STATS 116: Homework 5

Due: Thursday, August 3, 2023 at 10:00 pm PDT on Gradescope

There are 7 problems on this assignment, each worth 8 points, although subparts within a problem may not be equally weighted. Credit will be assigned primarily based on reasoning and work, not the final answer. You do not need to simplify arithmetic expressions unless otherwise noted. While you may discuss the problems on this assignment other students, you must write up your own solutions. As per the syllabus, you may occasionally use the Internet or other public resources to clarify concepts with citation when this information is used as part of your own solution to a homework problem. However, you may not search for direct solutions to any problems assigned for homework or exams. For example, you can ask ChatGPT to clarify a particular concept from lecture that may be related to a problem, but you cannot feed it any part of a course assignment or a substantively similar version.

1. (BH 5.37) Let T be the time until a radioactive particle decays, and suppose (as is often done in physics and chemistry) that $T \sim \text{Expo}(\lambda)$.

   (a) The half-life of the particle is the time at which there is a 50% chance that the particle has decayed (in statistical terminology, this is the median of the distribution of T). Find the half-life of the particle.

   Setting $P(T > t) = \exp(-\lambda t) = 1/2$ and solving for t, we get that the half-life is $t = \log 2/\lambda$.

   (b) Show that for ϵ a small, positive constant, the probability that the particle decays in the time interval $[t, t+\epsilon]$, given that it has survived until time t, does not depend on t and is approximately proportional to ϵ. Hint: $\exp(x) \approx 1 + x$ if $x \approx 0$.

   Using the definition of conditional probability and the Expo(λ) CDF,

$$\begin{aligned} P(T \in [t, t+\epsilon] \mid T \geq t) &= \frac{P(T \in [t, t+\epsilon], T \geq t)}{P(T \geq t)} \\ &= \frac{P(t \leq T \leq t+\epsilon)}{P(T \geq t)} \\ &= \frac{\exp(-\lambda t) - \exp(-\lambda(t+\epsilon))}{\exp(-\lambda t)} = 1 - \exp(-\lambda\epsilon) \end{aligned}$$

   which does not depend on t. For $\epsilon > 0$ small, the above is approximately $1 - (1 - \lambda\epsilon) = \lambda\epsilon$ by the hint.

(c) Now consider n radioactive particles, with i.i.d. times until decay $T_1, \dots, T_n \sim \text{Expo}(\lambda)$. Let L be the first time at which one of the particles decays. Find the CDF of L, $\mathbb{E}(L)$, and $\text{Var}(L)$.

As shown in class, $L = \min(T_1, \dots, T_n) \sim \text{Expo}(n\lambda)$. Hence L has CDF

$$F_L(\ell) = \begin{cases} 0 & \ell \leq 0 \\ 1 - \exp(-n\lambda\ell) & \ell > 0 \end{cases}$$

for all $\ell > 0$ with $\mathbb{E}(L) = 1/(n\lambda)$ and $\text{Var}(L) = 1/(n^2\lambda^2)$, by the known CDF, mean, and variance of the Exponential distribution.

2. Let f be a PDF supported on a finite interval (a, b).

(a) Suppose I pick a point uniformly at random in the region in $\mathbb{R}^2$ bounded by the x-axis, the graph of f, and the vertical lines $x = a$ and $x = b$. What is the PDF of the x-coordinate of that point?

Let R be the given region. Then

$$|R| = \int_a^b \int_0^{f(x)} dydx = \int_a^b f_X(x) = 1$$

so the coordinates (X, Y) of the point have constant PDF $1/|R| = 1$ on R. By marginalization, for each $x \in [a, b]$ the PDF of X evaluates to

$$f_X(x) = \int_0^{f(x)} f_{XY}(x, y)dy = \int_0^{f(x)} 1dy = f(x)$$

Hence the x-coordinate has PDF f.

(b) Now let g be another PDF supported on (a, b) such that $f(x) \leq Mg(x)$ for all $x \in [a, b]$. Suppose Y has PDF g and $U \sim \text{Unif}(0, 1)$ independently of Y. What is the conditional PDF of Y given the event that $U \leq f(Y)/(Mg(Y))$? Hint: You may consider first computing the conditional CDF by integrating the joint density of (Y, U).

