An Introduction to Bayes' Rule

David Wong

28/01/2016

Introduction

This short guide will introduce the following concepts:

- Venn diagrams
- Conditional Probability
- Standard probability notation
- Bayes' Rule

Probability Basics

A probability ***event*** is the set of outcomes of a test or experiment. For instance, if our experiment is to toss a fair coin, a possible event could be 'coin lands on a head'. For the sake of brevity, let's call this event A. The probability of a coin landing on a head is written formally as:

$$P(A) = 0.5 \tag{1}$$

Venn Diagram

Consider a case in which we survey 150 people about whether they like jam or peanut butter (or both). We find that 70 like peanut butter, 50 like jam, and 30 like both (no-one dislikes both!). We can represent this as the following Venn diagram:

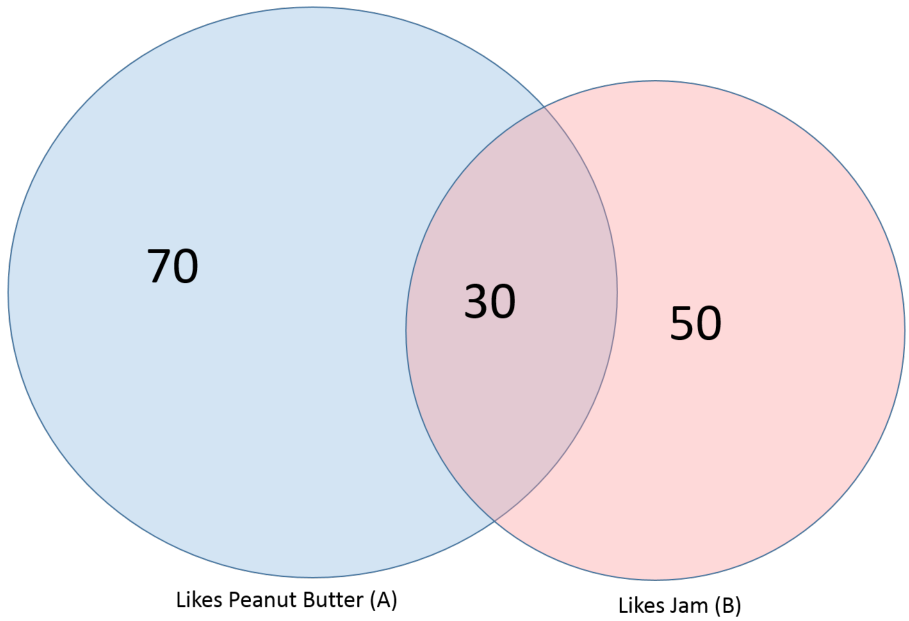

Figure 1: Venn diagram showing the number who like peanut butter (Blue), jam (red), both (overlap in purple)

For brevity again, let us say that *event A is 'likes peanut butter'* and *event B is 'likes jam'*. We can then say the probability of *likes peanut butter* is:

$$P(A) = \frac{70 + 30}{150} = 0.66 \tag{2}$$

This represents the blue circle (the number of people who like peanut butter, divided by the total population (all of the surveyed people)). Similarly the probability of *likes jam* is:

$$P(B) = \frac{50 + 30}{150} = 0.53 \tag{3}$$

Which represents the red circle. The probability that someone *likes peanut butter AND jam* is:

$$P(A \cap B) = \frac{30}{150} = 0.20 \tag{4}$$

This is the overlapping segment in the middle of the Venn diagram. In this case $\cap$ represents AND. Also note that the order of the events does not matter (*likes peanut butter AND jam* is the same as *likes jam AND peanut butter*).

Conditional Probability

A conditional probability occurs when we want to find the probability of an event if some other event has already occurred. For instance, what is the probability that someone likes jam ***given that*** they like peanut butter? The notation for given that is represented as a vertical bar, |, so we are attempting to find P(A|B).

The Venn diagram allows us to see that only 100 of the original 150 like peanut butter (blue circle), and of those, 30 like jam (overlapping segment). So:

$$P(A|B) = \frac{30}{100} = 0.30 \tag{5}$$

Noting from before that the red circle is P(B) and the overlap is P(AnB), we can rewrite the conditional probability as:

$$P(A|B) = \frac{Overlap}{Blue\,Circle} = \frac{P(A \cap B)}{P(B)} \tag{6}$$

Similarly, we can write the probability of people who like jam *given that* they like peanut butter as:

$$P(B|A) = \frac{Overlap}{Red\,Circle} = \frac{P(A \cap B)}{P(A)} \tag{7}$$

By rearranging 6 and 7, we can eliminate $P(A \cap B)$ so that:

$$P(B)P(A|B) = P(A \cap B) = P(A)P(B|A) \tag{8}$$

Rearranging again leads us straight to **Bayes' Rule**:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \tag{9}$$

Dependent and Independent Events

Two events are independent if one event does not affect the other. For instance, if toss a coin twice, the result of the second coin toss does not depend on the first coin toss. Events can appear, intuitively, to be independent when they are not. Let us again consider the peanut butter/jam example. The probabilty of someone liking jam is 0.53 (equation 2). However, if we know beforehand that the person likes peanut butter, the chance of them liking jam decreases (as in Equation 5). In this example, knowing something about a person's peanut butter preference changes how likely they are to like jam!

In a truly independent case, knowing whether someone likes peanut butter would have not impact on whether they liked jam. Mathematically, this can be written as:

$$P(B) = P(B|A) \tag{10}$$

Whilst this is a toy example, it is important to bear in mind that in real-life, two seemingly independent events may end up becoming dependent if they are dependent on a third, unobserved factor (I'd highly recommend the book Freakonomics for further information on this topic).