

David Varodayan
CS 109

Problem Set #1
January 8, 2020

Problem Set #1 Due: 1:00pm on Friday, January 17

With problems by Mehran Sahami and Chris Piech

For each problem, briefly explain/justify how you obtained your answer. Brief explanations of your answer are necessary to get full credit for a problem even if you have the correct numerical answer. The explanations help us determine your understanding of the problem whether or not you got the correct answer. Moreover, in the event of an incorrect answer, we can still try to give you partial credit based on the explanation you provide. It is fine for your answers to include summations, products, factorials, exponentials, or combinations; you don't need to calculate those all out to get a single numeric answer.

Note: all assignment submissions will be made online through Gradescope. You can find information on signing up to submit assignments through Gradescope on the class webpage. If you handwrite your solutions, you are responsible for making sure that you can produce **clearly legible** scans of them for submission. You may use any word processing software you like for writing up your solutions. On the CS109 webpage we provide a template file and tutorial for the \LaTeX system, if you'd like to use it.

1 Written problems

Submit your solutions to these written problems as a single pdf file on Gradescope.

1. Introduce yourself! Fill out this Google form to tell me a bit about you:

<https://forms.gle/tRvuhnyXfV1LXX7t5>

(No need to copy the answers into your Gradescope submission; you can select an arbitrary page or write “done” so there is something to select.)

2. 12 computers are brought in for servicing (and machines are serviced one at a time). Of the 12 computers, 3 are PCs, 5 are Macs, and 4 are Linux machines. Assume that all computers of the same type are indistinguishable (i.e., all the PCs are indistinguishable, all the Macs are the indistinguishable, etc.).
 - a. In how many distinguishable ways can the computers be ordered for servicing?
 - b. In how many distinguishable ways can the computers be ordered if the first 5 machines serviced must include all 4 Linux machines?
 - c. In how many distinguishable ways can the computers be ordered if 1 PC must be in the first 3 and 2 PCs must be in the last 3 computers serviced?

3. At the local zoo, a new exhibit consisting of 3 different species of birds and 3 different species of reptiles is to be formed from a pool of 7 bird species and 5 reptile species. How many exhibits are possible if
 - a. there are no additional restrictions on which species can be selected?
 - b. 2 particular bird species cannot be placed together (e.g., they have a predator-prey relationship)?
 - c. 1 particular bird species and 1 particular reptile species cannot be placed together?
4. A substitution cipher is derived from orderings of the alphabet. How many ways can the 26 letters of the English alphabet (21 consonants and 5 vowels) be ordered if each letter appears exactly once and:
 - a. There are no other restrictions?
 - b. The letters Q and U must be next to each other (but in any order)?
 - c. All five vowels must be next to each other?
 - d. No two vowels can be next to each other?
5. Given all the start-up activity going on in high-tech, you realize that applying combinatorics to investment strategies might be an interesting idea to pursue. Say you have \$20 million that must be invested among 4 possible companies. Each investment must be in integral units of \$1 million, and there are minimal investments that need to be made if one is to invest in these companies. The minimal investments are \$1, \$2, \$3, and \$4 million dollars, respectively for company 1, 2, 3, and 4. How many different investment strategies are available if
 - a. an investment must be made in each company?
 - b. investments must be made in at least 3 of the 4 companies?
6. How many ways can you split a class of 99 students into 33 project groups of 3 students each? Neither the order of the groups nor the order of students within groups matters.
7. You are counting cards in a card game that uses two standard decks of cards. There are 104 cards total. Each deck has 52 cards (13 values each with 4 suits). Cards are only distinguishable based on their suit and value, not which deck they came from.
 - a. In how many distinct ways can the cards be ordered?
 - b. You are dealt two cards. How many distinct pairs of cards can you be dealt? Note: the order of the two cards you are dealt does not matter.
8. Determine the number of vectors (x_1, x_2, \dots, x_n) such that each x_i is a non-negative integer and $\sum_{i=1}^n x_i \leq k$, where k is some constant non-negative integer. Note that you can think of n (the size of the vector) and k as constants that can be used in your answer.

