Taxies and Bayes

Marcello Di Bello

Suppose there are two taxi companies, Green Cabs Inc. and Blue Cabs Inc., whose vehicles are respectively painted green and blue. Green Cabs Inc. covers 85 percent of the market and Blue Cabs Inc. covers the rest. There are no other taxi companies around. On a misty day a cab hits and injures a passerby, but it drives off. A witness reports that it was a blue cab. The witness is right only 80 percent of the time. This means that his *reliability* equals 0.8 in the sense that he gets the color right 80 percent of the time. Given the witness report, what is the probability that the taxi cab involved in the accident was in fact blue? Let G be the hypothesis that the taxi was green and B the hypothesis that the taxi was blue; let W_b be our evidence, namely the witness reporting that the taxi was blue. The problem is to determine the value of $P(B|W_b)$.

To apply Bayes' theorem, we need the values of $P(G)$, $P(W_b|B)$, and $P(W_b)$. First, Green Cabs Inc. covers 85 percent of the market, so the probability $P(G)$, without taking into consideration further evidence, equals 0.85. Notice that $P(G) = 1 - P(B)$, so $P(B)$ equals 0.15. Second, we know that the witness is correct 80 percent of the time. This measures his reliability: when he sees a taxi, he can identify the color correctly 80 percent of the time. Thus, $P(W_b|B)$ equals 0.8; also, $P(W_b|G)$ equals 0.2 because the witness is wrong 20 percent of the time.[1] Finally, what is the probability of the evidence $P(W_b)$? By the rule of total evidence, we have that $P(W_b) = P(W_b|B)P(B) + P(W_b|G)P(G)$. That is, $P(W_b) = 0.8 * 0.15 + 0.2 * 0.85 = 0.29$. We are now ready to apply Bayes' theorem:

$$P(B|W_b) = \frac{P(W_b|B)P(B)}{P(W_b|B)P(B) + P(W_b|G)P(G)} = \frac{0.8 * 0.15}{0.29} =\approx 0.41.$$

The interesting result here is that, even if the witness is right 80 percent of the time, the probability that the taxi was in fact blue given the testimony is still quite low. The reason is that the probability of B regardless of the evidence is low and equals 0.15. The table below shows that by varying the probability of B we can arrive at different probabilities of B given the witness testimony, holding fixed the witness reliability:

$P(B)$	$P(G)$	$P(W_b\|B)$	$P(B\|W_b)$
0.15	0.85	0.8	0.41
0.25	0.75	0.8	0.57
0.35	0.65	0.8	0.68
0.45	0.55	0.8	0.76
0.50	0.50	0.8	0.80
0.55	0.45	0.8	0.83
0.65	0.35	0.8	0.88
0.75	0.25	0.8	0.92
0.85	0.15	0.8	0.95

[1] The assumption is that the witness' reliability is 80 percent, i.e. the witness gets the color of the taxi right 80 percent of the time. Let W_g abbreviate 'the witness reports that the taxi was green.' We have $P(W_b|B) = P(W_g|G) = 0.8$. Now, since the witness can report that the taxi was blue or that it was green, W_g and W_b are exhaustive and mutually exclusive, whence $P(W_b|G) = 1 - P(W_g|G) = 0.2$.

Given Bayes' theorem, people should assign a probability of 0.41 to B once they learn about W_b, or at least this is how people should update their probability assignment according to Bayes' update. Do people actually follow Bayes' rule and Bayes' update? They don't. Psychological experiments have shown that ordinary people, when confronted with scenarios such as Taxi Cabs, typically answer that the probability that the taxi was blue equals 0.8. Why? There are different explanations. One is that there is some ambiguity in the expression "the witness is correct 80 percent of the time." It might mean: when the witness sees a cab, he gets the colour right 80 percent of the time; that is, $P(W_b|B) = 0.8$ and $P(W_b|G) = 0.2$. But the expression could also be taken to mean: when the witness claims that the taxi is blue, he is correct 80 percent of the time; that is, $P(B|W_b) = 0.8$. If people adopt the second interpretation, their answer would naturally be that there is an 80 percent chance that the taxi was blue, given the witness report.

A second, related explanation is that people tend to confuse two different probabilities: $P(A|B)$ and $P(B|A)$. This tendency is know as *inversion fallacy*. In our example the confusion would be between $P(W_b|B)$ and $P(B|W_b)$. People might take both probabilities to be 0.8, though the former equals 0.8 and the latter 0.41 (at least, according to Bayes' rule). As the table indicates, $P(W_b|B)$ and $P(B|W_b)$ are interchangeable only if the prior probability of B equals 0.5. And, presumably, the reason why people confuse the two probabilities is that they disregard prior probabilities and only take into account the witness report. This tendency is known as *base rate neglect*.

Some scholars have challenged the conclusion that 0.41 is the correct probability assignment. They have argued that ordinary people might not be committing a fallacy, after all. The problem boils down to the question: Why should our probability assignments (or our degrees of belief) conform to the Kolmogorov axioms of probability and Bayes' theorem? A standard answer is that if one does otherwise, he would be subject to a *Dutch Book*. This means that he would be willing to take bets in which he loses money no matter what happens.