Lecture 19 - Inclusion-Exclusion

6.042 - April 29, 2003

1 The Cardinality of a Union

This lecture addresses the problem of counting the number of elements in a union of finite sets; that is, we want a convenient way to compute

$$|A_1 \cup A_2 \cup \ldots \cup A_n|$$

where $A_1, A_2, \ldots, A_n$ are finite sets that are not necessarily disjoint. In principle, this is really easy: you just plug numbers into the *inclusion-exclusion formula*, which we'll provide shortly, and the answer pops out. In practice, this formula is rather complicated and has applications many and devious.

1.1 Simple Special Cases

In its full generality, the inclusion-exclusion formula it is messy to write down and hard to understand intuitively. So we'll first consider some simple special cases: unions of just two and three sets.

Two Sets

Suppose that 30 people are happy, 20 people are fat, and 15 people are fat and happy. How many people are fat *or* happy? Let A be the set of happy people, and B be the set of fat people. In these terms, we are asking for the size of the set $A \cup B$. The sum rule says that if A and B are disjoint sets, then:

$$|A \cup B| = |A| + |B|$$

However, if the sets A and B intersect, then this formula is wrong; each element in the intersection of A and B is counted twice, once in the $|A|$ term and once in the $|B|$ term. Thus, we can not simply add the 30 happy people to the 20 fat people, because this double-counts the 15 people who are fat *and* happy. We can fix this problem by adding a corrective term to our equation:

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

An element in the intersection of A and B is now counted once by the $|A|$ term, counted a second time by the $|B|$ term, and then subtracted once by the $|A \cap B|$ term. The net effect is that every element in $A \cup B$ is counted just once, as it should be. This revised equation is actually the inclusion-exclusion formula for two sets.

We can solve our problem by plugging numbers into this formula:

$$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ &= 30 + 20 - 15 \\ &= 35 \end{aligned}$$

Thus, 35 people are fat, happy, or both.

Three Sets

For a union of three disjoint sets, the sum rule says:

$$|A \cup B \cup C| = |A| + |B| + |C|$$

Once again, this equation does not hold if the sets intersect. For example, an element in the intersection of A, B, and C would be triple-counted on the right side. We can get a valid equation by adding a bunch of corrective terms:

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

This is the inclusion-exclusion formula for three sets. Remarkably, the expression on the right accounts for each element in the the union of A, B, and C exactly once. For example, suppose that x is an element of all three sets, A, B, and C. Then x is counted three times (by the $|A|$, $|B|$, and $|C|$ terms), subtracted off three times (by the $|A \cap B|$, $|A \cap C|$, and $|B \cap C|$ terms), and then counted once more (by the $|A \cap B \cap C|$ term). The net effect is that x is counted just once.

1.2 The General Inclusion-Exclusion Formula

The general inclusion-exclusion formula is easier to describe in words than in mathematical notation:

> The size of a union of sets is equal to the sum of the sizes of the individual sets, *minus* the sizes of all pairwise intersections, *plus* the sizes of all three-way intersections, *minus* the sizes of all four-way intersections, etc.

The name "inclusion-exclusion" comes from this process of alternately adding and subtracting. As an example, let's see how the inclusion-exclusion formula for three sets follows from the description above:

$$
\begin{aligned}
|A \cup B \cup C| &= |A| + |B| + |C| \\
&\quad - |A \cap B| - |A \cap C| - |B \cap C| \\
&\quad + |A \cap B \cap C|
\end{aligned}
$$

Sure enough, we first add the sizes of the individual sets. Then we subtract off the sizes of all pairwise intersections. Then we add the size of the only three-way intersection. Since there are only three sets, no larger intersections are possible, and so we can stop.

In mathematical notation, the general inclusion-exclusion formula is quite a mess. Two equivalent formulations are given below.

$$
\begin{aligned}
|A_1 \cup A_2 \cup \ldots \cup A_n| &= \sum_{k=1}^{n} (-1)^{k+1} \cdot \left(\sum_{S \subseteq \{1,\ldots,n\}, |S|=k} \left| \bigcap_{i \in S} A_i \right| \right) \\
&= \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|+1} \cdot \left| \bigcap_{i \in S} A_i \right|
\end{aligned}
$$

These formulas may be a bit intimidating, but they amount to the same thing as the english text given above. These formulas can be proved by induction, but we shall not do so here.

