Spring 2022 18.600 Final Exam Solutions

1. (10 points) Alice is writing a computer program for a class she is unsure about taking. The program has 4 bugs but Alice is an excellent debugger. Every hour she makes a major change to the code. Each change has a 2/3 chance of eliminating a bug and a 1/3 chance of adding a new bug. She plans to keep working until either she reaches 0 bugs (at which point she will submit her code) or 8 bugs (at which point she will abandon the code and drop the class). Formally, let X_n be the number of bugs after n hours. Then $X_0 = 4$ and if $X_n \in \{1, 2, \dots, 7\}$ then X_{n+1} is $X_n - 1$ (with probability 2/3) or $X_n + 1$ (with probability 1/3). If $X_n \in \{0, 8\}$ then $X_{n+1} = X_n$ with probability 1. Let T be the total number of hours Alice spends coding, i.e., T is the smallest n with $X_n \in \{0, 8\}$.

(a) Which of the following is a martingale? Circle the martingales, no need to explain.

(i) $M_n = \begin{cases} X_n + n/3 - 4 & n \le T \\ 0 & n > T \end{cases}$

(ii) $M_n = \begin{cases} X_n + n/3 & n \le T \\ X_T + T/3 & n > T \end{cases}$

(iii) $M_n = 2^{X_n}$

(iv) $M_n = 17$

ANSWER: (i) is not a martingale because M_T may not be zero, but M_{T+1} is always zero; so one cannot say $E[M_{n+1}|\mathcal{F}_n] = M_n$ if n has a positive chance of being equal to T. (ii), (iii) and (iv) are martingales: *given* what one knows at stage n, the expectation of M_{n+1} is M_n.

(b) Compute the probability the code is fixed, i.e. $P(X_T = 0)$. **Hint:** use one of the martingales from (a). **ANSWER:** Call the answer p and write $M_n = 2^{X_n}$. Then $M_0 = 16$. Also M_T is 1 with probability p and 256 otherwise. Optional stopping gives $E[M_T] = M_0$ so $(1-p) \cdot 256 + p \cdot 1 = 16$ so $256 - 256p + p = 16$ and $p = 240/255 = 16/17$.

(c) Find the expected number of hours Alice works, i.e. $E[T]$. **Hint:** first use (b) to compute $E[X_T]$. Then use another martingale from (a). **ANSWER:** (b) gives $E[X_T] = 8 \cdot P(X_T = 8) = 8/17$. Taking M_n from (ii). Optional stopping theorem (OST) gives $E[M_T] = E[X_T] + \frac{1}{3}E[T] = \frac{8}{17} + \frac{1}{3}E[T] = M_0 = 4$. Solving gives $E[T] = 3(4 - \frac{8}{17}) = 12 - \frac{24}{17} = 10 + \frac{10}{17}$. *Technical point:* since M_n is not bounded, the "bounded martingale" OST does not directly apply. But if we define $\tilde{T} = \min\{T, N\}$ for some fixed N, then the OST gives $E[M_{\tilde{T}}] = M_0$. One can then argue that $E[M_{\tilde{T}} - M_T]$ tends to zero as $N \to \infty$ using the fact that $P(T > K)$ decays exponentially in K.

2. Bob is a 7th grade student whose school assigns grades randomly. He has 7 classes. In each class, he gets an A with probability 1/2, a B with probability 1/4, a C with probability 1/8, and a D or F each with probability 1/16 (independently of all other grade assignments). Let G_i be the grade that Bob gets in his ith class, so that $G = (G_1, G_2, \dots G_7)$ constitutes Bob's entire report card.

(i) Find the entropy $H(G_1)$ and $H(G)$. (The answers are rational numbers: give them explicitly.) **ANSWER:** $H(G_1) = \frac{1}{2}(-\log\frac{1}{2}) + \frac{1}{4}(-\log\frac{1}{4}) + \frac{1}{8}(-\log\frac{1}{8}) + 2\frac{1}{16}(-\log\frac{1}{16}) = \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 3 + 2\frac{1}{16}\cdot 4 = \frac{15}{8}$. G is a list of 7 i.i.d. instances of G_1 so we have $H(G) = 7H(G_1) = 105/8$.

(ii) Bob's nosy sister Carol wants to know Bob's grades. She decides to find out by asking Bob a series or yes-or-no questions. Give a strategy that allows Carol to determine G with the smallest possible expected number of yes-or-no questions. How many questions does she expect to ask? **ANSWER:** She can ask (annoyingly) "Did you get an A? Did you get a B? Did you get a C? Did you get a D?" for each class, stopping when she knows the answer. The expected number of questions is exactly $H(G) = 105/8$.

