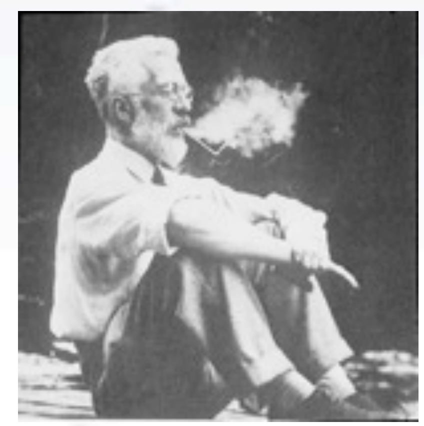
Applied Statistics

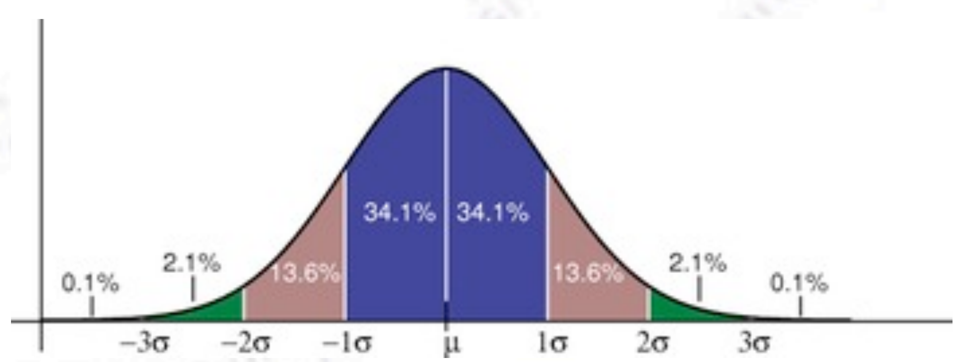
Bayes' Theorem








Troels C. Petersen (NBI)



"Statistics is merely a quantization of common sense"

Problem

Suppose a drug test can be characterized as follows:

• 99% positive results for users (99% sensitive, i.e. 1% Type I errors).

• 99% negative results for non-users (99% specific, i.e. 1% Type II errors).

If 0.5% of a population is using the drug, and a random person tests positive, what is the chance that he is using the drug?

Problem

Suppose a drug test can be characterized as follows:

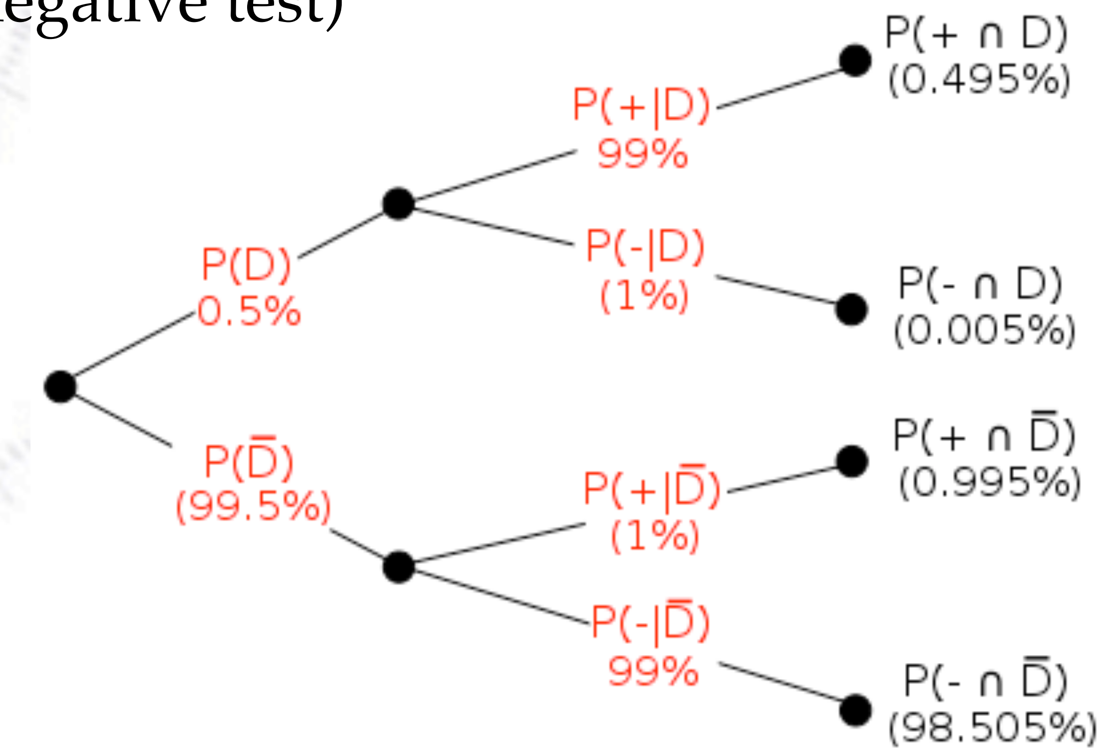
- 99% positive results for users (99% sensitive, i.e. 1% Type I errors).
- 99% negative results for non-users (99% specific, i.e. 1% Type II errors).

