EECS 126 Probability and Random Processes University of California, Berkeley: Spring 2017
Kannan Ramchandran March 21, 2017

Midterm Exam 2 (**Solutions**)

Last name	First name	SID

Name of student on your left:
Name of student on your right:

- DO NOT open the exam until instructed to do so.
- The total number of points is **110**, but a score of $\geq$ **100** is considered perfect.
- You have 10 minutes to read this exam without writing anything and 105 minutes to work on the problems.
- Box your final answers.
- Partial credit will not be given to answers that have no proper reasoning.
- **Remember to write your name and SID on the top left corner of every sheet of paper.**
- **Do not write on the reverse sides of the pages.**
- All eletronic devices must be turned off. Textbooks, computers, calculators, etc. are prohibited.
- No form of collaboration between students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- **You must include explanations to receive credit.**

Problem	Part	Max	Points	Problem	Part	Max	Points
1	(a)	20		2		15	
	(b)	12		3		20	
	(c)	8		4		25	
	(d)	10					
Total						110	

Problem 1. (a) (20 points, 4 points each part) Evaluate the statements with *True* or *False*. Give brief explanations in the provided boxes. Anything written outside the boxes will not be graded.

(1) Given a random variable $M \sim \text{Geom}(\frac{1}{10})$ and i.i.d. exponential random variables $X_i \sim \text{Exp}(1)$, the distribution of the sum $X_1 + X_2 + \cdots + X_M$ is Erlang of order 10 with rate 1.

True or False: False
Explanation: It is exponential

(2) Recall fountain codes as introduced in Lab 4. You would like to recover the packets x_1, x_2, x_3, x_4 and receive the following equations:

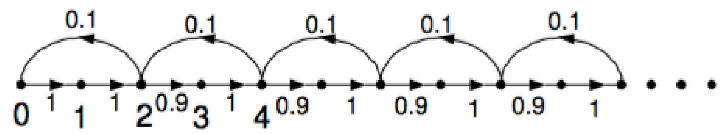
$$\begin{aligned} &x_2 \oplus x_4 = 1 \\ &x_1 = 1 \\ &x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 1 \\ &x_3 \oplus x_4 = 1 \end{aligned}$$

The peeling decoder is able to recover all the packets that can possibly be recovered.

True or False: False
Explanation: An optimal decoder could recover all the packets, but peeling can recover only one

(3) Give an example of a discrete time Markov Chain which is transient and has period 3.

Solution:



(4) Consider a Poisson process with rate 1. Suppose that the 117th arrival comes at time T. The joint distribution of the first 116 arrivals is the same as the joint distribution of $U_1, U_2, \ldots, U_{116}$, where U_i are i.i.d. $U[0,T]$ random variables.

True or False: False
Explanation: It has the joint distribution of the order statistics of 116 uniform random variables.

(5) For a random variable X with a well-defined $M_X(s)$, we have:

$$\mathrm{Var}(X) = \frac{\mathrm{d}^2}{\mathrm{d}s^2} \ln M_X(s)\bigg|_{s=0}$$

True or False: True
Explanation:

(b) (12 points) After attending an EE126 lecture, you went back home and started playing *Twitch Plays Pokemon*. Suddenly, you realized that you may be able to analyze *Twitch Plays Pokemon*.

Figure 1: A snapshot of 'Twitch Plays Pokemon' - 1

(i) (6 points) The player in the top left corner performs a random walk on the 8 checkered squares and the square containing the stairs. At every step the player is equally likely to move to any of the squares in the four cardinal directions (North, West, East, South) if there is a square in that direction. Find the expected number of moves until the player reaches the stairs in Figure 1.

Solution:

Using symmetry, the 9 states can be grouped as follows.

a	b	c
b	d	e
c	e	f

Now, you can also observe that state d is equivalent to state c.

a	b	c
b	c	e
c	e	f

With the above states, one can write down the following first-step equations.

$$T_a = 1 + T_b$$
$$T_b = 1 + \frac{1}{3}T_a + \frac{2}{3}T_c$$
$$T_c = 1 + \frac{1}{2}T_b + \frac{1}{2}T_e$$
$$T_e = 1 + \frac{2}{3}T_c + \frac{1}{3}T_f$$
$$T_f = 0$$

Solving the above equations gives:

$$T_a = 18,\ T_b = 17,\ T_c = 15,\ T_d = 11$$

Thus, the player has to make 18 moves to go downstairs on average.

Figure 2: A snapshot of 'Twitch Plays Pokemon' - 2

(ii) (6 points) The player randomly walks in the same way as in the previous part. Find the probability that the player reaches the stairs in the bottom right corner in Figure 2.

