Math 310: Elementary Combinatorics

Sections 1 & 2

Penn State University

Spring 2009

Group Project:
The Inclusion-Exclusion Principle

1 Introduction

One of the first methods of counting we learned was the sum rule. This rule depended on our ability to separate objects into smaller collections so that no one object was placed into two or more collections (i.e., we partitioned a set into disjoint subsets). What happens though if it is more convenient to partition a set into subsets that are not necessarily disjoint? For this project, you will examine a method that systematically compensates for objects that are placed into multiple subsets by applying both the sum and difference rules. This method is known as the Principle of Inclusion-Exclusion.

To give you an idea of the types of objects we will be able to count using this principle, consider the following variations of problems we have already encountered.

- Seat n brother/sister pairs in a row of chairs or at a circular table so that no child is sitting next to his/her sibling.
- Distribute n distinct pieces of candy to m children so that every child gets at least one piece of candy or equivalently, determine the number of onto functions from $[n]$ to $[m]$.
- Distribute r identical pieces of candy to k children so that no child gets n or more pieces of candy.
- Deliver n letters addressed to n different people so that everyone receives exactly one letter but nobody receives the letter that was addressed to them.
- r-combinations of the multiset $\{n_1 \cdot 1, n_2 \cdot 2, \ldots, n_k \cdot k\}$.

2 Rules and Regulations

Students will work in groups consisting of no more than 4 people. Each group will submit one typed report, including a cover page with project title and names of group members. Your report should have an introduction, giving a brief description of the project and any other information or basic definitions/assumptions you will be using that are relevant to the entire project. Each of the three parts of the project should also include a brief introduction describing the specific experiment and any other information that is relevant to that part of the project.

The group report will be graded out of 75 points (15% of your overall grade). Each project consists of three parts (see below). Part I is worth 20 points, Part II is worth 20 points and Part III is worth 20 points. Each of these parts will be graded for mathematical detail and accuracy. The remaining 15 points will be awarded based on presentation (clarity and consistency of writing style, grammatical correctness, appropriate use of notation, terminology, and examples, etc.)

Make sure to tell me who is in your group and which project you will be working on by **Friday, March 27**. During the week of **April 13 through April 17** or before, your group must show me a rough draft of your report, at which point I will make any necessary comments and/or suggestions. **The deadline to submit your completed project is Monday, April 27.**

3 The Project

Part I:

In this first part of the group project, you will discover the principle of Inclusion-Exclusion while at the same time determining the number onto functions from $[n]$ to $[m]$ or equivalently, the number of ways to distribute n distinct pieces of candy to m children so that every child gets at least one piece of candy.

Write up an introduction to the principle of inclusion-exclusion, using the following questions as a guide. Feel free to organize the content in any way you see fit, making sure to address each of the following questions. Do not simply answer each question. I have phrased each question in the context of onto functions, but feel free to write up your introduction in the context of distributing candy as described in the previous paragraph.

1. What is the number of onto functions $f : [n] \to [1]$ for any $n \geq 1$. Or, how many ways are there to distribute n distinct pieces of candy to 1 child? Recall that $[n] = \{1, 2, 3, \ldots, n\}$.

2. Use the difference rule to find a formula for the number of onto functions $f : [n] \to [2]$ for any $n \geq 2$. Or, how many ways are there to distribute n distinct pieces of candy to 2 children so that both children get at least one piece of candy?

3. Use the sum rule and your answer to the previous question to compute the number of functions $f : [n] \to [3]$ that are *not* onto. Find a formula for the number of onto functions $f : [n] \to [3]$ for any $n \geq 3$. Write your answer as a sum/difference where each term is the product of a binomial coefficient and the power of an integer.

4. Repeat the previous problem for functions $f : [n] \to [4]$.

5. Conjecture a formula for the number of onto functions from $[n]$ to $[m]$.

We say that a function $f : [n] \to [m]$ has property P_i if $f(j) \neq i$ for all $1 \leq j \leq n$. In other words, if i is not in the image of f, then f has property P_i. For example, the function $f : [2] \to [4]$ defined by $f(x) = x^2$ has properties P_2 and P_3 since $f(1) = 1$ and $f(2) = 4$.

In the context of distributing candy, we say that the distribution has property P_i if the ith youngest child did *not* receive any candy.

