Bayes' Formula

MATH 107: Finite Mathematics

University of Louisville

March 26, 2014

Test Accuracy

A motivating question

A rare disease occurs in 1 out of every 10,000 people. A test for this disease is 99.9% accurate (that is, it correctly determines whether the disease is present or not 99.9% of the time). You tested positive for the disease.

Is the probability you actually have the disease about...

- 1%?
- 10%?
- 50%?
- 90%?
- 99%?

Believe it or not, the majority of positive test results are false, despite this test's high accuracy!

Justification of that result

There are four possible classifications for any person: they could have the disease or not, and could test positive or negative.

Let's look at how we would expect 10,000,000 people to be classified:

	Test positive	Test negative	Total
Have disease	999	1	1,000
No disease	9,999	9,989,001	9,999,000
Total	10,998	9,989,002	10,000,000

1 of every 10,000 people has the disease.

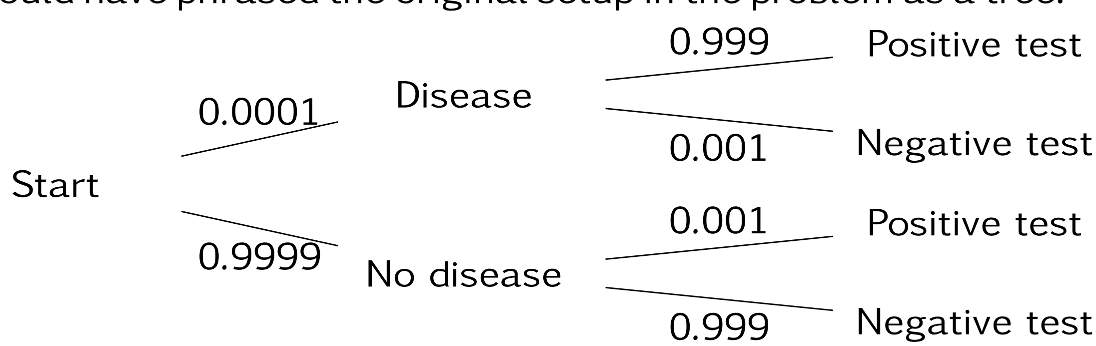
In each category, 1 of every 1,000 people gets the wrong test result.

We know we're one of the 10,998 people who test positive; but only 999 of those have the disease, so our chances of having the disease are:

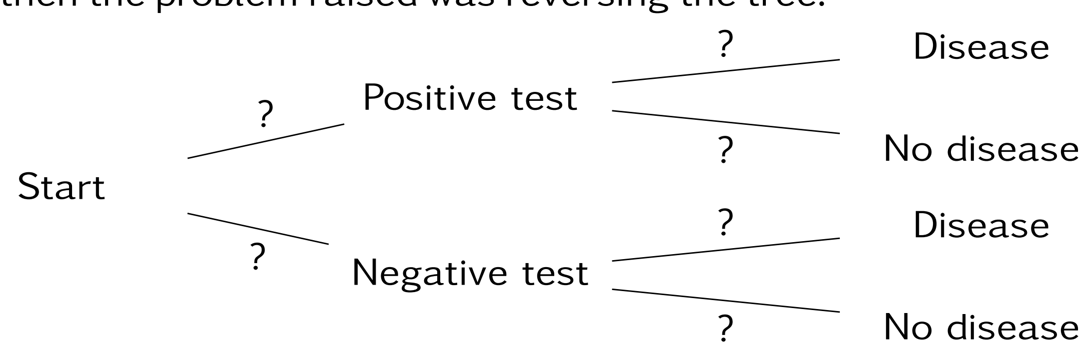
$$\frac{999}{10998} \approx 9\%$$

What we're doing here: as trees

We could have phrased the original setup in the problem as a tree:



And then the problem raised was reversing the tree:



What we're doing here: as events

We could identify the disease as event D, and a positive test as T.

The scenario told us that $P(D) = 0.0001$, and that $P(T|D) = P(T'|D') = 0.999$.

The question at the end of the scenario was to evaluate $P(D|T)$.

Thus, in a nutshell, what we want to do is *reverse* conditional probabilities which we know.

Deriving a formula

Suppose we know $P(A)$, $P(B|A)$, and $P(B|A')$. How do we find $P(A|B)$?

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(A \cap B) + P(A' \cap B)} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')(1 - P(A))} \end{aligned}$$

This is one of several variations on what is known as *Bayes' formula*.

Using Bayes' formula

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')(1 - P(A))}$$

In our disease example, the disease (event A) had probability 0.0001, while the test (event B) occurred with conditional probabilities $P(B|A') = 0.001$ and $P(B|A) = 0.999$, so:

$$P(A|B) = \frac{0.999 \times 0.0001}{0.999 \times 0.0001 + 0.001 \times 0.9999} \approx 0.0908$$

as we had previously worked out.

Detecting a fake

An unfair coin problem

We have two coins, one of which is fair and the other of which lands heads-up 75% of the time, but we don't know which is which. We pick a random coin and flip it eight times and get the results HHHHHTHT. This *looks* like it's the unfair one, but how certain can we be?

