reason and argument

John Kulvicki

Bayes's Rule

The last set of notes introduced you to how to manipulate probability schemata. This is useful because, like the sentential schemata that make up parts of these probability schemata, they can help us understand the claims we make better. Among other things, we learned how to understand probabilities of conjunctions, disjunctions, and negations based on the probabilities of the things conjoined, disjoined, and negated:

$$\begin{aligned} Pr(a \,\&\, b) &= Pr(a)Pr(b|a) = Pr(b)Pr(a|b) \\ Pr(a \vee b) &= Pr(a) + Pr(b) - Pr(a \,\&\, b) \\ Pr(-a) &= 1 - Pr(a) \end{aligned}$$

Take a look at the formula for the conjunction. Conjuncts can come in any order, which is why both of those formulas are equally good at giving the value of the probability of a conjunction. Those formulas are also helpful because they allow us to derive an important formula, Bayes' Rule:

$$Pr(a|b) = Pr(a)Pr(b|a)/Pr(b)$$

This is important because (1) it articulates the probability of a, conditioned on b, in terms of the probability of b conditioned on a as well as the unconditioned probabilities of a and b, and (2) we often want to know about the stuff described in (1). As a helpful exercise, try to work out some other formulae related to Bayes' rule. For example, it should be easy to work out a similar rule concerning probabilities conditioned on two statements, like $Pr(a|(b \,\&\, c))$.

Consider a pair of coins, one of them fair, one of them unfair. For the fair coin the probability of getting heads is ½, which is the same as the probability of getting tails. For the unfair coin, we get heads 2/3 of the time, and tails only 1/3 of the time. Now imagine that I've got just these two coins in my pocket, and I take one out and flip it three times. All tosses come out heads. Let's abbreviate as follows:

3: All three tosses come out heads.
F: The coin is fair.

Coin tosses are independent events, and since the probability of getting heads with the fair coin is ½, we know that the probability of getting heads three times in a row is ½* ½ * ½ = 1/8, or:

$Pr(3|F) = 1/8$

We also know that

$Pr(3|-F) = 2/3*2/3*2/3 = 8/27$

And since I pulled a coin at random from my pocket (and there are only two in my pocket!), we know that:

$Pr(F) = ½ = Pr(-F)$

Now imagine that someone offers an argument:

All three coin tosses came up heads.
Therefore, the coin is unfair.

This is not a deductively valid argument, but is it any good? For example, is

Pr(Conclusion | Premises) > Pr(-Conclusion | Premises)?

In this case:

$Pr(-F|3) > Pr(F|3)$?

We can use Bayes's rule and substitute for each side:

$$Pr(3|-F)Pr(-F)/Pr(3) > Pr(3|F)Pr(F)/Pr(3)$$

Notice that the denominators cancel each other out, and that Pr(F) = Pr(-F), so they cancel out as well. Overall, then, we want to know whether:

$Pr(3|-F) > Pr(3|F)$, or $8/27 > 1/8$ Yes.

The argument is not bad. The premise lends inductive support to the conclusion. What is the value of Pr(-F|3)? That is, what is the probability that the coin is unfair, given that I tossed it and got three heads in a row? We can think of the coin tosses as evidence that bears on an hypothesis to the effect that the coin is unfair, and Bayes's rule will let us figure out how much support that evidence lends to the hypothesis:

$$Pr(-F|3) = Pr(3|-F)Pr(-F)/Pr(3) = (8/27*1/2)/Pr(3) = (4/27)/Pr(3)$$

Usually, when using Bayes Rule in this manner, the one thing that's hard to discern is the unconditioned probability of getting your evidence, which is the term that will be in the denominator on the right-hand-side of the equation. What is the *unconditioned* probability that all three tosses come up heads? What is Pr(3)? We know that the tosses involve one or another of two coins. All tosses are either made with the fair coin or with the unfair coin. What we want to know is: what is the probability that we get three heads and the coin is fair or three heads and the coin is unfair? We want to know, that is:

$$Pr((3\&F) \vee (3\&-F))$$

Notice that this expression is logically equivalent to 3. If 3 is true, then the complex expression above is true, and if 3 is false, so is the complex one. In effect, we are just using a more complex expression for 3, which helps us because we can use it to figure out Pr(3). The complex should be easy to evaluate because we have formulas for the probabilities of disjunctions and conjunctions. So,

$$Pr((3\&F) \vee (3\&\text{-}F)) =$$
$$Pr(3\&F) + Pr(3\&\text{-}F) - Pr((3\&F)\&(3\&\text{-}F))$$

