Homework 2: More Counting and Probability

January 20, 2025

Zifeng Wang
CSE 312(12973)/Section B

---

1. Combinatorial Identities (16 points)

   Prove each of the following identities using a combinatorial argument (i.e., an argument that counts two different ways); an algebraic solution will be marked substantially incorrect.

   (a) $\sum_{i=0}^{x} \binom{x}{i}\binom{y}{i} = \binom{x+y}{x}$. You may assume that $y \geq x \geq 0$

   (b) $\sum_{i=x}^{y} \binom{y}{i}\binom{i}{x} = \binom{y}{x}2^{y-x}$. You may assume that $y \geq x \geq 0$

Ans:

(a) We aim to count the number of ways to select x people (distinguishable) from a group of x boys and y girls ($y \geq x \geq 0$). This can be done in two different ways:
Way 1: Direct Counting
There are x+y people in total, and we choose x of them directly. The number of ways to do this is: $\binom{x+y}{x}$.
Way 2: Counting for Group
The number of ways to choose i people from y girls is $\binom{y}{i}$.
The number of ways to choose (x - i) people from x boys is $\binom{x}{x-i}$ which is also equal $\binom{x}{i}$.
Thus, for each i, the total number of ways to choose this specific configuration is $\binom{x}{i}\binom{y}{i}$. Summing over all possible values of i gives the total count: $\sum_{i=0}^{x} \binom{x}{i}\binom{y}{i}$
Since both methods count the same thing-the total number of ways to select x people from x boys and y girls we conclude: $\sum_{i=0}^{x} \binom{x}{i}\binom{y}{i} = \binom{x+y}{x}$

(b) There are y people (distinguishable) competing for an internship. There are two rounds of selection: an OA (Online Assessment) and a VO (Virtual Onsite): The OA will eliminate the weakest x people, leaving (y - x) people remaining; During the interview, any of the remaining (y - x) people can be chosen as final candidates. The company is flexible in hiring any number of people between x and (y - x).
Way 1: counting in order

We first select x people from the y total people $\binom{y}{x}$
The remain (y - x) people pass the VO have half-half opportunity get internship so 2^{y-x}
Product Rule: total way is $\binom{y}{x}2^{y-x}$
Way 2: counting by loser
There are possible i people didn't get internship ($x \leq i \leq y$) so $\binom{y}{i}$
And in that i people x people is eliminate by OA: $\binom{i}{x}$
Thus, for each i, the total number of ways to choose this specific configuration is $\binom{y}{i}\binom{i}{x}$. Summing over all possible values of i gives the total count: $\sum_{i=x}^{y}\binom{y}{i}\binom{i}{x}$
Since both methods count the same thing, the total number of ways different people get internship, we conclude: $\sum_{i=x}^{y}\binom{y}{i}\binom{i}{x} = \binom{y}{x}2^{y-x}$.

2. Pigeon Chess (8 points)

The year is 2028 and Elon Musk wants to test his next generation Neuralink chips on 25 pigeons. He builds a 5x5 chessboard and places a pigeon on each square of the chessboard to begin with. At noon, Elon has programmed each chip to make the pigeons randomly walk to a valid adjacent square horizontally or vertically (but not diagonally). Argue that the probability that two or more pigeons end up on the same square is 1.
Hint: It's important to the problem that it's pigeons and not some other bird.

Ans:
There are 25 pigeons on the chess board at first, 13 are at white square and 12 are at black square. According to the problem, pigeons randomly walk to a valid adjacent square horizontally or vertically (but not diagonally). The pigeons on white squares will move to black square and vice versa.

Pigeons: 13 pigeons on the white square
Pigeonholes: 12 black square

Applying the (generalized) pigeonhole principle, there is at least one black square where have at least $\lceil \frac{13}{12} \rceil = 2$ pigeons.

3. Stuff into stuff (12 points)

   (a) We have 20 (distinguishable) cats and 55 (distinguishable) boxes. How many different ways are there to assign the (distinguishable) cats to the (distinguishable) boxes? (Any number of cats can go into any of the 55 boxes.)

   (b) We have 35 identical (indistinguishable) rubber ducks. How many different ways are there to place the rubber ducks into 15 (distinguishable) bathtubs? (Any number of rubber ducks can go into any of the bathtubs.)

   (c) We have 56 identical (indistinguishable) rubber ducks. How many different ways are there to place the rubber ducks into 12 (distinguishable) bathtubs, if each bathtub is required to have at least three rubber ducks in it?

