Stat 244, Aut 2017: HW3

due: Tuesday 10.17.17

1. In class, we had the following example: $X \sim \text{Poisson}(100)$ is the number of photons emitted by an X-ray beam, and then Y is the number of photons that successfully pass through an object (i.e. the person being imaged). In that example, the distribution of Y, if we know how many photons X were sent into the object, is given by $Y \mid X \sim \text{Binomial}(X, 0.4)$, i.e. given that X many photons are sent into the object, each one has a 40% chance of making it through and passing out the other side.

   Calculate the probability mass function for (X, Y), that is, $p(k, \ell) = P(X = k, Y = \ell)$ (as a function of k and ℓ).

   Solution: The mass function is

$$p(k, \ell) = P(X = k, Y = \ell) = P(X = k) \cdot P(Y = \ell \mid X = k) = \frac{100^k}{k!} e^{-100} \cdot \binom{k}{\ell} 0.4^\ell 0.6^{k-\ell}$$

   for all $k = 0, 1, 2, \ldots$ and all $\ell = 0, \ldots, k$. (This last part—the support of the distribution—must be specified for the answer to be complete.)

2. (a) Let X and Y be random variables with $X \sim \text{Exponential}(\lambda_1)$ and $Y \sim \text{Exponential}(\lambda_2)$. Suppose that X and Y are independent. Let $Z = \max\{X, Y\}$. Calculate the CDF of Z.

   Solution: For $z < 0$, $F_Z(z) = 0$ since Z is always nonnegative. For $z \geq 0$,

$$\begin{aligned} F_Z(z) = P(Z \leq z) = P(\max\{X, Y\} \leq z) = P(X \leq z, Y \leq z) &= P(X \leq z) P(Y \leq z) \\ &= (1 - e^{-\lambda_1 z})(1 - e^{-\lambda_2 z}) \ . \end{aligned}$$

   (b) Let $X \sim \text{Exponential}(1)$ and let $Y \sim \text{Bernoulli}(0.5)$. Let $Z = X + Y$. Calculate the CDF of Z.

   Solution: For $z \leq 0$, $F_Z(z) = 0$ since Z is always nonnegative. For $z \geq 0$,

$$\begin{aligned} F_Z(z) = P(Z \leq z) &= P(Z \leq z \cap Y = 0) + P(Z \leq z \cap Y = 1) \\ &= P(Y = 0) \cdot P(Z \leq z \mid Y = 0) + P(Y = 1) \cdot P(Z \leq z \mid Y = 1) \\ &= P(Y = 0) \cdot P(X \leq z) + P(Y = 1) \cdot P(X \leq z - 1) \\ &= \begin{cases} 0.5 \cdot (1 - e^{-z}) + 0.5 \cdot (1 - e^{-(z-1)}), & \text{if } z \geq 1, \\ 0.5 \cdot (1 - e^{-z}), & \text{if } 0 < z < 1. \end{cases} \end{aligned}$$

3. You have n random number generators, where the ith one draws a number uniformly at random from the interval $[0, t_i]$. Here the t_i's are arbitrary positive integers. Let X_i be the number drawn by the ith random number generator.

   (a) What is the expected value of the sum, $S = X_1 + \cdots + X_n$? (Your answer will be in terms of $t_1, \ldots, t_n$.)

   Solution: For each i, we know that $E(X_i) = t_i/2$ by symmetry, or alternately we could calculate this expectation using the density $f_{X_i}(x) = \frac{1}{t_i}$ on the interval $[0, t_i]$:

$$E(X_i) = \int_{x=0}^{t_i} x \cdot \frac{1}{t_i} \, dx = \left[\frac{1}{2t_i} x^2 \right]_{x=0}^{t_i} = t_i/2 \ .$$

Then we have

$$E(S) = E(X_1) + \cdots + E(X_n) = \frac{t_1 + \cdots + t_n}{2}\ .$$

(b) What is the expected value of Y, which counts how many of the X_i's are ≤ 1? (Your answer will be in terms of $t_1, \ldots, t_n$.)

Solution: First define Z_i to be the indicator variable for the event that $X_i \leq 1$, that is, $Z_i = 1$ if $X_i \leq 1$ and $Z_i = 0$ if $X_i > 1$. Then Z_i is Bernoulli(p_i) where

$$p_i = P(X_i \leq 1) = \frac{1}{t_i}.$$

Then $Y = Z_1 + \cdots + Z_n$ and so

$$E(Y) = E(Z_1) + \cdots + E(Z_n) = \sum_{i=1,\ldots,n} \frac{1}{t_i}.$$

4. Let $X \sim$ Exponential(λ) and let t be a constant with $0 < t < \lambda$.

(a) What is $\mathbb{E}\left[e^{tX}\right]$?

