Bayes Rule

Chase Brady

October 5, 2009

Conditional probabilities are not in general commutative, that is

$$P(A|B) \neq P(B|A).$$

On occasion, we are given one of these probabilities, when what we really would like to know is the other. For example, suppose a parent is considering enrolling their 8th grader in the Great Readers reading improvement program. In the company's brochure, they are told that out of all the local 10th graders who had scored in the top 10% on a recent reading skills test, 35% had been through the Great Readers program. This might sound impressive, but it is incomplete information. Suppose you found out that 30% of the students who did *not* score in the top 10% on this reading exam had been through the Great Readers program. Suddenly 35% of the top 10% doesn't sound quite so good, does it? The 35% figure is a conditional probability, i.e. the probability that a student went through the Great Readers program given that they scored in the top 10% on the reading skills test. We will denote this $P(GR|\text{top })$. What we really want to know is the probability that a child who has been through the Great Readers program will score in the top 10%, i.e. $P(\text{top }|GR)$.

Given sufficient information, we can use *Bayes Rule* to obtain a reversed conditional probability. The formal statement of Bayes rule is

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^C) \cdot P(A^C)}$$

where A^C is the complement of A. For our example above, we want

$$P(\text{top}|\text{GR}) = \frac{\text{P(GR|top)} \cdot \text{P(top)}}{\text{P(GR|top)} \cdot \text{P(top)} + \text{P(GR|top}^\text{C})) \cdot \text{P(top}^\text{C})}$$

$P(GR|\text{top}) = .35$ (The probability that a student in the top 10% has been through the Great Readers program.)
$P(GR|\text{top}^\text{C})) = .3$ (The probability that a student *not* in the top 10% has been through the Great Readers program.)
$P(\text{top}) = .1$ (The probability that a student is in the top 10%) and
$P(\text{top}^\text{C}) = .9$ (The probability that a student is not in the top 10%) so the calculation becomes

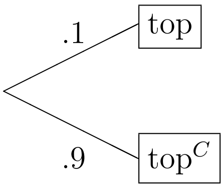
$$P(\text{top}|\text{GR}) = \frac{.35 \times .1}{.35 \times .1 + .3 \times .9} \approx .115 = 11.5\%$$

If the Great Readers program did not exist, the probability that a randomly chosen student would score in the top 10% would be, of course 10%, so 11.5% is not much of an improvement.

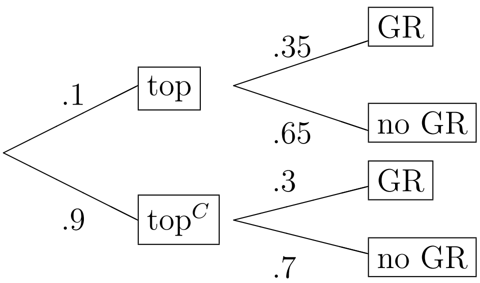
This same problem can also be approached using a tree diagram and the conditional probability rule

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

We begin our tree diagram with the non-conditional probabilities



A student either scored in the top 10% or the didn't, so the probabilities must sum to 1. For each branch, we draw two more branches corresponding to whether or not they had been through the Great Readers program. We put the conditional probabilities here.



$P(\text{top } \cap \text{ GR }) = .35 \times .1 = .035$

$P(\text{top } \cap \text{ no GR }) = .65 \times .1 = .065$

$P(\text{top}^C \ \cap \text{ GR }) = .3 \times .9 = .27$

$P(\text{top}^C \ \cap \text{ no GR }) = .7 \times .9 = .63$

The multiplication rule dictates we multiply the probabilities along the branches to get the joint probabilities at the end, as indicated in the figure. To get $P(GR)$, we add the probabilities for the first and third branches (the ones involving GR).

$$P(GR) = P(\text{top } \cap \text{ GR }) + P(\text{top}^C \ \cap \text{ GR }) = .035 + .27 = .305$$

