Midterm - S520

Keith Hickman

October 2, 2017

## R Markdown

**Problem 1** I toss six fair coins.

*a. What is the probability I get exactly five heads?*

We're looking for the probability of choosing five outcomes from 6 possibilities, or , and we are choosing exactly that number of outcomes divided by possible outcomes.

We can solve with a fraction, or with the dbinom function in R.

and thus or slightly less than 10%

To solve/confirm in R:

A <- dbinom(5,6,.5)  
A

## [1] 0.09375

*b. If I get at least five heads(B), what is the conditional probability I get exactly five heads(A)??*

The formula for conditional probability is . We know the probability of obtaining exactly five heads (.0938), which is event .

To solve for the probability of obtaining at least 5 heads or event , which is the same as the probabilily of obtaining 5 or 6 heads, , we can use from problem a and add to it which is the same as 1-pbinom(4, 6, .5) = 0.109375.

B <- 1-pbinom(4, 6, .5)  
B

## [1] 0.109375

Because set A is completely contained within Set B (A intersect B is the same as A), we can solve for the conditional probability by dividing

A/B

## [1] 0.8571429

Thus there is an 85.7% chance that I get exactly five heads if I get at least five heads.

*c. Are the events "I get exactly five heads" and "I get at least five heads" independent?* No. Here the events are independent because getting at least five heads (B) does not depend on getting exactly five heads (A) and vice versa.

**Problem 2:** 48% percent of U.S. adults have visited a library in the past year. I select two U.S. adults at random. Let X be a discrete random variable representing the number of the adults I selected who have visited a library in the last year.

*a) Find*  This is another binomial variable problem in that there are two outcomes in the underlying sample space, either 1:has visited, or 0:has not visited. Our sample space , and . Thus, we can test for P(X) by using the dbinom function in R for each X value:

dbinom(0,2,.48)

## [1] 0.2704

dbinom(1,2,.48)

## [1] 0.4992

dbinom(2,2,.48)

## [1] 0.2304

*b) Write down an expression for , the cumulative distribution function (CDF) of X, for all y-values from to .*

*c) Find the expected value and variance of X.*

or we can use the sum of the weighted values:

The Variance of X = or

**Problem 3**

Let X be a continuous random variable with the probability density function

*a) Find the cdf for all values from negative infinity to infinity*:

First, we can verify that X is a pdf because all values for every and the area under the graph = 1. and .

The cdf of X:

*b) Find the median.* **The Median of is 5**.

The median is the of X, or the quantile of X. We can use geometry to solve by using the area under the graph to find the quantile. The area of a region between 0 and of function x and above the horizontal axis is .5. Since the region between [0,8] is a rectangle with a base of 8 and ht of .1, the area is .8, thus lies on the interval [0,8]. The length of the rectangle's base that gives us is as, .

*c) Find the expected value* **The Expected Value of X is 4.8**.

We can solve using the weighted averages of the centers of the two blocks from the pdf: and the center of [0,7] = 3.5 and the center of [8,12] = 10 Thus, we have

**Problem 4**

*a) Is X discrete, continuous, or neither? Explain.*

X is a continuous variable as the set of values X is not countable, and the values lie on the interval [0,4].

*b) Find the standard deviation of X.*

The standard deviation is equal to . Var(x) = . Plugging in and =

*(I know the notation on isn't correct)*

in R:

EX2 <- (49/24)^2  
EXsq <- 48/8  
VarX <- EXsq - EX2  
StdDev <- sqrt(VarX)  
VarX

## [1] 1.831597

StdDev

## [1] 1.353365

*c) Let A and B be independent random variables with the same distribution as X. Let Y = A - B. What are the expected value and standard deviation of Y?*

Here, both EX and Standard Deviation of Y are 0. Since and are independent standard normal distributions, we can subtract the expected values and standard deviations of as follows:

The formula for standard deviation is similar:

**Problem 5**

1. **The median of X is 1.41**. of X equals the .5 quantile of X. Looking at the cdf of X, the .5 quantile will fall obviously somewhere between 0 and 1, or we can say that .
2. **The IQR of X is .85.** The IQR is or of X.

Solve for .75 or q3:

Solve for .25 or q1

Solve for

**The IQR for X = .85**

**Problem 6**

Let X be standard normal random variable. Let .

1. What is the median of Y?

**The median of Y is .383**. We need to find or . This is the difference between the cdf at -.5 and .5. In R:

pnorm(.5) - pnorm(-.5)

## [1] 0.3829249

1. What is ? This is the difference of the distribution CDF at 2 and at 1.

pnorm(2) - pnorm(1)

## [1] 0.1359051

Or 13.6%

1. Find the .95 quantile of Y. Since -y goes to infinity, we can drop that as impossible to calculate. We're looking then for We can use qnorm which is the inverse of the cdf.

qnorm(.975)

## [1] 1.959964

**Problem 7**

1) *I select an adult U.S. woman at random. What is the probability she is under five feet (152.4 cm) tall, or solve for*

We can solve with the pnorm function in R by passing in the value, mean, and standard deviation of the variable. The pnorm function is used for finding the quantile of values from negative infinity up to X.

pnorm(152.4, 162.1, 7.3)

## [1] 0.09196241

There is a **9.2%** chance of selecting a woman under five feet tall.

1. *I select an adult U.S. man at random. What is the probability he is over six feet (182.88 cm) tall?*

Here, since we're interested in , we need to subtract from 1.

prob6 <- 1 - pnorm(182,88, 175.9, 7.5)  
prob6

## [1] 0.2965343

Thus, there is a **29.6%** chance that a given male will be taller than 6 feet.

1. *I select four adult U.S. men at random. What is the probability at least one of them is over six feet tall?*

We have a binomial distribution (either is taller than 6'(:1) or is not (:0)). Here, we can pass in the probability that a given man is taller than 6' obtained using pnorm above (prob6), as the probability that one out of the four males selected will be taller than six feet using the pbinom function in R. If is "at least one of four males is taller than six feet" then is the probability that none of the four men is taller than six feet. The probability of is pbinom(0,4,prob6). Thus we can subtract that value from 1 to obtain the :

1- pbinom(0, 4, prob6)

## [1] 0.7551096

There is a **75.5%** chance that one of the four men selected will be taller than 6 feet.

1. I select at random an adult U.S. man and, independently, an adult U.S. woman. What is the probability the woman is taller?

Another way of wording this problem is to say, what's the probability I select a woman from a quantile of *and* that the quantile I select the female from is greater than the quantile of for the man I select.

The difference between the expected value of the two variables is or , or

1 - pnorm(175.9, 162.7, 7.3)

## [1] 0.0352862

There is a **3%** chance that the woman I select will be taller than than the man I select.

**Problem 8**

*1) Suppose this test is taken by a student who does not have psychic powers. What is the probability she gets 15 or more correct?*

Here, we have a binomial distribution, and are seeking the . For this we can use dbinom in R:

d <- dbinom(15, 20, .5)  
d

## [1] 0.01478577

1. *Suppose this test is taken by eighty students who do NOT have psychic powers. What is the probability at least one student gets 15 or more correct?*

We can use the previously-established probability that one person gets at least 15 questions corrected (1%) as the probability for our second dbinom function. Here, we want to know the probability that at least one student out of 80 will get at least 15 answers correct, or

dbinom(1, 80, d)

## [1] 0.3646342

There is a 36.4% probability that at least one student out of 80 will get 15 or more answers correct.

*This was a thoroughly, enjoyably difficult midterm - thanks!!*