



Homework 3

🔍 course	
📁 area	ucla
🕒 created	@February 2, 2023 12:01 PM
🕒 updated	@February 3, 2023 9:20 PM
📌 ✓ tasks	
📌 📄 courses	
🔗 URL	
📁 tags	math

Definitions



Big Ideas

▼ Question 1

Deck of cards, draw until Ace. X is r.v. number of cards until Ace. Find the PMF of X

▼ Solution

We know the maximum cards we can draw till an ace is 49 cards (all 48 other and the first ace). So we can express the PMF (by its definition) as

$$k \in \{1, \dots, 49\} \implies p_X(k) = \mathbb{P}(X = k)$$

Logically, this can be understood to mean the probability of no aces in $k - 1$ AND (intersection) an ace on the k^{th} draw. We can write the conditional probability as

$$\mathbb{P}(\text{ace on } k | \text{no aces in first } k - 1) \cdot \mathbb{P}(\text{no aces in first } k - 1)$$

Therefore,

$$p_X(k) = \mathbb{P}(X = k) = \frac{4}{52 - (k - 1)} \cdot \frac{\binom{48}{k-1}}{\binom{52}{k-1}}$$

▼ Question 2

X is d.r.v. with PMF

$$p_X(x) = \frac{c}{(1+x)(2+x)} \quad \forall x \in \mathbb{W}, c \in \mathbb{R}$$

▼ a) Find $\mathbb{P}(a \leq X \leq b)$ for $0 \leq a < b$

We can use the following proposition to find the probability of an interval

$$\mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a) = \sum_{y \leq b \in S} p_X(y) - \sum_{y \leq a \in S} p_X(y)$$

In our case, we can express this as

$$\mathbb{P}(a \leq X \leq b) = \sum_{x=0}^b p_X(x) - \sum_{x=0}^a p_X(x)$$

▼ b) Find c to make the PMF genuine

We can use the proposition that the PMF must sum to 1 over the set for which the random variable is defined, thus

$$1 = \sum_{x=0}^{\infty} \frac{c}{x^2 + 3x + 2}$$

The telescoping series test tells us this sum converges to 1 if and only if $c = 1$.

▼ c) $\mathbb{P}(X \geq 4 | X \geq 1)$

We can note that the complement of the first term in the conditional probability is limited by the second term such that

$$\mathbb{P}(X \geq 4 | X \geq 1) = 1 - \mathbb{P}(1 \leq X \leq 4) - \mathbb{P}(X = 4)$$

Because this probability is genuine for $c = 1$ as shown in part (b), we can use part (a) to find:

$$\mathbb{P}(X \geq 4 | X \geq 1) = 1 - \left(\sum_{x=1}^3 p_X(x) \right) = 1 - \frac{3}{10} = \frac{7}{10}$$

▼ d) $\mathbb{E}[X]$

We can use the formula for expected value to find

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} x p_X(x) = \sum_{x=0}^{\infty} \frac{x}{x^2 + 3x + 2} \longrightarrow \text{diverges}$$

This sum diverges to infinity, thus the expected value depends on the interval we choose to consider. This is because the set for which the random variable takes values from is not finite and countable.

▼ Question 3

$$p_X(x) = C2^{-|x|} \quad x \in \mathbb{Z}.$$

Find $\mathbb{P}(-10 \leq X \leq 10)$

▼ Solution

We know the sum of the PMF over the set it is defined is equal to 1 (by the definition of the PMF). So, we can write

$$1 = \sum_{x=-\infty}^{\infty} C2^{-|x|} = \sum_{x=1}^{\infty} C2^{-|x|} + \sum_{x=-\infty}^{-1} C2^{-|x|} + C$$

Because we have split it to this sum of sums, we can drop the absolute value since their summing “domain” is entirely encompassed without the absolute value. And, since we are given that the sum from 1 to infinity of 2^{-x} is 1, we can write:

$$1 = \sum_{x=1}^{\infty} C2^{-x} + \sum_{x=-\infty}^{-1} C2^{-x} + C = C \cdot 1 + C \cdot 1 + C$$

