CS174 Spr 99 Lecture 10 Summary John Canny

More about Inclusion/Exclusion

Recall that the inclusion/exclusion formula gives probability of an outcome having a certain number of properties E_i. The formula gives the probability for a disjunction of the properties in terms of conjunctions which are usually easier to compute. The formula is:

$$\Pr[E_1 \cup E_2 \cup \cdots \cup E_n] = \sum_{i=1}^{n} \Pr[E_i] - \sum_{1\leq i<j\leq n} \Pr[E_i \cap E_j] + \sum_{1\leq i<j<k\leq n} \Pr[E_i \cap E_j \cap E_k] - \cdots$$

Example 1

Lets consider the integers from 1 to 10,000. The properties are defined as follows:

E_1 = property that an integer is divisible by 4
E_2 = property that an integer is divisible by 5
E_3 = property that an integer is divisible by 6

The probabilities of individual divisability are easily computed:

$\Pr[E_1]$ = probability the number is divisible by 4 = $1/4$
$\Pr[E_2]$ = probability the number is divisible by 5 = $1/5$
$\Pr[E_3]$ = probability the number is divisible by 6 = $1/6$

The probabilities for double divisibility are equivalent to divisability by the lcm of the two factors:

$\Pr[E_1 \cap E_2]$ = probability the number is divisible by 4 and 5 = $1/20$
$\Pr[E_2 \cap E_3]$ = probability the number is divisible by 5 and 6 = $1/30$
$\Pr[E_1 \cap E_3]$ = probability the number is divisible by 4 and 6 = $1/12$

Finally, the probability of divisibility by all 3 factors is:

$\Pr[E_1 \cap E_2 \cap E_3]$ = probability the number is divisible by 4, 5 and 6 = $1/60$

Substituting in the inclusion/exclusion formula:

$$\begin{aligned}
\Pr[E_1 \cup E_2 \cup E_3] &= \textstyle\sum_{i=1}^{3} \Pr[E_i] - \sum_{1\leq i<j\leq 3} \Pr[E_i \cap E_j] + \Pr[E_1 \cap E_2 \cap E_3] \\
\Pr[E_1 \cup E_2 \cup E_3] &= (1/4 + 1/5 + 1/6) - (1/20 + 1/30 + 1/12) + 1/60 \\
&= ((15 + 12 + 10) - (3 + 2 + 5) + 1)/60 \\
&= 28/60 \;=\; 7/15
\end{aligned}$$

Example 2

Another important application of inclusion/exclusion is to permutations with constraints. Consider the following dinner guest problem: On the first night, n guests are seated around a circular table in some order. Number them clockwise from 1 to n. On the second night, the same guests need to be seated at the same table, but such that no guest has the same neighbor on his or her right side (hence also on the left side). That is, the new permutation has no consecutive elements $i(i+1)$. Since the arrangement is circular, we include the first and last elements as a pair. That is, the permutation should not be $((i+1),\ldots,i)$. Also the pair $n1$ is forbidden as n and 1 were neighbors on the first night.

To handle this problem, we define properties E_i where E_i is true if $i(i+1)$ occurs in the new permutation. Since we have a circular arrangement, E_n is the property that $n1$ occurs. The first step is to compute $\Pr[E_i]$. Now a permutation with property E_i can be thought of as a permutation of $1,2,\ldots,(i(i+1)),\ldots,n$, where we treat $(i(i+1))$ as a single item. That is, we permute $n-1$ objects. There are $(n-1)!$ ways to do this, so we have:

$$\Pr[E_i] = (n-1)!/n! = 1/n$$

Now consider $\Pr[E_i \cap E_j]$. There are two cases depending on whether $i+1=j$ or $i+1<j$. In the first case $j=i+1$, we are permuting the following objects: $1,2,\ldots,(i(i+1)(i+2)),\ldots,n$ where $i(i+1)(i+2)$ is a single item. There are $n-2$ objects in this case, so the number of permutations is $(n-2)!$. In the second case $j>i+1$ we are permuting the objects: $1,2,\ldots,(i(i+1)),\ldots,(j(j+1)),\ldots,n$, where $i(i+1)$ and $j(j+1)$ are single objects. But there are still $n-2$ of objects total, so the number of permutations is still $(n-2)!$. So in either case:

