CHAPTER

SIX

Inclusion-Exclusion

In this chapter, we study a classic enumeration technique known as Inclusion-Exclusion. In its simplest case, it is absolutely intuitive. Its power rests in the fact that in many situations, we start with an exponentially large calculation and see it reduce to a manageable size. We focus on three applications that every student of combinatorics should know: (1) counting surjections; (2) derangements; and (3) the Euler ϕ-function.

6.1 Introduction

We start this chapter with an elementary example.

Example 6.1. Let X be the set of 63 students in an applied combinatorics course at a large technological university. Suppose there are 47 computer science majors and 51 male students. Also, we know there are 45 male students majoring in computer science. How many students in the class are female students not majoring in computer science?

Although the Venn diagrams that you've probably seen drawn many times over the years aren't always the best illustrations (especially if you try to think with some sort of scale), let's use one to get started. In Figure 6.1, we see how the groups in the scenario might overlap. (Of course, there really are no computer science majors who are neither male nor female, but empty regions really aren't a problem.)

Now we can see that we're after the number of students in the magenta region on the far right, which is the female students not majoring in computer science. To compute this, we can start by subtracting the number of male students (the blue region) from the total number of students in the class and then subtracting the number of computer science majors (the yellow region). However, we've now subtracted the green region (the male computer science majors) *twice*, so we must add that number back. Thus, the

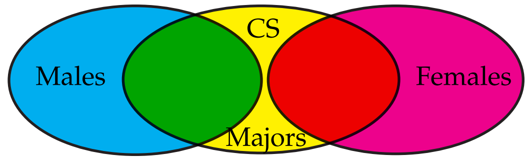

FIGURE 6.1: A VENN DIAGRAM FOR AN APPLIED COMBINATORICS CLASS

number of female students in the class who are not majoring in computer science is

$$63 - 51 - 47 + 45 = 10.$$

Example 6.2. Another type of problem where we can readily see how such a technique is applicable is a generalization of the problem of enumerating integer solutions of equations. In chapter 2, we discussed how to count the number of solutions to an equation such as

$$x_1 + x_2 + x_3 + x_4 = 100,$$

where $x_1 > 0$, $x_2, x_3 \geq 0$ and $2 \leq x_4 \leq 10$. However, we steered clear of the situation where we add the further restriction that $x_3 \leq 7$. The previous example suggests a way of approaching this modified problem.

First, let's set up the problem so that the lower bound on each variable is of the form $x_i \geq 0$. This leads us to set up the problem as

$$x_1' + x_2 + x_3 + x_4' = 97$$

with $x_1', x_2, x_3, x_4' \geq 0$, $x_3 \leq 7$, and $x_4' \leq 8$. (We'll then have $x_1 = x_1' + 1$ and $x_4 = x_4' + 2$ to get our desired solution.) Now we can count the solutions we want by starting with the number of solutions to $x_1' + x_2 + x_3 + x_4' = 97$ with all variables nonnegative, which is $C(100, 3)$, and subtracting the number in which $x_3 > 7$ ($C(92, 3)$) and the number in which $x_4' > 8$ ($C(91, 3)$). But now we've subtracted the number in which both $x_3 > 7$ and $x_4' > 8$ twice, so we need to add that number ($C(83, 3)$) back to even up accounts. Thus, the number of solutions is

$$\binom{100}{3} - \binom{92}{3} - \binom{91}{3} + \binom{83}{3} = 6516.$$

From these examples, you should start to see a pattern emerging that leads to a more general setting. In full generality, we will consider a set X and a family $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ of *properties*. We intend that for every $x \in X$ and each $i = 1, 2, \ldots, m$, either x satisfies P_i or it does not. There is no ambiguity. Ultimately, we are interested in determining the number of elements of X which satisfy *none* of the properties in $\mathcal{P}$.

Let's consider three examples of sets of properties. These properties will come back up during the remainder of the chapter as we apply inclusion-exclusion to some more involved situations.

Example 6.3. Let m and n be fixed positive integers and let X consist of all functions from $[n]$ to $[m]$. Then for each $i = 1, 2, \ldots, m$, and each function $f \in X$, we say that f satisfies P_i if i is not in the range of f.

As a specific example, suppose that $n = 5$ and $m = 3$. Then the function given by the table

i	1	2	3	4	5
$f(i)$	2	3	2	2	3

satisfies P_1 but not P_2 or P_3.

Example 6.4. Let m be a fixed positive integer and let X consist of all bijections from $[m]$ to $[m]$. Elements of X are called *permutations*. Then for each $i = 1, 2, \ldots, m$, and each permutation $\sigma \in X$, we say that σ satisfies P_i if $\sigma(i) = i$.

