Final Exam Announcements

Exam: (closed-book; three <u>*handwritten*</u> 2-sided 8.5x11 formula sheets and calculator permitted.

Time: 3 Hours

Content: All topics discussed in
Lectures up to and including Lecture 25
Textbook chapters 1-7
Recitations 1-14, Tutorials 1-14 and Problem Sets 1-10

Marathon Office Hours: The TAs will jointly hold office hours at selected times twice a day before the exam. Check the course webpage on the last day of Lectures.

Optional Review Session: The TAs will provide a two-hour review session during which the review problems below will be solved. We suggest working through these problems before attending the review session. For students unable to attend, solutions will be posted on the course web page immediately after the review session. Details for the review session are:

Time: 2 hour session

Also attached is a practice final exam from the previous semester and its solution. Working through the exam before looking at the solutions is a good way to identify areas meriting additional study.

Review Problems

1. Joe Lucky remembers to play the lottery on any given week with probability p, independently of whether he played on any other week. Each time he plays, he has a probability q of winning, again independently of everything else. During a fixed time period of n weeks, let X be the number of weeks that he played the lottery and Y the number of weeks that he won.

   (a) What is the probability that he played the lottery on any particular week, given that he did not win anything on that week?

   (b) Find the conditional PMF $p_{Y|X}(y \mid x)$.

   (c) Find the joint PMF $p_{X,Y}(x, y)$.

   (d) Find the marginal PMF $p_Y(y)$. *Hint:* One possibility is to start with the answer to part (c), but the algebra can be messy. But if you think intuitively about the procedure that generates Y, you may be able to guess the answer.

   (e) Find the conditional PMF $p_{X|Y}(x \mid y)$. Do this algebraically using previous answers.

   (f) Rederive the answer to part (e) by thinking as follows. For each one of the $n - Y$ weeks that he did not win, the answer to part (a) should tell you something.

In all parts of this problem, make sure to indicate the range of values to which your PMF formula applies.

2. A 6.041 graduate opens a new casino in Las Vegas and decides to make the games more challenging from a probabilistic point of view. In a new version of roulette, each contestant spins the following kind of roulette wheel. The wheel has radius r and its perimeter is divided into 20 intervals, alternating red and black. The red intervals (along the perimeter) are twice the width of the black intervals (also along the perimeter). The red intervals all have the same length and the black intervals all have the same length. After the wheel is spun, the center of the ball is equally likely to settle in any position on the edge of the wheel; in other words, the angle of the final ball position (marked at the ball's center) along the wheel's perimeter is distributed uniformly between 0 and 2π radians.

   (a) What is the probability that the center of the ball settles in a red interval?

   (b) Let B denote the event that the center of the ball settles in a black interval. Find the conditional PDF $f_{Z|B}(z)$, where Z is the distance, along the perimeter of the roulette wheel, between the center of the ball and the edge of the interval immediately clockwise from the center of the ball?

   (c) What is the unconditional PDF $f_Z(z)$?

   Another attraction of the casino is the Gaussian slot machine. In this game, the machine produces independent identically distributed (IID) numbers $X_1, X_2, \ldots$ that have normal distribution $\mathcal{N}(0, \sigma^2)$. For every i, when the number X_i is positive, the player receives from the casino a sum of money equal to X_i. When X_i is negative, the player pays the casino a sum of money equal to $|X_i|$.

   (d) What is the standard deviation of the net total gain of a player after n plays of the Gaussian slot machine?

   (e) What is the probability that the absolute value of the net total gain after n plays is greater than $2\sqrt{n}\sigma$?

3. Let a continuous random variable X be uniformly distributed over the interval $[-1, 1]$. Derive the PDF $f_Y(y)$ where:

   (a) $Y = \sin(\frac{\pi}{2}X)$

   (b) $Y = \sin(2\pi X)$

4. People who wish to use a particular mailbox arrive at the box in a Poisson manner with average arrival rate of λ customers per hour. Independently, each user of the mailbox wishes to mail <u>either</u> one letter (with probability $\frac{2}{3}$) <u>or</u> one parcel (with probability $\frac{1}{3}$). Although the users arrive one at a time in a Poisson manner, the ith user is accompanied by N_i <u>non-user</u> friends, where the N_i's are independent and identically-distributed random variables with an associated transform

$$M_N(s) = \frac{1}{2} + \frac{1}{3}e^s + \frac{1}{6}e^{2s} \qquad .$$

(a) Determine the expected value of Y, the total number of people arriving at the mailbox during a three hour interval.

