CHAPTER 3

The Principle of Inclusion and Exclusion

3.1. The Main Theorem

The principle of inclusion-exclusion was used by De Moivre in 1718 to calculate the number of derangements on n elements.

Let A be a finite set and for each $i \in \{1, 2, ..., n\}$, let A_i be a subset of A. We would like to know how many elements there are in the set

$$A \setminus \cup_{i=1}^{n} A_i.$$

That is, we would like to know the number of elements remaining in A after we have removed the elements of A_i for each $i = 1, 2, ..., n$. To this end, we define for each subset I of $[n] = \{1, 2, ..., n\}$,

$$A_I = \cap_{i \in I} A_i.$$

That is, A_I consists of elements belonging to all A_i, $i \in I$. If I is the empty set $\emptyset$, we make the convention $A_\emptyset = A$. The principle of inclusion and exclusion is contained in the following theorem.

THEOREM 3.1.1. *The number of elements not belonging to any A_i, $1 \leq i \leq n$ is given by*

$$\sum_{I \subseteq [n]} (-1)^{|I|} |A_I|.$$

PROOF. The sum is equal to

$$\sum_{I \subseteq [n]} (-1)^{|I|} \sum_{a \in A_I} 1 = \sum_{a \in A} \sum_{I : I \subseteq [n], a \in A_I} (-1)^{|I|}.$$

Let S_a be the set of indices i such that $a \in A_i$. Then, the inner sum is over all subsets of S_a. If S_a is empty, this sum is 1. Otherwise, by the Binomial Theorem, it is equal to

$$\sum_{j=0}^{|S_a|} (-1)^j \binom{|S_a|}{j} = (1-1)^{|S_a|} = 0,$$

when S_a is non-empty. Hence, the sum is equal to the number of elements a for which S_a is empty. Since the number of elements not

S. M. Cioabă and M. R. Murty, *A First Course in Graph Theory and Combinatorics*

belonging to any A_i is precisely the number of elements a for which S_a is empty, this completes the proof. ∎

This simple principle is one of the most powerful in all of mathematics and has important consequences which we will present in the next sections.

3.2. Derangements Revisited

It will be recalled that in Chapter 2, we derived a recurrence relation for the number of derangements of a set with n elements. We then "guessed" a formula and proved it by induction. We now give a more credible approach to the derivation of this formula. Let A be the set $\{1, 2, ..., n\}$. The number d_n counts the number of permutations without any fixed points. For each i, $1 \leq i \leq n$, let A_i be the subset of permutations fixing i. Then, the number of derangements is the number of permutations not belonging to any of the A_i, $1 \leq i \leq n$. For each subset I of $\{1, 2, .., n\}$, the number of elements of A_I is clearly $(n - |I|)!$. By the inclusion-exclusion principle, we obtain the following result.

THEOREM 3.2.1.

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

This is precisely the formula we established by induction in the previous chapter. As noted earlier, this result has the curious consequence that if a group of 100 people each wrote their names on a card and these cards were then collected and shuffled and a card is handed back to each person, then the probability that a person would receive their own original card back is very close to $1 - 1/e$.

3.3. Counting Surjective Maps

Let us now count the number of surjective functions from an n-set to a k-set. The total number of functions from $[n]$ to $[k]$ is clearly k^n.

THEOREM 3.3.1. *The number of surjective functions* $f : [n] \to [k]$ *is*

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

PROOF. For each $1 \leq i \leq k$, let A_i be the set of functions from an n-set to a k-set that do not have i in their range. Then, A_I has cardinality $(k - |I|)^n$. By the inclusion-exclusion principle, the result is now immediate. ∎

COROLLARY 3.3.2. *If k and n are nonnegative integers, then*

$$\sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n = \begin{cases} 0 & \text{if } k < n; \\ n! & \text{if } k = n. \end{cases}$$

PROOF. If $n < k$, there are no surjective functions from an n-set to a k-set. The number of surjective maps from an n-set to another n-set is clearly $n!$. The result now follows from the previous theorem. ∎

3.4. Stirling Numbers of the First Kind

We now introduce **Stirling numbers of the first kind**, denoted $s(n,k)$. The Stirling numbers of first kind and of second kind (which will be defined in the next section) are named after James Stirling whose formula for $n!$ is contained in the previous chapter. Recall that every permutation has a unique decomposition (up to rearrangement) as a product of disjoint cycles. We define $s(n,k)$ by the rule that $(-1)^{n-k}s(n,k)$ is the number of permutations of S_n which can be written as a product of k-disjoint cycles. Clearly, $s(n,n) = 1$ since the only permutation that has n disjoint cycles in its cycle decomposition is the identity permutation. It is also clear that

$$\sum_{k=1}^{n}(-1)^{n-k}s(n,k) = \sum_{k=1}^{n}|s(n,k)| = n!.$$

We now establish a recurrence for $s(n,k)$.

