NAME: ______________________________

Spring 2022 18.600 Final Exam: 100 points

Carefully and clearly show your work on each problem (without writing anything that is technically not true) and put a box around each of your final computations.

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 $X_n \in \{1, 2, \ldots, 7\}$ 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) $M_n = \begin{cases} X_n + n/3 - 4 & n \leq T \\ 0 & n > T \end{cases}$

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

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

(iv) $M_n = 17$

(b) Compute the probability the code is fixed, i.e. $P(X_T = 0)$. **Hint:** use one of the martingales from (a).

(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).

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 $G = (G_1, G_2, \ldots G_7)$ constitutes Bob's entire report card.

(i) Find the entropy $H(G_1)$ and $H(G)$. (The answers are rational numbers: give them explicitly.)

(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?

(iii) Compute the probability that Bob gets exactly 3 A's, 2 B's and 2 C's.

(iv) Give the conditional probability that Bob gets exactly 3 A's *given* that all Bob's grades are A or B.

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 $B - A$.

(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 $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.

(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.

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 $E[X]$.

(b) Let C be the price of the *cheapest* of the 5 wines. Compute the expectation $E[C]$ and variance $\mathrm{Var}(C)$.

(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 $A = \max\{A_1, A_2, \ldots, A_5\}$.

(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 $D = D_1 + D_2 + D_3 + D_4$.

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.

(b) Compute the probability $P(X_5 = 5)$.

(c) Over the long term, what fraction of time does Eve go to sleep at each of the 5 times?

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.

(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.

(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.)

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 $K \in [0, 1]$.

(b) Compute the derivatives C' and C'' on the interval $[0, 1]$.

(c) Compute the expectation $E[\sin(X)]$.

(d) Derive the moment generating function $M_X(t) := E[e^{tX}]$. Show your work.

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 $f_{X,Z}(x,z)$.

(b) Compute the marginal probability distribution $f_Z(z)$.

(c) Compute the conditional probability $P(Z > 1 | X > 1/2)$.

(d) Compute the mean and variance of Z.

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

(a) Compute the probability $P(X_1 - X_2 + X_3 - X_4 + X_5 > 5)$.

(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 $M_N(t)$.

(c) Compute the probability that we have both $X_1 < X_2 < X_3 < X_4$ *and* $X_1 < X_5$.

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 $\mathrm{Var}(S)$.

(b) Compute the moment generating function $M_S(t)$.

(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.

(d) Let S_n be the total length of the first n songs heard. Compute the correlation coefficient $\rho(S_{100}, S_{400})$.