requires*

$$\pi(\sigma)P(\sigma \to \sigma') = \pi(\sigma')P(\sigma' \to \sigma).$$

(1) ***Metropolis:*** $A_M(\sigma \to \sigma') = \min\{1, e^{-\beta\Delta H}\}$, *where* $\Delta H = H(\sigma') - H(\sigma)$. *Then*

$$\frac{A_M(\sigma \to \sigma')}{A_M(\sigma' \to \sigma)} = e^{-\beta\Delta H} = \frac{\pi(\sigma')}{\pi(\sigma)},$$

so detailed balance holds.
(2) ***Glauber:*** $A_G(\sigma \to \sigma') = (1 + e^{\beta\Delta H})^{-1}$. *Similarly,*

$$\frac{A_G(\sigma \to \sigma')}{A_G(\sigma' \to \sigma)} = e^{-\beta\Delta H} = \frac{\pi(\sigma')}{\pi(\sigma)},$$

thus Glauber dynamics also satisfies detailed balance.

Solution 7.3. *Check that the Markov chains set up by Metropolis and Glauber dynamics for the Ising model are both primitive.*

In the single–spin–flip scheme, any configuration σ *can reach any* σ' *by flipping spins one by one (finite number of steps). Each transition has strictly positive probability:*

$$Q(\sigma \to \sigma') > 0, \quad A_M > 0, \quad A_G \in (0, 1),$$

hence $P^\tau(\sigma, \sigma') > 0$ *for some finite* τ. *Therefore, both Metropolis and Glauber dynamics define* primitive *Markov chains.*

8 Homework 8 (Nov 17th)

Solution 8.1. *The state space is* $\mathscr{S} = \mathscr{X} \times \{1, \dots, L\}$. *For* $s = (x, i)$ *and* $s' = (y, j)$, *the transition kernel is*

$$P\big((x,i),(y,j)\big) = \alpha_0\,\mathbf{1}_{\{j=i\}}\,T_i(x,y) + (1-\alpha_0)\,\mathbf{1}_{\{y=x\}}\left[\alpha(i,j)a_{ij}(x)\mathbf{1}_{\{j\neq i\}} + r_i(x)\mathbf{1}_{\{j=i\}}\right],$$

where T_i *is the MCMC transition at level* i,

$$a_{ij}(x) = \min\left\{1,\ \frac{\pi_{st}(x,j)\,\alpha(j,i)}{\pi_{st}(x,i)\,\alpha(i,j)}\right\},$$

and the remaining–stay probability is

$$r_i(x) = 1 - \sum_{k\neq i}\alpha(i,k)a_{ik}(x).$$

Solution 8.2. *The state space is* $\mathscr{S} = \mathscr{X}^L$. *For* $x = (x_1, \dots, x_L)$ *and* $y = (y_1, \dots, y_L)$,

$$P(x,y) = \alpha_0\prod_{\ell=1}^{L}T_\ell(x_\ell,y_\ell) + (1-\alpha_0)\,\frac{1}{L-1}\sum_{i=1}^{L-1}\left[a_i(x)\,\mathbf{1}_{\{y=x^{(i\leftrightarrow i+1)}\}} + (1-a_i(x))\,\mathbf{1}_{\{y=x\}}\right],$$

where T_ℓ *is the MCMC kernel at temperature level* ℓ,

$$a_i(x) = \min\left\{1,\ \frac{\pi_i(x_{i+1})\pi_{i+1}(x_i)}{\pi_i(x_i)\pi_{i+1}(x_{i+1})}\right\},$$

and $x^{(i\leftrightarrow i+1)}$ *denotes the vector obtained by swapping the* i*th and* $(i+1)$*th coordinates of* x.

9 Homework 9 (Nov 17th)

Solution 9.1. *As $\beta \to \infty$, the Boltzmann distribution*

$$\pi_\beta(x) \propto e^{-\beta V(x)}$$

concentrates on the set of global minimizers of V.

If V has finitely many isolated minimizers $\{x_1, \ldots, x_m\}$, then

$$\pi_\beta \Longrightarrow \frac{1}{m} \sum_{k=1}^{m} \delta_{x_k},$$

i.e. the limiting distribution is the uniform distribution over all global minimizers.

10 Homework 10 (Nov 17th)

Solution 10.1. *Let*

$$W(x,t) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right).$$

Direct differentiation gives

$$\frac{\partial W}{\partial t} = \left(\frac{x^2}{4D^2t^2} - \frac{1}{2Dt}\right) DW, \qquad \frac{\partial^2 W}{\partial x^2} = \left(\frac{x^2}{4D^2t^2} - \frac{1}{2Dt}\right) W.$$

Thus $\partial_t W = D\, \partial_x^2 W$. As $t \to 0$, $W(x,t) \to \delta(x)$ in distribution.

Solution 10.2. *The reflecting-barrier solution is*

$$W_r(x,t;x_1) = \frac{1}{\sqrt{4\pi D t}} \left[e^{-x^2/(4Dt)} + e^{-(2x_1-x)^2/(4Dt)}\right].$$

Each term solves $\partial_t W = D\partial_x^2 W$, hence the sum does too. Initial condition: $W_r(x,0) = \delta(x)$. Neumann boundary:

$$\partial_x W_r(x_1, t) = 0$$

because the two exponentials have opposite derivatives at $x = x_1$.

Solution 10.3. *The absorbing-barrier solution is*

$$W_a(x,t;x_1) = \frac{1}{\sqrt{4\pi D t}} \left[e^{-x^2/(4Dt)} - e^{-(2x_1-x)^2/(4Dt)}\right].$$

It satisfies $\partial_t W_a = D\partial_x^2 W_a$ since each term does. Initial condition: $W_a(x,0) = \delta(x)$. Dirichlet boundary:

$$W_a(x_1, t) = 0$$

because the two terms cancel at $x = x_1$.

11 Homework 11 (Nov 17th)

Solution 11.1. *We have $\mathbb{E}[\xi_1] = \frac{2}{3} - \frac{1}{3} = \frac{1}{3} \neq 0$, hence*

$$\mathbb{E}[S_{\lfloor Nt \rfloor}] = \tfrac{1}{3} N t.$$

Thus

$$Z_t^N = \frac{S_{\lfloor Nt \rfloor}}{N^\alpha}$$

has mean

$$\mathbb{E}[Z_t^N] = \tfrac{1}{3} t\, N^{1-\alpha}.$$

To obtain a finite nontrivial limit we must have $1 - \alpha = 0$, *i.e.* $\alpha = 1$. *Moreover,*

$$Z_t^N = \frac{1}{N}\Big(\tfrac{1}{3} Nt + O(\sqrt{N})\Big) \to \tfrac{1}{3} t.$$

Hence the limit process is $Z_t = \frac{1}{3} t$.

Solution 11.2. *(a) Since* $W_t \sim N(0,t)$, *we can write* $W_t \overset{d}{=} \sqrt{t}\, W_1$ *with* $W_1 \sim N(0,1)$. *Hence*

$$\mathbb{E} W_t^4 = t^2 \mathbb{E} W_1^4 = t^2 \cdot 3 = 3t^2.$$

(b) Let $X = W_t - W_s + W_z$. *Then* $\mathbb{E}X^2 = \mathrm{Var}(X)$ *and, for a Wiener process,* $\mathrm{Cov}(W_u, W_v) = \min(u,v)$. *Thus*

$$\mathbb{E}(W_t - W_s + W_z)^2 = \mathrm{Var}(X) = t + s + z + 2\big[-\min(t,s) + \min(t,z) - \min(s,z)\big].$$

Solution 11.3. *Since* $X \sim N(0, A)$, *its density is*

$$f_X(x) = \frac{1}{(2\pi)^{n/2} \det(A)^{1/2}} \exp\!\left(-\frac{1}{2} x^\top A^{-1} x\right).$$

Then

$$\mathbb{E}\exp\big(-\tfrac{1}{2} X^\top B X\big) = \frac{1}{(2\pi)^{n/2} \det(A)^{1/2}} \int_{\mathbb{R}^n} \exp\!\left(-\frac{1}{2} x^\top (A^{-1} + B) x\right) dx.$$

Using the Gaussian integral $\int_{\mathbb{R}^n} e^{-\frac{1}{2} x^\top M x} dx = (2\pi)^{n/2} \det(M)^{-1/2}$ *for* $M \succ 0$, *we get*

$$\mathbb{E} e^{-\frac{1}{2} X^\top B X} = \det(A)^{-1/2} \det(A^{-1} + B)^{-1/2}.$$

Since $\det(A^{-1} + B) = \det(A^{-1}) \det(I + AB)$, *this simplifies to*

$$\boxed{\mathbb{E} e^{-\frac{1}{2} X^\top B X} = \det(I + AB)^{-1/2}}.$$

Solution 11.4. *We show that (1)–(3) are equivalent to (1')–(3').*

(1)–(3) $\Rightarrow$ ***(1')–(3').*** *Assume* $(W_t)_{t \geq 0}$ *is a Gaussian process with* $\mathbb{E}W_t = 0$ *and* $\mathrm{Cov}(W_s, W_t) = s \wedge t$.