For example, the permutation σ of [5] given in by the table

i	1	2	3	4	5
$\sigma(i)$	2	4	3	1	5

satisfies P_3 and P_5 and no other P_i.

Note that in the previous example, we could have said that σ satisfies property P_i if $\sigma(i) \neq i$. But remembering that our goal is to count the number of elements satisfying none of the properties, we would then be counting the number of permuations satisfying $\sigma(i) = i$ for each $i = 1, 2, \ldots, n$, and perhaps we don't need a lot of theory to accomplish this task—the number is one, of course.

Example 6.5. Let m and n be fixed positive integers and let $X = [n]$. Then for each $i = 1, 2, \ldots, m$, and each $j \in X$, we say that j satisfies P_i if i is a divisor of j.

At first this appears to be the most complicated of the sets of properties we've discussed thus far. However, being concrete should help clear up any confusion. Suppose that $n = m = 15$. Which properties does 12 satisfy? The divisors of 12 are 1, 2, 3, 4, 6, and 12, so 12 satisfies P_1, P_2, P_3, P_4, P_6, and P_{12}. On the other end of the spectrum, notice that 7 satisfies only properties P_1 and P_7, since those are its only divisors.

6.2 The Inclusion-Exclusion Formula

Now that we have an understanding of what we mean by a property, let's see how we can use this concept to generalize the process we used in the first two examples of the previous section.

Let X be a set and let $\mathcal{P} = \{P_1, P_2, \ldots, P_m\}$ be a family of properties. Then for each subset $S \subseteq [m]$, let $N(S)$ denote the number of elements of X which satisfy property P_i for all $i \in S$. Note that if $S = \emptyset$, then $N(S) = |X|$, as every element of X satisfies every property in S (which contains no actual properties).

In the examples of the previous section, we subtracted off $N(S)$ for the sets S of size 1 and then added back $N(S)$ for the set of properties of size 2, since we'd subtracted the number of things with both properties twice. Suppose that we had three properties P_1, P_2, and P_3. How would we count the number of objects satisfying none of the properties? First we would subtract for each of P_1, P_2, and P_3. But then we would have accounted for the objects satisfying both P_1 and P_2, both P_2 and P_3, and both P_1 and P_3 twice, so we need to add them back. However, now we've included those satisfying all three properties too many times! They were excluded three times when we subtracted $N(\{i\})$ for $i = 1, 2, 3$ but then we put them back three times when we added $N(\{1, 2\})$, $N(\{2, 3\})$, and $N(\{1, 3\})$. Thus, we must yet subtract $N(\{1, 2, 3\})$ to get the desired number:

$$N(\emptyset) - N(\{1\}) - N(\{2\}) - N(\{3\}) + N(\{1, 2\}) + N(\{2, 3\}) + N(\{1, 3\}) - N(\{1, 2, 3\}).$$

We can generalize this as the following theorem:

Theorem 6.6 (Principle of Inclusion-Exclusion)**.** *The number of elements of X which satisfy none of the properties in $\mathcal{P}$ is given by*

$$\sum_{S \subseteq [m]} (-1)^{|S|} N(S). \tag{6.1}$$

Proof. We proceed by induction on the number m of properties. If $m = 1$, then the formula reduces to $N(\emptyset) - N(\{1\})$. This is correct since it says just that the number of elements which do not satisfy property P_1 is the total number of elements minus the number which do satisfy property P_1.

Now assume validity when $m \leq k$ for some $k \geq 1$ and consider the case where $m = k + 1$. Let $X' = \{x \in X : x \text{ satisfies } P_{k+1}\}$ and $X'' = X - X'$ (i.e., X'' is the set of elements that do not satisfy P_{k+1}). Also, let $\mathcal{Q} = \{P_1, P_2, \ldots, P_k\}$. Then for each subset $S \subseteq [k]$, let $N'(S)$ count the number of elements of X' satisfying property P_i for all $i \in S$. Also, let $N''(S)$ count the number of elements of X'' satisfying property P_i for each $i \in S$. Note that $N(S) = N'(S) + N''(S)$ for every $S \subseteq [k]$.

Let X'_0 denote the set of elements in X' which satisfy none of the properties in $\mathcal{Q}$ (in other words, those that satisfy only P_{k+1} from $\mathcal{P}$), and let X''_0 denote the set of elements of X'' which satisfy none of the properties in $\mathcal{Q}$, and therefore none of the properties in $\mathcal{P}$.

