

Bayes Rule Spreadsheet

If you're rational, you're Bayesian.

by Kenneth R. French

January 25, 2018

Bayes rule is a way to update your model of the world when you have new information. Suppose we are interested in assessing the probability that a specific hypothesis is true. We start with an initial assessment, called our prior, which is based on all the data we have observed, books we have read, and our other life experiences. How should we update our initial assessment when we observe new data?

$$\text{Posterior Assessment} = \text{Prior Assessment} \times \text{New Information} \quad (1)$$

Bayes rule is a transformation of a simple identity. We can express the probability that two things are true—it is raining and there is a traffic jam—in at least three ways,

$$P(\text{Rain and Jam}) = P(\text{Jam}|\text{Rain})P(\text{Rain}) = P(\text{Rain}|\text{Jam})P(\text{Jam}). \quad (2)$$

The probability that it is both raining and there is a traffic jam is the conditional probability of a traffic jam given that it is raining times the unconditional probability that it is raining. The joint probability of rain and traffic is also the probability that it is raining conditional on a traffic jam times the unconditional probability of a traffic jam.

If we start with the second half of Equation (2),

$$P(\text{Jam}|\text{Rain})P(\text{Rain}) = P(\text{Rain}|\text{Jam})P(\text{Jam}), \quad (3)$$

and divide by the unconditional probability of rain,

$$P(\text{Jam}|\text{Rain}) = P(\text{Jam}) \times P(\text{Rain}|\text{Jam}) / P(\text{Rain}), \quad (4)$$

we can update our unconditional assessment of a traffic jam when we observe it is raining.

Equation (4) is Bayes Rule. Our prior assessment in Equation (4) is the unconditional probability of a traffic jam, $P(\text{Jam})$. We adjust this assessment for the new information that it is raining by multiplying the unconditional probability $P(\text{Jam})$ by the ratio of the conditional probability that is raining when we observe a traffic jam and the unconditional probability of rain,

$$P(\text{Rain}|\text{Jam}) / P(\text{Rain}). \tag{5}$$

The ratio in Equation (4) is what we learn about the likelihood of a traffic jam when we observe it is raining.

We can generalize Equation (4) if we interpret a traffic jam as any hypothesis and rain as generic data. Then Bayes Rule is

$$P(\text{Hypothesis}|\text{Data}) = P(\text{Hypothesis}) \times P(\text{Data}|\text{Hypothesis}) / P(\text{Data}). \tag{6}$$

The spreadsheet implements the Rule.