18.600 Midterm 2, Fall 2023: 50 minutes, 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.**

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

1. (10 points) A certain town contains 4800 adults, who all *totally intend* to vote in an upcoming election. However, the residents of this community often become distracted and forget to vote. The probability that each individual actually votes is .25 (independently of what the others do). Let X be the total number of people who vote.

(a) Compute $E[X]$ and $\text{Var}[X]$ and $\text{SD}[X]$.

(b) Use the de Moivre-Laplace limit theorem to approximate $P(1140 < X < 1260)$ (i.e., the probability that the *percentage* of individuals voting is between 23.75 and 26.25). Give an explicit numerical value. (To find this value, you may use the approximations $\Phi(-1) \approx .16$ and $\Phi(-2) \approx .023$ and $\Phi(-3) \approx .0013$ where $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.)

2. (20 points) In other election news: Alice, Bob, Carol and Dave are running for president. But they have little appeal to the voters, and it is known that with probability 1 they will all drop out of the race eventually. Let A, B, C and D denote the number of years from now until Alice, Bob, Carol and Dave (respectively) drop out. Assume that A, B, C and D are independent exponential random variables with parameter 1.

(a) Compute the expected amount of time until all 4 drop out. That is, find $E[\max\{A, B, C, D\}]$.

(b) Write $X = \frac{A+B+C+D}{4}$. (In other words, X is the *average* of the 4 random campaign durations.) Compute the probability density function for X.

(c) Compute the conditional probability that Alice drops out before Bob *given* that Bob drops out before Carol.

(d) Let $Y = \min\{A, B, C, D\}$ be the time at which the first candidate drops out. Compute $E[Y^3]$.

3. (15 points) Bob has recently moved to a new high school, and he plans to politely invite a fellow student to a school dance after introducing himself in person. Each person Bob invites will say yes with probability p, independently of what anybody else does, but Bob does not know *a priori* what p is. Based on his personal Bayesian prior, Bob thinks p is a uniform random variable on $[0, 1]$. **NOTE:** If it helps, you may use the fact that a Beta (a, b) random variable has expectation $a/(a + b)$ and density $x^{a-1}(1 - x)^{b-1}/B(a, b)$, where $B(a, b) = (a - 1)!(b - 1)!/(a + b - 1)!$.

(a) *Given* that Alex, Barbara, Carli, Deborah and Emily (the first five people Bob invites) decline his invitation, what is his revised *conditional* probability density function for p?

(b) *Given* that the first five invitations were declined, what is the conditional probability that Fiona (the sixth person Bob invites) accepts his invitation?

(c) Fiona accepts Bob's invitation but later that day (while saving a sloth from a bus) she is severely injured. So she asks Bob to invite someone else. *Given* that the first five invitees declined *and* the sixth accepted, what is the conditional probability that Gemma (the seventh person Bob invites) accepts his invitation?

4. (20) Alice is participating in an athletic competition where she will receive an overall score X. There are 16 factors that contribute to her score (sleep quality, shoe design, training intensity, protein consumption, luck during the first stage, luck during the second stage, etc.) Assume that X can be written as $X_1 + X_2 + \ldots + X_{16}$ where each X_i represents the net contribution of the ith factor to her score. Assume further that the X_i are i.i.d. random variables, each with variance 1 and mean zero.

(a) Alice wants a sense of "how related" the overall score is to a single individual score. To that end, compute the correlation coefficient $\rho(X_1, X)$.

(b) The next week, Alice competes in a similar competition. In this case, the first 9 factors stay the same but the other 7 (those involving competition-day luck) are resampled independently. In other words, her score for the second competition is

$$\tilde{X} = X_1 + X_2 + \ldots + X_9 + \tilde{X}_{10} + \tilde{X}_{11} + \ldots + \tilde{X}_{16}$$

where the random variables X_i and $\tilde{X}_i$ are all i.i.d. with variance 1 and mean zero. Compute the correlation coefficient $\rho(X, \tilde{X})$.

(c) Compute the conditional expectation $E[\tilde{X}|X]$ as a function of X.

5. (10 points) Each customer who walks into Terry's TV Emporium purchases a TV with probability 1/100, independently of what the others do. During the course of the day, 200 customers come by. Let Z be the total number of TVs purchased during that time.

(a) Compute the moment generating function $M_Z(t)$. Give an exact formula, not a Poisson approximation.

(b) Use a Poisson random variable to approximate $P(Z \geq 2)$.

6. (15 points) Let D denote the unit circle $\{(x,y) : x^2 + y^2 \leq 1\}$. Suppose that the pair (X, Y) has joint density function given by

$$f_{X,Y}(x,y) = \begin{cases} g\left(\sqrt{x^2+y^2}\right) & (x,y) \in D \\ 0 & (x,y) \notin D \end{cases},$$

for some fixed continuous function $g : [0,1] \to (0,\infty)$.

(a) Write $R = \sqrt{X^2 + Y^2}$. Compute the cumulative distribution function F_R in terms of g. Then compute the probability density function f_R.

(b) Compute the conditional density function $f_{X|Y=0}(x)$ in terms of g.

(c) Write $Z = X + Y$. Compute the conditional expectation $E[Z|Y]$ as a function of Y. Does the answer depend on g?

7. (10 points) Suppose X_1 through X_{10} are i.i.d. random variables, each with probability density function $f(x) = \frac{1}{\pi(1+x^2)}$.

(a) Compute the probability $P(X_1 + X_2 > X_3 + X_4 - 4)$.

(b) Compute $P(X_1^2 + X_2^2 \leq 1)$ as a double integral. (You don't have to evaluate the integral explicitly.)