NAME: ______________________________

Spring 2019 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) 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.

(b) Compute the conditional expectation $E\big[(X_1 - Y_1)|X\big]$ as a function of the random variable X.

(c) Compute $E[(X - Y)^2]$.

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.

(b) Compute the expected number of islanders who *at some point* have exactly 25 apples.

(c) Compute the expected number of islanders who *never* have more than 10 apples.

(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.)

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.

(b) Compute the conditional probability that the person interrogated on the first day is guilty *given* that the person confessed.

(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 $E[N]$.

(d) What is the overall probability that the person locked up at the end is guilty?

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 $\mathrm{Var}[D]$.

(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 $N^* = \sum_{i=1}^{8} N_i^*$.)

(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.)

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.

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

(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 $E[\min\{n : A_n = 5\}]$?

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.)

(b) Compute the derivatives C' and C'' on the interval $[0, 10]$.

(c) Compute the expectation $E[X^3]$.

7. (10 points) Let $X_1, X_2, X_3, \dots$ 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 $Y_0, Y_1, Y_2, \dots$ a martingale?

(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.)

(c) Compute the correlation coefficient $\rho(X_1 + X_2 + X_3, X_2 + X_3 + X_4)$.

(d) Give the probability density function for $X_1 + X_2 + X_3$.

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

(a) Compute the joint probability density $f_{X,Y}(x, y)$.

(b) Compute the marginal probability distribution $f_X(x)$.

(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.)

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.)

(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 $H(X)$.

(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?

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 $P(\max\{X_1, X_2\} > \max\{X_3, X_4, X_5\})$.

(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 $M_N(t)$.

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