Spring 2019 18.600 Final Exam Solutions

1. (10 points) A broccoli vendor is choosing a website photo: Photo 1 (attractive people eating broccoli at the beach) or Photo 2 (close-up of broccoli with salmon and quinoa). Assume that one of the photos is more "effective" and that a site visitor on average spends \$16 if shown the "more effective" photo and \$14 if shown the "less effective" photo (with standard deviation \$10 in each case). To find out which photo is best, the vendor implements an "A/B test" that involves trying each photo on 50 visitors. Denote the dollar amounts spent by those shown the more effective photo (whichever that is) by $X_1, X_2, \ldots, X_{50}$ and the amounts spent by those shown the less effective photo by $Y_1, Y_2, \ldots, Y_{50}$. Formally, for each $i \in \{1, 2, \ldots, 50\}$, we have $\mathrm{Var}[X_i] = \mathrm{Var}[Y_i] = 100$ and $E[X_i] = 16$ and $E[Y_i] = 14$, and the X_i are identically distributed (as are the Y_i) and all 100 random variables are independent of each other. Write $X = \sum_{i=1}^{50} X_i$ and $Y = \sum_{i=1}^{50} Y_i$.

(a) Use the central limit theorem to approximate the probability that $X > Y$ (so that the vendor correctly identifies the more effective photo). You may use the function $\phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ in your answer. **ANSWER:** $(X - Y) = \sum_{i=1}^{50}(X_i - Y_i)$ is approximately normal with expectation $50(16 - 14) = 100$ and variance $\mathrm{Var}(X) + \mathrm{Var}(-Y) = 50 \cdot 100 + 50 \cdot 100 = 100000$ and hence standard deviation 100. Since 0 is one standard deviation below $E[(X - Y)]$ we have

$$P(X - Y) > 0 \approx \int_{-1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \phi(-1) = \phi(1) \approx .84.$$

(b) Compute the conditional expectation $E\big[(X_1 - Y_1)|X\big]$ as a function of the random variable X. **ANSWER:** We know $E[X_1 + \ldots + X_{50}|X] = E[X|X] = X$. And $E[X_j|X]$ should be the same for each j by symmetry. Hence $E[X_1|X] = X/50$. Since Y_1 and X are independent $E[Y_1|X] = E[Y_1] = 14$. By additivity of conditional expectation, $E[(X_1 - Y_1)|X] = X/50 - 14$.

(c) Compute $E[(X - Y)^2]$. **ANSWER:** $\mathrm{Var}(X - Y) = E[(X - Y)^2] - E[(X - Y)]^2$, so $E[(X - Y)^2] = \mathrm{Var}(X - Y) + E[(X - Y)]^2 = 10000 + 10000 = 20000$.

2. (10 points) An explorer discovers an island with 10 islanders, each of whom has 10 apples. The explorer proposes a game. At each step of the game the explorer will do the following:

(i) First choose uniformly at random one of the $\binom{10}{2}$ possible *pairs* of islanders.

(ii) Then, if both islanders in the pair have *at least one* apple, toss a fair coin to declare one of them the "winner" and transfer one apple from the loser to the winner. (If either islander is already out of apples, do nothing.)

The above is repeated until one islander (the "overall winner") has all the apples. (This happens *eventually* with probability 1. You don't have to prove this; take it as given.) Let A_n^j be the number of apples the jth person has after n steps, and let T be the number of steps before the game ends. So $A_0^j = 10$ for $j \in \{1, 2, \ldots, 10\}$ and $A_T^j = \begin{cases} 100 & j \text{ is overall winner} \\ 0 & \text{otherwise} \end{cases}$

(a) When $j \in \{1, 2, \ldots, 10\}$ is fixed, is the sequence $A_0^j, A_1^j, A_2^j, \ldots$ a martingale? Explain why or why not. **ANSWER:** Yes. We put ourselves in the shoes of somebody who has seen the first n steps and knows A_n^j. The only way that we have $A_{n+1}^j \neq A_n^j$ is if a pair (j,k) is chosen and both jth and kth islanders have a positive number of apples. But if this happens, then the jth islander is just as likely to gain as to lose an apple. So $E[A_{n+1}^j|\mathcal{F}_n] = A_n^j$.

(b) Compute the expected number of islanders who *at some point* have exactly 25 apples. **ANSWER:** Write $S = \min\{n : A_n^j \in \{0, 25\}\}$. Write $p = P(A_S^j = 25)$. Then the optional stopping theorem implies $10 = 25p + (1-p) \cdot 0$ and solving gives $p = .4$. Each person has a .4 chance reach 25, so (by additivity of expectation) the expected number is $.4 \cdot 10 = 4$.