(2') *For any* $s, t \geq 0$, *the increment* $W_{s+t} - W_s$ *is Gaussian (since the process is Gaussian) with mean*

$$\mathbb{E}(W_{s+t} - W_s) = 0,$$

and variance

$$\mathrm{Var}(W_{s+t} - W_s) = \mathrm{Var}(W_{s+t}) + \mathrm{Var}(W_s) - 2\,\mathrm{Cov}(W_{s+t}, W_s) = (s+t) + s - 2s = t.$$

Hence $W_{s+t} - W_s \sim N(0,t)$.

(1') *Let* $t_0 < t_1 < \cdots < t_n$ *and set* $X_0 = W_{t_0}$, $X_k = W_{t_k} - W_{t_{k-1}}$ *for* $k \geq 1$. *For* $i \neq j$ *one checks, using* $\mathrm{Cov}(W_s, W_t) = s \wedge t$, *that* $\mathrm{Cov}(X_i, X_j) = 0$ *(a simple computation with* $s \wedge t$*). Since the vector* $(X_0, \ldots, X_n)$ *is Gaussian and its covariance matrix is diagonal, the* X_i *are independent. Thus (1') holds. Condition (3') is identical to (3).*

(1')–(3')* ⇒ *(1)–(3). *Assume now (1')–(3').*

From (2') with $s = 0$ *we get* $W_t \sim N(0,t)$*, so* $\mathbb{E}W_t = 0$ *and* $\operatorname{Var}(W_t) = t$*. By (1') the increments* $W_{t_0}, W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}$ *are independent and by (2') each increment is Gaussian. Hence any finite vector* $(W_{t_1}, \dots, W_{t_n})$ *is a linear combination of independent Gaussian random variables, so it is multivariate Gaussian. Thus* (W_t) *is a Gaussian process.*

To compute the covariance, let $0 \le s \le t$*. Write* $W_t = W_s + (W_t - W_s)$*. Then*

$$\operatorname{Cov}(W_s, W_t) = \operatorname{Cov}(W_s, W_s) + \operatorname{Cov}\big(W_s, W_t - W_s\big).$$

By (1') the increment $W_t - W_s$ *is independent of* W_s*, so the second covariance is* 0*, hence*

$$\operatorname{Cov}(W_s, W_t) = \operatorname{Var}(W_s) = s.$$

For $t < s$ *the same argument gives* $\operatorname{Cov}(W_s, W_t) = t$*, so* $\operatorname{Cov}(W_s, W_t) = s \wedge t$*. Finally, (3) is the same as (3'). Thus (1)–(3) hold.*

12 Homework 12 (Dec 1st)

Solution 12.1. *Let* $(W_t)_{t \ge 0}$ *be a Wiener process.*

(i) $Y_t = \frac{1}{\sqrt{c}} W_{ct}$***.*** *For* $0 \le s < t$*,*

$$Y_t - Y_s = \frac{1}{\sqrt{c}}(W_{ct} - W_{cs}) \sim \mathcal{N}(0, t - s),$$

and the increments are independent by those of W*. Continuity follows from that of* W*. Hence* Y *is a Wiener process.*

(ii) $Z_t = W(T) - W(T - t)$***.*** *For* $0 \le s < t \le T$*,*

$$Z_t - Z_s = W(T - s) - W(T - t) \sim \mathcal{N}(0, t - s),$$

with independent increments since they correspond to disjoint time intervals of W*. Continuity is inherited from* W*, so* Z *is a Wiener process.*

(iii) $X_t = tW_{1/t}$***,*** $X_0 = 0$***.*** *For* $s, t \in (0, 1]$*,*

$$\operatorname{Cov}(X_s, X_t) = st \operatorname{Cov}(W_{1/s}, W_{1/t}) = st \min\left\{\frac{1}{s}, \frac{1}{t}\right\} = \min\{s, t\}.$$

Thus (X_t) *is a centered Gaussian process with covariance* $\operatorname{Cov}(X_s, X_t) = \min\{s, t\}$*, the same as Brownian motion. Hence* X *has the same finite-dimensional distributions as* (W_t) *and, with* $X_0 = 0$*, is a Wiener process.*

Solution 12.2. *Let* W *be Brownian motion and let* $\Delta = \{0 = t_0 < \cdots < t_m = t\}$ *be a partition. Define*

$$Q_t^\Delta := \sum_{k=0}^{m-1} \big(W_{t_{k+1}} - W_{t_k}\big)^2.$$

Write $\Delta W_k := W_{t_{k+1}} - W_{t_k}$ *and* $\Delta t_k := t_{k+1} - t_k$*. Then the increments* ΔW_k *are independent and* $\Delta W_k \sim \mathcal{N}(0, \Delta t_k)$*.*

Proposition 2.2. *We have* $\mathbb{E}[\Delta W_k^2] = \Delta t_k$*, hence*

$$\mathbb{E}[Q_t^\Delta] = \sum_{k=0}^{m-1} \Delta t_k = t.$$

Moreover,

$$Q_t^\Delta - t = \sum_{k=0}^{m-1} \big(\Delta W_k^2 - \Delta t_k\big),$$

and the summands are independent with mean 0*, so*

$$\mathbb{E}(Q_t^\Delta - t)^2 = \sum_{k=0}^{m-1} \mathrm{Var}(\Delta W_k^2).$$

If $X \sim \mathscr{N}(0,\sigma^2)$ *then* $\mathbb{E}[X^4] = 3\sigma^4$*, so* $\mathrm{Var}(X^2) = 3\sigma^4 - \sigma^4 = 2\sigma^4$*. With* $\sigma^2 = \Delta t_k$ *this gives* $\mathrm{Var}(\Delta W_k^2) = 2(\Delta t_k)^2$*, hence*

$$\mathbb{E}(Q_t^\Delta - t)^2 = 2\sum_{k=0}^{m-1} (\Delta t_k)^2.$$

In particular, since $\sum(\Delta t_k)^2 \le |\Delta| \sum \Delta t_k = |\Delta|\, t$*, we get* $\mathbb{E}(Q_t^\Delta - t)^2 \le 2t|\Delta| \to 0$ *as* $|\Delta| \to 0$*, i.e.* $Q_t^\Delta \to t$ *in* L^2*.*

HW2 sharpening (dyadic partition). *Fix* $t > 0$ *and set* $t_k = k2^{-n}t$ *for* $k = 0, 1, \ldots, 2^n$*. Define*

$$Y_n(t) := \sum_{k=0}^{2^n-1} \big(W_{t_{k+1}} - W_{t_k}\big)^2.$$

Applying Proposition 2.2,

$$\mathrm{Var}(Y_n(t)) = \mathbb{E}(Y_n(t) - t)^2 = 2\sum_{k=0}^{2^n-1} \left(\frac{t}{2^n}\right)^2 = \frac{2t^2}{2^n}.$$

By Chebyshev, for any $\varepsilon > 0$*,*

$$\mathbb{P}\big(|Y_n(t) - t| > \varepsilon\big) \le \frac{\mathrm{Var}(Y_n(t))}{\varepsilon^2} = \frac{2t^2}{\varepsilon^2\, 2^n}.$$

Since $\sum_{n=1}^\infty 2^{-n} < \infty$*, Borel–Cantelli implies*

$$\mathbb{P}\big(|Y_n(t) - t| > \varepsilon \text{ i.o.}\big) = 0 \qquad \text{for every } \varepsilon > 0.$$

Hence $|Y_n(t) - t| \to 0$ *almost surely, i.e.* $Y_n(t,\omega) \to t$ *a.s.*

Solution 12.3. *Let* $C[0,\infty)$ *be the space of real-valued continuous functions on* $[0,\infty)$*. For* $x, y \in C[0,\infty)$ *define*

$$d(x,y) := \sum_{n=1}^\infty 2^{-n}\Big(\|x-y\|_{\infty,[0,n]} \wedge 1\Big), \qquad \|f\|_{\infty,[0,n]} := \sup_{t\in[0,n]} |f(t)|.$$

(A) Completeness. *Let* (x_k) *be a* d*-Cauchy sequence in* $C[0,\infty)$*. Fix* $m \in \mathbb{N}$*. We claim that* (x_k) *is Cauchy in the uniform norm on* $[0,m]$*.*

