18.600 Midterm 2, Spring 2021: Solutions

1. (10 points) Suppose 300 students want to book vaccine appointments during a high demand period. Each student succeeds in booking an appointment with probability $\frac{3}{4}$ independently of all others. Let X be the number who book appointments.

(a) Compute $E[X]$ and $\mathrm{Var}[X]$. **ANSWER:** X is binomial with $n = 300$, $p = 3/4$ and $q = 1/4$. We find $E[X] = np = 225$ and $\mathrm{Var}[X] = npq = 300 \cdot 3/16 = 900/16$.

(b) Use the de Moivre-Laplace limit theorem to approximate $P(217.5 < X < 247.5)$ (i.e., the probability that the *percentage* of students obtaining appointments is between 72.5 and 82.5). Give a numerical 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:** We calculate the standard deviation $\mathrm{SD}(X) = \sqrt{\mathrm{Var}[X]} = 30/4 = 7.5$. Since 217.5 is one SD below the mean and 247.5 is three SDs above the mean, the desired probability is $\Phi(3) - \Phi(-1) \approx$ $(1 - .0013) - .16 = .9987 - .16 = .8387$.

2. (20 points) At time zero, Alice, Bob, and Carol are (respectively) first, second and third in line at a taxi stand. At this stand, taxis arrive at times (measured in minutes) corresponding to a Poisson point process with parameter $\lambda = .5$. Alice will take the first taxi that arrives, Bob the second and Carol the third. Let A, B and C be (respectively) the times at which Alice's, Bob's and Carol's taxis arrive. (These are the first, second and third points in the Poisson point process.)

(a) Compute $E[C]$ and give a probability density function for C. **ANSWER:** This is Gamma with $n = 3$, $\lambda = .5$. The density is $.5e^{-.5x}(.5x)^2/2!$ on $[0, \infty)$ and $E[C] = n/\lambda = 6$.

(b) Compute $E[B^3]$. **Hint:** Try writing $X := B - A$ and work with $B = A + X$. **ANSWER:** Since A and X are i.i.d. exponential random variables with $\lambda = .5$ we have $E[B^3] = E[A^3 + 3A^2X + 3AX^2 + X^3] = E[A^3] + 3E[A^2X] + 3E[AX^2] + E[X^3]$. By scaling, the expectation is $\lambda^{-3} = 8$ times what it would be if we had $\lambda = 1$. Using $\int_0^\infty e^{-x}x^n dx = n!$ and independence of Z and X and we get $8(3! + 3 \cdot 2! \cdot 1! + 3 \cdot 1! \cdot 2! + 3!) = 8 \cdot 24 = 192$.

(c) Compute the probability that Bob ends up waiting at least twice as long (overall) as Alice does — i.e., compute $P(B \geq 2A)$. **ANSWER:** This is equivalent to $P(X > A)$ which is 1/2 since X and A are i.i.d.

(d) Compute the conditional expectations $E[B|C]$ and $E[C|B]$. **ANSWER:** Write $X_1 = A,\ \ X_2 = B - A,\ \ X_3 = C - B$ for the three i.i.d. waiting interval lengths. We know $E[X_1 + X_2 + X_3|C] = E[C|C] = C$ and by symmetry $E[X_i|C] = C/3$ for each i. Hence $E[B|C] = E[X_1 + X_2|C] = \frac{2}{3}C$. Next, $E[C|B] = E[B + X_3|B] = B + 2$ (since B and X_3 are independent and $E[X_3] = \lambda^{-1} = 2$ *a priori*).

3. (10 points) Alice sees a carnival game that requires her to knock over a bottle by throwing a ball. Alice knows that each time she attempts to knock over the bottle, she will succeed with probability p (independently of everything else). But *a priori* Alice does not know what p is. (That is, she does not know how difficult the game will be for her.) She believes *a priori* that p is a uniform random variable on $[0,1]$. She buys a ticket that allows her to make three attempts.

(a) *Given* that Alice fails on her first two attempts to knock over the bottle, what is her conditional probability density function for p? **ANSWER:** This is β with parameters $a-1=0$ and $b-1=2$. Answer comes to $(1-x)^2/B(1,3) = 3(1-x)^2$.

(b) *Given* that Alice fails on her first two attempts to knock over the bottle, what is the conditional probability that she also fails on her third attempt? **ANSWER:** We know from the problem sets that her conditional probability of success is the conditional expectation of p, which is $a/(a+b) = 1/4$. Hence her conditional probability of failure is $3/4$.

4. (15 points) George wants to inoculate his army against smallpox by deliberately injecting smallpox pus into the arms of 40,000 soldiers. Of the 40,000 people deliberately infected in this way, each one dies independently with probability .02. Let X be the total number of deaths.

