CS 70 Discrete Mathematics and Probability Theory
Summer 2021

HW 5

1 Throwing Frisbees

Shahzar and his $n-1$ friends stand in a circle and play the following game: Shahzar throws a frisbee to one of the other people in the circle randomly, with each person being equally likely, and thereafter, the person holding the frisbee throws it to someone else in the circle, again uniformly at random. The game ends when someone throws the frisbee back to Shahzar.

(a) What is the expected number of times the frisbee is thrown through the course of the game?

(b) What is the expected number of people that never get the frisbee during the game?

Solution:

(a) Let X be the random variable that corresponds to the number of times the frisbee was thrown. We can recognize that X is a shifted geometric random variable, as after the first throw, the probability that the game ends is always exactly $\frac{1}{n-1}$. Thus, the expected number of throws is 1 more than the expected value of a geometric random variable with parameter $p = \frac{1}{n-1}$, hence the expected value is $1+n-1 = \boxed{n}$.

Alternate Solution: Note that X is nonnegative and integer-valued, so we can use the tail-sum formula. Recall that the tail-sum formula states that

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{P}[X \geq i],$$

so it suffices to compute $\mathbb{P}[X \geq i]$ for each $i \geq 1$. The game clearly always consists of at least one throw, so $\mathbb{P}[X \geq 1] = 1$. For each subsequent i, the probability that the game continues for i or more throws is the probability that Shahzar doesn't have the frisbee thrown to him in the first $i-1$ throws. Since Shahzar throws the frisbee first, the first throw obviously goes to someone different. From each point thereafter, the probability that the next throw goes to someone other than Shahzar is $\frac{n-2}{n-1}$ as there are $n-1$ possible people and $n-2$ people who aren't Shahzar. Thus, the probability that Shahzar doesn't get the frisbee in the first $i-1$ throws is

$$\mathbb{P}[X \geq i] = \left(\frac{n-2}{n-1}\right)^{i-2}.$$

Finally, we can substitute this expression into the tail-sum formula and compute the resulting geometric series:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{P}[X \geq i] = 1 + \sum_{i=2}^{\infty} \left(\frac{n-2}{n-1}\right)^{i-2} = \boxed{n}.$$

(b) Like in the part (a), let X be the random variable that corresponds to the number of times the frisbee is thrown. Moreover, number the people $1, 2, \dots, n-1$ and let Y_i be the indicator that person i never gets the frisbee thrown to them. Observe that if we let $Y = Y_1 + \cdots Y_{n-1}$, then Y is exactly the number of people who never get the frisbee.

Since X can only take on values in the set $\{2, 3, \dots\}$, we can use the Law of Total Probability to get that

$$\mathbb{P}[Y_i = 1] = \sum_{j=2}^{\infty} \mathbb{P}[Y_i = 1 \cap X = j].$$

We can compute the probability that we have a game consisting of j throws and that person i doesn't get the frisbee in that game as follows. In order for such a scenario to occur, for the first throw, we only need to guarantee that person i doesn't get the frisbee. For the next $j-2$ throws, we have to guarantee that person i doesn't get the frisbee *and* Shahzar doesn't get the frisbee (as otherwise the game would end prematurely). Lastly, we have to guarantee that the last throw goes to Shahzar so that the game consists of the correct number of throws. The probability that person i doesn't get the frisbee in the first throw is $\frac{n-2}{n-1}$. Given that person i doesn't have the frisbee, the probability that the next person to have the frisbee is neither person i nor Shahzar is $\frac{n-3}{n-1}$. Finally, the probability that Shahzar gets the frisbee on any given throw is $\frac{1}{n-1}$. Thus, our probability is

$$\mathbb{P}[Y_i = 1 \cap X = j] = \left(\frac{n-2}{n-1}\right)\left(\frac{n-3}{n-1}\right)^{j-2}\left(\frac{1}{n-1}\right).$$

Plugging this into the summation and using the geometric series formula, we get that

$$\mathbb{P}[Y_i = 1] = \sum_{j=2}^{\infty} \left(\frac{n-2}{n-1}\right)\left(\frac{n-3}{n-1}\right)^{j-2}\left(\frac{1}{n-1}\right) = \frac{n-2}{2n-2}.$$

Finally, we can use linearity of expectation to get our final answer:

$$\mathbb{E}Y = \sum_{i=1}^{n-1} \mathbb{E}Y_i = \sum_{i=1}^{n-1} \frac{n-2}{2n-2} = \boxed{\frac{n-2}{2}}.$$

2 Various Variance Problems

(a) Suppose that X and Y are both binomial random variables with parameters n and p and $\operatorname{Var}(X-Y) = 2$. Find $\operatorname{cov}(X,Y)$ in terms of n and p.