(c) Compute the expected number of islanders who *never* have more than 10 apples. **ANSWER:** Same analysis as (b) but replace 0 and 25 with 0 and 11. We find that the expected number who reach 0 before 11 is $10 \cdot \frac{1}{11} = 10/11$.

(d) Compute the probability that the overall winner is someone who at some point in the game only had one apple. (Hint: let B_j be the event that the jth islander's apple count drops to 1 before subsequently rising to 100. Observe that $B_1, B_2, \ldots, B_{10}$ are disjoint.) **ANSWER:** By similar analysis to (b) and (c), the jth islander has a $90/99 = 10/11$ chance to reach 1 before 100, and *given* that has a 1/100 chance to reach 100. Multiplying, we find $P(B_j) = 1/110$, and since the B_j are disjoint we have $P(\cup_{j=1}^{10} B_j) = 1/11$.

3. (10 points) Detective Irene has effective techniques for inducing people to confess to crimes. When Irene interrogates a guilty person, that person confesses with probability .9. Unfortunately, Irene's techniques (extended isolation, claiming confession in best interest, etc.) sometimes lead innocent people to confess. When Irene interrogates an innocent person, that person confesses with probability .1. There is a group of 10 people, and it is known that exactly one is guilty of a crime. Irene has a plan to catch the guilty party. Each day, she will pick one of the 10 people (uniformly at random) and interrogate that person. She will continue this every day until somebody confesses, at which point the investigation will end and the confessing individual will be locked up. (Note: the same person may be interrogated multiple times. But a person's probability of confessing during an interrogation is always the same — i.e., .9 if guilty, .1 if innocent — independently of what has happened before.)

(a) Compute the probability that a confession is obtained on the first day. **ANSWER:** $.9 \cdot .1 + .1 \cdot .9 = .18$

(b) Compute the conditional probability that the person interrogated on the first day is guilty *given* that the person confessed. **ANSWER:** $.9 \cdot .1/.18 = 1/2$.

(c) Let N be the total number of interrogations performed (including the final interrogation, the one that produces the confession). Compute $P(N = k)$ for $k \in \{1, 2, 3, \ldots\}$ and compute $E[N]$. **ANSWER:** $P(N = k) = (1 - .18)^{k-1} \cdot .18$. This is geometric with parameter $p = .18$ so $E[N] = 1/p = 100/18 = 50/9$.

(d) What is the overall probability that the person locked up at the end is guilty? **ANSWER:** *Given* that the confession occurred on the kth day, the conditional probability that the person was guilty is 1/2, by similar argument as in (b). Since this is true for any day, the overall probability is 1/2.

4. (10 points) Suppose 8 people toss their hats into a bin. The hats are randomly shuffled (all shufflings equally likely) and returned to the people, one hat per person. But there is an additional twist: while in

the bin, each hat has a 1/2 probability (independently of all else) of falling into a muddy corner of the bin and getting dirty.

(a) Let D be the number of hats that get dirty. Compute $E[D]$ and $\mathrm{Var}[D]$. **ANSWER:** D is binomial with $n = 8$ and $p = 1/2$. We have $E[D] = np = 8/2 = 4$ and $\mathrm{Var}[D] = npq = 8(1/2)(1/2) = 2$.

(b) Let N be the number of people who get back their own hat. Let N^* be the number of people who get their own hat back *and* find that hat to be clean (i.e., not dirty). Compute $E[N^*]$ and $\mathrm{Var}[N^*]$. (In case this notation helps: let N_i^* be 1 if the ith person gets own hat *and* finds it clean, and 0 otherwise, so that $N^* = \sum_{i=1}^{8} N_i^*$.) **ANSWER:** $E[N^*]^2 = \sum_{i=1}^{8}\sum_{j=1}^{8} E[N_i^* N_j^*]$. There are 64 terms, and 8 of them have $i = j$, 56 have $i \neq j$. So $E[N^*]^2 = \sum_{i=j} P(N_i^* = 1) + \sum_{i\neq j} P(N_i^* = N_j^* = 1) = 8 \cdot (1/8)(1/2) + 56(1/56)(1/2)(1/2) = 1/2 + 1/4 = 3/4$. Similar arguments give $E[N^*] = 8(1/8)(1/2) = 1/2$. So $\mathrm{Var}[N^*] = 3/4 - (1/2)^2 = 1/2$.

(c) Let C be the total number of hats that stay clean. Compute the correlation coefficients $\rho(C, D)$ and $\rho(N, C)$. (Hint: this problem should not require a lot of computation.) **ANSWER:** $\rho(C, D) = \rho(C, 8 - C) = \rho(C, -C) = -1$. (Note generally that if X and Y are random variables and c is constant then $\rho(X, Y) = \rho(X, Y + c)$.) And $\rho(N, C) = 0$ since N and C are independent.

