Spring 2021 18.600 Final Exam: 100 points

1. **Carefully and clearly *show your work* on each problem (without writing anything that is technically not true). In particular, if you use any known facts (or facts proved in lecture) you should state clearly what fact you are using and why it applies.**

2. **No calculators, books, notes or other resources may be used except for the story sheet at** http://math.mit.edu/~sheffield/2018600/storysheet.pdf

3. **Simplify your answers as much as possible (but answers may include factorials and $\binom{n}{k}$ expressions — no need to multiply them out).**

NAME: ________________________________

1. (10 points) An online class contains 5 first-years, 5 sophomores, 5 juniors and 5 seniors. An instructor randomly divides the students into 4 numbered breakout rooms of 5 students each (with all such divisions being equally likely). Let S_i be the number of seniors in the ith breakout room. Alice and Bob are both seniors in the class.

(a) Compute the probability that Alice is in room 1 and Bob is in room 2.

(b) Compute $\mathrm{Cov}(S_1, S_2) = E(S_1 S_2) - E(S_1)E(S_2)$.

(c) Write $A = S_1 + S_2$ and compute the conditional expectation $E[S_1|A]$ as a function of A.

(d) Compute the correlation coefficient $\rho(S_1 + S_2, S_3 + S_4)$.

2. (10 points) Each customer at Alice's ice cream cone stand independently orders one scoop with probability 1/3, two scoops with probability 1/3 and three scoops with probability 1/3. For each $j \geq 1$ let X_j be the number of scoops the jth customer orders. Write $S = \sum_{i=1}^{150} X_i$ for the total number of scoops ordered by the first 150 customers. Compute the following:

(a) The characteristic function $\phi_{X_1}(t)$.

(b) The characteristic function $\phi_S(t)$.

(c) The variance $\mathrm{Var}(S)$.

(d) Use the central limit theorem to approximation the probability $P(S \leq 320)$. You may use the function $\Phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ in your answer.

3. (10 points) A persistent rock climber is climbing a difficult wall in a rock-climbing gym. The wall has four hand-holds of increasing heights. At each time the climber is in one of five states:

1. State 0: On the mat below the climbing wall.
2. State 1: Holding onto the first hold of the climbing wall.
3. State 2: Holding onto the second hold.
4. State 3: Holding onto the third hold.
5. State 4: At the top celebrating.

Every five seconds (with precise regularity) the climber transitions from one state to another according to the following rules. Whenever the climber is in State 0, the climber transitions to State 1. Whenever the climber is in state i for $i \in \{1, 2, 3\}$ the climber transitions to State $i + 1$ with probability 1/3 and back to State 0 with probability 2/3. Whenever the climber is in State 4, the climber transitions back to State 0.

(a) Write the Markov chain transition matrix corresponding to the climber's state evolution.

(b) If the climber is in State 0 at time 0, what is the probability the climber will be back in State 0 after four transition steps?

(c) Over the long term, what fraction of time does the climber spend in each of the 5 states?

4. (10 points) On Open-Minded Planet, there are 10^6 people, each of whom holds *either* Opinion A (which happens to be wrong) or Opinion B (which happens to be correct). At time zero 13 unlucky people have acquired Opinion A but $999,987$ have Opinion B. During each subsequent time unit, two people (chosen uniformly at random) have a discussion. If they have the same opinion, nothing changes. If they disagree, they talk until one convinces the other—and both parties end up adopting Opinion A with probability $1/2$ and Opinion B with probability $1/2$ (where this "fair coin toss" is independent of prior tosses). Let X_n be the number of people holding Opinion A after n conversations. One can show (no need to prove) that with probability one X_n *eventually* reaches 0 or 10^6 at which point everyone agrees.

(a) Is the sequence X_n a martingle? Why or why not?

(b) What is the probability X_n eventually reaches 10^6 (so a wrong opinion becomes consensus)?

Slightly Biased Planet is the same as above except that when two people of opposite opinions argue, they end up both adopting Opinion A with probability $.52 = 13/25$ and Opinion B with probability $.48 = 12/25$. Some pervasive bias makes the wrong opinion *slightly* easier to argue. Let X_n denote the total number of people holding Opinion A after n conversations (again $X_0 = 13$).

(c) Is there a constant $c \in (0, 1)$ for which $M_n := c^{X_n}$ is a martingale? If so, what is c?

