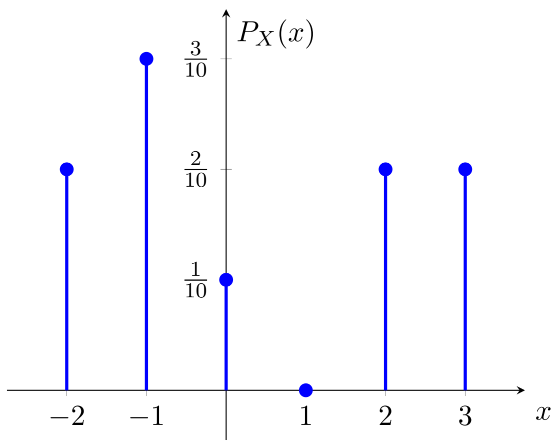
EK381 Probability, Statistics, and Data Science for Engineers — Boston University
Prof Carruthers — Summer 2025

Homework 3: finish by 5/30.

Reading: Notes: Chapter 2. **Videos:** 2.3 - 2.6

Problem 3.1 (**Video 2.3 - 2.6**, **Lecture Problem**)



Let X be a discrete random variable with probability mass function (PMF) as above. Let event $A = \{-2, 1, 3\}$.

(a) Given that $\{X \in A\}$ occurs, what is the conditional probability that $X > 1$, that is $\mathbb{P}[X > 1|X \in A]$?

Solution:

Let C be the values in the range of X such that $X > 1$. We want $\mathbb{P}[X \in C|X \in A]$.

Note first that $\mathbb{P}[\{x \in A\}] = \sum_{x\in A} P_X(x) = P_X(-2) + P_X(1) + P_X(3) = \frac{2}{10} + 0 + \frac{2}{10} = \frac{2}{5}$

By the definition of conditional probability,

$$\mathbb{P}[X \in C|X \in A] = \frac{\mathbb{P}[X \in C \cap A]}{\mathbb{P}[X \in A]}.$$

Now, note that $C \cap A = \{3\}$ and

$$\mathbb{P}[X \in C \cap A] = P_X(3) = \frac{2}{10} = \frac{1}{5}\ .$$

Combining this with our previous), we get

$$\mathbb{P}[\{X > 1\} | X \in A] = \frac{1/5}{2/5} = \frac{1}{2} \, .$$

(b) Determine $\mathbb{E}[X]$ and $\mathbb{E}[3X+2]$.

Solution:

The expected value is

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in R_X} x P_X(x) \\ &= (-2) \cdot \frac{2}{10} + (-1) \cdot \frac{3}{10} + 0 \cdot \frac{1}{10} + 1 \cdot 0 + 2 \cdot \frac{2}{10} + 3 \cdot \frac{2}{10} = \frac{3}{10} \, .\end{aligned}$$

By the linearity of expectation, $\mathbb{E}[3X+2] = 3\,\mathbb{E}[X] + 2 = 3 \cdot \frac{3}{10} + 2 = \frac{29}{10}$.

(c) Determine $\mathsf{Var}[X]$ and $\mathsf{Var}[2X-1]$.

Solution:

We will use the alternate variance formula, $\mathsf{Var}[X] = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2$. First, we need to calculate the second moment,

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{x \in R_X} x^2 P_X(x) \\ &= (-2)^2 \cdot \frac{2}{10} + (-1)^2 \cdot \frac{3}{10} + 0^2 \cdot \frac{1}{10} + 1^2 \cdot 0 + 2^2 \cdot \frac{2}{10} + 3^2 \cdot \frac{2}{10} = \frac{37}{10} \, ,\end{aligned}$$

and then subtract the square of the mean to obtain the variance

$$\mathsf{Var}[X] = \frac{37}{10} - \left(\frac{3}{10}\right)^2 = \frac{361}{100} \, .$$

We can use the identity $\mathsf{Var}[aX+b] = a^2 \mathsf{Var}[X]$ for the variance of a linear function to obtain

$$\mathsf{Var}[2X-1] = 2^2 \mathsf{Var}[X] = \frac{361}{25}$$

Problem 3.2 (**Video 2.3 - 2.6**, **Quick Calculations**)

Calculate each of the requested quantities.

(a) Your favorite sports team wins a game with probability $\frac{3}{5}$, independently of other games. Let X be the number of games they win out of 20. What kind of random variable is X? (Don't forget the parameters.) Calculate $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$.

Solution:

X is a Binomial$(20, \frac{3}{5})$ random variable. Therefore, we have that $\mathbb{E}[X] = 20 \cdot \frac{3}{5} = 12$ and $\mathsf{Var}[X] = 20 \cdot \frac{3}{5} \cdot \frac{2}{5} = \frac{120}{25} = \frac{24}{5}$. Using the alternate variance formula, we can solve for $\mathbb{E}[X^2] = \mathsf{Var}[X] + \big(\mathbb{E}[X]\big)^2 = \frac{24}{5} + (12)^2 = \frac{744}{5} = 148.8$.

