Section Notes for PSTAT 213C

Haosheng Zhou

Apr, 2025

Contents

Week 1

Review: Conditional Expectation

Exercise 1. *Let $\mathscr{F}_2 \subset \mathscr{F}_1 \subset \mathscr{F}$ be σ-fields on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$. For $X \in L^1$, show that*

$$\mathbb{E}(X - \mathbb{E}(X|\mathscr{F}_1))^2 \leq \mathbb{E}(X - \mathbb{E}(X|\mathscr{F}_2))^2. \tag{1}$$

If $X \in L^2$, what is the geometric interpretation?

Hints. Expand squares, it suffices to show that $\mathbb{E}[\mathbb{E}(X|\mathscr{F}_1)]^2 \geq \mathbb{E}[\mathbb{E}(X|\mathscr{F}_2)]^2$. Apply Jensen's inequality for the r.v. $\mathbb{E}(X|\mathscr{F}_1)$ under the conditional expectation $\mathbb{E}(\cdot|\mathscr{F}_2)$ yields $\mathbb{E}([\mathbb{E}(X|\mathscr{F}_1)]^2|\mathscr{F}_2) \geq [\mathbb{E}(\mathbb{E}(X|\mathscr{F}_1)|\mathscr{F}_2)]^2$. Taking expectations on both sides and using the tower property conclude the proof.

Geometrically, the distance from a fixed point to a vector space V (e.g., a plane) is no larger than the distance from the same point to a linear subspace of V (e.g., any line within the plane). $\square$

Exercise 2. *Define the conditional independence $\mathscr{F}_1 \perp \mathscr{F}_2|\mathscr{G}$ to hold if and only if*

$$\mathbb{E}(X_1X_2|\mathscr{G}) = \mathbb{E}(X_1|\mathscr{G}) \cdot \mathbb{E}(X_2|\mathscr{G}),\ \forall X_1 \in \mathscr{F}_1, X_2 \in \mathscr{F}_2. \tag{2}$$

Show that:

(i): $\mathscr{F}_1 \perp \mathscr{F}_2|\mathscr{G}$ if and only if $(\mathscr{F}_1 \vee \mathscr{G}) \perp \mathscr{F}_2|\mathscr{G}$, where $\mathscr{F}_1 \vee \mathscr{G} := \sigma(\mathscr{F}_1 \cup \mathscr{G})$.

(ii): If $\mathscr{G} \subset \mathscr{F}_1$, then $\mathscr{F}_1 \perp \mathscr{F}_2|\mathscr{G}$ if and only if $\mathbb{E}(X_2|\mathscr{F}_1) \in \mathscr{G}$ for $\forall X_2 \in \mathscr{F}_2$.

Hints. (i): Direction $\Leftarrow$ is trivial. Direction $\Rightarrow$ requires $\pi - \lambda$ theorem (first prove the statement for $X_1\mathbb{I}_G$, where $X_1 \in \mathscr{F}_1$ and $G \in \mathscr{G}$).

(ii): Direction $\Leftarrow$ is trivial (tower property). Direction $\Rightarrow$ requires proving $\mathbb{E}(X_2|\mathscr{F}_1) = \mathbb{E}(X_2|\mathscr{G})$, using the definition of conditional expectation. $\square$

Martingale

Exercise 3. *$\{X_n\}$ is a discrete-state Markov chain (with countable state space S) and transition probability matrix P. If there exists a bounded function $\psi : S \to \mathbb{R}$ such that*

$$\sum_{j\in S} P_{ij}\psi(j) = \psi(i),\ \forall i \in S, \tag{3}$$

check that $\{X_n\}$ is a martingale under the natural filtration.

A discrete-state Markov chain induces a difference operator L such that

$$(L\psi)(i) := \sum_{j\in S} P_{ij}\psi(j) - \psi(i),\ \forall i \in S. \tag{4}$$

What about a continuous-state Markov chain (e.g., Brownian motion)? That induces a differential operator on $\mathbb{R}^d$, which is one half the Laplacian $\Delta := \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$. With time evolution of the Markov chain introduced, this provides crucial connection between probability and PDE (generator, forward/backward equations).

