CS 109: Probability for Computer Scientists
Problem Set #1

Nguyen Huy Hieu

December 23, 2024

- Submit on Gradescope by 1:00pm Pacific on Friday, September 25th, for a small, "on-time" bonus.
- All students have a pre-approved extension, or "grace period" that extends until Monday 1:00pm Pacific, when they can submit with no penalty. **The grace period expires on 1:00 Pacific on Monday, September 27th**, after which we cannot accept further late submissions.
- **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.
- **Tagging written problems**: When you submit your written PDF on Gradescope you must tag your PDF, meaning that you must assign pages of your PDF as answers to particular questions so that we can properly grade your submission. For problem sets, we are deducting **2 points** for any submissions that do not have all questions tagged.
- **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.
- If you handwrite your solutions, you are responsible for making sure that you can produce **clearly legible** scans of them for submission. You may also 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.

Cited collaboration: N/A

1. How many ways can 10 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?

Answer. a. Since there are no restrictions, the answer is simple: $10! = ...$

b. If persons A and b must sit next to each other, we will count them as one person, which mean we have to arrange 9 people. In addition, for each permutation there are two posibility of A and B: AB and BA. Therefore the answer is: $9! \times 2 = ...$

c. If no two adults nor two children can sit next to each other, let call a children as A and an adult as B. There are two posible arrangement: ABABABABAB and BABABABABA For every 5 children or adults, there are $5!$ way to arrange them. So final the solution is: $5! \times 5! \times 2 = ...$

d. The ideas is similar to question b. For each couple we count them as 1 person and there are 2 way to arrange them. So the answer is: $5! \times 2^5 = ...$

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?

Answer. a. Since there are no restrictions, the answer is simple: $10! = ...$

b. If persons A and b must sit next to each other, we will count them as one person, which mean we have to arrange 9 people. In addition, for each permutation there are two posibility of A and B: AB and BA. Therefore the answer is: $9! \times 2 = ...$

c. If no two adults nor two children can sit next to each other, let call a children as A and an adult as B. There are two posible arrangement: ABABABABAB and BABABABABA For every 5 children or adults, there are $5!$ way to arrange them. So final the solution is: $5! \times 5! \times 2 = ...$

d. The ideas is similar to question b. For each couple we count them as 1 person and there are 2 way to arrange them. So the answer is: $5! \times 2^5 = ...$

3. 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. How many different investment strategies are available if

   a. an investment must be made in each company, and we must invest all \$20 million? Assume we have a minimal investment requirement per company, where we must invest a minimal investment of \$1, \$2, \$3, and \$4 million dollars for company 1, 2, 3, and 4, respectively.

   b. investments must be made in at least 3 of the 4 companies, and we must invest all \$20 million? Assume we have the same minimal investment as in part (a), where should we choose to invest in company n (for $n = 1, \ldots, 4$), we must invest a minimal of $\$n$ million.

   c. we must invest less than or equal to $\$k$ million dollars total among the 4 companies, where k is an integer such that $10 \leq k \leq 20$? Note that you can think of k as a constant that can be used in your answer. Assume in this part that we do not have a minimal investment.

Answer. a. Because we must invest a minimal 1, 2, 3 and 4 millions dollars to 4 companies, this mean that there are \$10 millions left. Number of ways to split these 10 millions into companies is: $\frac{(10+3)!}{10! \times 3!} = \ldots$

b. Consider 5 scenarios, there are $S_1 = \frac{(10+3)!}{10! \times 3!} = \ldots$ ways to invest in all 4 companies, there are $S_2 = \frac{(16+2)!}{16! \times 2!} = \ldots$ ways to invest only to companies 1, 2 and 3, there are $S_3 = \frac{(17+2)!}{17! \times 2!} = \ldots$ ways to invest only to companies 1, 2 and 4, there are $S_4 = \frac{(19+2)!}{19! \times 2!} = \ldots$ ways to invest only to companies 2, 3 and 4, there are $S_1 = \frac{(18+2)!}{18! \times 2!} = \ldots$ ways to invest only to companies 1, 3, 4. Overall, there are $S_1 + S_2 + S_3 + S_4 + S_5 = \ldots$ ways in total.

c. Simple: $\frac{(k+3)!}{k! \times 3!} = \ldots$

4. 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?

Answer. a. There are $100-(27+28+20-12-5-8+2)=48$ students not in any of the 3 classes. The probability of the chosen student is not in any of the 3 classes is $\frac{48}{100}=0.48$.

b. There are $27-12-5+2=12$ students only participate in Java class, there are $28-12-8+2=10$ students only participate in C++ class and there are $100-12-10-48=30$ student in Python only.

c. The answer is: $\binom{52}{1}\times\binom{99}{1}=...$

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

Answer. a.

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

    a. we will roll two different numbers *thrice* (three times) each?

    b. we will roll *exactly one* number *exactly* three times? Hint: Be careful of overcounting.

Answer.

a. There are $\binom{6}{2} = ...$ pair of values to roll and $\frac{6!}{3!.3!} = ...$ ways to arrange these 6 values of these two numbers. So the answer is ....

b.

7. 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?

Answer.

8. 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?

Answer. a. Since we only need to try second time if the first attempt is incorrect, the same analogy applied on the third attempt and so on. Therefore I'm using the chain rule to calculate the probability as:

$$\begin{aligned} P(E) =& P(E_1) + P(E_2|E_1^c) \times P(E_2) + P(E_3|E_1^c, E_2^c) \times P(E_3) + ...+ \\ & P(E_k|E_1^c, E_2^c, .., E_{k-1}^c) \times P(k_3) \\ =& \frac{1}{n} + \end{aligned} \tag{1}$$

where E_i is the event that the password is correct at attempted i.

b. The probability of the hacker successfully on the first attempt is $P(E_k|E_1^c, E_2^c, .., E_{k-1}^c) \times P(k_3) = ...$

9. 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.

Answer. First, we need to choose the k strings that hashed to first bucket, there are $\binom{m}{k}$ selection for this. Then we need to arrange $m - k$ remain strings into other N-1 buckets, there are $\frac{(m-k+N-2)!}{(k-m)!.(N-2)!}$. Overall the answer is $\binom{m}{k} \times \binom{m-k+N-2}{N-2}$

10. 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 3, 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%).

Answer.

11. **[Coding]** See the problem set writeup for this question. Submit your code on Gradescope using the associated coding assignment for this problem set.