

COMP3611: Machine Learning - Activity

Activity

Bayes' Rule - Real World Consequences

Marc de Kamps

October 24, 2023

Objectives

In this activity, we will see that Bayes' rule can be used to reason about every day situations. We will consider a medical case. Optionally, you can then move on to a legal case, which is quite hard - but interesting. In both cases professionals tend to give the wrong answer. Bayes' rule probably needs to be taught in medicine and in law school.

Both cases have information that come from the sources quoted in a New York times article: <https://opinionator.blogs.nytimes.com/2010/04/25/chances-are/>, which contains a number of interesting references. In the article, the answer is slightly off because they refrain from explain Bayes' rule. You, of course, will have to apply it correctly. Please read the article on breast cancer, but do not yet read about the O. J. Simpson trial.

Breast Cancer

We quote: *Eight out of every 1,000 women have breast cancer. Of these 8 women with breast cancer, 7 will have a positive mammogram. Of the remaining 992 women who don't have breast cancer, some 70 will still have a positive mammogram. Imagine a sample of women who have positive mammograms in screening. How many of these women actually have breast cancer?*

Exercise 1: Apply Bayes' rule to answer this question. To help you on its way: there are two stochastic variables: C which can attain two values: 0 or 1 ('no cancer' or 'cancer'), and T , which can attain two values: 0 or 1 ('negative test result' or 'positive test result'). The information above can be translated into probabilistic statements. For example, you have a prior $P(C = 1) = \frac{8}{1000}$ and have been given relevant conditional probabilities: $P(T = 1 | C = 1) = \frac{7}{8}$.

Translate the information in the statement into probability statements and use Bayes' rule to find the probability $P(C = 1 | T = 1)$. Comment on the outcome. Try to explain the result in plain English.

Answer:

We are interested in $P(C = 1 | T = 1)$, Bayes' rule is:

$$P(X | Y) = \frac{P(Y | X)P(X)}{\sum_x P(Y | X)P(X)}$$

In this case:

$$P(C = 1 | T = 1) = \frac{P(T = 1 | C = 1)P(C = 1)}{P(T = 1 | C = 0)P(C = 0) + P(T = 1 | C = 1)P(C = 1)}$$

All probabilities are either given or can easily be inferred: $P(C = 1) = \frac{8}{1000}$, $P(C = 0) = 1 - P(C = 1)$, $P(T = 1 | C = 1) = \frac{7}{8}$, $P(T = 1 | C = 0) = \frac{70}{992}$.

Working this out gives 9.1 %. Not a million miles from the answer cited in the article. I would expect you find this surprisingly low!

The reason is that cancer is relatively rare. The vast majority of those tested will not have it, but because there are many people in the general population, the number of false positives will be high compared to the number of true positives. The relatively low fraction of false positive results is offset by the large number of people tested who will not have cancer. In the article it is noted that even GPs get this estimate wrong.

But be careful! The question of whether to test or not is controversial and the simple picture suggested by this exercise changes as soon as any risk factor is involved. If ever offered a screening test, you should not take this simple exercise as an argument not to be tested! You should always discuss any doubts you might have with your GP or a specialist. (You may have to raise Bayes' rule with them).

Exercise 2 (optional): Most of you will now be too young to remember the process against O. J. Simpson. He was famous as a football player (American) and an actor. He was a prime suspect in the murder case of his wife, not least because he was followed in a helicopter fleeing the crime scene. He was nonetheless acquitted in a criminal trial. It has often been argued that the legal profession needs training in the use of Bayes' rule and this trial provides at least one example.

It was known that Simpson had battered his wife. His lawyer, Alan Dershowitz - a name you may remember from the more recent case against Jeffrey Epstein, argued that this point was irrelevant, because approximately 0.1 % of wife batterers go on to murder their wife and therefore this bit of information was irrelevant.

Before you go on, would you consider the information that Simpson had battered his wife (this was not disputed, apparently), relevant to the murder case, or do you accept Dershowitz's argument?

Now read the following letter to Nature by I. J. Good, ideally before reading the interpretation in the New York times article. <https://www.nature.com/articles/375541a0.pdf>

In order to understand it, you need to know what a Bayes' factor is. When there are two possible outcomes, guilty and not guilty, Bayes' rule reads:

$$P(G | X) = \frac{P(X | G)P(G)}{P(X)}$$

for one outcome and

$$P(\bar{G} | X) = \frac{P(X | \bar{G})}{P(X)}$$

When these are the only two possible outcomes, it is convenient to consider

$$\frac{P(G | X)}{P(\bar{G} | X)},$$

which is called the Bayes' factor.

- Why does he argue that Dershowitz's argument is deceptive?
- In Good's letter, he argues that correct interpretation of Dershowitz's information, even if this is accepted to be correct, means the probability that the husband has a history of battering his wife and the wife is murdered is at least $\frac{1}{2}$. This is not enough to ensure conviction. Is it relevant though?
- If interested, you may also consider reading a second letter by Good, also cited in the New York Times article, where he later argues that the above probability is at least 0.9.

Answer:

He argues that the relevant probability is not the fact that a wife batterer has a 0.1% probability of killing his wife. He argues that we should consider there was a history of battering *and* the wife was murdered. In the absence of any other information, this should be sufficient to put the husband in the frame at least. In the actual trial there was a lot of extra information. Far from being irrelevant, Good judged this history highly relevant.