18.600 Midterm 2, Fall 2023 Solutions

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 $\mathrm{Var}[X]$ and $\mathrm{SD}[X]$. **ANSWER:** $E[x] = np = 4800 \cdot \frac{1}{4} = 1200$. $\mathrm{Var}[X] = npq = 4800 \cdot \frac{1}{4} \cdot \frac{3}{4} = 900$. $\mathrm{SD}[X] = \sqrt{90} = 30$.

(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 $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.) **ANSWER:** This is the probability that X is between $E[X] - 2\mathrm{SD}[X]$ and $E[X] + 2\mathrm{SD}[X]$ which is roughly $\Phi(2) - \Phi(-2) = 1 - 2\Phi(-2) \approx 1 - .046 = .954$.

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 $E[\max\{A, B, C, D\}]$. **ANSWER:** This is the"radioactive decay" problem. The time until the first drops out is exponential with parameter 4 (hence has expectation 1/4) and the subsequent time until the second drops out is exponential with parameter 2 (hence has expectation 1/3) and so forth. Overall answer is $1/4 + 1/3 + 1/2 + 1 = \frac{25}{12}$.

(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. **ANSWER:** Set $S = A + B + C + D$. Then S is a gamma random variable, and $f_S(x) = x^3 e^{-x}/3!$ on $[0, \infty)$. Then $f_X(x) = f_{S/4}(x) = 4 f_S(4x) = 4(4x)^3 e^{-4x}/3! = \frac{256}{6} x^3 e^{-4x}$.

(c) Compute the conditional probability that Alice drops out before Bob *given* that Bob drops out before Carol.
ANSWER: $P(\text{Alice before Bob } \textit{and} \text{ Bob before Carol})/P(\text{Alice before Bob}) = \frac{1}{6}/\frac{1}{2} = 1/3$.

(d) Let $Y = \min\{A, B, C, D\}$ be the time at which the first candidate drops out. Compute $E[Y^3]$. **ANSWER:** Y is exponential with parameter 4. It has same law as $\frac{1}{4}Z$ where Z is exponential with parameter 1. So $E[Y^3] = E[(\frac{Z}{4})^3] = \frac{1}{64} E[Z^3] = 3!/64 = 3/32$.

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? **ANSWER:** This is a Beta random variable with $a - 1 = 0$ and $b - 1 = 5$. Plugging gives a density of $6(1 - x)^5$ on $[0, 1]$.

(b) *Given* that the first five invitations were declined, what is the conditional probability that Fiona (the sixth person Bob invites) accepts his invitation? **ANSWER:** As on the problem set, the probability is the expectation of the corresponding Beta random variable, which is $a/(a+b) = 1/7$.

(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? **ANSWER:** Same as (b) but with $(a-1) = 1$ and $(b-1) = 5$. Answer is $a/(a+b) = 2/8 = 1/4$.

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 $\rho(X_1, X)$. **ANSWER:** We can use bilinearity of covariance to find $\rho(X_1, X) = \mathrm{Cov}(X_1, X)/\sqrt{\mathrm{Var}(X_1)}\sqrt{\mathrm{Var}(X)} = 1/\sqrt{16} = 1/4$.

(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 $\rho(X, \tilde{X})$. **ANSWER:** Again, bilinearity of covariance lets us compute $\rho(X, \tilde{X}) = \mathrm{Cov}(X, \tilde{X})/\sqrt{\mathrm{Var}(X)\mathrm{Var}(\tilde{X})} = 9/16$.

(c) Compute the conditional expectation $E[\tilde{X}|X]$ as a function of X. **ANSWER:** $E[X_1 + X_2 + \ldots X_{16}|X] = E[X|X] = X$. By symmetry and additivity of conditional expectation we have $E[X_i|X] = X/16$ for each i. Applying addivitiy of conditional expectation again gives $E[\tilde{X}|X] = \frac{9}{16}X$. This is an example of regression to the mean.

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. **ANSWER:** If Z_i is the number of TVs purchased by the ith customer, then $M_{Z_1}(t) = E[e^{tZ_1}] = P(Z_1 = 0)e^{t\cdot 0} + P(Z_1 = 1)e^{t\cdot 1} = (.99 + .01e^t)$. When we add independent random variables, the moment generating functions are multiplied. So $M_Z(t) = \prod_{i=1}^{200} M_{Z_i}(t) = (.99 + .01e^t)^{200}$. Can also be written $(1 + .01(e^t - 1))^{200} \approx e^{2(e^t-1)}$.

(b) Use a Poisson random variable to approximate $P(Z \geq 2)$. **ANSWER:** Z is roughly Poisson with $\lambda = 200 \cdot \frac{1}{100} = 2$. So $P(Z \geq 2) = 1 - P(Z = 1) - P(Z = 0) \approx 1 - e^{-2}2^1/1! - e^{-2}2^0/0! = 1 - 3e^{-2}$.

6. (15 points) Let D denote the unit circle $\{(x,y) : x^2 + y^2 \le 1\}$. 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. **ANSWER:** Integrating in polar coordinates, $F_R(a) = P(R \le a) = \int_0^a \int_0^{2\pi} rg(r)d\theta dr = \int_0^a 2\pi rg(r)dr$ for $a \in [0,1]$. Differentiating gives $f_R(r) = 2\pi rg(r)$.

(b) Compute the conditional density function $f_{X|Y=0}(x)$ in terms of g. **ANSWER:** This is $\frac{f(x,0)}{C}$ where $C = \int_{-\infty}^{\infty} f(x,0)dx = \int_{-1}^{1} g(|x|)dx$. Plugging in g we get $\frac{g(|x|)}{C}$.

(c) Write $Z = X + Y$. Compute the conditional expectation $E[Z|Y]$ as a function of Y. Does the answer depend on g? **ANSWER:** $E[X+Y|Y] = E[X|Y] + E[Y|Y] = 0 + Y$. The first term is zero because, given any choice for Y, the conditional pdf for X is symmetric about 0, and hence the conditional expectation is zero.

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 $P(X_1 + X_2 > X_3 + X_4 - 4)$. **ANSWER:** $P(X_1 + X_2 > X_3 + X_4 - 4) = P\left(\frac{X_1+X_2-X_3-X_4}{4}\right) > -1$. Since $X_1, X_2, -X_3, -X_4$ are all Cauchy random variables, their average is Cauchy. And the probability that a Cauchy random variable is greater than -1 is $3/4$. (Recall spinning flashlight story.)

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

$$\int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi^2(1+x^2)(1+y^2)} dydx$$