EECS 126 Probability and Random Processes — UC Berkeley, Spring 2022
Kannan Ramchandran — May 13, 2022

Final

Last Name	First Name	SID

Left Neighbor Full Name	**Right Neighbor** Full Name	**Room Number**

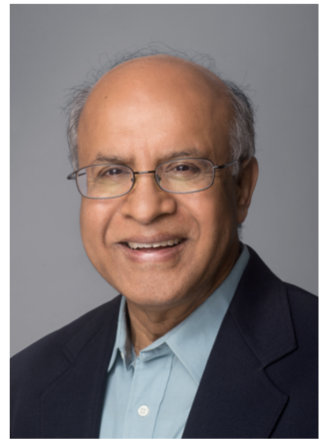
Rules.

- **Unless otherwise stated, all your answers need to be simplified and justified, and your work must be shown. Answers without sufficient justification will get no credit; answers without simplification may only get partial credit.**
- You have 170 minutes to complete the exam. (DSP students with $X\%$ time accommodation should spend $170 \cdot X\%$ time on the exam).
- This exam is not open book. You may reference three double-sided handwritten sheets of paper. No calculator or phones allowed.
- Collaboration with others is strictly prohibited. If you are caught cheating, you will receive a 0 on the final and will face disciplinary consequences.
- Write in your SID on every page to receive 1 point.

Problem	points earned	out of
SID/CAPTCHA		2
Problem 2		53
Problem 3		16
Problem 4		11
Problem 5		16
Problem 6		20
Problem 7		15
Problem 8		15
Problem 9		16
Total		164

1 CAPTCHA [1 point]

Bubble in the selection corresponding to Professor Kannan Ramchandran.



○ ○ ○ ○

2 Potpourri [7 + 7 + 7 + 7 + 8 + 7 + 10 points]

(a) **Exponential Sampling [7 points]**

Suppose you are able to generate values from $U \sim \text{Uniform}[0, 5]$. How can you simulate and sample values from an exponential distribution using U and $F_X(x)$, where $X \sim \text{Exponential}(\lambda)$? Provide complete justification for full credit.

(b) **Gob ears! [7 points]**

Sohom is watching the price of GBRS stock and notices that on each day, the stock either doubles or halves in price with equal probability. If X_n is the price of the stock on day n, then

$$X_{n+1} = \begin{cases} 2X_n & \text{w.p. } 1/2 \\ \frac{1}{2}X_n & \text{w.p. } 1/2. \end{cases}$$

Suppose $X_0 = 1$. Using the Central Limit Theorem, find a 95% confidence interval for X_{100}. You may approximate a 95% confidence interval with 2 standard deviations above and below the mean.

(c) **Chernoff Bounding [7 points]**

For a random variable $X \sim \text{Erlang}(2, 5)$, find the tightest upper bound on $\mathbb{P}(X \geq 1)$ given by the Chernoff bound. As a reminder, $\text{Erlang}(k, \lambda)$ is equivalent to the sum of k i.i.d. $\text{Exponential}(\lambda)$ random variables.

(d) **Entropic Kitchen [2 + 5 points]**

The number of dishes X in Akshit's sink at night is distributed as follows:

$$X = \begin{cases} 2 & \text{w.p. } \frac{1}{3} \\ 4 & \text{w.p. } \frac{1}{3} \\ 8 & \text{w.p. } \frac{1}{3}. \end{cases}$$

His roommate, Matt, either does all of the dishes ($Y = 1$) or none of them ($Y = 0$) according to the following model:

$$Y \sim \begin{cases} \text{Bernoulli}(\frac{2}{3}) & \text{if } X \leq 4 \\ \text{Bernoulli}(\frac{1}{3}) & \text{if } X > 4. \end{cases}$$

(i) Find $H(X)$, the entropy of X.

(ii) Akshit comes home and finds that the dishes are done. Find $H(X \mid Y = 1)$.

(e) **Comparing Gaussians [4 + 4 points]**

Answer True/False for the following two parts. If True, justify your answer. If False, provide a counterexample.

(i) If X and Y are jointly Gaussian, do they have Gaussian marginal distributions?