Exercise 4. *$\{X_n\}$ is a sequence of i.i.d. r.v. with an unknown density f. The density f is known to be either p or q (both are known and strictly positive), which results in a likelihood ratio test*

$$\begin{cases} H_0 : f = q \\ H_1 : f = p \end{cases}, \tag{5}$$

with a test statistic

$$Y_n := \prod_{i=1}^n \frac{p(X_i)}{q(X_i)}. \tag{6}$$

Show that $\{Y_n\}$ is a martingale under the natural filtration of $\{X_n\}$ if H_0 is true.

The rejection region of this test: $Y_n \geq a$. The simplest example of sequential hypothesis testing (continue collecting evidence until a conclusion can be drawn).

Does $\{Y_n\}$ admit a limit? If one terminates the test only when a rejection is made, can we control the probability of rejecting H_0 if H_0 is true (type-I error)? Will be answered next week.

Exercise 5. *A village contains $N+1$ people, one of whom suffers from an infectious illness that cannot be cured. S_t is the number of susceptibles at time t, I_t is the number of infectives and D_t is the number of deads such that $S_t + I_t + D_t = N+1$, while $S_0 = N, I_0 = 1, D_0 = 0$.*

Once a person is infected, his remaining lifespan is a random time that follows $\mathcal{E}(\mu)$. Once a susceptible interacts with an infective, he gets infected after a random time that follows $\mathcal{E}(\lambda)$. It is assumed that at any fixed time, any two individuals within the system interact, and the random times are all independent.

(i): Specify the state space S_X and the dynamics of the continuous-time discrete-state Markov chain (S_t, I_t).

(ii): Specify the generator G of (S_t, I_t).

(iii): If $G\psi = 0$ for some $\psi : S_X \to \mathbb{R}$, show that $Y_t := \psi(S_t, I_t)$ is a martingale under the natural filtration of $\{(S_t, I_t)\}$.

(iv): Find one such $\psi(s,i) = \alpha(s)\beta(i)$ that reveals the martingale structure of the epidemic model.

This epidemic model must result in one of the two situations: either everyone dies due to illness or the illness dies out itself before infecting everybody. How to calculate the probability of those situations happening? Will also be answered next week.

Hints. (i): $S_X = \{(s,i) : s + i \leq N+1, s \geq 0, i \geq 0\}$. Given $(S_t, I_t) = (s,i)$, the next state transition is either a death $(s,i) \to (s, i-1)$ or an infection $(s,i) \to (s-1, i+1)$.

The time until the next death is the minimum of the lifespan of each infective, which is the minimum of i independent $\mathcal{E}(\mu)$ r.v. Consequently, death rate $DR_{(s,i)} = i\mu$.

The time until the next infection is the minimum of is independent $\mathcal{E}(\lambda)$ r.v. (since each of the s susceptible interacts with all i infectives). Consequently, infection rate $IR_{(s,i)} = is\lambda$.

Therefore, the holding rate of state (s,i) is $i(\mu+s\lambda)$. Whenever a state transition happens, there is $\frac{i\mu}{i\mu+is\lambda}$ probability transiting to $(s,i-1)$, there is $\frac{is\lambda}{i\mu+is\lambda}$ probability transiting to $(s-1,i+1)$.

(ii): By definition,

$$G_{(s,i),(s,i-1)} = i\mu, \quad G_{(s,i),(s-1,i+1)} = is\lambda, \quad G_{(s,i),(s,i)} = -i(\mu+s\lambda), \tag{7}$$

while all other entries are zero.

(iii): Prove the martingale property within an infinitesimal interval $[t, t+\Delta t]$.

(iv): Assume $\alpha(N) = 1$ and solve $G\psi = 0$ to get

$$\psi(s,i) = \prod_{k=s+1}^{N} \frac{kB\lambda - (1-B)\mu}{k\lambda B^2} B^i, \tag{8}$$

which is a valid solution for an arbitrary $B > 0$. □

Remark. *One obtains the freedom of choosing $B > 0$ in the example above that ensures*

$$Y_t = \prod_{k=S_t+1}^{N} \frac{kB\lambda - (1-B)\mu}{k\lambda B^2} B^{I_t} \tag{9}$$

being a martingale.

