Homework 7: Sets, PIE, PP, Counting

Problem 1 Pigeonhole Principle

New England (consisting of the states MA, VT, ME, NH & CT) is home to over 14 million people. Show that there are at least 3 people in New England with the same four initials who were also born in the same month of the year (though not necessarily in the same year). By "four initials" we mean the first two letters of the first name and the first two letters of the last name (e.g. for Sally Post the four initials would be "SAPO")

Problem 2 Pigeonhole Principle for Sequences Mod n

i. Let $a_1, a_2, \cdots a_n$ be a sequence of n integers. Show that either there is an a_i that is divisible by n or there are two different numbers a_i, a_j such that their difference $a_i - a_j$ is divisible by n.

ii. ★ Let $x_1, x_2, \cdots x_n$ be a sequence of n integers. Show that there is a consecutive subsequence $x_i, x_{i+1}, \cdots x_j$ ("chunk [i:j]") such that the sum of its terms $S_{ij} = \sum_{k=i}^{j} x_k$ is divisible by n.

Problem 3 Printer Arrangements

An interior designer is considering function and appeal of different arrangements of printers for a library. He has to arrange 3 black & white printers and 5 color printers in a row. Note: all 8 printers are unique—the black & white printers are each different models, and similarly for the color printers.

i. In how many ways can the printers be arranged so that all the color printers are next to one another?

ii. In how many ways can the printers be arranged so that all the color printers are next to one another, or all the black & white printers are next to one another, or both?

iii. In how many ways can the printers be arranged so that no two black & white printers are adjacent?

Problem 4 A TV Station's Audience

Suppose that a polling company conducts a survey of the audience of a TV station. The company has surveyed a sample of 1000 of the station's subscribers which consists of 450 men, 500 people 35 years of age or younger, and where 400 people of the sample have a college degree. They have also found that 150 of the men in the sample have a college degree, 250 women are over 35 years old, and that among the college graduates 250 were over 35. Finally, they were surprised at the relatively large number of uneducated men over 35 who subscribed to the station's programming: 150.

i. An advertiser targets females or people who have a college degree. How many people like that are in the sample?

ii. Another advertiser targets men or people who do not have a college degree. How many people like that are there?

iii. The most lucrative demographics are the women who are at most 35 years of age, and who have a college degree. How many of such people are in the sample?

Problem 5 ★ Restricted possibilities - poor's man count

In class, recall that we solved problems similar to the following: How many 6-symbol passwords can you generate with 26 letters and 10 digits, such that every password contains at least one digit? We showed that one way to do this is to subtract the 26^6 invalid all-letter passwords from the count of all possible passwords using any 6 letters or numbers, 36^6.

However, student Jack Sparrow started on a different solution: if there has to be a mandatory digit d, that can be in anyone of the six positions 1,2,3,4,5,6. If d is in position 1, then there are 10 possibilities for that digit value, and 36^5 valid possibilities for the rest of 5 symbols, total $10 * 36^5$. Similarly if d is in second position, there are $10 * 36^5$ valid passwords; same for d in the third position and so on.

Jack needs to make a union of all these valid possibilities. He adds the sizes to $6 * 10 * 36^5$ total, but now he needs to subtract some, because many of them have been counted more than once. For example, any string with exactly 3 digits (looks like `dLddLL` or `dddLLL` or `dLdLdL` where L is any letter) have been counted three times, once for each digit position.

Your task is to help Jack correctly subtract the possibilities counted more than once, and to prove that the resulting calculation gives the correct answer $36^6 - 26^6$.

i. Consider the example of inclusion-exclusion with 4 random sets S_1, S_2, S_3, S_4. We could write the inclusion-exclusion principle as

$$|\cup_{i=1}^4 S_i| = \sum_{1\le i\le 4} |S_i| - \sum_{1\le i,j\le 4} |S_i \cap S_j| + \sum_{1\le i,j,t\le 4} |S_i \cap S_j \cap S_t| - \sum_{1\le i,j,t,z\le 4} |S_i \cap S_j \cap S_t \cap S_z|$$

In this example the first group (with plus sign) is a sum of 4 set sizes, and the second group (with minus sign) is a sum of 6 pair-intersections. The third group is a sum (with plus sign) over ... how many intersections? How about the fourth group? Where can you find all four numbers together in Pascal's Triangle?

ii. Now for Jack's problem: Write the inclusion-exclusion principle on these 6 sets:
A_i= {6-symbol sequences with digit in position i }; $1 \le i \le 6$
You should group the intersections in the formula by how many sets are intersected, just like it is shown in the example in the previous part for 4 sets. For each number of intersection symbols from zero (sets not being intersected) to 5 (all sets intersected), give the number of terms you would have if you wrote out all the terms with that many intersection symbols. (For example, the answer is 6 for the zero-intersection-symbols case.)

iii. Let's look at the terms with intersections of $k = 3$ sets. Think about what kinds of passwords are in these intersections in Jack's problem. How would you describe the passwords in the intersections of three sets in the above inclusion-exclusion?