(b) At 3 P.M. today, we shall begin counting letters and parcels as they are brought to the mailbox. What is the probability we shall see a total of exactly three parcels by the time the fifth letter arrives?

(c) Determine the probability that exactly three out of the next eight users will mail parcels.

(d) If we make an equally likely selection from all people (users and non-users) who arrive at the mailbox in the last week, what is the probability we select a person who was accompanied when he or she arrived at the mailbox

(e) Determine either the PMF or the transform for K, the total number of people arriving at the mailbox during a particular hour.

(f) For a time selected by random incidence, determine the transform for T, the total time interval from the arrival of the third previous user until the arrival of the fifth future user of the mailbox.

5. Mary loves gambling. She starts out with \$200. She can bet either \$100 or \$200 (assuming she has sufficient funds), and wins with probability p.

(a) Assuming that she stops when she runs out of money or when she has reached \$400, what is the optimal betting strategy? (i.e., how much should she bet when she has \$100, \$200, \$300? The amount she bets does not have to be the same amount at each time.)

(b) What is the expected number of transitions until she either runs out of money or reaches \$400 for $p = 0.75$ under the optimal strategy?

6. Wombats and dingos arrive in a Poisson manner to a particular water hole in the Australian Outback. The arrival rates of wombats and dingos are 2 and 4 per hour, respectively. Each animal will stay and drink until the next animal arrives to take over. No other animals visit the water hole.

(a) What is the expected number of animals (wombats or dingos) that visit the water hole in a 24-hour period?

(b) Given that a wombat is currently drinking, what is the probability that the next animal to visit is a dingo?

Crocodile Dundee arrives at the water hole at a random time and leaves the water hole immediately after the $900th$ animal he sees departing the water hole.

(c) How long, on average, will Crocodile Dundee have to wait to see a dingo?

(d) Consider the first dingo that Crocodile Dundee sees. How long, on average, does this dingo spend at the water hole?

(e) What does Chebyshev's inequality tell you about the probability that Crocodile Dundee stays at the water hole for between 140 and 160 hours?

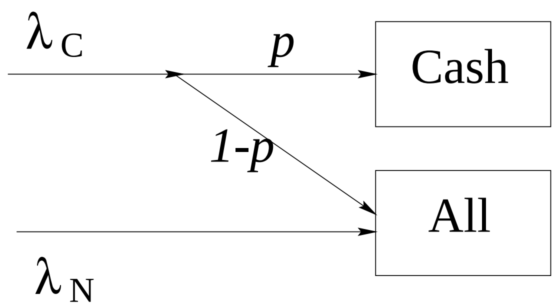
(f) Using the Central Limit Theorem, what is the probability that Crocodile Dundee stays at the water hole for between 140 and 160 hours?

SPRING 2002 FINAL EXAM

Problem 0. (2 points) Write your TA's name on the front of the booklet.

Problem 1. (46 points)

The student center lobby has two ATMs, one for cash-only transactions, and one for all transactions.



Customers (type-C) with cash-only transactions arrive according to a Poisson process with rate λ_C per hour. Each type-C customer chooses the cash-only ATM with probability p, and the all-transaction ATM with probability $(1-p)$, independently from other customers.

Customers (type-N) with non-cash transactions arrive according to an independent Poisson process with rate λ_N per hour, and choose the all-transaction ATM.

(a) (5 points) Give a formula for the mean and variance of the number of customers with cash-only transactions that arrive between 9 a.m. and 11 a.m.

(b) (5 points)A customer just arrived at the all-transaction ATM. Whatt is the probability that he is a type-C customer?

(c) (5 points) Given that there was **exactly one** type-C customer between 9 a.m. and 10 a.m., find the PDF of the arrival time of that customer.

(d) (5 points) Given that there was **at least one** type-C customer between 9 a.m. and 10 a.m., find the PDF of the arrival time of the first type-C customer.

(e) (6 points) Find the PDF of the square of the time elapsed between the arrival of the first and the second type-C customer.