(iii) Compute the probability that Bob gets exactly 3 A's, 2 B's and 2 C's. **ANSWER:** $\frac{7!}{3!2!2!} \cdot (\frac{1}{2})^3(\frac{1}{4})^2(\frac{1}{8})^2$

(iv) Give the conditional probability that Bob gets exactly 3 A's *given* that all Bob's grades are A or B. **ANSWER:** $\binom{7}{3}(\frac{2}{3})^3(\frac{1}{3})^4$ since after conditioning, each grade is A with probability 2/3, B otherwise.

3. Alvin and Beatrice are applying to out-of-state medical schools. Each applies to 50 schools; because of their strong credentials, each of Alvin and Beatrice has (independently of all else) a 10 percent chance of being offered a position at each school. So Alvin and Beatrice each *expect* to receive 5 acceptances. Let A be the number to which Alvin is accepted and let B be the number to which Beatrice is accepted.

(a) Compute the mean and variance of the difference $B - A$. **ANSWER:** $E[B-A] = E[B] - E[A] = 0$ and $\mathrm{Var}(B-A) = \mathrm{Var}(B) + \mathrm{Var}(A) = 2 \cdot (50 \cdot \frac{1}{10} \cdot \frac{9}{10}) = 9$.

(b) Use a normal approximation to estimate $P[B - A > 1.5]$. (This is the probability that Beatrice's acceptance number exceeds Alvin's by at least 2.) You may use $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$. **ANSWER:** The probability $B - A$ is $\frac{1}{2}$ SDs above mean, roughly $1 - \Phi(1/2) = \Phi(-1/2) \approx .31$.

(c) Use a Poisson random variable to estimate the probability that there is at least one medical school at which *both* Alvin and Beatrice are accepted. **ANSWER:** Expected number of both-accepted schools is $\lambda = 50 \cdot (\frac{1}{10})^2 = .5$. Probability is roughly $1 - e^{-\lambda} = 1 - e^{-.5} \approx .39$.

4. Carol is going to a fancy dinner with important clients. She wants to order a bottle of wine for the table. There are 5 types of wine, and each is independently assigned a random price which (in dollars) is an exponential random variable with parameter $\lambda = 1/100$. Let X_i be the price of the ith type of wine. Alice does not know or care much about fine wine, but in order to convey an appropriate impression (not too extravagant, not too cheap) she plans to order a bottle of the 3rd most expensive wine, whatever that turns out to be. Let X be the price of the bottle she orders.

(a) Compute the expectation $E[X]$. **ANSWER:** Viewing "wine prices" as "times" this is the "radioactive decay" problem from lecture. Time till first event is exponential with parameter 5λ, subsequent time till second is exponential with parameter 4λ, and subsequent time till third exponential with parameter 3λ. Overall expectation is $\frac{1}{5\lambda} + \frac{1}{4\lambda} + \frac{1}{3\lambda} = 100(\frac{1}{5} + \frac{1}{4} + \frac{1}{3}) = 20 + 25 + 33 + \frac{1}{3} = 78 + \frac{1}{3}$.

(b) Let C be the price of the *cheapest* of the 5 wines. Compute the expectation $E[C]$ and variance $\mathrm{Var}(C)$. **ANSWER:** C exponential with parameter 5λ so $E[C] = 20$ and $\mathrm{Var}(C) = 400$.

(c) Suppose there are 5 fancy appetizers whose prices $A_1, A_2, \ldots, A_5$ are independent uniform random variables on $[0, 100]$. Carol's client David plans to request that the most expensive appetizer be ordered to share with the table. Compute the probability density function for $A = \max\{A_1, A_2, \ldots, A_5\}$. **ANSWER:** If $x \in [0, 100]$ then $F_A(x) = P(A \le x) = (x/100)^5$ so $f_A(x) = 5 \cdot \frac{1}{100} \cdot (x/100)^4$.

(d) There are 4 more reasonably priced desserts whose prices D_1, D_2, D_3, D_4 are i.i.d. normal random variables, each with mean 10 and variance 1. Alice decides to order one of each and allow people to share. Compute the probability density function for $D = D_1 + D_2 + D_3 + D_4$. **ANSWER:** $\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ with $\sigma^2 = 4$, $\mu = 40$.

5. (10 points) Eve has a complicated relationship with sleep. She stays up late and her bedtime shifts a lot. Every night she goes to sleep at either 1am, 2am, 3am, 4am or 5am. Let $X_n \in \{1, 2, 3, 4, 5\}$ be the hour at which she goes to sleep on the nth night. Assume $X_0 = 1$. Furthermore:

If $X_n = 1$ then X_{n+1} is 1 with probability 1/2 and 2 with probability 1/2.