As we shall see later, martingales have very nice structures and are the easiest stochastic processes to investigate. Unfortunately, processes of general interests (e.g., branching processes, diffusions, asymmetric random walks) are typically not martingales. This example shows how one should discover the hidden martingale structure to be able to apply technical tools that are specifically designed for martingales (e.g., optional stopping theorem, maximal inequalities, etc.). ***The martingale structure is crucial but never for free!***

Week 2

Concentration Inequality

Theorem (Azuma-Hoeffding)**.** *$\{Y_n\}$ is a martingale under filtration $\{\mathscr{F}_n\}$ and there exists a sequence of real numbers $\{K_n\}$ such that $|Y_n - Y_{n-1}| \leq K_n$ a.s. for $\forall n$. Then*

$$\mathbb{P}(|Y_n - Y_0| \geq x) \leq 2e^{-\frac{1}{2}\frac{x^2}{\sum_{i=1}^n K_i^2}}, \ \forall x > 0. \tag{10}$$

Exercise 6. *Given n objects with independent identically distributed random sizes $X_1, ..., X_n$ on $[0,1]$, let B_n be the minimum number of bins (of size 1) required to pack $X_1, ..., X_n$. Show that*

$$\mathbb{P}(|B_n - \mathbb{E}B_n| \geq x) \leq 2e^{-\frac{1}{2}\frac{x^2}{n}}, \ \forall x > 0. \tag{11}$$

Hints. Set $\mathscr{F}_i := \sigma(X_1, ..., X_i)$ and $Y_i := \mathbb{E}(B_n|\mathscr{F}_i)$ as a martingale. Prove $|Y_{i+1} - Y_i| \leq 1$ by considering the minimum number of bins (of size 1) required to pack $X_1, ..., X_n$ without packing X_i (leave-one-out). □

Exercise 7. *Let $P_i := (U_i, V_i)$ such that $P_1, ..., P_n$ are independent and uniformly distributed points in $[0,1]^2$. Let D_n denote the length of the shortest tour that passes through each point exactly once and returns to the starting point. Show that there exists a constant $A > 0$, such that*

$$\mathbb{P}(|D_n - \mathbb{E}D_n| \geq x) \leq 2e^{-\frac{Ax^2}{\log n}}, \ \forall x > 0. \tag{12}$$

Hints. Set $\mathscr{F}_i := \sigma(P_1, ..., P_i)$ and $Y_i := \mathbb{E}(D_n|\mathscr{F}_i)$ as a martingale.

Consider the leave-P_i-out shortest path that has length $D_n(i)$. Note $D_n(i) \leq D_n \leq D_n(i) + 2Z_i$ for $i \leq n-1$, where Z_i is the shortest distance from P_i to one of the points in $\{P_{i+1}, ..., P_n\}$ (why?). We derive $|Y_i - Y_{i-1}| \leq 2\max\{\mathbb{E}(Z_i|\mathscr{F}_i), \mathbb{E}(Z_i|\mathscr{F}_{i-1})\}$. Now prove that $\max\{\mathbb{E}(Z_i|\mathscr{F}_i), \mathbb{E}(Z_i|\mathscr{F}_{i-1})\} \leq \frac{C}{\sqrt{n-i}}$ for some constant $C > 0$. □

Week 3

Application of OST

Exercise 8 (Doob's Maximal Inequality)**.** *Let $\{X_n\}$ be a sub-MG, prove that*

$$\mathbb{P}\left(\max_{1\le m\le n} X_m \ge \lambda\right) \le \frac{\mathbb{E}X_n^+ \mathbb{I}_{\max_{1\le m\le n} X_m \ge \lambda}}{\lambda} \le \frac{\mathbb{E}X_n^+}{\lambda},\ \forall \lambda > 0. \tag{13}$$

