EECS 126 Probability and Random Processes
Kannan Ramchandran

UC Berkeley, Fall 2022
September 29, 2022

Midterm 1

Last Name	First Name	SID

Left Neighbor First and Last Name	**Right Neighbor** First and Last Name

Rules.

- **Write in your SID on every page to receive 1 point.**
- **Unless otherwise stated, all your answers need to be justified and your work must be shown. Answers without sufficient justification will get no credit.**
- You have 10 minutes to read the exam and 70 minutes to complete the exam. (DSP students with $X\,\%$ time accommodation should spend $10 \cdot X\,\%$ time on reading and $70 \cdot X\,\%$ time on completing the exam).
- This exam is closed-book. You may reference one double-sided handwritten sheet of paper. No calculator or phones allowed.
- Collaboration with others is strictly prohibited. If you are caught cheating, you may fail the course and face disciplinary consequences.

Problem	out of
SID	1
Problem 1	41
Problem 2	16
Problem 3	22
Problem 4	24
Problem 5	22
Total	126

1 Potpourri [41 points]

(a) **Beat Stanford [9 points]**

Cal and Stanford are playing a best-of-three basketball match. The games are played at home, away and home respectively from the perspective of Cal. We know that Cal has a $\frac{2}{3}$ probability of winning at our home court and a $\frac{1}{2}$ probability of winning at an away court, independent of all other games. Given that Cal wins the best-of-three, what is the probability that all three games are needed to decide the champion? (In a best-of-three, the teams play a total of at most three games, and whoever wins two games is the champion. If a team wins both of the first two games, there is no need to play the third game.)

There are 3 ways Cal can win the best-of-three, namely CC, CSC and SCC. Since they are disjoint events,

$$\mathrm{P(Cal\ wins)} = \mathrm{P(CC)} + \mathrm{P(CSC)} + \mathrm{P(SCC)}.$$

Of the 3 outcomes, there are 2 that correspond to all 3 games being played, so

$$\mathrm{P(Need\ 3\ games} \cap \mathrm{Cal\ wins)} = \mathrm{P(CSC)} + \mathrm{P(SCC)}.$$

We can calculate the probability for each of the 3 outcomes,

$$\mathrm{P(CC)} = \frac{2}{3}\cdot\frac{1}{2} = \frac{1}{3}, \qquad \mathrm{P(CSC)} = \frac{2}{3}\cdot\frac{1}{2}\cdot\frac{2}{3} = \frac{2}{9}, \qquad \mathrm{P(SCC)} = \frac{1}{3}\cdot\frac{1}{2}\cdot\frac{2}{3} = \frac{1}{9}.$$

By Bayes' theorem,

$$\mathrm{P(Need\ 3\ games|Cal\ wins)} = \frac{\mathrm{P(Need\ 3\ games} \cap \mathrm{Cal\ wins)}}{\mathrm{P(Cal\ wins)}} = \frac{\frac{2}{9}+\frac{1}{9}}{\frac{1}{3}+\frac{2}{9}+\frac{1}{9}} = \frac{1}{2}.$$

(b) **Working Monkeys [7 points]**

Sohom and his army of monkeys are working on writing a new edition of the EECS 126 textbook. At each timestep from $i = 0, 1, 2, \ldots$, Sohom writes down a random character X_i drawn uniformly at random from the alphabet $\{\text{`}a\text{'}, \text{`}b\text{'}, \text{`}c\text{'}, \ldots, \text{`}z\text{'}\}$, independent of all other events. There are 26 letters in the English alphabet. Show that with probability 1, Sohom writes down an 'a' eventually. In other words, if A_i is the event $\{X_i = \text{`}a\text{'}\}$, show that $\lim_{n\to\infty} \mathrm{P}(\bigcup_{i=1}^{n} A_i) = 1$.

We can compute

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \mathrm{P}\left(\bigcap_{i=1}^{n} A_i^{\mathsf{c}}\right) = 1 - \prod_{i=1}^{n} \mathrm{P}(A_i^{\mathsf{c}}) = 1 - \left(\frac{25}{26}\right)^n \longrightarrow 1 \quad \text{as } n \to \infty.$$

$\square$

(c) **Bad Beats [10 points]**

Poker extraordinaire Aadil is looking to see the odds that he loses money at the table tonight. He has been playing so dominantly that he will need to lose at least 90% of the n remaining rounds to lose money. The probability that Aadil loses each round is $\frac{1}{2}$, independent of all other rounds. Using Chebyshev's Inequality, show that the probability of Aadil losing money is at most $\frac{25}{32n}$.

