Spring 2018 18.600 Final Exam Solutions

1. (10 points) Suppose that the pair (X, Y) is uniformly distributed on the unit circle $\{(x, y) : x^2 + y^2 \le 1\}$.

(a) Give the joint probability density function $f(x, y)$ for the pair (X, Y). **ANSWER:**

$$f(x,y) = \begin{cases} \frac{1}{\pi} & x^2+y^2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) Compute the marginal law $f_X(x)$. **ANSWER:**

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \begin{cases} \frac{2\sqrt{1-x^2}}{\pi} & |x| \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

(c) Compute the conditional expectation $E[X^2|Y]$ as a function of Y. **ANSWER:** If $y \in [-1, 1]$, then $E[X^2|Y = y]$ equals $E[Z^2]$ for Z uniform on $[-a, a]$ with $a = \sqrt{1-y^2}$. Here $E[Z^2] = \int_{-a}^{a} \frac{1}{2a} x^2 dx = x^3/3 \,\big|_{-a}^{a} = a^2/3$. So $E[X^2|Y = y] = (1 - y^2)/3$ and $E[X^2|Y] = (1 - Y^2)/3$. (We do not define $E[X^2|Y]$ for $Y \notin [-1, 1]$.)

2. (10 points) Ivan is waiting for the bus. There are three kinds of buses, whose arrival times are three independent Poisson point processes. During each hour, on average one expects to see 3 yellow buses, 2 blue buses, and 1 red bus at the bus stop where Ivan is waiting.

(a) Give the probability density function for T where T is the length of time until the first bus of *any* kind shows up. **ANSWER:** This is exponential with parameter $\lambda = 6$, so density is $f_T(x) = 6e^{-6x}$ for $x \ge 0$.

(b) Compute the probability that there are exactly two yellow buses during the first half hour. **ANSWER:** Number of yellow buses in $1/2$ hour is Poisson with parameter $\lambda = 3/2$. So probability is $e^{-\lambda}\lambda^k/k!$ with $k = 2$, which is $e^{-3/2}(3/2)^2/2$.

(c) Suppose that a yellow bus would get Ivan home in 30 minutes (from the time it arrives at the bus stop until the time it gets to his house) while either a red or a blue bus would get Ivan home in 15 minutes (from the time it picks Ivan up at the bus stop until the time it gets to his house). If Ivan wants to minimize the expected amount of time until he gets home, and a yellow bus arrives first, should he get on the yellow bus or should he wait for a red/blue bus to come? Give a sentence or two of explanation. **ANSWER:** If he holds out for a red/blue bus, it will take 20 minutes to arrive (in expectation) but only shorten transit duration for 15 minutes. So he is better off (in expectation) taking the bird in hand (i.e., getting on the yellow bus).

3. (10 points) Harriet and Helen are real estate agents. Each week, each agent is sent to close a deal, and one of the two agents closes a deal 10 percent of the time, while the other closes it 20 percent of the time. To set notation, let X_j be 1 if the more capable agent closes the deal the jth week and 0 otherwise. Let Y_j be 1 if the less capable agent closes the deal on the jth week, and 0 otherwise. So each X_j is equal to 1 with probability $2/10$ and 0 with probability $8/10$. Each Y_j is equal to 1 with probability $1/10$ and 0 with probability $9/10$. Assume that the random variables $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are all independent of each other. For each $j \geq 1$ write $Z_j = X_j - Y_j$. Write $S_n = \sum_{j=1}^n Z_j$.

(a) Compute the mean, variance and standard deviation of Z_1. **ANSWER:** $E[X_1] = .2$ and $E[Y_1] = .1$ so $E[Z_1] = E[X_1 - Y_1] = .1$. Using independence, $\mathrm{Var}(Z_1) = \mathrm{Var}(X_1 - Y_1) = \mathrm{Var}(X_1) + \mathrm{Var}(-Y_1) = \mathrm{Var}(X_1) + \mathrm{Var}(Y_1) = .09 + .16 = .25$. So $\mathrm{SD}(Z_1) = .5$.

(b) Compute the mean, variance and standard deviation of S_{100}. **ANSWER:** By additivity of mean and variance (for independent sums) we get $E[S_{100}] = 10$ and $\mathrm{Var}(S_{100}) = 25$ so $\mathrm{SD}(S_{100}) = 5$.