(a) Compute $E[X]$ and $\mathrm{SD}(X)$. **ANSWER:** This is binomial with $n = 40,000$ and $p = .02$. So $E[X] = np = 800$ and $\mathrm{SD}(X) = \sqrt{npq} = \sqrt{40,000 \cdot .0196} = 200 \cdot .14 = 28$.

(b) Use de Moivre-Laplace to estimate $P(772 \leq X \leq 828)$. (This corresponds to the death rate being between 1.93 and 2.07 percent.) You may use the approximations $\Phi(-1) \approx .16$ and $\Phi(-2) \approx .023$ and $\Phi(-3) \approx .0013$ where $\Phi(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$. **ANS:** $\Phi(1) - \Phi(-1) \approx .68$.

(c) Suppose that 100 of the 40,000 people are high-level officers. Use a Poisson approximation to estimate the probability that exactly 3 high-level officers die from the smallpox injection. **ANSWER:** Roughly $e^{-\lambda}\lambda^k/k!$ with $k = 3$ and $\lambda = .02 \times 100 = 2$. Comes to $e^{-2}2^3/3! = \frac{4}{3e^2}$. The numbers $p = .02$ and $n = 40000$ appear in the George Washington story recounted here. It is just a coincidence that these numbers make $\sqrt{npq}$ an integer.

5. (15 points) Carol is taking four exams for a class. Her scores on those exams, denoted X_1, X_2, X_3, X_4 are i.i.d. standard normal random variables. (The scores on these exams can be both positive and negative.) So each X_i has probability density function given by $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Write $S = X_1 + X_2 + X_3 + X_4$ for Carol's total exam score.

(a) Compute the probability density function f_S. **ANSWER:** S is normal with mean zero and variance 4. So $f_S(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} = \frac{1}{2\sqrt{2\pi}} e^{-x^2/8}$.

(b) Let $Z = X_1^4$. Compute the moment generating function $M_Z(t)$ as an integral. (You don't have to evaluate the integral explicitly.) **ANSWER:** $M_Z(t) = E[e^{tZ}] = E[e^{tX_1^4}] = \int_{-\infty}^{\infty} f(x) e^{tx^4} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{tx^4} dx$

(c) Carol wants to know how correlated her first exam score is with her overall score. Compute the correlation coefficient $\rho(X_1, S)$. **A:** $\rho(X_1,S)=\frac{\mathrm{Cov}(X_1,S)}{\sqrt{\mathrm{Var}[X_1]\mathrm{Var}[S]}}=\frac{\mathrm{Cov}(X_1,X_1+X_2+X_3+X_4)}{\sqrt{1\cdot 4}}=\frac{1}{2}$.

6. (15 points) Suppose that the pair (X, Y) is uniformly distributed on the diamond $D = \{(x, y) : |x| + |y| \le 1\}$. That is, the joint density function is given by

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2} & (x,y) \in D \\ 0 & (x,y) \notin D \end{cases}.$$

(a) Compute the marginal density function f_X. **ANSWER:** $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$ which comes to 0 if $|x| > 1$ and $1 - |x|$ if $|x| \le 1$.

(b) Compute the probability $P(3X < 7Y)$. **ANSWER:** By 180-degree rotational symmetry, the part of D above the line $3x = 7y$ has the same area as the part below that line. Answer is 1/2.

(c) Compute the conditional density function $f_{X|Y=.5}(x)$. **ANSWER:** The conditional density is uniform on $[-.5, .5]$. Hence $f_{X|Y=.5}(x)$ is 1 on $[-.5, .5]$ and 0 elsewhere.

7. (15 points) Suppose that X and Y are independent Cauchy random variables, so both have density function $f(x) = \frac{1}{\pi(1+x^2)}$. Write $Z = X^7$.

(a) Compute the cumulative distribution function $F_Z(a)$ in terms of a. **ANSWER:** $P(Z \le a) = P(X^7 \le a) = P(X \le a^{1/7}) = F_X(a^{1/7}) = \frac{1}{2} + \frac{1}{\pi}\arctan(a^{1/7})$.

(b) Compute $P(XY \le 1)$ as a double integral. (You don't have to evaluate the integral explicitly.) **ANSWER:** Write $f(x, y) = \frac{1}{\pi(1+x^2)}\frac{1}{\pi(1+y^2)}$. Then
$P(XY \le 1) = \int_{-\infty}^{0}\int_{1/y}^{\infty} f(x, y)dxdy + \int_0^{\infty}\int_{-\infty}^{1/y} f(x, y)dxdy.$

(c) Compute the probability $P(X + Y > 2)$. Give an explicit number. **ANSWER:** This is $P(\frac{X+Y}{2} > 1) = P(X > 1)$ (since the mean of i.i.d. Cauchy random variables is itself Cauchy). Using the spinning flashlight interpretation, a quarter of the angle range corresponds to $X > 1$. Hence $P(X + Y > 2) = 1/4$.