CSE 312 Section 5

Important Discrete Distributions,
More practice with r.v.s

Administrivia

Announcements & Reminders

- HW3
  - Grades released on gradescope – check your submission to read comments
  - Regrade requests open ~24 hours after grades are released and close after a week

- HW4
  - Released last night
  - Due Wednesday 7/31 @ 11:59pm
- Quiz 4
  - Tomorrow in lecture
  - Discrete random variable zoo

Review & Questions

Any lingering questions from this last week?

Each week in section, we’ll be reviewing the main concepts from this week and putting them into action by going through some practice problems together. But before we get into that review, we’ll try to start off each section with some time for you to ask questions. Was anything particularly confusing this week? Is there anything we can clarify before we dive into the review? This is your chance to clear things up!

Problem 1 – Content Review Questions

1 – Content Review

a) False. True only if they are independent.

1 – Content Review

a) False. True only if they are independent.
b) $9Var(X)$

1 – Content Review

a) False. True only if they are independent.
b) $9Var(X)$
c) True, by LOE

1 – Content Review

a) False. True only if they are independent.
b) $9Var(X)$
c) True, by LOE
d) $3\mathbb{E}[A] + 4$, by LOE

Problem 2 – Pond Fishing

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

a) How many of the next 10 fish I catch are blue, if I catch and release
b) How many fish I had to catch until I catch my first green fish, if I catch and release
c) How many red fish I catch in the next five minutes, if I catch on average r red fish per minute
d) Whether or not my next fish is blue
e) How many of the next 10 fish I catch are blue, if I do not release the fish back to the pond after each catch
f) How many fish I have to catch until I catch three red fish, if I catch and release

Work on this problem with the people around you, and then we'll go over it together!

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

a) How many of the next 10 fish I catch are blue, if I catch and release

$$Bin(10, \frac{B}{N})$$

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

b) How many fish I had to catch until I catch my first green fish, if I catch and release

$$Geo(\frac{G}{N})$$

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

c) How many red fish I catch in the next five minutes, if I catch on average r red fish per minute

$$Poi(5r)$$

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

d) Whether or not my next fish is blue

$$Ber(B/N)$$

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

e) How many of the next 10 fish I catch are blue, if I do not release the fish back to the pond after each catch

$$HypGeo(N, B, 10)$$

2 – Pond Fishing

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + G + R = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

a) How many fish I have to catch until I catch three red fish, if I catch and release

$$NegBin(3, R/N)$$

Problem 5 – Best Coach Ever!!

5 – Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

a) How many matches do you expect to fight until you win 10 times and what kind of r.v. is this?
b) You only get to play 12 matches every year. To win a spot in the Annual Boxing Championship, a boxer needs to win at least 10 matches in a year. What is the probability that you will go to the Championship this year and what kind of random variable is the number of matches you win out of 12
c) Let p be your answer to part (b). How many times can you expect to go to the Championship in your 20 year career.

Work on this problem with the people around you, and then we'll go over it together!

5 – Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

a) How many matches do you expect to fight until you win 10 times and what kind of r.v. is this?

The number of matches you have to fight until you win 10 times can be modeled by $\sum_{i=1}^{10} X_i$ where $X_i \sim Geo(0.2)$ is the number of matches you have to fight to go from $i-1$ wins to i wins, including the match that gets you your ith win, where every match has a 0.2 probability of success.

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{10} X_i\right] = \sum_{i=1}^{10} \mathbb{E}[X_i] = \sum_{i=1}^{10} \frac{1}{0.2} = 50$$

5 – Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

b) You only get to play 12 matches every year. To win a spot in the Annual Boxing Championship, a boxer needs to win at least 10 matches in a year. What is the probability that you will go to the Championship this year and what kind of random variable is the number of matches you win out of 12

You can go to the championship if you win more than or equal to 10 times this year. Let Y be the number of matches you win out of the 12 matches. Note that $Y \sim Bin(12{,}0.2)$. Since the max number you can win is 12 (there are 12 matches), we are looking for $\mathbb{P}(10 \leq Y \leq 12)$. Thus, since Y is discrete, we are interested in

$$\mathbb{P}(Y = 10) + \mathbb{P}(Y = 11) + P(Y = 12) = \sum_{i=10}^{12} \binom{12}{i} 0.2^i (1 - 0.2)^{12-i}$$

5 – Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

c) Let p be your answer to part (b). How many times can you expect to go to the Championship in your 20 year career.

The number of times you go to the championship can be modeled by $Y \sim Binomial(20, p)$. So, $\mathbb{E}[Y] = 20 \cdot Y$

Problem 3

3

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.) Find $\mathbb{E}[X]$.

Work on this problem with the people around you, and then we'll go over it together!