1.3 One More Special Case

One more special case of the inclusion-exclusion formula deserves individual attention. (In fact, this special case is all that we'll need for the applications we consider in the next two sections.) Suppose that every collection of k sets has an intersection of size N_k. In other words:

$$
\begin{aligned}
|A_p| &= N_1 \quad \text{for all } p \\
|A_p \cap A_q| &= N_2 \quad \text{for all distinct } p \text{ and } q \\
|A_p \cap A_q \cap A_r| &= N_3 \quad \text{for all distinct } p,\ q, \text{ and } r \\
\text{etc.}
\end{aligned}
$$

In this case, the general inclusion-exclusion formula can be simplified as follows:

$$
\begin{aligned}
|A_1 \cup A_2 \cup \ldots \cup A_n| &= \sum_{k=1}^{n} (-1)^{k+1} \cdot \left(\sum_{S \subseteq \{1,\ldots,n\}, |S|=k} \left| \bigcap_{i \in S} A_i \right| \right) \\
&= \sum_{k=1}^{n} (-1)^{k+1} \cdot \left(\sum_{S \subseteq \{1,\ldots,n\}, |S|=k} N_k \right) \\
&= \sum_{k=1}^{n} (-1)^{k+1} \cdot \binom{n}{k} \cdot N_k
\end{aligned}
$$

The first equation is the first mathematical formulation of inclusion-exclusion given in the previous section. In the second step, we use the supposition that every k-way intersection has size N_k. The last step uses the fact that the number of k-way intersections among n sets is $\binom{n}{k}$.

2 Derangements

A *derangement* is a permutation of $1, 2, 3, \ldots, n$ that leaves no number in its original position. For example:

1	2	3	4	5	6	7	8	9	$\leftarrow$ original order
2	4	5	9	8	6	1	7	3	$\leftarrow$ not a derangement; 6 doesn't move
2	4	9	8	1	3	5	7	6	$\leftarrow$ a derangement; every number moves

What is the probability that a random permutation of $1, 2, 3, \ldots, n$ is a derangement?

There are a total of $n!$ permutations. If all these permutations are equally likely, then the answer is:

$$
\Pr\{\text{derangement}\} = \frac{\#\text{ of derangements}}{n!}
$$

We've now reduced the probability question to a counting problem: how many derangements are there of the sequence $1, 2, 3, \ldots, n$?

We can use inclusion-exclusion to solve this problem. Initially, we count *non-derangements*, permutations that leave at least one number in the same position. In particular, let A_p be the set of permutations that leave p in the p-th position. Then the number of non-derangements is simply:

$$
|A_1 \cup A_2 \cup \ldots \cup A_n|
$$

By definition, each permutation in A_p keeps p in the p-th position. However, the remaining $n-1$ numbers can be assigned arbitrarily to the remaining $n-1$ positions. This can be done in $(n-1)!$ ways, and so $|A_p| = (n-1)!$ for all p. Similarly, each permutation in $A_p \cap A_q$ keeps p in the p-th position and q in the q-th position, but the other $n-2$ numbers can be assigned arbitrarily to the remaining $n-2$ positions. Therefore, $|A_p \cap A_q| = (n-2)!$ for all distinct p and q. In general, every k-way intersection has size $(n-k)!$, since k numbers are fixed and the other $(n-k)$ numbers can be assigned arbitrarily to the remaining $(n-k)$ positions.

With this observation in hand, we can apply the special case of inclusion-exclusion described in Section 1.3 with $N_k = (n-k)!$. This gives:

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} \cdot \binom{n}{k} \cdot (n-k)!$$

Now the number of derangements is the total number of permutations minus the number of non-derangements:

$$\begin{aligned}
\#\text{ of derangements} &= n! - \sum_{k=1}^{n} (-1)^{k+1} \cdot \binom{n}{k} \cdot (n-k)! \\
&= \sum_{k=0}^{n} (-1)^{k} \cdot \binom{n}{k} \cdot (n-k)!
\end{aligned}$$

Conveniently, the $n!$ moves in and serves as the 0-th term of the summation.

