

Lecture 21 — Principle of Inclusion and Exclusion

Holden Lee and Yoni Miller

5/6/11

1 Introduction and first examples

We start off with an example.

Example 1.1: At Sunnydale High School there are

- 28 students in algebra class,
- 30 students in biology class, and
- 8 students in both classes.

How many students are in either algebra or biology class?

Solution. Let A denote the set of students in algebra class and B denote the set of students in biology class. To find the number of students in either class, we first add up the students in each class:

$$|A| + |B|.$$

However, this counts the students in both classes twice. Thus we have to subtract them once:

$$-|A \cap B|.$$

This shows

$$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ |A \cup B| &= 28 + 30 - 8 = 50, \end{aligned}$$

so there are 50 students in at least one of the two classes.

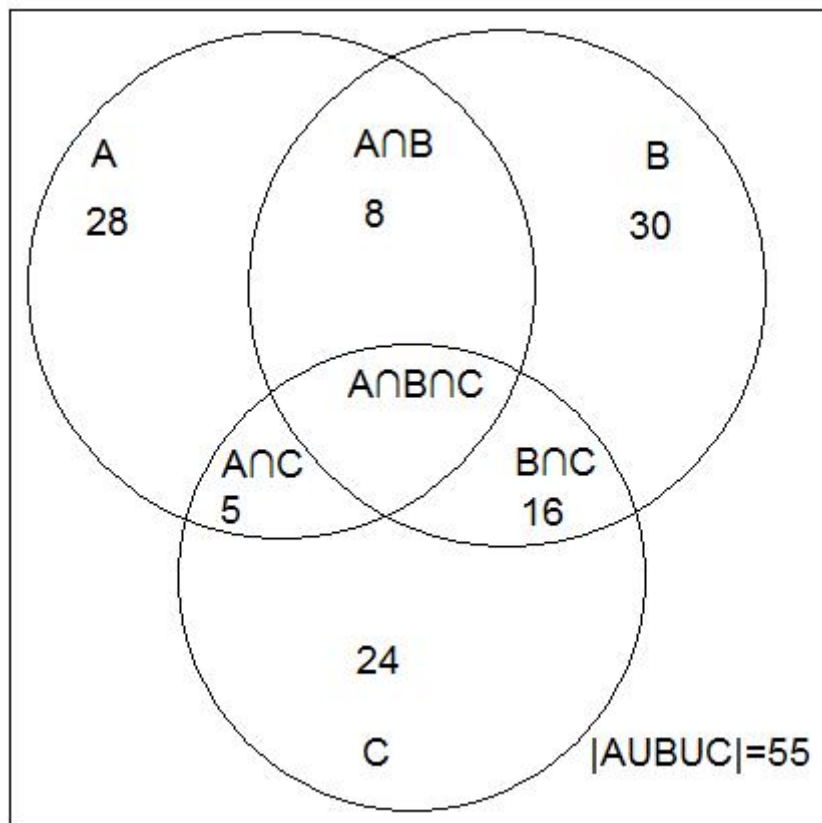
The same reasoning works with three sets.

Example 1.2: At Sunnydale High School there are

- 55 students in either algebra, biology, or chemistry class
- 28 students in algebra class
- 30 students in biology class
- 24 students in chemistry class
- 8 students in both algebra and biology

- 16 students in both biology and chemistry
- 5 students in both algebra and chemistry

How many students are in all three classes?



Solution. Let A, B, C denote the set of students in algebra, biology, and chemistry class, respectively. Then $A \cup B \cup C$ is the set of students in one of the three classes, $A \cap B$ is the set of students in both algebra and biology, and so forth. To count the number of students in all three classes, i.e. count $|A \cup B \cup C|$, we can first add all the number of students in all three classes:

$$|A| + |B| + |C|.$$

However, now we've counted the students in two classes too many times. So we subtract out the students who are in each pair of classes:

$$-|A \cap B| - |B \cap C| - |A \cap C|.$$

For students who are in two classes, we've counted them twice, then subtracted them once, so they're counted once. But for students in all three classes, we counted them 3 times, then subtracted them 3 times. Thus we need to add them again:

$$+|A \cap B \cap C|.$$

Thus

$$\begin{aligned} |A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C| \\ 55 &= 28 + 30 + 24 - 8 - 16 - 5 + |A \cap B \cap C|. \end{aligned}$$

Thus $|A \cap B \cap C| = 2$, i.e. there are 2 students in all three classes.

The same reasoning works with an arbitrary number of sets; we state the general result in the following theorem.

