EECS 126 Probability and Random Processes — UC Berkeley, Fall 2019
Shyam Parekh — November 5, 2019

Midterm 2

Last Name	First Name	SID

Rules.

- No form of collaboration between the students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- You have 10 minutes to read the exam and 90 minutes to complete it.
- The exam is not open book. No calculators or phones allowed.
- Unless otherwise stated, all your answers need to be justified.

Problem	points earned	out of
Problem 1		36
Problem 2		20
Problem 3		30
Problem 4		26
Problem 5		14
Total		126

1 Assorted Problems [36]

(a) **Bipartite Markov Chain** [4]

Consider the following undirected bipartite Markov chain. At each time step, you pick one of your neighbors uniformly at random to transition to. If we start with an arbitary distribution π_0, then the distribution of this Markov Chain after k transitions will always converge to the stationary distribution as $k \to \infty$. Justify your answer.

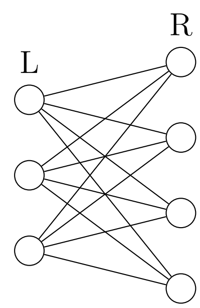


Figure 1: *

Complete Bipartite Graph for $m = 3, n = 4$

◯ True ◯ False

False. It is periodic with period 2, and thus does not always converge.

(b) **School Cancellations** [8]

In Berkeley, power outages happen according to a Poisson Process with a rate of λ_p and independently earthquakes happen according to a Poisson Process with a rate of λ_e. Marc Fisher cancels school with probability p_p if there is a power outage, and p_e if there is an earthquake. What is the expectation and the variance of the amount of time T between the previous school cancellation and the next school cancellation from today? You may assume that this trend has been going on since infinitely in the past.

The merged Poisson Process of school cancellations is PP($\lambda_p \cdot p_p + \lambda_e \cdot p_e$). By the random incidence paradox, the interarrival time of school cancellations is distributed as Exponential($\lambda_p \cdot p_p + \lambda_e \cdot p_e$)+Exponential($\lambda_p \cdot p_p + \lambda_e \cdot p_e$), where the two exponentials are independent.

$$E[T] = \frac{2}{\lambda_p \cdot p_p + \lambda_e \cdot p_e}$$
$$\text{var}(T) = \frac{2}{(\lambda_p \cdot p_p + \lambda_e \cdot p_e)^2}$$

(c) **Entropy of a Markov Chain [8]**

Consider a random walk along the integers (..., -3, -2, -1, 0, 1, 2, 3, ...). You start at 0 at time 0 and pick a direction (positive or negative with equal probability) to move in for time 1. At every time step after 1, you reverse direction with probability 0.25 and take a step in the new direction, or continue in the same direction otherwise. What is $H(X_0, \ldots, X_{n-1})$? You may leave your answer in terms of $H_b(p) = -p \log p - (1 - p) \log(1 - p)$.

The number of bits needed to encode is $H(X_0, \ldots, X_{n-1})$.

$$H(X_0, \ldots, X_{n-1}) = \sum_{k=0}^{n-1} H(X_k \mid X_{k-1}, \ldots, X_0)$$

We know that

- $H(X_0) = 0$ since we always start at 0.
- $H(X_1 \mid X_0) = H_b(\frac{1}{2}) = 1$ since it is equally likely for $X_1 = 1$ and $X_1 = -1$.
- $H(X_k \mid X_{k-1}, \ldots, X_0) = H(X_k \mid X_{k-1}, X_{k-2})$ since given the last two states we know exactly where we are and what direction we're moving in.
- $H(X_k \mid X_{k-1}, X_{k-2}) = H_b(0.25)$ since w.p 0.75, we continue in our direction and end up at $X_{k-1} + (X_{k-1} - X_{k-2})$, and w.p 0.25, we reverse direction and end up at X_{k-2}.

So the answer is $0 + 1 + (n - 2)H_b(0.25)$.

(d) **Poisson Process and Covariance [8]**

Consider a Poisson Process $\{N(s) : s \in [0, \infty)\}$ with rate λ.

Find the covariance of $N(t_1)$ and $N(t_2)$ for $t_2 > t_1 \geq 0$.

