EECS 126 Probability and Random Processes
Kannan Ramchandran

UC Berkeley, Fall 2022
November 3, 2022

Midterm 2

Last Name	First Name	SID

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

Rules.

- **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	30
Problem 2	15
Problem 3	15
Problem 4	15
Problem 5	20
Problem 6	30
Total	126

1 Scheduling Conflict [30 points]

Oh no, EECS 126 and CS 170 are having Homework Parties at Cory Courtyard at the same time! EECS 126 and CS 170 students arrive at Cory Courtyard independently according to two Poisson processes with rates λ_{126} and λ_{170} respectively. For simplicity, assume no student is taking the two classes at the same time.

(a) Let T_3 be the time that the third student arrives. What is $\mathrm{E}[T_3]$?

Andy walks by Cory Courtyard at time t and sees there are three students.

(b) What is the expected time between the last student arrival before Andy and the next student arrival after Andy?

(c) What is the probability that Andy sees more EECS 126 students than CS 170 students when he walks by?

(a) We can merge the two Poisson processes into one student arrival process of rate $\lambda_{126}+\lambda_{170}$. Since the interarrival times are independent, and each interarrival interval has an expected length of $(\lambda_{126}+\lambda_{170})^{-1}$, we have

$$\mathrm{E}[T_3] = \frac{3}{\lambda_{126}+\lambda_{170}}.$$

(b) This is slightly different from the Random Incidence Property since we are given there are three arrivals before time t, but a similar analysis technique can apply. The three arrivals before time t are distributed as the order statistics of three independent uniform random variables in $[0, t]$. The three student arrivals split the interval $[0, t]$ into four sub-intervals, each of which has the same expected length, so

$$\mathrm{E}[t - T_3 \mid N_t = 3] = \frac{1}{4}t.$$

Then the next student arrival takes $(\lambda_{126}+\lambda_{170})^{-1}$ time in expectation, giving us the answer

$$\frac{1}{4}t + \frac{1}{\lambda_{126}+\lambda_{170}}.$$

(c) By competing Exponentials, each student that arrives has probability $p := \lambda_{126}/(\lambda_{126}+\lambda_{170})$ of being a EECS 126 student, independent of all other students. Thus, the number of EECS 126 students in the first three arrivals is distributed as Binomial$(3, p)$. Then the probability that two or more EECS 126 students are in the first three student arrivals is

$$\frac{3\lambda_{126}^2\lambda_{170}+\lambda_{126}^3}{(\lambda_{126}+\lambda_{170})^3}.$$

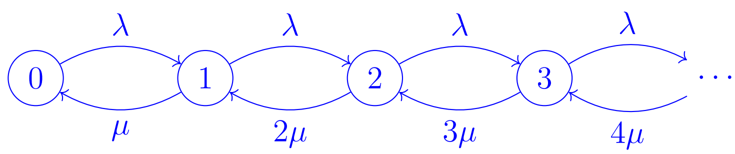
2 IKEA Shopping [15 points]

At a popular plush toy and furniture store, Axel observes that customers enter according to a Poisson process with rate λ, and *each customer* spends i.i.d. Exponential(μ) time in the store before leaving.

(a) Draw the state transition diagram for a CTMC to model the number of customers in the store.

(b) What is the stationary distribution of the CTMC? For simplicity, you do not need to normalize the distribution but can choose to leave it in terms of C, a normalizing constant.

(a) Let the states of the chain be $\mathbb{N}$, corresponding to the number of customers in the store. Observe that if the chain is in state n, it transitions to $n+1$ with rate λ and to $n-1$ with rate $n\mu$, which gives rise to the following transition diagram:



(b) To find the stationary distribution, we observe that the graph is a tree, so any stationary distribution must satisfy the detailed balance equations. Define $\rho = \frac{\lambda}{\mu}$. For all $n \geq 1$,

$$n\mu\pi(n) = \lambda\pi(n-1)$$
$$\pi(n) = \frac{\rho}{n}\pi(n-1)$$

Recursively applying this relation, we have $\pi(n) = \frac{\rho^n}{n!}\pi(0)$, which also holds for $\pi(0)$ since $\frac{\rho^0}{0!} = 1$. Thus,

$$\pi(n) = C \cdot \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!}.$$

For the optional task to find C, since the stationary distribution must sum to 1, we also have

$$1 = \sum_{n=0}^{\infty} \pi(n) = \sum_{n=0}^{\infty} \frac{\rho^n}{n!}\pi(0) = e^{\rho}\pi(0).$$

This gives $C = e^{-\rho}$.