(ii) Is the converse true? If X and Y both have Gaussian marginal distributions, are they jointly Gaussian?

(f) **Quadratic Estimation [7 points]**

Given zero-mean random variables X and Y, find the best quadratic estimator $Q[Y \mid X]$ if $\mathbb{E}[X^3] = 0$. Your answer can contain expectations of X and Y such as $\mathbb{E}[X]$ or $\mathbb{E}[XY^2]$. Be sure to simplify completely and justify your work.

 Student ID: ____________

(g) **German Tank Problem [3 + 7 points]**

Rohit is working for Allied intelligence during WWII. He is tasked with estimating N, the total number of German tanks. Every so often, the troops capture a tank (and don't release it) and record its serial number, which run from 1 to N and are all distinct and equally likely to be found. Suppose the army finds two distinct serial numbers X_1 and X_2.

(i) Find the MLE of N given X_1, X_2.

(ii) Suppose an expert says that the prior distribution is $N \sim$ Geometric(p). Derive the quadratic equation using the log likelihood to find the MAP of N given X_1, X_2. If your equation could produce a non-integral value, explain how you would fix this.

3 AlexBot [3 + 5 + 5 + 3 points]

AlexBot is taking a random walk on the non-negative integers $\{0, 1, 2, \ldots\}$. Let $(X_n)_{n\geq 0}$ be a Markov chain, where X_n is its position at time n. The bot is programmed to do the following at each time step. At a state $i > 0$, it goes to state $i + 1$ with probability p, state $i - 1$ w.p. q, and stays in state i w.p. $1 - p - q$, where $p < q$ and $p + q \leq 1$. At state 0, there is a toy cannon that launches the bot onto a random state $i > 0$ according to a distribution π with expectation $c < \infty$, independent of any other event. In other words, the transition matrix P is given by

$$P_{i,j} = \begin{cases} p & \text{if } i > 0 \text{ and } j = i + 1 \\ q & \text{if } i > 0 \text{ and } j = i - 1 \\ 1 - p - q & \text{if } i > 0 \text{ and } j = i \\ \pi(j) & \text{if } i = 0. \end{cases}$$

You may assume that $\pi(i)$ is nonzero for at least one $i > 0$.

a) Use the strong law of large numbers to show that from any state i, the bot would eventually reach state 0 as time goes to infinity.

b) Find the expected amount of time for the bot to reach state 0, starting from any state $i > 0$. *Hint*: What can you say about the relationship between expected time from state 1 to 0 and the expected time from state $k + 1$ to k for $k > 0$?

c) Find $\mathbb{E}[T]$, where $T = \min\{t > 0 : X_t = 0 \mid X_0 = 0\}$ is the first return time to state 0. Don't forget that riding the cannon also takes one time step, and recall that the expectation of π is c.

d) Which states of the Markov Chain, if any, are positive recurrent? Justify your answer, and show irreducibility if necessary.

4 Caffeinated [2 + 2 + 7 points]

Catherine opens a coffee shop. Customers arrive according a Poisson process of rate λ.

(a) Let A be the event that Catherine got 100 sales in the first 100 minutes. Let T_1 and T_2 be the times when Catherine gets her first sale and the 101st sale. What is the conditional expectation of T_2 given A?

(b) What is the conditional expectation of T_1 given A?

(c) Aadil comes to help Catherine. Both Catherine and Aadil have Exponential service times with rate μ. Catherine always serves customers first, and Aadil will only serve a customer if Catherine is busy. Assuming the coffee shop starts out empty at the start of the day, how long is Aadil expected to wait before he serves a customer?

5 Gamblers [7 + 4 + 5 points]

Kamyar and Zhiwei are at a casino and observe that gamblers arrive as a Poisson process with rate 1 to play on one of three slot machines. Arriving gamblers will taken any random machine that is available, or leave immediately if none are available. Gamblers who get on the machines stay for an Exponential(1) amount of time independently of each other.

(a) Consider a CTMC which tracks the number of machines that are available or in use. Find the stationary distribution of this CTMC.

