Chris Piech
CS 109

Problem Set #1

Problem Set #1

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 though 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.

0. This is an example of a great student answer. It will get full points because it was explained well. But there is a mistake. Your job is to find the mistake.

   Original Problem: 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. In how many distinct ways can the hits be serviced?

   Student answer: This problem can be solved using the divider method. You need to divide the $n = 8$ hits among the $r = 5$ computers. The number of ways of dividing is

$$\binom{n+r-1}{r-1} = \binom{8+5-1}{5-1} = \binom{11}{4} = 330.$$

1. A “DNA-turn" has 11 base pairs. Each base pair can take on one of four distinct values, {A, T, G, C}. How many distinct DNA-turns are there?

2. 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?

3. 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.

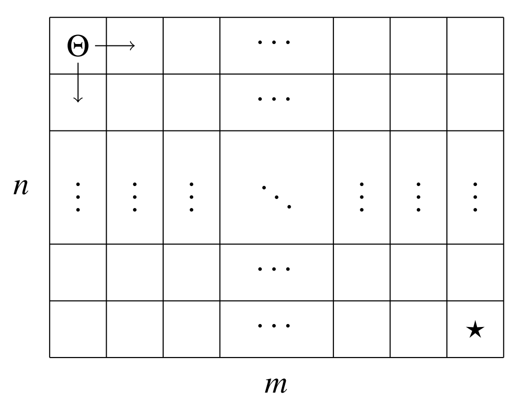
4. Given all the start-up activity going on in high-tech, you realize that applying combinatorics to investment strategies might be an interesting idea to pursue. Say you have \$20 million that must be invested among 4 possible companies. Each investment must be in integral units of \$1 million, and 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?

5. 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.

6. Determine the number of vectors $(x_1, x_2, \ldots, x_n)$ 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.

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):

   a. if there are no additional constraints?
   b. if the robot must start by moving to the right?

c. if 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. 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)?

2 Probability

9. 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 5th 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 5th try?

10. 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. 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.)

    b. two pairs? (This occurs when the cards have numeric values a, a, b, b, c, where a, b and c are all distinct.)

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

11. 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?

3 Coding

12. **C**onsider 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. Include your answer along with code used to compute it. Give your answer rounded to 3 places behind the decimal.

    Above and beyond: For extra credit, calculate the exact probability (without sampling)