If 0.5% of a population is using the drug, and a random person tests positive, what is the chance that he is using the drug?

The answer is 33.2%, i.e. NOT very high! The reason is the **prior probability**. False positives (0.995%) are large compared to true positives (0.495%).

(D = user, $\overline{\text{D}}$ = non=user, + = positive test, - = negative test)

$$
\begin{aligned}
P(D|+) &= \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|\bar{D})P(\bar{D})} \\
&= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995} \\
&= 33.2\%.
\end{aligned}
$$



Bayes' Theorem

Though Thomas Bayes was elected fellow of the Royal Society, his most famous paper was published posthumously.

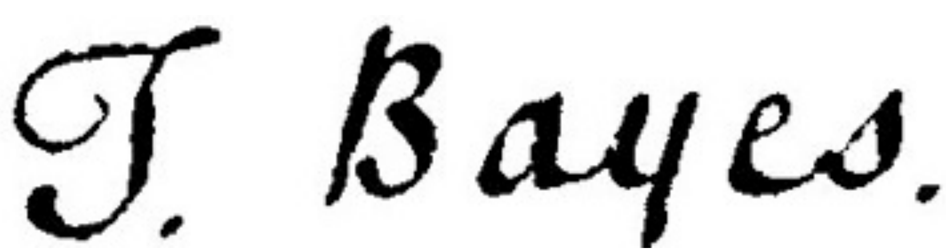
It was an attempt to answer a problem stated by **Abraham de Moivre**, and went by the name: *"Essay towards solving a Problem in the Doctrine of Chances" (1764).*

Bayes correctly realized, that in some cases one needs to know the **prior probability**.
Conditional probability (A given B) depends on the inverse (B given A).