If $X_n = 2$ then X_{n+1} is 1 with probability 1/2 and 3 with probability 1/2.

If $X_n = 3$ then X_{n+1} is 2 with probability 1/2 and 4 with probability 1/2.

If $X_n = 4$ then X_{n+1} is 3 with probability 1/2 and 5 with probability 1/2.

If $X_n = 5$ then X_{n+1} is 4 with probability 1/2 and 5 with probability 1/2.

(a) Write down the Markov chain transition matrix describing Eve's sleep pattern. **ANSWER:**

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

(b) Compute the probability $P(X_5 = 5)$. **ANSWER:** There are 2 ways to get from 1 to 5 in 5 steps (either $1 \to 1 \to 2 \to 3 \to 4 \to 5$ or $1 \to 2 \to 3 \to 4 \to 5 \to 5$) so probability is $2 \cdot (1/2)^5 = 1/16$.

(c) Over the long term, what fraction of time does Eve go to sleep at each of the 5 times? **ANSWER:** The stationary vector π with $\pi A = \pi$ is given by $(1/5, 1/5, 1/5, 1/5, 1/5)$.

6. (10 points) While touring a secure facility, 10 senators deposit their black iPhones at the front desk. After the tour, the phones are randomly returned—one per senator, with all 10! permutations being equally likely.

(a) Let S be the number of senators who get their own phone. Compute the mean and variance of S. If it helps you can let S_i be 1 if the ith phone goes to its owner, 0 otherwise. **ANSWER:** This is equivalent to the "hat problem." If X_i is 1 when ith person gets own phone, when $E[S] = E[\sum X_i] = 10 \cdot \frac{1}{10} = 1$. Also $E[X^2] = \sum_{i=1}^{10} \sum_{j=1}^{10} E[X_i X_j]$. The 10 "diagonal" $i = j$ terms contribute 1/10 each, while the 90 "off diagonal" $i \neq j$ terms contribute $\frac{1}{10} \cdot \frac{1}{9} = \frac{1}{90}$ each, so $E[X^2] = 2$ and $\text{Var}(X) = E[X^2] - E[X]^2 = 1$.

(b) Suppose there are 5 Democrat and 5 Republican senators. Let X the number of Democrat senators who end up with Republican cell phones. Compute $E[X]$ and $E[X^2]$. If it helps, you can let X_i be 1 if the ith Democrat gets a Republican cell phone, 0 otherwise. **ANSWER:** $E[X] = \sum_{i=1}^{5} E[X_i] = 5/2$. $E[X^2] = \sum_{i=1}^{5} \sum_{j=1}^{5} E[X_i X_j]$. The 5 "diagonal" $i = j$ terms contribute $\frac{5}{10} = \frac{1}{2}$ each while the 20 "off-diagonal" contribute $\frac{5}{10} \cdot \frac{4}{9} = \frac{2}{9}$ So $E[X^2] = 5/2 + 40/9$.

(c) Let Y be the number of Republican senators who end up with Democrat cell phones. Compute the correlation coefficient $\rho(X, Y)$. (Hint: this problem should not require a lot of computation.) **ANSWER:** $X = Y$ so $\rho(X, Y) = 1$.

7. (10 points) Let X be a uniform random variable on the interval $[0, 1]$. For each real number K write $C(K) = E[\max\{X - K, 0\}]$.

(a) Compute $C(K)$ as a function of K for $K \in [0, 1]$. **ANSWER:** $E[\max\{X - K, 0\}] = \int_0^1 \max\{x - K, 0\}dx = \int_K^1 x dx = (1 - K)^2/2$.

(b) Compute the derivatives C' and C'' on the interval $[0, 1]$. **ANSWER:** Write $C(x) = (1 - x)^2/2$. Then $C'(x) = -(1 - x) = x - 1$ and $C''(x) = 1$. Can compute this either directly or by recalling from lecture notes that $C'(x) = F_X(x) - 1$ and $C''(x) = f_X(x)$.

(c) Compute the expectation $E[\sin(X)]$. **ANSWER:** $\int_0^1 \sin(x)dx = -\cos(1) - \big(-\cos(0)\big) = 1 - \cos(1)$.

(d) Derive the moment generating function $M_X(t) := E[e^{tX}]$. Show your work. **ANSWER:** $E[e^{tX}] = \int_0^1 e^{tx}dx = \frac{1}{t}e^{1\cdot t} - \frac{1}{t}e^{0\cdot t} = (e^t - 1)/t$.

8. (10 points) Suppose X and Y are independent uniform random variables on $[0,1]$ and write $Z = X + Y$.

