ECS 132 Fall 2020: Assignment 1 - Total: 8 pts

October 10, 2020

Instructions

1. You may in no circumstances upload your homework to private tutoring websites such as CourseHero or Chegg. Remember all material related to this course is a property of the University of California and posting them is a violation of the copyright laws.
2. If you refer to a source (either a book or the internet), you must cite it.
3. You are highly urged to work on these problems on your own. See the homework grading policy. Not getting the answer correct has very low penalty. However, trying it and then figuring out where you went wrong will really help you understand the material better and you will be much better prepared for the exams. If you do discuss with others, you must list their names.
4. For what to submit (rmarkdown), due date, and where to submit (Gradescope), see the Homework Policy and Schedule.

1 Problem

Consider the Manhattan grid network shown in Figure 1. Suppose that starting at the point labelled A, you can go one step up or one step to the right at each move. This procedure is is continued until the point labelled B is reached. How many different paths from A to B are possible?

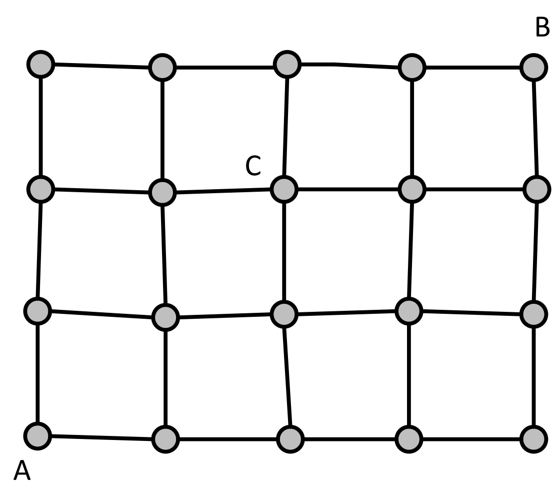

Figure 1: The Manhattan grid network.

1.1 Answer

Each path from A to B is a permutation of 7 steps consisting of 4 Rs (rights) and 3 Us (ups). The total number of ways is $\binom{7}{3}$. Imagine there are 7 positions. How many ways can you select 4 of them as R's. Note that order is not relevant. If the R's are labelled R_1, R_2, R_3, then only 1 order is relevant. Once the R's are selected, the Us are automatically selected. Also, note that $\binom{n}{k}$ is the same as $\binom{n}{n-k}$.

1.2 Points - Total: 1pt

- 1/2 pt for selecting correctly identifying that there are 7 possible moves and order does not matter.
- 1/2 pt for selecting the correct equation $\binom{n}{k}$: $\binom{7}{3(or4)}$ and finding the correct answer

2 Problem

Consider a pool of six I/O buffers. Assume that any buffer is just as likely to be available (or occupied) as any other. Compute the probabilities associated with the following events

1. A = "A least two but no more than five buffers are occupied"
2. B = "At least one buffer is occupied"

2.1 Answer

There are 6 I/O buffers and each o f them can be in one of two states available (0) and occupied (1). The state space S consist of 2^6 possible outcomes of the I/O buffer pool. For example $(0, 0, 0, 0, 0, 0)$ corresponds to the outcome that all 6 I/O buffers are available. Similarly, $(0, 1, 0, 0, 0, 1)$ correspond to the outcome that 2 I/O buffers (specifically 2 and 6 in this case) are occupied.

1. For event A we are looking for the outcomes in which 2, 3, 4, and 5 I/O buffers are occupied. Number of outcomes with i I/O buffers are occupied is $\binom{6}{i}$. Thus,

$$P(A) = \frac{\binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5}}{2^6} \tag{1}$$

2. For the event B, we can follow the same logic as before and

$$P(B) = \frac{\binom{6}{1} + \binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + + \binom{6}{6}}{2^6} \tag{2}$$

The other way think about is to consider the complement of event B denoted as $B^{\complement}$ which is 0 I/O buffers are available

$$\begin{aligned} P(B) &= 1 - P(B^{\complement}) && (3)\\ &= 1 - \frac{\binom{6}{0}}{2^6} && (4)\\ &= 1 - \frac{1}{2^6} && (5) \end{aligned}$$

2.2 Points - Total 2pts

2.2.1 A

- 1/2 pt for properly identifying and writing the equations for the number of occupied buffers in each case (each of the equations in the numerator in A)
- 1/2 pt for properly setting up the fraction in A

2.2.2 B

- 1/2 pt for properly identifying and writing the equations for the number of occupied buffers in each case (each of the equations in the numerator in B)
- 1/2 pt for properly setting up the fraction in B

3 Problem

Suppose your MP3 player contains 100 songs. 10 of those songs are by your favorite group (say Pink Floyd one of my favorite bands). Suppose the shuffle feature is used to play the songs in random order. Note that after a shuffle all the 100 songs are played before the next shuffle.

