STAT 230

June 6 Tutorial Solutions

Preview

In this tutorial, we will look at the relationship between a pf and cdf for a discrete random variable, and determine how to obtain one from the other.

We will consider three special discrete distributions, the Binomial, Hypergeometric, and Negative Binomial distributions.

Then, we will try to see if we can determine the appropriate distribution to be used for certain probability calculations, playing Name That Distribution!

Problem 1

Suppose you are given the following information regarding the cdf of a discrete random variable, X:

x	0	1	2	3	4	5	6
F(x)	0	0.05	0.2	0.3	0.55	0.8	1

a) Using the information given above, determine $P(2 \leq X \leq 5)$.

b) Determine the pf of the discrete random variable, X.

c) Verify the calculation in part a) using the pf from part b).

Problem 1 Solution

Suppose you are given the following information regarding the cdf of a discrete random variable, X:

x	0	1	2	3	4	5	6
F(x)	0	0.05	0.2	0.3	0.55	0.8	1

a) Using the information given above, determine $P(2 \leq X \leq 5)$.

Solution:

You have to be very careful here. You are given the cdf of X in table form.

We want $P(2 \leq X \leq 5)$. It's very tempting to think that $P(2 \leq X \leq 5) = F(5) - F(2)$, but this is incorrect. If you need to, please take a minute to think about this.

Instead, $P(2 \leq X \leq 5) = F(5) - F(1)$. So, $P(2 \leq X \leq 5) = 0.8 - 0.05 = 0.75$

We will verify this result using the pf in part c).

Problem 1 Solution (Cont'd)

Suppose you are given the following information regarding the cdf of a discrete random variable, X:

x	0	1	2	3	4	5	6
F(x)	0	0.05	0.2	0.3	0.55	0.8	1

b) Determine the pf of the discrete random variable, X.

Solution:

From the table, we see that $F(0) = 0$, so $P(X = 0) = f(0) = 0$.

With $F(1) = F(0) + P(X = 1)$, we have $P(X=1) = f(1) = F(1) – F(0) = 0.05$.

Also, $F(2) = F(1) + P(X = 2)$. This gives $P(X = 2) = f(2) = F(2) – F(1) = 0.2 – 0.05 = 0.15$.

$F(3) = F(2) + P(X = 3)$. This gives $P(X = 3) = f(3) = F(3) – F(2) = 0.3 – 0.2 = 0.1$.

$F(4) = F(3) + P(X = 4)$. This gives $P(X = 4) = f(4) = F(4) – F(3) = 0.55 – 0.3 = 0.25$.

$F(5) = F(4) + P(X = 5)$. This gives $P(X = 5) = f(5) = F(5) – F(4) = 0.8 – 0.55 = 0.25$.

Finally, $F(6) = F(5) + P(X = 6)$. This gives $P(X = 6) = F(6) – F(5) = 1 – 0.8 = 0.2$.

Problem 1 Solution (Cont'd)

Based on the calculations from the previous slide, here is the pf for the discrete random variable, X:

x	0	1	2	3	4	5	6
f(x)	0	0.05	0.15	0.1	0.25	0.25	0.2

You notice that all values of f(x) are ≥ 0, and the sum of the probabilities equals 1. As such we have a valid pf.

Problem 1 Solution (Cont'd)

Here is the pf in table form again for reference:

x	0	1	2	3	4	5	6
f(x)	0	0.05	0.15	0.1	0.25	0.25	0.2

c) Verify the calculation in part a) using the above pf from part b).

Solution:

In part a), we wanted to calculate $P(2 \leq X \leq 5)$.
We know that $P(2 \leq X \leq 5) = P(X=2) + P(X=3) + P(X=4) + P(X=5) = f(2) + f(3) + f(4) + f(5)$.
So, $P(2 \leq X \leq 5) = 0.15 + 0.1 + 0.25 + 0.25 = 0.75$ (same answer as in part a)).

Special Discrete Distribution: Hypergeometric Distribution Set-up

We have a collection of N objects that are classified into two types – “success” types (S) and “failure” types (F). From the N objects, there are r successes and (N-r) failures. Objects are selected **without replacement**.

We let X represent the number of successes obtained in a sample of n objects.

In this case, X has a Hypergeometric distribution with parameters N, r, and n.

The pf of X is given by:

f(x) = P(X = x) = $\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$, for max{0, n – N + r} ≤ x ≤ min{r, n} ; 0, otherwise

Problem 2

Fifteen televisions are to be purchased by a local sports bar from a large production run of 500.