Let $X \sim \text{Binomial}\left(n, \frac{1}{2}\right)$ be the number of rounds Aadil loses with $\mathrm{E}[X] = \frac{n}{2}$ and $\mathrm{var}(X) = \frac{n}{4}$. Note that $\text{Binomial}\left(n, \frac{1}{2}\right)$ is a symmetric distribution around its mean, which means that

$$\mathrm{P}(X \geq \mathrm{E}[X] + c) = \mathrm{P}(X \leq \mathrm{E}[X] - c) \quad \text{for all } c.$$

By Chebyshev's Inequality,

$$\begin{aligned}
\mathrm{P}\left(X \geq \frac{9}{10}n\right) &= \mathrm{P}\left(X \geq \mathrm{E}[X] + \frac{2n}{5}\right) \\
&= \frac{1}{2}\left[\mathrm{P}\left(X \geq \mathrm{E}[X] + \frac{2n}{5}\right) + \mathrm{P}\left(X \leq \mathrm{E}[X] - \frac{2n}{5}\right)\right] \\
&= \frac{1}{2} \cdot \mathrm{P}\left(\left|X - \mathrm{E}[X]\right| \geq \frac{2n}{5}\right) \\
&\leq \frac{1}{2} \cdot \frac{\frac{n}{4}}{\left(\frac{2n}{5}\right)^2} \\
&= \frac{25}{32n}
\end{aligned}$$

as desired.

□

(d) **Order! [9 points]**

Consider the sequence of i.i.d. random variables $X_1, X_2, \ldots, X_n$ where $X_i \sim \text{Uniform}[0,1]$ for all i. Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be their order statistics, where $X_{(k)}$ is the kth smallest element of the sequence. What is the joint PDF of $X_{(1)}$ and $X_{(n)}$, namely $f_{X_{(1)},X_{(n)}}(x_1, x_n)$ for $x_1 \le x_n$?

We pick one element from the n elements to be the smallest and one from the remaining $n-1$ elements to be the largest, which has $n(n-1)$ ways. Then, the rest $n-2$ elements need to go between x_1 and x_n, which has probability $(x_n - x_1)^{n-2}$. Then, the probability density that the two chosen elements (smallest and largest) take values x_1 and x_n is $1 \cdot 1 = 1$. Thus,

$$f_{X_{(1)},X_{(n)}}(x_1, x_n) = n(n-1)(x_n - x_1)^{n-2} \quad \text{for } x_1 \le x_n.$$

Alternative Method: We may also write the joint PDF as

$$f_{X_{(1)},X_{(n)}}(x_1, x_n) = f_{X_{(1)}}(x_1) \cdot f_{X_{(n)}|\{X_{(1)}=x_1\}}(x_n).$$

Then we compute the PDFs separately by differentiating the corresponding CDFs,

$$F_{X_{(1)}}(x_1) = 1 - (1 - x_1)^n \quad \Longrightarrow \quad f_{X_{(1)}}(x_1) = n(1 - x_1)^{n-1},$$

and

$$F_{X_{(n)}|\{X_{(1)}=x_1\}}(x_n) = \left(\frac{x_n - x_1}{1 - x_1}\right)^{n-1} \quad \Longrightarrow \quad f_{X_{(n)}|\{X_{(1)}=x_1\}}(x_n) = \frac{(n-1)(x_n - x_1)^{n-2}}{(1 - x_1)^{n-1}}.$$

Multiplying them together, we get the same result as above,

$$f_{X_{(1)},X_{(n)}}(x_1, x_n) = n(n-1)(x_n - x_1)^{n-2} \quad \text{for } x_1 \le x_n.$$

(e) **Encrypted Message [6 points]**

We are trying to send information across a channel using the fountain codes scheme **that we used in lab**. There are 4 data chunks, A, B, C, D, and Reina is able to successfully decode all of them using 4 packets, $(A, B), (A, B, D)$, (C) and a packet P whose content is unknown to us. What are all possible contents of P? Express your answer in tuples.

We directly have C. We need to either know A or B to decode all 4 chunks, so the possibilities for P are $(A), (B), (A, C), (B, C)$.

2 Café 126 [16 points]

When the popular EECS126 Café opens for business at 1:26PM, there is already a queue of $N \sim \text{Poisson}(\mu)$ people waiting outside. The café serves one person at a time, and the ith person's service time is X_i, where $X_i \sim \text{Exponential}(\lambda)$ for all i, independent of each other and N. Let $Y = \sum_{i=1}^{N} X_i$ be the total time needed to serve all customers that have been waiting in queue before Café 126 opens.

