EECS 126 Probability and Random Processes | UC Berkeley, Spring 2023
Thomas Courtade | March 23, 2023

Midterm 2

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.**
- **Start this exam by doing Problem 1 first.**
- **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 80 minutes to complete the exam. (DSP students with $X\%$ time accommodation should spend $80 \cdot X\%$ time on completing the exam).
- This exam is closed-book. You may reference one double-sided sheet of notes. No calculator or phones allowed.
- Remember the Berkeley Honor Code: "*As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others.*" Any violation of academic integrity will be taken seriously, and could result in disciplinary consequences.

Problem	out of
SID	1
Problem 1	6
Problem 2	13
Problem 3	18
Problem 4	20
Problem 5	20
Problem 6	22
Total	100

1 Answer This Question First [3 + 3 points]

a) What's the most interesting thing you've learned so far this semester (in this class, or any other)?

b) On a scale from 1-10 (10 being best), how well do you think you understand the material in this course?

2 Casino [8 + 5 points]

A casino has a new game where the host goes behind a curtain, flips a fair coin repeatedly, and records the number of tosses N it took to see the first heads. Your job, as the player of the game, is to determine the value of N by asking the host as few yes/no questions as possible.

a) How many yes/no questions, on average, do you need to ask in order to determine the value of N?

b) Give an optimal sequence of *simple* questions for determining N. By "optimal", we mean that you should ask fewest possible questions on average in order to determine N.

a) By construction $N \sim \mathrm{Geom}(1/2)$. By the source coding theorem, we need to ask at least

$$H(N) = \sum_{n \geq 1} (1/2)^n \log_2(2^n) = \sum_{n \geq 1} n(1/2)^n = \mathbb{E}[N] = 2$$

questions on average in order to determine N.

b) We can ask the sequence of questions "Is $N = 1$?", "Is $N = 2$?", "Is $N = 3$?", ... The number of questions we need to ask on average to determine N with this sequence of questions is precisely $\mathbb{E}[N] = 2$, so this sequence is optimal.

3 Simple Shuffles [6 + 8 + 4 points]

a) A square matrix P is called *doubly stochastic* if P is a stochastic matrix, and all columns sum to one. What is the stationary distribution of an irreducible finite-state DTMC with a doubly stochastic transition matrix P?

b) A robot shuffles a standard deck of 52 cards according to a simple algorithm: Starting with any initial ordering of the cards, a single card is uniformly chosen at random from the 52 cards in the deck, and is then re-inserted at another uniformly random location (it is possible for the card to be returned to the same place). This process is repeated over and over, ad infinitum. Argue that, as the number of these simple shuffles tends to infinity, the deck becomes perfectly shuffled.

c) If the deck in part (b) starts with ordering $(1, 2, \ldots, 52)$, how many operations will it take on average before the cards are again in the order $(1, 2, \ldots, 52)$?

a) Since all columns sum to one, we have $\mathbf{1}^T P = \mathbf{1}^T$, where $\mathbf{1}$ is the column vector of all-ones. By dividing through by the number of states, we see the uniform distribution is a stationary distribution.

b) This process may be modeled as a DTMC, with states equal to all permutations of $\{1, \ldots, 52\}$, corresponding to the ordering of the deck of cards. The chain is irreducible, since from any initial starting state, we can get to any other ordering of the cards in no more than 52 steps. The transition matrix for this DTMC is doubly stochastic, since all columns are permutations of one another (and therefore have the same column sums, which must be one). So, by the big theorem and part (a), the stationary distribution π exists, is unique, and is uniform across permutations. This chain has self-loops (the robot can select the top card and place it back on top), so is aperiodic, and hence the big theorem tells us the distribution over states converges to the stationary distribution, which is uniform over all permutations. In other words, as the number of shuffles tends to infinity, the deck will be equally likely to be in any of the possible orders (i.e., the deck becomes perfectly shuffled in the limit).

c) By the big theorem, the expected number of steps to return to the initial state is $52!$.

4 Double Heads [10 + 10 points]

Suppose you flip a fair coin repeatedly.

a) Given that the first flip turns heads, what is the expected number of tosses (including your first toss) you need to make until you see two heads in a row?

b) Again, given that the first flip turns heads, what is the probability you see two heads in a row before you see two tails in a row?

a) We can model this problem as a DTMC, with state space $\{HH, HT, TH, TT\}$ and transition matrix

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix},$$

where the columns (rows) are respectively correspond to states HH, HT, TH, TT. The state $s_1 s_2$ means we saw s_2 on the most recent flip, and s_1 on the flip before that. Now, we can imagine our first flip turning heads starts us in state HH, and we should compute the average number of flips to return to HH; the number of flips we are after is then 1 plus this number. By inspection, P admits stationary distribution $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ (you can also observe P is doubly stochastic). This chain is irreducible, so by the big theorem, the average number of flips to see two heads in a row, given our first flip is heads, is $1 + 4 = 5$.

