EEFCS 126 Probability and Random Processes University of California, Berkeley: Spring 2019
Kannan Ramchandran April 9, 2019

Midterm 2

Last Name	First Name	SID
Zhang	Albert	3034019226

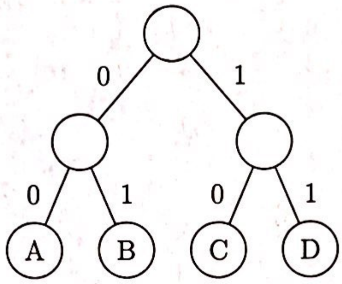
Rules.

- 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 100 minutes to complete it.
- The exam is not open book; we are giving you a cheat sheet. No calculators or phones allowed.
- Unless otherwise stated, all your answers need to be justified. Show all your work to get partial credit.
- Maximum you can score is 114 but 100 points is considered perfect.

Problem	points earned	out of
Problem 1		44
Problem 2		20
Problem 3		32
Problem 4		18
Total		114

Problem 1: Answer these questions briefly but clearly.

(a) [4] Which of the following frequencies for A,B,C and D can generate the following Huffman tree? (Select all that apply.)



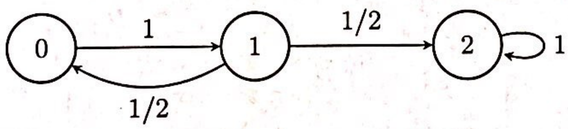
○ $p_A = 0.4$, $p_B = 0.3$, $p_C = 0.2$, $p_D = 0.1$

● $p_A = 0.35$, $p_B = 0.25$, $p_C = 0.2$, $p_D = 0.2$

● $p_A = 0.25$, $p_B = 0.25$, $p_C = 0.25$, $p_D = 0.25$

○ $p_A = 0.2$, $p_B = 0.35$, $p_C = 0.2$, $p_D = 0.25$

(b) [2+2+2] Consider the the Markov Chain $(X_n)_{n\in\mathbb{N}}$ whose transitions are given by



1. X_n converges almost surely. True False

   If true, it converges a.s. to (N/A if false): 2

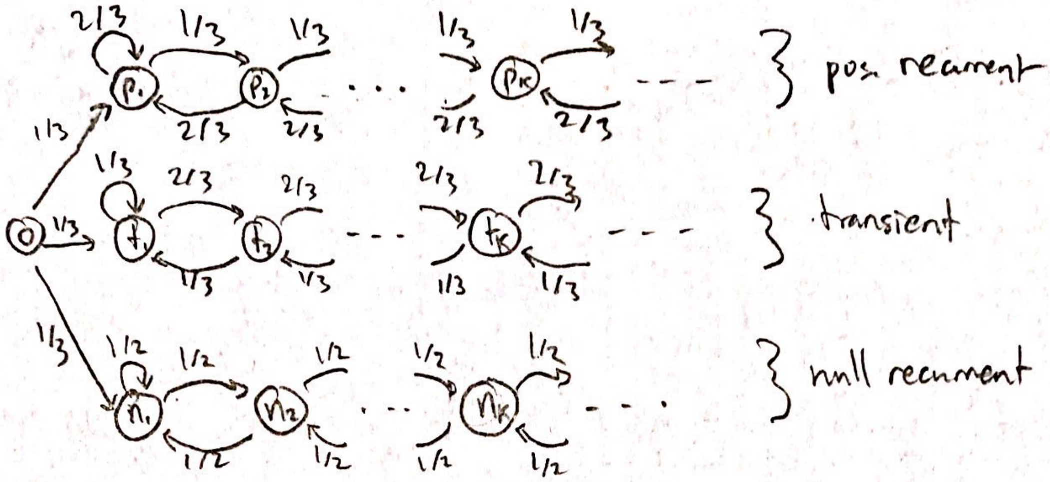
2. X_n converges in probability. True False

   If true, it converges i.p. to (N/A if false): 2

3. X_n converges in distribution. True False

   If true, it converges in distribution to (N/A if false): [0,0,1]

(c) [6] Construct a reversible Markov Chain which has one positive recurrent class, one transient class, and one null recurrent class.



(d) [6] Random Walk on a Random Graph: A particle performs a random walk on a graph with 3 vertices, labeled 1, 2 and 3, starting from state 1. However, at each time-step, before the particle makes a move, the edges of the graph are re-sampled according to an independent $G(3, 1/2)$ distribution. Once the edges have been sampled, the particle chooses a neighbor of its current state uniformly at random from the vertices connected to it. What is the expected number of time-steps before the particle hits state 3?