The last part of that formula is ugly, but luckily you can drop it. It can't happen that both of those conjunctions are true, because it can't happen that F and –F. So we are left with:

$$Pr(3\&F) + Pr(3\&\text{-}F) = Pr(3|F)Pr(F) + Pr(3|\text{-}F)Pr(\text{-}F)$$

And it turns out we have already computed all of those values

$$Pr(3|F)Pr(F) + Pr(3|\text{-}F)Pr(\text{-}F) = 1/8*1/2 + 8/27*1/2 =$$
$$1/16 + 4/27 = 91/432 = Pr(3)$$

Substituting this back into the formula:

$$Pr(\text{-}F|3) = (4/27)/(91/432) = 1728/2457 = \text{roughly } 7/10$$

That is to say, the probability that the coin is unfair, given three consecutive heads when it is tossed, is 7/10, or 70%. Because there is a big difference between the probabilities for each coin, and because the test was so interesting (all heads!), just three tosses give us excellent reasons for thinking the coin is unfair. You should work out how this number changes as you change (1) the results of the test (consider just one or two tosses, and also consider different patterns of heads and tails), and (2) the fairness of each coin. What if they are both unfair, in different ways, or both very close to fair?

Medical testing is one place where Bayes's Rule can be applied in eye-opening ways. Let's imagine a test for some disease, w, and consider the following abbreviations for claims about a subject, S.

w: S has the disease w.
p: S tests positive for having w.

Let's say, furthermore, that approximately 1 in 1000 people have w at any given time. So,

$Pr(w) = 1/1000$

The test for the disease is excellent, by medical standards. It is, for example, 100% *sensitive*, which means:

$Pr(p|w) = 1$ Sensitivity: 100%

Given that someone has the disease, they will test positive for it. If, for example $Pr(p|w) = 9/10$, then we say the test is 90% sensitive: it classifies 90% of those with the disease as such. Once a positive test result comes back, we must decide what to do. Should we commit those who pass the test to a painful and dangerous treatment program? Should we try that test again, just to be sure? Our decision should presumably depend upon how confident we are that person who tests positive actually has the disease, and our only source of confidence is this positive test, so what is the probability that S has w, given the positive test?

$Pr(w|p) = ?$

Luckily, we have the tools for answering this question. Or, at least we have the tools to know what further questions we need answered before we can answer this question. We know that:

$Pr(w|p) = Pr(p|w)Pr(w)/Pr(p)$

And we already know that:

$Pr(p|w) = 1$ (100% sensitivity)
$Pr(w) = 1/1000$ (*prior* probability of having w)

The prior probability is just the probability, unconditioned by the test result, someone has the disease. Our formula then becomes:

$Pr(w|p) = (1/1000)(1/Pr(p))$

Once again, we need to figure out the unconditioned probability of getting a positive result. Pr(p). Just as before, we want to know the probability of a disjunction, specifically:

$$Pr(p) = Pr((p \& w) \vee (p \& -w))$$

We want to know: what's the probability that you get a positive test and have the disease or you get a positive test but are not sick? That does seem to cover the space of options here, which is just what an unconditioned estimate of the probability should do. Also, we have an excellent formula for figuring out the probability of a disjunction based on the probabilities of its disjuncts:

$$Pr(p) = Pr((p \& w) \vee (p \& -w)) =$$
$$Pr(p \& w) + Pr(p \& -w) - Pr((p \& w) \& (p \& -w))$$