Let A be the event $U \leq f(Y)/Mg(Y)$. Following the hint, we consider the conditional CDF:

$$F_{Y|A}(k \mid A) = P(Y \leq k \mid A) = \frac{P(Y \leq k, A)}{P(A)}$$

Note (Y, U) has support $(a, b) \times (0, 1)$ with joint PDF satisfying $f_{YU}(y, u) = g(y)f_U(u) = g(y)$ for all $(y, u) \in (a, b) \times (0, 1)$. Thus for each $k \in (a, b)$ we have

$$P(Y \leq k, A) = \int_a^k \int_0^{f(y)/Mg(y)} f_{YU}(y, u)dudy = \int_a^k g(y) \int_0^{f(y)/Mg(y)} dudy = \int_a^k \frac{f(y)}{M} dy$$

Taking derivatives with respect to k, we get the conditional PDF

$$f_{Y|A}(k \mid A) = \frac{f(k)}{MP(A)}$$

Since PDFs must integrate to 1, and we've shown this conditional PDF is just f up to a normalizing constant, we conclude the conditional PDF $f_{Y|A}$ is in fact just f.

3. (BH 7.2) Alice, Bob, and Carl arrange to meet for lunch on a certain day. They arrive independently at uniformly distributed times between 1 pm and 1:30 pm on that day.

(a) What is the probability that Carl arrives first?

Let A, B, and C be the arrival times of Alice, Bob, and Carl, respectively, in minutes after 1pm. Then A, B, and C are i.i.d. Unif$(0, 30)$. By symmetry, the probability that Carl arrives first is $1/3$. Formally, we have

$$P(\min(A,B,C) = A) = P(\min(B,C,A) = B) = P(\min(C,A,B) = C)$$

by exchangeability of (A,B,C). But of course $\min(A,B,C) = \min(B,C,A) = \min(C,A,B)$, and ties have probability 0 since $A-B$, $B-C$, and $C-A$ are all continuous (e.g. by convolution) and hence none of them can equal 0 with positive probability. Thus the events $\min(A,B,C) = A$, $\min(B,C,A) = B$, and $\min(C,A,B) = C$ partition the entire sample space, so each of them has probability $1/3$.

For the rest of this problem, assume that Carl arrives first at 1:10 pm, and condition on this fact.

(b) What is the probability that Carl will have to wait more than 10 minutes for one of the others to show up? (So consider Carl's waiting time until at least one of the others has arrived.)

By considering the definition of conditional probability (in terms of conditional distributions given $C = 10$):

$$\begin{aligned}
P(A > 20, B > 20 \mid C = 10, A > 10, B > 10) &= \frac{P(A > 20, B > 20 \mid C = 10)}{P(A > 10, B > 10 \mid C = 10)} \\
&= \frac{P(A > 20)P(B > 20)}{P(A > 10)P(B > 10)} \\
&= \frac{1/3 \cdot 1/3}{2/3 \cdot 2/3} \\
&= \boxed{\frac{1}{4}}
\end{aligned}$$

where we use independence of A, B, and C for the second equality.

(c) What is the probability that Carl will have to wait more than 10 minutes for both of the others to show up? (So consider Carl's waiting time until both of the others has arrived.)

Similar to above, we have

$$\begin{aligned}
P(A > 20 \cup B > 20 \mid C = 10, A > 10, B > 10) &= 1 - P(A \leq 20, B \leq 20 \mid C = 10, A > 10, B > 10) \\
&= 1 - \frac{P(10 < A \leq 20, 10 < B \leq 20 \mid C = 10)}{P(A > 10, B > 10 \mid C = 10)} \\
&= 1 - \frac{P(10 < A \leq 20)P(10 < B \leq 20)}{P(A > 10)P(B > 10)} \\
&= 1 - \frac{1/3 \cdot 1/3}{2/3 \cdot 2/3} \\
&= \boxed{\frac{3}{4}}
\end{aligned}$$

(d) What is the probability that the person who arrives second will have to wait more than 5 minutes for the third person to show up?