(Note: $G(n,p)$ refers to the Erdos-Renyi random graph on n vertices where each edge exists independently with probability p, as introduced in lecture.)

12

$\frac{1}{4}\cdot 1+\frac{1}{4}\cdot\frac{1}{2}+\frac{1}{4}\cdot 1+\frac{1}{4}\cdot 0=\frac{5}{8}$

$\frac{1}{4}\cdot 1+\frac{1}{4}\cdot\frac{1}{2}+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0=\frac{3}{8}$

3

$$\mathbb{E}[\text{hitting time}] = \mathbb{E}[\text{Geom}(3/8)] = \boxed{\frac{8}{3}}$$

(e) [3+3] Interarrival Times: A factory's production line outputs items according to a Poisson Process with rate λ.

1. If each item is defective with probability 1/3, what is the distribution of the time between the arrivals of two successive defective items? What is its expectation?

$\text{Poisson}(\lambda) \longrightarrow \text{Poisson}(\lambda/3)$

$\Rightarrow X \sim \text{Expo}(\lambda/3), \quad \mathbb{E}[X] = 3/\lambda$

2. If every third item is defective, what is the distribution of the time between the arrivals two successive defective items? What is its expectation?

$\Rightarrow X \sim \text{Erlang}(3, \lambda)$

$\mathbb{E}[X] = 3\mathbb{E}[\text{Expo}(\lambda)] = 3/\lambda$

(f) [2+2+2] Convergence in Probability: Let $X_1, ..., X_n$ be independent continuous random variables Uniform in [0,1]. Let $Y_n = \max(\{X_1, ..., X_n\})$, and $Z_n = \max(\{X_1, ..., X_n\} \backslash Y_n)$. In other words Y_n is the largest element of the n variables and Z_n is the second largest element.

1. What is $\mathbb{P}(Y_n < 1 - \epsilon)$, for some $\epsilon > 0$?

$\prod_{i=1}^{n} \mathbb{P}[X_i < 1-\varepsilon] = (1-\varepsilon)^n$

2. What is $\mathbb{P}(Z_n < 1 - \epsilon)$, for some $\epsilon > 0$?

$\binom{n}{1}\mathbb{P}[\text{only largest } X_i \text{ in } (1-\varepsilon, 1)] + \mathbb{P}[\text{all } X_i \text{ in } (0, 1-\varepsilon)]$

$= n(1-\varepsilon)^{n-1}\varepsilon + (1-\varepsilon)^n$

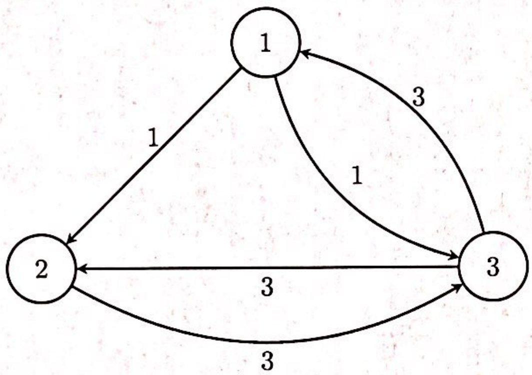
3. Show that $Y_n - Z_n$ converges in probability to 0.
Hint: You may use the fact that if $A_n \xrightarrow{\mathbb{P}} a$ and $B_n \xrightarrow{\mathbb{P}} b$ for constants a and b, then $A_n + B_n \xrightarrow{\mathbb{P}} a + b$

$\lim_{n\to\infty}(1-\varepsilon)^n = 0$, so $\mathbb{P}[Y_n < 1-\varepsilon] \xrightarrow{n\to\infty} 0 \Rightarrow Y_n \xrightarrow{\mathbb{P}} 1$

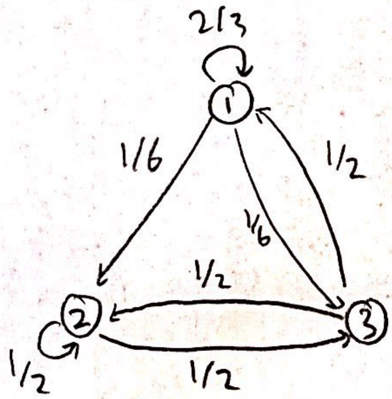
$\lim_{n\to\infty}\left(n(1-\varepsilon)^{n-1}\varepsilon + (1-\varepsilon)^n\right) = 0$, so $\mathbb{P}[Z_n < 1-\varepsilon] \xrightarrow{n\to\infty} 0 \Rightarrow Z_n \xrightarrow{\mathbb{P}} 1$

Thus, using Hint, we have $Y_n - Z_n \xrightarrow{\mathbb{P}} 1 - 1 = 0$.