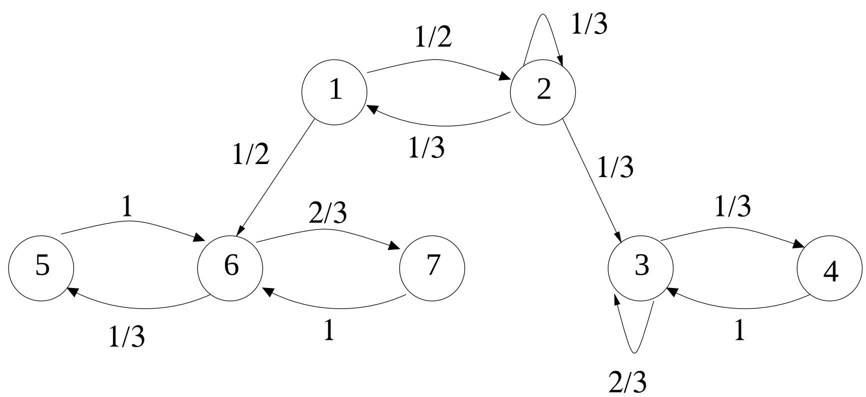
(f) (7 points) Give a formula for the probability that the number of type-C customers that arrive between 9 a.m. and 12 p.m. is at least twice the number of type-C customers that arrive between 9 a.m. and 10 a.m. (If your formula involves a complicated sum or integral, you do not need to evaluate it.)

(g) (6 points) Suppose that $\lambda_C = 2$ and that $p = 0.98$. Find a good numerical approximation for the probability that out of the first 100 type-C customers, exactly 95 choose the cash-only ATM.

(h) (7 points) Give a formula for the expected time until at least one customer of each type (that is, type-C and type-N) has arrived.

Problem 2. (52 points)

Consider the discrete-time Markov chain described by the following transition diagram:



(a) (5 points) Assume that the process starts at state 1, i.e. $X_0 = 1$. Find the probability that $X_2 = 5$.

(b) (5 points) Assume that the process starts at state 1, i.e. $X_0 = 1$. Find the probability that $X_2 = 1$ given that $X_3 = 6$.

(c) (5 points) Find approximately the probability $\mathbf{P}(X_n = 3 \mid X_0 = 3)$, when n is large.

(d) (5 points) Find approximately the probability $\mathbf{P}(X_n = 3 \mid X_0 = 1)$, when n is large.

(e) (5 points) Find the expected time until the state leaves the set $\{1, 2\}$, given that the process starts at state 1.

(f) (7 points) Given that the process starts at state 3, find the expected value and the variance of the time elapsed until state 4 is entered for the third time.

(g) (6 points) Suppose that the process starts at state 6. Find a good approximation for the probability that state 5 is visited at least 30 times in the first 144 transitions.

(h) (6 points) Suppose that the process starts at state 6. Let V_n be the number of visits to state 7 in the first n transitions. For each one of the expressions below, state whether it converges in probability, and also whether it converges with probability 1. Whenever the answer is yes, also state what the limit is.

   i. $\frac{V_n - E[V_n]}{\sqrt{n}}$

   ii. $\frac{V_n - E[V_n]}{n}$

(i) (8 points) Suppose that the process starts at time 0 at state 3. Each time (including time 0) that state 3 is visited, you win a prize, with probability 1/2. Each time that state 4 is visited, you win a prize, with probability 1/4. Given the current state, the event of winning a prize is assumed independent of everything else. Find the expected value of the time at which you win a prize for the first time.

SPRING 2002 FINAL EXAM SOLUTIONS

Problem 0: (2 pts) Answers may vary.

Problem 1:

(a) (5 pts) The number of type-C customers that arrive between 9am and 11am is Poisson with parameter $2\lambda_C$, which has mean $2\lambda_C$ and variance $2\lambda_C$.

(b) (5 pts) At the all-transaction ATM, type-C customers arrive with rate $(1-p)\lambda_C$ and type-N customers arrive with rate λ_N. So a customer at the all-transaction ATM is type-C with probability $\frac{(1-p)\lambda_C}{(1-p)\lambda_C+\lambda_N}$.

(c) (5 pts) Let X be the time elapsed between 9am and the arrival time of the first type-C customer. Given that there was exactly one type-C customer between 9am and 10am, that customer's arrival time is uniformly distributed between 9am and 10am. Therefore, X is uniformly distributed on the interval [0,1].

