Nikolay Kozlovskiy
CS 109

Problem Set #1

Problem Set #1 (Solved)

Important: If you want to view and work on the document in the LaTeX system, click on this url, create a copy of the project, and you're ready to go!

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. In this problem we show 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.

   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

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

   * Intresting thing, there is actually a typo in the example, the right answer should be $\binom{12}{4}$

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

Theory background: Step rule of counting

Solution: Each base pair value does not effect the number of outcomes of the next base pair thus the step rule of counting could be used. A turn has 11 base pairs with 4 distinct values which is resulted in 4^{11} distinct turns $(4 \times 4 \times 4 \cdots \times 4)$

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?

Theory background: Permutation of distinct objects, Step rule of counting

Solution:

a. All 26 letters appear once, thus each element is distinct, and they can be permuted in 26! ways

b. First step: let's consider Q and U placed together as some additional variable called α and also "delete" these two letters from the alphabet, thus we create 25 distinct objects (α and 24 English letters without Q and U) and its number of permutations equals 25!

   Second step: since Q and U in α could be placed in any order, there are 2! ways of doing so (or 2! variants of α)

   By the step rule of counting final number of permutations equals N = 25! · 2!

c. Again let's consider all five vowels placed together as some variable called β and "delete" these vowels from the alphabet. There are 22! ways of ordering distinct β and 21 consonants, and also within β there are 5! ways of ordering vowels, applying the step rule we get N = 22! · 5!

d. If no vowels can be placed together it means that each of these vowels should be placed in one of **22 places**: at the beginning/end of letter sequence or between 21 consonants. The uniqueness of letter sequence will vary on two conditions: unique ordering of consonants and unique placement of 5 vowels in these vacant 22 places

   First step: there are 21! ways of ordering constants

   Second step: the first vowel has 22 available places "to take", the second vowel has already 22-1 available places "to take" and so on, thus number of ways to put 5 vowels equals $22 \times 21 \times 20 \times 19 \times 18$

   $N = 21! \cdot 22 \times 21 \times 20 \times 19 \times 18$

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.

Theory background: Permutation of indistinct objects, Combinations, Mutually exclusive counting

Solution:

a. If we consider all 104 cards distinct then there are 104! ways of unique orderings. However, it is said that cards are not distinguishable by deck, which means that each card has its duplicate, thus to get the number of unique permutations the 104! has to be divided by $54 \cdot 2!$

b. First step: since we do not care about the order of the two cards, we can use combination theory to get the number of distinct pairs of cards. However, based on the definition of the combination, the set from which to choose the combinations should have only **distinct objects**, thus we only consider 52 cards (13 values 4 suits), $\binom{52}{2} = \frac{52\times51}{2} = 1326$

   Second step: In the previous step we calculated the number of distinct pairs of 52 cards (a.k.a one deck), but what the second deck can bring in terms of "enriching" the uniqueness of card pairs? The answer: with the second deck new pairs of identical cards can be formed (e.g. two Kings of Spades), there are 52 pairs of identical cards

   Third step: the outcomes of two previous steps are mutually exclusive, thus N = 1326 + 52 = 1378

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?

Theory background: Bucketing with indistinct Objects, Mutually exclusive counting

Solution:

a. Firstly, we need to take into account the spending restriction mentioned in text, which leads us to the fact that \$10 million dollars are already allocated and we can not do anything with them. The number of ways of investing the remaining \$10 million dollars is bucketing with indistinct objects problem and solved as: N = $\binom{10+4-1}{10}$ = 286

b. The condition “invest in at least 3 of the 4 companies” leads us to a few (but luckely definite) number of strategies:

   Invest in all 4: N = 286
   Invest in 1,2,3 companies: 6 millions - allocated, N = $\binom{14+3-1}{14}$ = 120
   Invest in 1,2,4 companies: 7 millions - allocated, N = $\binom{13+3-1}{13}$ = 105
   Invest in 1,3,4 companies: 8 millions - allocated, N = $\binom{12+3-1}{12}$ = 91
   Invest in 2,3,4 companies: 9 millions - allocated, N = $\binom{11+3-1}{11}$ = 78

   Since all 5 strategies have different set of companies involved, groups of outcomes are mutually exclusive, thus the N = 286 + 120 + 105 + 91 + 78 = 680

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.

Theory background: Bucketing into fixed-sized containers

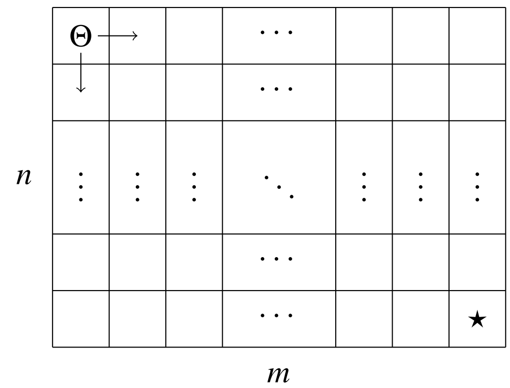
Solution: This is a straightforward application of the multinomial coefficient representation, $N = \frac{99!}{(3!)^{33}}$

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.