(g) [4] **Simulated CTMC:** Consider the Continuous-time Markov Chain (CTMC) shown below.



Construct a Discrete-time Markov Chain (DTMC) that has the same stationary distribution as the above chain.



(h) [6] **Cascaded BEC channel:** It is desired to transmit reliably over a composite channel comprising a cascade of two back-to-back BEC(p) channels (Binary Erasure Channels with erasure probability equal to p). What is the capacity (i.e. maximum rate at which you can transmit reliably) of this composite channel? (Recall that the capacity of a single BEC(p) channel is $(1-p)$ bits per channel use.)

$$p' = 1-(1-p)^2 = 2p-p^2$$

$$\sim \mathrm{BEC}(2p-p^2)$$

$$\Rightarrow \text{capacity is } \boxed{1-2p+p^2}$$

Problem 2 [20]: Frisbee Attempts

The probability that a frisbee player will catch the frisbee depends on the results of her last two attempts, and is given by

$$\mathbb{P}(\text{catch}) = \begin{cases} 1/2, & \text{if she dropped during both of her last two attempts} \\ 2/3, & \text{if she catches in exactly one of the two the last two attempts} \\ 3/4, & \text{if she catches in both of her last two attempts.} \end{cases}$$

(a) [3] Let $X_i = 1$ if you just caught the frisbee, and $X_i = 0$ otherwise; show with the help of an example that X_i is not a Markov chain.

If $X_1 = 1$, $X_2 = 1$

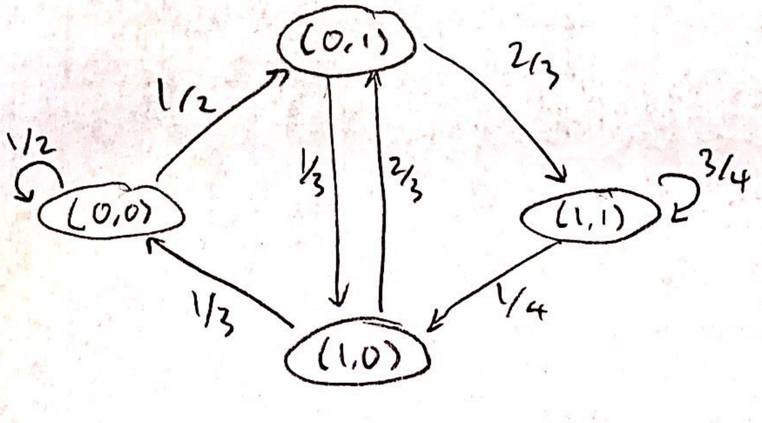
Then $\mathbb{P}[X_3=1 \mid X_2=1] = \mathbb{P}[X_3=1 \mid X_2=1, X_1=1] = 3/4$

If $X_1 = 0$, $X_2 = 1$ $\qquad \neq \Rightarrow$ not a valid MC.

Then $\mathbb{P}[X_3=1 \mid X_2=1] = \mathbb{P}[X_3=1 \mid X_2=1, X_1=0] = 2/3$

(b) [5] Let the states be denoted by ordered tuples (X_i, X_{i+1}); argue that this is a Markov chain. Draw the transition diagram.

Since each transition is dependent only on the previous "state" (the previous <u>two</u> attempts), this model satisfies the Markov Property.



(c) [5] Find the stationary distribution.

$\pi_{00} = \frac{1}{2}\pi_{00} + \frac{1}{3}\pi_{10}$

$\pi_{10} = \frac{1}{3}\pi_{01} + \frac{1}{4}\pi_{11}$

$\pi_{01} = \frac{1}{2}\pi_{00} + \frac{2}{3}\pi_{10}$

$\pi_{11} = \frac{2}{3}\pi_{01} + \frac{3}{4}\pi_{11}$

$\pi_{00} = \frac{2}{3}\pi_{10}$

$\pi_{01} = \pi_{10} = \frac{3}{2}\pi_{00}$

$\pi_{11} = \frac{8}{3}\pi_{01} = 4\pi_{00}$

$1 = \pi_{00} + \frac{3}{2}\pi_{00} + \frac{3}{2}\pi_{00} + 4\pi_{00} = 8\pi_{00}$

$\Rightarrow \pi_{00} = \frac{1}{8} \Rightarrow \pi = [\pi_{00}\ \pi_{01}\ \pi_{10}\ \pi_{11}] = [\frac{1}{8}\ \frac{3}{16}\ \frac{3}{16}\ \frac{1}{2}]$

(d) [4] Calculate the long-term fraction of times she catches the frisbee.