(b) Prove that if X and Y are independent random variables, then

$$\operatorname{Var}(XY) = \operatorname{Var}(X)\operatorname{Var}(Y) + \mathbb{E}[X]^2 \operatorname{Var}(Y) + \mathbb{E}[Y]^2 \operatorname{Var}(X).$$

Solution:

(a) Use the bilinearity of covariance:

$$\begin{aligned}\mathrm{Var}(X-Y) &= \mathrm{cov}(X-Y, X-Y)\\ &= \mathrm{cov}(X, X-Y) - \mathrm{cov}(Y, X-Y)\\ &= \mathrm{cov}(X,X) - 2\,\mathrm{cov}(X,Y) + \mathrm{cov}(Y,Y)\\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\,\mathrm{cov}(X,Y).\end{aligned}$$

Since X and Y are both binomial with parameters n and p, $\mathrm{Var}(X) = \mathrm{Var}(Y) = np(1-p)$, so

$$\mathrm{cov}(X,Y) = \frac{1}{2}(\mathrm{Var}(X) + \mathrm{Var}(Y) - \mathrm{Var}(X-Y)) = \boxed{np(1-p)-1}.$$

Alternate Solution: From the covariance formula, we know that

$$\mathrm{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y],$$

so we just need to figure out $\mathbb{E}[X]$, $\mathbb{E}[Y]$, and $\mathbb{E}[XY]$. The former two quantities are straightforward. We are given that both X and Y are binomial with parameters n and p, so $\mathbb{E}[X] = \mathbb{E}[Y] = np$. Now, to find the last quantity, we have to use the last piece of information given to us, namely, that the variance of $X-Y$ is equal to 2. We have that

$$\begin{aligned}2 &= \mathrm{Var}(X-Y)\\ &= \mathbb{E}[(X-Y)^2] - \mathbb{E}[X-Y]^2,\end{aligned}$$

but since $\mathbb{E}[X] = \mathbb{E}[Y]$, the second term above vanishes by linearity. Thus, we have

$$\begin{aligned}2 &= \mathbb{E}[(X-Y)^2]\\ &= \mathbb{E}[X^2 - 2XY + Y^2]\\ &= \mathbb{E}[X^2] + \mathbb{E}[Y^2] - 2\,\mathbb{E}[XY].\end{aligned}$$

Finally, to compute $\mathbb{E}[X^2]$, we can use the fact that $\mathrm{Var}(X) = np(1-p)$, as

$$np(1-p) = \mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - n^2p^2 \implies \mathbb{E}[X^2] = np(1-p) + n^2p^2.$$

We can do this exact algebra for Y, so

$$2 = \mathbb{E}[X^2] + \mathbb{E}[Y^2] - 2\,\mathbb{E}[XY] = 2np(1-p) + 2n^2p^2 - 2\,\mathbb{E}[XY],$$

hence $\mathbb{E}[XY] = n^2p^2 + np(1-p) - 1$, so our final answer is

$$\mathrm{cov}(X,Y) = n^2p^2 + np(1-p) - 1 - n^2p^2 = \boxed{np(1-p)-1}.$$

(b) Since X and Y are independent, we know that $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$ and we also know that $\mathbb{E}[X^2Y^2] = \mathbb{E}[X^2]\,\mathbb{E}[Y^2]$. Using this, we have the following algebra:

$$\begin{aligned}\mathrm{Var}(XY) &= \mathbb{E}[X^2Y^2] - \mathbb{E}[XY]^2\\ &= \mathbb{E}[X^2]\,\mathbb{E}[Y^2] - \mathbb{E}[X]^2\,\mathbb{E}[Y]^2.\end{aligned}$$