(b) Let X be Poisson(λ) and assume that $\mathbb{E}[X] = 2$.
Calculate λ, and $\mathbb{P}[X \le 3]$ and $\mathbb{P}[X \le 3 | X > 0]$.

Solution:

Since X is a Poisson(λ) random variable, we know that $\mathbb{E}[X] = \lambda$ and therefore $\lambda = 2$. We can calculate $\mathbb{P}[X \le 3]$ as

$$\begin{aligned}
\mathbb{P}[X \le 3] &= \big(P_X(0) + P_X(1) + P_X(2) + P_X(3)\big) \\
&= e^{-2}\left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}\right) \\
&= e^{-2}\left(\frac{1}{1} + \frac{2}{1} + \frac{4}{2} + \frac{8}{6}\right) \\
&= \frac{19}{3}e^{-2}
\end{aligned}$$

To calculate $\mathbb{P}[X \le 3 | X > 0]$ we use the definition of conditional probability:

$$\begin{aligned}
\mathbb{P}[X \le 3 | X > 0] &= \frac{\mathbb{P}[\{X \le 3\} \cap \{X > 0\}]}{\mathbb{P}[\{X > 0\}]} = \frac{\mathbb{P}[X \in \{1,2,3\}]}{\mathbb{P}[X > 0]} \\
\mathbb{P}[X \in \{1,2,3\} &= e^{-2}\left(\frac{2}{1} + \frac{4}{2} + \frac{8}{6}\right) = \frac{16}{3}e^{-2} \\
\mathbb{P}[X > 0] &= 1 - P_X(0) = 1 - e^{-2}\frac{2^0}{0!} = 1 - e^{-2} \\
\mathbb{P}[X \le 3 | X > 0] &= \frac{\frac{16}{3}e^{-2}}{1 - e^{-2}}
\end{aligned}$$

(c) Roll a six-sided die until the first 2 appears. Let X denote the number of rolls. What kind of a random variable is X? (Don't forget the parameters.) Calculate $\mathbb{E}[2X - 1]$ and $\mathsf{Var}[2X - 1]$.

Solution:

This is Geometric$(\frac{1}{6})$ random variable. We know that $\mathbb{E}[X] = \frac{1}{\frac{1}{6}} = 6$ and $\mathsf{Var}[X] = \frac{1-\frac{1}{6}}{\left(\frac{1}{6}\right)^2} = 30$. From the linearity of expectation, $\mathbb{E}[2X-1] = 2\mathbb{E}[X]-1 = 2\cdot 6-1 = 11$. We also know that the variance of a linear function is $\mathsf{Var}[2X-1] = 2^2\mathsf{Var}[X] = 4\cdot 30 = 120$.

(d) Let X be a random variable with $\mathbb{E}[X] = -1$ and $\mathsf{Var}[X] = 4$. Let $Y = -2X + 3$.
Calculate $\mathbb{E}[Y]$ and $\mathsf{Var}[Y]$.

Solution:

By the linearity of expectation, $\mathbb{E}[Y] = \mathbb{E}[-2X+3] = -2\mathbb{E}[X]+3 = (-2)\cdot(-1)+3 = 5$.
By the variance of a linear function, $\mathsf{Var}[Y] = \mathsf{Var}[-2X+3] = (-2)^2\mathsf{Var}[X] = 4\cdot 4 = 16$.

(e) Let X be a random variable with $\mathbb{E}[X] = 0$ and $\mathsf{Var}[X] = 2$.
Calculate $\mathbb{E}[X^2]$ and $\mathbb{E}[(2X-1)^2]$.

Solution:

Using the alternate variance formula, $\mathbb{E}[X^2] = \mathsf{Var}[X] + \big(\mathbb{E}[X]\big)^2 = 2 + 0^2 = 2$.
Using the linearity of expectation,
$\mathbb{E}[(2X-1)^2] = \mathbb{E}[4X^2 - 4X + 1] = 4\mathbb{E}[X^2] - 4\mathbb{E}[X] + 1 = 4\cdot 2 - 4\cdot 0 + 1 = 9$.

Problem 3.3 (**Video 2.5, 2.6**, **Lecture Problem**, **Spring 2022 Exam 1 Problem**)

You are measuring the number of spikes from a neuron in a one-second window. The resulting random variable X is Poisson(λ).

(a) After careful study, you have determined that the average number of spikes observed from this neuron in one second is $\mathbb{E}[X] = 2$. What is the probability that you see no spikes at all in a one-second window?