9. Consider an array x of integers with k elements (e.g., `int x[k]`), where each entry in the array has a distinct integer value between 1 and n , inclusive, and the array is sorted in increasing order. In other words, $1 \leq x[i] \leq n$, for all $i = 0, 1, 2, \dots, k-1$, and the array is sorted, so $x[0] < x[1] < \dots < x[k-1]$. How many such sorted arrays are possible?
10. You are running a web site that receives 8 hits (in a particular second of time). Your web site is powered by 5 computers, and each hit to your web site can be serviced by any one of the 5 computers you have, where each computer is capable of processing as many (or as few) requests as it is given.
 - a. In how many distinct ways could the 8 hits to your website be distributed among the 5 computers if all hits are considered identical?
 - b. In how many distinct ways could the 8 hits to your website be distributed among the 5 computers if the hits consisted of 5 identical requests for web page A and 3 identical requests for web page B (note that requests for web page A are distinguishable from requests for web page B)?
11. Say a university is offering 3 programming classes: one in Java, one in C++, and one in Python. The classes are open to any of the 100 students at the university. There are:
 - a total of 27 students in the Java class;
 - a total of 28 students in the C++ class;
 - a total of 19 students in the Python class;
 - 12 students in both the Java and C++ classes (note: these students are also counted as being in each class in the numbers above);
 - 4 students in both the Java and Python classes;
 - 11 students in both the C++ and Python classes; and
 - 3 students in all three classes (note: these students are also counted as being in each pair of classes in the numbers above).
 - a. If a student is chosen randomly at the university, what is the probability that the student is not in any of the 3 programming classes?
 - b. If a student is chosen randomly at the university, what is the probability that the student is taking *exactly one* of the three programming classes?
 - c. If two different students are chosen randomly at the university, what is the probability that at least one of the chosen students is taking at least one of the programming classes?
12. To get good performance when working binary search trees (BST), we must consider the probability of producing completely degenerate BSTs (where each node in the BST has at most one child). See Lecture Notes # 2, Example 2, for more details on binary search trees.
 - a. If the integers 1 through n are inserted in arbitrary order into a BST (where each possible order is equally likely), what is the probability (as an expression in terms of n) that the resulting BST will have completely degenerate structure?
 - b. Using your expression from part (a), determine the smallest value of n for which the probability of forming a completely degenerate BST is less than 0.001 (i.e., 0.1%).

13. Suppose that m strings are hashed (randomly) into N buckets, assuming that all N^m arrangements are equally likely. Find the probability that exactly k strings are hashed to the first bucket.
14. Say a hacker has a list of n distinct password candidates, only one of which will successfully log her into a secure system.
 - a. If she tries passwords from the list at random, deleting those passwords that do not work, what is the probability that her first successful login will be (exactly) on her k -th try?
 - b. Now say the hacker tries passwords from the list at random, but does **not** delete previously tried passwords from the list. She stops after her first successful login attempt. What is the probability that her first successful login will be (exactly) on her k -th try?
15. Say we send out a total of 20 distinguishable emails to 12 distinct users, where each email we send is equally likely to go to any of the 12 users (note that it is possible that some users may not actually receive any email from us). What is the probability that the 20 emails are distributed such that there are 4 users who receive exactly 2 emails each from us and 3 users who receive exactly 4 emails each from us?

2 Coding

Download the starter code `cs109_pset1.py` and fill out the function `q16` as described in it. You **MUST** use the starter code. When ready to submit, upload only `cs109_pset1.py` directly to Gradescope under “PSet1 - Coding” (not a zip file). **IMPORTANT:** Do NOT modify the file name, or the function name. Just make sure the function `q16` does as the problem describes, and returns a single probability.

16. Consider a game, which uses a generator that produces independent random integers between 1 and 100, inclusive. The game starts with a sum $S = 0$. The first player adds random numbers from the generator to S until $S > 100$, at which point they record their last random number x . The second player continues by adding random numbers from the generator to S until $S > 200$, at which point they record their last random number y . The player with the highest number wins; e.g., if $y > x$, the second player wins. Write a Python 3 program to simulate 100,000 games and output the estimated probability that the second player wins.