Now, we can compute that

$$\begin{aligned}\operatorname{Var}(X)\operatorname{Var}(Y) &= (\mathbb{E}[X^2]-\mathbb{E}[X]^2)(\mathbb{E}[Y^2]-\mathbb{E}[Y]^2)\\ &= \mathbb{E}[X^2]\,\mathbb{E}[Y^2]-\mathbb{E}[X]^2\,\mathbb{E}[Y^2]-\mathbb{E}[X^2]\,\mathbb{E}[Y]^2+\mathbb{E}[X]^2\,\mathbb{E}[Y]^2,\end{aligned}$$

which we may rewrite as

$$\operatorname{Var}(X)\operatorname{Var}(Y)+\mathbb{E}[X]^2\,\mathbb{E}[Y^2]+\mathbb{E}[X^2]\,\mathbb{E}[Y]^2-2\,\mathbb{E}[X]^2\,\mathbb{E}[Y]^2=\mathbb{E}[X^2]\,\mathbb{E}[Y^2]-\mathbb{E}[X]^2\,\mathbb{E}[Y]^2,$$

so we can transform the expression for $\operatorname{Var}(XY)$ as follows:

$$\begin{aligned}\operatorname{Var}(XY) &= \mathbb{E}[X^2]\,\mathbb{E}[Y^2]-\mathbb{E}[X]^2\,\mathbb{E}[Y]^2\\ &= \operatorname{Var}(X)\operatorname{Var}(Y)+\mathbb{E}[X]^2\,\mathbb{E}[Y^2]+\mathbb{E}[X^2]\,\mathbb{E}[Y]^2-2\,\mathbb{E}[X]^2\,\mathbb{E}[Y]^2\\ &= \operatorname{Var}(X)\operatorname{Var}(Y)+\mathbb{E}[X]^2(\mathbb{E}[Y^2]-\mathbb{E}[Y]^2)+\mathbb{E}[Y]^2(\mathbb{E}[X^2]-\mathbb{E}[X]^2)\\ &= \operatorname{Var}(X)\operatorname{Var}(Y)+\mathbb{E}[X]^2\operatorname{Var}(Y)+\mathbb{E}[Y]^2\operatorname{Var}(X),\end{aligned}$$

as desired.

3 Testing Model Planes

Amin is testing model airplanes. He starts with n model planes which each independently have probability p of flying successfully each time they are flown, where $0 < p < 1$. Each day, he flies every single plane and keeps the ones that fly successfully (i.e. don't crash), throwing away all other models. He repeats this process for many days, where each "day" consists of Amin flying any remaining model planes and throwing away any that crash. Let X_i be the random variable representing how many model planes remain after i days. Note that $X_0 = n$. Justify your answers for each part.

(a) What is the distribution of X_1? That is, what is $\mathbb{P}[X_1 = k]$?

(b) What is the distribution of X_2? That is, what is $\mathbb{P}[X_2 = k]$? Recognize the distribution of X_2 as one of the famous ones and provide its name and parameters.

(c) Repeat the previous part for X_t for arbitrary $t \geq 1$.

(d) What is the probability that at least one model plane still remains (has not crashed yet) after t days? Do not have any summations in your answer.

(e) Considering only the first day of flights, is the event A_1 that the first and second model planes crash independent from the event B_1 that the second and third model planes crash? Recall that two events A and B are independent if $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$. Prove your answer using this definition.

(f) Considering only the first day of flights, let A_2 be the event that the first model plane crashes *and* exactly two model planes crash in total. Let B_2 be the event that the second plane crashes on the first day. What must n be equal to in terms of p such that A_2 is independent from B_2? Prove your answer using the definition of independence stated in the previous part.

(g) Are the random variables X_i and X_j, where $i < j$, independent? Recall that two random variables X and Y are independent if $\mathbb{P}[X = k_1 \cap Y = k_2] = \mathbb{P}[X = k_1]\mathbb{P}[Y = k_2]$ for all k_1 and k_2. Prove your answer using this definition.

Solution:

(a) Since Amin is performing n trials (flying a plane), each with an independent probability of "success" (not crashing), we have $X_1 \sim \text{Binom}(n, p)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$, for $0 \leq k \leq n$.

(b) Each model plane independently has probability p^2 of surviving both days. Whether a model plane survives both days is still independent from whether any other model plane survives both days, so we can say $X_2 \sim \text{Binom}(n, p^2)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^{2k} (1-p^2)^{n-k}$, for $0 \leq k \leq n$.

(c) By extending the previous part we see each model plane has probability p^t of surviving t days, so $X_t \sim \text{Binom}(n, p^t)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^{tk} (1-p^t)^{n-k}$, for $0 \leq k \leq n$.

