Probability & Statistics Workbook Solutions

Discrete random variables

krista king MATH

DISCRETE PROBABILITY

1. Let X be a discrete random variable with the following probability distribution. Find $P(X \geq 3)$.

X	1	2	3	4	5
P(X)	0.35	0.25	0.20	0.15	?

Solution:

First, we need to find the $P(X = 5)$, which we'll do by subtracting all the other probabilities from 1.

$$P(X = 5) = 1 - 0.35 - 0.25 - 0.20 - 0.15$$

$$P(X = 5) = 1 - 0.95$$

$$P(X = 5) = 0.05$$

Then the probability that $X \geq 3$ is

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5)$$

$$P(X \geq 3) = 0.20 + 0.15 + 0.05$$

$$P(X \geq 3) = 0.40$$

2. Let B be a discrete random variable with the following probability distribution. Find μ_B and σ_B.

B	0	5	10	15
P(B)	1/5	1/5	2/5	1/5

Solution:

We'll weight each value of B by the probability that the value occurs, $P(B)$, in order to find the expected value μ_B.

$$\mu_B = E(B) = \sum_{i=1}^{4} B_i P(B_i) = 0\left(\frac{1}{5}\right) + 5\left(\frac{1}{5}\right) + 10\left(\frac{2}{5}\right) + 15\left(\frac{1}{5}\right)$$

$$\mu_B = 8$$

In order to find the standard deviation of B, σ_B, we have to find variance first.

$$\sigma_B^2 = \sum_{i=1}^{4} (B_i - \mu_B)^2 P(B_i)$$

$$\sigma_B^2 = (0-8)^2\left(\frac{1}{5}\right) + (5-8)^2\left(\frac{1}{5}\right) + (10-8)^2\left(\frac{2}{5}\right) + (15-8)^2\left(\frac{1}{5}\right)$$

$$\sigma_B^2 = 26$$

Then the standard deviation is

$$\sqrt{\sigma_B^2} = \sqrt{26}$$

$$\sigma_B \approx 5.099$$

■ 3. The table shows the distribution of size of households in the U.S. for 2016. Suppose we select a household of size at least 2 at random. What is the probability that this household has a size of at least 4?

Size of household	1	2	3	4	5	6	7+
P(size)	0.281	0.340	**?**	0.129	0.060	0.023	0.013

Solution:

Find the probability that the household is a 3-person household.

$$P(X = 3) = 1 - 0.281 - 0.340 - 0.129 - 0.060 - 0.023 - 0.013$$

$$P(X = 3) = 0.154$$

The probability of "at least 4" is

$$P(\text{at least } 4) = P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)$$

$$P(\text{at least } 4) = 0.129 + 0.060 + 0.023 + 0.013$$

$$P(\text{at least } 4) = 0.225$$

and the probability of "at least 2" is

$$P = (\text{at least } 2) = 1 - P(X = 1)$$

$$P = (\text{at least } 2) = 1 - 0.281$$

$$P = (\text{at least } 2) = 0.719$$

Then the probability that the household size is at least 4, given that the household size is at least 2, is

$$P(\text{at least } 4\text{, given at least } 2) = \frac{P(\text{size at least } 4)}{P(\text{size at least } 2)}$$

$$P(\text{at least } 4\text{, given at least } 2) = \frac{0.225}{0.719} \approx 0.313$$

4. A standard deck of cards is shuffled, and two cards are selected without replacement. Let R be the number of red cards selected. Construct a probability distribution for R.

Solution:

If we draw two cards, we can find the probability that either both are red $P(R = 2)$, that one is red $P(R = 1)$, or that neither are red $P(R = 0)$.

$$P(R = 0) = \frac{26}{52}\left(\frac{25}{51}\right) = \frac{25}{102}$$

$$P(R = 2) = \frac{26}{52}\left(\frac{25}{51}\right) = \frac{25}{102}$$

Then the probability that one card is red is

$$P(R = 1) = 1 - P(R = 2) - P(R = 0)$$

$$P(R = 1) = 1 - \frac{25}{102} - \frac{25}{102}$$

$$P(R = 1) = \frac{102}{102} - \frac{25}{102} - \frac{25}{102}$$

$$P(R = 1) = \frac{52}{102} = \frac{26}{51}$$

Which means we can build a probability distribution for R.

R	0	1	2
P(R)	25/102	52/102	25/102

5. A local restaurant features a wheel we can spin before paying the bill. The wheel is split into 8 equal size pieces. One of the sections gives us a $\$10$ discount on the bill, two sections give a $\$5$ discount, three sections give a $\$2$ discount, and the rest of the sections give no discount. Find the expected value for the discount given by the wheel.

Solution:

Let X be the amount of the discount. Then the expected value, or mean of the discount is

$$E(X) = \sum XP(X) = 10\left(\frac{1}{8}\right) + 5\left(\frac{2}{8}\right) + 2\left(\frac{3}{8}\right) + 0\left(\frac{2}{8}\right)$$

$$E(X) = \$3.25$$

6. John stops at the local gas station and decides to buy lottery tickets. Each ticket has a 20% chance of being a winner. He will buy a lottery ticket and check to see if it's a winner. If it's a winner, he'll collect his money and be done. If it's not a winner, he'll buy another. He'll repeat this until he gets a winning ticket. But if he hasn't won by his fifth ticket, he won't buy any more tickets. Let L be the number of lottery tickets John will buy, then find $E(L)$.

