Laplace's "Bayes Theorem"

"The greater the probability of an observed event given any one of a number of causes to which that event may be attributed, the greater the likelihood of that cause {given that event}.

$$P(C_i \,|\, E) \propto P(E \,|\, C_i)$$

Laplace's "Bayes Theorem"

"The greater the probability of an observed event given any one of a number of causes to which that event may be attributed, the greater the likelihood of that cause {given that event}. The probability of the existence of any one of these causes {given the event} is thus a fraction whose numerator is the probability of the event given the cause, and whose denominator is the sum of similar probabilities, summed over all causes.

$$P(C_i \mid E) = \frac{P(E \mid C_i)}{\sum_j P(E \mid C_j)}$$

Laplace's "Bayes Theorem"

"The greater the probability of an observed event given any one of a number of causes to which that event may be attributed, the greater the likelihood of that cause {given that event}. The probability of the existence of any one of these causes {given the event} is thus a fraction whose numerator is the probability of the event given the cause, and whose denominator is the sum of similar probabilities, summed over all causes. If the various causes are not equally probable *a priory*, it is necessary, instead of the probability of the event given each cause, to use the product of this probability and the *possibility of the cause itself*."

$$P(C_i \mid E) = \frac{P(E \mid C_i)\, P(C_i)}{\sum_j P(E \mid C_j)\, P(C_j)}$$

Laplace's "Bayes Theorem"

"The greater the probability of an observed event given any one of a number of causes to which that event may be attributed, the greater the likelihood of that cause {given that event}. The probability of the existence of any one of these causes {given the event} is thus a fraction whose numerator is the probability of the event given the cause, and whose denominator is the sum of similar probabilities, summed over all causes. If the various causes are not equally probable *a priory*, it is necessary, instead of the probability of the event given each cause, to use the product of this probability and the *possibility of the cause itself*."

$$P(C_i \mid E) = \frac{P(E \mid C_i)\, P(C_i)}{P(E)}$$

(*Philosophical Essai on Probabilities*)

[In general $P(E) = \sum_j P(E \mid C_j)\, P(C_j)$ (*weighted average*, with weigths being the probabilities of the conditions) if C_j form a *complete class of hypotheses*]

Laplace's "Bayes Theorem"

$$P(C_i \mid E) \;=\; \frac{P(E \mid C_i)\, P(C_i)}{P(E)} = \frac{P(E \mid C_i)\, P(C_i)}{\sum_j P(E \mid C_j)\, P(C_j)}$$

"This is the fundamental principle $^{(*)}$ **of that branch of the analysis of chance that consists of reasoning a posteriori from events to causes"**

(*) In his "Philosophical essay" Laplace calls 'principles' the 'fundamental rules'.

Laplace's "Bayes Theorem"

$$P(C_i \,|\, E) \;=\; \frac{P(E \,|\, C_i)\, P(C_i)}{P(E)} = \frac{P(E \,|\, C_i)\, P(C_i)}{\sum_j P(E \,|\, C_j)\, P(C_j)}$$

"This is the fundamental principle $^{(*)}$ **of that branch of the analysis of chance that consists of reasoning a posteriori from events to causes"**

(*) In his "Philosophical essay" Laplace calls 'principles' the 'fundamental rules'.

Note: denominator is just a normalization factor.

$$\Rightarrow \qquad P(C_i \,|\, E) \;\propto\; P(E \,|\, C_i)\, P(C_i)$$

Laplace's "Bayes Theorem"

$$P(C_i \,|\, E) \;=\; \frac{P(E \,|\, C_i)\, P(C_i)}{P(E)} = \frac{P(E \,|\, C_i)\, P(C_i)}{\sum_j P(E \,|\, C_j)\, P(C_j)}$$

"This is the fundamental principle $^{(*)}$ of that branch of the analysis of chance that consists of reasoning a posteriori from events to causes"

(*) In his "Philosophical essay" Laplace calls 'principles' the 'fundamental rules'.