3 An Experimental Algorithm [15 points]

Recall the Metropolis-Hastings algorithm, which allows us to generate samples from some distribution $\pi(x)$ by simulating an appropriate Markov chain. As in homework and lab, suppose we have access to unnormalized density $f(x) = C\pi(x)$ for unknown C and proposal distribution $g(x, \cdot)$, but we are experimenting with different choices of the acceptance function $A(x, y)$. For each of the following possible acceptance functions, determine whether or not samples from the resulting Metropolis-Hastings variant will follow the desired distribution $\pi(x)$ and briefly explain why.

Hint: At each step of the algorithm, at state x, we propose a candidate state y with probability $g(x, y)$ and accept it with probability $A(x, y)$, so $P(x, y) = g(x, y)A(x, y)$. Your goal is to determine whether the transition probabilities make the Markov Chain reversible under $\pi(x)$.

(a) $A(x, y) = \min\left\{1, \frac{f(y)g(y,x)}{f(x)g(x,y)}\right\}$

(b) $A(x, y) = \min\left\{\frac{1}{2}, \frac{1}{2} \cdot \frac{f(y)g(y,x)}{f(x)g(x,y)}\right\}$

(c) $A(x, y) = \max\left\{\frac{9}{10}, \min\left\{1, \frac{f(y)g(y,x)}{f(x)g(x,y)}\right\}\right\}$

(a) This is precisely the same acceptance function as in the original Metropolis-Hastings, which we showed to work in homework.

(b) This is equivalent to the lazy chain considered in Homework 8 Problem 5(d). The resulting transition matrix P' given by this choice of acceptance function obeys $P' = 1/2I + 1/2P$, where P is the transition matrix under the original algorithm, so the stationary distribution will be the same.

(c) This proposal function will not work. Consider the counterexample where $g(x, y) = g(y, x) = 0.5$, $\pi(x) = f(x) = 0.1$, $\pi(y) = f(y) = 0.2$. Then,

$$P(x, y) = 0.5 \cdot 1 = 0.5,$$

$$P(y, x) = 0.5 \cdot 0.9 = 0.45,$$

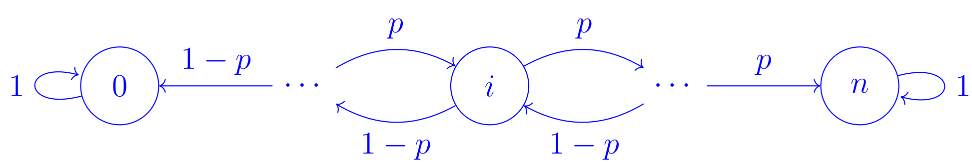
which doesn't satisfy $\pi(x)P(x, y) = \pi(y)P(y, x)$.

4 Brink of Ruin [15 points]

Alex, realizing his wallet is empty, finds himself entered into the following game. He starts with \$1. Then, on each turn, he flips a fair coin, winning one dollar if it comes up heads and losing one dollar otherwise. He also wins a *gambler's token* for every turn where he has only \$1, including the start. The game ends when he has no more money or has \$100 in total. What is the expected number of gambler's tokens Alex has when the game ends?

Hint: consider applying first-step analysis to β_i, the expected *number of visits* to state \$1 starting from $\$i$. To solve the recurrence relation that you find, use the fact that if b is the average of a and c, then $c - b = b - a$ is a common difference.

