S&DS 220: Homework 5

Due Friday February 16

Instructions

1. Complete the questions below. Upload your knitted PDF solutions to Gradescope by the due date.
2. Your solutions should be a combination of writing and R code. When writing, use complete sentences.
3. Previous homework assignments already had code chunks created for you. Now it is up to you to insert R code chunks within each problem as needed.
4. You should aim for clear and concise communication (in both words and R code).

Problem set questions

Question 1: (Exercise 3.1, 3.5, 3.23) PMFs, mean, and variance

Let X be a random variable with probability mass function given by

$$p(x) = \begin{cases} 1/4, & x = 0, \\ 1/2, & x = 1, \\ 1/8, & x = 2, \\ 1/8, & x = 3 \end{cases}.$$

Answer the following (without simulation).

(a) Verify that p is a valid probability mass function.

Solution.

(b) Find $P(X \geq 2)$.

Solution.

(c) Find $P(X \geq 2 | X \geq 1)$.

Solution.

(d) Find $P(X \geq 2 \cup X \geq 1)$.

Solution.

(e) Find the mean of X.

Solution.

(f) Find the variance and standard deviation of X.

Solution.

Question 2: (Exercise 3.39) Detecting cheating in video games with the binomial distribution

In October 2020, the YouTuber called "Dream" posted a speedrun of Minecraft and was accused of cheating. Answer the following (without simulation).

In Minecraft, when you trade with a piglin, the piglin gives you an ender pearl 4.7% of the time. Dream got 42 ender pearls after 262 trades with piglin.

Answer the following (without simulation). Recall that `pbinom` is the cdf and `dbinom` is the pmf of the binomial distribution in R.

(a) If you trade 262 times, what is the expected number of ender pearls you receive?

Solution.

(b) What is the probability of getting 42 or more ender pearls after 262 trades?

Solution.

When you kill a blaze, you have a 50% chance of getting a blaze rod. Dream got 211 blaze rods after killing 305 blazes.

(c) If you kill 305 blazes, what is the expected number of blaze rods you receive?

Solution.

(d) What is the probability of getting 211 or more blaze rods after killing 305 blazes?

Solution.

(e) Do you think Dream was cheating?

Solution.

Question 3: (Exercise 3.31, with some tweaks) Simulation and Poisson approximation

Suppose 27 people write their names down on slips of paper and put them in a hat. Each person then draws one name from the hat. Let N be the number of people who draw their own name (assuming no two people have the same name).

(a) Estimate the expected value and standard deviation of N.

Solution.

(b) The pmf of N can estimated from your simulation from part (a) using `proportions(table())`. Print this pmf and make a plot of the pmf using `plot()` with the argument `type = "h"`.

Solution.

(c) We can also approximate the distribution of N using a Poisson distribution. Since each person has a $1/27$ chance of drawing their name from the hat and all 27 people draw a name, the mean of N is $E(N) = 27 \cdot 1/27 = 1$. And so we can approximate N with a Poisson(1) distribution. Using `dpois`, give the values of the pmf of a Poisson(1) for `x = 0:8` (if the values are given in scientific notation, you can get rid of this by using `round()` with the argument `digits = 4`). Plot this pmf as in (b) using `plot()` with `type = "h"`.

Solution.

(d) Compare these values with the estimated pmf from (b). Do you think the Poisson approximation of N is good?

Solution.

Question 4: (Exercise 2.19) Scrabble

In the game of Scrabble, players make words using letter tiles. The data set `fosdata::scrabble` contains all 100 tiles.

Players begin the game by drawing seven tiles from a bag of 100 tiles. Estimate the probability that a player's first seven tiles contain no vowels. (Vowels are A, E, I, O, and U.)

Solution.

Question 5: (Exercise 3.39) More Scrabble!

In the game of Scrabble, players make words using letter tiles, see Exercise 2.19. The tiles consist of 42 vowels and 58 non-vowels (including blanks).

Hint: For sampling without replacement, see the Hypergeometric distribution in Section 3.6.3.

(a) If a player draws 7 tiles (without replacement), what is the probability of getting 7 vowels?

Solution.

(b) If a player draws 7 tiles (without replacement), what is the probability of 2 or fewer vowels?

Solution.

(c) What is the expected number of vowels drawn when drawing 7 tiles?

Solution.

(d) What is the standard deviation of the number of vowels drawn when drawing 7 tiles?

Solution.

Question 6: (Exercise 3.40 with some tweaks) Simulating deathrolling in World of Warcraft

Deathrolling in World of Warcraft works as follows. Player 1 tosses a 1000-sided die. Say they get x_1. Then player 2 tosses a die with x_1 sides on it. Say they get x_2. Player 1 tosses a die with x_2 sides on it. The player who loses is the player who first rolls a 1.

Coding hint: a `while` may be helpful in this exercise.

(a) Estimate the expected total number of rolls before a player loses.

Solution.

(b) Estimate the probability mass function of the total number of rolls. You can use `proportions(table( ))` to give your answer.

Solution.

(c) Plot the estimated probability mass function from (b), using `plot()` with the argument `type = "h"`.

Solution.

(d) Estimate the probability that player 1 wins.

Solution.