(c) Use the central limit theorem to approximate the probability that $S_{100} > 0$. (This is the probability that the more capable closer has closed more deals after 100 weeks.) You may use the function $\Phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ in your answer. **ANSWER:** Note that 0 is two standard deviations below the expectation of S_{100}. So using CLT approximation, $P(S_{100} > 0) \approx P(N > -2)$ where N is a standard normal random variable. This is $1 - \Phi(-2)$ or equivalently $\Phi(2) \approx .977$. **Remark:** This is similar to a problem set problem in which the probabilities were .5 and .6 instead of .1 and .2. In the problem set problem, the analog of $\mathrm{Var}(Z_1)$ was $.25 + .24 = .49$ and the analog of $\mathrm{SD}(Z_1)$ was .7. It would have taken 196 weeks in the problem set scenario to be able to distinguish between the candidates with as much confidence as we achieved with 100 weeks in this problem. But the take home message is that in both situations it takes a long time to be able to determine (with .977 confidence) which candidate is stronger based on empirical results.

4. (10 points) A future society has exactly seven software companies. (You can guess what they are.) I will call them Company 1, Company 2, Company 3, Company 4, Company 5, Company 6, and Company 7. Every software engineer either works for one of these companies or is unemployed. At the end of each month, a survey asks what each software engineer is doing. It happens that...

A software engineer who is unemployed one month has a $3/10$ chance to be unemployed the next month, and (for each $j \in \{1, 2, \ldots, 7\}$) a $1/10$ chance to be at Company j.

A software engineer at Company j one month (for any $j \in \{1, 2, \ldots, 7\}$) has a $69/70$ chance to remain at Company j the next month, and a $1/70$ chance to be unemployed.

Complete the following:

(a) Represent the employment activity as a Markov chain by writing down an 8 by 8 transition matrix with states numbered 0 (unemployment) or 1 to 7 (the companies). You don't have to write out all 64 terms. Just write the non-zero terms. **ANSWER:**

$$\begin{pmatrix} 3/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/70 & 69/70 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/70 & 0 & 69/70 & 0 & 0 & 0 & 0 & 0 \\ 1/70 & 0 & 0 & 69/70 & 0 & 0 & 0 & 0 \\ 1/70 & 0 & 0 & 0 & 69/70 & 0 & 0 & 0 \\ 1/70 & 0 & 0 & 0 & 0 & 69/70 & 0 & 0 \\ 1/70 & 0 & 0 & 0 & 0 & 0 & 69/70 & 0 \\ 1/70 & 0 & 0 & 0 & 0 & 0 & 0 & 69/70 \end{pmatrix}$$

(b) If a software engineer is currently unemployed, what is the probability that he or she will be unemployed in two months? **ANSWER:** For this to happen, either stay in state 0 both months or transition out to company state and then back to state zero. Adding the two probabilities gives $(3/10)^2 + 7(1/10) \cdot (1/70) = .09 + .01 = 1/10$.

(c) Compute $(\pi_0, \pi_1, \pi_2, \ldots, \pi_7)$ where π_i is the fraction of months spent in state i over the long term. [Hint: start by working out π_0.] **ANSWER:** Imagine we just focus on two states (U for unemployed or E for employed), and note that this is still a Markov chain. So we have

$$\begin{pmatrix} \pi_U & \pi_E \end{pmatrix} \begin{pmatrix} 3/10 & 7/10 \\ 1/70 & 69/70 \end{pmatrix} = \begin{pmatrix} \pi_U & \pi_E \end{pmatrix},$$

and first equation we get from matrix is $(3/10)\pi_U + (1/70)\pi_E = \pi_U$, which implies $(1/70)\pi_E = (7/10)\pi_U$, so $\pi_E = 49\pi_U$. Combining with $\pi_E + \pi_U = 1$ we find $\pi_E = .98$ and $\pi_U = .02$. Now using symmetry (all seven companies are equally likely) we find $\pi_0 = .02$ and $\pi_j = .98/7 = .14$ for $j \in \{1, 2, \ldots, 7\}$.

5. (10 points) Each of the customers calling a certain toaster customer service number fits into one of seven predictable categories:

1/4 want to purchase a new toaster.

1/4 want to complain about toast being too dark and crispy even on light settings.

1/8 want to complain about a toaster that is not connecting to the internet properly.

1/8 want to complain about a hand being stuck in a toaster.

1/8 want to complain about a toaster having started a fire.

1/16 want to know how to make cinnamon toast.

1/16 want to know how to adjust the springs so that the toast really flies into the air.

Let X be the category of the next customer who will call. We do not yet know X so we model it as a "category-valued" random variable, which takes one of the above seven values with the indicated probabilities.