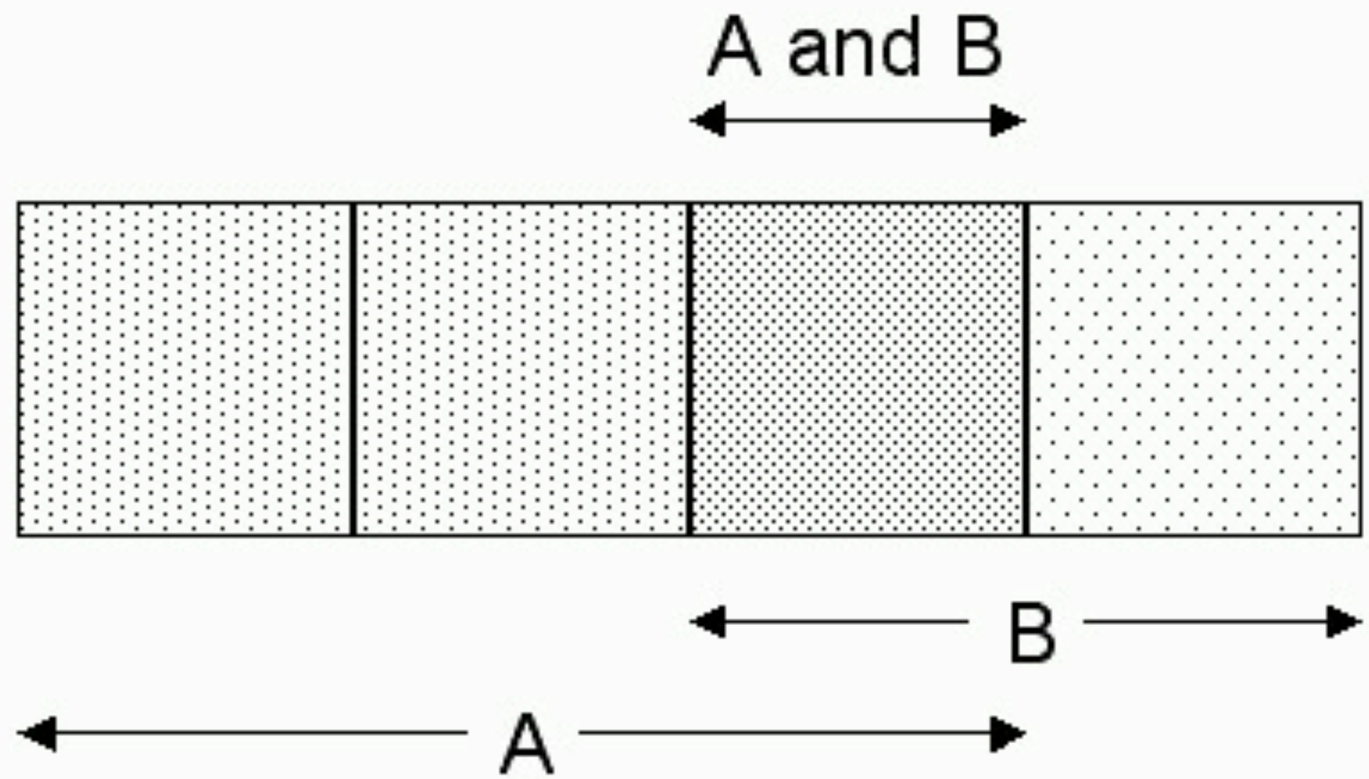
Bayes' Theorem was later proposed independently by **Pierre-Simon Laplace**, who also extended its use. Little did they know, that the theorem has founded an **interpretation of probability.**

$$P(h \mid D) = \frac{P(D \mid h)P(h)}{P(D)}$$

$P(h|D)$ = *posterior probability* of h
$P(h)$ = *prior probability* of h
$P(D|h)$ = probability of observing D given that h holds
$P(D)$ = probability of observing D



Bayes' Theorem illustrated



P(A) = 3/4

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

P(B) = 2/4

P(A and B) = P(AB) = 1/4

P(A|B) = P(AB) / P(B) = (1/4) / (2/4) = 1/2

P(B|A) = P(AB) / P(A) = (1/4) / (3/4) = 1/3

Bayes' Theorem

Reverend Thomas Bayes
(1702-1761)

$$\underbrace{P(\text{model}|\text{data}, I)}_{\text{Posterior Probability}} = \underbrace{P(\text{model}, I)}_{\text{Prior Probability}} \frac{\overbrace{P(\text{data}|\text{model}, I)}^{\text{Likelihood}}}{\underbrace{P(\text{data}, I)}_{\text{Normalizing constant}}}$$

Likelihood

describes how well the model predicts the data

Posterior Probability

represents the degree to which we believe a given **model** accurately describes the situation given the available **data** and all of our prior information **I**

Prior Probability

describes the degree to which we believe the model accurately describes reality based on all of our prior information.

Normalizing constant

Different versions...

The "original" version of Bayes' Theorem was stated as follows:

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

However, it can be expanded (using the total law of probability) to:

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

It is in this form, that Bayes' Theorem is most often used.

Interpretations

One way Bayes' Theorem is often used in normal thinking is:

$$P(\text{theory}|\text{data}) \propto P(\text{data}|\text{theory}) \cdot P(\text{theory}).$$

Here, P(data) has been omitted.

The trouble is, that it is hard to define a "degree of belief" in a theory.

Perhaps Glen Cowan sums it up best (chapter 1):

Bayesian statistics provides no fundamental rule for assigning the prior probability to a theory, but once this has been done, it says how one's degree of belief should change in the light of experimental data.