Theorem 1.3 (Principle of Inclusion and Exclusion (PIE)): Let A_1, \dots, A_n be sets. Define $S_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} |A_{i_1} \cap \dots \cap A_{i_k}|$, i.e. S_k is a sum over all choices of k subsets. For $k = 1$, it's simply

$$|A_1| + |A_2| + |A_3| + \dots + |A_n|$$

and for $k = n$, it represents one sum,

$$|A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n|.$$

Then

$$|A_1 \cup \dots \cup A_n| = S_1 - S_2 + S_3 + \dots + (-1)^{n-1} S_n.$$

Proof. Consider an element $x \in A_1 \cup \dots \cup A_n$. It is counted once on the left hand side, so we need to show that it is counted once on the right hand side as well. Suppose x is in the sets A_{j_1}, \dots, A_{j_m} , but not in the rest of the sets. There are $\binom{m}{k}$ sets $\{i_1, \dots, i_k\}$ such that $x \in A_{i_1} \cap \dots \cap A_{i_k}$, since $\{i_1, \dots, i_k\}$ has to be a subset of $\{j_1, \dots, j_m\}$. Thus x is counted $\binom{m}{k}$ times in the sum $\sum_{1 \leq i_1 < \dots < i_k \leq n} |A_{i_1} \cap \dots \cap A_{i_k}|$, and hence is counted

$$\binom{m}{1} - \binom{m}{2} + \dots + (-1)^{m-1} \binom{m}{m}$$

on the right hand side. By the Binomial Theorem, $1 - \binom{m}{1} + \binom{m}{2} - \dots + (-1)^m \binom{m}{m} = (1 - 1)^m$, so

$$\binom{m}{1} - \binom{m}{2} + \dots + (-1)^{m-1} \binom{m}{m} = 1,$$

showing x is counted once in the RHS. If $x \notin A_1 \cup \dots \cup A_n$, clearly x is not counted in either the LHS or the RHS. Since each element is counted the same number of times on either side, the theorem follows. \square

Example 1.4: Find the number of positive integers less than or equal to 1000 that are divisible by 7, 10, or 15.

Solution. For a positive integer k , let A_k denote the set of integers in $\{1, 2, \dots, 1000\}$ that are divisible by k . We want to find $|A_7 \cup A_{10} \cup A_{15}|$. Note that

$$A_k = \left\lfloor \frac{1000}{k} \right\rfloor$$

where $\lfloor y \rfloor$ denotes the greatest integer less than or equal to y . Indeed, the multiples of k that are less than 1000 are exactly $k, 2k, \dots, \lfloor \frac{1000}{k} \rfloor k$. Note also that $A_k \cap A_\ell = A_{\text{lcm}(k, \ell)}$ since a number is divisible by both k and ℓ if and only if it is divisible by $\text{lcm}(k, \ell)$. Using PIE, we get

$$\begin{aligned} |A_7 \cup A_{10} \cup A_{15}| &= |A_7| + |A_{10}| + |A_{15}| - |A_7 \cap A_{10}| - |A_7 \cap A_{15}| - |A_{10} \cap A_{15}| + |A_7 \cap A_{10} \cap A_{15}| \\ &= |A_7| + |A_{10}| + |A_{15}| - |A_{70}| - |A_{105}| - |A_{30}| + |A_{210}| \\ &= \left\lfloor \frac{1000}{7} \right\rfloor + \left\lfloor \frac{1000}{10} \right\rfloor + \left\lfloor \frac{1000}{15} \right\rfloor - \left\lfloor \frac{1000}{70} \right\rfloor - \left\lfloor \frac{1000}{105} \right\rfloor - \left\lfloor \frac{1000}{30} \right\rfloor + \left\lfloor \frac{1000}{210} \right\rfloor \\ &= 142 + 100 + 83 - 14 - 9 - 33 + 4 \\ &= 273. \end{aligned}$$

Suppose instead we want to find the sum of positive integers satisfying the above properties, not the number of positive integers. What do we do? We replace each term $|S|$ in the sum above by $\sum_{x \in S} x$, so, rather than contributing 1 to the value S , x contributes the value x to S . This is an example of the following more generalized version of inclusion-exclusion.

Theorem 1.5: Let $f : A_1 \cup \dots \cup A_n \rightarrow \mathbb{R}$ be a function. Then

$$\begin{aligned} \sum_{x \in A_1 \cup \dots \cup A_n} f(x) &= \sum_{k=1}^n \sum_{1 \leq i_1 < \dots < i_k \leq n} (-1)^k \sum_{x \in A_{i_1} \cap \dots \cap A_{i_k}} f(x) \\ &= \sum_{1 \leq i_1 \leq n} \sum_{x \in A_{i_1}} f(x) - \sum_{1 \leq i_1 < i_2 \leq n} \sum_{x \in A_{i_1} \cap A_{i_2}} f(x) + \dots + (-1)^{n-1} \sum_{x \in A_1 \cap \dots \cap A_n} f(x). \end{aligned}$$

Note that taking f to be identically 1, we get the original statement of PIE.

