18.05 Problem Set 2, Spring 2022

Problem 1. (20: 10,10 pts.) **('Binary' paradox)**
For this problem assume that puppies are equally probable to be male or female. Likewise for kittens.

Be sure to carefully justify your answers.

(a) Our dog Layla had two puppies. The older puppy is female. What is the probability that both puppies are female?

(b) Our cat Ariel had two kittens. At least one of them is male. What is the probability that both kittens are males?

Problem 2. (15 pts.) **(The blue taxi)**
In a city with one hundred taxis, 1 is blue and 99 are green. A witness observes a hit-and-run by a taxi at night and recalls that the taxi was blue, so the police arrest the blue taxi driver who was on duty that night. The driver proclaims their innocence and hires you to defend them in court. You hire a scientist to test the witness' ability to distinguish blue and green taxis under conditions similar to the night of accident. The data suggests that the witness sees blue cars as blue 99% of the time and green cars as blue 2% of the time.

Write a speech for the jury to give them reasonable doubt about your client's guilt. Your speech need not be longer than the statement of this question. Keep in mind that most jurors have not taken this course, so an illustrative table may be easier for them to understand than fancy formulas.

Problem 3. (20 pts.) **(Trees of cards)** There are 8 cards in a hat:

$$\{1\heartsuit, 1\spadesuit, 1\diamondsuit, 1\clubsuit, 2\heartsuit, 2\spadesuit, 2\diamondsuit, 2\clubsuit\}$$

You draw one card at random. If its rank is 1 you draw one more card; if its rank is two you draw two more cards. Let X be the sum of the ranks on the 2 or 3 cards drawn. Find $E[X]$. (Note: all the draws are done without replacement.)

Problem 4. (25: 5,10,5,5 pts.) **(Dice)** There are four dice in a drawer: one D4 (4 sides), one D6 (6-sides), and two D8 (8 sides). As usual, the sides of a die are numbered 1 to n, where n is the number of sides.

Your friend secretly grabs one of the four dice at random. Let S be the number of sides on the chosen die.

(a) What is the pmf of S?

(b) Now, without showing it to you, your friend rolls the chosen die and tells you the result. Let R be the result of the roll.

Use Bayes' rule to find $P(S = s \mid R = 3)$ for $s = 4, 6, 8$. Which die is most likely if $R = 3$? Terminology: You are computing the pmf of 'S given $R = 3$'.

(c) Which die is most likely if $R = 6$? Hint: You can either repeat the computation in (b), or you can reason based on your result in (b).

(d) Which die is most likely if $R = 7$? No computations are needed!

Problem 5. (10 pts.) **(Seating arrangement and relative height)** A total of n people randomly take their seats around a circular table with n chairs. No two people have the same height. What is the expected number of people who are shorter than both of their immediate neighbors?

Problem 6. (20: 3,2,10,5 pts.)
In this problem we will use R to simulate flipping a fair coin 50 times. We'll use the simulation to explore 'runs'. That is, sequences of all 1's or 0's.

(a) Make up a sequence of length 50 consisting of ones and zeros. Try to make the sequence look like it was randomly generated by flipping a coin.

(b) A run is a sequence of all 1s or all 0s. How long is the longest run in your answer to part (a)?

(c) Now we will use R to simulate 50 tosses of a fair coin and estimate the average length of the longest run. The code below simulates one trial. You will need to use a 'for loop' to run 10000 trials. On our MITx site you can find tutorials on both for loops and the rle() function used in the code. (As usual, choose the Course info tab and click on the R code link under course handouts).

Sample code illustrating the use of `rle()`.

```
# R code to simulate 50 flips of a fair coin and find the longest run
# rle stands for `run length encoding'.
# rle(trial)$lengths is a vector of the lengths of all the different runs
# in trial.
nflips = 50
trial = rbinom(nflips, 1, 0.5) # Note: binomial(1, 0.5) = bernoulli(0.5)
max_run = max(rle(trial)$lengths)
```

Sample code illustrating 'for loops'. (These may be useful in the code for this problem.)

```
# R code demonstrating `for' loops.
for (j in 1:5) {
  print(j^2)
}
```

(Should produce: `1, 4, 9, 16, 25.`)

```
sum = 0
for (j in 1:5) {
  sum = sum + j
}
print(sum)
```

(Should produce: `15.`)

Use R to simulate the average length of the longest run in 50 flips of a fair coin. Do this three times with 10000 trials each time and report the three results.

(d) A small modification of your code will let you estimate the probability of a run of 8 or more in 50 flips. Do this three times with 10000 trials each time. Report the three results.