MAT 307: Combinatorics

Lecture 4: Principle of inclusion and exclusion

Instructor: Jacob Fox

1 Principle of inclusion and exclusion

Very often, we need to calculate the number of elements in the union of certain sets. Assuming that we know the sizes of these sets, and their mutual intersections, the principle of inclusion and exclusion allows us to do exactly that.

Suppose that you have two sets A, B. The size of the union is certainly at most $|A| + |B|$. This way, however, we are counting twice all elements in $A \cap B$, the intersection of the two sets. To correct for this, we subtract $|A \cap B|$ to obtain the following formula:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

In general, the formula gets more complicated because we have to take into account intersections of multiple sets. The following formula is what we call the *principle of inclusion and exclusion.*

Lemma 1. *For any collection of finite sets $A_1, A_2, \ldots, A_n$, we have*

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left| \bigcap_{i \in I} A_i \right|.$$

Writing out the formula more explicitly, we get

$$|A_1 \cup \ldots A_n| = |A_1| + \ldots + |A_n| - |A_1 \cap A_2| - \ldots - |A_{n-1} \cap A_n| + |A_1 \cap A_2 \cap A_3| + \ldots$$

In other words, we add up the sizes of the sets, subtract intersections of pairs, add intersection of triples, etc. The proof of this formula is very short and elegant, using the notion of a *characteristic function.*

Proof. Assume that $A_1, \ldots, A_n \subseteq X$. For each set A_i, define the "characteristic function" $f_i(x)$ where $f_i(x) = 1$ if $x \in A_i$ and $f_i(x) = 0$ if $x \notin A_i$. We consider the following formula:

$$F(x) = \prod_{i=1}^{n} (1 - f_i(x)).$$

Observe that this is the characteristic function of the *complement* of $\bigcup_{i=1}^{n} A_i$: it is 1 iff x is not in any of the sets A_i. Hence,

$$\sum_{x \in X} F(x) = |X \setminus \bigcup_{i=1}^{n} A_i|. \tag{1}$$

Now we write $F(x)$ differently, by expanding the product into 2^n terms:

$$F(x) = \prod_{i=1}^{n} (1 - f_i(x)) = \sum_{I \subseteq [n]} (-1)^{|I|} \prod_{i \in I} f_i(x).$$

Observe that $\prod_{i \in I} f_i(x)$ is the characteristic function of $\bigcap_{i \in I} A_i$. Therefore, we get

$$\sum_{x \in X} F(x) = \sum_{I \subseteq [n]} (-1)^{|I|} \sum_{x \in X} \prod_{i \in I} f_i(x) = \sum_{I \subseteq [n]} (-1)^{|I|} |\bigcap_{i \in I} A_i|. \tag{2}$$

By comparing (1) and (2), we see that

$$|X \setminus \bigcup_{i=1}^{n} A_i| = |X| - |\bigcup_{i=1}^{n} A_i| = \sum_{I \subseteq [n]} (-1)^{|I|} |\bigcap_{i \in I} A_i|.$$

The first term in the sum here is $|\bigcap_{i \in \emptyset} A_i| = |X|$ by convention (consider how we obtained this term in the derivation above). Therefore, the lemma follows. □

2 The number of derangements

As an application of this principle, consider the following problem. A sequence of n theatergoers want to pick up their hats on the way out. However, the deranged attendant does not know which hat belongs to whom and hands them out in a random order. What is the probability that nobody gets their own hat? More formally, we have a random permutation $\pi : [n] \to [n]$ and we are asking what is the probability that $\forall i; \pi(i) \neq i$. Such permutations are called *derangements*.

Theorem 1. *The probability that a random permutation $\pi : [n] \to [n]$ is a derangement is $\sum_{k=0}^{n} (-1)^k / k!$, which tends to $1/e = 0.3678\ldots$ as $n \to \infty$.*

Proof. Let X be the set of all $n!$ permutations, and let A_i denote the set of permutations that *fix element* i, i.e.

$$A_i = \{\pi \in X \mid \pi(i) = i\}.$$

By simple counting, there are $(n-1)!$ permutations in A_i, since by fixing i, we still have $n-1$ elements to permute. Similarly, $\bigcap_{i \in I} A_i$ consists of the permutations where all elements of I are fixed, hence the number of such permutations is $(n - |I|)!$. By inclusion-exclusion, the number of permutations with some fixed point is

$$\begin{aligned} \left| \bigcup_{i \in I} A_i \right| &= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left| \bigcap_{i \in I} A_i \right| \\ &= \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)! \\ &= \sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!}. \end{aligned}$$

Hence, the probability that a random permutation has some fixed point is $\sum_{k=1}^{n} (-1)^{k+1} / k!$. By taking the complement, the probability that there is no fixed point is $1 - \sum_{k=1}^{n} (-1)^{k+1} / k! = \sum_{k=0}^{n} (-1)^k / k!$. In the limit, this tends to the Taylor expansion of $e^{-1} = \sum_{k=0}^{\infty} (-1)^k / k!$. □

3 The number of surjections

Next, consider the following situation. There are m teachers and n children, $m \geq n$. Each teacher gives one (random) child a cookie. What is the probability that all n children get at least one cookie?

Theorem 2. *The probability that all n children get cookies is $\sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (1 - k/n)^m$.*

Proof. We formalize this problem as follows. A function $f : [m] \to [n]$ is called a *surjection* if it covers all elements of $[n]$. There are n^m functions total; we are interested in how many of these are surjections. We denote by A_i the set of functions that leave element i uncovered, i.e.

$$A_i = \{f : [m] \to [n] \mid \forall j; f(j) \neq i\}.$$

The number of such functions is $(n-1)^m$, since we have $n-1$ choices for each of $f(1), f(2), \ldots, f(m)$. Similarly,

$$|\bigcap_{i \in I} A_i| = (n - |I|)^m$$

because we have $|I|$ forbidden choices for each function value. By inclusion-exclusion, we get that the number of functions which are *not surjections* is

$$|\bigcup_{i=1}^{m} A_i| = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \binom{n}{|I|} (n - |I|)^m.$$

By taking the complement, the number of surjections is

$$n^m - |\bigcup_{i=1}^{m} A_i| = \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n - k)^m.$$

Dividing by n^m, we get the desired probability. □