# Day 2: Bayes Rule

Stephen R. Proulx

## Understanding Bayes' Rule

Goals for today:
* Learn the form of Bayes' Rule
* Develop an understanding of how the components of Bayes' Rule affect the posterior probability
* Launch the shiny app in R and manipulate the input
* Write out your interpretation of the calculation

### Bayes' Rule

The general form of Bayes' Rule is shown here:

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

It gives as output the probability of event A given that we have observed event B. It takes as input the total probabilities of A and B, and the opposite conditional probability, that B happens given an observation of A. There are a lot of ways to think about how these pieces come together, and I have drawn some of them out in the additional file and on the board during class.

Remember that a total probability can be found by summing the component conditional probabilities.

$$Pr(B) = \sum_i Pr(B|p_i)Pr(p_i)$$

#### Applied to cold/covid symptoms

Here's a concrete example that we'll explore.

$$Pr(\text{covid}|\text{symptoms}) = \frac{Pr(\text{symptoms}|\text{covid})Pr(\text{covid})}{Pr(\text{symptoms})}$$

#### Symptoms and Cause

$$Pr(\text{covid}|\text{symptoms}) = \frac{Pr(\text{symptoms}|\text{covid})Pr(\text{covid})}{Pr(\text{symptoms})}$$

Which we can refine a bit by noting that their are two routes to have symptoms, you have covid or you have the flu. (of course in reality there are more reasons to have symptoms, it could be allergies, or paranoia.... But for these calculations we assume only two routes.)

$$Pr(\text{covid}|\text{symptoms}) = \frac{Pr(\text{symptoms}|\text{covid})Pr(\text{covid})}{Pr(\text{symptoms}|\text{covid})Pr(\text{covid}) + Pr(\text{symptoms}|\text{flu})Pr(\text{flu})}$$

Some example probabilities for rare covid:

$$Pr(\text{symptoms}|\text{covid}) = 0.3 Pr(\text{covid}) = 0.001 Pr(\text{symptoms}|\text{flu}) = 1 Pr(\text{flu}) = 0.02$$

If covid is rare, how likely is it that I have covid given that I have symptoms?

$$Pr(\text{covid}|\text{symptoms}) = \frac{0.3 * 0.001}{0.3 * 0.001 + 1 * 0.02} = 0.015$$

Now consider if covid is quite common:

$$Pr(\text{covid}) = 0.1$$

If covid is common, how likely is it that I have covid given that I have symptoms?

$$Pr(\text{covid}|\text{symptoms}) = \frac{0.3 * 0.1}{0.3 * 0.1 + 1 * 0.02} = 0.6$$

**Breakout: explore the shiny app**

Use this link to start your VM: https://bit.ly/EEMB174.
Add a directory to store your work. You will want to use the "export" function to save copies of your work on your own computer.

Navigate to the "In Class Material" folder and find today's date. Click on the file "app.R". Once it is loaded into the code tab you can run this shiny app by pressing the "Run App" button.

**One final note:**

Here we considered a situation where someone has flu-like symptoms, and wants to estimate the chances that the symptoms are due to covid. We see that the frequency of both flu and covid in the population as a whole has a "surprising' ' effect on the caclulations. This is very hard to accept based on most of the mathemtatical training you have recieved before. But it is true.

We did this example thinking about symptoms, but it is exactly the same problem if we look at tests for the presence of covid. Just replace "has sypmtoms" with "has a positive test result". If people who do not have covid can possibly test positive, and people who do can possibly test negative, then we have the same exact equation, just with different names for the categories.

This was why, in the early covid days when antibody testing was being used to assess covid prevalence in the population, there was a strong cautionary warning that testing positive did not mean that you had covid, even when the chance of a false positive of the test was low. When almost no one has had covid, almost every postive antibody test is due to a false positive!

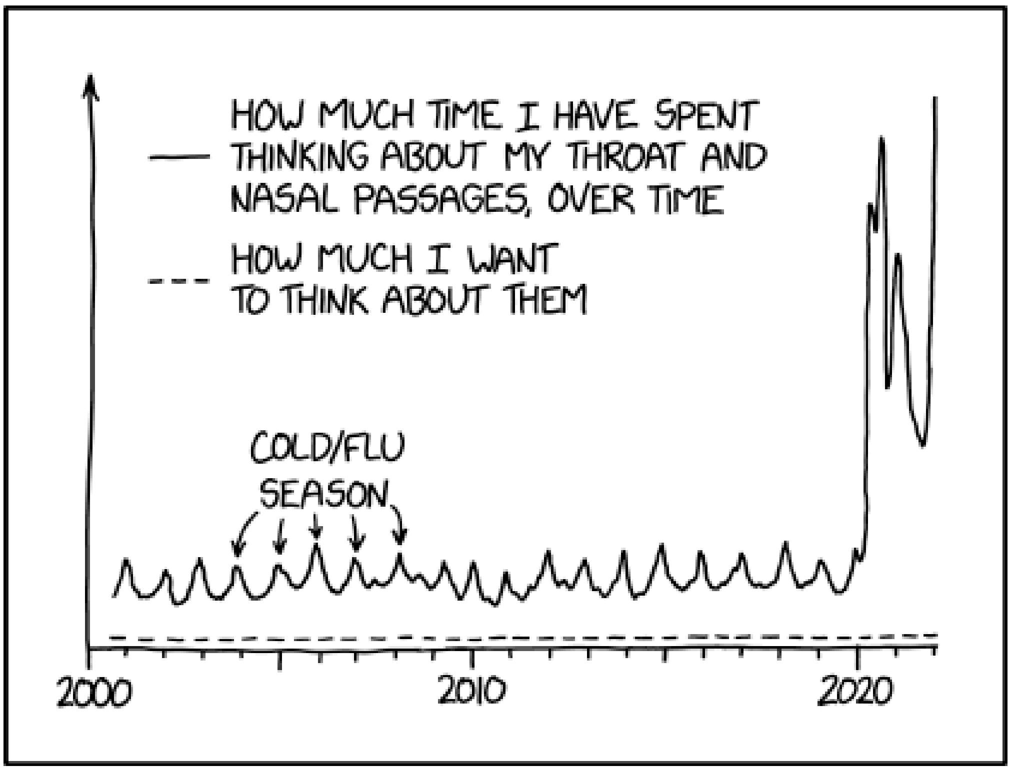


Figure 1: xkcd https://xkcd.com/2563/