We have the crucial fact that $N(t_1), N(t_2) - N(t_1)$ are independent. Thus we have the following:

$$\begin{aligned}
\text{Cov}\Big(N(t_1), N(t_2)\Big) &= \text{Cov}\Big(N(t_1), N(t_1) + N(t_2) - N(t_1)\Big) \\
&= \text{Cov}\Big(N(t_1), N(t_1)\Big) + \text{Cov}\Big(N(t_1), N(t_2) - N(t_1)\Big) \\
&= \text{Var}\Big(N(t_1)\Big) \\
&= \lambda t_1
\end{aligned}$$

The last equality is due to the fact that $N(t_1) \sim \text{Poisson}(\lambda t_1)$.

(e) **Metropolis Hastings [8]**

Answer the following True/False questions about the MCMC lab. Briefly justify your answer **in one or two sentences**.

Metropolis Hastings allows us to generate random samples from a distribution $p(x)$ even if it's intractable to compute.

◯ True ◯ False

True. It does this by creating a Markov chain whose stationary distribution equals $p(x)$.

Any Metropolis Hastings Markov chain can be made aperiodic.

◯ True ◯ False

True. This can be done by giving every state a self loop with the same probability i.e. lazy chain from homework.

Burn in time is how long the Markov chain takes to converge to the state whose stationary distribution probability is maximal.

◯ True ◯ False

False. In general Markov chains don't converge to a single state. Burn in time is how long it takes for the initial distribution to converge to the stationary distribution.

In the Traveling Salesman Problem, you want to find the shortest path to visit n cities. Even though this problem has $n!$ possible answers and is NP-hard, we can approximately solve it with Metropolis Hastings. Our plan is to design a Markov chain such that if x^* is the best path, the stationary distribution probability $\pi(x^*)$ will be maximal. Then by running the Markov chain for a long time and looking at the most commonly visited state, we can infer the best path. If $L(x)$ is the length of a path x, then for this problem we could let our directly proportional estimate $f(x)$ be equal to $L(x)$.

◯ True ◯ False

False. We need $f(x)$ to be decreasing in $L(x)$. Here $f(x^*)$ would actual be minimal, when we need it to be maximal.

2 Convergence [20]

(a) **Convergent Balls [10]**

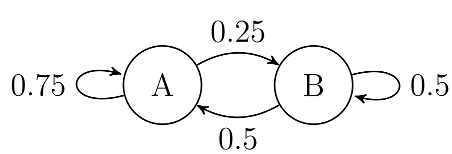
Kevin has a basket of k balls. Each ball is either white or red. Initially, all of the balls in his basket is red. Let $X_i^{(n)}$ denote the if the i-th ball is red at time step n. And let $Y_n = \sum_{i=1}^{k} X_i^{(n)}$ therefore be the number of balls in Kevin's basket that are red at time step n. At every time step, Kevin takes a ball out uniformly at random, and he replaces the ball with a white ball. Prove that Y_n converges in probability to 0.

$$\begin{aligned}
\lim_{n\to\infty} P\Big(|Y_n - 0| > \epsilon\Big) &= \lim_{n\to\infty} P\Big(Y_n \geq 1\Big) \\
&\leq \lim_{n\to\infty} \frac{E[Y_n]}{1} \\
&= \lim_{n\to\infty} \mathrm{E}\Big[\sum_{i=1}^{k} X_i^{(n)}\Big] \\
&= \lim_{n\to\infty} k\,\mathrm{E}[X_1^{(n)}] \\
&= \lim_{n\to\infty} k\Big(1 - \frac{1}{k}\Big)^n \\
&= 0
\end{aligned}$$

Many students attempted to split $P(Y_n \geq 1)$ into a product. However, it is important to note that the $X_i^{(n)}$ are not independent.

(b) **Markov Chain Groups [10]**

Consider the following Markov Chain



Given a finite length sequence $x_1, \ldots, x_n$ sampled from this Markov chain, we can form groups for samples that are in the same state. For example, if our sequence is

$$\underbrace{A, A, A}_{1}, \underbrace{B, B}_{2}, \underbrace{A, A}_{3}, \underbrace{B, B, B, B, B}_{4}$$

Then we have 4 groups of size 3, 2, 2, 5. Let G_n be

$$G_n = \frac{n}{\text{number of groups among } x_1, \ldots, x_n}$$

In the example, $G_{3+2+2+5} = G_{12} = \frac{12}{4} = 3$. What does G_n converge to?