(a) Compute the entropy $H(X)$. **ANSWER:** $\sum p_i(-\log p_i) = (1/4)2 + (1/4)2 + (1/8)3 + (1/8)3 + (1/8)3 + (1/16)4 + (1/16)4 = 21/8$.

(b) Imagine that the call center determines the category by asking the customer a sequence of yes or no questions, using a strategy that minimizes the expected number of questions asked. What is the expected number of questions asked? **ANSWER:** Since all probabilities are powers of two, answer is $H(X) = 21/8$.

(c) Suppose $X_1, X_2, X_3, \ldots, X_{100}$ are i.i.d. random variables, each with the same law as X, representing the calls received over an entire day. Compute $H(X_1, X_2, X_3, \ldots, X_{100})$. **ANSWER:** By additivity of entropy (for independent random variables) this is $100H(X_1) = 2100/8$.

(d) Let Y be the total number of callers during the day who want to purchase a new toaster. So Y is a random integer between 0 and 100. Write $Z = Y^2$. Compute the entropy $H(Z)$. You can write this as a sum; no need to compute an explicit number. **ANSWER:** Observe that $H(Z) = H(Y)$ since the map $x \to x^2$ is one to one when we restrict to the set $\{0, 1, 2, \ldots, 100\}$. Next note that Y is a binomial random variable with $P(Y = k) = (1/4)^k(3/4)^{100-k}\binom{100}{k}$ for $k \in \{0, 1, \ldots, 100\}$ so

$$H(Z) = H(Y) = -\sum_{k=0}^{100}(1/4)^k(3/4)^{100-k}\binom{100}{k}\log\Big((1/4)^k(3/4)^{100-k}\binom{100}{k}\Big).$$

6. (10 points) Let $X_1, X_2, \ldots$ be i.i.d. random variables, each with density function $\frac{1}{\pi(1+x^2)}$.

(a) Write $S_0 = 0$ and (for $n \geq 1$) write $S_n = \sum_{j=1}^n X_j$. Is the sequence S_n a martingale? Explain why or why not. (Recall that the definition of martingale requires that $E[\,S_n|] < \infty$ for all n, and explicitly note whether this is true.) **ANSWER:** This is not technically a martingale since $E[\,S_1|] = \int_{-\infty}^{\infty} \frac{1}{\pi(1+x^2)}|x|dx = \infty$.

(b) Compute the probability density function for S_{10}. **ANSWER:** We recall that $S_{10}/10$ has the same law as X_1, so S_{10} has the same law as $10X_1$. So density is $\frac{1}{10\pi(1+(x/10)^2)}$.

(c) Compute the probability $P(S_{10} > 10)$. **ANSWER:** This is same as probability that $X_1 > 1$, which is 1/4, as is easily seen from spinning flashlight story.

7. (10 points) Let $X_1, X_2, \ldots$ be an infinite sequence of i.i.d. random variables, each equal to -2 with probability $1/2$ and 2 with probability $1/2$. Write $S_0 = 0$ and (for $n \geq 1$) write $S_n = \sum_{j=1}^n X_j$.

(a) Compute the probability that the sequence $S_0, S_1, S_2, \ldots$ reaches 20 before the first time that it reaches -30. **ANSWER:** Let T be the first time one of these points is hit. Then by optional stopping theorem $S_0 = E[S_T] = 20P(S_T = 20) + (-30)P(S_T = -30)$. We also know $P(S_T = 20) + P(S_T = -30) = 1$. Solving for the two unknowns, we find $P(S_T = 20) = .6$ and $P(S_T = -30) = .4$.

(b) Compute the probability that there exists *some* positive integer n for which $S_n = -20$. **ANSWER:** This is 1. The probability that S_n hits -20 before hitting K (for large positive K) is $K/(K+20)$. This is a lower bound on the probability that -20 is hit *ever*. Since we can make this lower bound as close to 1 as we want (by taking K large enough) the answer is 1.

(c) Compute the moment generating function $M_{X_1}(t)$. **ANSWER:** $E[e^{tX_1}] = \frac{1}{2}e^{2t} + \frac{1}{2}e^{-2t}$.

(d) Compute the moment generating function $M_{S_9}(t)$. **ANSWER:** $E[e^{t(X_1+\ldots+X_9)}] = \prod_{j=1}^0 E[e^{tX_j}] = \left(\frac{1}{2}e^{2t} + \frac{1}{2}e^{-2t}\right)^9$.