$$\Pr[E_i \cap E_j] = (n-2)!/n! = 1/(n(n-1))$$

In the general case where we consider k properties, no matter how the properties are ordered and what size "meta-objects" are created by combining integers, the total number of objects is reduced by the number of properties, which is k. So there will still be $n-k$ objects to be permuted. The number of permutations is $(n-k)!$, and so:

$$\Pr[E_{i_1} \cup \cdots \cup E_{i_k}] = (n-k)!/n!$$

In this case, we are interested in the probability that *none* of the E_i's occur. That probability is $1-\Pr[E_1 \cap \cdots \cap E_n]$:

$$\begin{aligned}
\Pr[noE_i] &= 1 - \textstyle\sum_{i=1}^{n} \Pr[E_i] + \cdots + (-1)^k \sum_{1\le i_1<\cdots<i_k\le n} \Pr[E_{i_1} \cap \cdots \cap E_{i_k}] \cdots \\
&= 1 - \tbinom{n}{1} 1/n + \tbinom{n}{2}(n-2)!/n! - \cdots + (-1)^k \tbinom{n}{k}(n-k)!/n! \cdots \\
&= \textstyle\sum_{k=0}^{n} \tbinom{n}{k}(n-k)!/n!(-1)^k \\
&= \textstyle\sum_{k=0}^{n} \frac{n!(n-k)!}{k!(n-k)!n!}(-1)^k \\
&= \textstyle\sum_{k=0}^{n} \frac{1}{k!}(-1)^k \\
&\approx e^{-1}
\end{aligned}$$

Example 3

Another standard use of inclusion/exclusion is for problems about strings. Suppose we choose a word w with k letters from an alphabet with n symbols at random. What is the probability that w has 2 consecutive letters the same?

For strings with k letters from an alphabet of size n, the number of possible strings is n^k. We define the property E_i to be true iff $w_i = w_{i+1}$ where w_i is the i^{th} character of the word. This is not a circular situation, so there are only $k-1$ possible E_i's.

The number of words that have $w_i = w_{i+1}$ is just n^{k-1}, because we can choose the letters in order, and every letter has n choices, except for w_{i+1} which has only one possible value. So

$$\Pr[E_i] = n^{k-1}/n^k = 1/n$$

For the conjunction of E_i and E_j, there are two cases like before. If $i+1=j$, then we have only one choice for w_{i+1} and w_{i+2}, and both of these must equal w_i. The total number of choices for all letters is n^{k-2}. In the second case $i+1<j$, there are still two letters that have only one choice, namely w_{i+1} and w_{j+1}. So the total number of choices for the word is still n^{k-2}. So no matter what case we have:

$$\Pr[E_i \cap E_j] = n^{k-2}/n^k = 1/n^2$$

In the general case, when there are m properties, each one constrains the value of one of the letters, and there will be n^{k-m} choices. So

$$\Pr[E_{i_1} \cap \cdots \cap E_{i_m}] = n^{k-m}/n^k = 1/n^m$$

Going back to the inclusion/exclusion formula:

$$\begin{aligned}
\Pr[E_1 \cup \cdots \cup E_{k-1}] &= \textstyle\sum_{i=1}^{k-1} \Pr[E_i] - \cdots - +(-1)^m \sum_{1\le i_1<\cdots<i_k\le k-1} \Pr[E_{i_1} \cap \cdots \cap E_{i_k}] \cdots \\
&= \textstyle\binom{k-1}{1} 1/n - \cdots - (-1)^m \binom{k-1}{m} 1/n^m + \cdots \\
&= 1 - \textstyle\sum_{m=0}^{k-1} \binom{k-1}{m} (-1/n)^m \\
&= 1 - (1-1/n)^{k-1}
\end{aligned}$$

General Inclusion/Exclusion Formula

The general inclusion/exclusion formula gives the probability that an outcome has *exactly* k of the properties $E_1, \ldots, E_n$. That is, let T_k be the property that an outcome has exactly k properties, and define S_m as:

$$S_m = \sum_{1\le i_1<\cdots<i_m\le n} \Pr[E_{i_1} \cap \cdots \cap E_{i_m}]$$

then the general inclusion/exclusion formula is:

$$\Pr[T_k] = S_k - \binom{k+1}{1} S_{k+1} + \binom{k+2}{2} S_{k+2} \cdots + (-1)^{n-k} \binom{n}{n-k} S_n$$

or

$$\Pr[T_k] = \sum_{m=0}^{n-k} (-1)^m \binom{k+m}{m} S_{k+m}$$

Proof

To do the proof we consider all the possible outcomes of the experiment. Let o be an outcome, and let $p(o)$ be the (finite) probability that o occurs. Then a probability like $\Pr[E_1 \cap E_2]$ is the sum of probabilities for all outcomes that have properties E_1 and E_2. The probability $\Pr[T_k]$ should be the sum of probabilities for all outcomes that have exactly k properties. To prove that it is, we will consider an outcome o and do a case analysis based on the number of properties m that o has.