Hints. $T := \inf\{n : X_n \ge \lambda\}$, check that $\mathbb{P}(T \le n) \le \frac{\mathbb{E}X_T^+ \mathbb{I}_{T\le n}}{\lambda} = \frac{\mathbb{E}X_{T\wedge n}^+ - \mathbb{E}X_n^+ \mathbb{I}_{T>n}}{\lambda}$. By OST, $\mathbb{E}X_{T\wedge n}^+ \le \mathbb{E}X_n^+$. □

Exercise 9. *If $\{X_n\}$ is a simple symmetric random walk, prove that: (i). $\{X_n\}$ is a MG (ii). $\{X_n^2 - n\}$ is a MG (quadratic-MG) (iii). $\{\frac{e^{\lambda X_n}}{(\cosh\lambda)^n}\}$ is a MG for $\forall\lambda > 0$, where $\cosh\lambda := \frac{e^\lambda + e^{-\lambda}}{2}$ (exponential-MG).*

Let T be the first exit time of (a, b), where $a < 0, b > 0$ for $a, b \in \mathbb{Z}$, i.e., $T = \inf\{n : X_n = a$ or $X_n = b\}$. Using the MGs above, calculate $\mathbb{P}(X_T = a)$ and $\mathbb{E}T$.

Let T_1 be the first hitting time to 1, find the distribution of T_1.

Bonus: Can you find the distribution of T?

Hints. By OST, $\mathbb{P}(X_T = a) = \frac{b}{b-a}$ and $\mathbb{E}T = -ab$. PGF of T_1 is $\mathbb{E}s^{T_1} = \frac{1-\sqrt{1-s^2}}{s}$.

To find the distribution of T, consider the MG $Y_n = \frac{\cos\left(\lambda(X_n - \frac{b+a}{2})\right)}{\cos^n\lambda}$ (subtract the midpoint $\frac{b+a}{2}$ to symmetrize the exit time). By OST, $\mathbb{E}(\cos\lambda)^{-T} = \frac{\cos(\lambda\frac{b+a}{2})}{\cos(\lambda\frac{b-a}{2})}$. To verify the validity of OST, apply DCT and check $\mathbb{E}(\cos\lambda)^{-T} < \infty$ (prove by Fatou's lemma). □

Exercise 10. *$\{X_n\}$ is a simple asymmetric random walk with probability p stepping upward and probability $q = 1-p$ stepping downward. Let T be the first exit time of (a, b), where $a < 0, b > 0$ for $a, b \in \mathbb{Z}$, i.e., $T = \inf\{n : X_n = a$ or $X_n = b\}$. Calculate $\mathbb{P}(X_T = a)$ and $\mathbb{E}T$.*

Hints. $Y_n = \left(\frac{q}{p}\right)^{X_n}$ is a MG. $Z_n = X_n - (p-q)n$ is a MG. □

Exercise 11. *Consider the sequential hypothesis testing model in Example 4, show that $\{Y_n\}$ converges and identify the limit. If one terminates the test iff the first rejection is made, provide an upper bound for the probability of type-I error.*

Hints. The limit is almost surely 1 if $p = q$, otherwise it's 0 (Jensen).

Doob's maximal inequality: $\mathbb{P}(\sup_n Y_n \ge a | H_0) \le \frac{1}{a}$. □

Exercise 12. *Consider the epidemic model in Example 5, calculate the probability that eventually there are still people alive, i.e., the illness kills itself before infecting everyone.*

Hints. $T := \inf\{t : I_t = 0\}$. Use OST and find a clever way to specify B: consider B_r such that $rB_r\lambda - \mu(1-B_r) = 0$. Plug into the OST for $\forall 1 \le r \le N$ to get a system of equations. □

Week 4

Practice Problems for Midterm Revision

Exercise 13. *Let $f:\mathbb{Z}^2\to\mathbb{R}_+$ be any non-negative function such that*

$$f(x,y)=\frac{f(x-1,y)+f(x+1,y)+f(x,y-1)+f(x,y+1)}{4},\ \forall (x,y)\in\mathbb{Z}^2. \tag{14}$$

Show that f must be a constant function.