b) We can do this by appealing to the equations for hitting probabilities. Set $h(HH) = 1$, and $h(TT) = 0$, where $h(s)$ equals the probability we'll hit HH before TT, given we start in state s for the DTMC formulated in part (a). The hitting probability equations for states $h(HT)$ and $h(TH)$ are

$$h(HT) = \frac{1}{2}h(TH) + \frac{1}{2}h(TT) = \frac{1}{2}h(TH)$$
$$h(TH) = \frac{1}{2}h(HT) + \frac{1}{2}h(HH) = \frac{1}{2}h(HT) + \frac{1}{2} = \frac{1}{4}h(TH) + \frac{1}{2}.$$

The last says $h(TH) = \frac{2}{3}$, which gives $h(HT) = \frac{1}{3}$. Since we flipped a heads on the first toss, we can imagine starting in state HT or HH, with probability $\frac{1}{2}$ each, determined by our second flip. The probability we'll see HH before TT is thus equal to

$$P\{\text{we see } HH \text{ before } TT \mid \text{first flip is } H\} = \frac{1}{2}h(HT) + \frac{1}{2}h(HH) = \frac{1}{6} + \frac{1}{2} = \frac{2}{3}.$$

(Extra page for Problem 4)

5 Leaky Roof [6 + 6 + 8 points]

a) Let $(U_i)_{i\geq 1} \sim_{\text{i.i.d.}} \text{Unif}(0,1)$. Find $\mathbb{E}[\max\{U_1,\dots,U_n\}]$.

b) Raindrops from a leaky roof fall into a small red bucket according to a Poisson process, having rate 1 drop/sec. You empty the red bucket into a larger blue bucket at a rate of once per minute, according to a Poisson process which is independent of the raindrop process. When you empty the red bucket for the first time, what is the probability that it contains 60 or more raindrops?

c) Assume the model of part (b). For a given time $t \geq 0$ (t has units of seconds), find the expected number of raindrops in the larger blue bucket. (Make the simplifying assumption that the buckets never overflow.) Hint: Use part (a).

a) Let $U = \max\{U_1,\dots,U_n\}$. We have $P(U \leq t) = \prod_{i=1} P(U_i \leq t) = t^n$, $0 \leq t \leq 1$. Hence, the density of U is equal to $f_U(t) = nt^{n-1}$, and we compute

$$\mathbb{E}[U] = n\int_0^1 t \cdot t^{n-1} dt = \frac{n}{n+1}.$$

b) Let $(R_t)_{t\geq 0}$ denote process that counts raindrops, and $(N_t)_{t\geq 0}$ denote the number of times you've emptied the small bucket up to time t. We can view the processes $(R_t)_{t\geq 0}$ and $(N_t)_{t\geq 0}$ as being obtained from splitting a rate 61/60 Poisson process, where arrivals are marked as raindrops (resp. bucket dumps) with probability 60/61 (resp. 1/61). Hence, the probability we are after is the probability that the first 60 arrivals are marked as raindrops, which is $(60/61)^{60}$. Just for your interest, we note this number can be approximated as

$$(60/61)^{60} = (1-1/61)^{60} \approx e^{-60/61} \approx 1/e.$$

c) Letting T_{N_t} denote the time at which you most recently emptied the smaller bucket, the total number of raindrops you will have emptied will be $R_{T_{N_t}}$, which conditioned on T_{N_t}, is Poisson(T_{N_t}). Hence,

$$\mathbb{E}[R_{T_{N_t}}] = \mathbb{E}[\mathbb{E}[R_{T_{N_t}}|T_{N_t}]] = \mathbb{E}[T_{N_t}].$$

Now, conditioned on N_t, the time T_{N_t} coincides with the largest order statistic of N_t Uniform$(0,t)$ random variables. Hence, by part (a) and the fact that $N_t \sim \text{Poisson}(\lambda t)$ (for $\lambda = 1/60$),

$$\mathbb{E}[T_{N_t}] = \mathbb{E}[\mathbb{E}[T_{N_t}|N_t = n]] = t\mathbb{E}[\frac{N_t}{1+N_t}] = t(1-\mathbb{E}[\frac{1}{1+N_t}]) = t\left(1 - \sum_{n\geq 0}\frac{1}{1+n}\frac{e^{-\lambda t}(\lambda t)^n}{n!}\right).$$