Solution:

For a Poisson(λ) random variable, we know that $\mathbb{E}[X] = \lambda$. Therefore, $\lambda = 2$ and we have that $\mathbb{P}[X = 0] = P_X(0) = \frac{2^0}{0!}e^{-2} = e^{-2}$.

(b) What is the probability that the number of spikes you see in a one-second window is *less than or equal* to average? (Recall from (a) that the average is 2.)

Solution:

$$P[X \leq 2] = P_X(0) + P_X(1) + P_X(2) = \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!}\right)e^{-2} = 5e^{-2}$$

(c) Calculate $\mathbb{E}[3X^2 + 2X + 1]$.

Solution:

Using the alternate variance formula, we know that $\mathbb{E}[X^2] = \mathsf{Var}[X] + \big(\mathbb{E}[X]\big)^2 = \lambda + \lambda^2 = 2 + 2^2 = 6$. By linearity of expectation, we have

$$\mathbb{E}[3X^2 + 2X + 1] = 3\mathbb{E}[X^2] + 2\mathbb{E}[X] + 1 = 3\cdot 6 + 2\cdot 2 + 1 = 23$$

(d) Given that the number of spikes in a one-second window is *less than or equal* to average, what is the conditional expected value of X?

Solution:

$$P_{X|B}(x) = \begin{cases} \dfrac{P_X(x)}{\sum_{x\in B} P_X(x)} & x \in B \\ 0 & x \notin B \end{cases} = \begin{cases} \frac{e^{-2}}{5e^{-2}} & x = 0 \\ \frac{2e^{-2}}{5e^{-2}} & x = 1 \\ \frac{2e^{-2}}{5e^{-2}} & x = 2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{1}{5} & x = 0 \\ \frac{2}{5} & x = 1, 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}[X|B] = \sum_{x\in B} x\, P_{X|B}(x) = 0 \cdot \frac{1}{5} + 1 \cdot \frac{2}{5} + 2 \cdot \frac{2}{5} = \frac{6}{5}$$

Problem 3.4 (**Lecture Problem**, **Video 2.5, 2.6**) You have started watching Game of Thrones (or House of the Dragon), and from the very beginning realize that a particular character is your favorite. However, you are aware of the show's reputation for killing off characters, and would like to calculate the probability your favorite is eliminated after a certain number of episodes. Specifically, you use the following model: for each episode, the probability that your character is eliminated is $1/3$, independently of all other episodes. Let X be the episode number where your character is eliminated.

(a) What kind of random variable is X? (Don't forget the parameters.)

Solution:

We can view the episodes as independent trials where a "success" is that your character is eliminated. Therefore, X is the number of trials until the first success, which is a Geometric($1/3$) random variable.

(b) What is the probability that your character is eliminated in the third episode?

Solution:

$$\mathbb{P}[X = 3] = P_X(3) = \frac{1}{3}\left(\frac{2}{3}\right)^{3-1} = \frac{4}{27}$$

(c) What is the probability that your character lasts at least two episodes?

Solution:

Note that using the complement property is easier than trying to calculate the result directly, which corresponds to the sum of a geometric series.

$$\mathbb{P}[X > 2] = 1 - \mathbb{P}[X \leq 2] = 1 - \big(P_X(1) + P_X(2)\big) = 1 - \left(\frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3}\right) = 1 - \frac{5}{9} = \frac{4}{9}$$

(d) Let B be the event that your character lasts at least two episodes. Determine the conditional PMF of X given event B.

Solution:

First, recall that the conditional PMF is defined as

$$P_{X|B}(x) = \begin{cases} \dfrac{P_X(x)}{\mathbb{P}[X \in B]} & x \in B \\ 0, & \text{otherwise.} \end{cases}$$

We already know that

$$P_X(x) = \begin{cases} \dfrac{1}{3}\left(\dfrac{2}{3}\right)^{x-1} & x = 1, 2, \ldots \\ 0, & \text{otherwise.} \end{cases}$$

for a Geometric(1/3) random variable. We also know that $\mathbb{P}[X \in B] = \mathbb{P}[X > 2] = \frac{4}{9}$ from part (b). Therefore,

$$P_{X|B}(x) = \begin{cases} \dfrac{3}{4}\left(\dfrac{2}{3}\right)^{x-1} & x = 3, 4, \ldots \\ 0, & \text{otherwise.} \end{cases}$$

(e) Given that your character lasts at least two episodes, what is the probability that they are eliminated in the third episode?

Solution:

$$\mathbb{P}[X = 3|X > 2] = P_{X|B}(3) = \frac{3}{4}\left(\frac{2}{3}\right)^{3-1} = \frac{1}{3}$$

(f) Given that your character lasts at least two episodes, what is the probability that they last four or more episodes?

