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.

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 $\mathrm{Var}[X_1]$. **ANSWER:** $E[X_1] = (69 + 70 + 71 + 72 + 73)/5 = 71$ and $\mathrm{Var}(X_1) = \frac{1}{5}\Big((-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2\Big) = 2.$

(b) Compute $E[A]$ and $\mathrm{Var}[A]$ and the standard deviation of A. **ANSWER:** $E[A] = 71$ and $\mathrm{Var}(A) = \frac{50\mathrm{Var}(X_1)}{50^2} = \frac{1}{25}$ and $\mathrm{SD}(A) = \sqrt{\mathrm{Var}(A)} = 1/5$.

(c) Compute the moment generating functions $M_{X_1}(t)$ and $M_A(t)$. **ANSWER:** $M_{X_1}(t) = E[e^{tX_1}] = \frac{1}{5}\Big(e^{69t} + e^{70t} + e^{71t} + e^{72t} + e^{73t}\Big)$. Write $S = (X_1 + X_2 + \ldots + X_{50})$. Then $M_S(t) = \Big(M_{X_1}(t)\Big)^{50}$ and $M_A(t) = M_{S/50}(t) = M_S(t/50) = \Big(M_{X_1}(t/50)\Big)^{50}$. Substituting in, we get $M_A(t) = \Big[\frac{1}{5}\Big(e^{69t/50} + e^{70t/50} + e^{71t/50} + e^{72t/50} + e^{73t/50}\Big)\Big]^{50}$.

(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 $P(A \le 70.5)$. You may use the function $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ in your answer. **ANSWER:** Recall $\mathrm{SD}(A) = 1/5 = .2$ and $E[A] = 71$. Thus 70.5 is 2.5 standard deviations below the mean so the answer is about $\Phi(-2.5)$.

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 $\rho(X, \tilde{X})$. **ANSWER:** Additivity of variance for i.i.d. random variables gives $\mathrm{Var}(X) = 100$ and $\mathrm{SD}(X) = 10$. Bilinearity of covariance gives $\mathrm{Cov}(X, \tilde{X}) = 28$ and $\rho(X, \tilde{X}) = \frac{28}{10 \cdot 10} = .28$.

(b) Bob wants to guess the second score given the first. Express $G = E[\tilde{X}|X]$ as a function of X. **ANSWER:** We know that $E[X_1 + X_2 + \ldots + \ldots X_{100}|X] = E[X|X] = X$. By symmetry we must have $E[X_i|X] = X/100$ for each i and additivity of expectation gives $E[\tilde{X}|X] = 28 \cdot X/100 = .28X$.

(c) Carol wants to know the **standard deviation** of Bob's guess. Compute $\mathrm{Var}(G)$ and $\mathrm{SD}(G)$. **ANSWER:** $\mathrm{Var}(G) = \mathrm{Var}(.28X) = (.28)^2\mathrm{Var}(X) = .0784 \cdot 100 = 7.84$ and $\mathrm{SD}(G) = 2.8$.

(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 $\text{Var}(G)$ and $\text{Var}(X)$. **ANSWER:** $\text{Var}(X) = \text{Cov}(G+R, G+R) = \text{Cov}(G,G) + 2\text{Cov}(G,R) + \text{Cov}(R,R) = \text{Var}(G) + \text{Var}(R)$ is found by using the hintt. Substituting values computed before gives $100 = 7.84 + \text{Var}(R)$ so $\text{Var}(R) = 100 - 7.84 = 92.16$ and $\text{SD}(R) = \sqrt{\text{Var}(R)} = 9.6$. **REMARK:** Imagine scores represent surplus fatalities. If Eve and Frank are new surgeons, and Frank had 4 more fatalities than Eve in Year 1, then you *expect* Frank to have $4 \cdot .28 = 1.08$ more fatalities during Year 2 alone. If you must choose a surgeon to rehire (redirecting other to a less intensive speciality) then choosing Eve saves an entire expected life. So Bob's guess is important! On the other hand, what David overheard was correct: 9.6 is almost 10. Bob's guess is *both* important *and* imprecise. In the real world, it is hard to find a simple and systematic way to score surgeons.