We compute

$$\begin{aligned}
P(|A - B| > 5 \mid C = 10, A > 10, B > 10) &= \frac{P(A > 10, B > 10, |A - B| > 5 \mid C = 10)}{P(A > 10, B > 10) \mid C = 10)} \\
&= \frac{P(A > 10, B > 10, |A - B| > 5)}{P(A > 10, B > 10)}
\end{aligned}$$

The region $A > 10$, $B > 10$, and $|A - B| > 5$ within the support $(0, 30)^2$ of (A, B) has area 225 (it is two disjoint right triangles, each of which has legs of lengths 15 and 15). Since the joint PDF of (A, B) is a constant $1/30^2 = 1/900$ in this support, we have

$$P(A > 10, B > 10, |A - B| > 5) = \frac{225}{900} = \frac{1}{4}$$

Thus the desired probability is

$$P(|A - B| > 5 \mid C = 10, A > 10, B > 10) = \frac{\frac{1}{4}}{P(A > 10)P(B > 10)} = \frac{\frac{1}{4}}{\frac{2}{3} \cdot \frac{2}{3}} = \boxed{\frac{9}{16}}$$

4. (BH 7.10) Let X and Y be i.i.d. Expo(λ), and $T = X + Y$.

(a) Find the conditional CDF of T given $X = x$. Be sure to specify where it is zero.

Let $x > 0$. The conditional CDF of T given $X = x$ is

$$F_{T|X}(t \mid x) = P(T \le t \mid X = x) = P(Y \le t - x \mid X = x) = P(Y \le t - x) = 1 - \exp(-\lambda(t - x))$$

for $t > x$ (and the conditional CDF is 0 for $t \le x$). Independence of X and Y was used to drop the condition $X = x$ in $P(Y \le t - x \mid X = x)$. In other words, the conditional distribution of T given $X = x$ is a shifted Exponential: it is the distribution of $x+Y$ with $Y \sim \text{Expo}(\lambda)$.

(b) Find the conditional PDF $f_{T|X}(t \mid x)$, and verify that it is a valid PDF (i.e. integrates to 1).

Differentiating, for $t > x$ we have

$$f_{T|X}(t \mid x) = \frac{\partial}{\partial t} F_{T|X}(t \mid x) = \lambda \exp(-\lambda(t - x))$$

This is a valid PDF since, letting $y = t - x$, we have

$$\int_x^\infty \lambda \exp(-\lambda(t - x))dt = \int_0^\infty \lambda \exp(-\lambda y)dy = 1$$

using our knowledge that the Expo(λ) PDF integrates to 1.

(c) Find the conditional PDF $f_{X|T}(x \mid t)$. Hint: This can be done using Bayes' rule without having to know the marginal PDF of T, by recognizing what the conditional PDF is up to a normalizing constant. Then the normalizing constant must be whatever is needed to make the conditional PDF valid.

By Bayes' rule, for $0 < x < t$ we have

$$f_{X|T}(x \mid t) \propto f_{T|X}(t \mid x) f_X(x) = \lambda \exp(-\lambda(t - x)) \cdot \lambda \exp(-\lambda x) = \lambda^2 \exp(-\lambda t)$$

This doesn't depend on x, suggesting that the PDF is constant on $(0, t)$. That is, $f_{X|T}(x \mid t)$ is $1/t$ for $x \in (0, t)$, and 0 otherwise.

(d) Use Bayes' rule to show that the marginal PDF of T is given by $f_T(t) = \lambda^2 t \exp(-\lambda t)$ for $t > 0$.

By Bayes' rule, for $t > 0$ we have

$$\begin{aligned} f_T(t) &= \frac{f_{T|X}(t \mid x) f_X(x)}{f_{X|T}(x \mid t)} \quad \text{for all } 0 < x < t \\ &= \frac{\lambda^2 \exp(-\lambda t)}{1/t} = \lambda^2 t \exp(-\lambda t) \end{aligned}$$

as desired.

5. (BH 7.63) A chicken lays n eggs. Each egg independently does or doesn't hatch, with probability p of hatching. For each egg that hatches, the chick does or doesn't survive (independently of the other eggs), with probability s of survival. Let $N \sim \text{Bin}(n, p)$ be the number of eggs which hatch, X be the number of chicks which survive, and Y be the number of chicks which hatch but don't survive (so $X + Y = N$). Find the marginal PMF of X, and the joint PMF of X and Y. Are X and Y independent?

