18.05 Problem Set 2, Spring 2022 Solutions

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?*

Solution: Listing the sex of older puppy first, our sample space is $\{MM, MF, FM, FF\}$. The event "the older puppy is female" is $\{FM, FF\}$ and the event "both puppies are female" is $\{FF\}$. Thus the probability that both puppies are females given the older is female is

$$P(FF|\{FM, FF\}) = \frac{P(FF)}{P(\{FM, FF\})} = \frac{1/4}{1/2} = \frac{1}{2}.$$

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

Solution: The event "at least one kitten is male" is $\{MM, MF, FM\}$ so the probability that both kittens are male is

$$P(MM|\{MM, MF, FM\}) = \frac{P(MM)}{P(\{MM, MF, FM\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

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.

Solution: The defense will try to make the case that this is likely to be a case of random mis-identification. So we look for the probability a random taxi the witness sees as blue is actually blue.

This is a question of 'inverting' conditional probability. We know

$$P(\text{the witness sees blue}|\text{the car is blue})$$

but we'd like to know

$$P(\text{the car is blue}|\text{the witness sees blue}).$$

Our first job is to translate this to symbols.

Let W_b = 'witness sees a blue taxi' and let W_g = 'witness sees a green car'. Further, let T_b = 'taxi is blue' and let T_g = 'taxi is green'. With this notation we want to find $P(T_b|W_b)$. We will compute this using Bayes' formula

$$P(T_b|W_b) = \frac{P(W_b|T_b) \cdot P(T_b)}{P(W_b)}.$$

All the pieces are represented in the following diagram.
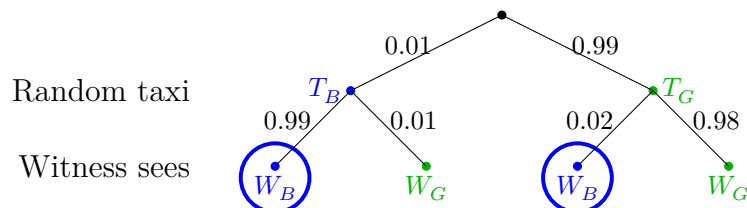
We can determine each factor in the right side of Bayes' formula:
We are given $P(T_b) = 0.01$ (so $P(T_g) = 0.99$).
We are given, $P(W_b|T_b) = 0.99$ and $P(W_b|T_g) = 0.02$.
We compute $P(W_b)$ using the law of total probability:

$$P(W_b) = P(W_b|T_b)P(T_b) + P(W_b|T_g)P(T_g) = 0.99 \times 0.01 + 0.02 \times 0.99 = 0.99 \times 0.03.$$

Putting all this in Bayes' formula we get

$$P(T_b|W_b) = \frac{0.99 \times 0.01}{0.99 \times 0.03} = \boxed{\frac{1}{3}}$$

Our speech:
Ladies and gentlemen of the jury. The prosecutor tells you that the witness is nearly flawless in their ability to distinguish whether a taxi is green or blue. They claims that this implies that beyond a reasonable doubt the taxi involved in the hit and run was blue. However probability theory shows without any doubt that a random taxi which the witness sees as blue is actually blue only 1/3 of the time. This is considerably more than a reasonable doubt. In fact it is more probable than not that the taxi involved in the accident was green. If the probability doesn't fit you must acquit!

Note: Confusing of $P(\text{testimony}|\text{guilty})$ for $P(\text{guilty}|\text{testimony})$ is known as the prosecutor's fallacy

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.)

Solution: The following tree shows all the possible ways we can draw cards. The first draw is the top row of nodes. The value of the draw is given in the circle. The probability of that draw is given along the edge. Likewise for the second and third draws. At the end of each path we give the value of X resulting from those draws.
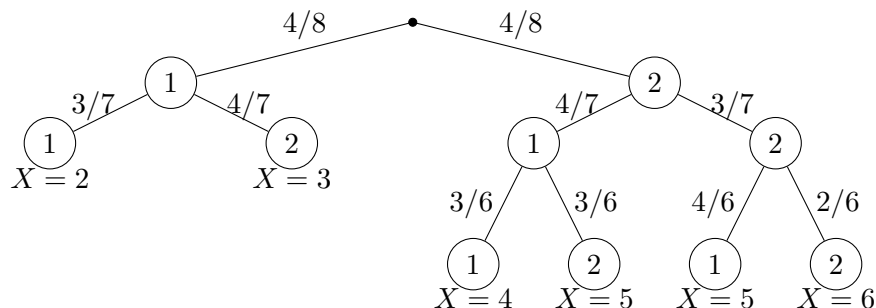
This gives us the probability distribution of X:

value x	2	3	4	5	6
pmf $p(x)$	3/14	2/7	1/7	2/7	1/14

So,

$$E[X] = 2 \cdot \frac{3}{14} + 3 \cdot \frac{2}{7} + 4 \cdot \frac{1}{7} + 5 \cdot \frac{2}{7} + 6 \cdot \frac{1}{14} = \frac{26}{7} \approx 3.7143.$$

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?*

Solution: Here is the pmf:

| value s	4	6	8
pmf $p(s)$	1/4	1/4	1/2

(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 $R = 3$'.