(d) (5 pts) Given that there was at least one type-C customer between 9am and 10am, we know that the first type-C customer arrived before 10am. So we want to find the PDF of X given that $X \leq 1$. For $0 \leq x \leq 1$, we have

$$f_{X|X\leq 1}(x) = \frac{f_X(x)}{\mathbf{P}(X \leq 1)} = \frac{\lambda_C e^{-\lambda_C x}}{1-e^{-\lambda_C}}.$$

(e) (6 pts) Let Y be the time elapsed between the arrival of the first and the second type-C customer, so Y is exponentially distributed with rate λ_C. We may find the PDF of $Z = Y^2$ by finding the CDF of Z first (notice that $Z = Y^2$ is monotonic because Y is nonnegative). For $z \geq 0$, we have

$$F_Z(z) = \mathbf{P}(Z \leq z) = \mathbf{P}(Y^2 \leq z) = \mathbf{P}(Y \leq \sqrt{z}) = 1 - e^{-\lambda_C\sqrt{z}}$$

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \frac{d}{dz}(1-e^{-\lambda_C\sqrt{z}}) = \frac{\lambda_C}{2\sqrt{z}}e^{-\lambda_C\sqrt{z}}$$

(f) (7 pts) Let A be the event that the number of type-C customers that arrive between 10am and 12pm is greater than or equal to the number of type-C customers that arrive between 9am and 10am.

$$\begin{aligned}
\mathbf{P}(A) &= \sum_{k=0}^{\infty} \mathbf{P}(A|k \text{ arrivals between 9am and 10am})\mathbf{P}(k \text{ arrivals between 9am and 10am}) \\
&= \sum_{k=0}^{\infty} \mathbf{P}(\text{at least } k \text{ arrivals between 10am and 12pm})\mathbf{P}(k \text{ arrivals between 9am and 10am}) \\
&= \sum_{k=0}^{\infty}\sum_{j=k}^{\infty} \frac{(2\lambda_C)^j e^{-2\lambda_C}}{j!}\frac{\lambda_C^k e^{-\lambda_C}}{k!}
\end{aligned}$$

(g) (6 pts) Let S_{100} be binomial with 100 trials and probability of success .02 (we need the probability of success to be small in order to use the Poisson approximation of the binomial, so in this case a success represents choosing the all-transaction ATM). We may approximate S_{100} as a Poisson random variable with parameter $np = 100(.02) = 2$.

$$\begin{aligned}\mathbf{P}(\text{out of 100 type-C customers, exactly 95 choose the cash-only ATM}) &= \mathbf{P}(S_{100} = 5) \\ &\approx \frac{2^5 e^{-2}}{5!}\end{aligned}$$

(h) (7 pts) Let T_1 be the time until the first customer arrives, so T_1 is exponentially distributed with rate $\lambda_C + \lambda_N$. Let T_2 be the time elapsed between the first customer arrival and the first arrival of a customer of opposite type, so the time until at least one customer of each type has arrived is $T_1 + T_2$.

$$\begin{aligned}\mathbf{E}[T_1 + T_2] &= \mathbf{E}[T_1] + \mathbf{E}[T_2] \\ &= \mathbf{E}[T_1] + \mathbf{E}[T_2|\text{first arrival is type-C}]\mathbf{P}(\text{first arrival is type-C}) \\ &\quad +\mathbf{E}[T_2|\text{first arrival is type-N}]\mathbf{P}(\text{first arrival is type-N}) \\ &= \frac{1}{\lambda_C + \lambda_N} + \frac{1}{\lambda_N}\frac{\lambda_C}{\lambda_C + \lambda_N} + \frac{1}{\lambda_C}\frac{\lambda_N}{\lambda_C + \lambda_N}\end{aligned}$$

Problem 2:

(a) (5 pts) The only way to be at state 5 after two transitions, if we are currently at state 1, is $1 \to 6 \to 5$. So $\mathbf{P}(X_2 = 5|X_0 = 1) = \frac{1}{2}\frac{1}{3}$.

