

The expert amateur

Getting really good at becoming not too bad

Mehul Motani

Kent Ridge AI

How sick are you?

You're feeling sick and so you go to the doctor, who runs a barrage of tests. It turns out that you test positive for a very very rare disease that affects 0.1% of the population.

In shock, you ask the doctor how likely is it that you have the disease. The doctor says that the test is fairly good — the test correctly identifies 99% of those who have the disease and incorrectly identifies 1% of those who do not have the disease.

Q1: How likely is it that you have the disease?

Being concerned, you go to another doctor, who also runs a barrage of tests (using another lab). It turns out you also test positive for the same very very rare disease.

Q2: How likely is it that you have the disease now?

Bayes Theorem

$$P(+, D) = P(+|D)P(D) = P(D|+)P(+)$$

$$P(D|+) = \frac{Pr(+|D)P(D)}{P(+)} \Rightarrow P(D|+) = \frac{Pr(+|D)P(D)}{P(+|D)P(D) + P(+|ND)P(ND)}$$

$$Q1 : P(D|+) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.01 \times 0.999} = 0.090$$

$$Q2 : P(D|+) = \frac{0.99 \times 0.09}{0.99 \times 0.09 + 0.01 \times 0.91} = 0.907$$

Bayes Theorem makes sense

- Imagine 1000 people. Based on the given statistics, 1 person has the disease on average. This person will get identified as sick with very high probability.
- The test will also identify as sick 10 people who do not have the disease.
- So you are part of a group of 11 people, of which only one has the disease — so the likelihood you have the disease is 1 in 11 —> 9%
- After the second positive test, the likelihood of you having the disease goes up to 91% which also makes sense. It is unlikely that both labs are wrong.
- However, note that the probability is still not as high as the probability that the test is right given the disease.

History of Bayes Theorem (from Wikipedia)

- Bayes' theorem was named after [Thomas Bayes](#) (1701–1761), who studied how to compute a distribution for the probability parameter of a [binomial distribution](#) (in modern terminology).
- Bayes's unpublished manuscript was significantly edited by [Richard Price](#) before it was posthumously read at the [Royal Society](#). Price edited^[8] Bayes's major work "[An Essay towards solving a Problem in the Doctrine of Chances](#)" (1763), which appeared in [Philosophical Transactions](#),^[9] and contains Bayes' theorem. Price wrote an introduction to the paper which provides some of the philosophical basis of [Bayesian statistics](#). In 1765, he was elected a Fellow of the Royal Society in recognition of his work on the legacy of Bayes.^{[10][11]}
- The French [mathematician Pierre-Simon Laplace](#) reproduced and extended Bayes's results in 1774, apparently unaware of Bayes's work.^{[note 1][12]} The [Bayesian interpretation](#) of probability was developed mainly by Laplace.^[13]

MODIFIED BAYES' THEOREM:

$$P(H|X) = P(H) \times \left(1 + P(C) \times \left(\frac{P(X|H)}{P(X)} - 1\right)\right)$$

H: HYPOTHESIS
X: OBSERVATION
P(H): PRIOR PROBABILITY THAT H IS TRUE
P(X): PRIOR PROBABILITY OF OBSERVING X
P(C): PROBABILITY THAT YOU'RE USING BAYESIAN STATISTICS CORRECTLY

DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)

THIS NEUTRINO DETECTOR MEASURES WHETHER THE SUN HAS GONE NOVA.

THEN, IT ROLLS TWO DICE. IF THEY BOTH COME UP SIX, IT LIES TO US. OTHERWISE, IT TELLS THE TRUTH.

LET'S TRY.
DETECTOR! HAS THE SUN GONE NOVA?

ROLL
YES.

$$P\left(\begin{matrix} \text{I'M NEAR} \\ \text{THE OCEAN} \end{matrix} \middle| \begin{matrix} \text{I PICKED UP} \\ \text{A SEASHELL} \end{matrix}\right) = \frac{P\left(\begin{matrix} \text{I PICKED UP} \\ \text{A SEASHELL} \end{matrix} \middle| \begin{matrix} \text{I'M NEAR} \\ \text{THE OCEAN} \end{matrix}\right) P\left(\begin{matrix} \text{I'M NEAR} \\ \text{THE OCEAN} \end{matrix}\right)}{P\left(\begin{matrix} \text{I PICKED UP} \\ \text{A SEASHELL} \end{matrix}\right)}$$

FREQUENTIST STATISTICIAN:

THE PROBABILITY OF THIS RESULT HAPPENING BY CHANCE IS $\frac{1}{32} = 0.03125$. SINCE $p < 0.05$, I CONCLUDE THAT THE SUN HAS EXPLODED.

BAYESIAN STATISTICIAN:

BET YOU \$50 IT HASN'T.

STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND DON'T HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.