Lecture 13: the pigeonhole principle, and counting in two ways

October 26, 2026

1 The Pigeonhole Principle

Recall this fact about finite sets:

Fact. If there is an injection from X to Y, then $|X| \leq |Y|$.

The contrapositive of this fact is:

Theorem 1.1. *Let X and Y be finite sets. If $|X| > |Y|$, then there is no injection from X to Y.*

This is called the **pigeonhole principle**. The name comes from thinking of a function as an assignment of pigeons to holes: if you have n pigeons and k holes with $n > k$, then no assignment can be injective, so at least one hole holds more than one pigeon.

Example 1.2. Everyone has fewer than 100,000 hairs on their head, and there are 500,000 people in upper Manhattan. So there are at least two people up there with exactly the same number of hairs on their head.

Example 1.3. If there are only 100 possible final grades and 300 students in the class, then at least two students get the same final grade.

In fact, with 300 students and 100 grades we can say more: at least *three* students share a grade. If every grade went to at most two students, the class could hold at most $2 \cdot 100 = 200$ students. This leads to the generalized pidgeonhole principle:

Theorem 1.4 (Generalized PHP)**.** *Let X and Y be finite sets and let $k \geq 1$. If $|X| > k \cdot |Y|$, then every function $f : X \to Y$ has some $y \in Y$ with $|f^{-1}[\{y\}]| > k$.*

Taking $k = 1$ recovers the ordinary pigeonhole principle.

The pigeonhole principle is useful for proving that some object must exist, without having to produce it.

Theorem 1.5. *Let $A \subseteq [100]$ with $|A| = 10$. Then there are subsets $B \neq C$ of A whose elements have the same sum.*

Proof. Let $\mathcal{S}$ be the set of all subsets of A, and define

$$f : \mathcal{S} \to \mathbb{N} \qquad \text{via} \qquad f(B) = \sum_{x \in B} x,$$

the sum of the elements of B (so $f(\varnothing) = 0$).

How large can a sum get? Every element of A is at most 100, and A has only 10 elements, so the largest a sum could possibly be is

$$100 + 99 + 98 + \cdots + 91 = 955.$$

So f really maps into $\{0, 1, \ldots, 955\}$, a set of 956 elements. Meanwhile A has 10 elements, so it has $2^{10} = 1024$ subsets.

Since $1024 > 956$, the pigeonhole principle says f is not injective: there are subsets $B \neq C$ with $f(B) = f(C)$, which is to say their elements have the same sum. □

Remark. In fact, we can argue that there are two *disjoint* subsets B and C of A that have the same sum. Why?

2 Counting in Two Ways

Counting also gives us a way to prove **combinatorial identities** — equations that come out of counting arguments.

For example, last week we said that

$$\binom{n}{k} = \binom{n}{n-k},$$

because the two sides "count the same thing": to pick out a size-k subset of $[n]$, we can either choose which k elements are in it ($\binom{n}{k}$ ways) or choose which $n-k$ elements are left out ($\binom{n}{n-k}$ ways). Both describe the same collection of subsets, so the two counts are equal.

This style of argument is called **counting in two ways**, or **double counting**: find one set, count it in two different ways, and conclude that the two answers are equal.

Proposition 2.1. *For* $1 \leq k \leq n$, $\displaystyle\binom{n}{k} \cdot k = n \cdot \binom{n-1}{k-1}$.

This is called the **chairperson identity**, for reasons the proof makes clear.

Proof. Both sides count the same set: the ways to choose a committee of k people out of n, and then name one member of that committee its chairperson.

Counting it the first way:

- choose the k people on the committee — $\binom{n}{k}$ ways
- choose which one of them is the chairperson — k ways

so by the multiplication principle there are $\binom{n}{k} \cdot k$ such committees.

Counting it the second way, starting from the chairperson instead:

- choose the chairperson, who can be any of the n people — n ways
- choose the other $k-1$ members from the $n-1$ people left — $\binom{n-1}{k-1}$ ways

so there are $n \cdot \binom{n-1}{k-1}$ of them.

Both procedures count the same committees, so the two counts are equal. □

Proposition 2.2. *For $n \geq 0$,* $2^n = \sum_{i=0}^{n} \binom{n}{i}$.

Proof. Both sides count the subsets of $[n]$.

For the left side, a subset is determined by n independent yes-or-no decisions:

1. is 1 in the subset? 2 ways
2. is 2 in the subset? 2 ways

$\vdots$

n. is n in the subset? 2 ways

By the multiplication principle there are $2 \cdot 2 \cdots 2 = 2^n$ subsets.

For the right side, partition the subsets of $[n]$ by their size: let $\mathcal{S}_i$ be the collection of subsets of size i, for $i = 0, 1, \ldots, n$. Every subset has exactly one size, so these do partition the subsets, and $|\mathcal{S}_i| = \binom{n}{i}$ by definition. By the addition principle,

$$\#\{\text{subsets of } [n]\} = \sum_{i=0}^{n} |\mathcal{S}_i| = \sum_{i=0}^{n} \binom{n}{i}.$$

The two counts are of the same set, so they are equal. □

Proposition 2.3.

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Hint: think about choosing a committee of k people from a group of n humans and 1 alien.

Counting in two ways is a bit of a game: the work is in finding a set whose two natural descriptions give you the two sides of the identity.