(b) Assume the chain has been running for a long time. If a gambler arrives, what is the probability that the first machine is open and the gambler starts playing on it?

(c) Suppose the casino has just opened, so all the slot machines are available. Set up equations to find the expected amount of time before a gambler leaves immediately after arriving (due to all machines being taken). You may leave the equations unsolved.

 Student ID: ____________________

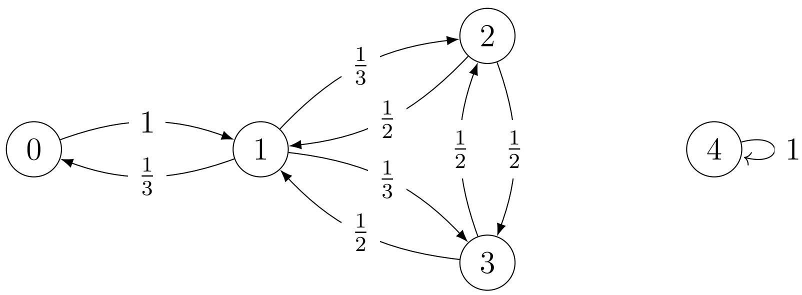
6 Random MCs [6 + 8 + 6 points]

Consider the Erdős–Rényi random graph $G \sim \mathcal{G}(n, p(n))$ on the state space $S := \{0, \dots, n-1\}$, where each edge in G exists independently with probability $p(n)$. Every such random graph uniquely describes a Markov chain on S as follows:

For vertices with positive (nonzero) degree $\deg(i)$, the transition probabilities are given by

$$P(i,j) = \begin{cases} \frac{1}{\deg(i)} & \text{if } (i,j) \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

where $E(G)$ denotes the edge set of G. Otherwise, when $\deg(i) = 0$, $P(i,i) = 1$. For example, consider the following Markov chain created using the process above:



In the following parts, a "random Markov chain" refers to the corresponding Markov chain from sampling a random graph $G \sim \mathcal{G}(n, p(n))$, **not** just the specific example given above. **Justify** any properties or equations of Markov chains and random graphs that you use.

(a) As $n \to \infty$, find the probability that a random MC has at least one stationary distribution. You may leave your answer in terms of $p(n)$ if necessary.

(b) Consider a "typical" vertex i in a "typical" random graph G. Specifically, let $|E(G)|$ be equal to the expected number of edges in a random graph drawn from $\mathcal{G}(n, p(n))$, and let $\deg(i)$ be the expected degree of a randomly chosen vertex in $G \sim \mathcal{G}(n, p(n))$. Compute the expected return time in the associated random MC for vertex i: $\mathbb{E}_i(T_i^+) = \mathbb{E}[\min\{t \geq 1 : X_t = i\} \mid X_0 = i]$.

(c) Suppose that $p(n) \in o(\frac{1}{n})$, i.e. as $n \to \infty$, $n \cdot p(n) \to 0$. As $n \to \infty$, find the probability that a random MC is reversible. (*Hint*: Recall that this is the **subcritical phase**, where the probability that $G \sim \mathcal{G}(n, p(n))$ contains no cycles tends to 1 as $n \to \infty$.)

7 Hypothesis Testing [5 + 3 + 2 + 5 points]

Consider a random variable Y that follows one of two distributions. Let X be a binary random variable indicating the true distribution of Y:

$$X = \begin{cases} 0, & Y \sim \mathcal{N}(0,1) \\ 1, & Y \sim U \cdot \mathcal{N}(2,1) + (1-U) \cdot \mathcal{N}(-2,1), \text{ where } U \sim \text{Bernoulli}(\frac{1}{2}). \end{cases}$$

Follow the steps below to construct a Neyman–Pearson decision rule $\hat{X}(Y)$ that maximizes $\mathbb{P}(\hat{X}(Y) = 1 \mid X = 1)$ under the constraint that $\mathbb{P}(\hat{X}(Y) = 1 \mid X = 0) \leq \beta$.

(a) Find the likelihood ratio $L(y)$ for $y \in \mathbb{R}$. Simplify $L(y)$ and show that $L(y) = (e^{2y} + e^{-2y})/(2e^2)$.