Proof. The proof is the same as above, except now an element x contained in m of the sets contributes $f(x)$ to the left hand side and $\left(\binom{m}{1} - \binom{m}{2} + \dots + (-1)^{m+1} \binom{m}{m}\right) f(x)$ to the right hand side. \square

2 Permutations and surjections

Example 2.1: As n people walk into a dinner party, they remove their hats. When they leave, the hats are randomly returned to them, so that all $n!$ matchings between people and hats are equally likely.

What is the probability that nobody gets his or her own hat back?

Solution. Labeling the people and the hats $1, 2, \dots, n$, a matching between people and hats corresponds to a permutation π on $1, 2, \dots, n$, where $\pi(i) = j$ if person i receives person j 's hat. So an equivalent formulation is the following: Given a random permutation on $\{1, 2, \dots, n\}$ (that is, a one-to-one function on $\{1, 2, \dots, n\}$), what is the probability that there is no fixed point, i.e. no i such that $\pi(i) = i$?

Let A be the set of permutations having a fixed point, and D be the set of permutations having no fixed point (these permutations are called *derangements*). We need to find $|A|$. To use PIE, we want to express A as a union of sets whose sizes are easy to calculate. Note that π having a fixed point means either 1 is a fixed point, or 2 is a fixed point, or 3 is a fixed point, or so on. Hence

$$A = A_1 \cup \dots \cup A_n.$$

where A_i is the set of permutations π such that i is a fixed point, i.e. $\pi(i) = i$. By PIE, we have

$$|A| = \sum_{k=1}^n \sum_{1 \leq i_1 < \dots < i_k \leq n} (-1)^{k-1} |A_{i_1} \cap \dots \cap A_{i_k}|. \tag{1}$$

An element π of $A_{i_1} \cap \dots \cap A_{i_k}$ must have $\pi(i_1) = i_1, \dots, \pi(i_k) = i_k$, but can permute the other $n - k$ elements in any way. Hence

$$|A_{i_1} \cap \dots \cap A_{i_k}| = (n - k)!$$

Since there are $\binom{n}{k}$ sets of size k , (1) becomes

$$|A| = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} (n-k)! = \sum_{k=1}^n \frac{(-1)^{k-1}}{n!} k!$$

$$|D| = n! - |A| = \sum_{k=0}^n \frac{(-1)^k n!}{k!}.$$

Thus the desired probability is

$$\frac{|D|}{n!} = \sum_{k=0}^n \frac{(-1)^k}{k!}.$$

Readers familiar with calculus may recognize that truncated series for $\frac{1}{e}$ appears above:

$$\frac{1}{e} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots$$

We calculated that

$$|D| = n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \dots + \frac{(-1)^n}{n!} \right).$$

Hence

$$|D| - \frac{n!}{e} = n! \sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!} \in \left(-\frac{n!}{(n+1)!}, \frac{n!}{(n+1)!} \right) \subseteq \left(-\frac{1}{2}, \frac{1}{2} \right).$$

Hence $|D|$ equals $\frac{n!}{e}$ rounded to the nearest integer, i.e. $|D| = \lfloor \frac{n!}{e} + \frac{1}{2} \rfloor$ and the probability equals

$$p = \left\lfloor \frac{1}{e} + \frac{1}{2n!} \right\rfloor.$$

In particular, as the number of hats n goes to infinity, the probability that all people get their hats back approaches $\frac{1}{e}$.

Note: What makes this proof so fascinating, is that it shows us whether it's five people who derange their hats or a thousand or billion, it will generally be the same probability.

Example 2.2: Suppose $m \geq n$. Find the number of surjective functions $f : \{1, 2, \dots, m\} \rightarrow \{1, 2, \dots, n\}$, i.e. functions f such that $f(x) = k$ has a solution for any $k \in \{1, 2, \dots, n\}$.