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 $A = (X_1 + X_2 + \ldots + X_{100})/100$. **ANSWER:** By special property of Cauchy random variables, the average has the law of X_1, itself, so $f_A(x) = f_{X_1}(x) = \frac{1}{\pi(1+x^2)}$.

(b) Compute the joint probability density function $f(x,y)$ of the pair (X_1, X_2). **ANSWERS:** By independence, we just multiply the corresponding densities, so $f(x,y) = \frac{1}{\pi(1+x^2)} \cdot \frac{1}{\pi(1+y^2)}$.

(c) Compute the probability density function of $Z = X_1 + 2X_2 + 3X_3 + 4X_4 + 10$. **ANSWER:** By the special Cauchy property, kX_1 has the same law as the sum of k independent copies of X_1. So Z has the law of $S + 10$ where $S = \sum_{i=1}^{10} X_i$, which in turn has the same law as $10X_1$. Thus $f_S(x) = \frac{1}{10} f_{X_1}(x/10)$ and $f_A(x) = \frac{1}{10} f_{X_1}\big((x-10)/10\big) = \frac{1}{10\pi\left(1+\left(\frac{x-10}{10}\right)^2\right)}$

(d) Compute the probability $P(X_1 > 2X_2 + 3)$. **ANSWER:** By special Cauchy property and the fact that by symmetry $2X_2$ has same law as $-2X_2$, we know that $X_1 - 2X_2$ has the same law as $3X_1$. So answer is $P(3X_1 > 3) = P(X_1 > 1) = 1/4$. (Recall spinning flashlight story.)

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 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).
2. 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).

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

$$M = \begin{pmatrix} 1/4 & 1/12 & 1/12 & 1/12 & 1/8 & 1/8 & 1/8 & 1/8 \\ 1/12 & 1/4 & 1/12 & 1/12 & 1/8 & 1/8 & 1/8 & 1/8 \\ 1/12 & 1/12 & 1/4 & 1/12 & 1/8 & 1/8 & 1/8 & 1/8 \\ 1/12 & 1/12 & 1/12 & 1/4 & 1/8 & 1/8 & 1/8 & 1/8 \\ 1/16 & 1/16 & 1/16 & 1/16 & 1/2 & 1/12 & 1/12 & 1/12 \\ 1/16 & 1/16 & 1/16 & 1/16 & 1/12 & 1/2 & 1/12 & 1/12 \\ 1/16 & 1/16 & 1/16 & 1/16 & 1/12 & 1/12 & 1/2 & 1/12 \\ 1/16 & 1/16 & 1/16 & 1/16 & 1/12 & 1/12 & 1/12 & 1/2 \end{pmatrix}$$

(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. **ANSWER:** Entry $(5,5)$ of M^2, which is $4 \cdot \frac{1}{8} \cdot \frac{1}{16} + \frac{1}{2} \cdot \frac{1}{2} + 3 \cdot \frac{1}{12} \cdot \frac{1}{12} = \frac{1}{32} + \frac{1}{4} + \frac{1}{48}$.

(c) In the long run, what fraction of the time does Alfred spend on each of the 8 topics? **ANSWER:** Each time Alfred is in *worry* state he stays in *some worry state* with probability $1/2$ and switches to dream with probability $1/2$. Each time he is in dream state, he stays in *some dream state* with probability $3/4$ and switches to worry state with probability $1/4$. Solving this as 2-state problem, we find Alfred spends $1/3$ his time worrying and $2/3$ his time dreaming. There is no distinction between 4 worry states, so each has same probability, ditto for dream states, so stationary vector is $(1/6, 1/6, 1/6, 1/6, 1/12, 1/12, 1/12, 1/12)$.

(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? **ANSWER:** He has a $1/4$ chance to stop dreaming at each step, so the answer is expectation of geometric random variable with parameter $1/4$, which is 4.