Solution:

We could find the probability of winning on each of the first four tickets.

$$P(L=1) = 0.2$$

$$P(L=2) = (0.2)(0.8) = 0.16$$

$$P(L=3) = (0.2)(0.8)(0.8) = 0.128$$

$$P(L=4) = (0.2)(0.8)(0.8)(0.8) = 0.1024$$

If we continue this pattern, we might think that the probability of winning on the fifth ticket would be

$$P(L=5) = (0.2)(0.8)(0.8)(0.8)(0.8) = 0.08192$$

But the question tells us that John will never buy more than five tickets. Because he's guaranteed to buy one, two, three, four, or five tickets, the probability that he's going to purchase one of those numbers of tickets must be 100%. So the probability that he purchases five tickets is actually

$$P(L = 5) = 1 - P(L \leq 4)$$

$$P(L = 5) = 1 - (0.2 + 0.16 + 0.128 + 0.1024)$$

$$P(L = 5) = 1 - 0.5904$$

$$P(L = 5) = 0.4096$$

Then the expected value for the number of tickets he'll buy, L, is

$$E(L) = 1(0.2) + 2(0.16) + 3(0.128) + 4(0.1024) + 5(0.4096)$$

$$E(L) = 0.2 + 0.32 + 0.384 + 0.4096 + 2.048$$

$$E(L) = 3.3616$$

TRANSFORMING RANDOM VARIABLES

1. We use the formula

$$^\circ F=\frac{9}{5}{}^\circ C+32$$

to convert from Celsius to Fahrenheit. August is the hottest month in Hawaii with a mean temperature of $27^\circ C$. What is the mean temperature in Hawaii in $^\circ F$.

Solution:

We'll plug $27^\circ C$ into the conversion formula to get the corresponding value in Fahrenheit.

$$\mu_{^\circ F}=\frac{9}{5}\mu_{^\circ C}+32=\frac{9}{5}(27)+32=\frac{243}{5}+32=80.6^\circ$$

2. Let Z be a random variable with $\sigma_Z^2=49$. Let $W=(1/2)Z-10$. Find σ_W.

Solution:

We've been given the variance of Z, so we need to use it first to find the standard deviation of Z.

$$\sqrt{\sigma_Z^2} = \sqrt{49}$$

$$\sigma_Z = 7$$

Standard deviation is effected by scaling, but not by shifting. So when we convert from Z to W using

$$W = \frac{1}{2}Z - 10$$

we need to multiply by $1/2$, but we don't need to shift by 10. So we can say that the standard deviation σ_W is

$$\sigma_W = \frac{1}{2}\sigma_Z$$

$$\sigma_W = \frac{1}{2}(7)$$

$$\sigma_W = \frac{7}{2}$$

■ 3. The students in each 8th period classroom were asked to donate money for a school fundraiser, and the class that raised the most money was awarded a pizza party. The school secretary recorded the amount raised by each class and made a five-number summary for the data.

Min	Q1	Median	Q3	Max
4.50	15.25	22.00	38.75	95.50

If a donor commits to matching equally the students' donations, create a new five-number summary of the total amount raised, including the donor's contribution.

Solution:

The donor is essentially scaling the data set, because she's doubling the students' donations. And we know that scaling a data set scales all the values in the five-number summary.

Therefore, after putting the donor's contribution together with the students' donations, we can give a the new five-number summary for the data set as

Min	Q1	Median	Q3	Max
4.5(2)	15.25(2)	22(2)	38.75(2)	95.5(2)

Min	Q1	Median	Q3	Max
9	30.5	44	77.5	191

■ 4. The number of items sold at a concession stand is normally distributed with $\mu = 323$ and $\sigma = 30$. The average price per item sold is $\$1.25$. Different student clubs volunteer to work the concession stand throughout the year and get to keep half of their sales to go towards their club's activities. What is the probability that a club will get to keep more than $\$220$ in sales?

Solution:

Let N be the number of items sold and S be the amount of money the club gets to keep. Then we could write an equation for the amount of money they keep as

$$S = \frac{1}{2}(1.25)(N)$$

We know the mean of N is $\mu_N = 323$, so we can use the conversion equation to find the mean of S.

$$\mu_S = \frac{1}{2}(1.25)(\mu_N) = \frac{1}{2}(1.25)(323) \approx 201.88$$

Using $\sigma_N = 30$, we can find the standard deviation of S in the same way.

$$\sigma_S = \frac{1}{2}(1.25)(\sigma_N) = \frac{1}{2}(1.25)(30) = 18.75$$

With a mean of $\mu_S = 201.88$ and a standard deviation of $\sigma_S = 18.75$, we can find the probability that the club will take home more than $\$220$. We'll find the z-score associated with $\$220$.

$$z = \frac{220 - 201.88}{18.75} = 0.9664 \approx 0.97$$

Since we're looking at 0.97 standard deviations above the mean, we get 0.8340 from the z-table, which tells us that the probability that the club takes home more than $\$220$ is $1 - 0.8340 = 0.166 \approx 17\,\%$. They have an approximately $17\,\%$ chance of taking home more than $\$220$.

5. The average length of a full-term new born baby is 20 inches with variance 0.81 inches. What are the mean and standard deviation of the length of a full-term new born, expressed in centimeters? Use 1 in $= 2.54$ cm.

Solution:

To convert between inches and centimeters, we'll say

$$\mu_{\text{length in cm}} = 2.54\mu_{\text{length in in}}$$

$$\mu_{\text{length in cm}} = 2.54(20)$$

$$\mu_{\text{length in cm}} = 50.8 \text{ cm}$$

We'll use the same conversion formula to convert the standard deviation.

$$\sigma_{\text{length in cm}} = 2.54\sigma_{\text{length in in}}$$

$$\sigma_{\text{length in cm}} = 2.54\sqrt{0.81}$$

$$\sigma_{\text{length in cm}} = 2.286 \text{ cm}$$

6. The weights of full-term new born babies are normally distributed with $\mu = 120$ ounces and $\sigma = 20$ ounces. Describe the shape, center, and spread

for the weights of full-term new born babies as measured in pounds. Use 1 pound $= 16$ ounces.

Solution:

We can use a conversion formula to convert the mean from pounds to ounces.

$$\mu_{\text{weight in pounds}} = \frac{1}{16}\mu_{\text{weight in ounces}}$$

$$\mu_{\text{weight in pounds}} = \frac{1}{16}(120)$$

$$\mu_{\text{weight in pounds}} = 7.5 \text{ pounds}$$

Now we'll convert the given standard deviation.

$$\sigma_{\text{weight in pounds}} = \frac{1}{16}\sigma_{\text{weight in ounces}}$$

$$\sigma_{\text{weight in pounds}} = \frac{1}{16}(20)$$

$$\sigma_{\text{weight in pounds}} = 1.25 \text{ pounds}$$

The distribution of weights of full-term new born babies remains normally distributed, even after converting from ounces to pounds. The mean is $\mu = 7.5$ pounds and the standard deviation is $\sigma = 1.25$ pounds.