(b) Consider $y_1 > y_2 > 0$. Show that $L(y_1) > L(y_2)$.

(c) Argue that for $y_1 < y_2 < 0$, $L(y_1) > L(y_2)$.

(d) Construct a Neyman–Pearson decision rule. Leave your answer(s) for the decision boundary in terms of Φ^{-1}, the inverse CDF of the standard normal distribution.

8 Some Estimation [3 + 5 + 3 + 4 points]

Consider the following joint PDF of two random variables X and Y:

$$f_{X,Y}(x,y) = \begin{cases} k(|x| + y), & -1 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine k.

(b) Find the minimum mean squared error estimator of Y given X, i.e. $\mathrm{MMSE}[Y \mid X = x]$.

(c) Determine $\mathrm{cov}(X, Y)$. *Hint*: You shouldn't need to compute integrals.

(d) Determine the linear least squares estimator of Y given X, i.e. $\mathbb{L}[Y \mid X = x]$.

9 Basketball III [4 + 6 + 6 points]

Captain America and Superman return for one final game of basketball! Unfortunately, there is a tall fence between you and the basketball court, so you must repeatedly jump up and down at times $n = 1, 2, 3, \ldots$ to watch the game.

To figure out who wins in the end, you decide to track the ball's x-coordinate, where $X_0 = 0$ represents the center of the court. The state space equations are as follows. Note that the dynamics model A_n is not time-homogeneous.

$$X_n = A_n X_{n-1} + V_n \qquad A_n = (-1)^n,\ V_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 2);$$

$$Y_n = X_n + W_n \qquad W_n \overset{\perp\!\!\!\perp}{\sim} \mathcal{N}(0, n).$$

For your reference, the (now slightly modified) Kalman filter equations are given below.

$$\hat{X}_{n|n-1} \leftarrow A_n \hat{X}_{n-1|n-1} \qquad \sigma^2_{n|n-1} \leftarrow A_n^2 \sigma^2_{n-1|n-1} + \sigma^2_V$$

$$\tilde{Y}_n \leftarrow Y_n - \hat{X}_{n|n-1} \qquad K_n \leftarrow \sigma^2_{n|n-1} / (\sigma^2_{n|n-1} + \sigma^2_{W_n})$$

$$\hat{X}_{n|n} \leftarrow \hat{X}_{n|n-1} + K_n \tilde{Y}_n \qquad \sigma^2_{n|n} \leftarrow (1 - K_n)\, \sigma^2_{n|n-1}.$$

(a) Find the distribution of Y_n. Your answer should not depend on any other random variables.

(b) Find $\hat{X}_{n|n}$ as a summation in terms of A_k, K_k, and Y_k for $k = 1, \ldots, n$.

(c) Find $\hat{X}_{2|2}$ as a function of Y_1 and Y_2, assuming we initialize $\hat{X}_{0|0} \leftarrow 0$ and $\sigma^2_{0|0} \leftarrow 0$.

You may use the fact that $K_2 = 4/7$.

10 Cheat Sheet

- $X \sim \text{Bernoulli}(p)$, $p \in [0,1]$.

  PMF: $p_X(x) = p^x(1-p)^{1-x}$, $x \in \{0,1\}$.

  MGF: $M_X(s) = 1 - p + p \exp s$.

  Moments: $\mathbb{E}[X] = p$, $\operatorname{var} X = p(1-p)$.

- $X \sim \text{Binomial}(n, p)$, $n \in \mathbb{Z}_+$, $p \in [0,1]$.

  PMF: $p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x \in \{0, \dots, n\}$.

  MGF: $M_X(s) = (1 - p + p \exp s)^n$.

  Moments: $\mathbb{E}[X] = np$, $\operatorname{var} X = np(1-p)$.

- $X \sim \text{Geometric}(p)$, $p \in (0,1)$.

  PMF: $p_X(x) = pq^{x-1}$, $x \in \mathbb{Z}_+$, $q = 1 - p$.

  MGF: $M_X(s) = (p \exp s)/(1 - q \exp s)$, $s < \ln(1/q)$.