Indeed, for any $\varepsilon \in (0,1)$ *choose* K *such that for all* $k, \ell \ge K$*,* $d(x_k, x_\ell) < 2^{-m}\varepsilon$*. Then*

$$2^{-m}\big(\|x_k - x_\ell\|_{\infty,[0,m]} \wedge 1\big) \le d(x_k, x_\ell) < 2^{-m}\varepsilon,$$

hence $(\|x_k - x_\ell\|_{\infty,[0,m]} \wedge 1) < \varepsilon$*. Since* $\varepsilon < 1$*, this implies*

$$\|x_k - x_\ell\|_{\infty,[0,m]} < \varepsilon,$$

so (x_k) *is Cauchy in* $(C[0,m], \|\cdot\|_{\infty,[0,m]})$. *Because* $(C[0,m], \|\cdot\|_\infty)$ *is complete, there exists a continuous function* $x^{(m)} \in C[0,m]$ *such that*

$$\|x_k - x^{(m)}\|_{\infty,[0,m]} \to 0 \quad \text{as } k \to \infty.$$

Compatibility. *If* $m < r$, *then on* $[0,m]$ *we have both* $x_k \to x^{(m)}$ *uniformly and* $x_k \to x^{(r)}$ *uniformly, hence* $x^{(r)} = x^{(m)}$ *on* $[0,m]$ *by uniqueness of uniform limits. Therefore the family* $\{x^{(m)}\}$ *is consistent.*

Define $x : [0,\infty) \to \mathbb{R}$ *by setting* $x(t) := x^{(m)}(t)$ *for any* $m > t$. *This is well-defined by compatibility, and* x *is continuous on each* $[0,m]$, *hence continuous on* $[0,\infty)$, *i.e.* $x \in C[0,\infty)$.

Convergence in d. *Fix* $\varepsilon > 0$. *Choose* M *so that* $\sum_{n>M} 2^{-n} < \varepsilon/2$. *For each* $1 \le n \le M$, *since* $\|x_k - x\|_{\infty,[0,n]} \to 0$, *choose* K *such that for all* $k \ge K$,

$$\|x_k - x\|_{\infty,[0,n]} < \frac{\varepsilon}{2} \quad \text{for all } 1 \le n \le M.$$

Then for $k \ge K$,

$$d(x_k, x) = \sum_{n=1}^{\infty} 2^{-n}\big(\|x_k - x\|_{\infty,[0,n]} \wedge 1\big) \le \sum_{n=1}^{M} 2^{-n}\|x_k - x\|_{\infty,[0,n]} + \sum_{n>M} 2^{-n} \le \frac{\varepsilon}{2}\sum_{n=1}^{M} 2^{-n} + \frac{\varepsilon}{2} < \varepsilon.$$

Hence $x_k \to x$ *in* $(C[0,\infty), d)$, *so the space is complete.*

(B) Separability. *Let* $\mathscr{D}$ *be the set of functions* $f \in C[0,\infty)$ *such that: for some* $N \in \mathbb{N}$,

- f *is piecewise linear on each interval* $[j, j+1]$ *for* $j = 0, 1, \ldots, N-1$, *with breakpoints at rational points in* $[0,N]$ *and values in* $\mathbb{Q}$;
- $f(t) = 0$ *for all* $t \ge N$.

Then $\mathscr{D}$ *is countable: it is a countable union over* N *of functions determined by finitely many rational breakpoints and finitely many rational values.*

We show $\mathscr{D}$ *is dense in* $(C[0,\infty), d)$. *Let* $x \in C[0,\infty)$ *and* $\varepsilon > 0$. *Choose* M *such that* $\sum_{n>M} 2^{-n} < \varepsilon/2$. *On* $[0,M]$, *by uniform continuity of* x *on the compact interval* $[0,M]$, *there exists a partition* $0 = s_0 < \cdots < s_L = M$ *fine enough so that the piecewise linear interpolation* $\tilde{f}$ *of* x *on this partition satisfies*

$$\|x - \tilde{f}\|_{\infty,[0,M]} < \delta,$$

where $\delta > 0$ *will be chosen. Approximating the finitely many breakpoints* s_i *by rationals and the finitely many values* $x(s_i)$ *by rationals, we obtain a piecewise linear* $f \in \mathscr{D}$ *(also set* $f(t) = 0$ *for* $t \ge M$*) such that*

$$\|x - f\|_{\infty,[0,M]} < \delta.$$

Then for $1 \le n \le M$, $\|x - f\|_{\infty,[0,n]} \le \|x - f\|_{\infty,[0,M]} < \delta$, *hence*

$$d(x, f) \le \sum_{n=1}^{M} 2^{-n}\delta + \sum_{n>M} 2^{-n} \le \delta + \frac{\varepsilon}{2}.$$

Choose $\delta = \varepsilon/2$ *to get* $d(x,f) < \varepsilon$. *Thus* $\mathscr{D}$ *is dense, and the space is separable.*

Combining (A) and (B), $(C[0,\infty), d)$ *is a complete, separable metric space.*

Solution 12.4. *Let $(W_u)_{u\geq 0}$ be a Brownian motion and fix $0 \leq s < t$. Put*

$$m := \frac{s+t}{2}, \qquad A := W_m - W_s, \qquad B := W_t - W_m.$$

By independent increments, A and B are independent, and

$$A \sim \mathscr{N}(0, m-s), \qquad B \sim \mathscr{N}(0, t-m).$$

Since $m = \frac{s+t}{2}$, we have $m - s = t - m = \frac{t-s}{2}$, hence

$$A, B \text{ i.i.d. } \mathscr{N}\left(0, \frac{t-s}{2}\right).$$

Now note that

$$W_m = W_s + A, \qquad W_t = W_s + A + B.$$

Condition on the event $\{W_s = x,\ W_t = y\}$. Then $A+B = y-x$ is fixed, and we need the conditional law of A given $A+B = y-x$ where A, B are i.i.d. centered Gaussians with variance $\sigma^2 := \frac{t-s}{2}$.

For i.i.d. $A, B \sim \mathscr{N}(0, \sigma^2)$, the vector $(A, A+B)$ is jointly Gaussian with

$$\mathbb{E}[A] = 0, \quad \mathbb{E}[A+B] = 0, \quad \operatorname{Var}(A) = \sigma^2, \quad \operatorname{Var}(A+B) = 2\sigma^2, \quad \operatorname{Cov}(A, A+B) = \operatorname{Var}(A) = \sigma^2.$$

Hence, by the Gaussian regression formula,

$$A \,\big|\, (A+B) = c \sim \mathscr{N}\left(\frac{\operatorname{Cov}(A, A+B)}{\operatorname{Var}(A+B)}\,c,\ \operatorname{Var}(A) - \frac{\operatorname{Cov}(A, A+B)^2}{\operatorname{Var}(A+B)}\right) = \mathscr{N}\left(\frac{c}{2},\ \frac{\sigma^2}{2}\right).$$

Taking $c = y - x$ and $\sigma^2 = \frac{t-s}{2}$ gives

$$A \,\big|\, (A+B) = y - x \sim \mathscr{N}\left(\frac{y-x}{2},\ \frac{t-s}{4}\right).$$

Therefore, since $W_m = W_s + A$ and $W_s = x$,

$$W_m \mid (W_s = x, W_t = y) \sim \mathscr{N}\left(x + \frac{y-x}{2},\ \frac{t-s}{4}\right) = \mathscr{N}\left(\frac{x+y}{2},\ \frac{t-s}{4}\right).$$

This is exactly the desired conditional distribution.

13 Homework 13 (Dec 16th)

Solution 13.1. *Let $\Delta = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ be a partition and write $\Delta W_j := W_{t_{j+1}} - W_{t_j}$, $\Delta t_j := t_{j+1} - t_j$, and $t_{j+\frac{1}{2}} := \frac{t_j + t_{j+1}}{2}$.*

(1) Midpoint approximation. *Use the identity (split term method)*

$$(\Delta W_j)^2 = \left(W_{t_{j+1}}^2 - W_{t_j}^2\right) - 2W_{t_j}\Delta W_j = 2W_{t_{j+\frac{1}{2}}}\Delta W_j + \Big(W_{t_{j+1}} - 2W_{t_{j+\frac{1}{2}}} + W_{t_j}\Big)\Delta W_j.$$

Equivalently, the cleanest split is

$$W_{t_{j+\frac{1}{2}}}\Delta W_j = \frac{1}{2}\left(W_{t_{j+1}}^2 - W_{t_j}^2\right) - \frac{1}{2}\left(W_{t_{j+1}} - W_{t_{j+\frac{1}{2}}}\right)^2 + \frac{1}{2}\left(W_{t_{j+\frac{1}{2}}} - W_{t_j}\right)^2,$$

which follows by expanding squares with $a = W_{t_{j+1}} - W_{t_{j+\frac{1}{2}}}$, $b = W_{t_{j+\frac{1}{2}}} - W_{t_j}$ and using $W_{t_{j+\frac{1}{2}}}\Delta W_j = \frac{1}{2}(a+b)(a-b) + \frac{1}{2}(W_{t_{j+1}}^2 - W_{t_j}^2)$.