Let S be a subset of the collection of properties $P = \{P_1, P_2, \ldots, P_m\}$. Define $N_a(S)$ to be the number of functions that have *at least* all of the properties in S (they may have other properties as well). Let $N_e(S)$ be the number of functions that have *exactly* the properties in S (and no more). So for example, there are two functions $f : [2] \to [4]$ that have exactly properties P_2 and P_3 and thus $N_e(\{P_2, P_3\}) = 2$ in this case.

6. Consider all of the functions $f : [n] \to [m]$ and define properties P_1 through P_m as above. Compute $N_a(\{P_1\})$ and $N_a(\{P_1, P_2\})$. Find an explicit formula for $N_a(S)$ for any S. Explain why your answer depends only on the cardinality of S, and not the actual elements of S.

7. For $m = 1, 2, 3$, find a formula for the number of onto functions $f : [n] \to [m]$ in terms of only the quantities $N_a(S)$ for various values of S. Make sure to include the quantity $N_a(P)$ in your answer. Use your answers to questions 1,2, and 3 as a guide.

8. How can you describe the number of onto functions $f : [n] \to [m]$ in terms of $N_e(S)$ for some set S? In other words, if f is onto, what set of properties would it have?

9. Using your answers to the last two questions, what connection have you found between the quantities N_e and N_a? It may be easiest to express this relationship as a summation over all subsets $S \subseteq P$. The identity you have just discovered is true in general and not just in the specific case of onto functions with the properties $P_1, \ldots, P_m$ as described above. It is known as the Principle of Inclusion-Exclusion. State this principle in full generality.

10. In general, $N_a(S)$ depends on which elements are in S and not just the number of elements in S. In the case when $N_a(S)$ only depends on the cardinality of S, we will denote this quantity by $N_a(k)$ (i.e., the number of elements that have at least k properties), where $k = |S|$. How would you state the Principle of Inclusion-Exclusion when $N_a(S)$ depends only on the cardinality of S. Verify the conjecture made in Problem 5.

11. Count the number of onto functions from $[n]$ to $[n]$ in two different ways. What identity have you discovered?

Part II:

Consider the situation where n brother/sister pairs sit in a row of $2n$ chairs for $n \geq 2$. In how many ways can the children sit down so that no brother/sister pair is sitting next to each other? Verify your answer for $n = 2$. What if additionally, the seating arrangement must alternate gender? (Hint: Seat the girls first. Then find the number of ways to select k chairs in a row of $2n$ chairs such that no two of the k chairs are next to each other. Use this to determine which k boys must sit next to their sister and at the same time, which chair each of these k boys will sit in.) Verify your answer for $n = 2$ or $n = 3$.

Now answer the previous questions under the assumption that the children are sitting down at a circular table instead of a row of chairs.

Part III:
Answer any three of the following four problems using inclusion-exclusion.

1. Find a formula for the number of r-combinations of the multiset

   $$\{n_1 \cdot 1, n_2 \cdot 2, n_3 \cdot 3\}.$$

   Use your answer to find the number of 18-combinations of the multiset $\{8 \cdot a, 4 \cdot b, 12 \cdot c\}$.

2. How many ways are there to distribute r identical pieces of candy to k children so that no child receives n or more pieces of candy? In the case when $n = 1$, what binomial identity have you discovered?

3. Determine the number of ways that n letters addressed to n different people can be delivered so that everyone receives a letter but nobody receives the letter that was addressed to them. Verify your answer for $n = 3$. If the n letters are distributed at random, what is the probability (see the other project description for a brief introduction to calculating probabilities) that nobody receives the letter that was addressed to them? What happens to the value of this probability as n tends to infinity?

4. The Euler ϕ-function, $\phi(n)$, yields the number of positive integers less than n that are relatively prime to $n \geq 2$. In other words, it is the number of positive integers less than n that do no have any factors in common with n. For example, $\phi(14) = 6$ since each of the numbers in the set $\{1, 3, 5, 9, 11, 13\}$ do not have any factors in common with 14. If p is prime, then $\phi(p) = p - 1$ since all of the integers from 1 to $p - 1$ are relatively prime to p.

   Suppose that the prime factorization of the integer n is given by:

   $$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$

   where $p_1, p_2, \ldots, p_k$ are distinct prime numbers and $e_i \geq 1$ for $i = 1, 2, \ldots, k$. Show that

   $$\phi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_k}\right).$$