Stat 244, Aut 2017: HW4

due: Tuesday 10.24.17

1. Let X and Y be random variables with variances σ_X^2 and σ_Y^2, and correlation ρ. Calculate $\mathrm{Cov}(X+Y, X-Y)$.

   Solution: By linearity of covariance,

$$\begin{aligned}\mathrm{Cov}(X+Y, X-Y) &= \mathrm{Cov}(X,X) - \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) - \mathrm{Cov}(Y,Y) \\ &= \sigma_X^2 - \sigma_X\sigma_Y\rho + \sigma_Y\sigma_X\rho - \sigma_Y^2 = \sigma_X^2 - \sigma_Y^2 \ .\end{aligned}$$

2. Suppose that you roll a fair die repeatedly until the kth time that you obtain a 1. Let X be the number of times you rolled the die (i.e. on your Xth roll, you obtained your kth 1). Let Y be the number of 6's that you observed during this process. What is the conditional distribution of Y given X?

   Solution: If the kth 1 was obtained on the nth roll, this means that out of the previous $n-1$ many rolls, we saw exactly $k-1$ many 1's. Therefore there were $n-k$ many rolls where we saw a 2 or 3 or 4 or 5 or 6. Since each of these is equally likely, and these rolls are independent from each other, this means that $Y \mid X \sim \mathrm{Binomial}(X-k, 0.2)$, or equivalently, $Y \mid (X=n) \sim \mathrm{Binomial}(n-k, 0.2)$. If we want a conditional PMF, then we have

$$p_{Y|X}(m|n) = \binom{n-k}{m} 0.2^m 0.8^{n-k-m}$$

   for the range of values $m = 0, 1, 2, \ldots, n-k$, by just using the PMF of a Binomial.

3. Suppose that X and Y are independent Geometric(p) random variables and $W = \min\{X, Y\}$. Calculate

$$P(X = x \mid W = x)$$

   where x is some fixed positive integer. Hint: first decide whether this probability should be equal to $1/2$, larger than $1/2$, or smaller than $1/2$.

   Solution: Fix any integer $x \geq 1$. We consider three cases:

   - $X = Y = x$. This occurs with probability $P(X=x)P(Y=x) = \big((1-p)^{x-1}\cdot p\big)^2$.
   - $X = x$ and $Y > x$. This occurs with probability $P(X=x)P(Y>x) = \big((1-p)^{x-1}\cdot p\big)\cdot(1-p)^x$.
   - $X > x$ and $Y = x$. This occurs with probability $P(X>x)P(Y=x) = \big((1-p)^{x-1}\cdot p\big)\cdot(1-p)^x$.

   So, in total,

$$P(W=x) = \big((1-p)^{x-1}\cdot p\big)^2 + 2\left((1-p)^{x-1}\cdot p\right)\cdot(1-p)^x$$

   while

$$P(X=x, W=x) = \big((1-p)^{x-1}\cdot p\big)^2 + \big((1-p)^{x-1}\cdot p\big)\cdot(1-p)^x$$

   and therefore

$$P(X=x \mid W=x) = \frac{P(X=x, W=x)}{P(W=x)} = \frac{1}{2-p} \ .$$

4. Let A and B be independent Exponential(1) random variables, and let C be a random sign, i.e. $P(C = +1) = P(C = -1) = 0.5$, drawn independently from A and B. Let $X = A \cdot C$ and let $Y = B \cdot C$.

   (a) Calculate $E(X)$, $E(Y)$, Var(X), Var(Y), and Cov(X, Y). You may use the following facts: for an Exponential(λ) random variable, its expected value is $1/\lambda$ and its variance is $1/\lambda^2$.

   Solution: First, since A and C are independent, $E(X) = E(A) \cdot E(C) = 0$, since $E(C) = 0$ clearly. Similarly $E(Y) = 0$.
   For variance, Var$(X) = E(X^2) - E(X)^2 = E(X^2) = E(A^2C^2) = E(A^2)$, since $C^2 = 1$ always. We also know from the stated fact that $1 = $ Var$(A) = E(A^2) - E(A)^2 = E(A^2) - 1^2$, and so $E(A^2) = 2$. So, Var$(X) = 2$. Similarly Var$(Y) = 2$.
   Finally, we calculate covariance. First, $E(XY) = E(AC \cdot BC) = E(A \cdot B \cdot C^2)$. Since $C^2 = 1$ always, this is equal to $E(AB) = E(A)E(B) = 1^2 = 1$, since $A \perp\!\!\!\perp B$ and the expected value of an Exponential(1) random variable is 1. Therefore, Cov$(X, Y) = E(XY) - E(X)E(Y) = 1$.

