

CS201 – Bayes' Theorem – Excerpts from Wikipedia

http://en.wikipedia.org/wiki/Bayes%27_theorem

http://en.wikipedia.org/wiki/Bayesian_inference

Bayes's theorem is stated mathematically as the following simple form:^[1]

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}.$$

For an epistemological interpretation:^[1]

For proposition A and evidence or background B ,

- $P(A)$, the **prior probability**, is the initial degree of belief in A .
- $P(A|B)$, the **conditional probability**, is the degree of belief in A having accounted for B .
- the quotient $P(B|A)/P(B)$ represents the support B provides for A .

Another form of Bayes's Theorem that is generally encountered when looking at two competing statements or hypotheses is:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B|A) P(A) + P(B|\neg A) P(\neg A)}.$$

For an epistemological interpretation:

For proposition A and evidence or background B ,^[3]

- $P(A)$, the **prior probability**, is the initial degree of belief in A .
- $P(\neg A)$, is the corresponding probability of the initial degree of belief against A : $1 - P(A) = P(\neg A)$
- $P(B|A)$, the **conditional probability** or likelihood, is the degree of belief in B , given that the proposition A is true.
- $P(B|\neg A)$, the **conditional probability** or likelihood, is the degree of belief in B , given that the proposition A is false.
- $P(A|B)$, the **posterior probability**, is the probability for A after taking into account B for and against A .

Frequentist example

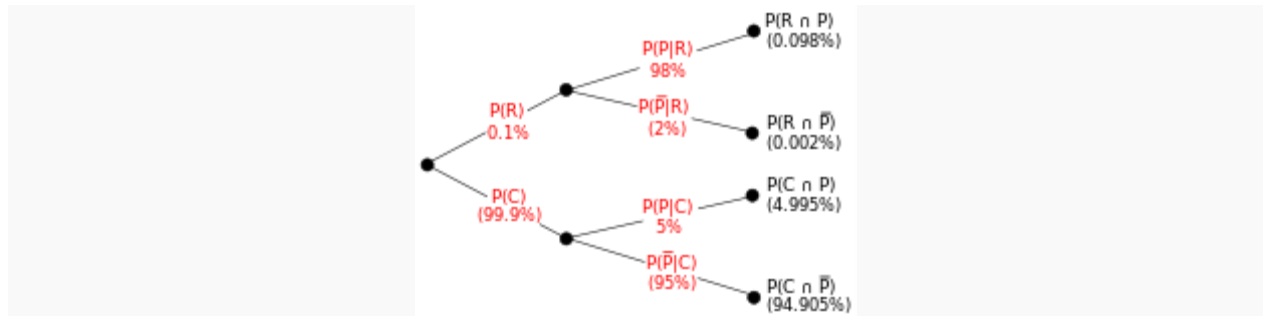
An entomologist spots what might be a rare subspecies of beetle, due to the pattern on its back. In the rare subspecies, 98% have the pattern, or $P(\text{Pattern}|\text{Rare}) = 98\%$. In the common subspecies, 5% have the pattern. The rare subspecies accounts for only 0.1% of the population. How likely is the beetle having the pattern to be rare, or what is $P(\text{Rare}|\text{Pattern})$?

From the extended form of Bayes's theorem (since any beetle can be only rare or common),

$$P(\text{Rare}|\text{Pattern}) = \frac{P(\text{Pattern}|\text{Rare})P(\text{Rare})}{P(\text{Pattern}|\text{Rare})P(\text{Rare}) + P(\text{Pattern}|\text{Common})P(\text{Common})}$$

$$= \frac{0.98 \times 0.001}{0.98 \times 0.001 + 0.05 \times 0.999}$$

$$\approx 1.9\%.$$



Tree diagram illustrating frequentist example. R, C, P and P bar are the events representing rare, common, pattern and no pattern. Percentages in parentheses are calculated. Note that three independent values are given, so it is possible to calculate the inverse tree (see figure above).

