Discrete Distributions

Hypergeometric Distribution and Poisson Distribution

Noara Razzak

April 30, 2020

Hypergeometric Distribution

A statistical experiment is defined as a hypergeometric experiment when a discrete random variable X is characterized by:

1. A fixed number of trials.

2. The probability of success is not the same from trial to trial.

We sample from two groups of items when we are interested in only one group. X is defined as the number of successes out of the total number of items chosen.

Hypergeometric Distribution

$$Y \sim \text{Hypergeometric}(r, b, n)$$

- r = the number of items in the group of interest
- b = the number of items in the group not of interest
- n = the number of items chosen
- N = the total number of items in both groups

PMF of Hypergeometric Distribution

$$P(Y = x) = \frac{\binom{r}{x}\binom{b}{n-x}}{\binom{N}{n}}$$

Expectation

$$E[Y] = \frac{nr}{r+b},$$

Hypergeometric Distribution

- You take samples from two groups.
- You are concerned with a group of interest, called the first group.
- You sample without replacement from the combined groups. For example, you want to choose a debate team from a combined group of 11 men and 13 women. The team consists of ten players.

Hypergeometric Distribution

- Each pick is not independent, since sampling is without replacement. In the example, the probability of picking a woman first is 13/24 . The probability of picking a man second is 11/23 if a woman was picked first. It is 10/23 if a man was picked first. The probability of the second pick depends on what happened in the first pick.
- You are not dealing with Bernoulli Trials.

Hypergeometric Distribution

Example 1 You are president of an on-campus special events organization. You need a committee of seven students to plan a special birthday party for the president of the college. Your organization consists of 18 women and 15 men. You are interested in the number of men on your committee. This is a Hypergeometric problem because you are choosing your committee from two groups (men and women)

Hypergeometric Distribution

Example 1

1. In words, define the random variable X.
2. List the values that x may take on.
3. Find r, b and n.
4. Give the distribution of X.
5. If the members of the committee are randomly selected, what is the probability that your committee has four men $P(X = 4)$?
6. If the members of the committee are randomly selected, what is the probability that your committee has more than two men?

Hypergeometric Distribution

Example 1 - ***Solution***

1. X is the number of men chosen in the committee.
2. $x = 0, 1, 2, 3, 4, 5, 6$ and 7.
3. $r = 15$, $b = 18$, $n = 7$
4. $X \sim H(15, 18, 7)$
5. $P(X = 4) = \frac{\binom{15}{4}\binom{18}{3}}{\binom{33}{7}} = 0.2607$

Hypergeometric Distribution

Example 1 - Solution

6. For more than two men we need to calculate $P(X=0)+P(X=1)+P(X=2)$. Then we subtract the answer from 1.

   - $P(X=0)=\frac{\binom{15}{0}\binom{18}{7}}{\binom{33}{7}}=0.007$
   - $P(X=1)=\frac{\binom{15}{1}\binom{18}{6}}{\binom{33}{7}}=0.065$
   - $P(X=2)=\frac{\binom{15}{2}\binom{18}{5}}{\binom{33}{7}}=0.211$
   - $P(X>2)=1-[P(X=0)+P(X=1)+P(X=2)]=0.717$

Hypergeometric Distribution

Example 2 A school site committee is to be chosen randomly from six men and five women. The committee will consist of four members chosen randomly.

Hypergeometric Distribution

Example 2

1. In words, define the random variable X.
2. List the values that x may take on.
3. Find r, b and n.
4. Give the distribution of X.
5. If the members of the committee are randomly selected, what is the probability that your committee has two men $P(X = 2)$?

Hypergeometric Distribution

Example 2 - Solution

1. X is the number of men chosen in the committee.
2. $x = 0, 1, 2, 3$ and 4.
3. $r = 6,\ b = 5,\ n = 4$
4. $X \sim H(6, 5, 4)$
5. $P(X = 4) = \dfrac{\binom{6}{2}\binom{5}{2}}{\binom{11}{4}} = 0.4545$

Poisson Distribution

A statistical experiment is defined as a Poisson experiment when a discrete random variable X counts the number of times,n a certain event will occur in a specific interval.

Poisson Distribution

If we perform n independent trials, the total number of successes Y follows a **Poisson distribution**:

$$Y \sim \text{Poisson}(\mu)$$

PMF of Poisson Distribution

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

Expectation

$$E[Y] = np,$$

Variance

$$Var(Y) = \mu$$

Note: $p =$ The probability that the event occurs in a given interval is the same for all intervals.

Poisson Distribution

- The probability, p that the event occurs in a given interval is the same for all intervals.
- The events occur with a known mean and independently of the time since the last event.
- There is a interval of interest which provides us with μ.

Poisson Distribution

Example 3 The average number of loaves of bread put on a shelf in a bakery in a half-hour period is 12. Of interest is the number of loaves of bread put on the shelf in five minutes. The time interval of interest is five minutes.

Poisson Distribution

Example 3

1. In words, define the random variable X.
2. List the values that x may take on.
3. Find the mean μ of loaves of bread produced in 5 minutes.
4. Give the distribution of X.
5. What is the probability that the number of loaves, selected randomly, put on the shelf in five minutes is three?
6. What is the probability that the number of loaves, selected randomly, put on the shelf in five minutes is less than three?

Poisson Distribution

Example 3 - Solution

1. X is the number of bread made within the five minute interval.
2. $x = 0, 1, 2, 3, 4, 5, 6, 7...$
3. Notice that 12 loaves of bread are produced in 30 minutes. Therefore, 2 loaves are made in 5 minutes. Mean: $\mu = 2$.
4. $X \sim P(2)$
5. $P(X = 3) = \dfrac{(e^{-2})(2^3)}{3!} = 0.1804$

Poisson Distribution

Example 3 - Solution

6. For less than three loaves of bread we need to calculate $P(X = 0) + P(X = 1) + P(X = 2)$.

- $P(X = 0) = \frac{(e^{-2})(2^0)}{0!} = 0.1353$
- $P(X = 1) = \frac{(e^{-2})(2^1)}{1!} = 0.2706$
- $P(X = 2) = \frac{(e^{-2})(2^2)}{2!} = 0.2706$
- $P(X < 3) = [P(X = 0) + P(X = 1) + P(X = 2)] = 0.6766$

Poisson Distribution

Example 4 Text message users receive or send an average of 41.5 text messages per day. Of interest is the number of messages received in an hour.

Poisson Distribution

Example 4

1. In words, define the random variable X.
2. List the values that x may take on.
3. Find the mean μ of text messages received in an hour.
4. Give the distribution of X.
5. What is the probability that the number of text messages received in an hour is two?

Poisson Distribution

Example 4 - Solution

1. X is the number of text messages received in an hour
2. $x = 0, 1, 2, 3, 4, 5, 6, 7...$
3. Notice that 41.5 text messages are received in 24 hours. Therefore $41.5/24 = 1.73$ messages are received in an hour. Mean: $\mu = 1.73$.
4. $X \sim P(1.73)$
5. $P(X = 2) = \dfrac{(e^{-1.73})(1.73^2)}{2!} = 0.2655$

This ends our discussion of Hypergeometric and Poisson Distribution. Next class we will move on to Continuous Distributions.