Lecture 4: More Combinations, and the Inclusion-Exclusion Principle

Anup Rao

April 8, 2019

We continue our discussion of binomial coefficients. We also discuss the inclusion-exclusion principle.

Combinations

So far, we have been counting the number sequences one can generate using a set. Let us now turn to counting the number of sets.

Identities involving binomial coefficients

The function $\binom{n}{k}$ satisfies a number of interesting identities. Usually, these identities are very easy to prove once you guess where they come from.

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

It is is easy to check this fact just by comparing the formulas, but there is a nice interpretation of it too. The point is that for every set S of size k, the complement S^c is a set of size $n-k$. Moreover, the complement uniquely determines the set S. So, the number of ways of picking S is the same as the number of ways to pick its complement.

Fact 2. $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

To see this, we count the number of sets $S \subseteq [n]$ of size k in two steps. In the first step, we count all the sets that contain 1. There are exactly $\binom{n-1}{k-1}$ such sets, because we are picking $k-1$ elements from $\{2, \ldots, n\}$—a set of size $n-1$. In the second step, we count all the sets that do not contain 1. There are exactly $\binom{n-1}{k}$ such sets, because we have to pick all k of the elements from $\{2, 3, \ldots, n\}$. This counts all of the sets of size k.

Fact 3. $\binom{n}{k} = \frac{n-k+1}{k} \cdot \binom{n}{k-1}$.

To prove it, compute

$$\binom{n}{k} / \binom{n}{k-1} = \frac{n! \cdot (k-1)! \cdot (n-k+1)!}{k! \cdot (n-k)! \cdot n!} = \frac{n-k+1}{k}.$$

Since $n-k+1 \geq k$ when $k \leq \frac{n+1}{2}$ and $n-k+1 \leq k$ when $k \geq \frac{n+1}{2}$, the largest binomial coefficient is the middle one(s), when k is about

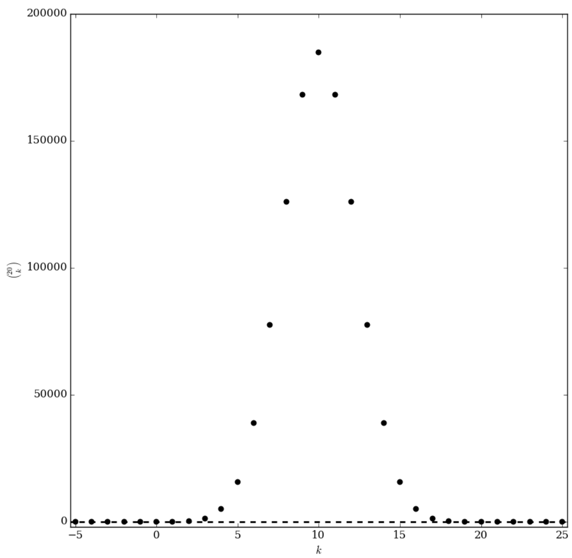

Figure 1: The largest binomial coefficients are the ones near the middle.

$\frac{n+1}{2}$. When n is even, the largest coefficient is $\binom{n}{n/2}$. When n is odd, the largest coefficients are $\binom{n}{(n+1)/2}$ and $\binom{n}{(n-1)/2}$.

For example, we calculated the number of paths in the grid that do not cross the diagonal as

$$\binom{2n}{n} - \binom{2n}{n-1} = \binom{2n}{n}(1 - \frac{n}{2n-n+1}) = \frac{1}{n+1} \cdot \binom{2n}{n}.$$

The binomial coefficients get their name from the following identity:

Fact 4. $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} \cdot x^i y^{n-i}$.

The proof follows by expanding $(x+y)^n = (x+y)(x+y)\dots(x+y)$ using the distributive law. The term $x^i y^{n-i}$ shows up whenever x is chosen from i of these product terms, and y is chosen from the remaining $n-i$. So, there are exactly $\binom{n}{i}$ for this term to show up in the product.

A special case of the above identity is:

Fact 5. $2^n = \sum_{i=0}^{n} \binom{n}{i}$.

This follows from setting $x = 1 = y$. Another way to see the same identity is to observe that 2^n counts all the subsets of $[n]$, while $\binom{n}{i}$ counts the number of subsets of size i.

