Notes: Bayes' Rule

CS 3130/ECE 3530: Probability and Statistics for Engineers

September 7, 2017

Bayes' Rule: So far we've talked about conditional probabilities $P(A|B) = P(A \cap B)/P(B)$. Bayes' Rule allows us to switch the order of A and B, that is, to compute $P(B|A)$. Here's how it works. First, use the multiplication rule two different times:

$$P(A \cap B) = P(A|B)P(B)$$

and

$$P(A \cap B) = P(B \cap A) = P(B|A)P(A)$$

This tells us that the two right hand sides above are equal:

$$P(A|B)P(B) = P(B|A)P(A)$$

Solving for $P(B|A)$, we get **Bayes' Rule**:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

In-Class Problem: In the urn problem from last time, if you picked a black ball, what is the probability that you had picked the first urn (the 4 black, 3 white urn)?

In-Class Problem: In the power supply problem from last time, if the auxillary power fails, what is the probability that the main power also failed?

Bayes' Rule is crucial in Machine Learning, where a main problem is to learn a model that describes something in the real world. There are typically many possible models, and the question becomes: What is the best model to describe the data that I have?

We can write this using Bayes' Rule as:

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}$$

Here M is the event "my model describes the real world", and D is the event "this is the data I've collected." So, in English the left-hand side, $P(M|D)$, is "the probability that my model describes the real world given that this is the data I've collected."