  Moments: $\mathbb{E}[X] = p^{-1}$, $\operatorname{var} X = q/p^2$.

- $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$.

  PMF: $p_X(x) = \lambda^x \exp(-\lambda)/x!$, $x \in \mathbb{N}$.

  MGF: $M_X(s) = \exp(\lambda(\exp s - 1))$.

  Moments: $\mathbb{E}[X] = \lambda$, $\operatorname{var} X = \lambda$.

  X, Y independent, $X \sim \text{Poisson}(\lambda)$, $Y \sim \text{Poisson}(\mu) \implies X + Y \sim \text{Poisson}(\lambda + \mu)$.

- $X \sim \text{Uniform}[a, b]$, $a < b$.

  PDF: $f_X(x) = (b-a)^{-1}$, $x \in [a, b]$.

  MGF: $M_X(s) = (\exp(sb) - \exp(sa))/(s(b-a))$.

  Moments: $\mathbb{E}[X] = (a+b)/2$, $\operatorname{var} X = (b-a)^2/12$.

- $X \sim \text{Exponential}(\lambda)$, $\lambda > 0$.

  PDF: $f_X(x) = \lambda \exp(-\lambda x)$, $x > 0$.

  CDF: $F_X(x) = (1 - \exp(-\lambda x))$, $x \geq 0$.

  MGF: $M_X(s) = \lambda/(\lambda - s)$, $s < \lambda$.

  Moments: $\mathbb{E}[X] = \lambda^{-1}$, $\operatorname{var} X = \lambda^{-2}$.

- $X \sim \mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$.

  PDF: $f_X(x) = (\sqrt{2\pi}\sigma)^{-1} \exp(-(x-\mu)^2/(2\sigma^2))$.

  CDF: $F_X(x) = \Phi((x-\mu)/\sigma)$.

  MGF: $M_X(s) = \exp(\mu s + \sigma^2 s^2/2)$.

  Moments: $\mathbb{E}[X] = \mu$, $\operatorname{var} X = \sigma^2$.

  X, Y independent, $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma_2^2) \implies X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

- $X \sim \text{Erlang}(k, \lambda)$, $k \in \mathbb{Z}_+$, $\lambda > 0$.

  Sum of k i.i.d. Exponential(λ).

  PDF: $f_X(x) = \lambda^k x^{k-1} \exp(-\lambda x)/(k-1)!$, $x \geq 0$.

Tail Sum: For $X \geq 0$, $\mathbb{E}[X] = \int_0^\infty \Pr(X \geq x)\,\mathrm{d}x$.

Variance: $\operatorname{var} X = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Sum: $\operatorname{var} \sum_{i=1}^n X_i = \sum_{i=1}^n \operatorname{var} X_i + \sum_{i \neq j} \operatorname{cov}(X_i, X_j)$.

Covariance: $\operatorname{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$.

Correlation: $\rho(X, Y) = \operatorname{cov}(X, Y)/\sqrt{(\operatorname{var} X)(\operatorname{var} Y)}$.

Entropy: $H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x)$.

Order Statistics: $f_{X^{(i)}}(x) = n\binom{n-1}{i-1} f(x) F(x)^{i-1} (1 - F(x))^{n-i}$.

MGF: $M_X(s) = \mathbb{E}[\exp(sX)]$.

Markov: For $X \geq 0$, $x > 0$, $\Pr(X \geq x) \leq \mathbb{E}[X]/x$.

Chebyshev: For $x > 0$, $\Pr(|X - \mathbb{E}[X]| \geq x) \leq (\operatorname{var} X)/x^2$.

Chernoff: For $t > 0$, $\Pr(X \geq x) = \Pr(e^{tX} \geq e^{tx})$. For $t > 0$, $\Pr(X \leq x) = \Pr(e^{-tX} \geq e^{-tx})$.

LLSE: $\mathbb{L}[Y \mid X] = \mathbb{E}[Y] + \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)}(X - \mathbb{E}[X])$.

MMSE: $\text{MMSE}[Y \mid X] = \mathbb{E}[Y|X]$.