Summing over j yields

$$\sum_{j=0}^{n-1} W_{t_{j+\frac{1}{2}}} \Delta W_j = \frac{1}{2}\big(W_t^2 - W_0^2\big) + \frac{1}{2}\sum_{j=0}^{n-1}\Big[(W_{t_{j+\frac{1}{2}}} - W_{t_j})^2 - (W_{t_{j+1}} - W_{t_{j+\frac{1}{2}}})^2\Big].$$

Since $W_0 = 0$, it remains to show the last sum goes to 0 in L^2. Let

$$R_\Delta := \sum_{j=0}^{n-1}\Big[(W_{t_{j+\frac{1}{2}}} - W_{t_j})^2 - (W_{t_{j+1}} - W_{t_{j+\frac{1}{2}}})^2\Big].$$

For each j, the two increments are independent $\mathscr{N}(0, \Delta t_j/2)$, so the difference has mean 0 and variance

$$\mathrm{Var}\Big((W_{t_{j+\frac{1}{2}}} - W_{t_j})^2 - (W_{t_{j+1}} - W_{t_{j+\frac{1}{2}}})^2\Big) = 2\,\mathrm{Var}\Big((W_{t_{j+\frac{1}{2}}} - W_{t_j})^2\Big) = 2\cdot 2\Big(\frac{\Delta t_j}{2}\Big)^2 = (\Delta t_j)^2,$$

using $\mathrm{Var}(X^2) = 2\sigma^4$ for $X \sim \mathscr{N}(0,\sigma^2)$. Moreover, for different j these terms are independent (disjoint increments), hence

$$\mathbb{E}[R_\Delta^2] = \mathrm{Var}(R_\Delta) = \sum_{j=0}^{n-1}(\Delta t_j)^2 \le |\Delta| \sum_{j=0}^{n-1} \Delta t_j = |\Delta|\, t \xrightarrow[|\Delta|\to 0]{} 0.$$

Therefore $R_\Delta \to 0$ in L^2, and so

$$\sum_{j=0}^{n-1} W_{t_{j+\frac{1}{2}}}(W_{t_{j+1}} - W_{t_j}) \underset{L^2}{\longrightarrow} \frac{1}{2}W_t^2.$$

(2) Right endpoint approximation. *Use the split*

$$W_{t_{j+1}}\Delta W_j = \frac{1}{2}\big(W_{t_{j+1}}^2 - W_{t_j}^2\big) + \frac{1}{2}(\Delta W_j)^2,$$

since $W_{t_{j+1}} = W_{t_j} + \Delta W_j$. Summing over j gives

$$\sum_{j=0}^{n-1} W_{t_{j+1}}\Delta W_j = \frac{1}{2}W_t^2 + \frac{1}{2}\sum_{j=0}^{n-1}(\Delta W_j)^2.$$

By Proposition 2.2 (quadratic variation), $\sum_j (\Delta W_j)^2 \to t$ in L^2 as $|\Delta| \to 0$. Hence

$$\sum_{j=0}^{n-1} W_{t_{j+1}}(W_{t_{j+1}} - W_{t_j}) \underset{L^2}{\longrightarrow} \frac{1}{2}W_t^2 + \frac{t}{2}.$$

This proves the claimed limits in $L^2(\Omega)$.

Solution 13.2. ***Notation.*** *Define the (probabilists') Hermite polynomials*

$$h_n(x) := (-1)^n e^{x^2/2}\frac{d^n}{dx^n}\big(e^{-x^2/2}\big),$$

and for $a > 0$ define

$$H_n(x,a) := a^{n/2}h_n\left(\frac{x}{\sqrt{a}}\right), \qquad H_n(x,0) := x^n.$$

(a) Generating functions.

First prove

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} h_n(x) = \exp\left(ux - \frac{u^2}{2}\right).$$

Let $G(u,x) := \exp(ux - u^2/2)$*. Using the definition of* h_n *and Taylor expansion,*

$$e^{-x^2/2} G(u,x) = \exp\left(ux - \frac{x^2}{2} - \frac{u^2}{2}\right) = \exp\left(-\frac{(x-u)^2}{2}\right).$$

Differentiate in u *at* $u=0$*:*

$$\frac{\partial^n}{\partial u^n}\left(e^{-x^2/2} G(u,x)\right)\Big|_{u=0} = \frac{\partial^n}{\partial u^n} \exp\left(-\frac{(x-u)^2}{2}\right)\Big|_{u=0} = (-1)^n \frac{d^n}{dx^n} e^{-x^2/2}.$$

Multiplying by $e^{x^2/2}$ *gives*

$$\frac{\partial^n}{\partial u^n} G(u,x)\Big|_{u=0} = h_n(x).$$

Hence the Taylor series of G *in* u *yields the desired generating function.*

For $a > 0$*,*

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} H_n(x,a) = \sum_{n=0}^{\infty} \frac{(\sqrt{a}\,u)^n}{n!} h_n\left(\frac{x}{\sqrt{a}}\right) = \exp\left(\sqrt{a}\,u \frac{x}{\sqrt{a}} - \frac{(\sqrt{a}\,u)^2}{2}\right) = \exp\left(ux - \frac{au^2}{2}\right).$$

Also $H_n(x,0) = x^n$ *is consistent since the right-hand side becomes* e^{ux} *at* $a = 0$*.*

(b) PDE and derivative identities.

Let

$$F(u;x,a) := \sum_{n=0}^{\infty} \frac{u^n}{n!} H_n(x,a) = \exp\left(ux - \frac{au^2}{2}\right).$$

Differentiate F*:*

$$\partial_x F = uF, \qquad \partial_{xx} F = u^2 F, \qquad \partial_a F = -\frac{u^2}{2} F.$$

Thus

$$\left(\frac{1}{2}\partial_{xx} + \partial_a\right) F = \left(\frac{1}{2}u^2 - \frac{1}{2}u^2\right) F = 0.$$

Comparing coefficients of $u^n/n!$ *gives, for each* n*,*

$$\left(\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial a}\right) H_n(x,a) = 0.$$

Also, from $\partial_x F = uF$*,*

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} \partial_x H_n(x,a) = u \sum_{n=0}^{\infty} \frac{u^n}{n!} H_n(x,a) = \sum_{n=1}^{\infty} \frac{u^n}{(n-1)!} H_{n-1}(x,a).$$

Comparing coefficients yields

$$\frac{\partial}{\partial x} H_n(x,a) = n\, H_{n-1}(x,a).$$

(c) Relation (3.3) via Itô + induction.

Let W be Brownian motion and define the iterated Itô integrals

$$I_n(t) := \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} dW_{t_n} \cdots dW_{t_1}, \qquad I_0(t) := 1.$$

Note the recursion (by definition/Fubini for stochastic integrals)

$$I_n(t) = \int_0^t I_{n-1}(s)\, dW_s, \qquad n \geq 1.$$

Now apply Itô's formula to the space-time function $(x,a) \mapsto H_n(x,a)$ at $(x,a) = (W_t, t)$:

$$dH_n(W_t,t) = \partial_x H_n(W_t,t)\, dW_t + \Big(\partial_t H_n(W_t,t) + \frac{1}{2}\partial_{xx} H_n(W_t,t)\Big)\, dt.$$

Using (b) with $a = t$ gives $\partial_t H_n + \frac{1}{2}\partial_{xx} H_n = 0$, hence

$$dH_n(W_t,t) = \partial_x H_n(W_t,t)\, dW_t = n H_{n-1}(W_t,t)\, dW_t.$$

Since $H_n(W_0,0) = H_n(0,0) = 0$ for $n \geq 1$, integrating yields

$$H_n(W_t,t) = n \int_0^t H_{n-1}(W_s,s)\, dW_s.$$

Divide by $n!$:

$$\frac{1}{n!} H_n(W_t,t) = \int_0^t \frac{1}{(n-1)!} H_{n-1}(W_s,s)\, dW_s.$$

Induction. *For $n = 0$, $\frac{1}{0!} H_0(W_t,t) = 1 = I_0(t)$. Assume $\frac{1}{(n-1)!} H_{n-1}(W_t,t) = I_{n-1}(t)$. Then the previous display gives*

$$\frac{1}{n!} H_n(W_t,t) = \int_0^t I_{n-1}(s)\, dW_s = I_n(t).$$

Thus for all $n \geq 0$,

$$I_n(t) = \frac{1}{n!} H_n(W_t,t).$$

Finally, using the definition $H_n(x,a) = a^{n/2} h_n(x/\sqrt{a})$ with $a = t$,

$$I_n(t) = \frac{1}{n!}\, t^{n/2}\, h_n\left(\frac{W_t}{\sqrt{t}}\right),$$

which is exactly relation (3.3).

