Discussion 8B Recap

Tyler Zhu

April 4, 2020

1 Principle of Inclusion-Exclusion

Theorem 1 (Inclusion-Exclusion Principle for Two Events)**.** For any two events A, B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$.

Everyone knows intuitively why this is true; draw a Venn diagram, and you'll see that the first two terms count the intersection twice, so we need to subtract it out once. This doesn't count as a proof, but it's enough for our purposes.

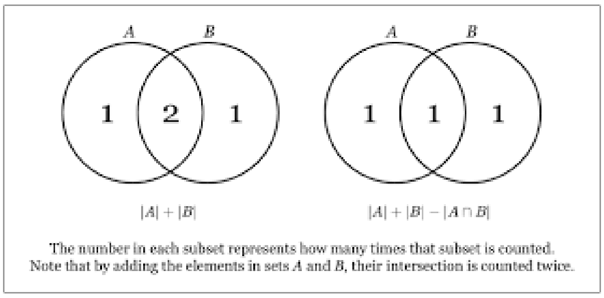

Figure 1: Inclusion-Exclusion Principle illustrated for two events. Source: Brilliant.

Similarly, we can do the same thing with three events. We just have to be careful with which events we're over and under subtracting. In this case, we have

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

See the figure for an explanation.

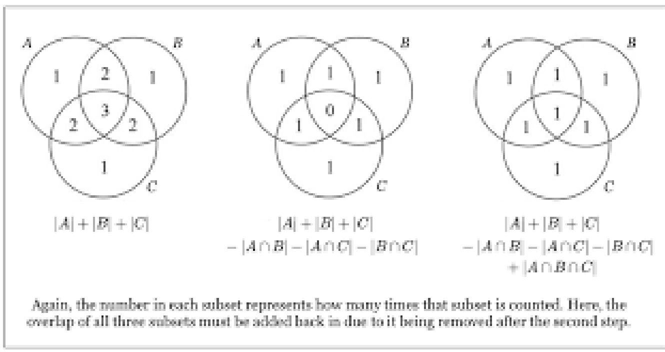

Figure 2: Inclusion-Exclusion Principle illustrated for three events. Source: Brilliant.

In general, this pattern of alternating between adding and subtracting and so on holds up. Here's the full-fledged form of the Inclusion-Exclusion Principle (but just remember the general pattern and you'll be fine).

Theorem 2 (Inclusion-Exclusion Principle). For sets $A_1, A_2, \ldots, A_n$,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}(-1)^{k-1}\left(\sum_{1\le i_1<\cdots<i_k\le n} \mathbb{P}(A_{i_1}\cap\cdots\cap A_{i_k})\right).$$

Unfortunately, as n gets larger, this formula becomes increasingly computationally intensive to use as we're computing $O(2^n)$ intersections to find the union. Anyone who's taken an algorithms class knows that it's infeasible to have an algorithm that runs in $O(2^n)$ time in terms of an input with size n. We're in a very similar spot here,

1.1 Bonferonni's Inequalities (Out of Scope)

It's interesting to ponder the above dilemma, which the notes suggest how we might resolve this. Instead of computing all of the intersections of size $j = 1, \ldots, n$, we can instead compute them up to some size k and use those values to give us an upper/lower bound on the actual union. To be precise, let

$$S_1 := \sum_{i=1}^{n} \mathbb{P}(A_i), \quad S_2 := \sum_{1\le i<j\le n} \mathbb{P}(A_i \cap A_j),$$

and in general,

$$S_k := \sum_{1\le i_1<\ldots.<i_k\le n} \mathbb{P}(A_{i_1}\cap\cdots\cap A_{i_k})$$

be the sum of the probabilities of all intersections of k events.

Then, for odd k,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{j=1}^{k}(-1)^{j-1}S_j,$$

and for even k, the reverse inequality holds. In other words, as we compute more S_j, we bounce around the true value of the union, which makes sense, since in the Inclusion-Exclusion formula we're adding and subtracting alternately. These are known as the **Bonferroni Inequalities**. The proof is not hard; it comes down to proving our intuition (that these alternating sums actually do jump around the true union value) is true. You might notice that taking $k = 1$ here gives us the union bound, while taking $k = n$ gives back the Inclusion-Exclusion formula (albeit a bound). Thus, this is a generalization of both of these facts. Beautiful piece of mathematics!

2 Union Bound

The union bound states that the probability of the union of events is upper bounded by the sum of their probabilities.

Theorem 3 (Union Bound). Given events $A_1, A_2, \ldots, A_n$,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} \mathbb{P}(A_i).$$

The reason is that worst case, all events are mutually exclusive, so equality actually holds. Then if some events have intersections, that's fine as that can only decrease the left hand side, never increase it.

The bound throws away a lot of information about potential intersections between events, but for a large majority of applications it won't matter, and this is good enough.