Ans:

(a) To assign 20 distinguishable cats to 55 distinguishable boxes, each cat can independently be placed in any of the 55 boxes. For the first cat, there are 55 choices of boxes. For the second cat, there are also 55 choices of boxes, and so on.
Thus, the total number of ways to assign the cats to the boxes is: 55^{20}

(b) This is a problem of distributing 35 identical rubber ducks into 15 distinguishable bathtubs. This can be solved using the stars and bars theorem. We use (15 - 1) bars separate different bathtubs. We line the 35 rubber ducks and 14 separate bars, there are total (35 + (15 - 1)) = 49 places. We chose 14 place for separate bars from total 49 places.
Thus, the number of ways to place 35 identical rubber ducks into 15 distinguishable bathtubs is: $\binom{35+15-1}{15-1} = \binom{49}{14}$

(c) In this case, we are distributing 56 identical rubber ducks into 12 distinguishable bathtubs, with the condition that each bathtub must contain at least 3 rubber ducks. To handle the condition, we can first place 3 rubber ducks in each of the 12 bathtubs. This will use up: $12 \cdot 3 = 36$ ducks.
Now, we are left with: 56 - 36 = 20 ducks. These 20 remaining ducks can be freely placed into the 12 bathtubs, with no restrictions. Use stars and bars theorem: $\binom{20+12-1}{12-1} = \binom{31}{11}$

4. Sample Spaces and Probabilities (online)

Ans: Gradescope Question

5. Miscounting (14 points)

Consider the question: How many 7-card poker hands (order doesn't matter) are there that contain at least three pairs (pairs means two cards of the same value).
Here is how we might compute this:
To compute the number of hands, apply the product rule using the following sequential process. First pick three ranks that have a pair. For the lowest rank of these, pick the suits of the two cards. Then for the middle rank of these, pick the suits of the two cards, then for the highest rank of these, pick the suits of the two cards. Then out of the remaining 52 - 6 = 46 cards, pick one. Therefore there are $\binom{13}{3} \cdot \binom{4}{2} \cdot \binom{4}{2} \cdot \binom{4}{2} \cdot \binom{46}{1}$ hands.
In this problem, you will find what is wrong with this solution.

(a) Is there overcounting in the solution? That is, is there a hand that can be produced by multiple outcomes of the sequential process? If there is, give one concrete example of such a hand and two outcomes of the process that produce it. If there is not, briefly (1-2 sentences) explain why there isn't.

(b) Is there undercounting in the problem? That is, is there a hand that cannot be produced by any outcomes of the sequential process? If there is, give one concrete example of such a hand and briefly explain why no outcome produces it. If there is no such hand, briefly explain why all hands are produced at least once.

(c) Correct the calculation – in this part you should produce a correct overall formula by subtracting/dividing out any errors that would fit in (a) and adding/multiplying in any errors that would fit in (b).

(d) Find the answer differently – take a different approach to counting this problem (e.g. use a different sequential process). Verify that you get the same number (via a different formula) than the last part.

Ans:

(a) Yes, there is overcounting in the original solution. For example, the sequence of steps for selecting the suits results in counting the same hand more than once. "ace of hearts, ace of diamonds, king of spades, king of diamonds, 7 of spades, 7 of hearts" three pairs and final chose ace of speed; Overcounting case is "ace of speed, ace of diamonds, king of spades, king of diamonds, 7 of spades, 7 of hearts" three pairs and final chose ace of hearts. Also "ace of speed, ace of hearts, king of spades, king of diamonds, 7 of spades, 7 of hearts" three pairs and final chose ace of diamonds will also be a overcounting case for this one. This overcounting only happen when the last is chose as the same rank as one of the first three pairs' rank we chose. This will create 3 same output but count at same time.

(b) No, there is no undercounting. Since we must have three distinct values with at least two cards of that value each. Every hand with exactly three pairs can be formed by choosing: The three ranks for the pairs. The suits for the two cards of each rank. A seventh card from the remaining deck. Thus, all such hands are produced at least once by the sequential process. The given way overcounting the all possible situation no undercounting.