Coin flip example

"Assume that you are presented with three coins, two of them fair and the other a counterfeit that always lands heads. If you randomly pick one of the three coins, the probability that it's the counterfeit is 1 in 3. This is the prior probability of the hypothesis that the coin is counterfeit. Now after picking the coin, you flip it three times and observe that it lands heads each time. Seeing this new evidence that your chosen coin has landed heads three times in a row, you want to know the revised posterior probability that it is the counterfeit. The answer to this question, found using Bayes's theorem (calculation mercifully omitted), is 4 in 5. You thus revise your probability estimate of the coin's being counterfeit upward from 1 in 3 to 4 in 5."

The calculation ("mercifully supplied") follows:

$$P(\text{Biased coin}) = \frac{1}{3}$$

$$P(\text{Fair coin}) = \frac{2}{3}$$

$$P(H|\text{Fair coin}) = \frac{1}{2}$$

$$P(HHH|\text{Fair coin}) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

$$P(HHH|\text{Biased coin}) = 1 \times 1 \times 1 = 1$$

$$\begin{aligned} P(\text{Biased coin}|HHH) &= \frac{P(HHH|\text{Biased coin})P(\text{Biased coin})}{P(HHH|\text{Biased coin})P(\text{Biased coin}) + P(HHH|\text{Fair coin})P(\text{Fair coin})} \\ &= \frac{1 \times \frac{1}{3}}{1 \times \frac{1}{3} + \frac{1}{8} \times \frac{2}{3}} = \frac{\frac{1}{3}}{\frac{10}{24}} = \frac{4}{5} \end{aligned}$$

Drug testing

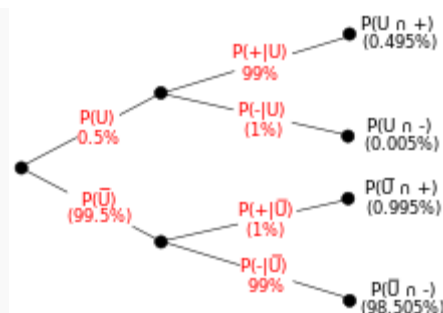
Suppose a drug test is 99% **sensitive** and 99% **specific**. That is, the test will produce 99% true positive results for drug users and 99% true negative results for non-drug users. Suppose that 0.5% of people are users of the drug. If a randomly selected individual tests positive, what is the **probability** he or she is a user?

$$\begin{aligned}
 P(\text{User}|+) &= \frac{P(+|\text{User})P(\text{User})}{P(+|\text{User})P(\text{User}) + P(+|\text{Non-user})P(\text{Non-user})} \\
 &= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995} \\
 &\approx 33.2\%
 \end{aligned}$$

Despite the apparent accuracy of the test, if an individual tests positive, it is more likely that they do *not* use the drug than that they do.

This surprising result arises because the number of non-users is very large compared to the number of users; thus the number of false positives (0.995%) outweighs the number of true positives (0.495%). To use concrete numbers, if 1000 individuals are tested, there are expected to be 995 non-users and 5 users. From the 995 non-users, $0.01 \times 995 \approx 10$ false positives are expected. From the 5 users, $0.99 \times 5 \approx 5$ true positives are expected. Out of 15 positive results, only 5, about 33%, are genuine.

Note: The importance of specificity can be illustrated by showing that even if sensitivity is 100% and specificity is at 99% the probability of the person being a drug user is $\approx 33\%$ but if the specificity is changed to 99.5% and the sensitivity is dropped down to 99% the probability of the person being a drug user rises to 49.8%. Even at 99% sensitivity and 99% specificity the probability of a person being a drug user is 47.5%.