Solution. Let S be the set of surjective functions. As in the previous example, we find it more convenient to consider the complement of the desired set, then subtract it from the set of all functions. So let $A = S^c$ be the set of non-surjective functions. We can express A as a union

$$A = A_1 \cup \dots \cup A_n$$

where A_k is the set of functions $f : \{1, 2, \dots, m\} \rightarrow \{1, 2, \dots, n\}$ whose range does not include k . Then by PIE,

$$|A| = \sum_{k=1}^n \sum_{1 \leq i_1 < \dots < i_k \leq n} (-1)^{k+1} |A_{i_1} \cap \dots \cap A_{i_k}|. \tag{2}$$

Now $|A_{i_1} \cap \cdots \cap A_{i_k}|$ consists of all functions that miss the values i_1, \dots, i_k . This leaves $n - k$ possibilities for each of $f(1), \dots, f(m)$, so there are $(n - k)^m$ such functions. Since there are $\binom{n}{k}$ subsets of size k , (2) becomes

$$|A| = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} (n - k)^m.$$

Since there are n^m functions from $\{1, \dots, m\}$ to $\{1, \dots, n\}$, we get

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

3 Problems

- (AHSME 1983/26) The probability that event A occurs is $\frac{3}{4}$; the probability that event B occurs is $\frac{2}{3}$. Let p be the probability that both A and B occur. Find the smallest interval necessarily containing p .
- At Sunnydale High School there are
 - 44 students in either algebra, biology, or chemistry class
 - 25 students in biology class
 - 23 students in chemistry class
 - 13 students in both algebra and biology
 - 9 students in both biology and chemistry
 - 10 students in both algebra and chemistry
 - 6 students in all three classes.

How many students are in algebra class?

- Four couples are sitting in a row. Find the number of arrangements in which no person is sitting next to his or her partner. What if they are sitting in a circle?
- Find the number of five-digit combinations from the set $\{1, 2, 3, 4, 5\}$ in which:
 - Some digit appears at least three times.
 - No digit appears more than twice.
- Let $\{p_1, \dots, p_r\}$ be a set of primes. Show that the number of positive integers less than or equal to n and relatively prime to p_1, \dots, p_r is equal to

$$x \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right) + 2^r \theta$$

for some $-1 < \theta < 1$.

- Find the sum of all positive integers less than or equal to 1000 that are not divisible by 7, 10, or 15.

7. Find the number of solutions to

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 15$$

such that each x_i is an integer and $0 \leq x_i \leq 4$.

8. Specially marked boxes of ChocoBalls cereal contain one of four puzzle pieces, each with probability $\frac{1}{4}$. Once you have all four puzzle pieces, you can mail them in and receive a coupon for more ChocoBalls cereal.
- If Jim buys n specially marked boxes of ChocoBalls cereal, what is the probability that he gets all four puzzle pieces?
 - What is the expected number of boxes that Jim has to buy to get all four puzzle pieces?
9. Find the number of permutations π of $\{1, 2, \dots, n\}$ such that $\pi(i+1) \neq \pi(i) + 1$ for all $1 \leq i \leq n$.
10. Let a_n be the number of permutations π of $\{1, 2, \dots, 2n\}$ such that $\pi(i+1) - \pi(i) \neq n$ for each $1 \leq i \leq 2n - 1$. Find a_n and show that $\lim_{n \rightarrow \infty} \frac{a_n}{(2n)!} = \frac{1}{e}$.
11. (TST 2004/2) Assume n is a positive integer. Consider sequences a_0, a_1, \dots, a_n such that $a_i \in \{1, 2, \dots, n\}$ for each i and $a_n = a_0$.
- Call such a sequence *good* if for all $i = 1, 2, \dots, n$, $a_i - a_{i-1} \not\equiv i \pmod{n}$. Suppose n is odd. Find the number of good sequences.
 - Call such a sequence *great* if for all $i = 1, 2, \dots, n$, $a_i - a_{i-1} \not\equiv i, 2i \pmod{n}$. Suppose that n is an odd prime. Find the number of great sequences.

(Note: We solved this problem using generating functions and complex numbers in lecture 11. Try to do it with PIE this time.)

12. (IMO 1991/3) Let $S = \{1, 2, 3, \dots, 280\}$. Find the minimal natural number n such that in any n -element subset of S there are five numbers that are pairwise relatively prime.
13. (USAMO 1994/5) Let $|U|$, $\sigma(U)$, and $\pi(U)$ denote the number of elements, the sum, and the product, respectively, of a finite set U of positive integers. (If U is the empty set, $|U| = 0$, $\sigma(U) = 0$, and $\pi(U) = 1$.) Let S be a finite set of positive integers. As usual, let $\binom{n}{k}$ denote $\frac{n!}{k!(n-k)!}$. Prove that

$$\sum_{U \subseteq S} (-1)^{|U|} \binom{m - \sigma(U)}{|S|} = \pi(S)$$

for all integers $m \geq \sigma(S)$.

References

- [1] The Inclusion-Exclusion Principle, 2011. <http://www.math.ust.hk/~mabfchen/Math232/Inclusion-Exclusion.pdf>