(c) Since the overcounting happens when the last card chosen is of the same rank as one of the pairs already selected, we can subtract the overcounted cases. These cases can be counted by choosing one of the 6 remaining cards from the same rank as one of the first three pairs, and adjusting for the fact that these cases appear 3 times (as the example show in (a)). In order to avoid this overcounting, the final correct formula is $\binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{46}{1} - \binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{6}{1} \cdot \frac{2}{3}$

(d) We can divide the situation to two parts:
1. three different rank pairs with another rank: $\binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{40}{1}$
2. three different rank pairs with a same rank (one have three suit the other two have two suit): $\binom{13}{1} \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot \binom{4}{2}^2$
Thus the all possible 7 hands is $\binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{40}{1} + \binom{13}{1} \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot \binom{4}{2}^2$
Get from Wolfram Alpha:
$\binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{46}{1} - \binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{6}{1} \cdot \frac{2}{3} = 2841696 - 247104 = 2594592$
$\binom{13}{3} \cdot \binom{4}{2}^3 \cdot \binom{40}{1} + \binom{13}{1} \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot \binom{4}{2}^2 = 2471040 + 123552 = 2594592$
So get the same number (via a different formula) than the last part.

6. Guessing Game (8 points)

You are taking a multiple-choice section of the SAT exam. With probability p, you know the answer to the question (and always get it correct). With probability 1 - p, you don't know the answer and guess randomly among the 4 possible options (of which exactly one is correct).

(a) Calculate the probability you get a question correct. Please define events and state which rules/laws you are using to do the calculation.

(b) Given that you got a question correct, what is the probability that you actually knew it (i.e., that you didn't get it correct by guessing)? Please define events and state which rules/laws you are using to do the calculation.

Ans:
Let K be the event that I know the answer to the question.
Let C be the event that answer the correct answer.

The probability of I know the answer is $\mathbb{P}(K) = p$.
The probability of the event I do not know the answer is $\mathbb{P}(\overline{K}) = 1 - p$.
If I know the answer, I can definitely answer correct the answer: $\mathbb{P}(C|K) = 1$
If I don't know the answer, the probability of the problem I can guess correct is $\frac{1}{4}$: $\mathbb{P}(C|\overline{K}) = \frac{1}{4}$

(a) The question ask for $\mathbb{P}(C)$ that the probability you get a question correct. By Law of Total Probability:

$$\mathbb{P}(C) = \mathbb{P}(C|K) \cdot \mathbb{P}(K) + \mathbb{P}(C|\overline{K}) \cdot \mathbb{P}(\overline{K}) = 1 \cdot p + \frac{1}{4} \cdot (1 - p) = \frac{3p + 1}{4}$$

(b) The question ask for $\mathbb{P}(K|C)$ that if you got a question correct, what is the probability that you actually knew.
By Bayes' Rule:

$$\mathbb{P}(K|C) = \frac{\mathbb{P}(C|K) \cdot \mathbb{P}(K)}{\mathbb{P}(C)} = \frac{1 \cdot p}{\frac{3p+1}{4}} = \frac{4p}{3p + 1}$$

7. Real-World: Is this a pigeon? (6 points)

   These 'real world' problems are a bit different from other problems! Because of that we're showing you what our rubric is going to be for this first one so you can see our expectations.

   (a) Your source and some background for the application.

   (b) Relate your proof to the theorem definition from lecture - what are the pigeons and what are the pigeonholes?

   (c) What does the theorem tell us about them?

Ans:

(a) There are total 4 possible kinds of \$12.99 sandwich are sold in Microsoft Café: Chicken Gouda, Muffuletta, Flank Steak, and Plant-Based (check the menu if you pass by, it's real believe me). I am lazy that don't want to make breakfast and lunch by myself, so I will buy one as a brunch every morning from Monday to Friday (it's true at least last week lol).

(b) Let pigeons be the sandwiches I buy throughout the weekdays.
Let pigeonholes be the types of sandwiches available.

(c) Since there are 5 days from Monday to Friday, n = 5. Since there are 4 kinds of sandwiches, k = 4. The Pigeonhole Principle guarantees that: $\lceil \frac{5}{4} \rceil = 2$. At least one type of sandwich will be purchased at least 2 times during the week. In fact, I ate 3 Chicken Gouda, 1 Muffuletta, 1 Flank Steak last week. This is confirmed by the results

8. Feedback (1 points)

- I spent approximately 8 hours on this assignment.
- I spent the most time on Problem 5: Miscounting
- It's really good way to thought about the Pigeonhole problem. I want to think a difficult one but I failed :(And even thought about an easy real one took me a lot of times.
- I also thought about the Pigeon Chess a lot, Until I really look at the chessboard not only the 5*5 square. The color tell things, maybe the hint should be change since we already know it's Pigeonhole problem through the problem title.