   (b) Calculate $P(X \le t \mid Y \le t)$, where $t > 0$ is a constant. Next, plot your answer as a function of t and/or calculate its value for a few different values of t, and describe what you observe.

   Solution: First,

$$P(Y \le t) = P(C = -1) + P(C = +1, B \le t) = 0.5 + 0.5(1 - e^{-t}).$$

   Next,

$$P(X \le t, Y \le t) = P(C = -1) + P(C = +1, A \le t, B \le t) = 0.5 + 0.5(1 - e^{-t})^2.$$

   So,

$$P(X \le t \mid Y \le t) = \frac{1 + (1 - e^{-t})^2}{1 + (1 - e^{-t})}.$$

   This probability is 1 when $t \to 0$, and approaches 1 as $t \to \infty$. For $t \approx 0.5$, it's smaller, around 0.82. To explain this:

   - If t is very large, then $X \le t$ is almost certain to be true, so the probability is close to 1
   - If t is very small (close to zero), then knowing $Y \le t$ means that probably $C = -1$, since it's much less likely than having $C = +1$ and $B \le t$. If this is the case, then $X \le t$ will certainly occur, since $X \le 0$ when $C = -1$. So the probability is again close to 1
   - If $t \approx 0.5$, then knowing that $Y \le t$ tells us that perhaps $C = -1$, but it's also plausible that $C = +1$ and $B \le t$. In the second case, it's far from certain that we would have $X \le t$ also—so the probability is noticeably smaller than 1.

   .

5. Let X and Y be random variables supported on $[0, 1] \times [0, 1]$, with joint density

$$f(x, y) = C \cdot (x^2 + y^2)$$

   on this region. Here C is a constant.

   (a) Calculate C.

   Solution: We know that the density has to integrate to 1. We have

$$\int_{x=0}^{1} \int_{y=0}^{1} (x^2 + y^2)\ dy\ dx = \int_{x=0}^{1} \left[yx^2 + \frac{1}{3}y^3 \right]_{y=0}^{1}\ dx = \int_{x=0}^{1} \left(x^2 + \frac{1}{3} \right)\ dx$$
$$= \left[\frac{1}{3}x^3 + \frac{1}{3}x \right]_{x=0}^{1} = \frac{2}{3}\ .$$

So, $C = 1.5$.

(b) Calculate the marginal density $f_X(x)$.

Solution: The support of X is $[0,1]$. On this region,

$$f_X(x) = \int_{y=0}^{1} f_{X,Y}(x,y)\ dy = \int_{y=0}^{1} 1.5(x^2+y^2)\ dy = \left[1.5x^2y + 0.5y^3\right]_{y=0}^{1} = 1.5x^2 + 0.5\ .$$

(c) Calculate the conditional density $f_{Y|X}(y|x)$ for $(x,y) \in [0,1]^2$.

Solution: The conditional density formula gives us

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1.5x^2+1.5y^2}{1.5x^2+0.5} = \frac{3x^2+3y^2}{3x^2+1} = \frac{3x^2}{3x^2+1} + \frac{3}{3x^2+1}\cdot y^2\ ,$$

on the range $0 \le y \le 1$.

(d) Are the variables X and Y positively correlated, negatively correlated, or uncorrelated? For this portion of the problem, it is not necessary to calculate Corr(X,Y) / Cov(X,Y) or to do any exact calculations (although this would also be fine)—it is sufficient to examine your calculations for $f_{Y|X}(y|x)$ and explain what you see. It may help to plot $f_{Y|X}(y|x)$ for various values of x.