COMBINATIONS OF RANDOM VARIABLES

■ 1. X and Y are independent random variables with $E(X) = 48$, $E(Y) = 54$, $SD(X) = 3$ and $SD(Y) = 5$. Find $E(X - Y)$ and $SD(X - Y)$.

Solution:

To find the expected value of the difference, we find the difference of the expected values.

$$E(X - Y) = E(X) - E(Y) = 48 - 54 = -6$$

To find the standard deviation of the difference, we have to square both standard deviations in order to get the variances. We get $SD^2(X) = 3^2 = 9$ and $SD^2(Y) = 5^2 = 25$. Then we can find the standard deviation of the difference.

$$SD(X - Y) = \sqrt{SD^2(X) + SD^2(Y)} = \sqrt{3^2 + 5^2} = \sqrt{34} \approx 5.831$$

■ 2. A and B are independent random variables with $E(A) = 6.5$, $E(B) = 4.4$, $SD(A) = 1.6$, and $SD(B) = 2.1$. Find $E(4A + 2B)$ and $SD(4A + 2B)$.

Solution:

The expected value of the sum of variables is the sum of the expected values.

$$E(4A + 2B) = 4E(A) + 2E(B) = 4(6.5) + 2(4.4) = 34.8$$

Then we'll find the standard deviation of the combination. When we scale a variable by some constant k, its standard deviation gets scaled by k as well. So the standard deviations of $4A$ and $2B$ are

$$SD(A) = 1.6$$

$$SD(4A) = 4(1.6)$$

$$SD(4A) = 6.4$$

and

$$SD(B) = 2.1$$

$$SD(2B) = 2(2.1)$$

$$SD(2B) = 4.2$$

Then the variances of the variables $4A$ and $2B$ are

$$SD^2(4A) = 6.4^2$$

$$SD^2(4A) = 40.96$$

and

$$SD^2(2B) = 4.2^2$$

$$SD^2(2B) = 17.64$$

The variance of a combination is the sum of the variances, so

$$SD^2(4A) + SD^2(2B) = 40.96 + 17.64$$

$$SD^2(4A) + SD^2(2B) = 58.6$$

Then the standard deviation of the combination is

$$\sqrt{SD^2(4A) + SD^2(2B)} = \sqrt{58.6}$$

$$\sqrt{SD^2(4A) + SD^2(2B)} \approx 7.66$$

■ 3. The time it takes students to complete multiple choice questions on an AP Statistics Exam has a mean of 55 seconds with a standard deviation of 12 seconds. If the exam consists of 40 multiple choice questions, find the mean total time to finish the exam. Then find the standard deviation in the total time. What assumption must be made?

Solution:

We have to assume that the questions are independent. Then we can say that the mean finishing time is

$$\mu_{Q_1} + \mu_{Q_2} + \mu_{Q_3} + \ldots + \mu_{Q_{40}} = 40(55) = 2{,}200 \approx 36.67 \text{ minutes}$$

and that the variance of the finishing time is

$$\sigma^2_{Q_1} + \sigma^2_{Q_2} + \sigma^2_{Q_3} + \ldots + \sigma^2_{Q_{40}} = 40(12^2) = 40(144) = 5{,}760 \text{ seconds}$$

such that the standard deviation of the finishing time is

$$\sigma = \sqrt{5{,}760} \approx 75.89 \approx 1.26 \text{ minutes}$$

■ 4. Let M represent the height of a male over 21 years of age and let W represent the height of a female over 21 years of age. Let D represent the difference between their heights ($D = M - W$). Let $E(M) = 70$ inches, $\sigma_M = 2.8$ inches, $E(W) = 64.5$ inches and $\sigma_W = 2.4$ inches.

What is the mean and standard deviation of the difference between the two heights?

Solution:

To find the mean of the difference, we'll find the difference of the means.

$$E(M - W) = E(M) - E(W) = 70 - 64.5 = 5.5 \text{ inches}$$

We'll find variance in order to get standard deviation. The variances are $\sigma_M^2 = 2.8^2 = 7.84$ and $\sigma_W^2 = 2.4^2 = 5.76$. Therefore, the standard deviation of the difference is

$$\sigma(M - W) = \sqrt{\sigma_M^2 + \sigma_W^2} = \sqrt{7.84 + 5.76} = \sqrt{13.6} \approx 3.69 \text{ inches}$$

■ 5. The Ironman is a challenge in which a competitor swims 2.4 miles, then bikes 112 miles, and finally runs 26.2 miles. Suppose the times for each of the legs are normally distributed with the given means and standard deviations.

Swim: $\mu_S = 76$ minutes and $\sigma_S = 18$ minutes

Bike: $\mu_B = 385$ minutes and $\sigma_B = 32$ minutes

Run: $\mu_R = 294$ minutes and $\sigma_R = 25$ minutes

What percent of the competitors finish the Ironman in under 710 minutes?

Solution:

Let T be the total time to complete all three legs of the Ironman. Then the mean finishing time is

$$\mu_T = \mu_S + \mu_B + \mu_R = 76 + 385 + 294 = 755 \text{ minutes}$$

Assuming the legs are independent random variables, then we can find the sum of the variances to get the variance of the sum.

$$\sigma_T^2 = \sigma_S^2 + \sigma_B^2 + \sigma_R^2 = 18^2 + 32^2 + 25^2 = 1{,}973$$

Then the standard deviation of finishing time is

$$\sigma_T = \sqrt{1{,}973} \approx 44.42 \text{ minutes}$$

Since S, B, and R are normally distributed, T will also be normally distributed. To find the probability that a finisher will finish in under 710 minutes, we'll find the z-score associated with 710 minutes.

$$z = \frac{710 - 755}{44.42} \approx -1.01$$

If we look up a z-score of $z = -1.01$ in a z-table, we get 0.1562, which means there's an approximately 15.62% chance that a finisher finishes in under 710 minutes.

6. We buy a scratch-off lottery ticket for $\$1$ at the local gas station. If we get three hearts in a row on the scratch-off, the state will pay us $\$500$. Let X be the amount the state pays us and let X have the following probability distribution.

