EECS 126 Probability and Random Processes University of California, Berkeley: Spring 2016
Kannan Ramchandran March 31, 2016

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)	12		2		20	
	(b)	8		3		20	
	(c)	9		4		25	
	(d)	8					
	(e)	8					
Total						110	

Problem 1. (a) (12 points) Evaluate the statements with *True* or *False*. Give brief explanations in the provided boxes. Anything written outside the boxes will not be graded.

(1) A Discrete-Time Markov Chain that is not irreducible has no stationary distribution.
True or False: **False**
Explanation:
If the DTMC is not irreducible, it has no unique stationary distribution, but can have many stationary distributions. For example, consider $P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. This chain is not irreducible, but has stationary distribution $\begin{bmatrix} p & 1-p \end{bmatrix}$ for any $p \in [0, 1]$.

(2) Convergence in probability implies convergence almost surely.
True or False: **False**
Explanation:
Consider an arrival process where the set of times is partitioned into intervals of the form $I_k = \{2^k, 2^k + 1, \ldots 2^{k+1} - 1\}$ such that exactly one arrival occurs in each $I_k, k \geq 0$. Now consider the random variable Y_n which is 1 if there is an arrival at time n, and 0 otherwise. Now, note that $P(Y_n = 1) = \frac{1}{2^k}$ if $n \in I_k$. As $k \to \infty$, the size of $I_k \to \infty$, so $\lim_{n\to\infty} P(Y_n = 0) = 1$ and Y_n converges to 0 in probability. However, note that there are infinitely many occurrences of Y_n which are equal to 1, and the event $\{\lim_{n\to\infty} Y_n = 0\}$ has probability 0, so this sequence does not converge almost surely.

(3) If buses have been arriving to Cory Hall according to a Poisson process with rate λ for an infinite amount of time and you arrive at 11:00AM, then the distribution of the interarrival time from the last bus that arrived before 11:00AM to the next bus to come is exponentially distributed with rate λ.
True or False: **False**
Explanation: The distribution is Erlang.

(b) (8 points) Consider a random variable X with moment generating function (MGF) $M_X(s) = a_2 s^2 + a_1 s + a_0$ where a_1, a_2 are such that $a_1 + a_2 = 1$ and $E[X] = \mathrm{Var}(X)$. Determine a_0, a_1, a_2.

Solution: First note that $M_X(0) = 1$, so $a_0 = 1$. Now, note that $E[X] = \frac{d}{ds} M_X(s)\big|_{s=0} = a_1$ and $\mathrm{Var}(X) = E[X^2] - E[X]^2 = \frac{d^2}{ds^2} M_X(s)\big|_{s=0} - a_1^2 = 2a_2 - a_1^2$. Thus, we have: $2a_2 - a_1^2 = a_1$, so $a_1 = 2(1-a_1) - a_1^2$. Solving the quadratic gives $a_1 = \frac{-3 \pm \sqrt{17}}{2}$. Note that since $a_1 = E[X] = \mathrm{Var}(X)$, and the variance of a random variable is nonnegative, we take the positive root. Thus $a_1 = \frac{\sqrt{17}-3}{2}$, and $a_2 = 1 - a_1 = 1 - \frac{\sqrt{17}-3}{2}$.

(c) (9 points) Alice would like to encode a $100\,\text{MB}$ file using a fountain code in order to send the file to Bob. She divides her file into 5 $20\,\text{MB}$ chunks and uses the following degree distribution: at the ith transmission, if $1 \leq i \leq 5$, she uniformly at random selects i of the five chunks and sends the mod 2 sum (or XOR) of these i chunks, while if $i > 5$, she uniformly at random selects 1 of the five chunks and sends that chunk. Assume that Bob uses a peeling decoder, as described in Lab 4. Find the probability that Bob is able to decode 3 packets after the 3rd transmission.

Solution: Note that the maximum number of packets Bob can decode is 3, and we count the three cases separately.

The probability is:

$$\frac{\binom{5}{1} \cdot \binom{4}{1} \cdot \binom{3}{1}}{\binom{5}{1} \cdot \binom{5}{2} \cdot \binom{5}{3}} = \frac{6}{50}$$

(d) (8 points) You have a set of three coins: $A, B,$ and C stacked in your hand. At each time instant, you shuffle the coins by taking the middle coin and putting it on top of the stack with probability $\frac{1}{2}$ and on the bottom of the stack with probability $\frac{1}{2}$.

(i) (4 points) Draw the state transition diagram.

Solution: See the re-labeling below, and note the transition probabilities.

(ii) (4 points) Starting from the order A, B, C find the expected number of shuffles until the coins are in the order C, B, A? (It is not necessary to solve numerically, just set up the equations)

Solution: The possible states are the permutations $\{(A,B,C),(B,A,C),(A,C,B),(B,C,A),(C,A,B),(C,B,A)\}$. For simplicity, consider the following labeling:

1	C, B, A
2	B, C, A
3	B, A, C
4	A, B, C
5	A, C, B
6	C, A, B

Letting $\beta(i)$ be the expected time to reach state 1 from state i, we are interested in finding $\beta(4)$. We thus have the first step equations:

$$\beta(i) = 1 + \frac{1}{2}\beta(i+1) + \frac{1}{2}\beta(i-1), i \neq 1$$
$$\beta(1) = 0$$

Where $6+1$ wraps around to state 1 and $1-1$ wraps around to state 6. Solving the system gives $\beta(4) = 9$.