(d) We consider the complement, the probability that no model planes remain after t days. By the previous part we know this to be $\mathbb{P}[X_t = 0] = \binom{n}{0} p^{t(0)} (1-p^t)^{n-0} = (1-p^t)^n$. So the probability of at least model plane remaining after t days is $1-(1-p^t)^n$.

(e) No. $\mathbb{P}[A_1 \cap B_1]$ is the probability that the first three model planes crash, which is $(1-p)^3$. But $\mathbb{P}[A_1]\mathbb{P}[B_1] = (1-p)^2(1-p)^2 = (1-p)^4$. So $\mathbb{P}[A_1 \cap B_1] \neq \mathbb{P}[A_1]\mathbb{P}[B_1]$ and A_1 and B_1 are not independent.

(f) $\mathbb{P}[A_1 \cap B_1]$ is the probability that only the first model plane and second model plane crash, which is $(1-p)^2 p^{n-2}$. $\mathbb{P}[A_1]$ is the probability that the first model plane crashes, and exactly one of the remaining $n-1$ model planes crashes, so $\mathbb{P}[A_2] = (1-p) \cdot \binom{n-1}{1}(1-p)p^{n-1-1} = (n-1)(1-p)^2 p^{n-2}$. Trivially, we have $\mathbb{P}[B_2] = 1-p$, so $\mathbb{P}[A_2]\mathbb{P}[B_2] = (n-1)(1-p)^3 p^{n-2}$ which is equal to $\mathbb{P}[A_2 \cap B_2] = (1-p)^2 p^{n-2}$ only when $(n-1)(1-p) = 1$, or when $n = \frac{1}{1-p} + 1$.

(g) No. Let $k_1 = 0$ and $k_2 = 1$. Then $\mathbb{P}[X_i = k_1 \cap X_j = k_2] = 0$ because you can't have 1 plane at the end of day 2 if there are no planes left at the end of day 1. But $\mathbb{P}[X_i = k_1] > 0$ and $\mathbb{P}[X_j = k_2] > 0$ so $\mathbb{P}[X_i = k_1]\mathbb{P}[X_j = k_2] > 0$. Since $\mathbb{P}[X_i = k_1]\mathbb{P}[X_j = k_2] \neq \mathbb{P}[X_i = k_1 \cap X_j = k_2]$, they are not independent.

4 Unreliable Servers

In a single cluster of a Google competitor, there are a huge number of servers n, each with a uniform and independent probability of going down in a given day. On average, 4 servers go down in the cluster per day. As each cluster is responsible for a huge amount of internet traffic, it is fair to assume that n is a very large number. Recall that as $n \to \infty$, a $\text{Binom}(n, \lambda/n)$ distribution will tend towards a $\text{Poisson}(\lambda)$ distribution.

(a) What is an appropriate distribution to model the number of servers that crash on any given day for a certain cluster? What is its parameter?

(b) Compute the expected value and variance of the number of crashed servers on a given day for a certain cluster.

(c) Compute the probability that fewer than 3 servers crashed on a given day for a certain cluster.

(d) Compute the probability at least 3 servers crashed on a given day for a certain cluster.

Solution:

(a) Because each server goes down independently of the other servers, and with the same probability, the number of servers that crash on a given day follows a binomial distribution $\text{Binom}(n, p)$, where n is the number of servers and p is the probability of each individual server crashing on any given day. Since on average, 4 servers crash per day, we have $p = \frac{4}{n}$. We are given that the number of servers in the cluster is large, so $n \gg p$ and we can model the number of servers that crash as a Poisson distribution with $\lambda = 4$.

(b) Recall that the expectation and variance of a Poisson distribution with parameter λ are both equal to λ and in this case $\lambda = 4$.

(c) To compute the probability that fewer than 3 servers went down, we must add the probabilities that 0 servers go down, 1 server goes down, and the probability that 2 servers go down. The PMF of the Poisson distribution is

$$\mathbb{P}[X = i] = \frac{\lambda^i}{i!} e^{-\lambda}.$$

Thus

$$\mathbb{P}[X = 0 \text{ or } X = 1 \text{ or } X = 2] = e^{-4} + 4e^{-4} + \frac{4^2}{2} e^{-4} = e^{-4} + 4e^{-4} + 8e^{-4} = 13e^{-4}.$$

(d) $1 - \mathbb{P}[\text{fewer than 3 servers crashed}] = 1 - 13e^{-4}$.