(b) (5 pts) There are three ways to be at state 6 after 3 transitions, if we are currently at state 1: $1 \to 2 \to 1 \to 6$, $1 \to 6 \to 5 \to 6$, and $1 \to 6 \to 7 \to 6$.

$$\mathbf{P}(X_2 = 1|X_3 = 6, X_0 = 1) = \frac{\mathbf{P}(X_2 = 1, X_3 = 6|X_0 = 1)}{\mathbf{P}(X_3 = 6|X_0 = 1)} = \frac{\frac{1}{2}\frac{1}{3}\frac{1}{2}}{\frac{1}{2}\frac{1}{3}\frac{1}{2} + \frac{1}{2}\frac{1}{3}1 + \frac{1}{2}\frac{2}{3}1}$$

(c) (5 pts) For n is large, $\mathbf{P}(X_n = 3|X_0 = 3) = \pi_3$. Noticing states 3 and 4 form a birth-death chain, we solve the following set of equations to find the steady-state probability of being at state 3:

$$\begin{aligned}\frac{1}{3}\pi_3 &= \pi_4 \\ \pi_3 + \pi_4 &= 1\end{aligned}$$

which gives us $\pi_3 = \frac{3}{4}$.

(d) (5 pts) For n is large, $\mathbf{P}(X_n = 3|X_0 = 1) = a_{13}\pi_3$, where a_{13} is the probability of being absorbed by state 3's recurrent class if we are currently at state 1. We solve the following set of equations to find a_{13}:

$$\begin{aligned}a_{13} &= \frac{1}{2}a_{23} \\ a_{23} &= \frac{1}{3}a_{23} + \frac{1}{3}a_{13} + \frac{1}{3}\end{aligned}$$

which gives us $a_{13}\pi_3 = \frac{1}{3}\frac{3}{4} = \frac{1}{4}$.

(e) (5 pts) We solve the following set of equations to find the expected time until absorption from state 1.

$$\begin{aligned}\mu_1 &= 1 + \frac{1}{2}\mu_2 \\ \mu_2 &= 1 + \frac{1}{3}\mu_1 + \frac{1}{3}\mu_2\end{aligned}$$

which gives us $\mu_1 = \frac{7}{3}$.

(f) (7 pts) Given we are currently at state 3, the time until we enter state 4 is geometric with parameter $p = \frac{1}{3}$. Given we are currently at state 4, the time until we enter state 3 is 1. So the time until we enter state 4 for the third time is the sum of three independent geometric random variables plus 2, which has mean $3\frac{1}{p} + 2 = 11$ and variance $3\frac{1-p}{p^2} = 18$.

(g) (6 pts) Suppose the process starts at state 6. Then at all even times we are at state 6. At all odds times we are at state 5 with probability $\frac{1}{3}$ or at state 7 with probability $\frac{2}{3}$, independently of where we were two transitions earlier. So the number of visits to state 5 in the first 144 transitions is a binomial random variable with $\frac{144}{2} = 72$ trials and probability of success $\frac{1}{3}$. Using the Central Limit Theorem with a half-correction, we have:

$$\mathbf{P}(S_{72} \geq 30) = \mathbf{P}\left(\frac{S_{72} - E[S_{72}]}{\sigma_{S_{72}}} \geq \frac{29.5 - 72\frac{1}{3}}{\sqrt{72\frac{1}{2}\frac{2}{3}}}\right) \approx \mathbf{P}(Z \geq 1.375) = 1 - \Phi(1.375) = .0838.$$

(h) (6 pts) V_n is a binomial random variable with $\lceil \frac{n}{2} \rceil$ trials and probability of success $\frac{2}{3}$. Because a binomial random variable is the sum of independent, identically distributed random variables, we know that $\frac{V_n - E[V_n]}{\sqrt{n}}$ approaches a normal distribution as $n \to \infty$ (by the Central Limit Theorem), and therefore does not converge, and that $\frac{V_n - E[V_n]}{n}$ converges to 0 in probability and with probability 1 (by the Weak and Strong Laws of Large Numbers respectively).

(i) (8 pts) Let h_i be the expected time until we win a prize given we are currently at state i. Using the total expectation theorem conditioned on whether we win or lose and on where the chain transitions to, we set up the following set of equations:

$$\begin{aligned}h_3 &= \frac{1}{2}(0) + \frac{1}{2}\left(\frac{2}{3}(1 + h_3) + \frac{1}{3}(1 + h_4)\right) \\ h_4 &= \frac{1}{4}(0) + \frac{3}{4}(1 + h_3)\end{aligned}$$

which gives us $h_3 = \frac{15}{13}$.