Spring 2025 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.

NAME: ______________________________

1. (10 points) Imagine a town where weather changes little from day to day. Each day, the daily high temperature is a random element of $\{69, 70, 71, 72, 73\}$ with all five values equally likely. The values are independent from one day to the next. Let $X_1, X_2, \ldots, X_{50}$ be the daily high temperatures of the next 50 days. Let $A = (X_1 + X_2 + \ldots + X_{50})/50$ be the *average* daily high over that period.

(a) Compute $E[X_1]$ and $\text{Var}[X_1]$

(b) Compute $E[A]$ and $\text{Var}[A]$ and the standard deviation of A.

(c) Compute the moment generating functions $M_{X_1}(t)$ and $M_A(t)$.

(d) The residents of this town despise cool weather. Riots will break out if the average daily high of the next 50 days is at or below 70.5. Use the central limit theorem to estimate the probability $P(A \leq 70.5)$. You may use the function $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ in your answer.

2. (10 points) On Surgery Planet, each new surgeon is given a score X at the end of Year 1 and another score at $\tilde{X}$ at the end of Year 2. The scores measure how positive the outcomes were for the surgeon's patients that year. We model the Year 1 score as $X = X_1 + X_2 + \ldots + X_{100}$ and the Year 2 score as $\tilde{X} = X_1 + X_2 + \ldots + X_{28} + \tilde{X}_{29} + \tilde{X}_{30} + \ldots + \tilde{X}_{100}$, where the X_i and $\tilde{X}_i$ are i.i.d. normal with mean 0, variance 1. We interpret $X_1, X_2, \ldots, X_{28}$ as components intrinsic to the particular doctor (surgical dexterity, medical knowledge, hygiene habits, etc.) while the remaining 72 components change from year to year (e.g. due to random fluctuations in the patients who show up). **Note:** you can use the arithmetic $.28^2 + .96^2 = .0784 + .9216 = 1$ in this problem.

(a) Alice wants to know the correlations. Compute $\mathrm{Var}(X)$, $\mathrm{SD}(X)$, $\mathrm{Cov}(X, \tilde{X})$ and $\rho(X, \tilde{X})$.

(b) Bob wants to guess the second score given the first. Express $G = E[\tilde{X}|X]$ as a function of X.

(c) Carol wants to know the **standard deviation** of Bob's guess. Compute $\mathrm{Var}(G)$ and $\mathrm{SD}(G)$.

(d) David overhears somebody say, "Bob's guesses are super imprecise! The standard deviation of the error $R = \tilde{X} - G$ is *almost as large* as the standard deviation of $\tilde{X}$ itself!" David wonders if that is really true. Compute the **variance and standard deviation** of R. **Hint:** You can use without proof that $\mathrm{Cov}(G, R) = 0$ and $\mathrm{Var}(X) = \mathrm{Var}(G + R) = \mathrm{Cov}(G + R, G + R)$. You can also use bilinearity of covariance and the known values of $\mathrm{Var}(G)$ and $\mathrm{Var}(X)$.

3. (10 points) Kiyoshi samples i.i.d. random variables $X_1, X_2, \ldots, X_{100}$. Each X_i has probability density function $f(x) = \frac{1}{\pi(1+x^2)}$.

(a) Compute the probability density function of the average $A = (X_1 + X_2 + \ldots + X_{100})/100$.

(b) Compute the joint probability density function $f(x, y)$ of the pair (X_1, X_2).

(c) Compute the probability density function of $Z = X_1 + 2X_2 + 3X_3 + 4X_4 + 10$.

(d) Compute the probability $P(X_1 > 2X_2 + 3)$.

4. (10 points) Alfred has 4 epic worries (climate change, AI apocalypse, etc.)— labeled 1 to 4. He also has 4 awesome dreams (true love, TikTok fame, etc.) — labeled 5 through 8. Each minute he muses about 1 of these 8 topics. At the end of the minute he updates his thinking as follows:

1. If Alfred is currently thinking about an awesome dream, then with probability 1/2 he keeps thinking about the same dream. With probability 1/4 he switches to a *different* dream (all 3 equally likely). With probability 1/4 he switches to an epic worry (all 4 equally likely).
2. If Alfred is currently thinking about an epic worry, then with probability 1/4 he keeps thinking about the same worry. With probability 1/4 he switches to a *different* worry (all 3 equally likely). With probability 1/2 he switches to an awesome dream (all 4 equally likely).

(a) Write down the 8×8 Markov transition matrix describing Alfred's state of mind.

(b) If Alfred is currently thinking about true love, what is the probability that Alfred will be thinking about true love in two minutes? You can leave your answer as a sum of fractions.

(c) In the long run, what fraction of the time does Alfred spend on each of the 8 topics?

(d) If Alfred is currently thinking about true love, what is the *expected* number of Markov transition steps until he starts thinking about one of his epic worries?

5. (10 points) Alice has for some reason started thinking about the Roman Empire on average once an hour. The times $X_1, X_2, \ldots$ at which these thoughts occur form a Poisson point process with parameter $\lambda = 1$. Denote the *waiting times* by $W_1 = X_1$ and by $W_n = X_n - X_{n-1}$ if $n \geq 2$.

(a) Compute the expectation $E[X_2^3] = E\big[(W_1 + W_2)^3\big]$.

(b) Compute the probability density function for X_7.

(c) Compute the probability Alice has exactly 5 Roman Empire thoughts in the first 7 hours.