X	\$0	\$500
P(X)	0.999	0.001

Suppose we buy one of these scratch-off tickets every day for a week (7 days). Find the expected value and standard deviation of our total winnings.

Solution:

The expected value of our winnings on any one ticket is

$$E(X) = 0(0.999) + 500(0.001) = \$0.50$$

Find the standard deviation of the winnings by taking the sum of the variances, weighted by the associated probabilities.

$$SD(X) = \sqrt{(0-0.5)^2(0.999) + (500-0.5)^2(0.001)} = \sqrt{249.75} \approx 15.80$$

Let W be the amount the state pays us for 7 lottery tickets. The expected value of the total winnings for 7 lottery tickets is therefore $E(W) = 7(0.5) = \$3.50$. The standard deviation of the total winnings is

$$SD(W) = \sqrt{(15.80)^2 + (15.80)^2 + \ldots + (15.80)^2}$$

$$SD(W) = \sqrt{7(15.80)^2}$$

$$SD(W) = \sqrt{1{,}747.48}$$

$$SD(W) \approx \$41.80$$

These are the mean and standard deviation of total winnings. They don't account for the price we paid for the scratch-off tickets. If we want to account for the cost of the tickets in order to calculate profit, instead of just winnings, then we'd use the probability distribution

X	-\$1	\$499
P(X)	0.999	0.001

in order to calculate the mean and standard deviation of our profit.

PERMUTATIONS AND COMBINATIONS

1. Calculate the binomial coefficient.

$$\binom{12}{7}$$

Solution:

Use the combination formula

$$\binom{n}{k} = {}_nC_k = \frac{n!}{k!(n-k)!}$$

Plug in $n = 12$ and $k = 7$.

$$\binom{12}{7} = {}_{12}C_7 = \frac{12!}{7! \cdot 5!} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

$$\binom{12}{7} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8}{5 \cdot 4 \cdot 3 \cdot 2}$$

$$\binom{12}{7} = 792$$

2. Calculate ${}_{10}P_3$.

Solution:

Use the permutation formula

$$_nP_k = \frac{n!}{(n-k)!}$$

Plug in $n = 10$ and $k = 3$.

$$_{10}P_3 = \frac{10!}{(10-3)!} = \frac{10!}{7!}$$

$$_{10}P_3 = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

$$_{10}P_3 = 10 \cdot 9 \cdot 8$$

$$_{10}P_3 = 720$$

■ 3. How much greater is $_5P_2$ than $_5C_2$?

Solution:

We'll calculate both values, then find the difference.

$$_5P_2 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1} = 5 \cdot 4 = 20$$

$$_5C_2 = \frac{5!}{2!(5-2)!} = \frac{5!}{2! \cdot 3!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} = \frac{5 \cdot 4}{2} = 10$$

The difference between ${}_5P_2$ and ${}_5C_2$ is

$${}_5P_2 - {}_5C_2 = 20 - 10 = 10$$

■ 4. The high school girls' basketball team has 8 players, 5 of whom are seniors. They need to figure out which senior will be captain and which senior will be co-captain. To make it fair, they choose two players out of a hat. The first drawn will be captain and the second will be co-captain. How many different captain/co-captain pairs are possible?

Solution:

Since the order matters, we have to calculate the permutations. There are 5 seniors we can choose from, and 2 spots to put them in.

$${}_nP_k = \frac{n!}{(n-k)!} = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1} = 5 \cdot 4 = 20$$

There are 20 possible captain/co-captain pairs.

■ 5. How many different ways can the letters in the word "SUCCESS" be rearranged?

Solution:

Since the order matters, we have to calculate the permutations. There are 7 letters we can choose from, and 7 spots to put them in.

$$_nP_k = \frac{n!}{(n-k)!} = \frac{7!}{(7-7)!} = \frac{7!}{0!} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{1} = 5{,}040$$

But since the letter S repeats three times in the word and the letter C repeats twice, the actual number of unique rearrangements will be less than $5{,}040$. We have to divide by $3!$ for the S and by $2!$ for the C.

$$\frac{5{,}040}{3! \cdot 2!} = \frac{5{,}040}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = \frac{5{,}040}{12} = 420$$

There are 420 possible arrangements of the letters.

■ 6. Mrs. B's kindergarten class has 14 students and Mr. G's kindergarten class has 16 students. Three students will be selected at random from each of these classrooms to ride on a float in the school parade coming up next week. How many different groups of 6 can be chosen to ride the float?

Solution:

Since order doesn't matter, this is a combination question. We need to find the combination for Mrs. B's class, and then the combination for Mr. G's class. Then we'll multiply those together get the total number of combinations.

$$_nC_k \cdot {_nC_k} = {_{14}C_3} \cdot {_{16}C_3} = 364 \cdot 560 = 203{,}840$$

BINOMIAL RANDOM VARIABLES

■ 1. We toss a fair coin 15 times and record the number of tails. Is this experiment modeled by a binomial random variable? If it isn't, explain why. If it is, determine its parameters n and p and express the binomial random variable as $X \sim B(n, p)$.

Solution:

Yes, this experiment results in a binomial random variable. Let X be the number of tails observed out of 15 tosses. We know that $p = 0.5$ for each trial because there are only two possible outcomes, heads or tails. Therefore, $X \sim B(15,0.5)$.

■ 2. We randomly select students from our school until we find a student in the school band. Assume there are 900 students in the school and 80 participate in the school band. Is this experiment modeled by a binomial random variable? If it isn't, explain why. If it is, determine its parameters n and p and express the binomial random variable as $X \sim B(n, p)$.

Solution:

No, this experiment does not result in a binomial random variable. We do have a fixed probability of success,

$$p = \frac{80}{900} = \frac{4}{45} \approx 0.09$$

and the trials can be considered independent because we have a large population. But we're not using a fixed number of trials, because we're continuing to select students until we find one in the band, and we don't know how many trials that will take.

■ 3. Let $X \sim B(n, p)$ be a binomial random variable with $n = 12$ and $p = 0.08$. Find $P(X = 4)$.