THEOREM 3.4.1.

$$s(n+1,k) = -ns(n,k) + s(n,k-1).$$

PROOF. Of the permutations of S_{n+1} with k disjoint cycles, we consider those in which $(n+1)$ appears as a one cycle and those in which it does not. The number in the first group is clearly $(-1)^{n-(k-1)}s(n,k-1)$. For the number in the second group, we may view the elements as permutations of S_n with k disjoint cycles into which we have inserted $(n+1)$. For a cycle of S_n of length j, there are j places into which we can insert $(n+1)$ giving j new permutations. Now if σ is a permutation of S_n with k-cycles of lengths $j_1, ..., j_k$, we can interpolate $(n+1)$ into this in $j_1 + \cdots + j_k = n$ ways. Thus, the number of elements in the second group is $n(-1)^{n-k}s(n,k)$.

Thus,

$$(-1)^{n+1-k}s(n+1,k) = (-1)^{n-(k-1)}s(n,k-1) + n(-1)^{n-k}s(n,k).$$

This simplifies to give the stated recurrence. ∎

For $t \in \mathbb{R}$, we denote $(t)_n = t(t-1)\dots(t-n+1)$. Using the previous result, we can prove the following:

THEOREM 3.4.2.

$$(t)_n = \sum_{k=1}^{n} s(n,k)t^k.$$

PROOF. Again, we use induction on n. For $n = 1$, the result is clear. Assume that the result is established for $n \leq m$. Then,

$$(t)_{m+1} = (t)_m(t-m) = \left(\sum_{k=1}^{m} s(m,k)t^k\right) \cdot (t-m).$$

The coefficient of t^k on the right is

$$s(m,k-1) - ms(m,k)$$

which is precisely $s(m+1,k)$ by the previous theorem. This completes the proof. ∎

3.5. Stirling Numbers of the Second Kind

We denote by $S(n,k)$ the number of partitions of an n-set into k-blocks. These numbers are called the **Stirling numbers of the second kind**. We will try to relate these numbers to the discussion of the surjective functions. Observe that if we have a surjective map f from an n-set to a k-set, the "fibers", namely $f^{-1}(j) := \{i : i \in [n], f(i) = j\}$, for $1 \leq j \leq k$ form a partition of the n-set into k-blocks. Conversely, given a partition of a n-set into k-blocks, there are clearly $k!S(n,k)$ ways of defining a surjective map from the n-set to a k-set because we can view each block as the fiber of the image of such a map and there are $k!$ ways of assigning the image. Putting this together with Theorem 3.3.1 gives the following result.

THEOREM 3.5.1.

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

We again see how one can deduce Corollary 3.3.2. Indeed, if $k > n$, there are no ways of partitioning an n-set into k-blocks as each block must contain at least one element. For $k = n$, we clearly have $S(n,n) = 1$.

This formula also allows us in yet another way to deduce the generating function for the Bell numbers B_n which we derived in the previous chapter. Indeed, we clearly have

$$B_n = \sum_{k=0}^{n} S(n,k).$$

On the other hand, notice that

$$\sum_{n=0}^{\infty} \frac{S(n,k)t^n}{n!} = \frac{1}{k!}\sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n.$$

Upon interchanging the summation, we get that this equals

$$\frac{1}{k!}\sum_{j=0}^{k}(-1)^j \binom{k}{j} e^{(k-j)t}.$$

Using the binomial theorem, we can simplify the right hand side and deduce:

THEOREM 3.5.2.

$$\sum_{n=0}^{\infty} \frac{S(n,k)t^n}{n!} = \frac{1}{k!}(e^t - 1)^k.$$

Combining this fact with the formula relating B_n with the Stirling numbers of the second kind easily gives us again the generating function

$$\sum_{n=0}^{\infty} \frac{B_n t^n}{n!} = e^{e^t - 1}.$$

Even though we have an explicit formula for the $S(n,k)$'s, it will be useful to derive the following recurrence relation.