Solution:

$$\mathbb{E}\left[e^{tX}\right] = \int_{x=0}^{\infty} e^{tx} \cdot \lambda e^{-\lambda x}\ dx = \lambda \int_{x=0}^{\infty} e^{-(\lambda - t)x}\ dx = \lambda \cdot \left[-\frac{1}{\lambda - t} e^{-(\lambda-t)x}\right]_{x=0}^{\infty} = \frac{\lambda}{\lambda - t}.$$

(b) Use the Markov inequality to prove a bound on $P(e^{tX} \geq a)$ (here $a > 0$ is any positive number, while we assume $0 < t < \lambda$ as before).

Solution:

$$P(e^{tX} \geq a) \leq \frac{\mathbb{E}\left[e^{tX}\right]}{a} = \frac{\lambda}{a(\lambda - t)}.$$

(c) Now reformulate this into a bound on $P(X \geq b)$ (here $b > 0$ is any positive number, and again $0 < t < \lambda$). How does this compare to the true value of the probability $P(X \geq b)$?

Solution:

$$P(X \geq b) = P(e^{tX} \geq e^{tb}) \leq \frac{\lambda}{e^{tb}(\lambda - t)} = e^{-tb} \cdot \frac{\lambda}{\lambda - t}.$$

To compare, from the CDF of the exponential distribution we know that

$$P(X \geq b) = 1 - P(X \leq b) = 1 - F(b) = 1 - (1 - e^{-\lambda b}) = e^{-\lambda b}.$$

Obviously the upper bound is larger for any $t \in (0, \lambda)$. For example if we set $t = \lambda/2$ we would obtain an upper bound of $2e^{-\lambda b/2}$.

5. (a) Suppose you flip a fair coin 10 times. Let X be the total number of times that you see the sequence HT. What is $P(X = 0)$?

Solution: To get $X = 0$, we need to never have a Heads followed by a Tails. The only way this is possible is to have a stretch of all tails for the first n flips followed by only Heads for the next $10 - n$ flips, where we could have $n = 0, 1, \ldots, 10$. Each of these possibilities has probability 0.5^{10}, so

$$P(X = 0) = \sum_{n=0}^{10} P(n \text{ Tails and then } (10-n) \text{ Heads}) = 11 \cdot 0.5^{10} = 0.0107 \ .$$

(b) What is $E(X)$? (Hint: think of X as a sum.)

Solution: Let A_i be the event that coin flips #i and #$(i+1)$ give the sequence HT. Then we have

$$X = \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_9}$$

where $\mathbb{1}_A$ is the indicator variable for the event A, i.e. $\mathbb{1}_A$ is a random variable equal to 1 if the event A occurs and 0 if the event A does not occur.
For each i, we have $P(A_i) = 0.25$ and so

$$E(X) = E(\mathbb{1}_{A_1}) + \cdots + E(\mathbb{1}_{A_9}) = 9 \cdot 0.25 = 2.25 \ .$$

(c) Suppose you draw cards one at a time from a standard deck, without replacement, and record their colors. Let Y be the total number of times that you see the sequence RB (i.e. a red card followed by a black card). What is $P(Y = 0)$?

Solution: This can only happen if we draw all the black cards first then all the red cards. There are 26! possible orderings for the black cards, and 26! for the red cards, as compared to 52! possible orders for the entire deck of cards, so the probability is

$$\frac{(26!)^2}{52!} \approx 2 \cdot 10^{-15}$$

(d) What is $E(Y)$? (Hint: think of Y as a sum.)

Solution: Let A_i be the event that cards #i and #$(i+1)$ give the sequence RB. Then we have

