Probabilistic Diagnosis

99% accurate TB testing
A great-sounding diagnostic test for TB: if you have TB, the test is guaranteed to detect it. If you don't have TB, the test says so 99% of the time.
99% accurate TB testing

A great-sounding diagnostic test for TB: if you have TB the test is guaranteed to detect it. If you don't have TB, the test says so 99% of the time.

Your doctor gives you the test, and it says you have TB!

Do you have TB?

What is the probability that you have TB given that a 99% accurate test says you do?

Pr[TB | +] = ?

“+” for [test positive]

99% accurate TB testing

The test says TB!

TB is a serious disease and the test is at least 99% accurate. How worried should you be? What is the probability that you actually have TB?

Do you have TB?

Pr[ + | TB] = 1

Pr[ + | not TB] = \frac{1}{100}

false positive rate only 1%
Do you have TB?

\[
\Pr[\text{TB} \mid +] = \frac{\Pr[\text{TB AND } +]}{\Pr[+]} \\
= \frac{\Pr[+ \mid \text{TB}] \cdot \Pr[\text{TB}]}{\Pr[+]} \\
= \frac{1}{\Pr[+]} \\
= \frac{\Pr[+ \mid \text{TB}] \cdot \Pr[\text{TB}]}{\Pr[+]} \\
= \frac{\Pr[\text{TB}] \cdot \Pr[+]}{\Pr[+]} \\
\]

You do or you don’t

\[
\Pr[+] = \Pr[+ \mid \text{TB}] \cdot \Pr[\text{TB}] + \Pr[+ \mid \text{not TB}] \cdot \Pr[\text{not TB}] \\
\]

Total Probability Rule
You do or you don’t

\[
\Pr[+] = \Pr[+ \mid TB] \cdot \Pr[TB] \\
+ \Pr[+ \mid \text{not } TB] \cdot \Pr[\text{not } TB] \\
= 1 \cdot \Pr[TB] \\
+ \frac{1}{100} \cdot \Pr[\text{not } TB]
\]

Probability of Testing Positive

\[
\Pr[+] = \Pr[+ \mid TB] \cdot \Pr[TB] \\
+ \Pr[+ \mid \text{not } TB] \cdot \Pr[\text{not } TB] \\
= \frac{99}{100} \Pr[TB] + \frac{1}{100}
\]

You do or you don’t

\[
\Pr[+] = \Pr[+ \mid TB] \cdot \Pr[TB] \\
+ \Pr[+ \mid \text{not } TB] \cdot \Pr[\text{not } TB] \\
= 1 \cdot \Pr[TB] \\
+ \frac{1}{100} \cdot (1 - \Pr[TB])
\]

Do you have TB?

\[
\Pr[TB \mid +] = \frac{\Pr[TB]}{\Pr[+]}
\]

\[
= \frac{\frac{99}{100} \Pr[TB] + \frac{1}{100}}{\Pr[+]}
\]
Do you have TB?

\[
\Pr[\text{TB} | +] = \frac{\Pr[\text{TB}]}{\Pr[+]} = \frac{100\Pr[\text{TB}]}{99\Pr[\text{TB}] + 1}
\]

What is \(\Pr[\text{TB}]\)?

11,000 TB cases reported

CDC got reports of 11,000 cases of TB in US in 2011. Will be lots of unreported. So estimate:

\[
\Pr[\text{TB}] \approx \frac{1}{10,000}
\]

Do you have TB?

\[
\Pr[\text{TB} | +] = \frac{100\Pr[\text{TB}]}{99\Pr[\text{TB}] + 1} \approx \frac{100}{10000} \approx \frac{1}{100}
\]

Unlikely you have TB

Because of relatively high false positive rate (1%) compared to TB rate (0.01%), chance of having TB remains small (1%)!
Unlikely you have TB
99% accurate test is not so
good here.

A “more accurate” test
99% accurate test is not so
good here. In fact, there’s
a trivial test that is 99.99%
accurate:
always say “No TB”

Bayes Rule
\[
\Pr[TB | +] = \frac{\Pr[+ | TB] \cdot \Pr[TB]}{\Pr[+]} \\
\Pr[B | A] = \frac{\Pr[A | B] \cdot \Pr[B]}{\Pr[A]}
\]

99% accuracy still useful
99% accurate test did
increase your probability
of TB 100 times.
99% accuracy still useful

99% accurate test did increase your probability of TB 100 times. If you only had 5M medicine doses for a population of 350M, whom should you medicate?

99% accuracy still useful

Medicate the 3.5M who test positive, and you’re likely to cure nearly all the cases.