5 Short Answer

(a) Let X be uniform on the interval $[0,2]$, and define $Y = 2X + 1$. Find the PDF, CDF, expectation, and variance of Y.

(b) Let X and Y have joint distribution

$$f(x,y) = \begin{cases} cxy + 1/4 & x \in [1,2] \text{ and } y \in [0,2] \\ 0 & \text{otherwise.} \end{cases}$$

Find the constant c. Are X and Y independent?

(c) Let $X \sim \text{Exp}(3)$.

(i) Find probability that $X \in [0, 1]$.

(ii) Let $Y = \lfloor X \rfloor$. For each $k \in \mathbb{N}$, what is the probability that $Y = k$? Write the distribution of Y in terms of one of the famous distributions; provide that distribution's name and parameters.

(d) Let $X_i \sim \text{Exp}(\lambda_i)$ for $i = 1, \ldots, n$ be mutually independent. It is a (very nice) fact that $\min(X_1, \ldots, X_n) \sim \text{Exp}(\mu)$. Find μ.

Solution:

(a) Let's begin with the CDF. It will first be useful to recall that

$$F_X(t) = \mathbb{P}(X \leq t) = \begin{cases} 0 & t \leq 0 \\ \frac{t}{2} & t \in [0,2] \\ 1 & t \geq 1 \end{cases}.$$

Since Y is defined in terms of X, we can compute that

$$\begin{aligned} F_Y(t) = \mathbb{P}(Y \leq t) &= \mathbb{P}[2X + 1 \leq t] \\ &= \mathbb{P}\left[X \leq \frac{t-1}{2}\right] \\ &= F_X\left(\frac{t-1}{2}\right) \\ &= \begin{cases} 0 & t \leq 1 \\ \frac{t-1}{4} & t \in [1,5] \\ 1 & t \geq 5 \end{cases} \end{aligned}$$

where in the third line we have used the PDF for X. We know that the PDF can be found by taking the derivative of the CDF, so

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \begin{cases} \frac{1}{4} & t \in [1,5] \\ 0 & \text{else} \end{cases}.$$

By linearity of expectation $\mathbb{E}[Y] = \mathbb{E}[2X+1] = 2\,\mathbb{E}[X] + 1 = 3$, and similarly

$$\text{Var}(Y) = \text{Var}(2X+1) = 4\,\text{Var}(X) = 4 \cdot \frac{4}{12} = \frac{4}{3}.$$

(b) To find the correct constant, we use the fact that a PDF must integrate to one. In particular,

$$1 = \int_1^2 \int_0^2 (cxy + 1/4)\, dy\, dx = 3c + 1/2,$$

so $c = 1/6$. In order to check independence, we need to first find the marginal distributions of X and Y:

$$\begin{aligned} f_X(x) &= \int_0^2 f(x,y)\, dy = 1/2 + x/3 \\ f_Y(y) &= \int_1^2 f(x,y)\, dx = 1/4 + y/4. \end{aligned}$$

Since $f_X(x)f_Y(y) = 1/8 + y/8 + x/12 + xy/12 \neq 1/4 + xy/6 = f(x,y)$, the random variables are not independent.

(c) (i) Since $X \sim \text{Exp}(3)$, the CDF of X is $F(x) = 1 - e^{-3x}$. Thus we have

$$\mathbb{P}[X \in [0,1]] = \int_0^1 f(x)\,dx = F(1) - F(0) = (1 - e^{-3}) - (1 - e^0) = 1 - e^{-3}.$$

(ii) Similarly, if $Y = \lfloor X \rfloor$, then $Y = k$ exactly when $X \in [k, k+1)$, so

$$\begin{aligned}\mathbb{P}[Y = k] &= \mathbb{P}[X \in [k, k+1)] \\ &= \int_k^{k+1} f(x)\,dx \\ &= F(k+1) - F(k) \\ &= (1 - e^{-3(k+1)}) - (1 - e^{-3k}) \\ &= e^{-3k} - e^{-3(k+1)} \\ &= e^{-3k}\left(1 - e^{-3}\right) = (e^{-3})^k\left(1 - e^{-3}\right).\end{aligned}$$

In other words, $Y = W - 1$ for $W \sim \text{Geometric}(1 - e^{-3})$.