3

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.) Find $\mathbb{E}[X]$.

Let X_i be 1 if bin i is empty, and 0 otherwise.

$$X = \sum_{i=1}^{n} X_i$$

$$\mathbb{E}[X_i] = 1 * \mathbb{P}(X_i = 1) + 0 * \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = \left(\frac{n-1}{n}\right)^m$$

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n * \left(\frac{n-1}{n}\right)^m$$

Problem 4

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

(b) Find $E[X]$.

(c) Find $Var(X)$.

Work on this problem with the people around you, and then we'll go over it together!

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

First let us define the range of X. A three sided-die can take on values {1,2,3}. Since X is the sum of two rolls, the range of X is $\Omega_X = \{2, 3, 4, 5, 6\}$.

We must define two random variables R_1, R_2 with R_1 being the roll of the first die, and R_2 being the roll of the second die. Then, $X = R_1 + R_2$. Note that $\Omega_{R1} = \Omega_{R2} = \{1,2,3\}$\$. With that in mind we can find the PMF of X:

Work on this problem with the people around you, and then we'll go over it together!

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

This gives us the following:

$$
\begin{aligned}
p_X(k) = \mathbb{P}(X = k) &= \sum_{i \in \Omega_{R1}} \mathbb{P}(R_1 = i, R_2 = k - i) \\
&= \sum_{i \in \Omega_{R1}} \mathbb{P}(R_1 = i) \cdot \mathbb{P}(R_2 = k - i) \qquad \text{(By independence of the rolls)} \\
&= \sum_{i \in \Omega_{R1}} \frac{1}{3} \cdot p_{R2}(k - i) \\
&= \frac{1}{3} \left(p_{R2}(k-1) + p_{R2}(k-2) + p_{R2}(k-3) \right)
\end{aligned}
$$

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

At this point, we can evaluate the pmf of X for each value in the range of X, noting that $p_{R2}(k - i) = 0$ if $\mathrm{k} - i \notin \Omega_{R2}$, $1/3$ otherwise. We get:

$$p_X(k) = \begin{cases} 1/9 & k = 2 \\ 2/9 & k = 3 \\ 3/9 & k = 4 \\ 2/9 & k = 5 \\ 1/9 & k = 6 \\ 0 & \text{otherwise} \end{cases}$$

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(b) Find $E[X]$.

There are two ways to find the expected value of X. We could apply the *definition of expectation* using the PMF found in part (a). This gives us

$$\mathbb{E}[X] = \sum_{k=2}^{6} k p_X(k) = 2 \cdot \frac{1}{9} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{3}{9} + 5 \cdot \frac{2}{9} + 6 \cdot \frac{1}{9} = \boxed{4}$$

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(b) Find $E[X]$.

Alternatively, we can use *linearity of expectation* here. Let R_1 be the roll of the first die, and R_2 the roll of the second. Then, $X = R_1 + R_2$.
By linearity of expectation, we get:

$$\mathbb{E}[X] = \mathbb{E}[R_1 + R_2] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$$

We compute:

$$\mathbb{E}[R_1] = \sum_{i \in \Omega_{R1}} i \cdot \mathbb{P}(R_1 = i) = \sum_{i \in \Omega_{R1}} i \cdot \frac{1}{3} = \frac{1}{3}(1 + 2 + 3) = 2$$

Similarly, $E[R_2] = 2$, since the rolls are independent.

Plugging into our expression for the expectation of X gives us:

$$E[X] = 2 + 2 = \boxed{4}$$

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(c) Find $\text{Var}[X]$.

We know from the definition of variance that

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

We can compute the $\mathbb{E}[X^2]$ term as follows:

$$\mathbb{E}[X^2] = \sum_{x=2}^{6} x^2 p_X(x) = \frac{2^2 \cdot 1 + 3^2 \cdot 2 + 4^2 \cdot 3 + 5^2 \cdot 2 + 6^2 \cdot 1}{9} = \frac{52}{3}$$

Plugging this into our variance equation gives us

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{52}{3} - 4^2 = \boxed{\frac{4}{3}}$$

3

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(c) Find $\mathrm{Var}[X]$.

We know from the definition of variance that

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

We can compute the $\mathbb{E}[X^2]$ term as follows:

$$\mathbb{E}[X^2] = \sum_{x=2}^{6} x^2 p_X(x) = \frac{2^2 \cdot 1 + 3^2 \cdot 2 + 4^2 \cdot 3 + 5^2 \cdot 2 + 6^2 \cdot 1}{9} = \frac{52}{3}$$

Plugging this into our variance equation gives us

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{52}{3} - 4^2 = \boxed{\frac{4}{3}}$$

That's All, Folks!

Thanks for coming to section this week!
Any questions?