THEOREM 3.5.3.

$$S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

PROOF. In partitioning the n-set $\{1, 2, ..., n\}$ into k blocks, we have two possibilities. Either n is in a singleton block by itself or it is not. In the first case, the number of such decompositions clearly corresponds to $S(n-1, k-1)$. In the second case, we take the decomposition of an $(n-1)$-set into k-blocks, and we now have k choices into which we may place n. This gives the recursion. ∎

We may use this recursion to give another 'generating form' for the numbers $S(n,k)$ for n fixed and varying k. To this end, we recall the notation $(t)_n = t(t-1)(t-2)...(t-n+1)$.

THEOREM 3.5.4.

$$t^n = \sum_{k=1}^{n} S(n,k)(t)_k.$$

PROOF. The proof is by induction on n. For $n = 1$, the result is clear. Suppose that we have proved the formula for $n \leq m$. Then, we write

$$t^{m+1} = t^m \cdot t = \sum_{k=1}^{m} S(m,k)(t)_k((t-k)+k)$$

by the induction hypothesis. Because $(t)_k(t-k) = (t)_{k+1}$, we deduce that

$$t^{m+1} = \sum_{k=1}^{m} S(m,k)(t)_{k+1} + \sum_{k=1}^{m} kS(m,k)(t)_k.$$

By changing variables on the first sum, and noting that $S(m,m+1) = 0$, we may write the right hand side as

$$\sum_{k=1}^{m+1} \{S(m,k-1) + kS(m,k)\}(t)_k = \sum_{k=1}^{m+1} S(m+1,k)(t)_k$$

by the recursion of Theorem 3.5.3. This completes the proof. ∎

COROLLARY 3.5.5. *If A and B are the $n \times n$ matrices whose (i,j)-th entries are given by $s(i,j)$ and $S(i,j)$ respectively, then $B = A^{-1}$.*

PROOF. Let V be the vector space of polynomials of degree $\leq n$, with constant term zero. Then A and B are the transition matrices from the two bases:

(1) $t, t^2, ..., t^n$;
(2) $(t)_1, (t)_2, ..., (t)_n$.

The result now follows from linear algebra. ∎

COROLLARY 3.5.6. *The following are equivalent:*

(1) $g_n = \sum_{k=1}^{n} S(n,k) f_k$;
(2) $f_n = \sum_{k=1}^{n} s(n,k) g_k$.

PROOF. This is immediate from matrix inversion. ∎

If we define f_0 and g_0 so that $f_0 = g_0$, and

$$F(t) = \sum_{n=0}^{\infty} \frac{f_n t^n}{n!}$$

and

$$G(t) = \sum_{n=0}^{\infty} \frac{g_n t^n}{n!}$$

where g_n and f_n are related as in Corollary 3.5.6, then we can determine the relationship between these two generating functions as follows.

$$G(t) = f_0 + \sum_{n=1}^{\infty}\sum_{k=1}^{n} S(n,k) f_k \frac{t^n}{n!}.$$

Interchanging summation, we deduce

$$G(t) = f_0 + \sum_{k=1}^{\infty} f_k \frac{(e^t-1)^k}{k!} = F(e^t-1).$$

which implies the following result.

COROLLARY 3.5.7. *If f_n and g_n are related as in Corollary 3.5.6, then*

$$G(t) = F(e^t - 1).$$

This allows us to deduce the generating function for Stirling numbers of the first kind. Let $g_k = 1$ and $g_n = 0$ for $n \neq k$. Then, $f_n = s(n,k)$. By Corollary 3.5.7, we get

$$\frac{t^k}{k!} = F(e^t - 1).$$

Putting $x = e^t - 1$ gives

$$\sum_{n=0}^{\infty} \frac{s(n,k)x^k}{k!} = \frac{(\log(1+x))^k}{k!}.$$

3.6. Exercises

EXERCISE 3.6.1. There are 13 students taking math, 17 students taking physics and 18 students taking chemistry. We know there are 5 students taking both math and physics, 6 students taking physics and chemistry and 4 students taking chemistry and math. Only 2 students out of the total of 50 students are taking math, physics and chemistry. How many students are not taking any courses at all ?