This is the symmetric gambler's ruin problem with $p = \frac{1}{2}$ and $n = 100$, which can be represented as a discrete-time Markov chain on the state space $\{0, \dots, n\}$ with transition diagram



We wish to find $\mathrm{E}_1[N_1]$, the expected total number of visits to state 1 starting from 1. Letting $\beta_i := \mathrm{E}_i[N_1]$, we have the following system of equations by first-step analysis:

$$\beta_1 = 1 + \frac{1}{2}\beta_2$$

$$\beta_i = \frac{1}{2}\beta_{i-1} + \frac{1}{2}\beta_{i+1}, \qquad 2 \le i \le 99$$

along with $\beta_0 = 0$ and $\beta_{100} = 0$. Using the hint, we observe that β_i is precisely the average, or center, of β_{i-1} and β_{i+1}, and we can denote the following common difference Δ:

$$\beta_i - \beta_{i+1} = \beta_{i-1} - \beta_i.$$

Then, with $\beta_{99} = \frac{1}{2}\beta_{98} = \beta_{98} - \beta_{99} = \Delta$, we have by telescoping sums

$$\beta_k = \beta_k - \beta_{100} = \sum_{i=k}^{99}(\beta_i - \beta_{i+1}) = (100 - k)\Delta = (100 - k)\beta_{99}.$$

We can use $\beta_1 = 99\beta_{99} = 1 + \frac{1}{2}\beta_2 = 1 + 49\beta_{99}$ to find $\beta_{99} = \frac{2}{100}$, which gives the answer of

$$\mathrm{E}_1[N_1] = \beta_1 = \boxed{\frac{99}{50}}.$$

Alternate solution. We observe that if θ is the probability that the chain eventually revisits state 1 starting from 1, then

$$(N_1 \mid X_0 = 1) \sim \text{Geometric}(1-\theta).$$

The probability that state 1 is visited exactly $k \in \mathbb{Z}^+$ times is $\theta^{k-1}(1-\theta)$ by the memorylessness given by the Markov property, which characterizes the memoryless Geometric distribution. We can then find θ as

$$\mathrm{P}(\textsf{visits state 2 from 1}) \cdot \mathrm{P}(\textsf{visits state 1 before 100} \mid \textsf{starts from state 2}) = \frac{1}{2} \cdot \frac{98}{99},$$

observing that the probability the chain visits state 1 before state 100, starting from state 2, is precisely the hitting probability of the symmetric gambler's ruin problem, which we recall from homework as $\frac{100-2}{100-1} = \frac{98}{99}$. Thus, as above, we find

$$\mathrm{E}[N_1 \mid X_0 = 1] = \frac{1}{1-\theta} = \frac{99}{50}.$$

5 Entropic Maneuvers [20 points]

Consider two (potentially dependent) random variables X and Y that each takes value in the set $\{1, 2, \dots, n\}$. Let $E = \mathbb{1}_{\{X \neq Y\}}$ be the indicator random variable for the event that X and Y are not equal, and let $p = \mathrm{P}(X \neq Y) = \mathrm{P}(E = 1)$.

(a) Show that $H(X, E \mid Y) = H(X \mid Y)$.

(b) Using the previous part, show that $H(X \mid Y) \leq p \log_2(n-1) + H(E)$.

Hint: Use the fact that the uniform distribution on a set of k elements has an entropy of $\log_2 k$ and has the maximum entropy among all distributions on the set.

(a) By the chain rule for entropy,

$$H(X, E \mid Y) = H(X \mid Y) + H(E \mid X, Y).$$

We see that $H(E \mid X, Y) = 0$: knowing X and Y allows us to determine E with certainty. Thus $H(X, E \mid Y) = H(X \mid Y)$.

(b) By another application of the chain rule,

$$\begin{aligned} H(X \mid Y) &= H(X, E \mid Y) \\ &= H(E \mid Y) + H(X \mid E, Y) \\ &= H(E \mid Y) + p \cdot H(X \mid E = 1, Y) + (1 - p) \cdot H(X \mid E = 0, Y). \end{aligned}$$