G_n can be written as

$$G = \frac{1}{2k}\sum_{i=1}^{k}\Big(A_i + B_i\Big)$$

where $k = \frac{n}{2}$. A few details:

- A_i is the size of the i-th A-group. In the example, $A_1 = 3$ and $A_2 = 2$.
- B_i is the size of the i-th B-group. In the example, $B_1 = 2$ and $B_2 = 4$.

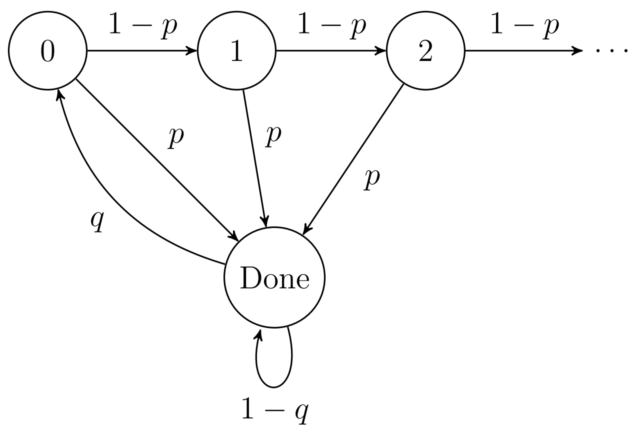
We know that

- $A_i \sim \text{Geometric}(0.25)$ and $B_i \sim \text{Geometric}(0.5)$.
- Because of the Markov Property, every A_i and B_i is independent of each other.
- There are some small issues with the last A_i not necessarily being geometric, as well as the fact that there might be one more A group than B group (or vice versa). However, they are negligible as $n \to \infty$.

By the SLLN, G_n converges almost surely to $\frac{E[A_1]+E[B_1]}{2} = \frac{\frac{1}{0.25}+\frac{1}{0.5}}{2} = 3$.

3 Discrete Time Markov Chains [30]

Consider the following Discrete Time Markov Chain.



(a) **Stationary Distribution [10]**

Find the stationary distribution of the chain.

Choosing cuts around all groups of vertices $\{0, 1 \dots i\}$ we get the following balance equations

$$(1-p)\pi(0) + p\pi(0) = q\pi(d) = \pi(0)$$

$$(1-p)\pi(1) + p\pi(1) + p\pi_0 = q\pi_d \Rightarrow \pi_1 = q\pi_d(1-p)$$

$$\dots$$

$$\pi_i = q(1-p)^i \pi_d$$

Normalizing the distribution we get

$$\pi_d = \frac{p}{q+p}$$

$$\pi_i = q(1-p)^i \frac{p}{q+p}$$

(b) **Hitting Time Backwards [10]**

What is the expected hitting time from state n to state 0?

$\frac{1}{p} + \frac{1}{q}$. You can observe this by seeing that if you start from n, the expected time to be done is the expectation of a geometric series. Moreover, the expected time to leave done is also a geometric series of parameter q!

(c) **Hitting Time Forward [10]**

Suppose $p = q = \frac{1}{2}$. What is the expected hitting time from state 0 to state n?

Hint: Use parts (a) and (b).

Let

- T_n^+ be the return time from state n to state n.
- $T_{0,n}$ be the hitting time from state 0 to state n.
- $T_{n,0}$ be the hitting time from state n to state 0.

Then

$$\begin{aligned} T_n^+ &= T_{n,0} + T_{0,n} \\ T_{0,n} &= T_n^+ - T_{n,0} \end{aligned}$$

Using our answers to part (a) and (b)

$$\begin{aligned} E[T_{0,n}] &= E[T_n^+] - E[T_{n,0}] \\ &= \frac{1}{\pi(n)} - \left(\frac{1}{p} + \frac{1}{q}\right) \\ &= \frac{1}{q(1-p)^n \frac{p}{p+q}} - \frac{1}{p} - \frac{1}{q} \end{aligned}$$

Substituting our values, we get

$$\begin{aligned} E[T_{0,n}] &= \frac{1}{\frac{1}{2}(1-\frac{1}{2})^n \frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}}} - \frac{1}{\frac{1}{2}} - \frac{1}{\frac{1}{2}} \\ &= 4 \cdot 2^n - 4 \\ &= 4(2^n - 1) \end{aligned}$$

