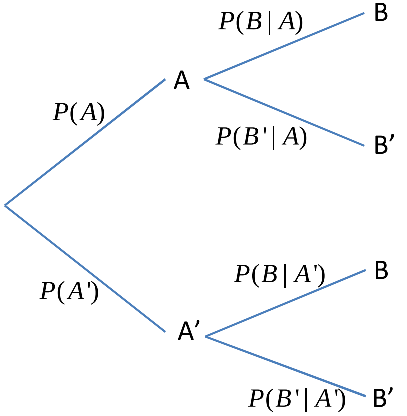
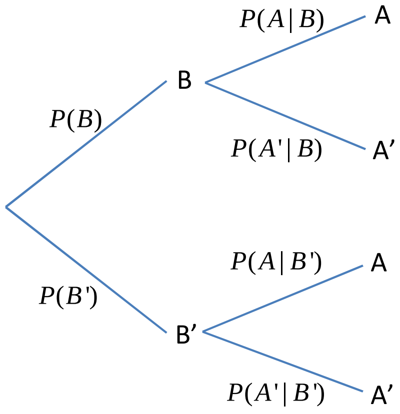
Conditional Probability, Bayes' Theorem and Inference

Consider two events A and B. Each can either occur or *not* occur. Event A *not occurring* is denoted by A' and B *not occurring* is denoted by B'. We can construct TWO tree diagrams to map out the possible permutations of outcomes. The probability of A occurring *given* B has occurred is P(A|B). Events A *and* B occurring is P(A & B). Note the order *does not matter* for the latter.

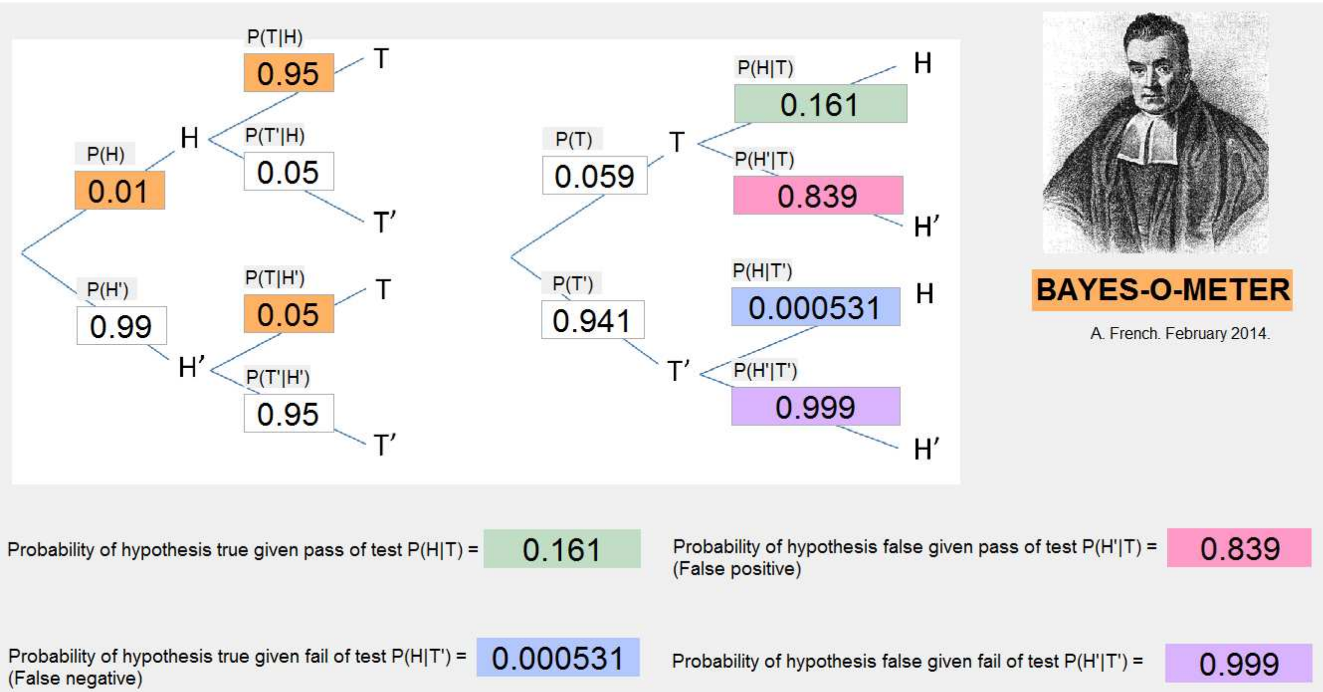




$$P(A \& B) = P(B \& A)$$
$$P(A)P(B\,|\,A) = P(B)P(A\,|\,B)$$
$$P(B\,|\,A) = \frac{P(B)P(A\,|\,B)}{P(A)}$$
$$P(B\,|\,A) = \frac{P(B)P(A\,|\,B)}{P(B)P(A\,|\,B) + P(B')P(A\,|\,B')}$$
$$\boxed{P(B\,|\,A) = \frac{1}{1 + \frac{P(B')P(A\,|\,B')}{P(B)P(A\,|\,B)}}}$$

This is called **Bayes' Theorem**, and allows P(B|A) to be computed from P(A|B), P(A) and P(B)

[Thomas Bayes](#)
1701-1761



An important application of **Bayes' Theorem** is *inference*.
Let H be a *hypothesis* e.g. 'a person has a particular disease' and T be a *test for the hypothesis to be true*.