Solution:

Consider 9 initial states and corresponding probabilities of reaching the 'good' stairs as follows. Using symmetry, one can obtain the following table.

p	$\frac{1}{2}$	$1-p$
q	$\frac{1}{2}$	$1-q$
0	$\frac{1}{2}$	1

With the above probabilities, one can write down the following first-step equations.

$$p = \frac{1}{2}q + \frac{1}{2} \cdot \frac{1}{2}$$
$$q = \frac{1}{3}p + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0$$

Solving the above equations gives:

$$p = 0.4, \ q = 0.3$$

Thus, Red is going to reach the good stairs with probability 0.4.

(c) (8 points) Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables and $X_i \sim U[-1,1]$. Does the sequence $Y_n = \max\{X_1, X_2, \ldots, X_n\}$ converge in probability? If so, what does it converge to?

Solution:

Consider $\epsilon \in [0,1]$. We see that:

$$\begin{aligned}\Pr(|Y_n - 1| \geq \epsilon) &= \Pr(\max\{X_1, \ldots, X_n\} \leq 1 - \epsilon) \\ &= \Pr(X_1 \leq 1 - \epsilon, \ldots, X_n \leq 1 - \epsilon) \\ &= \Pr(X_i \leq 1 - \epsilon)^n = \left(1 - \frac{\epsilon}{2}\right)^n\end{aligned}$$

Thus, $\Pr(|Y_n - 1| \geq \epsilon) \to 0$ and we are done.

(d) (10 points) Nodes in a wireless sensor network are dropped randomly according to a two-dimensional Poisson process. A two-dimensional Poisson process is defined as follows:

(i.) For any region of area A, the number of nodes in that region has a Poisson distribution with mean λA,

(ii.) The number of nodes in non-overlapping regions is independent.

Suppose that you are at an arbitrary location in the plane. Let X be the distance to the nearest node in the sensor network. Find $E[X]$. (Here distance between the points (x_1, y_1) and (x_2, y_2) is defined as $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$)

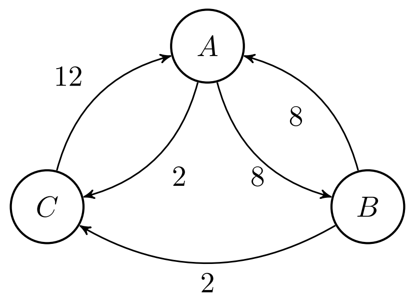
Solution: Given an arbitrary location, $X > t$ if and only if there are no special points in the circle of radius t around the given point. The expected number in that circle is $\lambda \pi t^2$, and since the number in that circle is Poisson with expected value $\lambda \pi t^2$, the probability that number is 0 is $e^{-\lambda \pi t^2}$. Thus $\Pr(X > t) = e^{-\lambda \pi t^2}$. Since $X \geq 0$, we have

$$E[X] = \int_0^\infty \Pr(X > t) dt = \int_0^\infty e^{-\lambda \pi t^2} dt$$

We can look this up in a table of integrals, or recognize its resemblance to the Gaussian PDF. If we define $\sigma^2 = 1/(2\pi\lambda)$, the above integral is

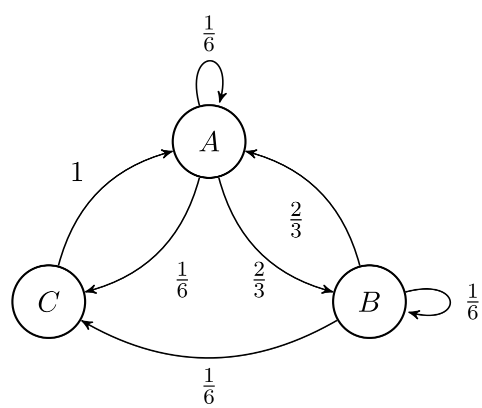
$$E[X] = \sigma\sqrt{2\pi} \int_0^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} dt = \frac{\sigma\sqrt{2\pi}}{2} = \frac{1}{2\sqrt{\lambda}}.$$

Problem 2. (15 points) Consider the continuous time Markov chain given in the figure.



(a.) (7 points) Draw a DTMC (discrete time Markov Chain) which has the same stationary distribution as the given CTMC and find the stationary distribution.

Solution:



The flow equations are given by:

$$
\begin{aligned}
10\pi_A &= 8\pi_B + 12\pi_C \\
12\pi_C &= 2\pi_A + 2\pi_B \\
10\pi_B &= 8\pi_A \\
\pi_A + \pi_B + \pi_C &= 1
\end{aligned}
$$

Thus, $\pi_A = \frac{10}{21}, \pi_B = \frac{8}{21}, \pi_C = \frac{1}{7}$

(b.) (8 points) Let τ denote the first time the CTMC is in state C after starting from state A at time 0. Find $M_\tau(s)$, the MGF of τ.