Theory background: Bucketing with indistinct objects

Solution: This is a classic problem in combinatorics, the main trick is too add a slack parameter x_{n+1} so it turns $\sum_{i=1}^{n} x_i \leq k$ inequality to $\sum_{i=1}^{n} x_i + x_{n+1} = k$ quality, and thus this problem turns into a bucketing with indistinct objects problem, one that is presented in e.g. 4 exercise. Where k is "number of millions" which can be invested in integral units of $1 million (remember the non-negative integer condition) and n+1 is "the number of companies", thus the solution is N = $\binom{k+n}{k}$

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.

Theory background: Permutations of indistinct objects, Bucketing with indistinct objects, Step rule of counting, Mutually exclusive counting

Solution:The first thing to understand is no matter what way a user chooses there always be (n-1) down steps and (m-1) right steps.

a. Consider the way as a sequence of (n-1) down steps and (m-1) right steps, by the permutations of indistinct objects rule the number of unique orderings of steps is N = $\frac{(n-1+m-1)!}{(n-1)!(m-1)!} = \frac{(n+m-2)!}{(n-1)!(m-1)!} or \binom{n-1+m-1}{n-1}$

b. The same logic as in the *a* point, but the experiment has now two steps: turn right and then create unique orderings of (n-1) down steps and (m-2) right steps, thus N = $1 \cdot \frac{(n+m-3)!}{(n-1)!(m-2)!}$

c. There are two strategies to order the steps to make **exactly 3 direction changes**

   If a user starts with down step: d_1 moves r_1 moves d_2 moves r_2 moves
   If a user starts with right step: r_1 moves d_1 moves r_2 moves d_2 moves
   Important: $d_1 + d_2$ = n - 1 and $r_1 + r_2$ = m - 1

   In both strategies (n-1) down moves and (m-1) right moves have to be stuffed into 2 segments (or containers) respectively and **each container has to have at least one step allocated**. By the bucketing with indistinct objects rule this could be done in

$\binom{n-1-2+2-1}{n-1-2} = n - 2$ ways for down moves group and $\binom{m-1-2+2-1}{m-1-2} = m - 2$ ways for right moves group, by the product rule of counting total paths for each strategy would be $(n - 2) \times (m - 2)$, since we have two mutually exclusive strategies the total number of paths equals $2 \times (n - 2) \times (m - 2)$

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)?

Theory background: Bucketing with indistinct objects, Step rule of counting

Solution:

a. This is a straightforward application of bucketing with indistinct objects, N = $\binom{8+4}{8} = \frac{12\times11\times10\times9}{4!} = 495$

b. Fist step is to calculate the number of ways to distribute 5 **indistinguishable** requests for web page A among 5 computers, $N_A = \binom{5+4}{5} = 126$

   Second step is to calculate the number of ways to distribute 3 **indistinguishable** requests for web page B among 5 computers, $N_B = \binom{3+4}{3} = 35$

   By the step rule of counting, the total number of ways to distribute the requests of group A and B is N = $N_A \cdot N_B = 4410$

   *This solution works if we do not care about the order of A and B requests

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?

Theory background: Independence, Probability of And with independent events, Probability of And with dependent events

Solution:

a. In this case when we delete a password after a failure the events are dependent. The probability of some outcome in the next event depends on the outcomes of the previous event because the pool of possible passwords (outcomes) changes, thus we should use **the chain rule**.

   The prob of a failure in the first attempt: $P(E_1) = \frac{n-1}{n}$

   The prob of a failure in the second attempt given that E_1 occurred: $P(E_2 \mid E_1) = \frac{n-2}{n-1}$

   The prob of a failure in the third attempt given that E_1 and E_2 occurred:$P(E_3 \mid E_1 \text{ and } E_2) = \frac{n-3}{n-2}$

   The prob of a failure in the fourth attempt given that E_1 and E_2 and E_3 occurred:$P(E_4 \mid E_1 \text{ and } E_2 \text{ and } E_3) = \frac{n-4}{n-3}$

   The prob of a success in the fifth attempt given that E_1 and E_2 and E_3 and E_4 occurred:$P(E_5 \mid E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } E_4) = \frac{1}{n-4}$

   $P(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } E_4 \text{ and } E_5) = \frac{1}{n}$

b. In this case all steps are independent since we do not delete previously tried passwords and thus the formula is $P(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } E_4 \text{ and } E_5) = \prod_{i=1}^{5} P(E_i)$

The prob of a failure for each of four first step is $P(E) = \frac{n-1}{n}$, the prob of a success in the fifth attempt is $P(E_5) = \frac{1}{n}$

$P(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } E_4 \text{ and } E_5) = \frac{(n-1)^4}{n^5}$

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.)