If 100 samples are injected with the disease and the test yields a positive outcome for 95 of these, then **P(T|H) = 0.95.**

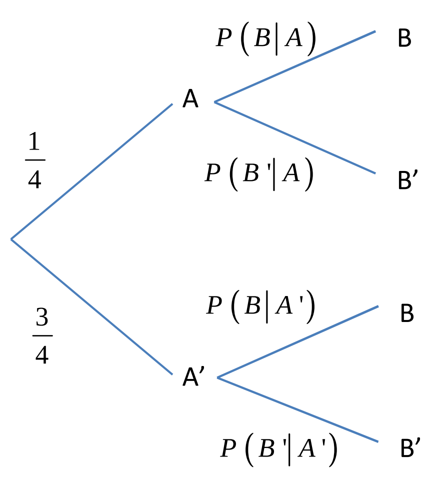
100 identical samples were then tested *without the disease being injected*. In this case, 5 were tested positive. Hence **P(T|H') = 0.05**

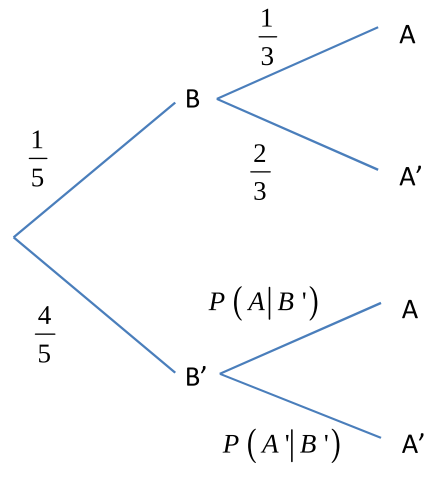
Based upon statistics of the disease, only 1% of patients who exhibit symptoms of the disease actually have the disease. Hence **P(D) = 0.01.**

The manufacturers of the test claim '95% accuracy' but this result is *misleading*. Applying Bayes' Theorem, the probability of having the disease *given* a positive test result is *only* 16.1%
i.e. **P(H|T) = 0.161**

Conditional probability examples

P(A|B) = 1/3, P(A) = 1/4 and P(B) = 1/5 : Find all the other probabilities in both tree diagrams corresponding to events A and B .





$$\frac{1}{4}P(B\,|\,A) = \left(\frac{1}{5}\right)\left(\frac{1}{3}\right)$$

$$\therefore \boxed{P(B\,|\,A) = \frac{4}{15}}$$

$$\therefore \boxed{P(B'\,|\,A) = \frac{11}{15}}$$

$$\frac{3}{4}P(B\,|\,A') = \left(\frac{1}{5}\right)\left(\frac{2}{3}\right)$$

$$\therefore \boxed{P(B\,|\,A') = \frac{8}{45}}$$

$$\therefore \boxed{P(B'\,|\,A') = \frac{37}{45}}$$

$$\frac{4}{5}P(A\,|\,B') = \left(\frac{1}{4}\right)P(B'\,|\,A) = \left(\frac{1}{4}\right)\left(\frac{11}{15}\right)$$

$$\therefore \boxed{P(A\,|\,B') = \frac{11}{48}}$$

$$\therefore \boxed{P(B'\,|\,A) = \frac{11}{15}}$$

Check!

Adapted from *The Signal and the Noise* by Nate Silver p247.
In this example H means "Terrorist attack", T means "planes crash into the World Trade Center"

PRIOR PROBABILITY		
Initial estimate of how likely it is that terrorists would crash planes into Manhattan skyscrapers	P(H)	0.005%
A NEW EVENT OCCURS: FIRST PLANE HITS WORLD TRADE CENTER		
Probability of plane hitting if terrorists are attacking Manhattan skyscrapers	P(T|H)	100%
Probability of plane hitting if terrorists are not attacking Manhattan skyscrapers (i.e. an accident)	P(T|H')	0.008%
POSTERIOR PROBABILITY		
Revised estimate of probability of terror attack, given first plane hitting World Trade Center	P(H|T)	38%

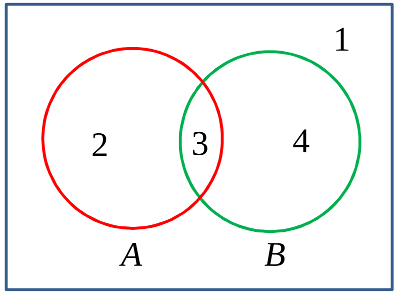
But then probability of terror attack, given *second* plane hitting the World Trade Center is **99.99%** since we re-do the analysis but set P(H) = 38%

Infamous criminal cases where convictions (or acquittals) have been made despite criticisms about the use of probability to justify the legal arguments ("The Prosecutors' Fallacy")

Sally Clark (1998) – Accused of murdering her two children rather than both dying of Sudden Infant Death Syndrome (SIDS).