(a) What is $\mathrm{E}[Y]$?

(b) What is $\mathrm{var}(Y)$?

(a) By the Law of Total Expectation,

$$\mathrm{E}[Y] = \mathrm{E}\left[\sum_{i=1}^{N} X_i\right] = \mathrm{E}\left[\mathrm{E}\left[\sum_{i=1}^{N} X_i \middle| N\right]\right] = \mathrm{E}\left[N \cdot \frac{1}{\lambda}\right] = \frac{\mu}{\lambda}.$$

(b) By the Law of Total Variance,

$$\begin{aligned}
\mathrm{var}(Y) &= \mathrm{E}\left[\,\mathrm{var}(Y|N)\right] + \mathrm{var}\left(\,\mathrm{E}[Y|N]\right) \\
&= \mathrm{E}\left[N \cdot \mathrm{var}(X_1)\right] + \mathrm{var}\left(N \cdot \mathrm{E}[X_1]\right) \\
&= \mathrm{E}[N]\,\mathrm{var}(X_1) + \mathrm{var}(N)\,\mathrm{E}[X_1]^2 \\
&= \frac{\mu}{\lambda^2} + \frac{\mu}{\lambda^2} \\
&= \frac{2\mu}{\lambda^2}.
\end{aligned}$$

3 Competing Geometrics [22 points]

Let $X_1 \sim \text{Geometric}(p_1)$ and $X_2 \sim \text{Geometric}(p_2)$ be two independent random variables.

(a) Let $Y = \min(X_1, X_2)$. Find the distribution of Y.

(b) Find $\text{P}(X_1 < X_2)$.

(c) Let $Z = \max(X_1, X_2) - \min(X_1, X_2)$. Write $M_Z(s)$ in terms of $M_{X_1}(s)$ and $M_{X_2}(s)$, where $M_Z(s)$ is the moment generating function of Z. *Hint: condition Z on each of the three events $\{X_1 < X_2\}$, $\{X_1 > X_2\}$ and $\{X_1 = X_2\}$.*

(a) We can find

$$\text{P}(Y > y) = \text{P}(X_1 > y) \cdot \text{P}(X_2 > y) = (1-p_1)^y(1-p_2)^y = \big((1-p_1)(1-p_2)\big)^y,$$

which is the complementary CDF of $\text{Geometric}(p_1 + p_2 - p_1p_2)$.

(b) Consider two independent Bernoulli Processes (coin flipping processes) with success probability p_1 and p_2, which give rise to X_1 and X_2 respectively. The event $\{X_1 < X_2\}$ is equivalent to $\{X_1 = Y\} \cap \{X_2 > Y\}$ since Y is the minimum of the two. Then, conditioned on Y (meaning that the first $Y-1$ trials fail in both processes and the Yth trial succeeds in at least one process), the probability that the successful trial comes from the first process but not the second is $\frac{p_1(1-p_2)}{p_1+p_2-p_1p_2}$. Since this expression is not dependent on Y, $X_1 < X_2$ is independent of Y. Thus,

$$\text{P}(X_1 < X_2) = \text{P}(X_1 < X_2 | Y) = \frac{p_1(1-p_2)}{p_1+p_2-p_1p_2}.$$

(c) Using the same logic as the previous part,

$$\text{P}(Z \overset{d}{=} X_2) = \text{P}(X_1 < X_2) = \frac{p_1(1-p_2)}{p_1+p_2-p_1p_2},$$

$$\text{P}(Z \overset{d}{=} X_1) = \text{P}(X_2 < X_1) = \frac{p_2(1-p_1)}{p_1+p_2-p_1p_2},$$

$$\text{P}(Z = 0) = \text{P}(X_1 = X_2) = \frac{p_1p_2}{p_1+p_2-p_1p_2},$$

where $\overset{d}{=}$ denotes equal in distribution (having the same distribution). Thus,

$$M_Z(s) = \frac{p_2(1-p_1)M_{X_1}(s) + p_1(1-p_2)M_{X_2}(s) + p_1p_2}{p_1+p_2-p_1p_2}.$$

4 Rounded Exponentials [24 points]

Let $X \sim \text{Exponential}(\lambda)$, and let Y be the value of X rounded to the nearest integer. Namely,

$$Y = \begin{cases} \lfloor X \rfloor & \text{if } X - \lfloor X \rfloor < 0.5 \\ \lceil X \rceil & \text{if } X - \lfloor X \rfloor \geq 0.5 \end{cases}$$

where $\lfloor x \rfloor = \max\{k : k \leq x, k \in \mathbb{Z}\}$ is the floor function and $\lceil x \rceil = \min\{k : k \geq x, k \in \mathbb{Z}\}$ is the ceiling function.