5. (10 points) Alice has for some reason started thinking about the Roman Empire on average once an hour. The times $X_1, X_2, \dots$ 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 $E[X_2^3] = E\big[(W_1 + W_2)^3\big]$. **ANSWER:** X_2 is Gamma with $\lambda = 1$, $n = 2$. Has density $f(x) = e^{-x}x$ on $[0, \infty]$ so $E[X_2^3] = \int_0^\infty f(x)x^3 dx = \int_0^\infty e^{-x}x^4 dx = 4! = 24$. **Also:** $E[X_2^3] = E\big[(W_1 + W_2)^3\big] = E[W_1^3 + 3W_1^2W_2 + 3W_1W_2^2 + W_2^3] = 6 + 6 + 6 + 6 = 24$.

(b) Compute the probability density function for X_7. **ANSWER:** This is Gamma with $n = 7$ and $\lambda = 1$, so answer is $e^{-x}x^6/6!$ on $[0, \infty)$.

(c) Compute the probability Alice has exactly 5 Roman Empire thoughts in the first 7 hours. **ANSWER:** This is $e^{-\lambda}\lambda^k/k!$ with $\lambda = 7$ and $k = 5$. Comes to $\frac{7^5}{5! e^7}$

(d) Compute the probability density function of $M = \min\{W_1, W_2, W_3\}$. **ANSWER:** M is exponential with parameter 3, so its density is $3e^{-3x}$ on $[0, \infty)$.

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 $X = (X_1, X_2, \dots, X_8)$ be the sequence of words.

(a) Compute the entropy $H(X_1)$ and $H(X)$. **ANSWER:** $H(X_1) = \sum(-p_i \log p_i)$ using log base 2. This comes to $\frac{5}{8} \cdot 3 + \frac{2}{16} \cdot 4 + \frac{8}{32} \cdot 5 = \frac{15}{8} + \frac{4}{8} + \frac{5}{8} = \frac{10}{8} = \frac{29}{8}$. Since the X_i are independent, we have $H(X) = 8 \cdot H(X_1) = 29$.

(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? **ANSWER:** We determine X_1 by asking questions where each questions cuts probability space in half. For example, first question can be "Is it one of Slay, Delulu, Cap or Rizz?" If answer is yes, next question can be "Is it either Slay or Delulu?" Continuing this way, we discover X_1 with $29/8$ question in expectation. Repeating for each i, we expect to need 29 questions to find X.

(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. **ANSWER:** K is binomial with $p = 1/8$ and $n = 8$. Hence

$$H(K) = -\sum p_i \log(p_i) = -\sum_{k=0}^{8} \binom{8}{k} (1/8)^k (7/8)^{8-k} \log\Big(\binom{8}{k} (1/8)^k (7/8)^{8-k}\Big).$$

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 $K \geq 0$. **ANSWER:** We have $f_X(x) = \frac{1}{2}$ on $[0, 2]$. So $E[\max\{X-K, 0\}] = \int_0^2 \frac{1}{2} \max\{x-K, 0\}dx = \frac{1}{2}\int_K^2 (x-K)dx = \frac{1}{2}\Big(\frac{x^2}{2} - Kx\Big)\Big|_K^2 = \frac{4-K^2}{4} - \frac{2K-K^2}{2}$ on $[0, 2]$, which comes to $\frac{K^2}{4} - K + 1$. Then we have $C(K) = 0$ for $K > 2$.

(b) Compute the derivatives C' and C'' on $[0, \infty)$. **ANSWER:** On $[0, 2]$ we can compute $C'(K) = \frac{K}{2} - 1$ and $C''(K) = 1/2$. For $K > 2$, we have $C'(K)$ and $C''(K)$ both equal to zero. Alternatively, recall general call function facts: $C'(K) = F_X(K) - 1$ and $C''(K) = f_X(K)$.

(c) Compute the expectation $E[X^3 + 3X]$. **ANS:** $\int_0^2 \frac{1}{2}(x^3 + 3x)dx = \frac{1}{2}(\frac{x^4}{4} + \frac{3x^2}{2})\Big|_0^2 = \frac{1}{2}(4 + 6) = 5$.

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). **ANSWER:** T has area $1/2$ so $f(x, y)$ will be 2 if $(x, y) \in T$ and 0 if $(x, y) \notin T$.