Hints. Construct a martingale and apply martingale convergence theorem. Use recurrence/transience properties of Markov chain if necessary. □

Exercise 14. *Let $\{X_n\}$ be a symmetric simple random walk starting at 0 and T be the first exit time of $(-a,a)$ for a positive integer a. Compute $\mathbb{E}T^2$.*

Hints. Recall the computation of $\mathbb{E}T$ where we use the square martingale $\{X_n^2-n\}$ because the n term would become T when applying OST. Therefore, we hope to construct a martingale that looks like $f(X_n)+bn^2+cn$ for some function f. Try the function $f(S_n)=S_n^4-6nS_n^2$ and figure out the values of b,c such that the martingale property holds. Apply OST to yield $\mathbb{E}T^2=\frac{5a^4-2a^2}{3}$. □

Exercise 15. *Let $\{Z_n\}$ be a branching process with $Z_0=1$. The offspring distribution has mean μ and variance $\sigma^2>0$. We wish to investigate the maximum population $\sup_n Z_n$ that has ever appeared in the history when the branching process is in its subcritical phase, i.e., $\mu<1$. For simplicity, we also assume that each individual has positive probability of giving birth to more than two children, i.e., $\mathbb{P}(Z_1\ge 2)>0$.*

(i): Let G denote the probability generating function of Z_1. Let η be the largest root of the equation $x=G(x)$. Show that $\eta>1$.

(ii): Show that under the filtration $\mathscr{F}_n:=\sigma(Z_0,...,Z_n)$, $Y_n:=\eta^{Z_n}$ is a martingale.

(iii): Show that $\mathbb{E}\sup_n Z_n\le\frac{\eta}{\eta-1}$.

Hints. For (iii), use the tail formula for expectation and apply Doob's maximal inequality. □

Exercise 16. *Let $\{X_n\}$ be an L^2 martingale, i.e., $X_n\in L^2$, $\forall n$ under some filtration $\{\mathscr{F}_n\}$ such that $X_0=0$. Clearly, $\{X_n^2\}$ is a sub-martingale and admits a unique Doob's decomposition*

$$X_n^2=M_n+A_n, \tag{15}$$

where $\{M_n\}$ is a martingale and $\{A_n\}$ is a predictable increasing process with $A_0=0$.

(i): Show that $\mathbb{E}\sup_n|X_n|^2\le 4\mathbb{E}A_\infty$.

(ii): Fix any $a>0$ and consider $T_a:=\inf\{n:A_{n+1}>a^2\}$. Show that T_a is a stopping time w.r.t. $\{\mathscr{F}_n\}$.

(iii): Show that

$$\mathbb{P}\left(\sup_n|X_n|>a\right)\le\mathbb{P}\left(A_\infty>a^2\right)+\mathbb{P}\left(\sup_n|X_{n\wedge T_a}|>a\right), \tag{16}$$

where

$$\mathbb{P}\left(\sup_n |X_{n\wedge T_a}| > a\right) \le \frac{\mathbb{E}(A_\infty \wedge a^2)}{a^2}. \tag{17}$$

(iv): Use (iv) to show that $\mathbb{E}\sup_n |X_n| \le 3\mathbb{E}\sqrt{A_\infty}$.

(v): Let T *be the first hitting time to* 1 *of a simple symmetric random walk that starts from* 0. *Prove that* $\mathbb{E}\sqrt{T} = \infty$.

Hints. For (i), apply Doob's L^2 inequality and use MCT.

For (iii), discuss if $\{T_a = \infty\}$ happens and use Doob's maximal inequality.

For (iv), use tail formula for expectation and apply Fubini's theorem.

For (v), consider Doob's decomposition of $\{X_{n\wedge T}\}$ and reach a contradiction by DCT. $\square$

Remark. *From optional stopping theorem, we prove that* $\mathbb{E}T = \infty$, *while the exercise above proves a stronger conclusion* $\mathbb{E}\sqrt{T} = \infty$. *Actually, the continuous limit of this stopping time is the first hitting time to* 1 *of Brownian motion, which follows an* α*-table law with* $\alpha = \frac{1}{2}$. *This agrees with the current observation* $\mathbb{E}\sqrt{T} = \infty$.