We will give a story proof that $X \sim \text{Bin}(n, ps)$. Consider any one of the n eggs. With probability p, it hatches. Given that it hatches, with probability s the chick survives. So the probability is ps of the egg hatching a chick which survives. Thus, $X \sim \text{Bin}(n, ps)$ with PMF

$$P(X = k) = \binom{n}{k}(ps)^k(1-ps)^{n-k}, \quad k = 0, 1, \ldots, n$$

By a similar argument, $Y \sim \text{Bin}(n, p(1-s))$.

The joint PMF of X and Y can be found by noting that for any nonnegative integers i, j, we have

$$\begin{aligned}
p_{XY}(i,j) = P(X = i, Y = j) &= P(X = i, N = i + j) \\
&= P(X = i \mid N = i + j)P(N = i + j) \text{ by definition of conditional probability} \\
&= \binom{i+j}{i} s^i (1-s)^j \cdot \binom{n}{i+j} p^{i+j}(1-p)^{n-i-j} \text{ since } X \mid N = n \sim \text{Bin}(n, s) \\
&= \frac{n!}{i!j!(n-i-j)!}(ps)^i(p(1-s))^j(1-p)^{n-i-j}
\end{aligned}$$

In particular $p_{XY}(0,0) = (1-p)^n$ while

$$P(X = 0)P(Y = 0) = (1-ps)^n(1-p(1-s))^n$$

which is not the same if $s \neq 1$, hence X and Y are not independent.

6. As in the setting of beta-binomial conjugacy, suppose we have a coin that lands heads with unknown probability p, which we express with a prior $p \sim \text{Unif}(0, 1)$. However, this time, instead of flipping the coin a fixed number of times n, we decide to collect data by continually flipping until the first time the coin lands heads. Let X be the number of tails observed before observing the first heads.

   (a) Specify the posterior distribution of p given $X = x$, for $x = 0, 1, \ldots$.

   Note $X \mid p \sim \text{Geom}(p)$. Thus the conditional density of p given $X = x$ is

$$f_{p|X}(k \mid x) \propto P(X = x \mid p = k)f_p(k) = (1-k)^x k$$

By pattern matching, this is the Beta$(2, x+1)$ distribution. This is intuitive as it corresponds to the same posterior distribution we'd get if we had flipped the coin $x+1$ times and gotten a sequence of x tails followed by 1 heads. The fact that we chose a random stopping time doesn't affect the posterior distribution.

(b) Find the mean and variance of p given $X = x$. That is, what are the mean and variance of p under the posterior distribution in part (a)? Note that your answer will depend on x. Hint: You can use without proof the fact that PDF of any known distribution covered in lecture integrates to 1.

Using the Beta$(2, x+1)$ PDF, we have

$$\mathbb{E}[p \mid X = x] = \int_0^1 y \frac{\Gamma(x+3)}{\Gamma(2)\Gamma(x+1)} y(1-y)^x dy = \frac{\Gamma(x+3)}{\Gamma(2)\Gamma(x+1)} \int_0^1 y^2(1-y)^x dy$$

Since the Beta$(3, x+1)$ PDF integrates to 1, we know

$$\frac{\Gamma(x+4)}{\Gamma(3)\Gamma(x+1)} \int_0^1 y^2(1-y)^x dy = 1$$

and hence

$$\mathbb{E}[p \mid X = x] = \frac{\Gamma(x+3)}{\Gamma(2)\Gamma(x+1)} \cdot \frac{\Gamma(3)\Gamma(x+1)}{\Gamma(x+4)} = \boxed{\frac{2}{x+3}}$$

by recalling the identity $\Gamma(a+1) = a\Gamma(a)$ for all $a > 0$. Similarly by LOTUS

$$\mathbb{E}[p^2 \mid X = x] = \int_0^1 y^2 \frac{\Gamma(x+3)}{\Gamma(2)\Gamma(x+1)} y(1-y)^x dy = \frac{\Gamma(x+3)}{\Gamma(2)\Gamma(x+1)} \int_0^1 y^3(1-y)^x dy$$

Since the Beta$(4, x+1)$ PDF integrates to 1, we know

$$\frac{\Gamma(x+5)}{\Gamma(4)\Gamma(x+1)} \int_0^1 y^3(1-y)^x dy = 1$$

and hence

$$\mathbb{E}[p^2 \mid X = x] = \frac{\Gamma(x+3)}{\Gamma(2)\Gamma(x+1)} \cdot \frac{\Gamma(4)\Gamma(x+1)}{\Gamma(x+5)} = \frac{6}{(x+3)(x+4)}$$

