CSCI 2244 – Homework 2

Out: Friday, September 6, 2019
Due: Friday, September 13, 2019, 11:59pm

This homework consists of written exercises and coding problems. You *must* type your solutions. See the "Assignments" section in the syllabus for advice about doing this. You should submit your homework via Canvas. In particular, you should upload a zip file called:

`FirstName_LastName_Homework2.zip`

Please use your full first name and last name, as they appear in official university records. The reason for doing so is that the TAs and I must match up these names with the entries in the gradebook.

This zip file should contain 2 files:

- written.pdf – containing your answers to all the tasks in section 1, and the results of running your code as requested by the task in section 2.
- birthday.py – containing the code you wrote for section 2. You should start with the starter code file by this name on the course website.

1 Written Exercises

For the first three tasks, you are asked about the number of ways to do something or the probability of an event. You should give your answer both as a formula that involves the product of factorials/binomial coefficients, and also a decimal answer. To compute the decimal answer, you can use Python 3 code for computing binomial coefficients and factorials, which is demonstrated in the starter code.

Task 1.1 (2 pts). A website decides to require that all users' passwords must consist of 5 lowercase letters from the English alphabet followed by exactly 2 numerical digits. For example, under this policy, "abcda12" is a valid password, but "a1a1aaa" is not. How many possible passwords are there? (Note: do not do this if you ever manage security for a website, it's a *terrible* idea.)

Task 1.2 (2 pts). You flip a fair coin 10 times. What is the probability that you get 3 or fewer heads?

Task 1.3 (3 pts). The BC student assembly has 40 senators. Suppose there are 30 seniors and 10 juniors on the senate. The assembly decides to form a committee with 5 members. The committee must have at least 3 seniors. How many valid committees are there?

Task 1.4 (3 pts). Show that if you flip n fair coins, the probability that there are an even number of heads is 1/2. (Hint: Imagine starting by flipping $n-1$ coins, and then flipping the last one. Think about how the situation compares to task 1.5 from the first homework. It might be helpful to think about the cases where $n = 2$ and $n = 3$.)

In class, we discussed ways to prove an equation of the form $A = B$ where A and B are formulas involving binomial coefficients. The two ways we discussed were:

1. Apply the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ and then use algebra to deduce the equation.

2. Show that A and B describe two ways to count all the elements of the same set S.

We showed how both methods could be used to prove that $\binom{n}{k} = \binom{n}{n-k}$ for $0 \le k \le n$. (This equation is called the **symmetry** property of binomial coefficients). We also used the second method to show that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \tag{1}$$

The idea there was that if X is a set with n elements, then both sides of this equation count how many subsets of X there are: the left side of the equation corresponds to saying that there $\binom{n}{0}$ subsets with 0 elements, $\binom{n}{1}$ subsets with 1 elements, $\binom{n}{2}$ subsets with 2 elements, and so on; the right hand side is observing that for each element of X, we can either include it or not in a subset, so there are 2^n subsets.

Another method for proving these equations is to use the **binomial theorem**, which says that given a natural number n and real numbers x and y, we have:

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

If we set $x = 1$ and $y = 1$ in this equation, then the left hand side becomes 2^n, while $\binom{n}{k} x^k y^{n-k}$ just becomes $\binom{n}{k}$. So, with these values of x and y, we get Equation 1. For the next few tasks, you will use these various methods to prove equations about binomial coefficients:

Task 1.5 (2 pts). For $n \ge 1$ and $k \ge 1$, show that

$$\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1}$$

Task 1.6 (3 pts). Show that

$$\sum_{k=0}^{n} k \binom{n}{k} = n 2^{n-1}$$

(Hint: Take the derivative of both sides of the binomial theorem equation with respect to x, and then substitute in an appropriate choice for x and y.)

2 Coding

In class, we discussed the birthday problem, which involves computing the probability that at least 2 people would share a common birthday in a group of n people. Besides being amusing and surprisingly counter-intuitive, this problem is similar to scenarios that come up in practical situations in computer science.

A natural variant is to consider what happens when we require that more than just two people have the same birthday. For example, we might want to know the probability that at least 4 people have the same birthday. It is possible to derive formulas for this, but in this exercise we will instead use simulations to help us approximate the answer.

As with last week's assignment, look at the starter code on the course website, which shows how to generate random numbers that you'll need to do the task.

Task 2.1 (8 pts). Write a function called `birthday(n, k, c)` which first generates `k` random integers with each integer sampled uniformly from the set $\{1, \ldots, \texttt{n}\}$. It should then return `True` if at least `c` of the integers are the same, and return `False` otherwise. That is, if there is some number z which occurs at least c times among the random integers, return `True`. For example, if $c = 3$ and the random numbers generated were $1, 1, 1, 2$, the result would be `True`, while for $c = 3$ and $1, 1, 2, 2$ the result would be `False`. For ease of debugging, you might want to break the code up into separate functions, one of which generates the random numbers, and another of which tests whether at least `c` are the same. That way, you can test your code for the latter separately from the random part.

Task 2.2 (4 pts). In this task, you will use your procedure to estimate the probabilities of events. Review section 2.1.3 from the textbook, especially the part on Monte Carlo simulation. Then, use your code from the previous task to estimate the probability of various birthday scenarios. In particular, you should try $n \in \{365, 1000\}$, and $k \in \{20, 40\}$, and $c \in \{2, 3\}$ and estimate the probability that `birthday(n, k, c)` returns true. For each combination of n, k, and c values, you should run 100000 trials to form this estimate. Report your estimated probabilities in the PDF you submit.