(d) Since the X_i are independent,

$$\begin{aligned}\mathbb{P}[\min(X_1, \ldots, X_n) \leq t] &= 1 - \mathbb{P}[X_1 > t, X_2 > t, \cdot \ldots \cdot X_n > t] \\ &= 1 - \mathbb{P}[X_1 > t] \cdot \mathbb{P}[X_2 > t] \cdot \ldots \cdot \mathbb{P}[X_n > t] \qquad \text{by independence} \\ &= 1 - e^{-\lambda_1 t} e^{-\lambda_2 t} \cdot \ldots \cdot e^{-\lambda_n t} \\ &= 1 - e^{-(\lambda_1 + \lambda_2 + \ldots + \lambda_n)t}.\end{aligned}$$

This is exactly the CDF of an $\text{Exp}(\lambda_1 + \lambda_2 + \ldots + \lambda_n)$ random variable, so $\mu = \lambda_1 + \ldots + \lambda_n$.

6 Arrows

You and your friend are competing in an archery competition. You are a more skilled archer than he is, and the distances of your arrows to the center of the bullseye are i.i.d. Uniform$[0,1]$ whereas his are i.i.d. Uniform$[0,2]$. To even out the playing field, you both agree that you will shoot one arrow and he will shoot two. The arrow closest to the center of the bullseye wins the competition. What is the probability that you will win? *Note*: The distances *from the center of the bullseye* are uniform.

Solution:

Let X be the distance of your arrow to the center, Y_1, Y_2 be the distances of your friend's arrows from the center, and $Y = \min(Y_1, Y_2)$ be the distance of the closest to the center of your friend's arrows. Then, for $x \in [0,1]$ and $y \in [0,2]$,

$$\mathbb{P}[X > x] = 1 - x \quad \text{and} \quad \mathbb{P}[Y > y] = \mathbb{P}[Y_1 > y, Y_2 > y] = \left(1 - \frac{y}{2}\right)^2.$$

Hence,

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}(1 - \mathbb{P}[X > x]) = -\frac{\mathrm{d}}{\mathrm{d}x}(1 - x) = 1, \qquad x \in [0,1].$$

Also,

$$\mathbb{P}[Y > X \mid X = x] = \left(1 - \frac{x}{2}\right)^2.$$

Then, by total probability,

$$\begin{aligned}\mathbb{P}[Y > X] &= \int_0^1 \mathbb{P}(Y > X \mid X = x) f_X(x)\,\mathrm{d}x = \int_0^1 \left(1 - \frac{x}{2}\right)^2 f_X(x)\,\mathrm{d}x = \mathbb{E}\left[\left(1 - \frac{X}{2}\right)^2\right] \\ &= \mathbb{E}\left[1 - X + \frac{X^2}{4}\right] = 1 - \frac{1}{2} + \frac{1}{12} = \frac{7}{12},\end{aligned}$$

since $\mathbb{E}[X] = 1/2$ and $\mathbb{E}[X^2] = \mathrm{Var}[X] + \mathbb{E}[X]^2 = 1/3$.

7 Waiting For the Bus

Edward and Jerry are waiting at the bus stop outside of Soda Hall.

Like many bus systems, buses arrive in periodic intervals. However, the Berkeley bus system is unreliable, so the length of these intervals are random, and follow Exponential distributions.

Edward is waiting for the 51B, which arrives according to an Exponential distribution with parameter λ. That is, if we let the random variable X_i correspond to the difference between the arrival time ith and $i-1$st bus (also known as the inter-arrival time) of the 51B, $X_i \sim \mathrm{Expo}(\lambda)$.

Jerry is waiting for the 79, whose inter-arrival times also follows Exponential distributions with parameter μ. That is, if we let Y_i denote the inter-arrival time of the 79, $Y_i \sim \mathrm{Expo}(\mu)$. Assume that all inter-arrival times are independent.

(a) What is the probability that Jerry's bus arrives before Edward's bus?

(b) After 20 minutes, the 79 arrives, and Jerry rides the bus. However, the 51B still hasn't arrived yet. Let D be the additional amount of time Edward needs to wait for the 51B to arrive. What is the distribution of D?

(c) Lavanya isn't picky, so she will wait until either the 51B or the 79 bus arrives. Find the distribution of Z, the amount of time Lavanya will wait before catching her bus.

(d) Khalil doesn't feel like riding the bus with Edward. He decides that he will wait for the second arrival of the 51B to ride the bus. Find the distribution of $T = X_1 + X_2$, the amount of time that Khalil will wait to ride the bus.