Solution: By plotting the conditional density $f_{Y|X}(y|x)$ as a function of y for fixed values of x, we see that this is a steeply increasing function of y for small x values (e.g. $x = 0$ or $x = 0.1$), and increases more slowly as a function of y for larger x values (e.g. $x = 0.8$ or $x = 1$). This means that when X is small, larger Y values are much more likely than small Y values; when X is large, Y is closer to uniformly distributed. So, they are negatively correlated.
Alternately we could calculate Cov(X,Y). We have

$$E(X) = \int_{x=0}^{1} x \cdot f_X(x)\ dx = \int_{x=0}^{1} x(1.5x^2+0.5)\ dx = [\frac{1.5}{4}x^4 + \frac{0.5}{2}x^2]_{x=0}^{1} = 0.625$$

and by symmetry, $E(Y) = 0.625$ also. And,

$$E(XY) = \int_{x=0}^{1}\int_{y=0}^{1} xy \cdot f_{X,Y}(x,y)\ dy\ dx = \int_{x=0}^{1}\int_{y=0}^{1} xy(1.5x^2+1.5y^2)\ dy\ dx$$

$$= \int_{x=0}^{1} [\frac{1.5}{2}x^3y^2 + \frac{1.5}{4}xy^4]_{y=0}^{1}\ dx = \int_{x=0}^{1} 0.75x^3 + 0.375x\ dx = [\frac{0.75}{4}x^4 + \frac{0.375}{2}x^2]_{x=0}^{1} = 0.375.$$

So, Cov$(X,Y) = E(XY) - E(X)E(Y) = 0.375 - 0.625^2 = -0.015625$.

6. In this question we will use a basic example to learn about Bayesian statistics, which models parameters with prior distributions (sometimes to indicate uncertainty in our beliefs). Suppose that you have a coin which may not be fair. Its parameter P, which is the chance of landing Heads, could in theory lie anywhere in the range $[0,1]$. You flip this coin one time and let $X = \mathbb{1}_{\text{Heads}}$, and now would like to draw some conclusions about P. Let's choose a prior distribution for this parameter, and assume that P is drawn from a Uniform$[0,1]$ distribution.

(a) Write down a hierarchical model for this scenario, which should be of the form

$$\begin{cases} \text{(some variable)} \sim \text{(some distribution)} \\ \text{(some other variable)} \mid \text{(the first variable)} \sim \text{(some distribution)} \end{cases}$$

Solution:

$$\begin{cases} P \sim \text{Uniform}[0,1] \\ X \mid P \sim \text{Bernoulli}(P) \end{cases}$$

(b) Now calculate the following: for any $t \in [0,1]$, find $P(P \leq t, X = 0)$ and $P(P \leq t, X = 1)$. To do these calculations, you are working with a joint distribution where one variable is discrete and one is continuous; intuitive rules will apply for combining integrals and sums, etc. For example it's fine to write $P(X = 1 \mid P = p) = p$.

Solution: Since the density of P is just $f_P(p) = 1$ for $p \in [0,1]$, and since $P(X = 0 \mid P = p) = (1-p)$,

$$\begin{aligned} P(P \leq t, X = 0) = \int_{p=0}^{t} f_P(p) \cdot P(X = 0 \mid P = p)\, dp &= \int_{p=0}^{t} 1 \cdot (1-p)\, dp \\ &= \left[p - \frac{1}{2}p^2\right]_{p=0}^{t} = t - t^2/2\ . \end{aligned}$$

And, similarly,

$$P(P \leq t, X = 1) = \int_{p=0}^{t} f_P(p) \cdot P(X = 1 \mid P = p)\, dp = \int_{p=0}^{t} 1 \cdot p\, dp = \left[\frac{1}{2}p^2\right]_{p=0}^{t} = t^2/2\ .$$

(c) Finally calculate the conditional distribution of P, given that you observe $X = 1$. To do this, start with the conditional CDF, i.e. $P(P \leq t \mid X = k)$, then get the density.

In Bayesian statistics, the conditional distribution of P given our observed value of X, is called the posterior distribution for P—meaning its distribution after observing the data (which in this case is X, the data from tossing the coin).

Solution: First we have $P(X = 1) = P(P \leq 1, X = 1) = 1^2/2 = 1/2$, using the work above. So,

$$F_{P|X}(t \mid 1) = P(P \leq t \mid X = 1) = \frac{P(P \leq t, X = 1)}{P(X = 1)} = \frac{t^2/2}{1/2} = t^2\ .$$

Therefore,

$$f_{P|X}(t \mid 1) = \frac{d}{dt} F_{P|X}(t \mid 1) = 2t\ .$$

The support for P given $X = 1$, is still $[0,1]$.