Then

$$\text{Var}(p \mid X = x) = \mathbb{E}[p^2 \mid X = x] - (\mathbb{E}[p \mid X = x])^2 = \frac{6}{(x+3)(x+4)} - \frac{4}{(x+3)^2} = \boxed{\frac{2x+2}{(x+3)^2(x+4)}}$$

7. Suppose Z is a Standard Normal random variable. Let S be a random sign, meaning that $P(S = 1) = P(S = -1) = 1/2$, and assume S and Z are independent. Define $Y = SZ$.

(a) Find the marginal distribution of Y.

Letting Φ be the CDF of the Standard Normal, we note that for any real y, we have

$$\begin{aligned}
P(Y \le y) &= P(Y \le y \mid S = 1)P(S = 1) + P(Y \le y \mid S = -1)P(S = -1) \text{ by LOTP} \\
&= P(Z \le y \mid S = 1) \cdot \frac{1}{2} + P(-Z \le y \mid S = -1) \cdot \frac{1}{2} \\
&= P(Z \le y) \cdot \frac{1}{2} + P(-Z \le y) \cdot \frac{1}{2} \text{ by independence of } S \text{ and } Z \\
&= \Phi(y) \text{ since } Z \text{ and } -Z \text{ are both Standard Normal}
\end{aligned}$$

Thus $Y \sim \mathcal{N}(0, 1)$.

(b) Are Y and S independent? Are Y and Z independent? Explain.

The event $S \le s$ for any s has probability 0 if $s < -1$ and probability 1 if $s \ge 1$. Otherwise it is equivalent to the event $S = -1$. Thus by Proposition 10.5, to show Y and S are independent it suffices to show $P(Y \le y, S \le -1) = P(Y \le y)P(S \le -1)$. But indeed

$$\begin{aligned}
P(Y \le y, S \le -1) &= P(Z \ge -y, S = -1) \\
&= P(Z \ge -y)P(S = -1) \text{ by independence of } Z \text{ and } S \\
&= \frac{1}{2}(1 - \Phi(-y)) \\
&= \frac{1}{2}\Phi(y) = P(S \le -1)P(Y \le y) \text{ by symmetry of } \Phi \text{ and part (a)}
\end{aligned}$$

However Y and Z are not independent. For instance, given $|Z| < 1$ we must have $|Y| < 1$, hence the conditional distribution of Y given $|Z| < 1$ cannot be the marginal distribution of Y, i.e. Standard Normal (which has positive probability of being outside $[-1, 1]$).

(c) Are Y and S exchangeable? Are Y and Z exchangeable? Explain.

Y and S are not exchangeable because they do not have the same marginal distribution. However, Y and Z are exchangeable. To see this, note that for any real numbers y and z, we have

$$\begin{aligned}
P(Y \le y, Z \le z) &= P(Y \le y, Z \le z, S = 1) + P(Y \le y, Z \le z, S = -1) \\
&= P(Z \le \min(y, z), S = 1) + P(-y \le Z \le z, S = -1) \\
&= P(Z \le \min(y, z))P(S = 1) + P(-y \le Z \le z)P(S = -1) \text{ by independence of } Z \text{ and } S \\
&= \frac{1}{2}(\Phi(\min(y, z)) + \max(0, \Phi(z) - \Phi(-y)))
\end{aligned}$$

Thus

$$
\begin{aligned}
P(Z \le y, Y \le z) = P(Y \le z, Z \le y) &= \frac{1}{2}(\Phi(\min(z, y)) + \max(0, \Phi(y) - \Phi(-z))) \text{ by the previous display} \\
&= \frac{1}{2}(\Phi(\min(y, z)) + \max(0, (1 - \Phi(-y)) - (1 - \Phi(z)))) \text{ by symmetry} \\
&= \frac{1}{2}(\Phi(\min(y, z)) + \max(0, \Phi(z) - \Phi(-y))) \\
&= P(Y \le y, Z \le z)
\end{aligned}
$$

showing that (Y, Z) and (Z, Y) have the same joint CDF.