Stats 116 Problem Set 6

Due: Wednesday, May 11 5:00 p.m. on Gradescope

Please show your work for each exercise. If you collaborate with someone else—this is fine—be sure to note that in your homework submission. You must each write up separate answer sets. Questions are either from Ross's *A First Course in Probability* or our are home-cooked special sauce.

Question 6.1 (Ross problem 6.34)**:** Let $X \sim \mathsf{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathsf{N}(\mu_2, \sigma_2^2)$ be independent normal random variables (with means μ_i and variances σ_i^2, respectively). Let $a \in \mathbb{R}$. Find the x such that

$$\mathbb{P}(X - Y > x) = P(X + Y > a).$$

Question 6.2 (Ross, problem 6.45)**:** The joint density of X and Y is

$$f(x, y) = cxy^3, \quad 0 \le x \le 1,\ 0 \le y \le x^{1/4}.$$

Find c (that is, the c such that $\int f(x,y)dxdy = 1$) and the conditional distribution of X given $Y = y$.

Question 6.3 (Ross problem 6.52)**:** Let $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \mathsf{Geom}(p)$, that is, X_i are geometric random variables with parameter $p \in (0, 1)$. Let $a \in \mathbb{N}$ be a positive integer. Find

(a) $P(\min\{X_1, \ldots, X_n\} \le a)$.

(b) $P(\max\{X_1, \ldots, X_n\} \le a)$.

Question 6.4 (Ross theoretical exercise 6.21)**:** Let W be the amount of moisture in the air on a given day and assume that $W \sim \mathsf{Gamma}(t, \beta)$, so that its density is $f(w) = \beta e^{-\beta w}(\beta w)^{t-1}/\Gamma(t)$ for $w \ge 0$. Assume that conditional on $W = w$, the number N of accidents during that day follows $N \sim \mathsf{Poisson}(w)$. Show that

$$W \mid N = n \sim \mathsf{Gamma}(t + n, \beta + 1),$$

that is, conditional on $N = n$, W has gamma distribution with parameters $(t + n, \beta + 1)$.

Question 6.5 (Projection onto a linear subspace)**:** Let $x \in \mathbb{R}^n$ be any vector, and let $S \subset \mathbb{R}^n$ be a subspace of $\mathbb{R}^n$, that is, S is a set such that $0 \in S$ and if $v, w \in S$, then the line $\{sv + tw \mid s \in \mathbb{R}, t \in \mathbb{R}\}$ is in S. Alternatively, we can always describe S as

$$S = \{v \in \mathbb{R}^n \mid Av = 0\}$$

for some matrix $A \in \mathbb{R}^{m \times n}$, where $m \le n$. (We assume that A has rank m, so that $AA^\top \in \mathbb{R}^{m \times m}$ is invertible.) Define the vector

$$\pi(x) = x - A^\top(AA^\top)^{-1}Ax.$$

(a) Show that $A\pi(x) = 0$, so that $\pi(x) \in S$.

(b) Show that for any $v \in S$, the vector $v - \pi(x)$ is orthogonal to $x - \pi(x)$. Draw a picture of this (in two dimensions).

(c) Recall that the ℓ_2, or Euclidean, norm of a vector $v \in \mathbb{R}^n$ is $\|v\|_2 = \sqrt{v^\top v} = \sqrt{\sum_{i=1}^n v_i^2}$. Show that $\pi(x)$ is the *projection of x onto S*, meaning that solves

$$\underset{v}{\text{minimize}} \quad \|v - x\|_2^2 \quad \text{subject to} \quad Av = 0.$$