(b) Compute $E[X|Y]$ as a function of Y. **ANSWER:** If a value of $Y = y$ in $[0, 1]$ is given, the conditional law of X is uniform on $[0, 1-y]$, and hence the conditional expectation is $(1-Y)/2$.

(c) Express the expectation $E[\cos(XY)]$ as a double integral. No need to evaluate the integral. **ANSWER:** $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)\cos(xy)dxdy = \int_0^1 \int_0^{1-y} 2\cos(XY)dxdy$

(d) Compute the probability $P(Y > 3X)$. **ANSWER:** The line $Y = 3X$ divides T into two pieces. You can check that the piece where $Y > 3X$ is larger is a triangle with area $1/8$, so $P(Y > 3X) = 1/4$

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)? **ANSWER:** There are $\binom{20}{5}$ ways to choose an undordered group of five students, four of which correspond to all-same-year groupings, so answer is $4/\binom{20}{5}$.

(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? **ANSWER:** 4 ways to choose the year, then 5 ways to choose which student to swap out and 15 ways to choose which student to swap in. So answer is $300/\binom{20}{5}$,

(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. **ANSWER:** Let A, B be events from (a), (b). $P(A|A \cup B) = P(A)/P(A \cup B) = 4/304 = 1/76$.

(d) What is the probability that the Governance Committee contains at least 1 student from each of the 4 years? **ANSWER:** 4 ways to choose year to have 2 people. Then have to choose people from each year. Answer is $4 \cdot \binom{5}{2} 5^3 / \binom{20}{5} = 5000/\binom{20}{5}$.

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) **YES!** Can check that each time Player 7 bets, Player 7's expected profit (given all that has happened in the past) is zero.

(ii) $\sum_{j=1}^{5} A_j(n)$ (i.e., the total number of chips the first 5 players have) **YES!** This is sum of five martingales, hence a martingale itself.

(iii) $(-1)^{A_7(n)}$ (i.e., 1 if Player 7 has even number of chips, -1 otherwise) **NO!** This starts at 1, and later can take only values -1 or 1 (hence expectation at future times is less than 1 and not equal initial value).

(iv) $A_1(n)^2$ (i.e., the square of Player 1's chip count) **NO!** If Player 1 is betting at stage n, then $A_1(n+1)$ is equal to $A_n(1)$ plus a random element $X \in \{-1, 1\}$. But then $A_1(n+1)^2 = A_1(n)^2 + 2XA_1(n) + X^2$, and the conditional expectation of this will be larger than $A_1(n)^2$.

(v) $A_1(n)^2 - A_2(n)^2$ (square of Player 1's chip count minus square of Player 2's chip count) **NO!** After either $A_1(n)$ or $A_2(n)$ reaches zero, the same logic as in (iv) shows that it cannot be a martingale at that point.

(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. **ANSWER:** There is a $90/99 = 10/11$ chance to get down to 1 before getting up to 100. Given that, there is a $1/100$ chance of getting up to 100, so overall probability is $(10/11) \cdot 1/100 = 1/110$.

(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. **ANSWER:** The 10 events are disjoint, so the probability of at least one is 10 times the probability of one. Answer is $1/11$.

(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 $E[X^2] = \frac{1}{k} \cdot (k-1)^2 + \frac{k-1}{k} \cdot 1^2 = \frac{k^2 - 2k + 1 + k - 1}{k} = \frac{k^2 - k}{k} = k - 1$. **ANSWER:** In above scenario $E\Big[\Big(A_1(1) + X\Big)^2 \Big| \mathcal{F}_n\Big] = E\Big[\Big(A_1(n)^2 + 2A_1(n)X + X^2\Big)^2 \Big| \mathcal{F}_n\Big] = A_1(n)^2 + 0 + (k-1)$. But in same scenario $\sum_{j=1}^{n} B(j)$ also increases by $k - 1$. Since this is the only scenario under which either $A_1(n)^2$ and $\sum_{j=1}^{n} B(j)$ can change, we conclude that $M(n)$ is a martingale.