Solution 13.3. ***(a)*** *$dX_t + \frac{1}{1+t} X_t\, dt = \frac{1}{1+t} dW_t$. Take the integrating factor $\mu(t) = 1 + t$ and set $Y_t = (1+t)X_t$. Then*

$$dY_t = (1+t)dX_t + X_t dt = dW_t,$$

so $Y_t = W_t$ (since $Y_0 = 0$). Hence

$$X_t = \frac{W_t}{1+t}.$$

(b) *$dX_t + X_t\, dt = e^{-t} dW_t$. Take the integrating factor $\mu(t) = e^t$ and set $Y_t = e^t X_t$. Then*

$$dY_t = e^t dX_t + e^t X_t dt = dW_t,$$

so $Y_t = X_0 + W_t$. Hence

$$X_t = e^{-t}(X_0 + W_t).$$

Solution 13.4. *Consider the d-dimensional OU SDE*

$$dX_t = AX_t\,dt + \sigma\,dW_t,$$

where $A \in \mathbb{R}^{d\times d}$, $\sigma \in \mathbb{R}^{d\times m}$, *and* (W_t) *is* m*-dimensional Brownian motion. Let*

$$m_t := \mathbb{E}[X_t], \qquad \Sigma_t := \operatorname{Cov}(X_t) = \mathbb{E}[(X_t - m_t)(X_t - m_t)^\top].$$

Mean. *Taking expectation in the SDE (the Itô term has zero mean) gives*

$$\frac{d}{dt}m_t = Am_t.$$

A stationary mean m_∞ *must satisfy* $\frac{d}{dt}m_t = 0$*, hence*

$$Am_\infty = 0.$$

Covariance. *Let* $M_t := X_t - m_t$*. Then* $dM_t = AM_t\,dt + \sigma\,dW_t$*. Apply Itô to* $M_tM_t^\top$*:*

$$d(M_tM_t^\top) = (dM_t)M_t^\top + M_t(dM_t)^\top + (dM_t)(dM_t)^\top.$$

Using $dM_t = AM_t\,dt + \sigma\,dW_t$ *and* $(dW_t)(dW_t)^\top = I_m\,dt$*,*

$$(dM_t)(dM_t)^\top = \sigma\,dW_t\,dW_t^\top\sigma^\top = \sigma\sigma^\top dt.$$

Taking expectations and noting $\mathbb{E}[dW_t] = 0$ *yields the Lyapunov ODE*

$$\frac{d}{dt}\Sigma_t = A\Sigma_t + \Sigma_tA^\top + \sigma\sigma^\top.$$

A stationary covariance Σ_∞ *must satisfy* $\frac{d}{dt}\Sigma_t = 0$*, hence*

$$A\Sigma_\infty + \Sigma_\infty A^\top + \sigma\sigma^\top = 0.$$

Therefore the stationary mean and covariance must satisfy

$$\boxed{Am_\infty = 0, \qquad A\Sigma_\infty + \Sigma_\infty A^\top + \sigma\sigma^\top = 0}$$

(the second equation is the continuous-time algebraic Lyapunov equation).

Solution 13.5. *Let* $\Delta = \{0 = t_0 < \cdots < t_n = T\}$ *and define the* backward *(right-endpoint) integral by the Riemann sums*

$$\int_0^T f(t,\omega) * dW_t := \lim_{|\Delta|\to 0}\sum_{j=0}^{n-1} f(t_{j+1},\omega)\,\Delta W_j, \qquad \Delta W_j := W_{t_{j+1}} - W_{t_j},$$

whenever the limit exists (e.g. in L^2*). Consider the backward SDE*

$$X_{t_{j+1}} - X_{t_j} = b(X_{t_j}, t_j)\,\Delta t_j + \sigma(X_{t_{j+1}}, t_{j+1})\,\Delta W_j, \qquad \Delta t_j := t_{j+1} - t_j.$$

Rewrite the stochastic term by expanding $\sigma(X_{t_{j+1}}, t_{j+1})$ *around* (X_{t_j}, t_j)*:*

$$\sigma(X_{t_{j+1}}, t_{j+1}) = \sigma(X_{t_j}, t_j) + \partial_x\sigma(X_{t_j}, t_j)\,(X_{t_{j+1}} - X_{t_j}) + o_p(|X_{t_{j+1}} - X_{t_j}|) + O(\Delta t_j),$$

so multiplying by ΔW_j and using $X_{t_{j+1}} - X_{t_j} = O_p(\sqrt{\Delta t_j})$ gives

$$\sigma(X_{t_{j+1}}, t_{j+1})\,\Delta W_j = \sigma(X_{t_j}, t_j)\,\Delta W_j + \partial_x \sigma(X_{t_j}, t_j)\,(X_{t_{j+1}} - X_{t_j})\Delta W_j + o_p(\Delta t_j).$$

Substitute $X_{t_{j+1}} - X_{t_j} = b(X_{t_j}, t_j)\Delta t_j + \sigma(X_{t_{j+1}}, t_{j+1})\Delta W_j$ into the product term; the drift part contributes $O_p(\Delta t_j^{3/2})$ and is $o_p(\Delta t_j)$, while the leading contribution is

$$(X_{t_{j+1}} - X_{t_j})\Delta W_j = \sigma(X_{t_{j+1}}, t_{j+1})(\Delta W_j)^2 + o_p(\Delta t_j) = \sigma(X_{t_j}, t_j)(\Delta W_j)^2 + o_p(\Delta t_j).$$

Hence

$$\sigma(X_{t_{j+1}}, t_{j+1})\,\Delta W_j = \sigma(X_{t_j}, t_j)\,\Delta W_j + \partial_x \sigma(X_{t_j}, t_j)\sigma(X_{t_j}, t_j)(\Delta W_j)^2 + o_p(\Delta t_j).$$

Plugging back into the increment equation,

$$X_{t_{j+1}} - X_{t_j} = b(X_{t_j}, t_j)\Delta t_j + \sigma(X_{t_j}, t_j)\Delta W_j + \partial_x \sigma(X_{t_j}, t_j)\sigma(X_{t_j}, t_j)(\Delta W_j)^2 + o_p(\Delta t_j).$$

Summing over j and letting $|\Delta| \to 0$, we use the quadratic variation $\sum_j (\Delta W_j)^2 \to T$ and, more generally,

$$\sum_j \partial_x \sigma(X_{t_j}, t_j)\sigma(X_{t_j}, t_j)(\Delta W_j)^2 \;\longrightarrow\; \int_0^T \partial_x \sigma(X_t, t)\sigma(X_t, t)\,dt$$

(in probability, or in L^1 under standard growth/Lipschitz conditions). The term $\sum_j o_p(\Delta t_j) \to 0$. Therefore the limiting continuous-time equation is

$$dX_t = \Big(b(X_t, t) + \partial_x \sigma(X_t, t)\,\sigma(X_t, t)\Big)\,dt + \sigma(X_t, t)\,dW_t,$$

i.e. the backward SDE $dX_t = b(X_t, t)\,dt + \sigma(X_t, t) * dW_t$ is equivalent to the Itô SDE with drift correction $\partial_x \sigma \cdot \sigma$.

14 Homework 14 (Dec 16th)

Solution 14.1. *Let $X_t \in \mathbb{R}^d$ solve a diffusion driven by an m-dimensional Brownian motion. Write $\sigma(x,t) \in \mathbb{R}^{d\times m}$ with entries $\sigma_{ik}(x,t)$, and set*

$$a(x,t) := \sigma(x,t)\sigma(x,t)^\top, \qquad a_{ij} = \sum_{k=1}^m \sigma_{ik}\sigma_{jk}.$$

Let $p(x,t)$ be the transition density of X_t (assume smooth and decaying so integration by parts is justified). For a test function $\varphi \in C_c^\infty$, define $\langle p_t, \varphi\rangle := \int_{\mathbb{R}^d} \varphi(x)p(x,t)\,dx$.