$\frac{1}{8}\cdot 0 + \frac{3}{16}\cdot 1 + \frac{3}{16}\cdot 0 + \frac{1}{2}\cdot 1 = \boxed{\frac{11}{16}}$

(e) [3] Is this chain reversible? Justify your answer.

Not Reversible; sufficient to check local balance fails for one equation:

$\frac{1}{16} = \pi_{00}\cdot\frac{1}{2} \neq \pi_{01}\cdot 0 = 0.$

Problem 3 [32]: please do this thanks ray

Students ask logistical questions on Piazza according to a Poisson process of rate λ. Ray checks Piazza according to an independent Poisson process of rate μ. Every time Ray checks Piazza, he answers all logistical questions instantaneously.

(a) [3] Let's say Ray checked Piazza for the first time at time t, and there were n unanswered logistical questions. What is the expected time when the first question showed up?

n questions are uniformly distributed in $[0,t]^n$;

So $\mathbb{E}[1^{st}\ ?] = \boxed{\frac{t}{n-1}}$

(b) [4] Ray just finished answering logistical questions, what is the expected number of logistical questions he will have to answer next time he checks Piazza?

$\mathbb{P}[n^{th}\ ?\ \text{comes before check} \mid (n-1)?\text{'s unanswered}] = \frac{\lambda}{\lambda+\mu}$

so # of ?'s Ray will answer is $X \sim \text{Geo}\left(\frac{\mu}{\lambda+\mu}\right) - 1$

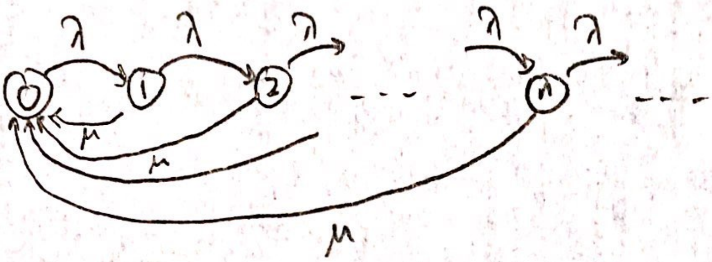
$\mathbb{E}[X] = \frac{\lambda+\mu}{\mu} - 1 = \boxed{\frac{\lambda}{\mu}}$

(c) [5] What is the expected number of times Ray will check Piazza before he is greeted with a nightmare of $\geq n$ unanswered logistical questions (including the time he's greeted with the nightmare)?

$$\sum_{k=0}^{n-1} (1-p)^k p = p\left(\frac{1-(1-p)^n}{p}\right) = 1-(1-p)^n \Rightarrow \text{Geom}((1-p)^n)$$

$$\mathbb{E}[\text{Geom}((1-p)^n)] = \boxed{\frac{1}{(1-p)^n}}$$

(d) **[3]** We can model the number of outstanding logistical requests as a CTMC with the natural numbers as the state space (0, 1, 2, ...). Draw this CTMC.



(e) **[6]** Find the long term fraction of time for which there are no outstanding logistical questions.

$$\lambda \pi_0 = \mu \left(\sum_{i=1}^{\infty} \pi_i \right) \qquad \sum_{i=0}^{\infty} \pi_0 \left(\frac{\lambda}{\lambda+\mu}\right)^i = 1$$

$$\Longleftrightarrow \frac{1}{1-\frac{\lambda}{\lambda+\mu}} \pi_0 = \frac{\lambda+\mu}{\mu} \pi_0 = 1$$

$$(\lambda+\mu)\pi_i = \lambda \pi_{i-1} \quad \text{for } i = 1, 2, 3, \dots \qquad \pi_0 = \frac{\mu}{\lambda+\mu}$$

$$\pi_i = \frac{\lambda}{\lambda+\mu} \pi_{i-1} \Rightarrow \pi_i = \left(\frac{\lambda}{\lambda+\mu}\right)^i \pi_0$$

$$\lambda \pi_0 = \mu \left(\sum_{i=1}^{\infty} \left(\frac{\lambda}{\lambda+\mu}\right)^i \pi_0 \right) = \cancel{\mu} \left(\frac{\frac{\lambda}{\lambda+\mu}}{\frac{\mu}{\lambda+\mu}} \right) \pi_0 = \lambda \pi_0$$

Thus $\pi_i = \left(\frac{\mu}{\lambda+\mu}\right)\left(\frac{\lambda}{\lambda+\mu}\right)^{i-1}$

(f) **[3]** Is this CTMC positive recurrent, null recurrent or transient? (If it depends on λ, μ, specify how so).

positive recurrent

(g) [8] A student just posted the third logistical question since Ray's last check. What is the expected time until the next moment when there are exactly two outstanding logistical questions?