Solution:

We're being asked to find the probability that we get exactly 4 successes in 12 trials, if the probability of success is $p = 0.08$.

$$P(X = 4) = \binom{12}{4}(0.08)^4(1 - 0.08)^8$$

$$P(X = 4) = (495)(0.08)^4(1 - 0.08)^8$$

$$P(X = 4) \approx 0.0104$$

■ 4. Let Y be the number of times we roll a 1 on a fair 6-sided die if we do 10 trials. Fill in the following probability distribution for Y, rounding each probability to 4 decimal places.

Y	0	1	2	3	4	5	6	7	8	9	10
P(Y)											

Solution:

With $n = 10$, $p = 1/6$, and $k = 0,\ 1,\ 2,\ 3,\ \dots\ 10$, find $P(Y = k)$ for each value of k using

$$P(k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k(1-p)^{n-k}$$

After rounding each value to 4 decimal places, the table is

Y	0	1	2	3	4	5	6	7	8	9	10
P(Y)	0.1615	0.3230	0.2907	0.1550	0.0543	0.0130	0.0022	0.0003	0.0000	0.0000	0.0000

5. For each binomial random variable, determine whether the shape of the probability distribution will be skewed right, skewed left, or symmetrical.

1. $X \sim B(n, p)$ with $n = 10$ and $p = 0.15$

2. $Y \sim B(n, p)$ with $n = 10$ and $p = 0.75$

3. $Z \sim B(n, p)$ with $n = 10$ and $p = 0.50$

Solution:

The probability distribution for X will be skewed right because the probability of success, $p = 0.15$, is less than 0.5.

The probability distribution for Y will be skewed left because the probability of success, $p = 0.75$, is greater than 0.5.

The probability distribution for Z will be symmetrical because the probability of success, $p = 0.50$, is exactly 0.5.

■ 6. Suppose an environmental biologist is studying juvenile sunfish mortality. He finds that only 30% of juvenile sunfish survive in a certain lake. Out of 8 randomly selected juvenile sunfish, what is the probability that exactly 3 will survive?

Solution:

We're finding the probability that we get exactly 3 successes in 8 trials.

$$P(X = 3) = \binom{8}{3}(0.3)^3(1 - 0.3)^5$$

$$P(X = 3) = (56)(0.3)^3(1 - 0.3)^5$$

$$P(X = 3) \approx 0.2541$$

POISSON DISTRIBUTIONS

1. A student is able to solve 10 practice problems per hour, on average. Find the probability that she can solve 12 in the next hour.

Solution:

We know this is a Poisson experiment with the following values:

$\lambda = 10$, the average number of practice problems solved in an hour

$x = 12$, the number of homework problems she wants to complete in the next hour

The Poisson probability is

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$P(12) = \frac{10^{12} e^{-10}}{12!}$$

$$P(12) \approx 0.0948$$

So the probability the student will solve 12 homework problems is approximately 0.0948 or 9.48%.

2. A student is able to solve 6 practice problems per hour, on average. Find the probability that she can solve at least 4 in the next hour.

Solution:

The probability that the student solves at least 4 practice problems is the probability that she doesn't solve either 0, 1, 2, or 3 practice problems. So the probability we need to find is

$$P(X \geq 4) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3)$$

We know this is a Poisson experiment with $\lambda = 6$, the average number of practice problems solved in an hour, so the Poisson probability is

$$P(X \geq 4) = 1 - \frac{6^0 e^{-6}}{0!} - \frac{6^1 e^{-6}}{1!} - \frac{6^2 e^{-6}}{2!} - \frac{6^3 e^{-6}}{3!}$$

$$P(X \geq 4) = 1 - \frac{e^{-6}}{1} - \frac{6e^{-6}}{1} - \frac{36e^{-6}}{2} - \frac{216e^{-6}}{6}$$

$$P(X \geq 4) = 1 - \frac{1}{e^6} - \frac{6}{e^6} - \frac{18}{e^6} - \frac{36}{e^6}$$

$$P(X \geq 4) = 1 - \frac{1}{e^6}(1 + 6 + 18 + 36)$$

$$P(X \geq 4) = 1 - \frac{61}{e^6}$$

$$P(X \geq 4) \approx 0.8488$$

So the probability the student will solve at least 4 homework problems is approximately 0.8488 or $84.88\,\%$.

3. A student is able to solve 5 practice problems per hour, on average. Find the probability that she solves at most 3 in the next hour.

Solution:

The probability that the student solves at most 3 practice problems is the probability that she solves either 0, 1, 2, or 3 practice problems. So the probability we need to find is

$$P(X \le 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

We know this is a Poisson experiment with $\lambda = 5$, the average number of practice problems solved in an hour, so the Poisson probability is

$$P(X \le 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X \le 3) = \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} + \frac{5^3 e^{-5}}{3!}$$

$$P(X \le 3) = \frac{e^{-5}}{1} + \frac{5e^{-5}}{1} + \frac{25e^{-5}}{2} + \frac{125e^{-5}}{6}$$

$$P(X \le 3) = \frac{1}{e^5} + \frac{5}{e^5} + \frac{25}{2e^5} + \frac{125}{6e^5}$$

$$P(X \le 3) = \frac{1}{e^5}\left(1 + 5 + \frac{25}{2} + \frac{125}{6}\right)$$

$$P(X \leq 3) = \frac{1}{e^5}\left(\frac{36}{6} + \frac{75}{6} + \frac{125}{6}\right)$$

$$P(X \leq 3) = \frac{118}{3e^5}$$

$$P(X \leq 3) \approx 0.2650$$

So the probability the student will solve at most 3 homework problems is approximately 0.2650 or 26.50%.

4. A baker is able to bake 50 loaves of bread per day, on average. Find the probability that he can bake 60 on Friday.

Solution:

We know this is a Poisson experiment with the following values:

$\lambda = 50$, the average number of loaves of bread baked per day

$x = 60$, the number of loaves of bread he wants to bake on Friday

The Poisson probability is

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$P(60) = \frac{50^{60}e^{-50}}{60!}$$

$$P(60) \approx 0.0201$$

So the probability the baker will bake 60 loaves of bread is approximately 0.0201 or 2.01%.

■ 5. A baker is able to bake 10 cakes per hour, on average. Find the probability that he can bake more than 5 in the next hour.