Now by the inductive hypothesis, we know

$$|X'_0| = \sum_{S \subseteq [k]} (-1)^{|S|} N'(S)$$

and

$$|X''_0| = \sum_{S \subseteq [k]} (-1)^{|S|} N''(S).$$

It follows that

$$\begin{aligned}
|X_0''| &= \sum_{S\subseteq[k]} (-1)^{|S|} N''(S) \\
&= \sum_{S\subseteq[k]} (-1)^{|S|} \left(N(S) - N'(S)\right) \\
&= \sum_{S\subseteq[k]} (-1)^{|S|} N(S) + \sum_{S\subseteq[k]} (-1)^{|S|+1} N(S\cup\{k+1\}) \\
&= \sum_{S\subseteq[k+1]} (-1)^{|S|} N(S).
\end{aligned}$$

□

6.3 Enumerating Surjections

As our first example of the power of inclusion-exclusion, consider the following situation: A grandfather has 15 distinct lottery tickets and wants to distribute them to his four grandchildren so that each child gets at least one ticket. In how many ways can he make such a distribution? At first, this looks a lot like the problem of enumerating integers solutions of equations, except here the lottery tickets are not identical! A ticket bearing the numbers 1, 3, 10, 23, 47, and 50 will almost surely not pay out the same amount as one with the numbers 2, 7, 10, 30, 31, and 48, so who gets which ticket really makes a difference. Hopefully, you have already recognized that the fact that we're dealing with lottery tickets and grandchildren isn't so important here. Rather, the important fact is that we want to distribute distinguishable objects to distinct entities, which calls for counting functions from one set (lottery tickets) to another (grandchildren). In our example, we don't simply want the total number of functions, but instead we want the number of surjections, so that we can ensure that every grandchild gets a ticket.

For positive integers n and m, let $S(n,m)$ denote the number of surjections from $[n]$ to $[m]$. Note that $S(n,m) = 0$ when $n < m$. In this section, we apply the Inclusion-Exclusion formula to determine a formula for $S(n,m)$. We start by setting X to be the set of all functions from $[n]$ to $[m]$. Then for each $f \in X$ and each $i = 1, 2, \dots, m$, we say that f satisfies property P_i if i is not in the range of f.

Lemma 6.7. *For each subset $S \subseteq [m]$, $N(S)$ depends only on $|S|$. In fact, if $|S| = k$, then*

$$N(S) = (m-k)^n.$$

Proof. Let $|S| = k$. Then a function f satisfying property P_i for each $i \in S$ is a string of length n from an alphabet consisting of $m-k$ letters. This shows that

$$N(S) = (m-k)^n.$$

□

Now the following result follows immediately from this lemma by applying the Principle of Inclusion-Exclusion, as there are $C(m,k)$ k-element subsets of $[m]$.

Theorem 6.8. *The number $S(n,m)$ of surjections from $[n]$ to $[m]$ is given by:*

$$S(n,m) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} (m-k)^n.$$

For example,

$$\begin{aligned} S(5,3) &= \binom{3}{0}(3-0)^5 - \binom{3}{1}(3-1)^5 + \binom{3}{2}(3-2)^5 - \binom{3}{3}(3-3)^5 \\ &= 243 - 96 + 3 - 0 \\ &= 150. \end{aligned}$$

Returning to our lottery ticket distribution problem at the start of the section, we see that there are $S(15,4) = 1016542800$ ways for the grandfather to distribute his 15 lottery tickets so that each of the 4 grandchildren receives at least one ticket.

6.4 Derangements

Now let's consider a situation where we can make use of the properties defined in Example 6.4. Fix a positive integer n and let X denote the set of all permutations on $[n]$. A permutation $\sigma \in X$ is called a *derangement* if $\sigma(i) \neq i$ for all $i = 1, 2, \ldots, n$. For example, the first permutation given below is a derangement, while the second is not.

i	1	2	3	4
$\sigma(i)$	2	4	1	3

i	1	2	3	4
$\sigma(i)$	2	4	3	1

If we again let P_i be the property that $\sigma(i) = i$, then the derangements are precisely those permutations which do not satisfy P_i for any $i = 1, 2, \ldots, n$.

Lemma 6.9. *For each subset $S \subseteq [n]$, $N(S)$ depends only on $|S|$. In fact, if $|S| = k$, then*

$$N(S) = (n-k)!$$

Proof. For each $i \in S$, the value $\sigma(i) = i$ is fixed. The other values of σ are a permutation among the remaining $n-k$ positions, and there are $(n-k)!$ of these. □

As before, the principal result of this section follows immediately from the lemma and the Principle of Inclusion-Exclusion.