\# ?'s asked after 3^{rd} ? before check

$N \sim \text{Geom}\left(\frac{\mu}{\lambda+\mu}\right) - 1$

$\mathbb{E}[\text{time til check}] = \mathbb{E}[N \cdot \text{Expo}(\lambda)] = \frac{\lambda}{\mu} \cdot \frac{1}{\lambda} = \frac{1}{\mu}$

$\tau_0 = \frac{1}{\lambda} + \tau_1 = \frac{1}{\lambda} + \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\tau_0 \Rightarrow \tau_0 = \frac{\lambda+\mu}{\lambda}\left(\frac{2\lambda+\mu}{\lambda(\lambda+\mu)}\right) = \frac{2\lambda+\mu}{\lambda^2}$

$\tau_1 = \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\tau_0 + \cancel{\frac{\lambda}{\lambda+\mu}\tau_2}$

$\tau_2 = 0$

Overall, we have $\mathbb{E}[\text{Time to } 2\text{?'s}] = \frac{1}{\mu} + \frac{2\lambda+\mu}{\lambda^2}$

Problem 4 [18]: Plants vs. Zombies

Efe decides the current version of Plants vs. Zombies is too easy for him, so he undertakes the task of writing a new version of the game. He's currently writing the randomizer for zombies entering the screen, and needs your help to analyze how many zombies will appear on each level. Suppose that at the beginning of each level Efe creates $Y \sim \text{Poisson}(\mu)$ zombie generators, and each zombie generator independently creates zombies according to a Poisson Process with parameter λ. Each level lasts for T seconds, where T is fixed. Let N be the number of zombies generated in a given level.

(a) [6] Find $\mathbb{E}[N]$.

$$\mathbb{E}[Y \cdot \text{Poisson}(\lambda T)] = \mathbb{E}[Y]\mathbb{E}[\text{Poisson}(\lambda T)]$$
$$= \mu \cdot \lambda T$$

(b) [6] Find $\text{var}(N)$.

$$\text{Var}(N) = \text{var}(\mathbb{E}[N \mid Y]) + \mathbb{E}[\text{var}(N \mid Y)]$$
$$= \text{var}(Y \cdot \lambda T) + \mathbb{E}[Y \text{var}(\text{Poisson}(\lambda T))]$$
$$= (\lambda T)^2 \text{var} Y + \mathbb{E}(\lambda T Y)$$
$$= (\lambda T)^2 \mu + \lambda T \mu$$
$$= \boxed{\lambda T \mu (\lambda T + 1)}$$

(c) [6] Now suppose that Efe runs his simulation for 100 levels. Assume you are told that $\lambda < 2$ zombies/sec and $\mu < 5$. Using the CLT and $\sum_{i=1}^{100} N_i/n$ as an estimator, construct a 95% confidence interval for $\mathbb{E}[N]$.
Hint: For $Z \sim \mathcal{N}(0,1), \mathbb{P}(-1.96 \le Z \le 1.96) = 0.95$.

$$\mathbb{P}\left[\left|\sum_{i=1}^{100} \frac{N_i}{100} - \mathbb{E}[N]\right| > 2\sigma\right)$$

$$\cancel{= \mathbb{P}\left[\left|\mathcal{N}\left(0, \frac{\text{Var}N}{100}\right)\right| > 2\sigma\right]}$$

$$= \boxed{\mathbb{P}\left[\left|\sum_{i=1}^{100} \frac{N_i}{100} - \mathbb{E}[N]\right| < 2\sqrt{\frac{11}{10}T}\right] \geq 0.95}$$

$$\text{Var}\left(\sum_{i=1}^{100} \frac{N_i}{100}\right) = \frac{\text{Var}\,N}{100}$$

$$= \frac{\lambda T \mu (\lambda T + 1)}{100}$$

$$< \frac{10T(10+1)}{100}$$

$$= \frac{11}{10}T.$$