Solution:

(a) Let $\{Y \in \mathrm{d}t\}$ be shorthand for $Y \in [t, t+\mathrm{d}t]$, the continuous version of the event that Y takes on value t. By total probability,

$$\begin{aligned}
\mathbb{P}(X_i > Y_i) &= \int_{t=0}^{\infty} \mathbb{P}(Y_i \in \mathrm{d}t \cap X_i > Y_i) \\
&= \int_{t=0}^{\infty} \mathbb{P}(Y_i \in \mathrm{d}t) \cdot \mathbb{P}(X_i > Y_i \mid Y_i \in \mathrm{d}t) \\
&= \int_{t=0}^{\infty} \mathbb{P}(Y_i \in \mathrm{d}t) \cdot \mathbb{P}(X_i > t) \\
&= \int_{t=0}^{\infty} f_{Y_i}(t) \cdot (1 - F_{X_i}(t))\,\mathrm{d}t \\
&= \int_{t=0}^{\infty} \mu e^{-\mu t}(e^{-\lambda t})\,\mathrm{d}t \\
&= \mu \int_{t=0}^{\infty} e^{-(\lambda+\mu)t}\,\mathrm{d}t \\
&= \frac{\mu}{\lambda+\mu} \int_{t=0}^{\infty} (\lambda+\mu)e^{-(\lambda+\mu)t}\,\mathrm{d}t \\
&= \frac{\mu}{\lambda+\mu},
\end{aligned}$$

where the integral in the second-to-last line evaluates to 1, since it is the total integral of the Exponential$(\lambda+\mu)$ density.

(b) We observe that $\mathbb{P}(D > d) = \mathbb{P}(X > 20 + d \mid X \geq 20)$. Then, we apply Bayes Rule:

$$\begin{aligned}
\mathbb{P}(X > 20 + d \mid X \geq 20) &= \frac{\mathbb{P}(X > 20 + d)}{\mathbb{P}(X \geq 20)} \\
&= \frac{1 - F_X(20 + d)}{1 - F_X(20)} \\
&= \frac{e^{-\lambda(20+d)}}{e^{-20\lambda}} \\
&= e^{-\lambda d}
\end{aligned}$$

Thus, the CDF of D is given by $\mathbb{P}(D \leq d) = 1 - \mathbb{P}(D > d) = 1 - e^{-\lambda d}$. This is the CDF of an exponential, so D is exponentially distributed with parameter λ.

One can also directly apply the memoryless property of the exponential distribution to arrive at this answer.

(c) Lavanya's waiting time is the minimum of the time it takes for the 51B and the time it takes for the 79 to arrive. Thus, $Z = \min(X,Y)$.

$$\begin{aligned}
\mathbb{P}(Z > t) &= \mathbb{P}(X > t \cap Y > t) \\
&= \mathbb{P}(X > t) \cdot \mathbb{P}(Y > t) \\
&= (1 - F_X(t))(1 - F_Y(t)) \\
&= (1 - (1 - e^{-\mu t}))(1 - (1 - e^{-\lambda t})) \\
&= e^{-\mu t} e^{-\lambda t} \\
&= e^{-(\mu+\lambda)t}
\end{aligned}$$

It follows that the CDF is Z, $\mathbb{P}(Z \le t) = 1 - e^{-(\mu+\lambda)t}$. Thus, Z is exponentially distributed with parameter $\mu + \lambda$.

(d) Let $t > 0$. By total probability,

$$\begin{aligned}
\mathbb{P}(T \le t) &= \mathbb{P}(X_1 + X_2 \le t) \\
&= \int_0^\infty \mathbb{P}(X_1 + X_2 \le t \mid X_1 \in \mathrm{d}x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) \\
&= \int_0^t \mathbb{P}(X_1 + X_2 \le t \mid X_1 \in \mathrm{d}x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) + \int_t^\infty 0 \cdot \mathbb{P}(X_1 \in \mathrm{d}x) \\
&= \int_0^t \mathbb{P}(X_2 \le t - X_1 \mid X_1 \in \mathrm{d}x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) + 0 \\
&= \int_0^t \mathbb{P}(X_2 \le t - x)) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) \\
&= \int_0^t F_{X_2}(t - x) \cdot f_{X_1}(x)\,\mathrm{d}x \\
&= \int_0^t \left(1 - e^{-\lambda(t-x)}\right) \cdot \lambda e^{-\lambda x}\,\mathrm{d}x \\
&= \int_0^t \lambda e^{-\lambda x} - \lambda e^{-\lambda t}\,\mathrm{d}x \\
&= \int_0^t \lambda e^{-\lambda x} - \lambda e^{-\lambda t} \int_0^t \mathrm{d}x \\
&= F_{X_1}(t) - \lambda e^{-\lambda t} \cdot t \\
&= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}
\end{aligned}$$