Theorem 6.10. *For each positive integer n, the number d_n of derangements of $[n]$ satisfies*

$$d_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)!.$$

For example,

$$\begin{aligned} d_5 &= \binom{5}{0}5! - \binom{5}{1}4! + \binom{5}{2}3! - \binom{5}{3}2! + \binom{5}{4}1! - \binom{5}{5}0! \\ &= 120 - 120 + 60 - 20 + 5 - 1 \\ &= 44. \end{aligned}$$

It has been traditional to cast the subject of derangements as a story, called the Hat Check problem. The story belongs to the period of time when men wore top hats. For a fancy ball, 100 men check their top hats with the Hat Check person before entering the ballroom floor. Later in the evening, the mischeivous hat check person decides to return hats at random. What is the probability that all 100 men receive a hat other than their own? It turns out that the answer is very close to $1/e$, as the following result shows.

Theorem 6.11. *For a positive integer n, let d_n denote the number of derangements of $[n]$. Then*

$$\lim_{n\to\infty} \frac{d_n}{n!} = \frac{1}{e}.$$

Equivalently, the fraction of all permutations of $[n]$ that are derangments approaches $1/e$ as n increases.

Proof. It is easy to see that

$$\begin{aligned} \frac{d_n}{n!} &= \frac{\sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)!}{n!} \\ &= \sum_{k=0}^{n} (-1)^k \frac{n!}{k!(n-k)!} \frac{(n-k)!}{n!} \\ &= \sum_{k=0}^{n} (-1)^k \frac{1}{k!}. \end{aligned}$$

Recall from Calculus that the Taylor series expansion of e^x is given by

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

and thus the result then follows by substituting $x = -1$. $\square$

Usually we're not as interested in d_n itself as we are in enumerating permutations with certain restrictions, as the following example illustrates.

Example 6.12. Consider the Hat Check problem, but suppose instead of wanting no man to leave with his own hat, we are interested in the number of ways to distribute the 100 hats so that precisely 40 of the men leave with their own hats.

If 40 men leave with their own hats, then there are 60 men who do not receive their own hats. There are $C(100, 60)$ ways to choose the 60 men who will not receive their own hats and d_{60} ways to distribute those hats so that no man receives his own. There's only one way to distribute the 40 hats to the men who must receive their own hats, meaning that there are

$$\begin{aligned}\binom{100}{60} d_{60} =& 42078873492228172128327462833391345210773815159514072218289944 \\ & 4678525002320680486289651537677289131789401969 20\end{aligned}$$

such ways to return the hats.

6.5 The Euler ϕ Function

After reading the two previous sections, you're probably wondering why we stated the Principle of Inclusion-Exclusion in such an abstract way, as in those examples $N(S)$ depended only on the size of S and not its contents. In this section, we produce an important example where the value of $N(S)$ *does* depend on $|S|$. Nevertheless, we are able to make a reduction to obtain a useful end result.

For a positive integer $n \geq 2$, let

$$\phi(n) = |\{m \in \mathbb{N} : 1 \leq m \leq n, \gcd(m, n) = 1\}|.$$

This function is usually called the Euler ϕ function or the Euler totient function and has many connections to number theory. We won't focus on the number-theoretic aspects here, only being able to compute $\phi(n)$ efficiently for any n.

For example, $\phi(12) = 4$ since the only numbers from $\{1, 2, \ldots, 12\}$ that are relatively prime to 12 are 1, 5, 7 and 11. As a second example, $\phi(9) = 6$ since 1, 2, 4, 5, 7 and 8 are relatively prime to 9. On the other hand, $\phi(p) = p - 1$ when p is a prime. Suppose you were asked to compute $\phi(321974)$. How would you proceed?

In chapter 3 we discussed a recursive procedure for determining the greatest common divisor of two integers, and we wrote a code for accomplishing this task. Let's assume that we have a function declared as follows:

```
int gcd(int m, int n);
```

that returns the greatest common divisor of m and n.

Then we can calculate $\phi(n)$ with this code snippet:

```
answer = 1;
for (m = 2; m < n; m++) {
    if (gcd(m,n) == 1) {
       answer++;
    }
}
return(answer);
```

A program called `phi.c` using the code snippet above answers almost immediately that $\phi(321974) = 147744$.

On the other hand, in just under two minutes the program reported that

$$\phi(319572943) = 319524480.$$

So how could we find $\phi(13691222573287670 73)$?