Case 1, o has less than k properties. In this case, o will not be counted in $\Pr[T_k]$, because it will not be counted in any of the sums $S_k, S_{k+1}, \ldots$. All of these sums count only outcomes with k or more properties.

Case 2, o has exactly k properties. In that case, o will be counted in S_k, but none of the other sums $S_{k+1}, S_{k+2}, \ldots, S_n$. Since S_k appears in T_k with coefficient 1, the outcome o is counted exactly once in $\Pr[T_k]$.

Case 3, o has some number $k+m$ of properties with $m > 0$. Then o will be counted in the sums $S_k, S_{k+1}, \ldots, S_{k+m}$. Notice that o will appear multiple times in most of these sums. For example, o will be counted $\binom{k+m}{k}$ times in S_k. To see this, label the properties that o has $E_{j_1}, \ldots, E_{j_{k+m}}$. The outcome o will contribute to every $\Pr[E_{i_1} \cap \cdots \cap E_{i_k}]$ in S_k such that $\{i_1, \ldots, i_k\}$ is a subset of $\{j_1, \ldots, j_{k+m}\}$. The number of k-subsets of a $(k+m)$-set is $\binom{k+m}{k}$.

More generally, the outcome o will contribute to other S_{k+p} in each $\Pr[E_{i_1} \cap \cdots \cap E_{i_{k+p}}]$ where $\{i_1, \ldots, i_{k+p}\}$ is a subset of $\{j_1, \ldots, j_{k+m}\}$. The number of such subsets is $\binom{k+m}{k+p}$. So the total contribution of o to $\Pr[T_k]$ is:

$$\underbrace{\binom{k+m}{k}}_{\text{for } S_k} - \underbrace{\binom{k+m}{k+1}\binom{k+1}{1}}_{\text{for } S_{k+1}} + \underbrace{\binom{k+m}{k+2}\binom{k+2}{2}}_{\text{for } S_{k+2}} + \cdots$$

Or expressed as a sum, the series becomes:

$$\sum_{p=0}^{m} (-1)^p \binom{k+m}{k+p} \binom{k+p}{p}$$

This sum looks complicated, but a sum like this which involves two binomial coefficients is often the result of a repeated binomial expansion (or a trinomial expansion, which we havent discussed). That is the case here, which we show with the next lemma:

Lemma

The coefficient of x^k in $(1 + (x-1))^{k+m}$ is

$$\sum_{p=0}^{m} (-1)^p \binom{k+m}{k+p} \binom{k+p}{p}$$

Proof

First break the expansion of $(1+(x-1))^{k+m}$ into parts of degree at least k and less than k. We can ignore the latter terms for getting the coefficient of x^k. That is:

$$(1+(x-1))^{k+m} = \sum_{i=0}^{k+m} \binom{k+m}{i} (x-1)^i = \sum_{q=0}^{k-1} \binom{k+m}{q} (x-1)^q + \sum_{p=0}^{m} \binom{k+m}{k+p} (x-1)^{k+p}$$

And we immediately ignore the first sum over q because the coefficient of x^k in $(x-1)^q$ must be zero for $q<k$. Now the coefficient of x^k in $(x-1)^{k+p}$ is $\binom{k+p}{k}(-1)^p$. And so the coefficient of x^k in $(1+(x-1))^{k+m}$ is

$$\sum_{p=0}^{m} \binom{k+m}{k+p} \binom{k+p}{k} (-1)^p$$

which is what we wanted to show.

Corollary

Since $(1+(x-1))^{k+m} = x^{k+m}$, it follows that the coefficient of x^k is either 1 if $m=0$ or 0 if $m>0$. That is the value of:

$$\sum_{p=0}^{m} (-1)^p \binom{k+m}{k+p} \binom{k+p}{p}$$

is either 1 if $m=0$ or 0 if $m>0$. Thus an outcome with exactly k properties (and therefore $m=0$) contributes 1 to $\Pr[T_k]$, while it contributes zero to $\Pr[T_k]$ if it has more than k properties ($m>0$).