(a) Let $\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ be the indicator random variable for the event that X is rounded down. Identify the distribution name of $\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ along with its parameter(s). Then, find its expectation and variance.

(b) Identify the distribution name of $\lceil X \rceil$ along with its parameter(s). Then, find its expectation and variance.

(c) Write Y in terms of $\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ and $\lceil X \rceil$. Then, find the expectation and variance of Y in terms of $\mathrm{E}\left[\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}\right]$, $\mathrm{var}\left(\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}\right)$, $\mathrm{E}\left[\lceil X \rceil\right]$ and $\mathrm{var}\left(\lceil X \rceil\right)$.

(a) By the memoryless property of X, we can note that $X - \lfloor X \rfloor$ is independent of $\lfloor X \rfloor$. Thus,

$$\mathrm{P}(X - \lfloor X \rfloor < 0.5) = \mathrm{P}(X - \lfloor X \rfloor < 0.5 | \lfloor X \rfloor = 0) = \frac{\mathrm{P}(X \leq 1/2)}{\mathrm{P}(X \leq 1)} = \frac{1 - e^{-\lambda/2}}{1 - e^{-\lambda}}.$$

Then,

$$\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}} \sim \text{Bernoulli}\left(\frac{1 - e^{-\lambda/2}}{1 - e^{-\lambda}}\right).$$

Call the parameter p. Then, $\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ has expectation

$$p = \frac{1 - e^{-\lambda/2}}{1 - e^{-\lambda}}$$

and variance

$$p(1-p) = \left(\frac{1 - e^{-\lambda/2}}{1 - e^{-\lambda}}\right)\left(\frac{e^{-\lambda/2} - e^{-\lambda}}{1 - e^{-\lambda}}\right) = \frac{e^{-\lambda/2} - 2e^{-\lambda} + e^{-3\lambda/2}}{(1 - e^{-\lambda})^2}.$$

(b) **Solution 1**: For integer k,

$$\mathrm{P}(\lceil X \rceil > k) = \mathrm{P}(X > k) = e^{-\lambda k},$$

which is the complementary CDF of $\text{Geometric}\left(1 - e^{-\lambda}\right)$. Call the parameter q. Then, $\lceil X \rceil$ has expectation

$$\frac{1}{q} = \frac{1}{1 - e^{-\lambda}}$$

and variance

$$\frac{1-q}{q^2} = \frac{e^{-\lambda}}{(1-e^{-\lambda})^2}.$$

Solution 2: We can think of each natural number $1, 2, 3, \ldots$ as a threshold. We ask whether X exceeds the first threshold, then repeatedly ask whether the leftover part of X exceeds the next threshold. By the memoryless property of the Exponential distribution, the leftover part of X beyond a certain threshold also has an Exponential(λ) distribution, so the probability that X exceeds each threshold is always $\mathrm{P}(X > 1)$. Since $\lceil X \rceil$ is equivalent to the number of tries it takes for X to fail to reach the next threshold, $\lceil X \rceil$ has the Geometric distribution with success probability $\mathrm{P}(X \le 1) = 1 - e^{-\lambda}$ (note that we are counting the number of trials to first fail, but the Geometric distribution counts the number of trials to first succeed, so we instead use the complementary probability as the success probability). The expectation and variance calculations are the same as above.

(c) $Y = \lceil X \rceil - \mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ since we first round X up regardless then subtract 1 if we actually need to round it down.

Since $X - \lfloor X \rfloor$ is independent of $\lfloor X \rfloor$, it is also independent of $\lfloor X \rfloor + 1 = \lceil X \rceil$ (which is true with probability 1). Thus, $\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ is also independent of $\lceil X \rceil$ because $\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}$ is a deterministic function of $X - \lfloor X \rfloor$. Thus,

