

Day 2: Bayes Rule

Stephen R. Proulx

1/4/2022

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

Chance, probability, and the butterfly effect

Breakout: Experience with chance - Discuss with group 3 experiences in the last month where an outcome could be considered chance. - Discuss how you could convince a space alien that such an event was in some sense random. - Pick one of these to report back on and write two sentences to paste into the chat.

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 conditions probability, that B happens given an observation of A. There are a lot of ways to think about how these pieces come together, and we explore that in the virtual whiteboard file for today.

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 there 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.)

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

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

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

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

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

Breakout Room Thought Question

If covid is really rare, how much more probable is it that our patient has covid given symptoms? How does it depend on the probability of symptoms in both groups?

$$Pr(\text{covid}|\text{symptoms}) = \frac{P_{sc}P_c}{P_{sc}P_c + P_{sf}P_f}$$

$$\frac{Pr(\text{covid}|\text{symptoms})}{P_c} \approx \frac{P_{sc}}{P_{sf}P_f}$$

Consider a situation where both the flu and covid have the same chance of producing symptoms: What does our equation tell us then?

Now consider a case where covid is less likely to produce symptoms (more asymptomatic cases): What does our equation say?

And finally, the less realistic case, where covid always produces symptoms, but flu rarely does.

Breakout: explore the shiny app

Use this link to start your VM: <http://bit.ly/3hyqx4Y>.

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 calculations. This is very hard to accept based on most of the mathematical training you have received 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 symptoms” 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 positive antibody test is due to a false positive!

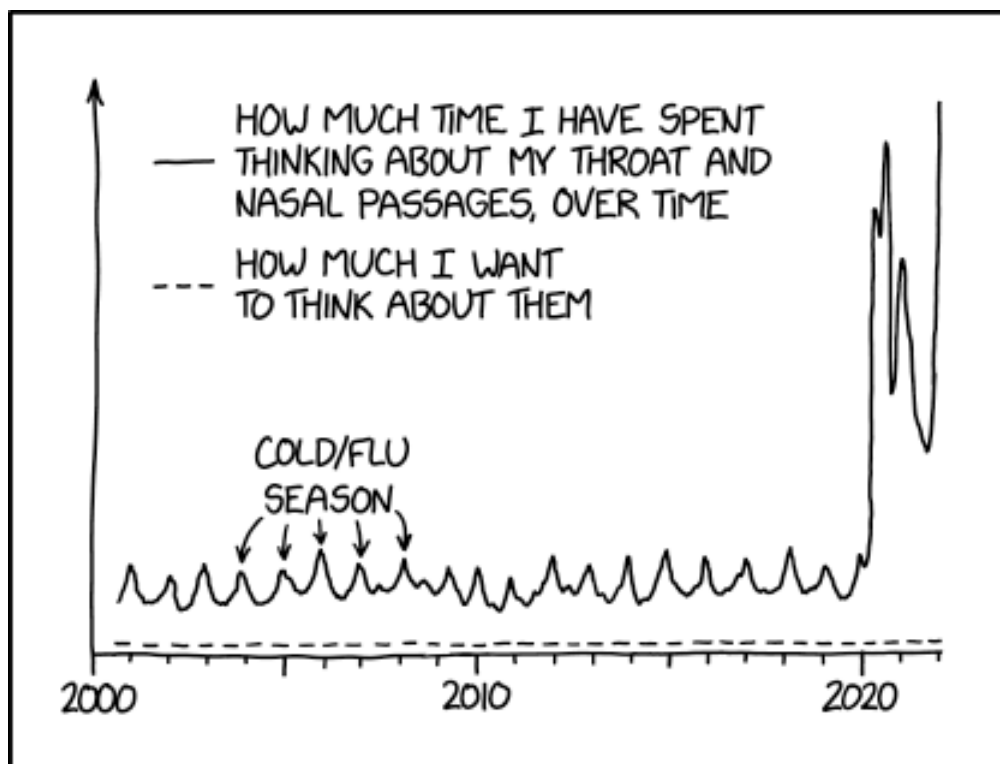


Figure 1: xkcd <https://xkcd.com/2563/>