(e) (8 points) Consider two irreducible, aperiodic Markov Chains with the same state space such that P_1, P_2 give the transition matrices and π_1, π_2 give the stationary distributions. We construct a process $X_n, n \geq 0$ as follows. Let $X_0 = 1$. Now, you flip a coin such that if the coin toss results in a heads, the rest of the transitions are made according to P_1, and if the coin toss results in a tails, the rest of the transitions are made according to P_2. Is $X_n, n \geq 0$ a Markov Chain? If so, determine the transition probabilities. If not, provide a counterexample.

Solution: It is not a Markov Chain. Consider:

$$P_1 = \begin{bmatrix} 1-\delta_1 & \delta_1 \\ 1-\delta_2 & \delta_2 \end{bmatrix}, \begin{bmatrix} \delta_1 & 1-\delta_1 \\ \delta_2 & 1-\delta_2 \end{bmatrix}$$

Where $\delta_1, \delta_2 << 1$. Now, we see that:

$$P(X_3 = 1 | X_2 = 1, X_1 = 1) > P(X_3 = 1 | X_2 = 1, X_1 = 2)$$

Problem 2. (20 points) Empty taxis pass by a street corner at a Poisson rate of two per minute and pick up a passenger if one is waiting there. Passengers arrive at the street corner at a Poisson rate of one per minute and wait for a taxi only if there are less than four persons waiting; otherwise they leave and never return. John arrives at the street corner at a given time. Find his expected waiting time, given that he joins the queue. Assume that the process is in steady state.

Solution:

Consider a continuous time Markov chain with states $X \in \{0, 1, 2, 3, 4\}$ which denotes the number of people waiting. For $n = 0, 1, 2, 3$, the transitions from n to $n+1$ have rate 1, and the transitions from $n+1$ to n have rate 2. The balance equations are then,

$$\pi(n) = \frac{1}{2}\pi(n-1), \; n = 1, 2, 3, 4.$$

Using the above equations and $\sum_{i=0}^{4} \pi(i) = 1$ we find that $\pi(i) = 2^{-i}\pi(0)$ and $\pi(0) = 16/31$. Since the expected waiting time for a new taxi is 0.5, the expected waiting time of John given that he joins the queue can be computed as follows.

$$E[T] = \frac{\pi(0)}{\pi(0)+\pi(1)+\pi(2)+\pi(3)} \times 0.5 + \frac{\pi(1)}{\pi(0)+\pi(1)+\pi(2)+\pi(3)} \times 1 + \frac{\pi(2)}{\pi(0)+\pi(1)+\pi(2)+\pi(3)} \times 1.5 + \frac{\pi(3)}{\pi(0)+\pi(1)+\pi(2)+\pi(3)} \times 2 = 26/30.$$

Problem 3. (20 points) The citizens of the country USD (the United States of Drumpf) vote in the following manner for their presidential election: if the country is liberal, then each citizen votes for a liberal candidate with probability p and a conservative candidate with probability $1-p$, while if the country is conservative, then each citizen votes for a conservative candidate with probability p and a liberal candidate with probability $1-p$. After the election, the country is declared to be of the party with the majority of the votes.

For part (a), assume that $p=\frac{3}{4}$, and use Chebyshev's inequality to obtain your results.

(a) (10 points) Suppose that 100 citizens of USD vote in the election and that USD is known to be Conservative. Bound the probability that it is declared to be a Liberal country.

Solution: Let X_i be the indicator that voter i votes as a Liberal. We are interested in bounding the quantity $P(S_{100} \geq 51)$ where $S_{100} = X_1 + X_2 + \cdots + X_{100}$. We have:

$$\begin{aligned} P(S_n \geq 51) &= P(X - 25 \geq 26) \\ &\leq P(|X-25| \geq 26) \\ &\leq \frac{\text{Var}(X)}{26^2} = \frac{75}{4 \cdot 26^2} \approx 0.03 \end{aligned}$$

(b) (10 points) For this part, we no longer assume that $p=\frac{3}{4}$, and would like to estimate the unknown p. Using the CLT, determine the number of voters necessary to determine p within an error of 0.01, with probability at least 0.95.

