Homework 2: More Counting and Probability

In general, your goal in an explanation is to write enough that a student from class who has attended lecture, but not thought about the problem yet, could understand your approach, verify your reasoning, and believe your answer is correct. While we do not usually need to see arithmetic, you must include enough work that in principle one could rederive your answer with only a scientific calculator.

Unless a problem states otherwise, you should leave your answer in terms of factorials, combinations, etc., for instance 26^7 or $26!/7!$ or $26 \cdot \binom{26}{7}$ are all good forms for final answers.

Remember that you must "select pages" to identify which problem is on which page as part of your gradescope upload.

Submission: You must upload a **PDF** of your written solutions to Gradescope under "HW2 [Written]". The use of $\LaTeX$ is *required*.

Due Date: This assignment is due at **11 PM Wednesday July 5**.
You will submit the written problems as a PDF to Gradescope. Please put each numbered problem on its own page of the PDF (this will make selecting pages easier when you submit), and ensure that your PDFs are oriented correctly (e.g. not upside-down or sideways). The coding problem will also be submitted to Gradescope.

Note that there will be **no leeway** with the submission time. Any submission to Gradescope after 11:00:00 PM on the deadline day will be considered late. We recommend submitting a few minutes before the deadline in order to avoid technical issues with submissions.

Collaboration: Please read the full collaboration policy. If you work with others, you must still write up your solution independently and name all of your collaborators somewhere on your assignment.

Full submission instructions are in the course syllabus.

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.

For the purposes of these problems, using commutativity of multiplication and addition (i.e. $ab = ba$, $a + b = b + a$), and distributivity/factoring ($a(b + c) = ab + ac$) are allowed as part of a combinatorial argument. Any other algebra facts (e.g. Pascal's Rule about combinations, the definition of combinations/permutations in terms of factorials, canceling numbers that appear in numerators and denominators) would make it an algebraic solution, not a combinatorial one.

(a) $\displaystyle\sum_{k=0}^{m} \binom{m}{k}\binom{n}{k} = \binom{m+n}{m}$. You may assume that $n \geq m \geq 0$. Hint: Start with the right-hand side and imagine you are choosing a team of m people from a group of people consisting of n Americans and m Canadians. (8 points)

(b) $\displaystyle\sum_{k=m}^{n} \binom{n}{k}\binom{k}{m} = \binom{n}{m}2^{n-m}$. Assume that $n \geq m \geq 0$. Hint: Think about choosing a committee from n people of which some subset are leaders.

2. Feeling Squabished [10 points]

At a dinner party, all of the n people present are to be seated at a circular table. Suppose there is a nametag at each place at the table and suppose that nobody sits down in their correct place. Use the pigeonhole principle to show that it is possible to rotate the table so that at least two people are sitting in the correct place. Be sure to specify precisely what the pigeons are, precisely what the pigeonholes are, and precisely what the mapping of pigeons to pigeonholes is.

3. Club Sandwich [12 points]

(a) We have 10 distinguishable clubs and 55 distinguishable students. How many different ways are there to assign the students to the clubs? (Any number of students can be a member of any of the 10 clubs.) (4 points)

(b) We have 42 identical (indistinguishable) sandwiches. How many different ways are there to place the sandwiches into 12 (distinguishable) boxes? (Any number of sandwiches can go into any of the boxes.) (4 points)

(c) We have 42 identical (indistinguishable) sandwiches. How many different ways are there to place the sandwiches into 7 (distinguishable) boxes, if each box is required to have at least three sandwiches in it? (4 points)

4. In PIE We Crust [10 points]

A router in the network has three paths along which it can send packets. It has 185 distinct packets that need to be randomly routed along one of the three paths, independently. In how many ways can the packets be distributed so that all of the paths have at least 1 packet traversing them, assuming the order of the packets distributed on one path doesn't matter?

5. What are the Odds? [18 points]

For each of the following scenarios,

(i) describe the sample space,

(ii) indicate how big it is (i.e., what its cardinality is),

(iii) and answer the probability question.

(a) You flip a fair coin 60 times. What is the probability of exactly 25 heads? (3 points)

(b) You roll two fair 6-sided dice, one red and one blue. What is the probability that the sum of the two values showing is 5? (3 points)

(c) You are given a random 5 card poker hand (selected from a single deck with 52 cards with 13 ranks and 4 suits). What is the probability you have a full-house (3 cards of one rank and 2 cards of another rank)? (3 points)

(d) 25 labeled balls are placed into 12 labeled bins (with each placement equally likely). What is the probability that bin 1 contains exactly 3 balls? (3 points)

(e) There are 25 computer scientists and 42 statisticians attending a certain conference. 3 of these 67 people are randomly chosen to take part in a panel discussion. What is the probability that at least one computer scientist is chosen? (3 points)

(f) You buy 12 donuts choosing from 3 different flavors (glazed, jelly, and chocolate). Donuts of the same type are indistinguishable. For this problem, each distinguishable collection of donuts is equally likely (for example, the probability of getting 12 chocolate donuts is equal to the probability of getting 4 glazed, 6 jelly, and 2 chocolate donuts in any order). What is the probability that you have at least one donut of each type? (3 points)

6. Miscounting [14 points]

Consider the question: How many **7-card** poker hands (order doesn't matter) are there that contain at least two 3-of-a-kinds (3-of-a-kind means three cards of the same value)? For example, this would be a valid hand: ace of hearts, ace of diamonds, ace of spaces, 7 of clubs, 7 of spades, 7 of hearts, and queen of clubs. (Note that a hand consisting of all 4 aces and three of the 7s is also valid.)

Here is how we might compute this:

> To compute the number of hands, apply the product rule using the following sequential process. First pick two ranks that have a 3-of-a-kind (e.g. ace and 7 in the example above). For the lower rank of these, pick the suits of the three cards. Then for the higher rank of these, pick the suits of the three cards. Then out of the remaining $52 - 6 = 46$ cards, pick one. Therefore there are
>
> $$\binom{13}{2} \cdot \binom{4}{3} \cdot \binom{4}{3} \cdot \binom{46}{1}$$
>
> 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. (4 points)

(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. (3 points)

(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). (3 points)

(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. (4 points)

7. The Probability of Coincidences [5 points]

How many times have you bumped into someone who you know somewhere you didn't expect to see them and thought - what are the chances? In this section, we'll take a look at exactly what the chances are using our knowledge of probability to do a back-of-the-napkin calculation.

You decide to check out the U-District Street Fair over the weekend. There are around $50,000$ people from the university visiting the fair. You can recognize around 50 people from the university. Throughout the day, you see 200 visitors, a subset of all visitors to the fair. Suppose that every size-200 subset of the population is equally likely to be seen. What are the odds that you see at least 1 person you know? Are the chances as low as you thought?

8. Is this a pigeon? [Extra Credit]

The pigeonhole principle is surprisingly powerful and can be used to prove a variety of things – find one application of the principle that interests you (it can be CS-related or not, serious or not – my only requirements are that (a) you must find it interesting and (b) we shouldn't have proved it already in this course) and explain the proof to us. Make sure to clearly relate your proof to the theorem definition from lecture - what are the pigeons, what are the pigeonholes and what does the theorem tell us about them? (For this question only, you can feel free to Google without limits, as long as you cite your sources).