Upon differentiating the CDF, we have

$$\begin{aligned}
f_T(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(T \le t) &= \lambda e^{-\lambda t} - \lambda e^{-\lambda t} + \lambda^2 t e^{-\lambda t} \\
&= \lambda^2 t e^{-\lambda t}, \qquad t > 0.
\end{aligned}$$

8 Variance of the Minimum of Uniform Random Variables

Let n be a positive integer and let $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}[0,1]$. Find $\text{Var}[Y]$, where

$$Y := \min\{X_1, \ldots, X_n\}.$$

Hint: You may need to perform integration by parts.

Solution:

We know that the density of Y is $f(y) = n(1-y)^{n-1}$, for $y \in [0,1]$, and $\mathbb{E}[Y] = (n+1)^{-1}$. It remains to compute (via integration by parts)

$$\begin{aligned}
\mathbb{E}[Y^2] &= \int_0^1 y^2 \cdot n(1-y)^{n-1}\,\mathrm{d}y \\
&= n\int_0^1 y^2(1-y)^{n-1}\,\mathrm{d}y \\
&= -y^2(1-y)^n\Big|_0^1 + 2\int_0^1 y(1-y)^n\,\mathrm{d}y \\
&= \frac{2}{n+1}\int_0^1 y(n+1)(1-y)^n\,\mathrm{d}y.
\end{aligned}$$

Since $g(y) := (n+1)(1-y)^n$ is the density of the minimum of $n+1$ i.i.d. Uniform$[0,1]$ random variables, we recognize the last integral as the expectation of this minimum, which is $1/(n+2)$. Thus,

$$\mathbb{E}[Y^2] = \frac{2}{(n+1)(n+2)}$$

and so

$$\begin{aligned}
\mathrm{Var}[Y] &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\
&= \frac{2}{(n+1)(n+2)} - \frac{1}{(n+1)^2} \\
&= \frac{2(n+1)-(n+2)}{(n+1)^2(n+2)} \\
&= \frac{n}{(n+1)^2(n+2)}.
\end{aligned}$$

Fun Fact: For a non-negative random variable X with density f_X, one can extend the tail sum formula to give

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^\infty x^2 f_X(x)\,\mathrm{d}x \\
&= \int_0^\infty \left(\int_0^x 2s\,\mathrm{d}s\right) f_X(x)\,\mathrm{d}x \\
&= \int_0^\infty 2s \int_s^\infty f_X(x)\,\mathrm{d}x\,\mathrm{d}s \\
&= \int_0^\infty 2s\mathbb{P}(X \geq s)\,\mathrm{d}s
\end{aligned}$$

and this gives another way to compute $\mathbb{E}[Y^2]$ in this problem. You can derive a similar formula to compute any moment $\mathbb{E}[X^k]$ for $k \in \mathbb{N}$.

Another Fun Fact: To avoid computing integrals, we can take an advanced approach through probability. We may first note that the expression inside the integral for $\mathbb{E}[Y^2]$, $y^2(1-y)^{n-1}$, resembles the density of the third smallest out of $n+2$ uniform random variables. That is, for W the third smallest out of $n+2$ iid uniform random variables,

$$\mathbb{P}(W \in \mathrm{d}y) = f_W(y)\mathrm{d}y = \frac{(n+2)!}{2!(n-1)!} y^2 \cdot 1\mathrm{d}y \cdot (1-y)^{n-1}.$$

Therefore

$$\begin{aligned}
\mathbb{E}[Y^2] &= n \int_0^1 y^2 (1-y)^{n-1} \mathrm{d}y \\
&= n \frac{2!(n-1)!}{(n+2)!} \int_0^1 \frac{(n+2)!}{2!(n-1)!} \cdot y^2 (1-y)^{n-1} \, \mathrm{d}y \\
&= \frac{2}{(n+2)(n+1)}.
\end{aligned}$$

where the integral in the second-to-last line is 1 since it is the total integral over the density of W.