Solution: For now, we let consider general error α and want the probability to be at least $1-\beta$. We are thus interested in:

$$P(|\frac{S_n}{n} - p| \geq \alpha)$$

Note that by the CLT, $\frac{S_n}{n} - p \approx \sqrt{\frac{p(1-p)}{n}} Z$ where $Z \sim N(0,1)$. Thus, we have:

$$\begin{aligned} P(|\frac{S_n}{n} - p| \geq \alpha) &\approx P(|Z| \geq \sqrt{\frac{n}{p(1-p)}} \cdot \alpha) \\ &\leq P(|Z| \geq 2\alpha\sqrt{n}) \end{aligned}$$

Now, we have:

$$\begin{aligned} P(|Z| \geq 2\alpha\sqrt{n}) &= 2P(Z \geq 2\alpha\sqrt{n}) \\ &= 2(1 - P(Z \leq 2\alpha\sqrt{n})) = \beta \end{aligned}$$

Now, we substitute in $\alpha = 0.01$, $\beta = 0.05$, and see that:

$$n = (\frac{1.96}{2 \cdot 0.01})^2 = 98^2 = 9604$$

Problem 4. (25 points) In this problem, we consider a scenario where we compute a sequence of functions, denoted by $\{f_1, f_2, \ldots\}$, using two machines, denoted by machine 1 and 2. For every i and j, computing f_j at machine i takes a random amount of time, denoted by $T_{i,j}$. We assume that the $T_{i,j}$'s are i.i.d. exponential random variables of rate 1 (per second).

We now assume that a machine is assigned an infinitely long list of functions, and that the machine computes the functions in the list one by one.

Alice wants to compute as many *distinct* functions as possible in t seconds. She assigns the odd-indexed functions $(f_1, f_3, f_5, \ldots)$ to machine 1 and the even-indexed functions $(f_2, f_4, f_6, \ldots)$ to machine 2, so that the computations performed by the two machines do not overlap. Each machines computes the functions on its own list one by one for t seconds. We denote the number of functions computed by machine 1 by $N_1(t)$, and we denote the number of functions computed by machine 2 by $N_2(t)$.

(a) (6 points) What is the distribution of the number of *distinct* functions computed for **t=200** seconds by machine 1 and machine 2?

Solution: Since the two lists do not overlap each other, it's simply $N_1(200) + N_2(200)$. Since Poisson(200) + Poisson(200) = Poisson(400), it's Poisson(400).

(b) (6 points) Conditioned on $N_1(200)+N_2(200) = 500$, what are the distributions of $N_1(200)$ and $N_2(200)$? Are they (conditionally) independent?

Solution: Both are Binomial($500, \frac{1}{2}$). They are conditionally dependent since $N_1(200)+N_2(200) = 500$.

Bob proposes a new idea, as described below. Both machines are assigned the same list of functions, say $(f_1, f_2, \ldots)$, and they concurrently compute the functions in the list one by one. As soon as one of the two machines completes a function computation, the other machine immediately cancels its ongoing task, and both machines start working on the next function on the list. This process is repeated for t seconds. Denote the number of computed functions for t seconds under this strategy by $B(t)$.

(c) (6 points) Assume $t = 200$. What is the distribution of $B(t)$?

Solution: Note that inter-arrival time is now exponentially distributed with rate 2. Thus, it's Poisson(400).

Bob starts implementing his strategy but, unfortunately, he realizes that his system does not support task cancellation, which is a crucial component of his strategy.

After struggling for a while, he comes up with a modified version of his strategy, which does not require task cancellation. The new strategy is the following. Both machines are assigned the same list of functions, say $(f_1, f_2, \ldots)$. In the beginning, both machine start concurrently computing f_1. A machine is called 'head' if it is computing f_i and the other one is computing f_j, and $i \geq j$. When a 'head' machine finishes a function computation, it proceeds to the next function on the list. When a non-'head' machine finishes a function computation, it skips down on the list and proceeds to the function being computed by the head machine.

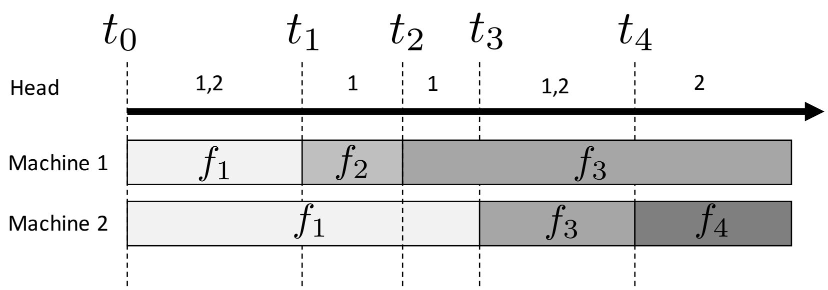

Figure 1: Illustration of the new strategy

See Fig. 1 for illustration. For $t_0 \leq t \leq t_1$, both machines are head. At $t = t_1$, machine 1 finishes computing f_1, and it starts computing f_2 since it is a head. Similarly, at $t = t_2$, machine 1 finishes computing f_2 and proceeds to f_3. At $t = t_3$, machine 2 finishes computing f_1, and it proceeds to f_3, the function being computed by the head. At $t = t_4$, machine 2 finishes computing f_3, and it proceeds to f_4, becoming a new head.

This process is repeated for t seconds.

(d) (7 points) Denote the number of computed functions for t seconds under the modified strategy by $B(t)$. Find $\lim_{t\to\infty} \frac{B(t)}{t}$.

Solution: $\frac{4}{3}$.

END OF THE EXAM.

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