Solution:

The probability that the baker bakes more than 5 cakes is the probability that he doesn't bake either 0, 1, 2, 3, 4, or 5 cakes. So the probability we need to find is

$$P(X > 5) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

$$-P(X = 3) - P(X = 4) - P(X = 5)$$

We know this is a Poisson experiment with $\lambda = 10$, the average number of cakes baked in an hour, so the Poisson probability is

$$P(X > 5) = 1 - \frac{10^0e^{-10}}{0!} - \frac{10^1e^{-10}}{1!} - \frac{10^2e^{-10}}{2!}$$

$$-\frac{10^3e^{-10}}{3!} - \frac{10^4e^{-10}}{4!} - \frac{10^5e^{-10}}{5!}$$

$$P(X > 5) = 1 - \frac{e^{-10}}{1} - \frac{10e^{-10}}{1} - \frac{100e^{-10}}{2}$$

$$-\frac{1{,}000e^{-10}}{6}-\frac{10{,}000e^{-10}}{24}-\frac{100{,}000e^{-10}}{120}$$

$$P(X>5)=1-\frac{1}{e^{10}}-\frac{10}{e^{10}}-\frac{50}{e^{10}}-\frac{500}{3e^{10}}-\frac{1{,}250}{3e^{10}}-\frac{2{,}500}{3e^{10}}$$

$$P(X>5)=1-\frac{1}{e^{10}}\left(1+10+50+\frac{500}{3}+\frac{1{,}250}{3}+\frac{2{,}500}{3}\right)$$

$$P(X>5)=1-\frac{4{,}433}{3e^{10}}$$

$$P(X>5)\approx 0.9329$$

So the probability the baker will bake more than 5 cakes is approximately 0.9329 or $93.29\,\%$.

6. A baker is able to frost 2 cakes per hour, on average. Find the probability that he frosts fewer than 5 cakes in the next hour.

Solution:

The probability that the baker frosts fewer than 5 cakes is the probability that he frosts either 0, 1, 2, 3, or 4 cakes. So the probability we need to find is

$$P(X<5)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$$

We know this is a Poisson experiment with $\lambda=2$, the average number of cakes frosted in an hour, so the Poisson probability is

$$P(X<5)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$$

$$P(X<5)=\frac{2^0e^{-2}}{0!}+\frac{2^1e^{-2}}{1!}+\frac{2^2e^{-2}}{2!}+\frac{2^3e^{-2}}{3!}+\frac{2^4e^{-2}}{4!}$$

$$P(X<5)=\frac{e^{-2}}{1}+\frac{2e^{-2}}{1}+\frac{4e^{-2}}{2}+\frac{8e^{-2}}{6}+\frac{16e^{-2}}{24}$$

$$P(X<5)=\frac{1}{e^2}+\frac{2}{e^2}+\frac{2}{e^2}+\frac{4}{3e^2}+\frac{2}{3e^2}$$

$$P(X<5)=\frac{1}{e^2}\left(1+2+2+\frac{4}{3}+\frac{2}{3}\right)$$

$$P(X<5)=\frac{7}{e^2}$$

$$P(X<5)\approx 0.9473$$

So the probability the baker will frost fewer than 5 cakes is approximately 0.9473 or 94.73%.

"AT LEAST" AND "AT MOST," AND MEAN, VARIANCE, AND STANDARD DEVIATION

■ 1. Assume X is a binomial random variable. Let $X \sim B(n, p)$ with $n = 15$ and $p = 0.45$. Find $P(X > 7)$.

Solution:

Since we're running $n = 15$ trials, and we want to find the probability that we get the first success *after* the 7th trial, we can express this as

$$P(X > 7) = P(X = 8) + P(X = 9) + \ldots + P(X = 15)$$

which is the same as

$$1 - P(X \leq 7)$$

Find $P(X \leq 7)$.

$$P(X \leq 7) = \binom{15}{0}(0.45)^0(1 - 0.45)^{15} + \binom{15}{1}(0.45)^1(1 - 0.45)^{14}$$

$$+\binom{15}{2}(0.45)^2(1 - 0.45)^{13} + \binom{15}{3}(0.45)^3(1 - 0.45)^{12}$$

$$+\binom{15}{4}(0.45)^4(1 - 0.45)^{11} + \binom{15}{5}(0.45)^5(1 - 0.45)^{10}$$

$$+\binom{15}{6}(0.45)^6(1 - 0.45)^9 + \binom{15}{7}(0.45)^7(1 - 0.45)^8$$

$$P(X \leq 7) = 0.55^{15} + 15(0.45)(0.55^{14})$$

$$+105(0.45^2)(0.55^{13}) + 455(0.45^3)(0.55^{12})$$

$$+1{,}365(0.45^4)(0.55^{11}) + 3{,}003(0.45^5)(0.55^{10})$$

$$+5{,}005(0.45^6)(0.55^9) + 6{,}435(0.45^7)(0.55^8)$$

$$P(X \leq 7) \approx 0.6535$$

Then $1 - P(X \leq 7)$ is

$$1 - P(X \leq 7) \approx 1 - 0.6535$$

$$1 - P(X \leq 7) \approx 0.3465$$

2. According to a 2017-2018 survey, $68\,\%$ of U.S. households own a pet. Suppose we select 12 households at random. What is the probability that fewer than 8 of them own a pet?

Solution:

Let X be the number of households that own a pet. Then we can express the variable as $X \sim B(12{,}0.68)$. The probability that we'll have fewer than 8 successes is

$$P(X < 8) = P(X \leq 7) = P(X = 0) + P(X = 1) + \ldots + P(X = 7)$$

Find $P(X < 8)$.

$$P(X<8)=\binom{12}{0}(0.68)^0(1-0.68)^{12}+\binom{12}{1}(0.68)^1(1-0.68)^{11}$$

$$+\binom{12}{2}(0.68)^2(1-0.68)^{10}+\binom{12}{3}(0.68)^3(1-0.68)^9$$

$$+\binom{12}{4}(0.68)^4(1-0.68)^8+\binom{12}{5}(0.68)^5(1-0.68)^7$$

$$+\binom{12}{6}(0.68)^6(1-0.68)^6+\binom{12}{7}(0.68)^7(1-0.68)^5$$

$$P(X<8)=(0.32^{12})+12(0.68)(0.32^{11})$$

$$+66(0.68^2)(0.32^{10})+220(0.68^3)(0.32^9)$$

$$+495(0.68^4)(0.32^8)+792(0.68^5)(0.32^7)$$

$$+924(0.68^6)(0.32^6)+792(0.68^7)(0.32^5)$$

$$P(X<8)\approx 0.3308$$

■ 3. According to a 2017-2018 survey, 68% of U.S. households own a pet. Suppose 200 households are selected at random. Find the expected value and standard deviation for the number of households that own a pet.