Alternate Solution

Fact: If $x_{n+1} = ax_n + b$, then $x_n = a^n x_0 + \sum_{i=0}^{n-1} a^i b = a^n x_0 + b\frac{1-a^n}{1-a}$

Let $B(n)$ be the expected hitting time from state 0 to state n. We can set up the following recurrence relation.

$$B(n+1) = B(n) + 1 + p \cdot \left(\frac{1}{q} + B(n+1)\right)$$

The idea is that we need to reach n first and then make 1 more move. With probability $1-p$ we move to $n+1$, but with probability p, we go to Done, and then take in expectation $\frac{1}{q}$ steps before we can try again.

$$\begin{aligned}\Big(1-p\Big)B(n+1) &= B(n)+1+\frac{p}{q}\\ \Big(1-p\Big)B(n+1) &= B(n)+\frac{p+q}{q}\\ B(n+1) &= \frac{1}{1-p}B(n)+\frac{p+q}{q(1-p)}\end{aligned}$$

For our problem, this means

$$\begin{aligned}B(n+1) &= \frac{1}{1-\frac{1}{2}}B(n)+\frac{\frac{1}{2}+\frac{1}{2}}{\frac{1}{2}(1-\frac{1}{2})}\\ &= 2B(n)+4\end{aligned}$$

Applying the fact and knowing that $B(0)=0$, we get that

$$\begin{aligned}B(n) &= 2^n\cdot 0+4\cdot\frac{1-2^n}{1-2}\\ &= 4(2^n-1)\end{aligned}$$

4 Last Arrival [26]

Suppose we have a two independent Poisson Processes X and Y with arrival rates λ and 2λ, respectively.

(a) **X or Y? [6]**

What is the probability that the last arrival for Y comes after the last arrival for X in some given time interval $(0, t)$?

We intended this answer to be $\frac{2}{3}$ since we can consider the probability that the last arrival is from Y. However, based on how this question was worded there's a subtle issue: knowing that there were arrivals from both X and Y in a finite time interval changes the probability. If the question asked the probability that the first arrival from Y came before the first arrival from X (and not in some finite time interval), then the answer would indeed by $\frac{2}{3}$. To make it also true for the last arrivals, we need to make the Poisson Process reversible, which requires that we look at some point infinitely far in the future.

(b) **Expected Arrivals Afterwards [10]**

Suppose we are given that $N_Y(t) = n$. What is the expected number of arrivals N from X after the last arrival from Y on $(0, t)$?

Let T be the time between the n-th arrival time for Y and t. A couple observations:

- The n-th arrival time from Y is distributed as the maximum of n independent $U[0, t]$ RVs.
- The number of arrivals from X in any time interval of length T is Poisson with parameter λT.

Therefore,

$$\begin{aligned} E[N] &= E[E[N \mid T]] \\ &= E[\lambda T] \\ &= \lambda E[T] \\ &= \lambda(t - \frac{n}{n+1}t) \\ &= \frac{\lambda}{n+1}t \end{aligned}$$

(c) **Chernoff Bound [10]**

Provide an upper bound using the Chernoff bound on the probability that the total number of arrivals from X and Y in the interval $(0, t)$ exceed $k\lambda t$ for some positive integer k. Express your answer in terms of k and simplify your answer as much as possible.

Hint: The MGF for a Poisson random variable X with parameter λ is $M_X(s) = \exp(\lambda(e^s - 1))$.

$$\begin{aligned} P(X+Y \geq k\lambda t) &\leq \min_s \frac{M_{X+Y}(s)}{e^{ks\lambda t}} \\ &= \min_s \frac{\exp(3\lambda t(e^s - 1))}{e^{ks\lambda t}} \\ &= \min_s \exp(\lambda t(3e^s - ks - 3)) \end{aligned}$$

The find the optimal s value, we can just minimize the exponent since exp is monotonically increasing.

$$\frac{d}{ds}\lambda t(3e^s - ks - 3) = 0 = 3\lambda t e^s - k\lambda t$$

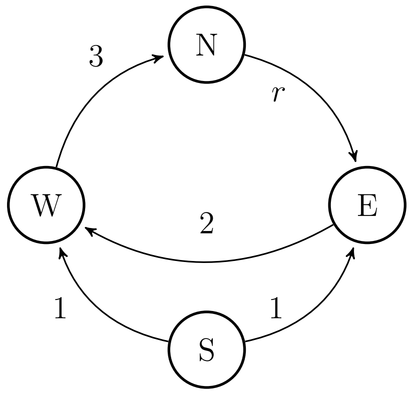
We can set $s = \ln \frac{k}{3}$ to satisfy the equation.