Theory background: Combinations, Step rule of counting, Definition of probability

Solution: We work in the space where we treat objects as distinct but unordered, thus for each sub-tasks the size of the sample space equals $|S| = \binom{52}{5} = 2598960$

a. To find $|E|$:
First step: choose the suit $\binom{4}{1} = 4$
Second step: once the suit is chosen there are $\binom{13}{5} = 1287$ options to choose five cards of the same suit
Third step: by step rule of counting $|E| = 4 \cdot 1287 = 5148$

$$P(E) = \frac{|E|}{|S|} = \frac{4 \cdot \binom{13}{5}}{\binom{52}{5}} = \frac{5148}{2598960} = 0.00198 = 0.198\%$$

b. To find $|E|$:
First step: choose random three unique ranks $\binom{13}{3} = 286$
Second step: out of three chosen ranks calculate the number of ways to choose two for a and b $\binom{3}{2} = 3$
Third step: for each chosen a and b, there are $\binom{4}{2} = 6$ ways to form a pair (in other words choose 2 suits out of 4)
Fourth step: for c value there are $\binom{4}{1} = 4$ ways to choose the suit
Fifth step: by step rule of counting $|E| = \binom{13}{3} \cdot \binom{3}{2} \cdot \binom{4}{2} \cdot \binom{4}{2} \cdot \binom{4}{1} = 286 \cdot 3 \cdot 6 \cdot 6 \cdot 4 = 123552$

$$P(E) = \frac{|E|}{|S|} = \frac{123552}{2598960} = 0.0475 = 4.75\%$$

c. To find $|E|$:
First step: choose random three unique ranks $\binom{13}{3} = 286$
Second step: out of three chosen ranks calculate the number of ways to choose one for a $\binom{3}{1} = 3$
Third step: for chosen a there is $\binom{4}{3} = 4$ ways to form three of a kind (in other words choose 3 suits out of 4)
Fourth step: for each b and c value there are $\binom{4}{1} = 4$ ways to choose their suit

Fifth step: by step rule of counting $|E| = \binom{13}{3} \cdot \binom{3}{1} \cdot \binom{4}{3} \cdot \binom{4}{1} \cdot \binom{4}{1} = 286 \cdot 3 \cdot 4 \cdot 4 \cdot 4 = 54912$

$P(E) = \frac{|E|}{|S|} = \frac{54912}{2598960} = 0.0211 = 2.11\%$

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?

Theory background: Combinations, Bucketing, Step rule of counting, Definition of probability

Solution:We work in the space where we treat objects as distinct but unordered, thus by rule of bucketing of distinct objects the sample space is $|S| = 12^{20}$

To find $|E|$:
First step: calculate the number of ways to select 4 users that receive 2 emails $\binom{12}{4}$
Second step: out of remaining 8 users, calculate the number of ways to choose 3 users that receive 4 emails $\binom{8}{3}$
Third step: by the rule of bucketing into fixed-sized containers apply multinominal coefficient $\binom{20}{2,2,2,2,4,4,4}$ to find the number of ways of emails' distribution among selected users (containers)
Fourth step: by step rule of counting $|E| = \binom{12}{4} \cdot \binom{8}{3} \cdot \binom{20}{2,2,2,2,4,4,4}$

$P(E) = \frac{|E|}{|S|} = \frac{\binom{12}{4} \cdot \binom{8}{3} \cdot \binom{20}{2,2,2,2,4,4,4}}{12^{20}}$ - if you want you can calculate, but I had enough calculation overall in previous exercises :)

3 Coding

12. 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. 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) - *Maybe not today, maybe not tomorrow and maybe not the next month, but only one thing is true, I'll be champion one day. I promise*

```python
import random

MIN_LIMIT = 1
MAX_LIMIT = 100
NUM_TRIALS = 100000

def run_simulations(num_trials: int = NUM_TRIALS) -> float:
    num_wins_for_second = 0

    for _ in range(num_trials):
        S = 0
        x, y = None, None

        while S <= 200:
            rand_int = random.randint(MIN_LIMIT, MAX_LIMIT)
            S += rand_int

            if x is None and S > 100:
                x = rand_int

        y = rand_int
        num_wins_for_second += int(y > x)

    return round(num_wins_for_second / num_trials, 3)

if __name__ == "__main__":
   prob_second_win = run_simulations()
   print(prob_second_win) # appr. 0.52
```