1) Stratonovich SDE $\Rightarrow$ PDE (2.11). *Assume the Stratonovich dynamics*

$$dX_t = b(X_t, t)\,dt + \sigma(X_t, t) \circ dW_t.$$

The Stratonovich generator acting on φ is (chain rule form)

$$L\varphi = b_i\,\partial_i\varphi + \frac{1}{2}\sum_{k=1}^m (\sigma_{\cdot k}\cdot\nabla)^2\varphi = b_i\,\partial_i\varphi + \frac{1}{2}\,\sigma_{ik}\,\partial_i\big(\sigma_{jk}\partial_j\varphi\big),$$

(using Einstein summation over repeated indices $i, j = 1, \dots, d$ and $k = 1, \dots, m$). By the Kolmogorov forward equation in weak form,

$$\frac{d}{dt}\langle p_t, \varphi\rangle = \langle p_t, L\varphi\rangle = \int p\,b_i\partial_i\varphi\,dx + \frac{1}{2}\int p\,\sigma_{ik}\partial_i(\sigma_{jk}\partial_j\varphi)\,dx.$$

Integrate by parts: first term

$$\int p\, b_i \partial_i \varphi\, dx = -\int \varphi\, \partial_i(b_i p)\, dx.$$

For the second term, integrate by parts twice:

$$\int p\, \sigma_{ik}\partial_i(\sigma_{jk}\partial_j\varphi)\, dx = -\int \sigma_{jk}\partial_j\varphi\, \partial_i(\sigma_{ik}p)\, dx = \int \varphi\, \partial_j\Big(\sigma_{jk}\, \partial_i(\sigma_{ik}p)\Big)\, dx.$$

Hence

$$\frac{d}{dt}\langle p_t, \varphi\rangle = \int \varphi\left[-\partial_i(b_i p) + \frac{1}{2}\, \partial_j\Big(\sigma_{jk}\, \partial_i(\sigma_{ik}p)\Big)\right] dx.$$

Since this holds for all φ, we obtain the PDE

$$\partial_t p + \partial_i(b_i p) = \frac{1}{2}\, \partial_j\Big(\sigma_{jk}\, \partial_i(\sigma_{ik}p)\Big).$$

In vector notation this is exactly

$$\boxed{\partial_t p + \nabla_x \cdot (bp) = \frac{1}{2}\, \nabla_x \cdot \big(\sigma\, \nabla_x \cdot (\sigma p)\big)}$$

with $(\nabla_x \cdot (\sigma\, \nabla_x \cdot (\sigma p)))_j = \partial_j(\sigma_{jk}\partial_i(\sigma_{ik}p))$, which is (2.11).

2) Backward (right-endpoint) integral $\Rightarrow$ PDE (2.12). *Assume the backward SDE*

$$dX_t = b(X_t, t)\, dt + \sigma(X_t, t) * dW_t.$$

From the earlier conversion (backward $\to$ Itô),

$$dX_t = \tilde{b}(X_t, t)\, dt + \sigma(X_t, t)\, dW_t, \qquad \tilde{b}_i = b_i + \partial_k \sigma_{ij}\, \sigma_{kj},$$

(where $\partial_k = \partial/\partial x_k$ and we sum over $j = 1, \ldots, m$). For the Itô SDE, the (Itô) generator is

$$L_{Ito}\varphi = \tilde{b}_i \partial_i \varphi + \frac{1}{2} a_{ij}\partial_{ij}\varphi, \qquad a_{ij} = \sigma_{ik}\sigma_{jk}.$$

The forward (Fokker–Planck) equation is the adjoint:

$$\partial_t p = -\partial_i(\tilde{b}_i p) + \frac{1}{2}\, \partial_{ij}(a_{ij} p).$$

Substituting $\tilde{b}_i = b_i + \partial_k \sigma_{ij}\sigma_{kj}$ and $a_{ij} = \sigma_{ik}\sigma_{jk}$ gives

$$\boxed{\partial_t p + \partial_i\Big[\big(b_i + \partial_k\sigma_{ij}\sigma_{kj}\big)p\Big] = \frac{1}{2}\, \partial_{ij}\big(\sigma_{ik}\sigma_{jk}\, p\big)}$$

which is exactly (2.12) (with $\partial_{ij} = \partial_i\partial_j$).

Remark 14.1. *Write the standard SDE and calculate the generator.*

Solution 14.2. *Consider the OU SDE on $\mathbb{R}^d$*

$$dX_t = BX_t\, dt + \sigma\, dW_t, \qquad a := \sigma\sigma^\top \in \mathbb{R}^{d\times d},$$

and assume the invariant density is Gaussian $\pi = \mathscr{N}(0, \Sigma)$ with $\Sigma \succ 0$.

Step 1: Fokker–Planck equation and probability current. *For an Itô diffusion* $dX_t = b(X_t)\,dt + \sigma\,dW_t$ *with constant* $a = \sigma\sigma^\top$*, the density* $p(x,t)$ *solves*

$$\partial_t p(x,t) = -\nabla\cdot\big(b(x)p(x,t)\big) + \frac{1}{2}\nabla\cdot\big(a\nabla p(x,t)\big) = -\nabla\cdot J(x,t),$$

where the probability current is

$$J(x,t) := b(x)p(x,t) - \frac{1}{2}a\nabla p(x,t).$$

Here $b(x) = Bx$*, so at stationarity* ($p = \pi$)*,*

$$J(x) := Bx\,\pi(x) - \frac{1}{2}a\nabla\pi(x).$$

Step 2: Detailed balance $\Leftrightarrow J \equiv 0$***.*** *Detailed balance (reversibility) means the net flux between any two states cancels; for diffusions this is equivalent to vanishing stationary current:*

$$\boxed{\text{Detailed balance} \iff J(x)\equiv 0.}$$

Step 3: Compute $\nabla\pi$ ***for*** $\pi = \mathcal{N}(0,\Sigma)$***.*** *Up to a constant,*

$$\pi(x) = \exp\left(-\frac{1}{2}x^\top\Sigma^{-1}x\right).$$

Hence

$$\nabla\log\pi(x) = -\Sigma^{-1}x, \qquad \nabla\pi(x) = \pi(x)\nabla\log\pi(x) = -\Sigma^{-1}x\,\pi(x).$$

Step 4: Convert $J \equiv 0$ ***into a matrix identity.*** *Substitute the gradient into the current:*

$$0 = J(x) = Bx\,\pi(x) - \frac{1}{2}a\big(-\Sigma^{-1}x\,\pi(x)\big) = \Big(B + \frac{1}{2}a\Sigma^{-1}\Big)x\,\pi(x).$$

Since this holds for all $x \in \mathbb{R}^d$ *and* $\pi(x) > 0$*, we obtain*

$$\boxed{B + \frac{1}{2}a\Sigma^{-1} = 0} \quad\iff\quad \boxed{B = -\frac{1}{2}a\Sigma^{-1}} \quad\iff\quad \boxed{a = -2B\Sigma}.$$

This is the detailed balance condition in matrix form.

15 Project 1

15.1 Introduction and Setting

The Potts model is a fundamental generalization of the Ising model in statistical mechanics, in which each lattice site can occupy one of q discrete states. It plays a key role in the study of phase transitions and critical phenomena, especially in two-dimensional systems.

In this project, we investigate the phase transition behavior of the two-dimensional q-state Potts model on a square lattice with periodic boundary conditions. The system is simulated using the Metropolis Monte Carlo algorithm, and various thermodynamic quantities are measured near the critical temperature.

The Hamiltonian of the Potts model is given by

$$H(\sigma) = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i,\sigma_j} - h \sum_i \sigma_i, \tag{2}$$

where $\sigma_i \in \{1, \dots, q\}$ denotes the state at lattice site i, $\delta_{\sigma_i,\sigma_j}$ is the Kronecker delta, J is the nearest-neighbor interaction strength, and h is an external magnetic field. The sum $\langle i,j \rangle$ runs over all nearest-neighbor pairs.

Throughout this study, we focus on the case $q = 3$ with the following parameter settings:

- Lattice size: $N \times N$ square lattice with $N = 100$,
- Number of states: $q = 3$,
- Interaction strength: $J = 1$,
- Boltzmann constant: $k_B = 1$,
- External field: $h = 0$ unless otherwise specified.

These choices allow us to systematically explore the thermodynamic behavior and phase transition properties of the two-dimensional three-state Potts model.

15.2 Question 1

Model 2D q-state Potts model on an $N \times N$ square lattice (PBC), coupling J. Each site $s_{ij} \in \{0, 1, \dots, q-1\}$.

- **Parameters:** $N = 128$, $q = 3$, $J = 1$.
- **Temperature scan:** $T \in [0.5, 1.5]$ (11 points), $\beta = 1/T$.
- **Initialization:** i.i.d. random states on the lattice.
- **Metropolis sweep:** repeat N^2 times: pick (i, j) uniformly, propose $s'_{ij} \neq s_{ij}$ uniformly from the other $q - 1$ states, accept with prob. $\min\{1, e^{-\beta \Delta E}\}$.
- **Energy:** $H = -J \sum_{\langle xy \rangle} \mathbf{1}[s_x = s_y]$ (count each nearest-neighbor bond once).
- **Thermalization:** $n_{\text{therm}} = 20000$ sweeps per T.
- **Sampling:** $n_{\text{sample}} = 50000$ sweeps per T; record H every $k = 10$ sweeps ($n_{\text{meas}} = n_{\text{sample}}/k$).
- **Observables (per site):**

$$u = \frac{\langle H \rangle}{N^2}, \qquad c = \frac{\beta^2 \operatorname{Var}(H)}{N^2}.$$

Results Figure 1 shows the internal energy per site u and the specific heat c of the 2D $q = 3$ Potts model as functions of temperature.

The internal energy decreases monotonically with decreasing temperature. At high temperatures the system is disordered, while at lower temperatures spin alignment becomes favorable, leading to a rapid drop in u. A noticeable change in slope occurs near $T \simeq 1$.

The specific heat exhibits a clear peak around $T \simeq 1$, indicating enhanced energy fluctuations and signaling a thermal phase transition. Due to the finite lattice size, the peak is smooth and finite.