Take a look at the messy term toward the end. We always need to consider that when figuring the probability of a disjunction of statements. In this case, the term is equal to zero. What are the odds that a test is positive, and one both has the disease and does not? Zero. Great. We have:

$$Pr(p) = Pr(p \& w) + Pr(p \& -w)$$

And we already know how to deal with probabilities of conjunctions. Let's expand the previous formula to get:

$$Pr(p) = Pr(p|w)Pr(w) + Pr(p|-w)Pr(-w)$$

We already know most of these values, which is excellent. The left-hand part of the sum is just the sensitivity multiplied by the prior probability, so 1 times 1/1000. The right hand side is a little tricky, but not bad. We know

$$Pr(-w) = 1 - Pr(w) = 1 - .001 = .999$$

What we do not know, and what I have not told you yet, is the value of $Pr(p|-w)$. Though my test is 100% sensitive, this probability asks after the *selectivity* of the test. How often does the test misfire? What are the chances that a healthy person will produce a positive test for having w? What is the *false positive* rate? My test has a selectivity of 95%. That means, in 95% of cases in which one is not sick, this test says so, but in 5% of cases, this test comes back positive. So:

$$Pr(p|{-w}) = .05 = 1/20$$

This is still a fairly excellent test, by medical standards. Now we have all the components of our formula: it's just a matter of plug and chug. First:

$$Pr(p) = Pr(p|w)Pr(w) + Pr(p|{-w})Pr({-w}) =$$
$$1/1000 + (999/1000)(1/20) = (20 + 999)/20000 = 1019/20000$$

This is just about equal to 1/20, which matters, as we will see.
Now that we know Pr(p), we can plug it into the formula and solve for the value we want:

$$Pr(w|p) = (1/1000)(1/Pr(p)) = (1/1000)(20000/1019) = 20/1019 =$$
just about 2%

This says something quite surprising to most people unfamiliar with this stuff: the probability that S has w, given that she tested positive for it, is 2%: 1 chance in 50. *50.* How could this be?

It's relatively rare for a person to have the disease w. Our best guess is that only 1 in 1000 have the disease. So, the base rate of this affliction is a mere .1%. The false positive rate of the test for the affliction is a *full order of magnitude* greater: 5%. Selectivity of 95% sounds pretty good, right? It isn't bad, but it's rather terrible in the face of a base rate that is so small. Notice, for example, that the value of Pr(p) is, as I said, "basically" 1/20. That is, the probability that you test positive for the disease is basically equal to the false positive rate! That's because those who test positive and also have the disease are a small fraction of the total cohort that tests positive.

The easiest way to see this, by my lights (though you might have a different way) is to imagine that we test 10,000 people. How many people test positive? Well, if the base rate is 1/1000, we should get 10 of them (the test gets all of them because it is 100% sensitive), and we should also get five percent of the remaining population. 5% of 10,000 is

500. The total? 510 people test positive, while only 10 of them (2%) are in fact sick. If we increase the selectivity of this test to 99%, leaving sensitivity at 100%, then the 10 sick people plus 100 others (1% of the total population) test positive. In that case, a positive test only results in a roughly 10% chance that you are in fact sick.

Here are a few things to try and work out for yourself. First, work out the numbers for my test if it does *not* have 100% sensitivity. Try some wild numbers, like 50% sensitivity and 90% sensitivity. This will give you a feel for how a test's efficacy changes for low base-rate problems as a function of sensitivity. Also, try some other figures for selectivity. Then make the base rate very high, say 30%, and see how changing sensitivity and selectivity alter the effectiveness of the test. These results are eye-opening.

Most medical tests have poorer selectivity and sensitivity than the test I imagine here. Given that, they are most helpful when applied to a population in which the base rate of a disease is high. As you increase the base rate, tests with limited selectivity and sensitivity become much more diagnostic. It's costly, psychologically, economically, and sometimes medically, to receive a positive test for a disease one does not have. So, it pays to get tested only when one is a member of a population with a reasonably high base rate of the disease. For example, it pays to get tested when you exhibit symptoms characteristic of a disease, since many people exhibiting those symptoms have it.

*