

Problem Set #1 Due: 1:00pm on Wednesday, July 1

With problems by Mehran Sahami, Chris Piech, and Lisa Yan

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 **L^AT_EX** system, if you'd like to use it.

Collaboration policy: You are encouraged to discuss problem-solving strategies with each other as well as the course staff, but you must write up your own solutions and submit individual work. Please cite any collaboration at the top of your submission.

Sanity check: For this problem set, we are offering an optional "sanity check" tool on Gradescope. It will let you check some of your numerical answers before the deadline. We hope it will prevent some stress in these uncertain times.

1 Written problems

Submit your solutions to these written problems as a single pdf file on Gradescope. When answering, please put your final answer up top with the explanation below.

1. How many ways can 10 (distinct) people be seated in a row if
 - a. there are no restrictions on the seating arrangement?
 - b. persons A and B must sit next to each other?
 - c. there are 5 adults and 5 children, and no two adults nor two children can sit next to each other?
 - d. there are 5 married couples and each couple must sit together?
2. 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 8 bird species and 6 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?

3. 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 20 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);
- 5 students in both the Java and Python classes;
- 8 students in both the C++ and Python classes; and
- 2 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? Please find a numerical answer.

4. 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. 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

- an investment must be made in each company?
- investments must be made in at least 3 of the 4 companies?

5. 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?

6. If we assume that all possible poker hands (comprised of 5 cards from a standard 52 card deck) are equally likely, what is the probability of being dealt:

- a flush? (A hand is said to be a flush if all 5 cards are of the same suit. Note that this definition means that *straight flushes* (five cards of the same suit in numeric sequence) are also considered flushes.)
- two pairs? (This occurs when the cards have numeric values a, a, b, b, c , where a, b and c are all distinct.)

c. four of a kind? (This occurs when the cards have numeric values a, a, a, a, b , where a and b are distinct.)

7. Imagine you have a robot (Θ) that lives on an $n \times m$ grid (it has n rows and m columns):

The robot starts in cell $(1, 1)$ and can take steps either to the right or down (**no left or up steps**). How many distinct paths can the robot take to the destination (\star) in cell (n, m) if

- there are no additional constraints?
- the robot must start by moving to the right?
- the robot changes direction exactly 3 times? Moving down two times in a row is not changing directions, but switching from moving down to moving right is. For example, moving [down, right, right, down] would count as having two direction changes.

8. Say we roll a six-sided die six times. What is the probability that

- we will roll two different numbers *thrice* (three times) each?
- we will roll *exactly one* number *exactly* three times? Hint: Be careful of overcounting.

9. A binary string containing M 0's and N 1's (in arbitrary order, where all orderings are equally likely) is sent over a network. What is the probability that the first r bits of the received message contain exactly k 1's?

10. 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?

11. Say a hacker has a list of n distinct password candidates, only one of which will successfully log her into a secure system.

- 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?
- 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?

12. 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.
13. **[Extra credit]** 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%).

2 Coding

Download the starter code `cs109_pset1.py` and fill out the function `q14` 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 `q14` does as the problem describes, and returns a single probability.

The Gradescope autograder will report your problem score; there are no hidden tests. No partial credit will be given if you do not pass the autograder, but you can submit an unlimited number of times. We will only use the score from your most recent submission.

14. 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.