(d) Compute the probability density function of $M = \min\{W_1, W_2, W_3\}$.

6. (10 points) Isabella communicates using only words that appear high in her online searches for *gen alpha slang*. Every word she uses is

(i) **Slay**, **Delulu**, **Cap**, **Rizz**, or **Periodt**: each with probability 1/8

(ii) **Bussin** or **Looksmaxxing**: each with probability 1/16

(iii) **Sus**, **Ate**, **Cheugy**, **Yeet**, **Simp**, **Sigma**, **Pookie** or **Drip**: each with probability 1/32.

Her choices are independent from one word to the next. She decides to send a message containing 8 words. Let $X = (X_1, X_2, \ldots, X_8)$ be the sequence of words.

(a) Compute the entropy $H(X_1)$ and $H(X)$.

(b) Suppose you want to determine X with a sequence of yes or no questions. What strategy minimizes the expected number of questions you have to ask? What is the expected number of questions needed to determine X in this case?

(c) Let K be the number of times that **Cap** appears in X. Compute the entropy $H(K)$. You can leave your answer as an unsimplifed sum.

7. (10 points) Let X be a uniform random variable on $[0, 2]$. For each real number K write $C(K) = E[\max\{X - K, 0\}]$.

(a) Compute $C(K)$ as a function of K for $K \geq 0$.

(b) Compute the derivatives C' and C'' on $[0, \infty)$.

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

8. (10 points) Let the pair (X, Y) be uniformly distributed on the triangle T of points (x, y) for which $x \geq 0$ and $y \geq 0$ and $x + y \leq 1$.

(a) Compute the probability density function $f(x, y)$ for the pair (X, Y).

(b) Compute $E[X|Y]$ as a function of Y.

(c) Express the expectation $E[\cos(XY)]$ as a double integral. No need to evaluate the integral.

(d) Compute the probability $P(Y > 3X)$.

9. (10 points) A classroom contains 5 first year students, 5 second year students, 5 third year students, and 5 fourth year students.

(a) The teacher random selects 5 students (with all 5-person sets equally likely) to form a Governance Committee. What is the probability that all 5 students are from the same year (i.e., they are either all first year, all second year, all third year, or all fourth year)?

(b) What is the probability the Governance Committee contains 4 people who all come from the same year, along with 1 student from a different year?

(c) What is the conditional probability that the Governance Committee contains 5 people from the same year *given* that one can find at least 4 people on the committee from the same year.

(d) What is the probability that the Governance Committee contains at least 1 student from each of the 4 years?

10. (10 points) 10 players gather for a game called **Simplified Poker**. Each player begins with 10 *chips*. When a player runs out of chips, that player is out of the game. During a round of play, if $k \geq 2$ players remain in the game, they each must wager 1 of their chips by placing it in a pile; poker hands are then randomly dealt and the player with the best hand claims the k chips in the pile. (Each of the k wagering players has a $1/k$ chance of having the best hand, independently of prior outcomes.) For $i \in \{1, 2, \ldots, 10\}$ let $A_i(n)$ be the number of chips the ith player has after n rounds. Note that $A_1(0) = A_2(0) = \cdots = A_{10}(0) = 10$. The game goes until the first time T at which 1 player (the winner) has all 100 chips. Nothing changes after time T. If $n > T$ we write $A_j(n) = A_j(T)$.

(a) Which of the following (viewed as sequence of random variables indexed by n) is a martingale? Just circle the corresponding letters.

(i) $A_7(n)$ (i.e., the number of chips Player 7 has)

(ii) $\sum_{j=1}^{5} A_j(n)$ (i.e., the total number of chips the first 5 players have)

(iii) $(-1)^{A_7(n)}$ (i.e., 1 if Player 7 has even number of chips, -1 otherwise)

(iv) $A_1(n)^2$ (i.e., the square of Player 1's chip count)

(v) $A_1(n)^2 - A_2(n)^2$ (square of Player 1's chip count minus square of Player 2's chip count)

(b) Compute the probability that Player 1 has an *epic comeback* — i.e., that the process $A_1(0), A_1(1), A_1(2), \ldots$ at some point gets all the way down to 1 before later rising up to 100.

(c) Compute the probability that *some* player has an epic comeback—i.e., that the winning player is somebody who at some point had only one chip.

(d) For $n \geq 1$, let $B(n) \in \{0, 1, 2, \ldots, 9\}$ be the number of wagers *against* Player 1 in the nth round. So if k players including Player 1 wager in the nth round then $B(n) = k - 1$. If $A_1(n) \in \{0, 100\}$ (so Player 1 is not betting) then $B(n) = 0$. Is $M(n) = A_1(n)^2 - \sum_{j=1}^{n} B(j)$ a martingale? Explain why or why not. **Hint:** imagine that $k \geq 2$ players—one of which is Player 1—remain in the game after the nth round. Observe that during the $(n+1)$th round, Player 1's chip count must increase by a random amount $X \in \{-1, k-1\}$. Can you compute $E\Big[\Big(A_1(n+1)\Big)^2 \Big| \mathcal{F}_n\Big] = E\Big[\Big(A_1(n) + X\Big)^2 \Big| \mathcal{F}_n\Big]$ in this scenario? If it helps, start by noting that $E[X^2] = \frac{1}{k} \cdot (k-1)^2 + \frac{k-1}{k} \cdot 1^2 = \frac{k^2-2k+1+k-1}{k} = \frac{k^2-k}{k} = k - 1$.