Strictly speaking, we're done. We now know exactly how many derangements there are, which is what we wanted. However, a little more work gives an approximate answer that is actually more enlightening:

$$\begin{aligned}
\sum_{k=0}^{n} (-1)^{k} \cdot \binom{n}{k} \cdot (n-k)! &= \sum_{k=0}^{n} (-1)^{k} \cdot \frac{n!}{k!\,(n-k)!} \cdot (n-k)! \\
&= n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!}
\end{aligned}$$

In the first step, we expand the binomial $\binom{n}{k}$ into factorials. In the second step, we cancel the $(n-k)!$ terms and pull the $n!$ outside the summation.

As n grows large, the remaining summation rapidly approaches a constant. Recall the Taylor expansion for e^x:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots$$

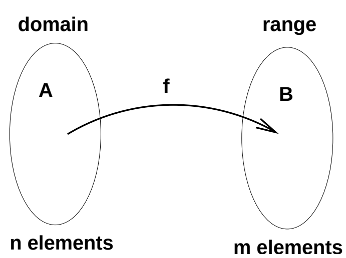
Setting $x = -1$ gives:

$$\begin{aligned} e^{-1} &= 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \ldots \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \end{aligned}$$

The sum in the derangements formula consists of the $n+1$ leading terms of this formula for e^{-1}. Therefore, the number of derangments is very close to $n!/e$. In probability terms, this means that a random permutation is a derangement with probability about $1/e$.

3 Surjective Functions

How many different functions $f : A \to B$ are there, where $|A| = n$ and $|B| = m$?



Since each of the n elements in the domain can be mapped to any one of m elements in the range, the number of functions $f : A \to B$ is simply m^n.

How many of these functions are surjective? This is a harder question. For one thing, we must remember what "surjective" means: a function $f : A \to B$ is surjective if every element of the range is mapped to by some element of the domain. Of course, if the domain is smaller than the range, than there are no surjective functions at all. So let's assume $n \geq m$.

Initially, we compute the number of *non-surjective* functions. For each element b in the range B, let F_b be the set of all functions $f : A \to B$ that do not map to b. (That is, the image of every function in F_b is contained in $B - \{b\}$.) Since every non-surjective function fails to map to some element $b \in B$, the set of all non-surjective functions is:

$$\bigcup_{b \in B} F_b$$

Once again, we can compute the size of this union using the special case of inclusion-exclusion described in Section 1.3. Consider a k-way intersection of sets:

$$F_{b_1} \cap F_{b_2} \cap \ldots \cap F_{b_k}$$

A function f in this intersection must map each of the n domain elements to one of $m - k$ elements in the range. In particular, for every $a \in A$, we must have $f(a) \in B - \{b_1, b_2, \ldots, b_k\}$. The number of functions with this property is $(m-k)^n$. Therefore, the size of each k-way intersection is $(m-k)^n$.

Plugging this result into the inclusion-exclusion formula gives:

$$\left| \bigcup_{b \in B} F_b \right| = \sum_{k=1}^{m} (-1)^{k+1} \cdot \binom{m}{k} \cdot (m-k)^n$$

This is the number of non-surjective functions; subtracting this from the total number of functions gives the number of surjective functions:

$$\begin{aligned} \text{\# of surjective functions} &= m^n - \sum_{k=1}^{m} (-1)^{k+1} \cdot \binom{m}{k} \cdot (m-k)^n \\ &= \sum_{k=0}^{m} (-1)^{k} \cdot \binom{m}{k} \cdot (m-k)^n \end{aligned}$$

Once again, the extra m^n term conveniently becomes the 0-th term of the summation. We're done!

This formula has an interesting special case. If the domain and range are the same size ($n = m$), then every surjection is a bijection and vice-versa. The number of bijections is $n!$, since the first domain element can be mapped n ways, the second can be mapped $n - 1$ ways, the third in $n - 2$ ways, and so forth. Therefore, the number of surjections is also $n!$. Equating this to the expression for the number of surjections derived above gives a puzzling identity:

$$n! = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \cdot (n-k)^n$$