8. (10 points) Suppose there are 6 cards with labels $1, 2, 3, 4, 5, 6$. The cards are shuffled to a uniformly random permutation and are then turned over one at a time. Each time one sees a card that is higher in numerical value than all other preceding cards, one excitedly shouts "Record!" (The first card to be turned over is automatically a record, since it has no competition.) For example, if the cards are turned over in the order $1, 2, 3, 4, 5, 6$ then one will shout "Record!" all 6 times, since each card has a larger number than all previous numbers. If the cards are turned over in the order $6, 3, 4, 5, 1, 2$ then one will shout "Record!" only once. Let R be the total number of times "Record!" is shouted. (So R is an integer-valued random variable taking values between 1 and 6.) To further set notation, let A_i be 1 if the ith card turned over is a record and 0 otherwise. Compute the following:

(a) The expectation $E[R]$. **ANSWER:** $E[R] = E[\sum_{i=1}^6 A_i] = \sum_{i=1}^6 P(A_i = 1) = 1 + 1/2 + 1/3 + 1/4 + 1/5 + 1/6 = 49/20$.

(b) The variance $\text{Var}[R]$. **ANSWER:** The easiest way to do this is to observe that the A_j are independent of each other. Hence $\text{Var}(R) = \text{Var}(\sum_{i=1}^6 A_i) = \sum_{i=1}^6 \text{Var}(A_i) = \sum_{i=1}^6 \frac{1}{i} \cdot \frac{i-1}{i} = 0/1 + 1/4 + 2/9 + 3/16 + 4/25 + 5/36$.

(c) Compute the conditional probability that the last three cards turned over are all records *given* that the first three cards are all records. **ANSWER:** Chance first three cards are in increasing order is $1/3!$. Chance all cards are records is $1/6!$. So conditional probability is $(1/6!)/(1/3!) = 3!/6! = 1/120$.

9. (10 points) Let $X_1, X_2, X_3 \dots$ be an infinite sequence of i.i.d. normal random variables, each with mean zero and variance one.

(a) Is the sequence $X_1, X_2, \dots$ a martingale? Explain why or why not in one sentence. **ANSWER:** No, because $E[X_n|\mathcal{F}_{n-1}] = 0$, which is not the same as X_{n-1}. In other words, a person at stage $n-1$ knows the true value of X_{n-1} but considers (given all information currently available) that the expectation of X_n is 0, and this is not generally the same as X_{n-1}.

(b) Compute the correlation coefficient of $A = \sum_{n=1}^{60} X_n$ and $B = \sum_{n=41}^{100} X_n$. **ANSWER:** Since A and B have mean zero, we know that

$$\mathrm{Cov}(A,B) = E[AB] = E[\sum_{j=1}^{60} X_j \sum_{k=41}^{100} X_k] = \sum_{j=1}^{60}\sum_{k=41}^{100} E[X_j X_k].$$

The only non-zero terms are when $j = k$, and there are 20 such terms, so $\mathrm{Cov}(A,B) = 20$. By additivity of variance for independent random variables, we have $\mathrm{Var}(A) = \mathrm{Var}(B) = 60$. So the correlation coefficient is

$$\mathrm{Cov}(A,B)/\sqrt{\mathrm{Var}(A)\mathrm{Var}(B)} = 20/60 = 1/3.$$

(c) Write down the probability density function for A. **ANSWER:** This is normal with mean zero and variance 60, so density is $\frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$ with $\sigma^2 = 60$, i.e., $\frac{1}{\sqrt{120\pi}} e^{-x^2/120}$

(d) Compute the variance of $A+B$. **ANSWER:** $\mathrm{Var}(A+B) = \mathrm{Var}(A) + \mathrm{Var}(B) + 2\mathrm{Cov}(A,B) = 60 + 60 + 40 = 160$.

10. (10 points) Let X, Y, Z be i.i.d. exponential random variables, each with parameter $\lambda = 1$.

(a) Compute the probability density function for $X+Y+Z$. **ANSWER:** This a gamma distribution with $n = 3$ and $\lambda = 1$. So density is $x^2 e^{-x}/2!$ on $[0, \infty)$.

(b) Compute the expectation and variance of $\max\{X, Y, Z\}$. **ANSWER:** This is the radioactive decay problem. Expectation is $1/3 + 1/2 + 1 = 11/6$. Variance is $1/9 + 1/4 + 1 = 49/36$.

(c) Compute the expectation and variance of $\min\{X, Y, Z\}$. **ANSWER:** The minimum is exponential with parameter $\lambda = 3$. So expectation is $1/3$ and variance $1/9$.