Probability and Statistics:
A Primer for Beginners and Pre-Beginners

The Journey Begins: Probability Theory

Part Four: Bonferroni's Inequality (and some other stuff)

Remember this?

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Well, there's one more interesting thing we can do with this. Remember that $P(A \cup B) \leq 1$. The probability of an event can't be greater than 1, after all. So let's take this pretty self-evident bit of knowledge and do some tinkering:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq 1$$
$$\Rightarrow P(A) + P(B) - 1 \leq P(A \cap B)$$
$$\Rightarrow \boldsymbol{P(A \cap B) \geq P(A) + P(B) - 1}$$

Bonferroni's Inequality

$$P(A \cap B) \geq P(A) + P(B) - 1$$

Simplified: The probability that two events will both happen is at least as high as the probability of the first event plus the probability of the second event minus 1.

Let's play a game.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Behind six of these curtains is a *brand-new car!* Behind five of these curtains is a deadly snake. You get to choose one. What's the probability of winning a brand-new car with a snake inside?

…well, we can't really say from the given information, but Bonferroni's inequality can give us a *lower bound* for that probability.

Let's play a game.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

$$P(A \cap B) \geq P(A) + P(B) - 1$$

$$\Rightarrow P(\text{🚗} \cap \text{🐍}) \geq P(\text{🚗}) + P(\text{🐍}) - 1$$

$$\Rightarrow P(\text{🚗} \cap \text{🐍}) \geq \frac{6}{9} + \frac{5}{9} - 1 = \frac{2}{9}$$

The probability you'll get a car with a snake inside is at least 2/9. Makes sense: with 6 cars, 5 snakes, and only 9 curtains, *at least* two curtains must be hiding both. Drive safe!

Two more things to round out the lesson

If some events $C_1, C_2,\ldots$ form a partition (remember, that means that they don't have any outcomes in common, but when combined, they have every outcome in the whole sample space Ω), then:

$$P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$$

Visual example

We're not going to prove this (you're welcome), but a diagram illustrates this pretty well. Let Ω be this rectangle, partitioned by C_1, C_2, C_3, C_4. Imagine that the probability of an event is the fraction of the rectangle that it occupies. That oval in the middle is A. Notice how A is comprised of its intersections with the C events?
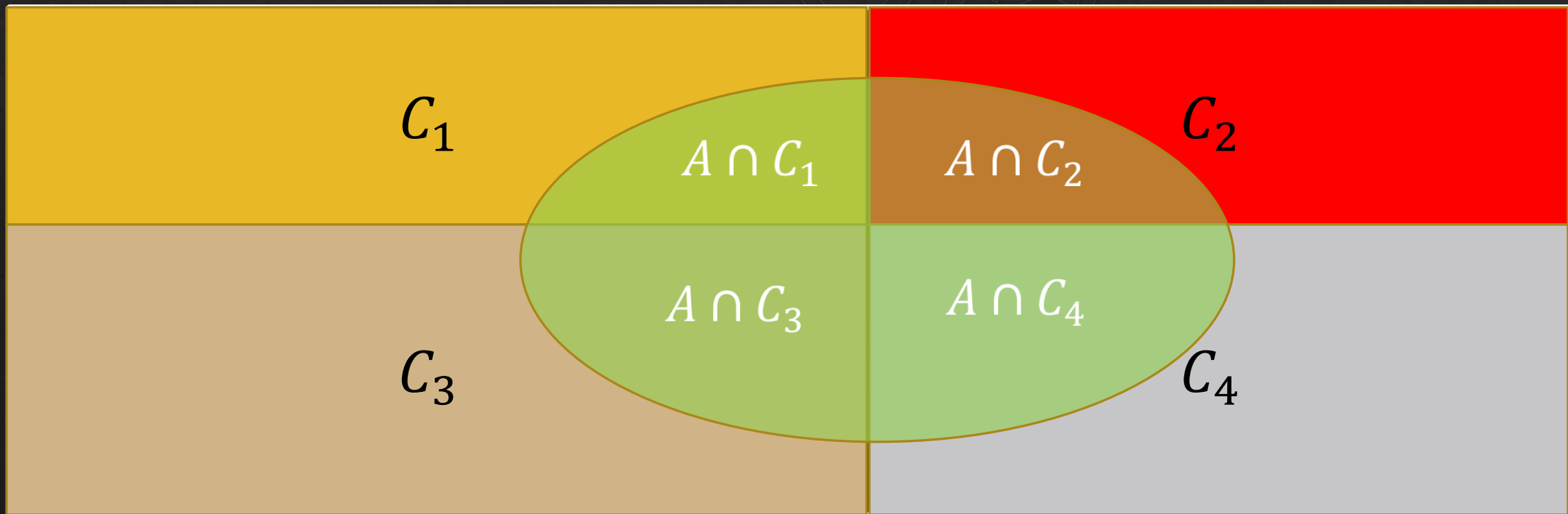
Last one (Boole's Inequality)

For *any* set of events $A_1, A_2,...$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$$

So the probability of the union of these events is always less than or equal to the sum of the individual events' probabilities. It's a little more abstract than the last one, but another diagram should make it clear.

Boole's Inequality illustrated

Again let Ω be this rectangle, with events A_1, A_2, A_3. Imagine one more time that the probability of an event is the fraction of the rectangle that it occupies.

If the three sets are totally disjoint, the probability of their union is equal to the sum of their probabilities:
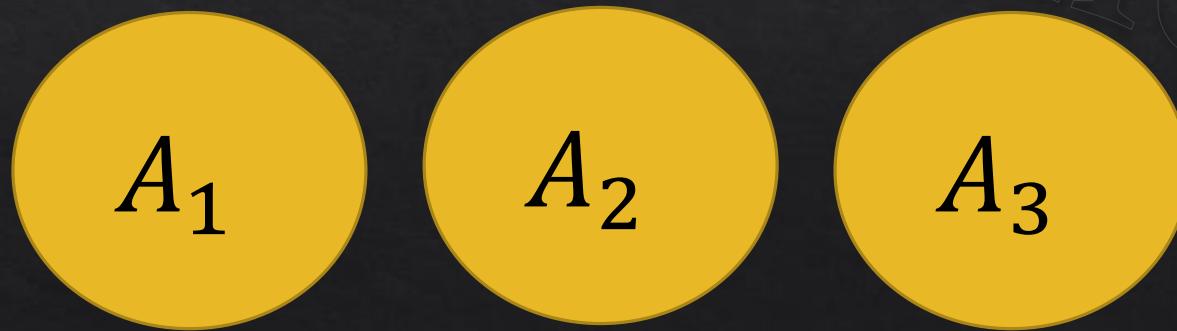
Boole's Inequality illustrated

But now, they're overlapping, so together, they take up less of the rectangle, and thus, in our analogy, the probability of their union is less than the sum of their individual probabilities. The probability of their union was at its maximum when they weren't overlapping at all.

A_1 A_2 A_3