Thus, we can find the constant as

$$C = \frac{1}{3}$$

Now, we can find the probability as

$$\mathbb{P}(-10 \leq X \leq 10) = \sum_{x=-10}^{10} \frac{1}{3} 2^{-|x|} = \sum_{x=1}^{10} \frac{1}{3 \cdot 2^{|x|}} + \sum_{x=-10}^{-1} \frac{1}{3 \cdot 2^{|x|}} + \frac{1}{3}$$

Computing this using the given information gives

$$\mathbb{P}(-10 \leq X \leq 10) = 2 \cdot \frac{341}{1024} + \frac{1}{3} = \frac{1535}{1536}$$

▼ Question 4

▼ a) Argument for the PMF of a hypergeometric distribution

For the given probability, we can first note that the cardinality of all the possible combinations of selecting n objects from the collection of N objects is simply $\binom{N}{n}$. Thus, we leave this in the denominator of the probability.

Now, the number of ways to select (choose) exactly x objects from collection N_1 is $\binom{N_1}{x}$ and similarly to choose the remaining $n - x$ objects from collection N_2 is $\binom{N_2}{n-x}$. Thus, by the multiplication principle we get

$$p_X(x) = \frac{\binom{N_1}{x} \binom{N_2}{n-x}}{\binom{N}{n}}$$

▼ b) Use the Vandermonde identity to prove $p_X(x)$ is a PMF

For the given function, it is a PMF if its sum for all the elements it takes values from is 1:

$$1 = \sum_{x \in S} p_X(x)$$

For our case, the maximum number of objects we can choose is n and the minimum is 0, so

$$1 = \sum_{x=0}^n \frac{\binom{N_1}{x} \binom{N_2}{n-x}}{\binom{N}{n}}$$

Now we can use the Vandermonde identity such that $k = x$ and $r = n$ and $m = N_1$ and $n = N_2$ such that:

$$\sum_{x=0}^n \frac{\binom{N_1}{x} \binom{N_2}{n-x}}{\binom{N}{n}} = \frac{\binom{N_1+N_2}{n}}{\binom{N}{n}} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1$$

Thus the function is a genuine PMF

▼ Question 5

Suppose there are 3 defective items in a collection of 50 items. A sample of size 10 is taken at random and without replacement. Let X denote the number of defective items in the sample. Find the probability that the sample contains:

▼ a) Exactly 1 defective item

We can model the random variable as the number of defective items per sample $X \sim \text{Binomial}(10, \frac{3}{50})$. For such a Binomial r.v. the PMF tells us the probability of exactly 1 defective item

$$\mathbb{P}(X = 1) = p_X(x = 1) = \binom{10}{1} \cdot \left(\frac{3}{50}\right)^1 \cdot \left(1 - \frac{3}{50}\right)^{10-1} \approx 0.3438$$

▼ b) At most 1 defective item

Using the setup from the previous part:

$$\mathbb{P}(X \leq 1) = p_X(x \leq 1) = \sum_{x=0}^1 \binom{10}{x} \cdot \left(\frac{3}{50}\right)^x \cdot \left(1 - \frac{3}{50}\right)^{10-x}$$

$$\mathbb{P}(X \leq 1) = \left(1 - \frac{3}{50}\right)^{10} + (a) \approx 0.8824$$

▼ Question 6

$f(t) = \log M_X(t)$ is smooth and well defined over $t \in (-\delta, \delta)$ s.t. $\delta > 0$

▼ a) $f'(0) = \mathbb{E}[X]$

We know the definition of the MGF as

$$M_X(t) = \mathbb{E}[e^{tX}] \implies \log M_X = \mathbb{E}[Xt] \quad (1)$$

$$\frac{d^r}{dt^r} M_X \Big|_{t=0} = \mathbb{E}[X^r] \quad (2)$$

Then, we can use (2) to take the derivative of this term to get the expected value

$$f'(0) = \frac{d}{dt} \log M_X \Big|_{t=0} = \frac{d}{dt} \log \mathbb{E}[e^{tX}] \Big|_{t=0} = \mathbb{E}[X e^{tX}] \Big|_{t=0} = \mathbb{E}[X]$$