Solution:

Let X be the number of households that own a pet. Then we can express the variable as $X \sim B(12,0.68)$. The expected value is

$$\mu_X = E(X) = (200)(0.68) = 136 \text{ households}$$

The variance is

$$\sigma_X^2 = Var(X) = (200)(0.68)(1 - 0.68) = 43.52$$

which means the standard deviation is

$$\sigma_X = SD(X) = \sqrt{43.52} \approx 6.597 \text{ households}$$

■ 4. $3\ \%$ of runners in the Boston Marathon do not finish. Suppose we select a SRS of 140 Boston Marathon runners. How many do we expect to finish the race?

Solution:

Let X be the number of runners who finish the Boston Marathon. Then we can say $X \sim B(140,0.97)$. Then the expected value is

$$\mu_X = E(X) = (140)(0.97) = 135.8 \text{ runners}$$

■ 5. We roll a fair die 6 times. What is the probability we'll observe an even number in at most 3 of the rolls?

Solution:

Let X be the number of even numbers observed. Then we can say $X \sim B(6,0.5)$.

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X \leq 3) = \binom{6}{0}(0.5)^0(1 - 0.5)^6 + \binom{6}{1}(0.5)^1(1 - 0.5)^5$$

$$+\binom{6}{2}(0.5)^2(1 - 0.5)^4 + \binom{6}{3}(0.5)^3(1 - 0.5)^3$$

$$P(X \leq 3) = (0.5^6) + 6(0.5)(0.5^5) + 15(0.5^2)(0.5^4) + 20(0.5^3)(0.5^3)$$

$$P(X \leq 3) = 0.5^6 + 6(0.5^6) + 15(0.5^6) + 20(0.5^6)$$

$$P(X \leq 3) = 42(0.5^6)$$

$$P(X \leq 3) \approx 0.6563$$

■ 6. We roll two fair 6-sided die 10 times and observe the sum. What is the probability of rolling a sum of 7 on at least six of the rolls?

Solution:

Let X be the number of times we roll a sum of 7. Since there are 36 possible rolls when we roll two die, and 6 of them result in a sum of 7,

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

the probability is

$$P(\text{sum of } 7) = \frac{6}{36} = \frac{1}{6}$$

Therefore we can express X as

$$X \sim B\left(10, \frac{1}{6}\right)$$

So the probability of rolling a sum of 7 at least 6 times out of 10 rolls is

$$P(X \geq 6) = P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)$$

or

$$1 - P(X \leq 5)$$

Find $P(X \leq 5)$.

$$P(X \leq 5) = \binom{10}{0}\left(\frac{1}{6}\right)^0\left(1 - \frac{1}{6}\right)^{10} + \binom{10}{1}\left(\frac{1}{6}\right)^1\left(1 - \frac{1}{6}\right)^9$$

$$+\binom{10}{2}\left(\frac{1}{6}\right)^2\left(1-\frac{1}{6}\right)^8+\binom{10}{3}\left(\frac{1}{6}\right)^3\left(1-\frac{1}{6}\right)^7$$

$$+\binom{10}{4}\left(\frac{1}{6}\right)^4\left(1-\frac{1}{6}\right)^6+\binom{10}{5}\left(\frac{1}{6}\right)^5\left(1-\frac{1}{6}\right)^5$$

$$P(X \le 5)=\left(\frac{5}{6}\right)^{10}+10\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^9$$

$$+45\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^8+120\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7$$

$$+210\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^6+252\left(\frac{1}{6}\right)^5\left(\frac{5}{6}\right)^5$$

$$P(X \le 5) \approx 0.9976$$

Then $1-P(X \le 5)$ is

$$1-P(X \le 5) \approx 1-0.9976$$

$$1-P(X \le 5) \approx 0.0024$$

or about a 0.24% chance.

BERNOULLI RANDOM VARIABLES

1. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." We buy 5 spins for $\$3.00$. If we land on "winner" on any of our 5 spins, we get to choose a stuffed animal. Is this an example of Bernoulli trials?

Solution:

The set of 5 spins can be considered Bernoulli trials because the spins are independent of one another, there are exactly two outcomes (land on a winning, or not), and the probability of success (landing on a winner) remains constant for each trial at $p = 2/8 = 1/4 = 0.25 = 25\%$.

2. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." We buy 5 spins for $\$3.00$. If we land on "winner" on any of our 5 spins, we get to choose a stuffed animal. Find the mean and standard deviation for each trial.

Solution:

We already know that the probability of winning on any single spin is $p = 2/8 = 1/4 = 0.25 = 25\,\%$, which means $\mu = p = 0.25$. The standard deviation will therefore be

$$\sigma = \sqrt{p(1-p)} = \sqrt{(0.25)(1-0.25)} = \sqrt{0.1875} \approx 0.4330$$

■ 3. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." We buy 5 spins for $\$3.00$. If we land on "winner" on any of our 5 spins, we get to choose a stuffed animal. Find the mean and standard deviation for the number of winners expected in a set of 5 spins.

Solution:

We already know that the probability of winning on any single spin is $p = 2/8 = 1/4 = 0.25 = 25\,\%$, which means $\mu = p = 0.25$. Therefore, for 5 spins the mean will be $\mu = np = 5(0.25) = 1.25$. And the standard deviation for 5 spins will be

$$\sigma = \sqrt{np(1-p)} = \sqrt{(5)(0.25)(1-0.25)} = \sqrt{0.9375} \approx 0.9682$$

■ 4. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." We buy 5 spins for $\$3.00$. If we land on "winner" on any of our 5 spins, we get to choose a stuffed animal. Find the probability of observing no winners in a set of 5 spins.