(a) Compute the joint probability density $f_{X,Y}(x,y)$ and the joint probability density function $f_{X,Z}(x,z)$. **ANSWER:** $f_{X,Y}$ is 1 on the box $[0,1]^2$ and zero elsewhere. $f_{X,Z}$ is 1 on the parallelogram of (x,z) pairs with $x \in [0,1]$ and $z \in [x, x+1]$ and zero elsewhere.

(b) Compute the marginal probability distribution $f_Z(z)$. **ANSWER:** As shown in lecture:

$$f_Z(z) = \begin{cases} z & z \in [0,1] \\ (2-z) & z \in (1,2] \\ 0 & z \notin [0,2] \end{cases}.$$

(c) Compute the conditional probability $P(Z > 1 | X > 1/2)$. **ANSWER:** One can draw a picture to see that the part of $[0,1]^2$ with $x + y > 1$ and $x > 1/2$ has area $3/8$. So $P(Z > 1, X > 1/2)/P(X > 1/2) = (3/8)/(1/2) = 3/4$.

(d) Compute the mean and variance of Z. **ANSWER:** $E[Z] = E[X] + E[Y] = 1$ and $\text{Var}(Z) = \text{Var}(X) + \text{Var}(Y) = 1/6$ because $\text{Var}(Y) = \text{Var}(X) = E[X^2] - E[X]^2 = \int_0^1 x^2 dx - (1/2)^2$.

9. (10 points) Let $X_1, X_2, X_3, \ldots X_5$ be i.i.d. each with probability density function given by $f(x) = \frac{1}{\pi(x^2+1)}$.

(a) Compute the probability $P(X_1 - X_2 + X_3 - X_4 + X_5 > 5)$. ANSWER: Since X_i has same law as $-X_i$, this is the same as $P(\sum_{i=1}^5 X_i > 5) = P(A > 1)$ where A is the average of the X_i. Since X_i are Cauchy, A is also Cauchy, so $P(A > 1) = 1/4$.

(b) Let N be the number of $j \in \{1,2,3,4,5\}$ for which $X_j > 1$. (So N is a random element of the set $\{0,1,2,3,4,5\}$.) Compute the moment generating function $M_N(t)$. **ANSWER:** Write $N = \sum_{j=1}^n N_j$ where $N_j = 1$ if $X_j > 1$, 0 otherwise. $M_{N_1}(t) = \frac{3}{4} + \frac{1}{4}e^t$ so $M_N(t) = \left(\frac{3}{4} + \frac{1}{4}e^t\right)^5$.

(c) Compute the probability that we have both $X_1 < X_2 < X_3 < X_4$ *and* $X_1 < X_5$. **ANSWER:** All 5! relative rankings of the five elements X_i are equally likely, but only 4 satisfy the conditions: namely, $X_1 < X_5 < X_2 < X_3 < X_4$ and $X_1 < X_2 < X_5 < X_3 < X_4$ and $X_1 < X_2 < X_3 < X_5 < X_4$ and $X_1 < X_2 < X_3 < X_4 < X_5$. So the answer is $4/120 = 1/30$.

10. (10 points) Sam is listening to random songs on his fancy head phones. Each song has length 3 minutes (with probability 1/3), 4 minutes (with probability 1/3) or 5 minutes (with probability 1/3). Let S be the combined duration (in minutes) of the first 24 songs he listens to.

(a) Compute $E[S]$ and $\text{Var}(S)$. **ANSWER:** $E[S] = 24 \cdot (\frac{1}{3}3 + \frac{1}{3}4 + \frac{1}{3}5) = 96$. Variance for one length of one song is $2/3$, so $\text{Var}(S) = \frac{2}{3} \cdot 24 = 16$, since variance is additive for independent random variables.

(b) Compute the moment generating function $M_S(t)$. **ANSWER:** $\left(\frac{e^{3t}+e^{4t}+e^{5t}}{3}\right)^{24}$.

(c) Use the central limit theorem to approximate $P(S > 100)$. You may use the function $\Phi(a) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ in your answer. **ANSWER:** Since 100 is one standard deviation above the mean, $P(S > 100) \approx 1 - \Phi(1) = \Phi(-1)$.

(d) Let S_n be the total length of the first n songs heard. Compute the correlation coefficient $\rho(S_{100}, S_{400})$. **ANSWER:** $\frac{\text{Cov}(S_{100},S_{100})}{\sqrt{\text{Var}(S_{100})\text{Var}(S_{400})}} = \frac{\text{Var}(S_{100})}{\sqrt{\text{Var}(S_{100})\text{Var}(S_{400})}} = \frac{\frac{2}{3}100}{\sqrt{\frac{2}{3}100\frac{2}{3}400}} = \frac{1}{2}$