Note: denominator is just a normalization factor.

$$\Rightarrow \qquad P(C_i \,|\, E) \;\propto\; P(E \,|\, C_i)\, P(C_i)$$

Most convenient way to remember Bayes theorem

Laplace's teaching

$$\frac{P(H_0 \mid \text{data})}{P(H_1 \mid \text{data})} = \frac{P(\text{data} \mid H_0)}{P(\text{data} \mid H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- We should possibly use the data, rather then the test variables 'θ' (χ^2 etc);
  [although in some case 'sufficient summaries' do exist]

Laplace's teaching

$$\frac{P(H_0\,|\,\text{data})}{P(H_1\,|\,\text{data})} = \frac{P(\text{data}\,|\,H_0)}{P(\text{data}\,|\,H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- We should possibly use the data, rather then the test variables 'θ' (χ^2 etc);
  [although in some case 'sufficient summaries' do exist]
- At least two hypotheses are needed!

Laplace's teaching

$$\frac{P(H_0 \,|\, \text{data})}{P(H_1 \,|\, \text{data})} = \frac{P(\text{data} \,|\, H_0)}{P(\text{data} \,|\, H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- We should possibly use the data, rather then the test variables 'θ' (χ^2 etc);
  [although in some case 'sufficient summaries' do exist]
- At least two hypotheses are needed!
- ...and also how they appear belivable *a priori*!

Laplace's teaching

$$\frac{P(H_0\,|\,\text{data})}{P(H_1\,|\,\text{data})} \;=\; \frac{P(\text{data}\,|\,H_0)}{P(\text{data}\,|\,H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- We should possibly use the data, rather then the test variables 'θ' (χ^2 etc);
  [although in some case 'sufficient summaries' do exist]
- At least two hypotheses are needed!
- ... and also how they appear belivable *a priori*!
- If $P(\text{data}\,|\,H_i) = 0$, it follows $P(H_i\,|\,\text{data}) = 0$:
  $\Rightarrow$ falsification (the 'serious' one) is a corollary of the theorem, rather than a principle.

Laplace's teaching

$$\frac{P(H_0 \,|\, \text{data})}{P(H_1 \,|\, \text{data})} = \frac{P(\text{data} \,|\, H_0)}{P(\text{data} \,|\, H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- We should possibly use the data, rather then the test variables 'θ' (χ^2 etc);
  [although in some case 'sufficient summaries' do exist]
- At least two hypotheses are needed!
- ...and also how they appear belivable *a priori*!
- If $P(\text{data} \,|\, H_i) = 0$, it follows $P(H_i \,|\, \text{data}) = 0$:
  $\Rightarrow$ falsification (the 'serious' one) is a corollary of the theorem, rather than a principle.
- There is no conceptual problem with the fact that $P(\text{data} \,|\, H_1) \rightarrow 0$ (e.g. 10^{-37}), provided the ratio $P(\text{data} \,|\, H_0)/P(\text{data} \,|\, H_1)$ is not undefined.

Bayes factor ('likelihood ratio')

$$\frac{P(H_0 \,|\, \text{data})}{P(H_1 \,|\, \text{data})} \quad = \quad \frac{P(\text{data} \,|\, H_0)}{P(\text{data} \,|\, H_1)} \times \frac{P(H_0)}{P(H_1)}$$

Bayes factor ('likelihood ratio')

$$\frac{P(H_0 \,|\, \text{data})}{P(H_1 \,|\, \text{data})} \;=\; \frac{P(\text{data} \,|\, H_0)}{P(\text{data} \,|\, H_1)} \times \frac{P(H_0)}{P(H_1)}$$

$$\text{Prob. ratio}\Big|_{\textit{posterior}} \;=\; \text{Bayes factor} \times \text{Prob. ratio}\Big|_{\textit{prior}}$$

(*prior*/*posterior* w.r.t. data)

Bayes factor ('likelihood ratio')

$$\frac{P(H_0\,|\,\text{data})}{P(H_1\,|\,\text{data})} \;=\; \frac{P(\text{data}\,|\,H_0)}{P(\text{data}\,|\,H_1)} \times \frac{P(H_0)}{P(H_1)}$$

$$\text{Prob. ratio}\big|_{posterior} \;=\; \text{Bayes factor} \times \text{Prob. ratio}\big|_{prior}$$

(*prior*/*posterior* w.r.t. data)

If H_0 and H_1 are 'complementary', that is $H_1 = \overline{H}_0$, then

posterior odds = Bayes factor × prior odds