Hint: show that for any $v \in S$, i.e., $v \in \mathbb{R}^n$ satisfying $Av = 0$, you have $\|v - x\|_2^2 = \|v - \pi(x)\|_2^2 + \|x - \pi(x)\|_2^2$, and conclude that unless $v = \pi(x)$, one has

$$\|v - x\|_2^2 > \|v - \pi(x)\|_2^2.$$

(d) **Extra credit:** Suppose that S is the 1 dimensional subspace $S = \{tu \mid t \in \mathbb{R}\}$, where $u \in \mathbb{R}^n$ is a unit vector (i.e. $\|u\|_2 = 1$). Let $A \in \mathbb{R}^{n-1 \times n}$ have rows forming an orthogonal basis for the space $S^\perp = \{v \in \mathbb{R}^n \mid Av = 0\} = \{v \in \mathbb{R}^n \mid u^\top v = 0\}$ perpendicular to u, that is, $AA^\top = I_{n-1}$, the $(n-1) \times (n-1)$ identity matrix. Write $A^\top A$ and $\pi(x)$ for this A in terms of the unit vector u.

Question 6.6 (Poisson distributions for radioactive decay)**:** In class we discussed that for particles decaying at rate λ, the probability that an individual particle survives until time t follows an exponential distribution, that is, $P(X \ge t) = e^{-\lambda t}$. We now show that when individual particles follow this decay probability at rate λ, then the number of particles decaying in an amount of time T follows a $\mathsf{Poisson}(\lambda T)$ distribution.

To that end, let $X_1, X_2, X_3, \ldots$ be the times between respective decay events, that is, X_i is the time between decay event i and decay event $i-1$ (so X_1 is the time of the first decay, and the X_i are independent and identically distributed random variables, each satisfying $X_i \sim \mathsf{Exp}(\lambda)$). Define the random variable

$$Y := \#\left\{\text{Decay events before time } T\right\},$$

so that we also have the identity

$$Y = \min_{i \in \mathbb{N}} \left\{ i \text{ such that } \sum_{j=1}^{i} X_j \le T, \ \sum_{j=1}^{i+1} X_j > T \right\}.$$

Show that

$$Y \sim \mathsf{Poisson}(\lambda T).$$

Hint: Use that

$$Y = i \ \text{ if and only if } \ \sum_{j=1}^{i} X_j \le T \text{ and } X_{i+1} > T - \sum_{j=1}^{i} X_j.$$

The identity that for any event A and any random variable Z with density f, we have $P(A) = P(A, Z \in \mathbb{R}) = \int P(A \mid Z = z) f(z) dz$ is likely to be useful as well.

Question 6.7* (Extra credit: Rigorously getting densities)**:** In this question, we make rigorous a few of the density approximations we made in defining densities. For a rectangle A with widths Δx and Δy, recall that $\text{Vol}(A) = \Delta x \Delta y$. Let $f : \mathbb{R}^2 \to \mathbb{R}_+$ be a function satisfying $\int_{\mathbb{R}^2} f(x,y) dx dy = 1$, so that it is a valid density. In this question, we will assume that f is *locally Lipschitz*, meaning that at any point $(x, y) \in \mathbb{R}^2$, there are $\delta > 0$ and $L < \infty$ such that $|x' - x| \leq \delta$ and $|y' - y| \leq \delta$ implies that

$$|f(x, y) - f(x', y')| \leq L \max\{|x - x'|, |y - y'|\}.$$

Let the pair of random variables (X, Y) have joint density f.

(a) Let x, y, δ, L be as above. Show that if A is the rectangle $A = [x - \Delta x/2, x + \Delta x/2] \times [y - \Delta y/2, y + \Delta y/2]$, where $0 < \Delta x \leq \delta$ and $0 < \Delta y \leq \delta$, then

$$|P((X, Y) \in A) - f(x, y)\Delta x \Delta y| \leq L(\Delta x)^2 \Delta y + L(\Delta y)^2 \Delta x$$

(b) Show that if A_Δ, defined for any $\Delta > 0$, are any collection of rectangles satisfying the containment

$$A_\Delta \subset [x - \Delta/2, x + \Delta/2] \times [y - \Delta/2, y + \Delta/2],$$

then

$$f(x, y) = \lim_{\Delta \to 0} \frac{P((X, Y) \in A_\Delta)}{\text{Vol}(A_\Delta)}$$