(d) What is the probability that X_n eventually reaches 10^6? *Numerically estimate* this probability by using the approximations $(1 - \frac{1}{13})^{13} \approx .36 \approx e^{-1}$ and $(1 - \frac{1}{13})^{10^6} \approx 0$.

5. (10 points) Let X_1, X_2, X_3, X_4 be independent exponential random variables with parameter $\lambda = 1$. Write $M = \max\{X_1, X_2, X_3, X_4\}$ and $W = \min\{X_1, X_2, X_3, X_4\}$.

(a) Compute the expectation and variance of W.

(b) Compute the expectation and variance of M.

(c) Compute the expectation $E[X_1 X_2^2 X_3^3 X_4^4]$.

6. (10 points) Suppose that (X, Y) has probability density function $f_{X,Y}(x, y) = \frac{1}{2\pi} e^{(-x^2-y^2)/2}$.

(a) Compute the conditional probability $P(X > Y | X > 0)$.

(b) Compute the probability density function of $Z = X + 2Y$.

(c) Compute the expectation $E[X^3Y^3 + X^2Y^2 + XY]$. Given an explicit numerical answer.

7. (10 points) In a city with two political tribes, one tribe expresses genuine frustration with the other by posting 10 articles a day in its news outlet. These articles are surprisingly similar from day to day. Each article is based on one of 6 standard "templates." If $X = (X_1, X_2, \ldots, X_{10})$ is the template sequence corresponding to a day's articles, then each X_i is independently equal to

I. ***"They treat us unfairly"*** with probability 1/4,

II. ***"They criticize us hypocritically"*** with probability 1/8,

III. ***"They have dark secret motives"*** with probability 1/8,

IV. ***"We stand up to them"*** with probability 1/4,

V. ***"We mock them"*** with probability 1/8,

VI. ***"We are smarter and nobler than they are"*** with probability 1/8.

(a) Alice does not have time to read all 10 articles, so she programs an AI to categorize the articles and send her just the sequence X. She wants to quantify how many bits of information she learns when she reads X. Compute the entropy $H(X_1)$ and $H(X)$.

(b) Suppose Alice wants to figure out the value of X_1 by asking a series of yes or no questions. Describe a strategy that minimizes the expected number of questions she needs to ask. How many questions does she expect to ask when she uses this strategy?

(c) Let K be the number of times Template V appears in the sequence X. Compute the entropy $H(K)$. You can leave the answer as an unsimplified sum.

8. (10 points) Suppose 10 students take a physical fitness exam. Each student's score on the exam is an independent uniform random variable on $[0, 1]$. Denote the scores by $X_1, X_2, \ldots, X_{10}$. Denote by H the highest score and by T the third highest score.

(a) Compute the probability density function $f_H(x)$.

(b) Compute the probability density function $f_T(x)$.

(c) Compute the conditional probability that the last five scores $X_6, X_7, \ldots, X_{10}$ are in increasing order (i.e., $X_6 < X_7 < \ldots < X_{10}$) *given* that the first six scores are in increasing order (i.e., $X_1 < X_2 < \ldots < X_6$).

(d) Compute the correlation coefficient $\rho\big(\sum_{j=1}^{8} X_j, \sum_{k=3}^{10} X_k\big)$.

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

(a) Compute the probability that all five X_i lie within the interval $[0, 1]$. Give an explicit number. (Recall spinning flashlight story.)

(b) Compute the probability density function for $\frac{1}{3}(X_1 + X_2 + X_3)$.

(c) Compute the probability density function for $Y = X_1 + X_2 + X_3$.

(d) Compute the probability density funtion for $Z = X_1 + 2X_2 + 3X_3 + 4X_4 + 5X_5$.

10. (10 points) A bag of microwave popcorn contains 500 kernels. For each kernel, suppose the length of time (in seconds) until the kernel pops is an independent gamma random variable with parameter $n = 150$ and $\lambda = 1$. Imagine the kernels are numbered from 1 to 500 and let X_i be the amount of time until the ith kernel pops (so the X_i are i.i.d.).

(a) Compute the probability density function for the "pop time average" $A := \frac{1}{500}\sum_{i=1}^{500} X_i$.

(b) Write $G(a) = F_{X_1}(a) = P(X_1 \leq a)$. (You don't have to compute this explicitly.) In terms of the function G, compute the probability density function for the time T at which the last popcorn kernel pops.

(c) In terms of G, compute the density function for the time at which the first kernel pops.