We have two events: A is the event that the chosen coin is in fact the unfair one, and B would be the event that 8 coin flips come out the way ours did. What we want to find is $P(A|B)$.

We know:

$$P(A) = \frac{1}{2}; P(B|A) = \left(\frac{3}{4}\right)^6 \left(\frac{1}{4}\right)^2; P(B|A') = \left(\frac{1}{2}\right)^8$$

Detecting a fake, continued

$$P(A) = \frac{1}{2}; P(B|A) = \frac{729}{2^{12}}; P(B|A') = \frac{1}{2^6}$$

We apply Bayes' formula:

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')(1 - P(A))} \\ &= \frac{\frac{729}{2^{12}} \times \frac{1}{2}}{\frac{729}{2^{12}} \times \frac{1}{2} + \frac{1}{2^6} \times \frac{1}{2}} \\ &= \frac{729}{793} \approx 92\% \end{aligned}$$

which is a good but not ironclad certainty that we have the right choice.

Another variation on Bayes' theorem

Sometimes instead of having a single prior result we want to find, we might have one of several mutually exclusive possibilities.
We might have mutually exclusive events $A_1, A_2, \ldots, A_n$ with total probability 1 and an event B, and want to find $P(A_1|B)$ from all the $P(A_i)$ and $P(B|A_i)$:

$$\begin{aligned} P(A_1|B) &= \frac{P(A_1 \cap B)}{P(B)} \\ &= \frac{P(B|A_1)P(A_1)}{P(B)} \\ &= \frac{P(B|A_1)P(A_1)}{P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)} \\ &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)} \end{aligned}$$

A multiple-possibility scenario

Two fake coins!

We now happen to have a collection of three coins; one fair, one of which lands heads-up $\frac{3}{4}$ of the time, and one of which lands tails-up $\frac{3}{4}$ of the time. We pick one at random, flip it 7 times, and get the result TTHHHTH. What is the chance we chose the fair coin?

Let A_1 be the event of picking the fair coin, A_2 the event of picking the heads-biased coin, A_3 the event of picking the tail-biased coin, and B the event of flipping our chosen coin to get TTHHHTH. We want to know $P(A_1|B)$.Clearly, each $P(A_i) = \frac{1}{3}$, and:

$$P(B|A_1) = \left(\frac{1}{2}\right)^7; P(B|A_2) = \left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right)^3; P(B|A_3) = \left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^4.$$

A multiple-possibility scenario, cont'd

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}; P(B|A_1) = \frac{1}{2^7}; P(B|A_2) = \frac{81}{4^7}; P(B|A_3) = \frac{27}{4^7}$$

We apply Bayes' formula:

$$\begin{aligned} P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)} \\ &= \frac{\frac{1}{2^7} \times \frac{1}{3}}{\frac{1}{2^7} \times \frac{1}{3} + \frac{81}{4^7} \times \frac{1}{3} + \frac{27}{4^7} \times \frac{1}{3}} \\ &= \frac{128}{236} \approx 54\% \end{aligned}$$

which is better than a blind guess (right 33% of the time), but not by much!

Another unequal-probability scenario

Two unfair coins. . . and lots of fair ones!

Our heads-up (75%) and tails-up (75%) coins are mixed in with eight normal coins. We pick one at random, flip it 6 times, and get all heads. What are the probabilities we got the heads-biased coin? The tails-biased coin? A fair coin?

Let A_1 be the event of picking a fair coin, A_2 the event of picking the heads-biased coin, A_3 the event of picking the tail-biased coin, and B the event of flipping six heads. We're interested in the three conditional probabilities $P(A_i|B)$.

$$P(A_1) = 0.8; P(A_2) = P(A_3) = 0.1$$

$$P(B|A_1) = \frac{1}{2^6}; P(B|A_2) = \frac{3^6}{4^6}; P(B|A_3) = \frac{1}{4^6}$$

Another unequal-probability scenario, cont'd

Applying Bayes' formula:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}$$
$$= \frac{\frac{1}{2^6} \times 0.8}{\frac{1}{2^6} \times 0.8 + \frac{729}{4^6} \times 0.1 + \frac{1}{4^6} \times 0.1} = \frac{256}{621} \approx 41\%$$

$$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}$$
$$= \frac{\frac{729}{4^6} \times 0.1}{\frac{1}{2^6} \times 0.8 + \frac{729}{4^6} \times 0.1 + \frac{1}{4^6} \times 0.1} = \frac{729}{1242} \approx 59\%$$

$$P(A_3|B) = \frac{P(B|A_3)P(A_3)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}$$
$$= \frac{\frac{1}{4^6} \times 0.1}{\frac{1}{2^6} \times 0.8 + \frac{729}{4^6} \times 0.1 + \frac{1}{4^6} \times 0.1} = \frac{1}{1242} \approx 0.08\%$$

Where we use this

In many problems, the real-world state is not yet known, and the test is *all we have*.

Even the probabilities associated with the real-world state are often unknown!

In practice, we often assume some simple probability distribution (called the *prior probability*) for the real-world state, and use Bayes' formula to figure out a distribution which better fits the evidence (called the *posterior probability*).

By using the posterior of one experiment to determine the prior of another, we can get a very accurate idea of the "true" underlying probabilities!

This technique is known as *Bayesian statistics*.