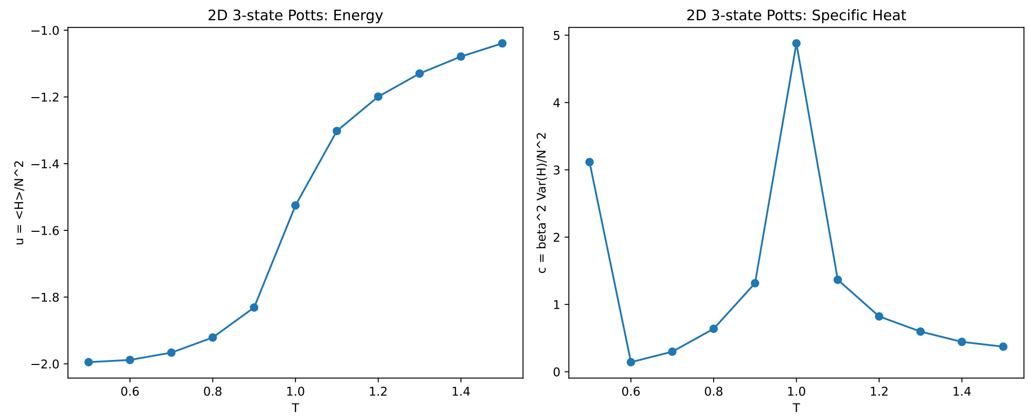


Figure 1: Internal energy (left) and specific heat (right) of the 2D $q = 3$ Potts model as functions of temperature.

15.3 Question 2

Model 2D q-state Potts model on an $N \times N$ square lattice (PBC), coupling J, with an external field h favoring a reference state s_{ref} (here $s_{\text{ref}} = 0$):

$$H = -J \sum_{\langle xy \rangle} \mathbf{1}[s_x = s_y] \; - \; h \sum_x \mathbf{1}[s_x = s_{\text{ref}}], \qquad s_x \in \{0, 1, \ldots, q-1\}.$$

- **Parameters:** $N = 128$, $q = 3$, $J = 1$, $s_{\text{ref}} = 0$.
- **Scan:** $T \in \{0.5, 1.0, 1.5\}$, $h \in [-2, 2]$ (21 points), $\beta = 1/T$.
- **Initialization:** i.i.d. random states.
- **Metropolis sweep:** repeat N^2 times: pick x uniformly, propose $s'_x \neq s_x$ uniformly from the other $q - 1$ states, accept with prob. $\min\{1, e^{-\beta \Delta E}\}$ (including interaction + field).
- **Thermalization:** $n_{\text{therm}} = 8000$ sweeps per (T, h).
- **Sampling:** $n_{\text{sample}} = 20000$ sweeps per (T, h); record every $k = 10$ sweeps.
- **Observable:** $\sigma_x = \mathbf{1}[s_x = s_{\text{ref}}]$, $M = \sum_x \sigma_x$, and

$$m = \frac{\langle M \rangle}{N^2} \in [0, 1] \qquad \text{(baseline } m = 1/q \text{ by symmetry at } h = 0\text{)}.$$

- **Continuation:** when scanning h, the final configuration at one h is used to initialize the next h (otherwise re-randomize at each h).

Results Figure 2 shows the magnetization $m = \langle M \rangle / N^2$ as a function of the external field h for different temperatures.

At $h = 0$, the magnetization is close to the symmetric value $m = 1/q$, indicating no preference among Potts states. As h increases, the field favors the reference state and m increases monotonically. This response becomes sharper at lower temperatures: for $T = 0.5$ the magnetization rises abruptly, while for higher T the curve is smoother due to stronger thermal fluctuations.

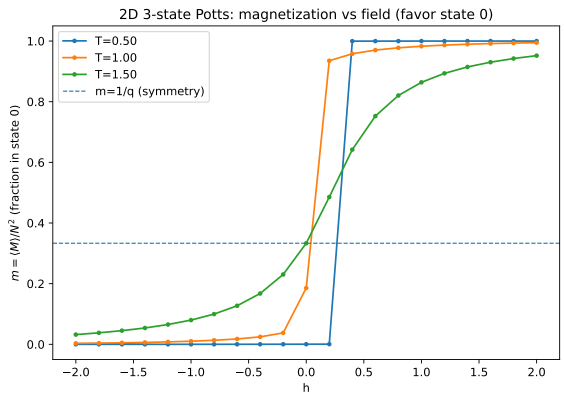


Figure 2: Magnetization $m = \langle M \rangle / N^2$ versus external field h for the 2D $q = 3$ Potts model at different temperatures. The dashed line indicates the symmetric value $m = 1/q$.

15.4 Question 3

Model 2D q-state Potts model on an $N \times N$ lattice (PBC), $q = 3$, $J = 1$, zero field:

$$H = -J \sum_{\langle xy \rangle} \mathbf{1}[s_x = s_y], \qquad s_x \in \{0, 1, 2\}.$$

- **Scan:** $T \in [0.5, 1.5]$ (11 points), $\beta = 1/T$.
- **Update:** Metropolis sweeps (each sweep N^2 proposals).
- **Thermalization:** $n_{\text{therm}} = 8000$ sweeps per T.
- **Sampling:** $n_{\text{sample}} = 20000$ sweeps; measure every $k_{\text{int}} = 20$ sweeps.
- **Indicator:** $\sigma_x = \mathbf{1}[s_x = s_{\text{ref}}]$ with $s_{\text{ref}} = 0$, $m = \langle \sigma \rangle$.
- **Correlation estimator:** for $k = 1, \ldots, N/2$,

$$\Gamma(k) = \frac{1}{4N^2} \sum_x \sum_{y \in S_x(k)} \left(\sigma_x \sigma_y - m^2 \right), \quad S_x(k) = \{x \pm k\hat{e}_x,\ x \pm k\hat{e}_y\}.$$

- **Fit:** estimate ξ from $\Gamma(k) \approx \Gamma_0 e^{-k/\xi}$ by a linear fit of $\log \Gamma(k)$ over $k \in [k_{\min}, k_{\max}]$ with $k_{\min} = 4$, $k_{\max} = N/2$.

Results Figure 3 shows the correlation length ξ as a function of temperature for the 2D $q = 3$ Potts model at zero field.

The correlation length increases rapidly as the temperature approaches $T \simeq 1$ from both sides, indicating the development of long-range correlations near the critical point. Away from this region, ξ remains finite and small, corresponding to short-range order in both the high- and low-temperature phases. The finite peak of ξ is due to the finite lattice size.

15.5 Question 4

Data Specific heat $c(T)$ and correlation length $\xi(T)$ obtained from previous simulations of the 2D $q = 3$ Potts model.

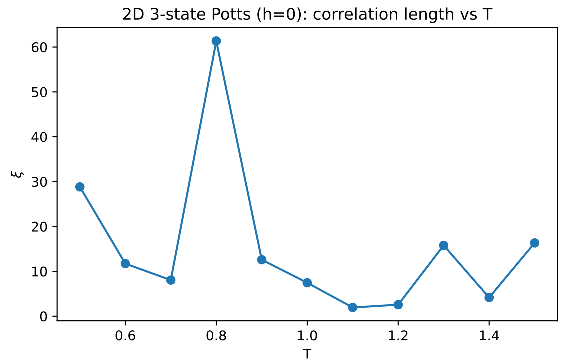


Figure 3: Correlation length ξ as a function of temperature $T$for the 2D $q = 3$ Potts model at zero external field.

- **Critical temperature:** fixed to $T^* = 1$.
- **Reduced temperature:** $\epsilon = |1 - T/T^*|$.
- **Selection:** use all valid data points with $\epsilon > 0$ and finite c, ξ (no scaling window imposed).
- **Scaling ansatz:**
$$c(\epsilon) \sim \epsilon^{-\gamma}, \qquad \xi(\epsilon) \sim \epsilon^{-\delta}.$$
- **Fit:** linear regression on log–log data, $\log y = a + b \log \epsilon$, with exponents $\gamma = -b_c$ and $\delta = -b_\xi$.