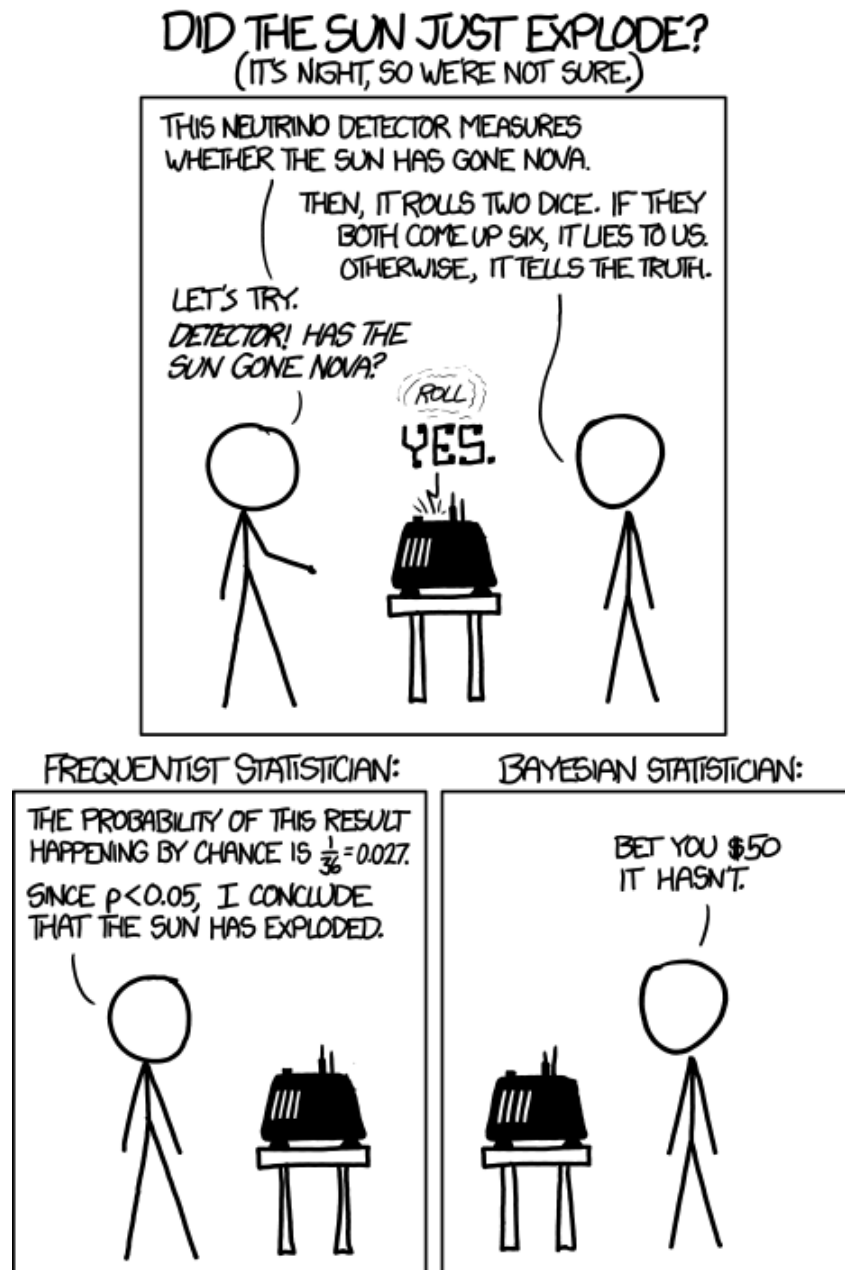
Tree diagram illustrating drug testing example. U, U bar, "+" and "-" are the events representing user, non-user, positive result and negative result. Percentages in parentheses are calculated.

Additional Examples:

<https://brilliant.org/discussions/thread/conditional-probability-bayes-theorem/>

<http://chemicalstatistician.wordpress.com/2013/10/30/detecting-an-unfair-die-with-bayes-theorem/>

XKCD: <http://xkcd.com/1132/>



<http://www.explainxkcd.com/wiki/index.php/1132: Frequentists vs. Bayesians>