Clearly, the program is useless to tackle this beast! It not only iterates $n-2$ times but also invokes a recursion during each iteration. Fortunately, Inclusion-Exclusion comes to the rescue.

Theorem 6.13. *Let $n \geq 2$ be a positive integer and suppose that n has m distinct prime factors: $p_1, p_2, \ldots, p_m$. Then*

$$\phi(n) = n \prod_{i=1}^{m} \frac{p_i - 1}{p_i}. \tag{6.2}$$

Proof. We present the argument when $m = 3$. The full result is an easy extension.

Our argument requires the following elementary proposition whose proof we leave as an exercise.

Proposition 6.14. *Let $n, k \geq 2$, and let $p_1, p_2, \ldots, p_k$ be distinct primes each of which divide n evenly (without remainder). Then the number of integers from $\{1, 2, \ldots, n\}$ which are divisble by each of these k primes is*

$$\frac{n}{p_1 p_2 \ldots p_k}.$$

Then Inclusion-Exclusion yields:

$$\begin{aligned}
\phi(n) &= n - \left(\frac{n}{p_1} + \frac{n}{p_2} + \frac{n}{p_3}\right) + \left(\frac{n}{p_1 p_2} + \frac{n}{p_1 p_3} + \frac{n}{p_2 p_3}\right) - \frac{n}{p_1 p_2 p_3} \\
&= n \frac{p_1 p_2 p_3 - (p_2 p_3 + p_1 p_3 + p_1 p_2) + (p_3 + p_2 + p_1) - 1}{p_1 p_2 p_3} \\
&= n \frac{p_1 - 1}{p_1} \frac{p_2 - 1}{p_2} \frac{p_3 - 1}{p_3}.
\end{aligned}$$

$\square$

Example 6.15. Maple reports that

$$1369122257328767073 = (3)^3(11)(19)^4(31)^2(6067)^2$$

is the factorization of 1369122257328767073 into primes. It follows that

$$\phi(1369122257328767073) = 1369122257328767073\ \frac{2}{3}\,\frac{10}{11}\,\frac{18}{19}\,\frac{30}{31}\,\frac{6066}{6067}.$$

Thus Maple quickly reports that

$$\phi(1369122257328767073) = 760615484618973600.$$

Example 6.16. Alice and Bob receive the same challenge from their professor, namely to find $\phi(n)$ when

$$\begin{aligned} n =\,&31484972786199768889479107860964368171543984609017931\\ &39001922159851668531040708539722329324902813359241016\\ &93211209710523.\end{aligned}$$

However the Professor also tells Alice that $n = p_1p_2$ is the product of two large primes where

$$p_1 = 470287785858076441566723507866751092927015824834881906763507$$

and

$$p_2 = 669483106578092405936560831017556154622901950048903016651289.$$

Is this information of any special value to Alice? Does it really make her job any easier than Bob's?

6.6 Exercises

1. List all the derangements of $[4]$. (For brevity, you may write a permutation σ as a string $\sigma(1)\sigma(2)\sigma(3)\sigma(4)$.)

2. A school has 147 third graders. The third grade teachers have planned a special treat for the last day of school and brought ice cream for their students. There are three flavors: mint chip, chocolate, and strawberry. Suppose that 60 students like (at least) mint chip, 103 like chocolate, 50 like strawberry, 30 like mint chip and strawberry, 40 like mint chip and chocolate, 25 like chocolate and strawberry, and 18 like all three flavors. How many students don't like any of the flavors available?

3. The State of Georgia is distributing \$173 million in funding to Fulton, Gwinnett, DeKalb, Cobb, and Clayton counties in even millions of dollars. In how many ways can this distribution be made, assuming that each county receives at least \$1 million, Clayton county receives at most \$10 million, and Cobb county receives at most \$30 million? What if Fulton county is to receive at least \$5 million instead?

4. How many integer solutions are there to the inequality

$$y_1 + y_2 + y_3 + y_4 < 184$$

with $y_1 > 0$, $0 < y_2 \leq 10$, $0 \leq y_3 \leq 17$, and $0 \leq y_4 < 19$?

5. A careless payroll clerk is placing employees' paychecks into pre-labeled envelopes and the envelopes get sealed before he realizes he didn't match the names on the paychecks with the names on the envelopes. If there are seven employees, in how many ways could he have placed the paychecks into the envelopes so that exactly three employees receives the correct paycheck?

6. Compute $\phi(756)$.

7. Given that $1625190883965792 = (2)^5(3)^2(13)(23)^3(5973)^2$, compute

$$\phi(1625190883965792).$$

8. Prove Proposition 6.14.