Suppose that 450 will last at least 5 years without needing repair.

Find the probability that at least 14 of the 15 televisions will last at least five years without needing to be repaired.

Problem 2 Solution

Fifteen televisions are to be purchased by a local sports bar from a large production run of 500.

Suppose that 450 will last at least 5 years without needing repair.

Find the probability that at least 14 of the 15 televisions will last at least five years without needing to be repaired.

Solution:

Let X represent the number of TVs in the sample of 15 that will last at least five years without needing to be repaired..

In this case, X has a Hypergeometric distribution with N=500, r=450, n=15.

The pf of X is given by:

f(x) = P(X = x) = $\frac{\binom{450}{x}\binom{50}{15-x}}{\binom{500}{15}}$**, for 0 ≤ x ≤ 15 ; 0, otherwise**

We want P(X ≥ 14) = P(X = 14) + P(X = 15) = $\frac{\binom{450}{14}\binom{50}{1}+\binom{450}{15}\binom{50}{0}}{\binom{500}{15}}$ **= 0.54688**

Special Discrete Distribution: Binomial Distribution Set-up

Suppose that we repeat independent Bernoulli trials consisting of two outcomes:

S (success) and F (failure), and P(S) = p in each trial, with $0 < p < 1$.

We are interested in the number of successes in n independent trials, where the probability of success stays constant from trial to trial.

In this case, X has a Binomial distribution. We write X~Binomial(n, p).

The pf of X is given by:

$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, for x = 0, 1, 2, ……., n ; 0, otherwise.

Problem 3

A large fast-food chain is having a contest. They have distributed millions of game pieces and will give away thousands of prizes!

Every time a customer purchases a sandwich, side order, and a drink (this is known as a "combo"), they will receive a game piece, for a chance to play and win a prize.

The fast food chain claims that there is a 1 in 5 chance of winning a prize with each game piece.

Note: We will assume that the probability of a win is constant for each game piece, and that the outcomes for each game piece are independent.

a) What is the probability that a customer will win exactly 3 prizes if they purchase 5 combos?

b) What is the probability that a customer will win at least one prize if they purchase 3 combos?

c) How many game pieces are needed so that a customer has at least a 95% chance of winning at least one prize? (Take a guess before this calculation is done!)

Problem 3 Solution

a) What is the probability that a customer will win exactly 3 prizes if they purchase 5 combos?

Solution:

Let X represent the number of prizes won if 5 combos are purchased.
In this case, X~Binomial(n = 5, p = 1/5).

The pf, f(x), is given by:
P(X = x) = $\binom{5}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{5-x}$**, for x = 0, 1, 2, 3, 4, 5;0, otherwise.**
We want P(X = 3) = f(3) = $\binom{5}{3}\left(\frac{1}{5}\right)^{3}\left(\frac{4}{5}\right)^{2} = \frac{32}{625}$ **= 0.0512**

Problem 3 Solution (Cont'd)

b) What is the probability that a customer will win at least one prize if they purchase 3 combos?

Solution:

Let X represent the number of prizes won if 3 combos are purchased.

In this case, X~Binomial(n = 3, p = 1/5)

and f(x) = P(X = x) = $\binom{3}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{3-x}$**, for x = 0, 1, 2, 3; 0, otherwise.**

We want $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{3}{0}\left(\frac{1}{5}\right)^{0}\left(\frac{4}{5}\right)^{3}$

$P(X \geq 1) = \frac{61}{125} = 0.488$

Problem 3 Solution (Cont'd)

c) How many game pieces are needed so that a customer has at least a 95% chance of winning at least one prize? (Take a guess before this calculation is done!)

Solution:

Let X represent the number of prizes won if n combos are purchased.

In this case, X~Binomial(n, p = 1/5), and f(x) = P(X = x) = $\binom{n}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{n-x}$**, for x = 0, 1, 2, …., n; 0, otherwise.**

In this case, we want to solve for n such that P(X $\geq$ 1) $\geq$ 0.95.

From part b), we know that P(X $\geq$ 1) = 1 – P(X = 0).