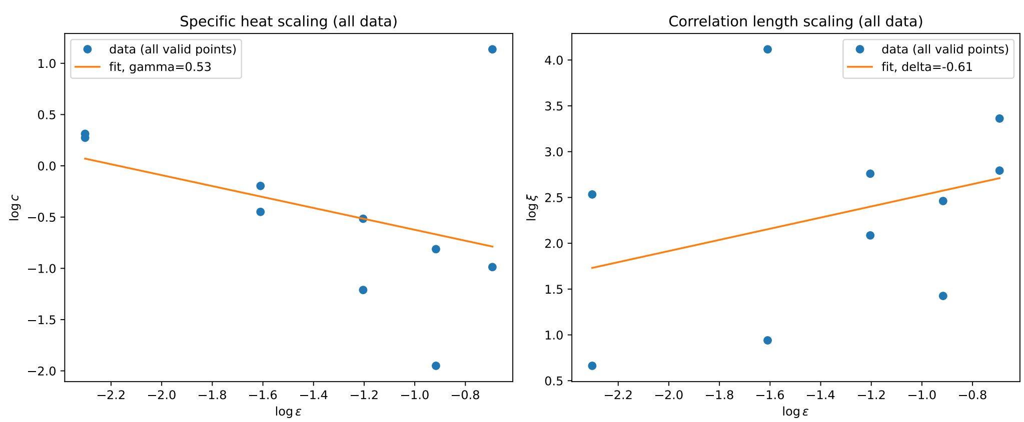


Figure 4: Log–log scaling of the specific heat c (left) and correlation length ξ (right) as functions of the reduced temperature $\epsilon = |1 - T/T^*|$ with $T^* = 1$. Solid lines indicate linear fits using all valid data points.

Results Figure 4 shows the log–log fits of the specific heat c and the correlation length ξ as functions of the reduced temperature $\epsilon = |1 - T/T^*|$ with $T^* = 1$.

Using all valid data points, both observables exhibit approximate power-law behavior. Linear regression yields the effective exponents $\gamma \simeq 0.53$ for the specific heat and $\delta \simeq 0.61$ for the correlation length. Deviations from ideal scaling are expected due to finite-size effects and the absence of a restricted asymptotic scaling window.

16 Project 2

16.1 Introduction

This project investigates the numerical simulation of exit times for a two-dimensional stochastic differential equation (SDE) driven by a nontrivial potential landscape. The SDE is given by

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\varepsilon}\,dW_t,$$

where $X_t = (X_t, Y_t) \in \mathbb{R}^2$, $\varepsilon > 0$ denotes the noise intensity, and W_t is a two-dimensional standard Wiener process.

The potential function $V(x, y)$ is constructed from a symmetric mixture of two Gaussian distributions centered at $(\pm 1, 0)$:

$$p^+(x, y) = \mathscr{N}\big((1, 0), I_2\big), \qquad p^-(x, y) = \mathscr{N}\big((-1, 0), I_2\big),$$

$$p(x, y) = \frac{1}{2}\big(p^+(x, y) + p^-(x, y)\big), \qquad V(x, y) = -\log p(x, y),$$

where I_2 denotes the 2×2 identity matrix. This construction yields a double-well potential with two metastable regions separated by a barrier near $x = 0$.

The main quantity of interest is the expected exit time

$$T(\varepsilon, x_0) = \mathbb{E}_{x_0}\big[\tau_b^\varepsilon\big], \qquad \tau_b^\varepsilon = \inf\{t \geq 0 :\ X_t = 0\},$$

defined as the first time the process reaches the boundary $x = 0$.

The objectives of this project are to:

- compute $T(\varepsilon, x_0)$ numerically for a fixed initial condition $x_0 = (1, 0)$;
- investigate the dependence of the exit time on the noise intensity ε;
- explore the influence of different initial positions x_0 on the exit behavior.

16.2 Parameter Settings

Unless otherwise specified, the numerical experiments are performed with the following parameter choices:

- **Time step:** $\Delta t = 10^{-3}$.
- **Noise intensity:** ε is varied over a range from relatively large to small values (e.g., $\varepsilon \in \{1.0, 0.6, 0.4, 0.3, 0.2, 0.15, 0.1\}$) in order to investigate its effect on the exit time.
- **Initial condition:** The default initial position is $x_0 = (1, 0)$. Additional experiments are conducted with different initial points to study the dependence of the exit time on x_0.
- **Stopping criterion:** The exit time τ_b^ε is defined as the first time the first coordinate reaches zero, i.e.,

  $$\tau_b^\varepsilon = \inf\{t \geq 0 : X_t = 0\}.$$

  In the numerical implementation, the exit event is detected when the discrete trajectory crosses $x = 0$, with linear interpolation used to improve accuracy.
- **Maximum simulation time:** A cutoff time $T_{\max}$ is imposed to avoid excessively long simulations. Typically, $T_{\max} = 200$ for moderate values of ε, and increased up to $T_{\max} = 400$ for smaller ε.
- **Monte Carlo sample size:** For each parameter configuration, $N = 2000$ independent trajectories are simulated to estimate the expected exit time.

16.3 Results

16.3.1 Effect of noise level ε

For $x_0 = (1, 0)$, the Monte Carlo estimate of the exit time $T(\varepsilon, x_0)$ increases monotonically as ε decreases. Specifically, $T(\varepsilon, x_0)$ grows from 1.67 at $\varepsilon = 1.0$ to 7.84 at $\varepsilon = 0.06$ (all standard errors below 0.14). This trend is also visible in Figure 5. [**<empty citation>**]

To examine scaling, we plot $\log \widehat{T}(\varepsilon, x_0)$ against $1/\varepsilon$ (Figure 6). The relationship is approximately linear, suggesting an exponential law of the form

$$T(\varepsilon, x_0) \approx C \exp\left(\frac{\Delta V}{\varepsilon}\right),$$

consistent with metastable barrier-crossing dynamics. See Figure 6. [**<empty citation>**]

16.3.2 Effect of initial condition x_0

Fixing $\varepsilon = 0.2$, we vary x_0 and observe that the exit time mainly depends on the x-coordinate. Starting closer to the boundary reduces the exit time: $\widehat{T}(0.2, (0.8, 0)) = 3.55$ versus $\widehat{T}(0.2, (1.2, 0)) = 4.34$. In contrast, changing the transverse coordinate has negligible influence in our experiments: $\widehat{T}(0.2, (1, 0)) \approx \widehat{T}(0.2, (1, 0.5)) \approx \widehat{T}(0.2, (1, 1)) \approx 3.96$.

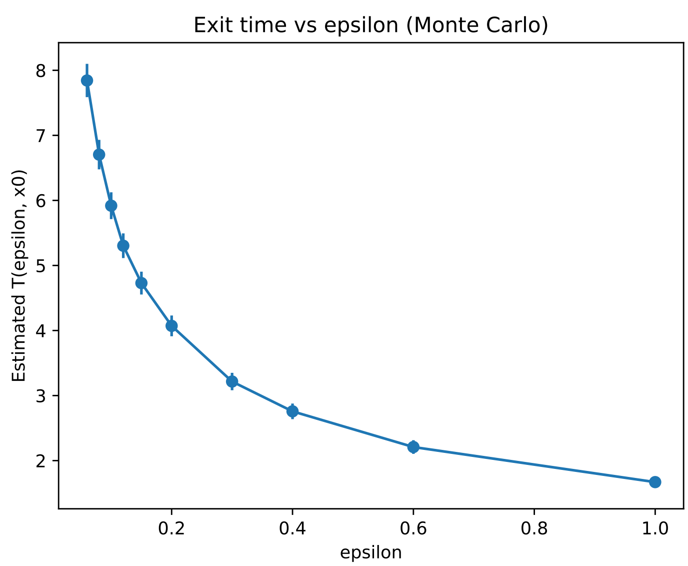


Figure 5: Estimated exit time $T(\varepsilon, x_0)$ versus ε with $x_0 = (1, 0)$.

Overall, the exit time increases rapidly as ε decreases, and $\log T(\varepsilon, x_0)$ is approximately linear in $1/\varepsilon$, indicating metastable barrier-crossing behavior.

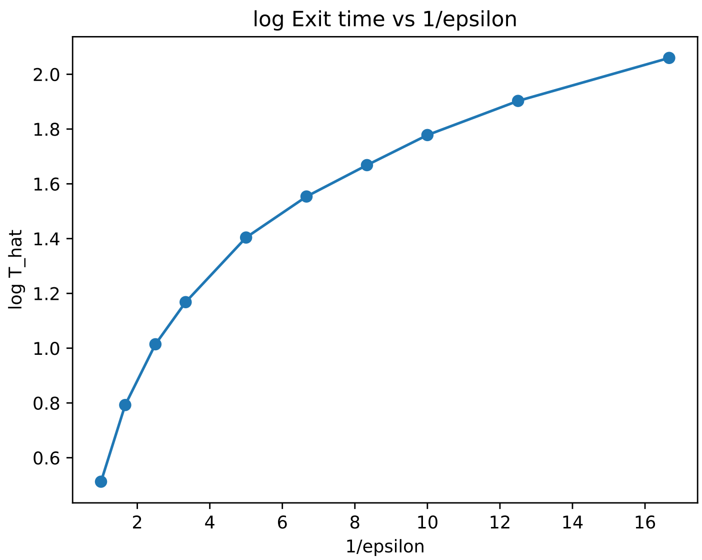


Figure 6: $\log \widehat{T}(\varepsilon, x_0)$ versus $1/\varepsilon$ with $x_0 = (1, 0)$. The near-linearity indicates $T(\varepsilon, x_0)$ grows approximately exponentially as $\varepsilon \to 0$.