Solution:

Let T_C, T_B be the events that the first transition is to state C and state B, respectively:

$$\begin{aligned} E[e^{s\tau}] &= E[e^{s\tau}|T_C]P(T_C) + E[e^{s\tau}|T_B]P(T_B) \\ &= \frac{2}{2-s}\frac{1}{5} + E[e^{s\tau}]\frac{4}{5} \end{aligned}$$

Where in the last equality, we note that the hitting time to state C from state A is the same as the hitting time to state C from state B. Thus, $M_\tau(s) = \frac{2}{2-s}$.

Problem 3. (20 points)

(a.) (10 points) Batteries from Company X last an amount of time exponentially distributed with rate λ. The company guarantees that $\lambda \geq 2$. Suppose that you have used n batteries from Company X and observed their lifetimes to be $X_1, X_2, \ldots, X_n$. Using Chebyshev's inequality, find the minimum n required to construct a confidence interval for the mean interval of the light bulb such that it has tolerance at most ϵ and confidence at least $1-\delta$. In other words, we want to find n such that

$$P\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \frac{1}{\lambda}\right| \geq \varepsilon\right) \leq \delta$$

Solution:

We require

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{\lambda}\right| \geq \varepsilon\right) \leq \delta.$$

By Chebyshev's Inequality, the above inequality will be satisfied so long as

$$\frac{\mathrm{Var}(n^{-1}\sum_{i=1}^n X_i)}{\varepsilon^2} = \frac{\mathrm{Var}(X_i)}{n\varepsilon^2} = \frac{1}{n\lambda^2\varepsilon^2} \leq \delta.$$

From our trustworthy source, we have the lower bound $\lambda \geq 2$, so we are happy if we choose

$$n \geq \frac{1}{4\delta\varepsilon^2}.$$

(b.) (10 points) Suppose now that you have taken 10000 samples of the batteries. With the same tolerance level ϵ, use the CLT to find your new confidence (you may again use the company guarantee that $\lambda \geq 2$). You may leave your answer in terms of Φ, the CDF of the standard normal distribution.

Solution: We now assume $n^{-1}\sum_{i=1}^n X_i \sim \mathcal{N}(\lambda^{-1}, n^{-1}\lambda^{-2})$. We now seek to bound the probability of an ε-deviation from the mean:

$$\begin{aligned}
P\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{\lambda}\right| \geq \varepsilon\right) &\approx P(|\mathcal{N}(\lambda^{-1}, n^{-1}\lambda^{-2}) - \lambda^{-1}| \geq \varepsilon) \\
&= P(|\mathcal{N}(0, n^{-1}\lambda^{-2})| \geq \varepsilon) \\
&= P(|\mathcal{N}(0,1)| \geq \varepsilon\lambda\sqrt{n}) \leq 2\Phi(-200\varepsilon).
\end{aligned}$$

Problem 4. (25 points, 5 points each) There are a set of M bins and a set of balls labeled $f_1, f_2, \ldots$. Let $f_{(i)} = \{f_j : j \equiv i \mod M\}$ be the set of balls thrown to bin i. The interarrival time between balls thrown to bin i is exponential with rate λ balls per second, and all interarrival times are independent. The situation is illustrated in Figure 3.

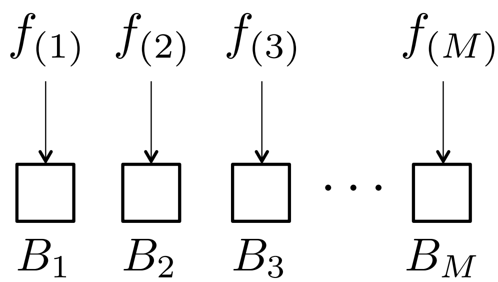


Figure 3: Balls throwing to bins

(a.) Alice and Bob decide to play a betting game. Bob bets that there will be at least one ball in bin B_1 after $\frac{1}{\lambda}$ seconds and Alice bets that there will be at least two balls in bin B_1 after $\frac{2}{\lambda}$ seconds. What is the probability that exactly one of Bob and Alice wins money?

Solution: Let E_1 be the event that Bob wins money and E_2 the event that Alice wins money. We are interested in $P(E_1, E_2^c) + P(E_1^c, E_2)$. Note that $P(E_1, E_2^c)$ is the probability that there is exactly one arrival in the first $\frac{1}{\lambda}$ seconds and no arrivals in the next $\frac{1}{\lambda}$ seconds, so $P(E_1, E_2^{(c)}) = e^{-2}$. Additionally, $P(E_1^c, E_2)$ is the probability that there are no arrivals in the first $\frac{1}{\lambda}$ seconds, and at least 2 arrivals in the next $\frac{1}{\lambda}$ seconds, which is given by $e^{-1}(1 - 2e^{-1})$. Thus, the probability that exactly one of them wins money is $e^{-1} - e^{-2}$.