▼ b) $f''(0) = \text{var}(X)$

We can take the derivative of the un-evaluated derivative from (a) to use (2) to get:

$$f''(0) = \frac{d^2}{dt^2} \log M_X \Big|_{t=0} = \frac{d}{dt} \log \mathbb{E}[X e^{tX}] \Big|_{t=0} = \mathbb{E}[X^2 e^{tX}] \Big|_{t=0} = \mathbb{E}[X^2] = \text{var}(X)$$

This is the condensed form of the variance

▼ Question 7

Let X equal an integer selected at random from the first m positive integers, $\{1, 2, \dots, m\}$. Find the value of m for which $\mathbb{E}[X] = \text{var}(X)$.

▼ Solution

Given the information about the random variable where X is the number of integers selected from m (for which we are only ever selecting 1), we can define the PMF as

$$p_X(x) = \mathbb{P}(X = x) = \frac{1}{m} \quad 1 \leq m \leq \infty \quad (1)$$

So, we can use the definition of expected value and variance to set up the equation:

$$\sum_{x=1}^m x p_X(x) = \sum_{x=1}^m x^2 p_X(x) - \left[\sum_{x=1}^m x p_X(x) \right]^2 \quad (2)$$

We can use the PMF in (1) to rewrite (2) as

$$\frac{1}{m} \sum_{x=1}^m x = \frac{1}{m} \sum_{x=1}^m x^2 - \frac{1}{m^2} \left[\sum_{x=1}^m x \right]^2 \quad (2)$$

We can use certain propositions of these sums to write

$$\frac{m+1}{2} = \frac{m^2+1}{12}$$

Using algebra we can find this equation is true for (omitting the negative possible solution because it is irrelevant) :

$$m = 7$$

▼ Question 8

$X_1, X_2 \sim \text{Unif}\{1, \dots, m\}$ s.t. $X = \max(X_1, X_2)$

▼ a) Find the PMF of X

From the definition of the PMF we can write

$$p_X(k) = \mathbb{P}(X = k) = \mathbb{P}(\max(X_1, X_2) = k)$$

For this pair of variables, we can show the values they take as

$$p_X(k) = (1, k), \dots, (k, k) \text{ and } (k, 1), \dots, (k, k) = \frac{2k-1}{m^2}$$

▼ b) Find the mean and variance

From the definition of expected value

$$\mathbb{E}[X] = \sum_{x \in S} x p_X(x) = \sum_{k=1}^m k \left(\frac{2k-1}{m^2} \right) = \frac{1}{m^2} \left[2 \sum_{k=1}^m k^2 - \sum_{k=1}^m k \right]$$

Using a sum convergence stated in discussion (and simplifying for readability):

$$\mathbb{E}[X] = \frac{(m+1)(4m-1)}{6m}$$

Now, we can find the variance as

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

For which, we can find

$$\mathbb{E}[X^2] = \sum_{x \in S} x^2 p_X(x) = \frac{1}{m^2} \left[2 \sum_{k=1}^m k^3 - \sum_{k=1}^m k^2 \right]$$

Using the given proposition and the one used previously (and simplifying):

$$\mathbb{E}[X^2] = \frac{(m+1)(3m^2 + m - 1)}{6m}$$

So,

$$\text{var}(X) = \frac{(m+1)(3m^2 + m - 1)}{6m} - \left(\frac{(m+1)(4m-1)}{6m} \right)^2$$

▼ c) Rolling 2 fair 6-sided die

We can represent this problem using our previous parts such that $m = 6$

$$\mathbb{E}[X] = \frac{(6+1)(4(6)-1)}{6(6)} = \frac{161}{36}$$

$$\text{var}(X) = \frac{(6+1)(3(6)^2 + (6) - 1)}{6(6)} - \left(\frac{161}{36}\right)^2 = \frac{2555}{1296}$$



SUMMARY