5. (10 points) A certain biotech company has a distinctive corporate culture. Each employee has a "level" of 1, 2, 3, 4, or 5. At the end of each year, each employee of level j is assigned a new level in the following way:

1. If $j \in \{1, 2, 3, 4\}$ then the new level is j with probability 1/2 and $j + 1$ with probability 1/2. ("All non-top-level employees have even odds of being promoted each year," reads the company brochure.)

2. If $j = 5$ then the new level is 1 probability 1. ("All top-level employees return to their bottom-level roots.")

(a) Interpret this as a Markov chains and write the corresponding transition matrix. **ANSWER:**

$$\begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(b) Over the long term, what fraction of the time does an employee spend in each of the 5 states?

$$\begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix} \cdot \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix}.$$

Solving gives

$$\begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix} = \begin{pmatrix} 2/9 & 2/9 & 2/9 & 2/9 & 1/9 \end{pmatrix}$$

(c) If an employee starts out in level 1, how many promotion cyles will it take in expectation before the employee reaches state 5? More formally: if A_n is the employee's rank during year n and we are given that $A_0 = 1$ then what is $E[\min\{n : A_n = 5\}]$? **ANSWER:** The number of cycles required for each promotion is geometric with parameter $p = 1/2$, hence expectation $1/p = 2$. Expected number of cycles for four promotions is $4 \cdot 2 = 8$.

6. (10 points) Let X be a uniform random variable on the interval $[0, 10]$. For each real number K write $C(K) = E[\max\{X - K, 0\}]$.

(a) Compute $C(K)$ as a function of K for $K \in [0, 10]$. (Hint: you might find that the integral you need to compute to find $C(K)$ is the area of a triangle.) **ANSWER:**
$C(K) = \int_0^{10} \frac{1}{10} \max\{x - K, 0\}dx = \frac{1}{10} \int_K^{10} (x - K)dx = \frac{1}{10}(10 - K)^2/2$.

(b) Compute the derivatives C' and C'' on the interval $[0, 10]$. **ANSWER:** Taking derivative directly gives $C'(x) = -\frac{1}{10}(10 - x) = x/10 - 1$ and $C''(x) = \frac{1}{10}$. Alternatively, one could remember the call function formulas derived in lecture: $C'(x) = F_X(x) - 1$ and $C''(x) = f_X(x)$.

(c) Compute the expectation $E[X^3]$. **ANSWER:** $\int_0^{10} \frac{1}{10} x^3 dx = \frac{1}{10} x^4/4\Big|_0^{10} = \frac{1}{10} 10^4/4 = 250$.

7. (10 points) Let $X_1, X_2, X_3, \ldots$ be independent exponential random variables, each with parameter $\lambda = 1$.

(a) Let c be a fixed constant and write $Y_n = \left(\sum_{i=1}^n X_i^3\right) - cn$. (So $Y_0 = 0$.) For which (if any) values of c is the sequence $Y_0, Y_1, Y_2, \ldots$ a martingale? **ANSWER:** $Y_n = \sum_{i=1}^n (X_i^3 - c)$ is a cumulative sum of i.i.d. terms and is a martingale only if the terms have expectation zero. Since in general $E[X^n] = n!$ we have $E[X_i^3 - c] = 6 - c$ which is zero if and only if $c = 6$.

(b) Compute the probability $P\big(X_1 + X_2 + X_3 < 2 \text{ and } X_1 + X_2 + X_3 + X_4 > 2\big)$. (Hint: try to come up with a Poisson point process interpretation of the question.) **ANSWER:** The points $X_1, X_1 + X_2, X_1 + X_2 + X_3, X_1 + X_2 + X_3 + X_4, \ldots$ form a Poisson point process of rate 1. So answer is the probability that a Poisson point process of rate 1 has three points in $[0, 2]$ (with the fourth point coming after 2). This is given by $e^{-\lambda}\lambda^k/k!$ with λ set to 2 and k set to 3.

(c) Compute the correlation coefficient $\rho(X_1 + X_2 + X_3, X_2 + X_3 + X_4)$. **ANSWER:** Each X_i has variance one. By independence and covariance bilinearity $\mathrm{Cov}(X_1 + X_2 + X_3, X_2 + X_3 + X_4) = 2$ and $\mathrm{Var}(X_1 + X_2 + X_3) = \mathrm{Var}(X_2 + X_3 + X_4) = 3$ so answer is $2/\sqrt{3 \cdot 3} = 2/3$.