(b.) Suppose that $\lambda = 1$. You peek into bin B_i at a random time t. Let X be the difference between the time of the most recent arrival to bin B_3 and the time of the next arrival to any of the M bins. Find $E[X^2]$.

Solution: Note that by the same argument as random incidence, the time to the most recent arrival to the bin is exponential with rate 1, and the time to the next arrival to any bin is exponential with rate M. We are interested in X, the sum of these two random variables. We note that since they are independent, we find $M_X(s)$ as the product of the MGFs of these random variables to see that:

$$M_X(s) = \left(\frac{1}{1-s}\right)\left(\frac{M}{M-s}\right) = \frac{M}{M + s^2 - (M+1)s}$$

We note that $E[X^2] = \frac{d^2}{ds^2} M_X(s)|_{s=0}$. Plugging in, one sees:

$$E[X^2] = -\frac{2}{M} + \frac{2(M+1)^2}{M^2} = \frac{2M^2 + 2M + 2}{M^2}$$

Your friend finds a switch which can send balls to two consecutive bins as in Figure 4.

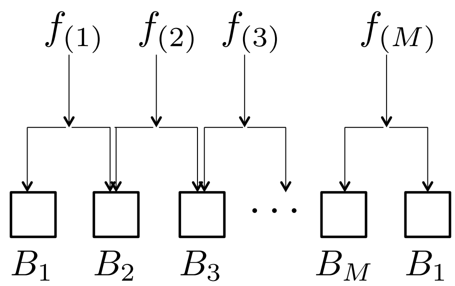


Figure 4: Balls throwing to bins in the new configuration

(c.) Call a bin a singleton if it has exactly one ball contained in it. What is the expected number of singletons at time T?

Solution:

$2MT\lambda e^{-2T\lambda}$, by linearity of expectation.

(d.) Suppose that your friend randomly chooses the setting on the switch so that the situation is equally likely to be that of Figure 3 and Figure 4. What is the variance of the total number of distinct balls in B_1 and B_2 after 1 second?

Solution:

Let N be the number of distinct balls in B_1 and B_2 and let S_1 be the setting of part a., S_2 be the setting of part b. We need $E[N^2]$ and EN^2.

First, we see that $E[N^2] = E[N^2|S_1]P(S_1) + E[N^2|S_2]P(S_2)$. Noting that $E[N^2|S_1] = \mathrm{Var}(N|S_1) + E[N|S_1]^2$, and the analagous setting for $E[N^2|S_2]$, we find that $E[N^2] = \frac{5}{2}\lambda + \frac{13}{2}\lambda^2$. We likewise see that $E[N] = \frac{5}{2}\lambda$, so $\mathrm{Var}(N) = \frac{5}{2}\lambda + \frac{1}{4}\lambda^2$.

(e.) Alice and Bob decide to play a new game. They flip the switch so that the set up is as in Figure 4, but with a twist. Now, the balls in $f_{(1)}$ arrive according to a Poisson process with rate λ, but all other sets of balls ($f_{(2)}, f_{(3)}, \dots$) arrive according to a Poisson process with rate 2λ. The game is as follows: Alice wins if at any point in time, B_3 contains 10 balls more than B_2. What is the probability that Alice wins?

Solution: Note that the balls which go to both B_2 and B_3 do not factor into the difference. Now, consider the merged Poisson process containing balls that go to B_2 but not B_3 and the balls that go to B_3 but not B_2. Now, note that each arrival goes to bin B_2 with probability $\frac{\lambda}{\lambda+2\lambda} = \frac{1}{3}$ and to bin B_3 with probability $\frac{2}{3}$. Now, set a Markov chain that tracks the difference between the number of balls in B_3 and B_2. Let state 10 be the starting state, state 20 the state at which Alice wins and state 0 the state at which Bob wins. The transitions are governed by $P_{i,i+1} = \frac{2}{3}, P_{i,i-1} = \frac{1}{3}$ for $i = \{1, 2, \dots, 19\}$ and $P_{0,0} = 1, P_{20,20} = 1$. Let p_i denote the probability of hitting state 20 when starting at state 10 and note that $p_i = \frac{2}{3}p_{i+1} + \frac{1}{3}p_{i-1}$. Additionally, one has that $p_{20} = 1, p_0 = 0$. Thus, one can see that $p_i = \frac{2^i - 1}{2^{-1}} p_1$. Thus, setting $p_k = 1$, one finds that $p_1 = \frac{2^{19}}{2^{20}-1}$. Plugging in, one sees that $p_{10} = \frac{2^{10}(2^{10}-1)}{2^{20}-1}$.

END OF THE EXAM.

Please check whether you have written your name and SID on every page.