Now, we plug s to get our upper bound.

$$P(X+Y \geq k\lambda t) \leq \exp(\lambda t(k - k\ln\frac{k}{3} - 3))$$

5 Continuous Time Markov Chains [14]

Consider the following CTMC for $r > 0$.



(a) **Stationary Distribution [7]**

For the above CTMC, find the stationary distribution in terms of r, where $r > 0$.

$\pi(S) = 0$ since it's transient. Considering just N, E, and W, we know that the jump chain has stationary distribution $\frac{1}{3}$ in every state. To convert this stationary distribution to that of the CTMC, we multiply $\frac{1}{3}$ by the expected time we stay in each state before leaving and then re-normalize. So it turns out

$$\begin{aligned}\pi &= \frac{1}{\frac{1}{r}+\frac{1}{2}+\frac{1}{3}}\begin{bmatrix}\frac{1}{r} & \frac{1}{2} & 0 & \frac{1}{3}\end{bmatrix}\\ &= \frac{6r}{6+2r+3r}\begin{bmatrix}\frac{1}{r} & \frac{1}{2} & 0 & \frac{1}{3}\end{bmatrix}\end{aligned}$$

Alternate Solution

Order our states N, E, S, W. Our Q matrix is the following.

$$Q = \begin{bmatrix} -r & r & 0 & 0 \\ 0 & -2 & 0 & 2 \\ 0 & 1 & -2 & 1 \\ 3 & 0 & 0 & -3 \end{bmatrix}$$

Use columns 1, 2, 3 and the normalizing equation $\sum \pi(i) = 0$ to get the following system.

$$\begin{aligned} -r\cdot\pi(N) + 3\pi(W) &= 0 \\ r\cdot\pi(N) - 2\pi(E) + \pi(S) &= 0 \\ -2\pi(S) &= 0 \\ \pi(N)+\pi(E)+\pi(S)+\pi(W) &= 1 \end{aligned}$$

- $\pi(S)$ must be 0.

- Equation 1 tells us that $\pi(N) = \frac{3}{r}\pi(W)$.
- Adding equation 1 and 2 tells us that $3\pi(W) = 2\pi(E)$ or that $\pi(W) = \frac{2}{3}\pi(E)$.

Combining into equation 4 we get that

$$\begin{aligned}
\frac{3}{r} \cdot \frac{2}{3} \cdot \pi(E) + \frac{2}{3} \cdot \pi(E) + \pi(E) &= 1 \\
\left(\frac{2}{r} + \frac{2}{3} + 1\right)\pi(E) &= 1 \\
\frac{6 + 2r + 3r}{3r}\pi(E) &= 1 \\
\pi(E) &= \frac{3r}{6 + 2r + 3r}
\end{aligned}$$

This leads to a final stationary distribution of

$$\pi = \begin{bmatrix} \frac{6}{6+2r+3r} & \frac{3r}{6+2r+3r} & 0 & \frac{2r}{6+2r+3r} \end{bmatrix}$$

(b) **Equivalent DTMC [7]**

Draw a DTMC with the same state space with the same stationary distribtion as the above CTMC. If necessary, draw multiple DTMCs depending on the value of r.

For any $q \geq \sup q(i)$, $P = \frac{Q}{q} + I$ defines a valid DTMC probability transition matrix with the same stationary distribution. Let our $q = r + 3$. Then

$$P = \begin{bmatrix} \frac{3}{r+3} & \frac{r}{r+3} & 0 & 0 \\ 0 & \frac{r+1}{r+3} & 0 & \frac{2}{r+3} \\ 0 & \frac{1}{r+3} & \frac{r+1}{r+3} & \frac{1}{r+3} \\ \frac{3}{r+3} & 0 & 0 & \frac{r}{r+3} \end{bmatrix}$$

is the DTMC we're looking for.