Solution:

This is easier to handle via the complement property.

$$\begin{aligned} \mathbb{P}[X > 4|X > 2] &= 1 - \mathbb{P}[X \leq 4|X > 2] \\ &= 1 - \left(P_{X|B}(3) + P_{X|B}(4)\right) \\ &= 1 - \left(\frac{3}{4}\left(\frac{2}{3}\right)^{3-1} + \frac{3}{4}\left(\frac{2}{3}\right)^{4-1}\right) \\ &= 1 - \left(\frac{1}{3} + \frac{2}{9}\right) = \frac{4}{9} \end{aligned}$$

Notice that this is the same value as we got for $\mathbb{P}[X > 2]$ in part (c). This is due to the "memoryless" property of the geometric distribution. Specifically, one can show that, if X is Geometric(p), the conditional PMF $P_{Y|B}(y)$ for $Y = X - c$ and event $B = \{c + 1, c + 2, \ldots\}$ itself corresponds to a Geometric(p) distribution. Intuitively, knowing that the success has not occurred yet does not help us predict when it will happen in the future, due to the independence of trials.

Problem 3.5 (Video 2.5, 2.6, **Fall 2020 Exam 1 Problem**)

You are practicing your free throws for an upcoming basketball game. Every throw is successful with probability $2/3$, independently of the others. Let X denote the number of successful throws out of 5.

(a) What kind of random variable is X? (Don't forget the parameters.)

Solution:

X is Binomial$(5, 2/3)$.

(b) What is the probability that you successfully make at least 3 out of the 5 free throws?

Solution:

$$\begin{aligned}\mathbb{P}[\{X \geq 3\}] &= P_X(3) + P_X(4) + P_X(5) \\ &= \binom{5}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2 + \binom{5}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^1 + \binom{5}{5}\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^0 \\ &= \frac{80+80+32}{3^5} = \frac{192}{3^5} = \frac{64}{81}\end{aligned}$$

(c) Given that you successfully make at least 3 out of 5 free throws, what is the probability that you successfully make exactly 3 out of 5?

Solution:

From part (b), we know that $\mathbb{P}[X = 3] = P_X(3) = \frac{80}{243}$. It follows that the conditional probability of interest is

$$\mathbb{P}[X = 3 | X \geq 3] = \frac{\mathbb{P}[\{X = 3\} \cap \{X \geq 3\}]}{\mathbb{P}[X \geq 3]} = \frac{\mathbb{P}[X = 3]}{\mathbb{P}[X \geq 3]} = \frac{80/243}{192/243} = \frac{80}{192} = \frac{5}{12}.$$

(d) What is the probability of scoring exactly 3 **consecutive** free throws within the set of 5?

Solution:

Let T_i be the event that the i^{th} throw is successful. There are three ways of making 3 consecutive free throws out of the 5 attempts:

$$T_1 \cap T_2 \cap T_3 \cap T_4^{\mathsf{c}} \cap T_5^{\mathsf{c}}, \qquad T_1^{\mathsf{c}} \cap T_2 \cap T_3 \cap T_4 \cap T_5^{\mathsf{c}}, \qquad T_1^{\mathsf{c}} \cap T_2^{\mathsf{c}} \cap T_3 \cap T_4 \cap T_5\ .$$

By independence, we have that

$$\begin{aligned}\mathbb{P}[T_1 \cap T_2 \cap T_3 \cap T_4^{\mathsf{c}} \cap T_5^{\mathsf{c}}] &= \mathbb{P}[T_1]\,\mathbb{P}[T_2]\,\mathbb{P}[T_3]\,\mathbb{P}[T_4^{\mathsf{c}}]\,\mathbb{P}[T_5^{\mathsf{c}}] \\ &= \frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3} = \frac{8}{243}\ .\end{aligned}$$

Similarly, we have that

$$\mathbb{P}[T_1^{\mathsf{c}} \cap T_2 \cap T_3 \cap T_4 \cap T_5^{\mathsf{c}}] = \frac{8}{243}$$
$$\mathbb{P}[T_1^{\mathsf{c}} \cap T_2^{\mathsf{c}} \cap T_3 \cap T_4 \cap T_5] = \frac{8}{243}\ .$$

By the additivity axiom, $\mathbb{P}[\{3 \text{ consecutive free throws made}\}] = 3 \cdot \frac{8}{243} = \frac{8}{81}$.

(e) You keep practicing with sets of 5 free throws. What is the average number of sets until your first set where you miss every free throw?

Solution:

The event $T_1^{\mathsf{c}} \cap T_2^{\mathsf{c}} \cap T_3^{\mathsf{c}} \cap T_4^{\mathsf{c}} \cap T_5^{\mathsf{c}}$ that you miss every free throw in a set in a set of has probability $\frac{1}{3^5} = \frac{1}{243}$.

The number of sets is a Geometric($\frac{1}{243}$) random variable. The average of this random variable is 243.