iv. Thinking further about the inclusion-exclusion applied to 6-symbol passwords: for any $1 \leq i, j, t \leq 6$, the intersection size of sets A_i, A_j, and A_t (defined earlier as the sets that definitely have digits in those places) does not depend on the triplet (i, j, t). Every intersection like this has the same size; thus you could write the size of the sum group as
`how-many-intersections * intersection-size`
How many passwords are there across all intersections of exactly three sets of passwords in the above inclusion-exclusion? (You don't need to write out the number; giving the terms to multiply will do.)

v. Generalize your calculation from the previous part, turning it into a formula that works for any k (number of intersected sets) between 1 and 6, producing the correct number of passwords overall in each case. Include $(-1)^{k+1}$ in your formula, so that it correctly generates either a positive or negative term in the inclusion/exclusion.

vi. Show that evaluating your formula for each term of the inclusion/exclusion produces the same answer as the original in-class approach: $36^6 - 26^6$.

Problem 6 One to one map of subsets

Assume finite set X, $|X| = n$, includes elements $a \neq b$. Take A={all subsets of X including a} and B={all subsets of X including b}.

i. What are the sizes $|A|$ and $|B|$?

ii. Construct a one-to-one map from A to B.

iii. Is (A, B) a partition of the powerset $\mathbb{P}(X)$?

Problem 7 Divisibility and indexed sets

Recall that $n \in \mathbb{Z}$ is divisible by k if there exists $b \in \mathbb{Z}$ such that $n = bk$. When counting multiples of k in a given range, it is often easier (and safer) to index the set. For example, the set of integers divisible by $k = 7$ between 1 and 60 is
$A = \{7, 14, 21..., 49, 56\} = \{7i | i \in \mathbb{Z}, 1 \leq i \leq 8\} = \{7i | i = 1 : 8\}$, the last expression being the set indexed by i from 1 to 8. Once a set is indexed <u>starting at 1</u>, it is easy to count: set A has 8 elements.
Another example: let's say we want to count the set of integers divisible by 13 between 100 and 300. We index the set as $B = \{104, 117, 130..., 299\} = \{13j + 91 | j \in \mathbb{Z}, 1 \leq j \leq 16\}$. Verify the first `13*1+91=104` and the last `13*16+91=299`. Since indices go from 1 to 16, we have $|B| = 16$.

i. How many positive integers from 1 to 500 are divisible by 2 and 3 but not 5?

ii. How many positive integers from 1 to 500 are divisible by 2 or 3 or 11?

iii. What is the least number of distinct integers we can choose between 1 and 500, that guarantees that at least one of them is divisible by 7?

Problem 8 Permutations and inversions
Consider sequence $A = (1, 2, 3, 4, 5, 6, 7)$.

- **i.** How many permutations of A do not have have either 1 or 2 at the beginning, at the end, or in the exact-middle ?
- **ii.** How many permutations of A contain the subsequence (1,2,3) in this original order (not necessarily in consecutive positions) ?
- **iii.** For a permutation of A, an inversion is a pair (a,b) that is not in the original order: a comes before b in the permutation, but b was before a in the original order. (If they get moved around but are still in the correct order relative to each other, they are not inverted.) What pairs are inversions in the permutation $B = (6, 4, 1, 7, 5, 2, 3)$? How many inversions?
- **iv.** What is the maximum number of inversions a permutation of A can have ? Which permutation achieves that?
- **v.** It turns out that for any two elements (a,b), half of all permutations invert the elements, and half leave them in the original order. What is the average number of inversions per permutation, across all permutations of A?
- **vi.** A permutation has a pivot p if all values smaller than p are listed before p in the permutation, and all values larger than p are listed after p. Show that p must be in its original position (p) and that p is not part of any inversion pair.
- **vii.** How many permutations of sequence A have 5 as pivot?

Problem 9 ★★ (optional, no credit) Divisible Sum
In how many ways one can pick a set of three positive numbers smaller than 100 with their sum a multiple of 7?

Problem 10 ★★★ (optional, no credit) Max number of tournament tri-cycles.
There are $2n + 1$ players in a round-robin tournament; each player plays every other player exactly once, no ties. We say the players (x, y, z) form a cyclic triplet if x beats y, y beats z, and z beats x. Determine the maximum number of cyclic triplets possible as a function of n.

Problem 10 ★★★ (optional, no credit) Number of permutaions with half-gap
For any given $n = gap$, consider the permutations of $1 : 2n$ indices as $(x_1, x_2, ...x_{2n})$ partitioned as
$GOOD =$ set the permutations where there are two consecutive values have gap n: $|x_{i+1} - x_i| = n$
$BAD =$ (complement) set of prermutations where no pair of consecutive values has gap n
Show that $|GOOD| > |BAD|$
part (B) ★★★★ estimate $|GOOD|$ as a proportion of the total $(2n)!$ permutations.