Solution:

The probability of spinning a winner is

$$P(\text{winner}) = p = \frac{2}{8} = \frac{1}{4} = 0.25$$

$$P(\text{non-winner}) = 1 - p = 1 - 0.25 = 0.75$$

Therefore, the probability of no winners in 5 spins is

$$P(\text{no winners in } 5 \text{ spins}) = (0.75)^5 = 0.2373$$

5. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." We buy 5 spins for $\$3.00$. If we land on "winner" on any of our 5 spins, we get to choose a stuffed animal. What is the probability of observing at least 1 winner in a set of 5 spins?

Solution:

If we observe at least one winner out of 5 spins, that means we're looking for the probability of getting 1, 2, 3, 4, or 5 winners. The only result we're excluding is the probability of getting 0 winners. Which means we could flip this problem around and calculate the probability of at least 1 winner as

$$P(W \geq 1) = 1 - P(W = 0)$$

$$P(W \geq 1) = 1 - 0.2373$$

$$P(W \geq 1) = 1 - 0.7627$$

■ 6. Our goal is to learn about the percentage of students with high ACT scores. We randomly select high school seniors and record their highest ACT score. Explain why these aren't Bernoulli trials. Then design a way to conduct the experiment differently so that they can be considered Bernoulli trials.

Solution:

These are not Bernoulli trials because actual ACT scores are recorded. This is a random variable, but the variable can take on many different values, not simply "success" or "failure."

To change the experiment so that we're running Bernoulli trials, we could define a specific range of ACT scores as "failures" and another range as "successes." For instance, we could define a success as a score of 28 or higher, and a failure as a score lower than 28 (27 or lower).

Then the probability of a senior having a score of 28 or higher will have some constant probability of success from trial to trial, so we now have an experiment in which we're using Bernoulli trials.

GEOMETRIC RANDOM VARIABLES

■ 1. We toss a coin until we get "tails." Does this experiment represent a geometric random variable? If it doesn't, explain why. If it does, determine its parameter p and express the variable as $X \sim$ Geom(p).

Solution:

Yes, this experiment results in a geometric random variable. Let X be the number of trials it takes to get our first "tails." We know that $p = 0.5$ for each trial because there are two equally likely outcomes when we flip a coin. So $X \sim$ Geom(0.5).

■ 2. We randomly select students from our school until we find a student in the school band. Assume there are 900 students in the school and 80 participate in the school band. Does this experiment represent a geometric random variable? If it doesn't, explain why. If it does, determine its parameter p and express the variable as $X \sim$ Geom(p).

Solution:

Yes, this experiment results in a geometric random variable. We do have a fixed probability of success,

$$p = \frac{80}{900} = \frac{4}{45} \approx 0.09$$

and the trials can be considered independent because we have a large population. We're selecting students until we find someone in the band. Therefore $X \sim$ Geom(0.09).

■ 3. Let $X \sim$ Geom(p) with $p = 0.25$. Find $P(X = 5)$.

Solution:

We're being asked to find the probability that we get our first success on the 5th trial, if the probability of success on any single trial is $p = 0.25$.

$$P(X = n) = p(1 - p)^{n-1}$$

$$P(X = 5) = (0.25)(1 - 0.25)^{5-1}$$

$$P(X = 5) = (0.25)(0.75)^4$$

$$P(X = 5) \approx 0.0791$$

■ 4. Suppose we roll a 6-sided fair die until we observe a 2. What is the probability that a 2 will be observed within the first 5 trials?

Solution:

The probability of success on any single trial is $p = 1/6$, which means the probability of failure is $1 - p = 1 - (1/6) = 5/6$. Therefore, the probability that we get a 2 within the first 5 trials is

$$P(X \leq 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$

$$P(X \leq 5) = \frac{1}{6}\left(\frac{5}{6}\right)^{1-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{2-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{3-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{4-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{5-1}$$

$$P(X \leq 5) = \frac{1}{6}\left(\frac{5}{6}\right)^{0} + \frac{1}{6}\left(\frac{5}{6}\right)^{1} + \frac{1}{6}\left(\frac{5}{6}\right)^{2} + \frac{1}{6}\left(\frac{5}{6}\right)^{3} + \frac{1}{6}\left(\frac{5}{6}\right)^{4}$$

$$P(X \leq 5) \approx 0.5981$$

5. Suppose we roll a 6-sided fair die until we observe a 2. What is the probability that a 2 won't be observed until at least the 6th trial?

Solution:

The probability of success on any single trial is $p = 1/6$, which means the probability of failure is $1 - p = 1 - (1/6) = 5/6$. Therefore, the probability that we get a 2 within the first 5 trials is

$$P(X \leq 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$

$$P(X \leq 5) = \frac{1}{6}\left(\frac{5}{6}\right)^{1-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{2-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{3-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{4-1} + \frac{1}{6}\left(\frac{5}{6}\right)^{5-1}$$

$$P(X \le 5) = \frac{1}{6}\left(\frac{5}{6}\right)^0 + \frac{1}{6}\left(\frac{5}{6}\right)^1 + \frac{1}{6}\left(\frac{5}{6}\right)^2 + \frac{1}{6}\left(\frac{5}{6}\right)^3 + \frac{1}{6}\left(\frac{5}{6}\right)^4$$

$$P(X \le 5) \approx 0.5981$$

Therefore, the probability that we don't observe a success until the 6th trial or later is

$$P(X \ge 6) \approx 1 - 0.5981$$

$$P(X \ge 6) \approx 0.4019$$

6. According to a 2017-2018 survey, 68% of U.S. households own a pet. Suppose we start randomly surveying households and asking whether they are pet owners. How many do we expect we will need to survey to find our first household that owns a pet?

Solution:

Let X be the trial when we find our first pet owner. We know that X is a geometric random variable with $X \sim$ Geom(0.68). Then the expected value is

$$\mu_X = E(X) = \frac{1}{p} = \frac{1}{0.68} \approx 1.471$$

So we could say that we expect we'll need to survey somewhere between 1 and 2 households in order to find our first pet-owning household.