Then by the conditional probability rule

$$P(\text{top}|\text{GR}) = \frac{\text{P(top } \cap \text{ GR)}}{\text{P(GR)}} = \frac{.035}{.305} \approx .115$$

In the example above, both events were either or cases, i.e. a student had either been through the Great readers program or they hadn't; they either scored in the top 10% or they didn't. Bayes rule can be extended to more complicated situations. If we have n different mutually exclusive alternative events $A_1, A_2, \ldots, A_n$, then the probability of B becomes

$$P(B) = P(B|A_1)\cdot P(A_1) + P(B|A_2)\cdot P(A_2) + \cdots + P(B|A_n)\cdot P(A_n) = \sum_{i=1}^{n} P(B|A_i)\cdot P(A_i)$$

Then Bayes rule becomes

$$P(A_j|B) = \frac{P(B|A_j)\cdot P(A_j)}{\sum_{i=1}^{n} P(B|A_i)\cdot P(A_i)}$$

Example 2: A bag contains a large number of balls, all the same size. 33% of them are red, 28% are white, and the rest are blue. They are either plastic or wood. 42% of the red balls are wood, and the rest are plastic. 65% of the white balls are wood, and the rest are plastic. 37% of the blue balls are wood, and the rest are plastic. If a ball is taken from the bag at random, what is the probability that it will be plastic? If it is plastic, what is the probability that it is red?

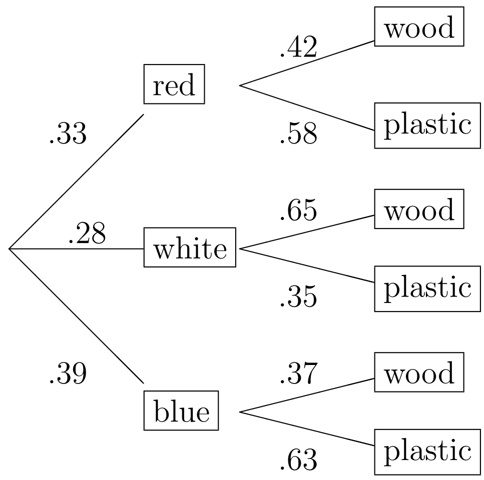
We want the probability that a randomly chosen ball is plastic, P(plastic), and the probability that it is red given that it is plastic, P(red|plastic). Writing the given information in the form of probabilities, we have

$$P(\text{red}) = .33 \qquad \text{P(white)} = .28 \qquad \text{P(blue)} = 1 - .28 - .33 = .39$$

$$P(\text{wood}|\text{red}) = .42 \qquad \text{P(wood|white)} = .65 \qquad \text{P(wood|blue)} = .37$$

$$P(\text{plastic}|\text{red}) = 1 - .42 = .58 \qquad \text{P(plastic|white)} = 1 - .65 = .35 \qquad \text{P(plastic|blue)} = 1 - .37 = .63$$

We could use the formula above, but a tree diagram might be a little clearer.



$P(\text{red} \cap \text{wood}) = .42 \times .33 = .139$

$P(\text{red} \cap \text{plastic}) = .58 \times .33 = .191$

$P(\text{white} \cap \text{wood}) = .65 \times .28 = .182$

$P(\text{white} \cap \text{plastic}) = .35 \times .28 = .098$

$P(\text{blue} \cap \text{wood}) = .37 \times .39 = .144$

$P(\text{blue} \cap \text{plastic}) = .63 \times .39 = .246$

Adding up the probabilities for the branches that result in plastic balls, we get

$$\begin{aligned} P(\text{plastic}) &= P(\text{red } \cap \text{ plastic }) + P(\text{white } \cap \text{ plastic }) + P(\text{blue } \cap \text{ plastic }) \\ &= .191 + .098 + .246 \\ &= .535 \end{aligned}$$

Using the conditional probability rule,

$$P(\text{red}|\text{plastic}) = \frac{\text{P(red } \cap \text{ plastic)}}{\text{P(plastic)}} = \frac{.191}{.535} = .357$$