$$X = \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_{51}} \ .$$

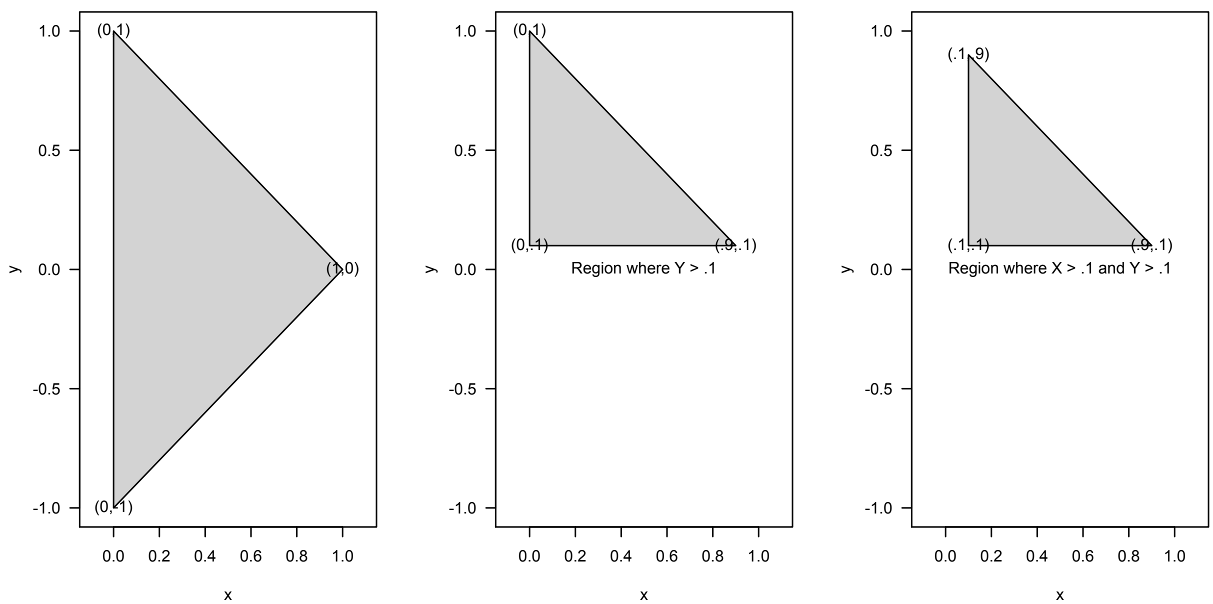
For each i, we have $P(A_i) = \frac{26}{52} \cdot \frac{26}{51}$ and so

$$E(Y) = E(\mathbb{1}_{A_1}) + \cdots + E(\mathbb{1}_{A_{51}}) = 51 \cdot \frac{26}{52} \cdot \frac{26}{51} = 13 \ .$$

6. Suppose that (X, Y) is a point chosen uniformly at random from the triangular region formed by connecting the points $(1, 0)$, $(0, 1)$ and $(0, -1)$.

   (a) Calculate $P(X > 0.1 \mid Y > 0.1)$.

   Solution:

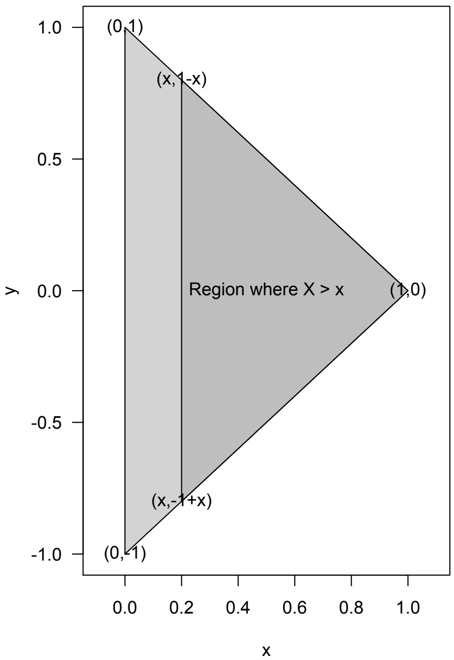


   Let T be the triangular region (left figure). First, the region $T \cap \{y > 0.1\}$ is a triangle with vertices $(0.9, 0.1)$, $(0, 0.1)$, $(0, 1)$ so it has area $.9 \cdot 1/2 = .45$. Next, the region $T \cap \{x > 0.1, y > 0.1\}$ is a triangle with vertices $(0.9, 0.1)$, $(0.1, 0.1)$, $(0.1, 0.9)$, so it has area $.8 \cdot .8/2 = .32$. The total area of the region T is 1. So the probability is

$$P(X > 0.1 \mid Y > 0.1) = \frac{P(X > 0.1, Y > 0.1)}{P(Y > 0.1)} = \frac{0.32}{0.45} = 0.711 \ .$$

   (b) What is the CDF of the variable X?

   Solution:



The range of possible values for X is the interval $[0, 1]$, so we only need to calculate $F_X(x)$ for $0 \le x \le 1$. We have

$$F_X(x) = P(X \le x) = 1 - P(X > x)$$
$$= 1 - \frac{\text{Area of the triangle with vertices (1,0), (x,1-x), and (x,-1+x)}}{\text{area of the triangle}} = 1 - \frac{(1-x)^2}{1} = 2x - x^2.$$