Another special case of the above identity is:

Fact 6. $\sum_{odd\ i} \binom{n}{i} = \sum_{even\ i} \binom{n}{i}$.

You can prove that by setting $x = -1, y = 1$, which proves that $0^n = \sum_{i=0}^{n} \binom{n}{i}(-1)^i$.

Asymptotic Estimates

It can be hard to understand how large the binomial coefficients and factorials can get. Here are some useful estimates.

We have Stirling's approximation, which says:

Fact 7. $n! \approx \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n$.

Stirling's approximation does give some estimate about the size of binomial coefficients. For example, when n is even and $k = n/2$, we have

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!} \\ &\approx \frac{\sqrt{2\pi n} \cdot (n/e)^n}{\sqrt{2\pi k} \cdot (k/e)^k \sqrt{2\pi(n-k)} \cdot ((n-k)/e)^{n-k}} \\ &= \sqrt{\frac{1}{\pi/2 \cdot n}} \cdot 2^n.\end{aligned}$$

In other words, the sets of size $n/2$ consist of about $1/\sqrt{\pi n/2}$ fraction of all the sets.

Another useful bound:

Fact 8. $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$.

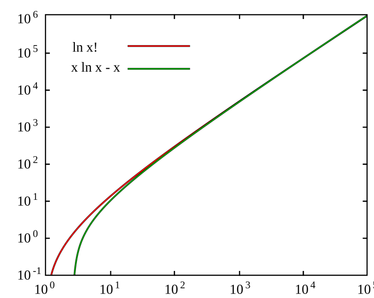

Figure 2: A plot showing the accuracy of stirling's approximation. Note that $\ln((x/e)^x) = x \ln x - x$.

Inclusion-Exclusion Principle

Often we want to count the size of the union of a collection of sets that have a complicated overlap. The inclusion exclusion principle gives a way to count them.

Given sets $A_1, \ldots, A_n$, and a subset $I \subseteq [n]$, let us write A_I to denote the intersection of the sets that correspond to elements of I:

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

Fact 9.

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

For example, if we have two sets A, B then the formula says

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

Intuitively, when we count $|A| + |B|$, we have overcounted all of the elements in the intersection, so subtracting those out gives the right value.

If we have three sets A, B, C then we have

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

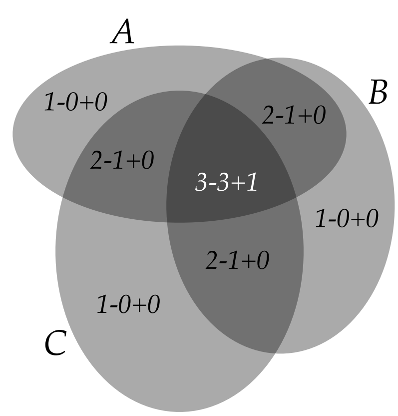

Figure 3: The union of 3 sets. The numbers show the number of times each region is counted by the inclusion-exclusion formula when it counts the size of the intersections of 1 set, 2 sets and 3 sets.

Intuitively, the sum $|A| + |B| + |C|$ overcounts all the elements that are in two of the sets, so we need to subtract the pairwise intersections. However, then we have undercounted all the elements that appear in all three sets, so we need to add those back in.

To prove that the formula is correct, let us consider how many times it counts each element $x \in \bigcup_{i \in [n]} A_i$. Suppose x occurs in k of the sets $A_1, \ldots, A_n$, and without loss of generality assume that the element is in $A_1, \ldots, A_k$.

If x is not in the union, then it is certainly never counted by the formula.

Then the number of times x is counted in the inclusion-exclusion formula is exactly

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \binom{k}{4} + \ldots + (-1)^{k+1}\binom{k}{k}.$$

We claim that this quantity is exactly 1! Indeed, we have seen that

see Fact 6

$$\sum_{\text{even } i} \binom{k}{i} = \sum_{\text{odd } i} \binom{k}{i},$$

which implies that

$$1 = \binom{k}{0} = \binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \binom{k}{4} + \ldots + (-1)^{k+1}\binom{k}{k}.$$

Thus, every element is counted exactly once in the formula, and the total is the number of elements in the union $\bigcup_{i \in [n]} A_i$.