O.J. Simpson (1994) – Accused of murdering his wife.

Bayes' Theorem and Venn diagrams



$P(A) = \frac{5}{10} = \frac{1}{2}$

$P(B) = \frac{7}{10}$

$P(B \mid A) = \frac{3}{5}$

$P(A \mid B) = \frac{3}{7}$

$P(A \,\&\, B) = P(A \cap B) = \frac{3}{10}$

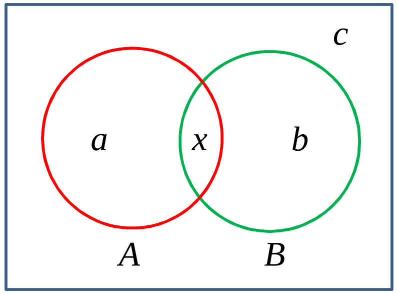
$P(A \text{ or } B) = P(A \cup B) = \frac{9}{10}$

$P(A)P(B \mid A) = \frac{1}{2} \times \frac{3}{5} = \frac{3}{10}$

$P(B)P(A \mid B) = \frac{7}{10} \times \frac{3}{7} = \frac{3}{10}$

i.e. **Bayes's Theorem** holds

$$P(A)P(B \mid A) = P(B)P(A \mid B)$$



Proof of Bayes' Theorem using Venn diagrams

$P(A) = \frac{x+a}{a+x+b+c}$

$P(B) = \frac{x+b}{a+x+b+c}$

$P(B \mid A) = \frac{x}{x+a}$

$P(A \mid B) = \frac{x}{x+b}$

$P(A \,\&\, B) = P(A \cap B) = \frac{x}{a+x+b+c}$

$P(A \text{ or } B) = P(A \cup B) = \frac{a+x+b}{a+x+b+c}$

$P(A)P(B \mid A) = \frac{x+a}{a+x+b+c} \times \frac{x}{x+a} = \frac{x}{a+x+b+c}$

$P(B)P(A \mid B) = \frac{x+b}{a+x+b+c} \times \frac{x}{x+b} = \frac{x}{a+x+b+c}$

$$\therefore \; P(A)P(B \mid A) = P(B)P(A \mid B)$$

Note:

$n(\varepsilon) = a + x + b + c$

$n(A) = a + x$

$n(B) = b + x$

$n(A \cap B') = a$

$n(B \cap A') = b$

Inverse function of Bayesian Inference Formula

Determine the test probability of success P(T|H) as a function of x = P(H|T).

Assume symmetry i.e. P(T|H') = 1 – P(T|H)

$$P(H \mid T) = \frac{1}{1 + \frac{P(H')P(T \mid H')}{P(H)P(T \mid H)}}$$

$p = P(T \mid H)$

$1 - p = P(T \mid H')$

$\alpha = P(H)$

$1 - \alpha = P(H')$

$x = P(H \mid T)$

$$x = \frac{1}{1 + \frac{(1-\alpha)(1-p)}{\alpha p}}$$

$$1 + \frac{(1-\alpha)(1-p)}{\alpha p} = \frac{1}{x}$$

$$\frac{(1-\alpha)(1-p)}{\alpha p} = \frac{1}{x} - 1 = \frac{1-x}{x}$$

$$(1-\alpha)(1-p) = \left(\frac{1-x}{x}\right)\alpha p$$

$$1 - \alpha = p\left\{\left(\frac{1-x}{x}\right)\alpha + 1 - \alpha\right\}$$

$$p = \frac{1-\alpha}{\left(\frac{1-x}{x}\right)\alpha + 1 - \alpha}$$

$$p = \frac{1-\alpha}{\alpha\left\{\frac{1-x}{x} - 1\right\} + 1}$$

$$p = \frac{1-\alpha}{\frac{\alpha}{x} - 2\alpha + 1}$$

$$p = \frac{x(1-\alpha)}{\alpha + x(1-2\alpha)}$$

Hence for the disease example above, the test needs to be 99.888% accurate to yield P(H|T) = 90%

$x = 0.9$

$\alpha = 0.01$

$$\therefore \; p = \frac{x(1-\alpha)}{\alpha + x(1-2\alpha)}$$

$$p = \frac{0.9 \times 0.99}{0.01 + 0.9 \times 0.98} \approx 0.99888$$

$$P(H \mid T) = \frac{1}{1 + \frac{(1-P(H))(1-P(T \mid H))}{P(H)P(T \mid H)}}$$

$p = P(T \mid H)$

$\alpha = P(H)$

$x = P(H \mid T)$

Assume symmetry i.e. P(T|H') = 1 – P(T|H)