We can simplify the series at the end using the Poisson pmf, by

$$\sum_{n\geq 0}\frac{1}{1+n}\frac{e^{-\lambda t}(\lambda t)^n}{n!} = \frac{1}{\lambda t}\sum_{n\geq 0}\frac{e^{-\lambda t}(\lambda t)^{n+1}}{(n+1)!} = \frac{1}{\lambda t}\left(1-e^{-\lambda t}\right).$$

Hence, altogether, we have

$$\mathbb{E}[R_{T_{N_t}}] = \mathbb{E}[T_{N_t}] = t\left(1 - \frac{1}{\lambda t}\left(1 - e^{-\lambda t}\right)\right) = t\left(1 - \frac{60}{t}\left(1 - e^{-t/60}\right)\right).$$

(Extra page for Problem 5)

6 Waiting in Line [6 + 8 + 8 points]

Consider the M/M/1 queue with arrival rate $\lambda > 0$, and service rate $\mu > 0$. Recall that this refers to a queuing system for which the arrival process to the system is a rate-λ Poisson process, the service times are i.i.d. $\text{Exp}(\mu)$, and there is one server which serves customers on a first-in first-out basis.

a) Let $(X_t)_{t\geq 0}$ denote the number of people in the system at time $t \geq 0$. Draw the state transition diagram for this CTMC, with clearly labeled states and arrows labeled with nonzero transition rates.

b) Under what conditions does a stationary distribution exist? Assuming these conditions are met, what is the stationary distribution?

c) You are a customer who arrives to the system after it's been running for a very long time (i.e., assume the system is in steady-state). What is the expected time you spend in the system before departing (total time spent waiting in the queue, and in service)?

a) State space is $\{0, 1, \ldots\}$, and transition rates are $q_{n,n+1} = \lambda$, $q_{n,n-1} = \mu$, for all $n \geq 0$.

b) We should try to solve $\pi Q = 0$ to find a stationary distribution. This reads as:

$$\begin{aligned}
\lambda\pi(0) &= \mu\pi(1) \quad \Rightarrow \quad \pi(1) = \frac{\lambda}{\mu}\pi(0) \\
\mu\pi(2) &= (\mu+\lambda)\pi(1) - \lambda\pi(0) = \frac{\lambda^2}{\mu}\pi(0) \quad \Rightarrow \quad \pi(2) = \left(\frac{\lambda}{\mu}\right)^2 \pi(0) \\
\vdots & \qquad \vdots \\
\pi(n) &= \left(\frac{\lambda}{\mu}\right)^n \pi(0).
\end{aligned}$$

Now, to solve for $\pi(0)$ (and therefore π) we write

$$1 = \sum_{n\geq 0}\pi(n) = \pi(0)\sum_{n\geq 0}\left(\frac{\lambda}{\mu}\right)^n = \begin{cases} \frac{1}{1-\frac{\lambda}{\mu}} & \text{if } \lambda < \mu \\ +\infty & \text{if } \lambda \geq \mu. \end{cases}$$

Hence, $\pi Q = 0$ admits a solution (with π a probability vector) if and only if $\mu > \lambda$ (i.e., the service rate must exceed the arrival rate). Moreover, in this case, we see that

$$\pi(n) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n, \quad n \geq 0.$$

Thus, π is a geometric distribution with success probability $1 - \lambda/\mu$, supported on $\{0, 1, \ldots\}$.

c) Since we were told to assume the system is in steady state, we can assume $\lambda < \mu$. If the system has been running for a very long time, we can assume it is stationary. If I enter the system at time t and there are already X_t people in the system, then the time I spend in the system is

$$T = \sum_{i=1}^{X_t+1} T_i, \quad \text{where } (T_i)_{i \geq 1} \sim_{\text{i.i.d.}} \text{Exp}(\mu)\,, \text{ independent of } X_t.$$

Indeed, each person in already in the queue needs to be served before me (with $\text{Exp}(\mu)$ service times, including the person currently in service due to memorylessness), and then I need to be served myself. So, we compute:

$$\mathbb{E}[T] = \mathbb{E}[\mathbb{E}[T|X_t]] = \mathbb{E}[(X_t+1)\frac{1}{\mu}] = \frac{1}{\left(1-\frac{\lambda}{\mu}\right)\mu} = \frac{1}{\mu-\lambda},$$

where we used the steady-state assumption for t large to imply $(X_t+1) \sim \text{Geom}(1-\lambda/\mu)$, supported on $\{1, 2, \dots\}$.

The above answer considering the case $\lambda < \mu$ is good enough for full credit. For completeness, we note that if $\lambda \geq \mu$, then my average waiting time will be $+\infty$. Indeed, the expected waiting time is monotone increasing in λ for fixed μ (by thinning the arrival process), so the answer above implies the claim by taking $\lambda \uparrow \mu$.

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

(Extra page for Problem 6)

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

(This page has been intentionally left blank)