EXERCISE 3.6.2. The **greatest common divisor** $\gcd(a,b)$ of two natural number a and b is the largest natural number that divides both a and b. If n is a natural number, denote by $\phi(n)$ the number of integers k with $1 \leq k \leq n$ and $\gcd(n,k) = 1$. Show that if p is a prime, then $\phi(p) = p-1$ and that if $p \neq q$ are two primes, then $\phi(pq) = (p-1)(q-1)$.

EXERCISE 3.6.3. If $n = p_1^{a_1} \dots p_r^{a_r}$ with p_i distinct primes, then show that

$$\phi(n) = n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right).$$

EXERCISE 3.6.4. How many integers less than n are not divisible by any of $2, 3$ and 5 ?

EXERCISE 3.6.5. How many 7 digit phone numbers contain at least 3 odd digits ?

EXERCISE 3.6.6. If $A_1, A_2, \ldots, A_n$ are finite sets, show that

$$\sum_{i=1}^{n} |A_i| - \sum_{i \neq j} |A_i \cap A_j| \leq |\cup_{i=1}^{n} A_i| \leq \sum_{i=1}^{n} |A_i|.$$

When does equality happen ?

EXERCISE 3.6.7. If n and r are non-negative integers with $0 \leq r \leq n$, denote by $f(n, r)$ the number of permutations of S_n with exactly r fixed points. Show that

$$\lim_{n \to \infty} \frac{f(n, r)}{n!} = \frac{1}{er!}.$$

EXERCISE 3.6.8. Show that

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} (n-j)^{n+1} = \binom{n+1}{2} n!.$$

EXERCISE 3.6.9. Let $s(n, k)$ denote the Stirling numbers of the first kind. Show that

$$x(x+1) \ldots (x+n-1) = \sum_{k=0}^{n} |s(n, k)| x^k.$$

EXERCISE 3.6.10. Using the previous identity, prove that the number of permutations with an even number of cycles (in their decomposition as a product of disjoint cycles) is equal to the number of permutations with an odd number of cycles.

EXERCISE 3.6.11. Let $S(n, k)$ denote the Stirling numbers of second kind. Show that

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

EXERCISE 3.6.12. Prove that

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} \binom{m+n-i}{k-i} = \begin{cases} \binom{m}{k} & \text{if } m \geq k \\ 0 & \text{if } m < k. \end{cases}$$

EXERCISE 3.6.13. Show that

$$|s(n, 1)| = (n-1)!.$$

Give two proofs.

EXERCISE 3.6.14. Prove that $S(n,1) = S(n,n) = 1$ and $S(n,2) = 2^{n-1} - 1$.

EXERCISE 3.6.15. Show that $S(n, n-1) = \binom{n}{2}$.

EXERCISE 3.6.16. Let $s(n)$ be the number of involutions in the symmetric group S_n. Show that

$$f(t) := \sum_{n \geq 0} \frac{s(n)t^n}{n!} = e^{t+\frac{t^2}{2}}.$$

EXERCISE 3.6.17. The Bernoulli numbers b_n are defined by the recurrence relation

$$\sum_{k=0}^{n} \binom{n+1}{k} b_k = 0$$

for $n \geq 1$ and $b_0 = 1$. Prove that

$$g(t) := \sum_{n \geq 0} \frac{b_n t^n}{n!} = \frac{t}{e^t - 1}.$$

These numbers were first studied by Jakob Bernoulli (1654-1705) in 1713.

EXERCISE 3.6.18. Show that $g(t) + \frac{t}{2}$ is an even function of t, where $g(t)$ is defined in the previous exercise.

EXERCISE 3.6.19. Show that $b_n = 0$ for each odd number $n \geq 3$.

EXERCISE 3.6.20. Let $(f_n)_{n \geq 0}$ and $(g_n)_{n \geq 0}$ be sequences, with exponential generating functions $F(X)$ and $G(X)$. Show that the following the statements

$$g_n = \sum_{k=0}^{n} \binom{n}{k} f_k$$

and

$$G(X) = e^X f(X)$$

are equivalent.