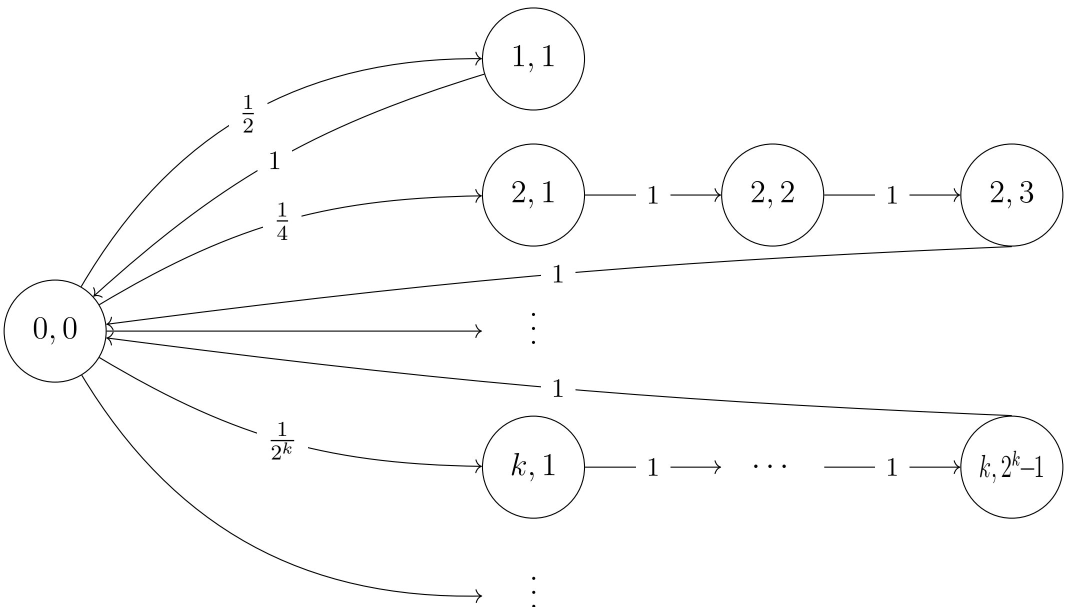
We see that $H(X \mid E = 0, Y) = 0$ since $E = 0$ implies $X = Y$, so there is no leftover uncertainty in X. Next, given $E = 1$ and Y, X takes value in $n - 1$ elements, namely everything but the value Y takes on, so it has a conditional entropy upper-bounded by $\log_2(n-1)$. Lastly, $H(E \mid Y) \leq H(E)$ since conditioning only decreases entropy. Putting these together, we have

$$H(X \mid Y) \leq p \log_2(n-1) + H(E)$$

as desired.

6 Bot on a Stroll [30 points]

The EECS 126 Bot is taking a walk on a Markov chain with state space $\mathbb{N} \times \mathbb{N}$, starting from state $(0, 0)$, as shown by the graph below.



From state $(0, 0)$, the bot chooses "path k" with probability 2^{-k} for $k = 1, 2, \ldots$ Each path k contains $2^k - 1$ states, which the bot will travel through in sequence then return to $(0, 0)$ deterministically.

(a) Is this Markov chain irreducible? Justify your answer.

(b) What is the period of this Markov chain?

(c) What is the expected time to return to state $(0, 0)$?

(d) Is this Markov chain positive recurrent, null recurrent, or transient? Justify your answer.

(a) The Markov chain is irreducible. The path from state (x_t, y_t) to (x_{t+1}, y_{t+1}) can be traced out by first following path x_t to reach $(0, 0)$ in $2^{x_t} - y_t$ steps, then jumping to $(x_{t+1}, 1)$ and following path x_{t+1} for $y_{t+1} - x_{t+1}$ steps.

(b) The period is 2 since the the time to travel path k is 2^k for $k = 1, 2, 3, \ldots$ with a GCD of 2.

(c) The probability to take path i is 2^{-i}, and the time it takes to return to $(0, 0)$ given that we take path i is $(2^i - 1) + 1 = 2^i$. Thus, the expected return time is

$$\sum_{i=1}^{\infty} 2^i \cdot 2^{-i} = \sum_{i=1}^{\infty} 1 = \infty.$$

 Student ID (write on every page for 1 point): ____________

(d) The Markov chain is null recurrent. From $(0, 0)$, no matter which path i the bot takes, it will always return to $(0, 0)$ in $2^i < \infty$ steps. Since it always returns to state $(0, 0)$, the Markov chain is recurrent. From the previous part, since the expected return time is infinite, we further classify the Markov chain as null recurrent.