(d) Give the probability density function for $X_1 + X_2 + X_3$. **ANSWER:** This is gamma with $n = 3$ and $\lambda = 1$. Answer is $x^2 e^{-x}/2$.

8. (10 points) Suppose that the pair (X, Y) is uniformly distributed on the unit circle $\{(x, y) : x^2 + y^2 \le 1\}$.

(a) Compute the joint probability density $f_{X,Y}(x, y)$. **ANSWER:** $1/\pi$ in the unit circle, 0 elsewhere.

(b) Compute the marginal probability distribution $f_X(x)$. **ANSWER:** $\frac{1}{\pi} 2\sqrt{1 - x^2}$

(c) Compute $E[R]$ where $R = \sqrt{X^2 + Y^2}$. (Hint: maybe use a polar coordinates integral. Or maybe find a way to compute F_R and/or f_R without doing that.) **ANSWER:**
$F_R(a) = P(R \le a) = (\pi a^2)/\pi = a^2$ so $f_R(x) = 2x$ so $E[R] = \int_0^1 2x^2 dx = 2/3$.

9. (10 points) Andrew and Alyssa want to have children, and are eager to have at least one girl and at least one boy. So they decide they will have children (one at a time) until the first time they either have at least one child of each gender *or* they have four children total. Thus, if we let X denote the gender sequence for this family, then the possible values for X are
$\{GB, GGB, GGGB, GGGG, BG, BBG, BBBG, BBBB\}$. Assume that each child born has a .5 chance to be a girl and .5 chance to be a boy, independently of what has happened before.

(a) Compute the entropy $H(X)$. (The answer is a rational number. Give it explicitly.) **ANSWER:** The log probability of an outcome is the number of children in that outcome. So the entropy is the expected number of children, which is $(1/2) \cdot 2 + (1/4) \cdot 3 + (1/4) \cdot 4 = 11/4$.

(b) Describe a strategy for asking a sequence of yes/no questions such that the *expected* number of questions one has to ask to learn the value of X is exactly $H(X)$. **ANSWER:** Optimal approach is to choose questions that have equal probability to be yes or no. There are many ways to do this, but one is to ask the questions "Is first child a girl? Is second child a girl? Is third child a girl? Is fourth child a girl?" but stop once X is known. (For example, if first answer is yes, second answer is no, then one can stop asking questions since it is then clear that $X = GB$ and there are no more children.)

(c) Let $Y \in \{G, B\}$ be the gender of the first child born. Compute $H(Y)$, $H_Y(X)$, and $H(X, Y)$. Is it true that $H(X, Y) = H(Y) + H_Y(X)$ in this setting? **ANSWER:** $H(Y) = 1$. And $H_Y(X) = \frac{1}{2}H_{Y=G}(X) + \frac{1}{2}H_{Y=B}(X)$. By symmetry of B and G this is the same as $H_{Y=G}(X)$ (which is the conditional entropy given the first child is a girl). By same logic as above, this is just the expected number of children (excluding the first) given that the first child is a girl, which comes out to $7/4$. And $H(X, Y) = H(X) = 11/4$ (since X determines Y, the pair (X, Y) contains no more information than X does alone). So the identity holds.

10. (10 points) Let X_1, X_2, X_3, X_4, X_5 be i.i.d. random variables, each with probability density function given by $f(x) = \frac{1}{\pi(x^2+1)}$.

(a) Compute the probability $P(\max\{X_1, X_2\} > \max\{X_3, X_4, X_5\})$. **ANSWER:** This just the probability that the largest of the five elements is either X_1 or X_2. Since each of the X_i is equally likely to be largest, the answer is $2/5$.

(b) Let N be the number of $j \in \{1, 2, 3, 4, 5\}$ for which $X_j > 0$. (So N is a random element of the set $\{0, 1, 2, 3, 4, 5\}$.) Compute the moment generating function $M_N(t)$. **ANSWER:** Let N_j be 1 if $X_j > 0$ and 0 otherwise. Then $M_{N_j}(t) = E[e^{tN_j}] = \frac{1}{2}e^0 + \frac{1}{2}e^t$. And $M_N(t) = M_{N_j}(t)^5 = \left(\frac{1}{2} + \frac{1}{2}e^t\right)^5$.

(c) Compute the probability $P(X_1 + X_2 > X_3 + X_4 + X_5 + 5)$. **ANSWER:** Note that by symmetry $X_1 + X_2 - X_3 - X_4 - X_5$ has same law as $X_1 + X_2 + X_3 + X_4 + X_5$ which in turn has same law as $5X_1$ (by a special property of Cauchy random variables). So answer is equivalent to $P(5X_1 > 5) = P(X_1 > 1)$ which in turn (recall spinning flashlight story) is $1/4$.