Solution: Bayes' rule says

$$P(S = s | R = 3) = \frac{P(R = 3 \mid S = s)P(S = s)}{P(R = 3)}.$$

We summarize what we know in a tree. In the tree the notation D_4 means the 4-sided die ($S = 4$), likewise R_3 means a 3 was rolled ($R = 3$). Because we only care about the case $R = 3$ the tree does not include other possible rolls.
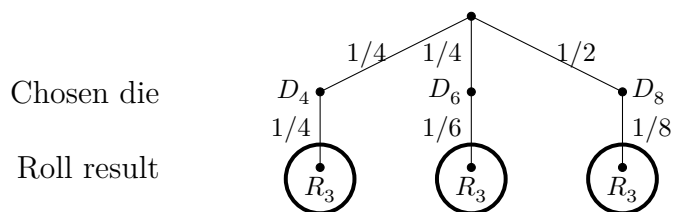
We have the following probabilities (you should identify them in the tree):

$$P(R = 3 \,|\, S = 4) = 1/4, \quad P(R = 3 \,|\, S = 6) = 1/6, \quad P(R = 3 \,|\, S = 8) = 1/8.$$

The law of total probability gives (again, see how the tree tells us this):

$$
\begin{aligned}
P(R = 3) &= P(R = 3 \,|\, S = 4)P(S = 4) + P(R = 3 \,|\, S = 6)P(S = 6) + P(R = 3 \,|\, S = 8)P(S = 8) \\
&= \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{6} \cdot \frac{1}{4} + \frac{1}{8} \cdot \frac{1}{2} \\
&= \frac{1}{6}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
P(S = 4 \,|\, R = 3) &= \frac{P(R = 3 \,|\, S = 4) \cdot P(S = 4)}{P(R = 3)} = \frac{(1/4) \cdot (1/4)}{1/6} = \boxed{\frac{3}{8}} \\
P(S = 6 \,|\, R = 3) &= \frac{P(R = 3 \,|\, S = 6) \cdot P(S = 6)}{P(R = 3)} = \frac{(1/6) \cdot (1/4)}{1/6} = \boxed{\frac{1}{4}} \\
P(S = 8 \,|\, R = 3) &= \frac{P(R = 3 \,|\, S = 8) \cdot P(S = 8)}{P(R = 3)} = \frac{(1/8) \cdot (1/2)}{1/6} = \boxed{\frac{3}{8}}.
\end{aligned}
$$

We see that the 4 and 8 sided dice are equally most likely 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).*

Solution: In a similar vein, we have

$$P(R = 6 | S = 4) = 0, \quad P(R = 6 | S = 6) = 1/6, \quad P(R = 6 | S = 8) = 1/8.$$

and

$$P(R = 6) = 0 \cdot \frac{1}{4} + \frac{1}{6} \cdot \frac{1}{4} + \frac{1}{8} \cdot \frac{1}{2} = \frac{5}{48}.$$

So,

$$P(S = 4 | R = 6) = 0, \quad P(S = 6 | R = 6) = \frac{1/24}{5/48} = \frac{2}{5}, \quad P(S = 8 | R = 6) = \frac{1/16}{5/48} = \frac{3}{5}.$$

The eight-sided die is more likely. Note, the denominator is the same in each probability, i.e. the total probability $P(R = 6)$, so all we really had to check was the numerator.

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

Solution: The only way to get $R = 7$ is if we picked an 8-sided die D8.

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?*

Solution: Label the seats 1 to n going clockwise around the table. Let X_i be the Bernoulli random variable with value 1 if the person in seat i is shorter than their neighbors. Then

$X = \sum_{i=1}^{n} X_i$ represents the total number of people who are shorter than both of their neighbors, and, by linearity of expected value,

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

Recall that this property of expected values holds even when the X_i are dependent, as is the case here!

Among 3 random people the probability that the middle one is the shortest is $1/3$. Therefore $X_i \sim \text{Bernoulli}(1/3)$, which implies $E[X_i] = 1/3$. Therefore the expected number of people shorter than both their neighbors is

$$E[X] = \sum_{i=1}^{n} E[X_i] = \frac{n}{3}.$$

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.*

Solution: Any sequence of 50 zeros and ones is valid.

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

Solution: Answers will vary because we can't know the sequence you chose. However, most people, when trying to be random, do not put in any long runs. Parts (c) and (d) show these actually happen frequently.

(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.

Solution: Here is my code with comments

```
nflips = 50
ntrials = 10000
total = 0 # We'll keep a running total of all the trials' longest runs
for (j in 1:ntrials) {
  # One trial consists of 50 flips
  trial = rbinom(nflips, 1, 0.5) # binomial(1, 0.5) = bernoulli(0.5)
  # rle() finds the lengths of all the runs in trials. We add the max to total
  total = total + max(rle(trial)$lengths)
}
# The average maximum run is the total/ntrials
ave_max = total/ntrials
print(ave_max)
```

My three runs of this code produced ave_max = 5.9863, 5.9696, 5.9804

(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.*

Solution: Instead of keeping a total we keep a count of the number of trials with a run of 8 or more

```
nflips = 50
ntrials = 10000
# We'll keep count of all the trials with a run of 8 or more
count = 0
for (j in 1:ntrials) {
  trial = rbinom(nflips, 1, 0.5) # binomial(1, 0.5) = bernoulli(0.5)
  count = count + (max(rle(trial)$lengths) >= 8)
}
# The probability of a run of 8 or more is count/ntrials
prob8 = count/ntrials
print(prob8)
```

My run of this code produced prob8 = 0.1596, 0.1637, 0.1632