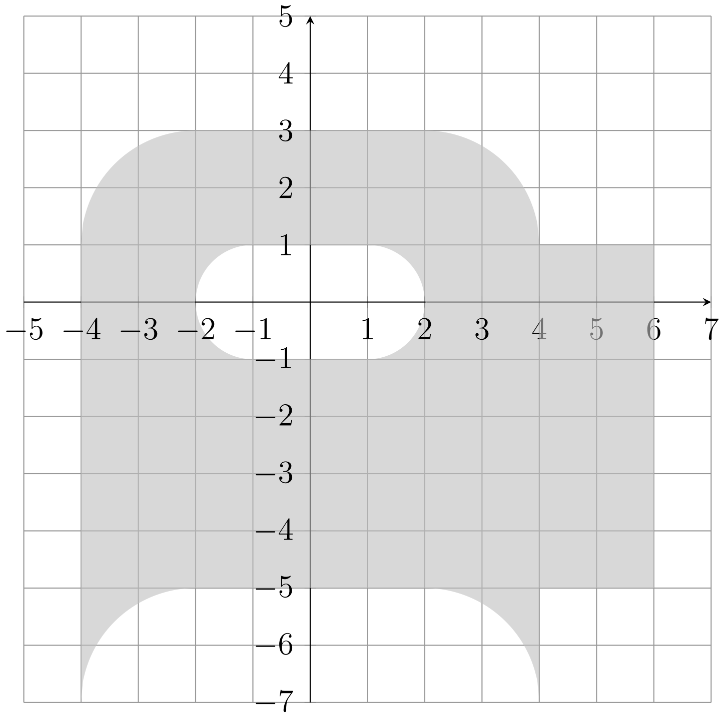
$$\mathrm{E}[Y] = \mathrm{E}\left[\lceil X \rceil\right] - \mathrm{E}\left[\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}\right] \qquad \text{by Linearity of Expectation}$$

and

$$\begin{aligned}
\operatorname{var}(Y) &= \operatorname{var}\left(\lceil X \rceil\right) + \operatorname{var}\left(-\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}\right) \qquad \text{by independence} \\
&= \operatorname{var}\left(\lceil X \rceil\right) + (-1)^2 \operatorname{var}\left(\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}\right) \\
&= \operatorname{var}\left(\lceil X \rceil\right) + \operatorname{var}\left(\mathbb{1}_{\{X - \lfloor X \rfloor < 0.5\}}\right).
\end{aligned}$$

5 Graphical Density? [22 points]

Consider the following shaded region in the x-y plane with uniform probability density ρ. The total area of the region is $72 - \pi$.



(a) **S**how your work: find ρ. You may leave answers to later subparts in terms of ρ.

(b) **U**sing ρ in part (a), find $\mathrm{E}[X]$ and $\mathrm{E}[X \mid Y = y]$ for $|y| \geq 1$.

(c) **S**till showing your work, find $f_X(x)$ for $|x| \geq 2$ and $f_{Y|X}(y \mid x)$ for $|x| \geq 2$.

(a) Since the probability over the region is uniform, the probability density ρ is the constant $\frac{1}{A}$, where A is the total area of the region. This gives $\rho = \frac{1}{72-\pi}$.

Remark. Slightly more formally, we are given $f_{X,Y}(x,y) \equiv \rho \cdot \mathbf{1}_R$; since $\iint f_{X,Y}\, dx\, dy = 1$, we find that $\rho \cdot \iint \mathbf{1}_R\, dx\, dy = \rho \cdot \mathsf{area}(R) = 1$.

(b) By the law of total expectation,

$$\mathrm{E}[X] = \mathrm{E}[X \mid |X| \leq 4] \cdot \mathrm{P}(|X| \leq 4) + \mathrm{E}[X \mid X > 4] \cdot \mathrm{P}(X > 4).$$

We note that since $f_{X,Y}$ is constant, any pair of points $(-x, y)$ and (x, y) contribute a total of zero to the expectation. Thus, by symmetry, $\mathrm{E}[X \mid |X| \leq 4] = 0$. Then the probability $\mathrm{P}(X > 4) = 12\rho$ is ρ times the area of the region lying in $\{X > 4\}$, and by symmetry again,

$\mathrm{E}[X \mid X > 4] = 5$. Therefore $\mathrm{E}(X) = 60\rho$. By the same argument,

$$\mathrm{E}[X \mid Y = y] = \begin{cases} 0 & y \in [1,3] \text{ or } y \in [-5,-7] \\ 1 & y \in [-1,-5]. \end{cases}$$

(c) We observe that the marginal density $f_X(x)$ is

$$f_X(x) = \int f_{X,Y}(x,y)\, dy = \rho \cdot \mathsf{length}(\{\text{points on the line } X = x\}).$$

Then we proceed by casework:

$$f_X(x) = \begin{cases} 8\rho & x \in [-4,-2] \text{ or } x \in [2,4] \\ 6\rho & x \in [4,6]. \end{cases}$$

The conditional distribution is then

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\rho}{f_X(x)} = \begin{cases} \frac{1}{8} & x \in [-4,-2] \text{ or } x \in [2,4] \\ \frac{1}{6} & x \in [4,6] \end{cases}.$$