We want P(X $\geq$ 1) = 1 – P(X = 0) $\geq$ 0.95 => 1 - $\binom{n}{0}\left(\frac{1}{5}\right)^{0}\left(\frac{4}{5}\right)^{n} \geq 0.95.$

We know that $\binom{n}{0}$ and $\left(\frac{1}{5}\right)^{0}$ **are each 1. This simplifies the problem.**

So, we have $\left(\frac{4}{5}\right)^{n} \leq 0.05.$ **Taking natural logs (i.e. ln) of both sides, this gives: n** $\geq \frac{\log(0.05)}{\log(\frac{4}{5})} = 13.425.$

Note that the sign changes, as we are dividing by log(4/5), which is a negative number.

So, we see that a minimum of 14 game pieces (combos) are needed.

Note: Don't empty out your bank account just yet! We don't know which prize you will win. We are just claiming that you will win at least one prize. It could be a side order or a drink!

Negative Binomial Distribution

Suppose that we repeat independent Bernoulli trials consisting of two outcomes:

S (success) and F (failure), and P(S) = p in each trial, with $0 < p < 1$.

We are interested in the number of failures obtained before the k^{th} success. In this case, X has a Negative Binomial distribution. We can write $X \sim NB(k, p)$.

The pf of X is given by:

$f(x) = P(X = x) = \binom{x+k-1}{x} p^k (1-p)^x$, for x = 0, 1, 2, ……. ; 0, otherwise.

Note: Take a minute and think about the range of the random variable. Remember that X represents the number of failures before the k^{th} success.

Some Comparisons

Binomial and Negative Binomial Distribution Comparisons

These two distributions should be easily distinguishable. They reverse what is specified or known in advance, and what is variable.

Binomial: The number of independent Bernoulli trials is known. However, we don't know the number of successes until the experiment has been conducted.

Negative Binomial: The number of successes is known in advance. This is k. However, we do not know the number of trials needed to obtain the k successes until the experiment has been conducted.

Problem 4

Suppose that there are two cameras that are identical in appearance. The flash mechanism fails on each camera from time to time though.

In fact, on any given attempt to take a picture, the flash mechanism on camera A fails with probability 0.1, while the mechanism on camera B fails with probability 0.05.

a) Suppose that camera A is selected and the photographer uses it to attempt to take pictures. What is the probability that the flash mechanism works successfully for the second time on the fifth attempt to take a picture?

b) Now suppose that the photographer randomly selects one of the cameras. What is the probability that the photographer needs exactly 10 attempts in order to have exactly 8 of them with the flash working properly? Assume that the 8^{th} came on their 10^{th} attempt.

c) Given that the photographer needed to take exactly 10 attempts in order to have exactly 8 of them with the flash working properly, what is the probability that camera A was used?

Problem 4 Solution

Suppose that there are two cameras that are identical in appearance. The flash mechanism fails on each camera from time to time though.

In fact, the flash mechanism on camera A fails with probability 0.1, while the mechanism on camera B fails with probability 0.05.

a) Suppose that camera A is selected and the photographer uses it to take pictures. What is the probability that the flash mechanism works successfully for the second time on the fifth attempt to take a picture?

Solution:

Let X represent the number of failures before the second (i.e. k = 2) success.

In this case, X~NB(k=2, p = 0.9).

So, f(x) = P(X = x) = $\binom{x+2-1}{x} p^2(1-p)^x$**, for x = 0, 1, 2, ……. ; 0, otherwise.**

So, to "visualize" this, in the 5 total attempts, there were 3 failures, and 1 success in the first 4 attempts, followed by a success on the 5th attempt.

So, we want f(3) = P(X = 3) = $\binom{4}{3} 0.9^2(1-0.9)^3 = 0.00324$

Problem 4 Solution (Cont'd)

b) Now suppose that the photographer randomly selects one of the cameras. What is the probability that the photographer needs exactly 10 attempts in order to have exactly 8 of them with the flash working properly? Assume that the 8^{th} came on their 10^{th} attempt.

Solution:

In order to calculate this probability, we will need to take the selection of each camera into consideration.

Let X_A represent the number of failures before the 8^{th} success assuming camera A is selected.

Let X_B represent the number of failures before the 8^{th} success assuming camera B is selected.

Here X_A ~ NB(k = 8, p = 0.9) and X_B ~ NB(k = 8, p = 0.95).

For each camera, there were 9 previous attempts, with 2 failures and 7 successes, then a success on the 10^{th} attempt.