1. What is the probability that the first PF song heard is the 5th song played?
2. Similar to the **simulation** of the Birthday paradox that we discussed in class, write a R code to simulate the problem and determine the probability that the first PF song heard is the 5th song played.

3.1 Answer

1. What is the probability that the first PF song heard is the 5th song played? The number of ways in which the first 5 songs can be played is $100 \times 99 \times 98 \times 97 \times 96$. The number of ways in which the first 4 songs are non-PF is $90 \times 89 \times 88 \times 87$. And the number of ways in which the 5th song is a PF song is 10. Thus, P(5th song is PF) is given by

$$\begin{aligned} P(\text{5th song is PF}) &= \frac{90 \times 89 \times 88 \times 87 \times 10}{100 \times 99 \times 98 \times 97 \times 96} \\ &= \frac{10 \times {}^{90}P_4}{{}^{100}P_5} \\ &= .0679 \end{aligned}$$

2.
```
m = 100000                   # number of trials
n = 100                      # total number of songs. We will assume that the
x = numeric(m)               # binary vector 1 if the  5th song is the 1st PF
for (i in 1:m)
{
b = sample(1:100, n, repl=F) # a random shuffle of 100 songs

if ((b[1] > 10) & (b[2] > 10) & (b[3] > 10) & (b[4] >10) & (b[5] <= 10)) {
    x[i] = 1
    }
```

```
else {
    x[i] = 0
    }
}
pmean = mean(x == 1)
print(pmean)
```

3.2 Points - Total: 3 pts

- 1 pt for recognizing and setting up the probability in 3.1. This means your numerator should be 100 $\times$ 99 $\times$ ... (1/2 credit if one error is made in numerator or denominator, or if the wrong decimal is computed- 0 points if errors are made in more than one place)
- 1 pt for a simulation that reasonably simulates the probability of song 5 being the first Pink Floyd song. This can be done as above, using a sample of 5 elements 1:100 and estimating the probability for the correct order being in play. Basically, you get the point if you successfully created a simulation
- 1 pt if your probability estimate from part 1 and part 2 are the same – you should have done this check to make sure that you did both 3.1 and 3.2 correctly. If these decimals match (or are very close, with reasonable rounding error), then you get this point.

4 Problem

Consider two specific memory locations that is within the address space of a computer program when it executes. With probability 0.5, the program will access the first location; with probability 0.4 it will access the second location and with probability 0.3 it will access both locations. What is the probability that the program will access neither location?

4.1 Answer

Let B_i denote the event that the computer program access location $i, i = 1, 2$. Then the probability that the program accesses at least one of the locations is given by

$$\begin{aligned} P(B_1 \cup B_2) &= P(B_1) + P(B_2) - P(B_1 B_2) && (6)\\ &= 0.5 + 0.4 - 0.3 && (7)\\ &= 0.6 && (8) \end{aligned}$$

The event that the computer program will access neither location is

$$\begin{aligned} P(B_1^{\complement} \cap B_2^{\complement}) &= P((B_1 \cup B_2)^{\complement}) && (9)\\ &= 1 - P(B_1 \cup B_2) && (10)\\ &= 1 - 0.6 && (11)\\ &= 0.4 && (12) \end{aligned}$$

4.2 Points - Total: 1 pt

- 1/2 pt if you successfully setup the union of the first and second locations (0.6)
- 1/2 pt if you then successfully found the probability of the complement (0.4)

5 Problem

John is 31 years old, single, outspoken, and has multiple talents. He majored in Computer Science. As a student, he was deeply concerned with issues of discrimination and social justice, and environment. Which of the following scenarios is more probable? **A**) John is a computer programmer. **B**) John is a computer programmer and is an activist in the environmental movement. Give a reason for your answer.

5.1 Answer

The event that John is both a computer programmer and environmental activist is a subset of the event that John is a computer programmer. Thus, it cannot be more likely than the event that John is a computer programmer.

If event X is a subset of event Y then

$$P(X) \leq P(Y)$$

Here, $P(X) < P(Y)$ since there is some non-zero probability that John is a computer programmer but not active in the environmental movement.

A similar problem was made famous by Amos Tversky and Daniel Kahneman, whose 1983 study showed that 85% of respondents erroneously thought it was more likely that John is a computer programmer and is an activist in the environmental movement. These are related to mental/implicit biases that can/should be avoided and probability provides a framework for analyzing and avoiding them `brilliant.org`.

5.2 Points - Total: 1 pt

- 1/2 pt for answering A
- 1/2 pt for making some mention of a subset or that the joint probability is implicitly lower (or equal) than either marginal probability