Each camera is equally likely to be selected, so, the required probability is given by:

$$\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\binom{9}{2}(\mathbf{0.1})^{\mathbf{2}}(\mathbf{0.9})^{\mathbf{8}}+\left(\frac{\mathbf{1}}{\mathbf{2}}\right)\binom{9}{2}(\mathbf{0.05})^{\mathbf{2}}(\mathbf{0.95})^{\mathbf{8}}$$

We'll leave the answer in this form for now. We will need to make use of this in the next part.

Problem 4 Solution (Cont'd)

c) Given that the photographer needed to take exactly 10 attempts in order to have exactly 8 of them with the flash working properly, what is the probability that camera A was used?

Solution:

Let A be the event that camera A was used.

Let B be the event that camera B was used. (B is also the same as $\overline{A}$ in this case)

Let C be the event that the photographer needed to take exactly 10 pictures in order to have exactly 8 of them with the flash working properly.

We want P(A|C).

P(A|C) = $\frac{P(A \cap C)}{P(C)}$

We actually calculated P(C) in part b). In this case, P(A ∩ C) = P(A)*P(C|A) = $\left(\frac{1}{2}\right)\binom{9}{2}(0.1)^2(0.9)^8$.

So, P(A|C) = $\frac{P(A \cap C)}{P(C)} = \frac{\left(\frac{1}{2}\right)\binom{9}{2}(0.1)^2(0.9)^8}{\left(\frac{1}{2}\right)\binom{9}{2}(0.1)^2(0.9)^8+\left(\frac{1}{2}\right)\binom{9}{2}(0.05)^2(0.95)^8} = \frac{(0.1)^2(0.9)^8}{(0.1)^2(0.9)^8+(0.05)^2(0.95)^8} = 0.7219$

Note: You may have recognized that this was a Bayes' Rule application!

Problem 5

Recall from Problem 2: Fifteen televisions are to be purchased by a local sports bar from a large production run of 500.

Suppose that 450 will last at least 5 years without needing repair.

Now, let's use the appropriate distribution to **approximate** the probability that at least 14 of the 15 televisions will last at least five years without needing to be repaired.

Problem 5 Solution

Recall from Problem 2: Fifteen televisions are to be purchased by a local sports bar from a large production run of 500.

Suppose that 450 will last at least 5 years without needing repair.

Now, let's find the **approximate** probability that at least 14 of the 15 televisions will last at least five years without needing to be repaired.

Solution:

Here, we let X represent the number of TVs in the sample of 15 that will last at least five years without needing to be repaired.

Let's see if it's appropriate to use the Binomial approximation to the Hypergeometric. Look at N/n = 500/15 = 33.333 (at least 20). Check!

The approximation should work well. Here, we assume X is approximately Binomial(n=15, p = 450/500 = 0.9).

The pf of X, f(x), is given by:

f(x) = P(X = x) = $\binom{15}{x}(0.9)^{x}(0.1)^{15-x}$**, for x = 0, 1, 2, …, 15 ; 0, otherwise.**

We want P(X ≥ 14) = P(X = 14) + P(X = 15) = $\binom{15}{14}(0.9)^{14}(0.1)^{15-14}+\binom{15}{15}(0.9)^{15}(0.1)^{15-15}=0.549043$

Remember that the EXACT probability was 0.54688. So the approximation worked quite well, as expected!

Name That Distribution!

Suppose that we are drawing cards from a standard 52 card deck with replacement. That is, the cards are shuffled well after each draw.

We want to calculate the following probabilities. What distribution would be used?

a) If 4 cards are drawn, what is the probability of drawing two kings?

b) What is the probability that the third club will be drawn on the 6^{th} draw?

c) Now, assuming we draw cards <u>without replacement</u>, what is the probability of drawing three hearts in 5 draws?

Solution to Name That Distribution!

Suppose that we are drawing cards from a standard 52 card deck with replacement. That is, the cards are shuffled well after each draw.

We want to calculate the following probabilities. What distribution would be used?

a) If 4 cards are drawn, what is the probability of drawing two kings?

Solution: A Binomial distribution with $n = 4$, and $p = 4/52 = 1/13$ would be used. We would calculate $f(2)$.

b) What is the probability that the third club will be drawn on the 6^{th} draw?

Solution: A Negative Binomial distribution would be used with $k = 3$ and $p = 13/52 = 1/4$. We would calculate $f(3)$.

c) Now, assuming we draw cards <u>without replacement</u>, what is the probability of drawing three hearts in 5 draws?

Solution: A Hypergeometric distribution would be used with $N = 52$, $r = 13$, and $n = 5$. We would calculate $f(3)$.