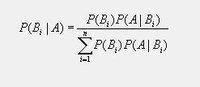
学过概率理论的人都知道条件概率的公式：P(AB)=P(A)P(B|A)=P(B)P(A|B)；即事件A和事件B同时发生的概率等于在发生A的条件下B发生的概率乘以A的概率。由条件概率公式推导出贝叶斯公式：P(B|A)=P(A|B)P(B)/P(A)；即,已知P(A|B)，P(A)和P(B)可以计算出P(B|A)。

贝叶斯公式：



在贝叶斯法则中，每个名词都有约定俗成的名称：

Pr(A)是A的先验概率或边缘概率。之所以称为"先验"是因为它不考虑任何B方面的因素。

Pr(A|B)是已知B发生后A的条件概率，也由于得自B的取值而被称作A的后验概率。

Pr(B|A)是已知A发生后B的条件概率，也由于得自A的取值而被称作B的后验概率。

Pr(B)是B的先验概率或边缘概率，也作标准化常量（normalized constant）。

按这些术语，Bayes法则可表述为：

后验概率 = (似然度 \* 先验概率)/标准化常量 也就是说，后验概率与先验概率和似然度的乘积成正比。

另外，比例Pr(B|A)/Pr(B)也有时被称作标准似然度（standardised likelihood），Bayes法则可表述为：

后验概率 = 标准似然度 \* 先验概率

要理解贝叶斯推断，必须先理解贝叶斯定理。后者实际上就是计算"条件概率"的公式。

所谓"条件概率"（Conditional probability），就是指在事件B发生的情况下，事件A发生的概率，用P(A|B)来表示。

全概率公式是问总体看来被偷的概率是多少，贝叶斯公式是知道了总体被偷了这件事，概率并不知道，问你个更有意思的问题，像是侦探断案：是哪个小偷的偷的，计算每个小偷偷的概率。贝叶斯公式有意思极了，简单说就是逆全概公式

例子

假设有一所学校，学生中60%是男生和40%是女生。女生穿裤子与裙子的数量相同；所有男生穿裤子。现在有一个观察者，随机从远处看到一名学生，因为很远，观察者只能看到该学生穿的是裤子，但不能从长相发型等其他方面推断被观察者的性别。那么该学生是女生的概率是多少？

用事件 G 表示观察到的学生是女生，用事件 T 表示观察到的学生穿裤子。于是，现在要计算的是条件概率 P(G|T) ，我们需要知道：

* P(G) 表示一个学生是女生的概率。由于观察者随机看到一名学生，意味着所有的学生都可能被看到，女生在全体学生中的占比是 40% ，所以概率是 P(G)=0.4 。注意，这是在没有任何其他信息下的概率。这也就是先验概率。后面我们还会详细讨论。
* P(B) 是学生不是女生的概率，也就是学生是男生的概率，这同样也是指在没有其他任何信息的情况下，学生是男生的先验概率。 B 事件是 G 事件的互补的事件，于是易得 P(B)=0.6 。
* P(T|G) 是在女生中穿裤子的概率，根据题目描述，女生穿裙子和穿裤子各占一半，所以 P(T|G)=0.5 。这也就是在给定 G 的条件下，T 事件的概率。
* P(T|B) 是在男生中穿裤子的概率，这个值是1。
* P(T) 是学生穿裤子的概率，即任意选一个学生，在没有其他信息的情况下，该名学生穿裤子的概率。根据全概率公式 P(T)=∑ni=1P(T|Ai)P(Ai)=P(T|G)P(G)+P(T|B)P(B) ，计算得到 P(T)=0.5×0.4+1×0.6=0.8。

根据贝叶斯公式

P(Ai|T)=P(T|Ai)P(Ai)∑ni=1P(T|Ai)P(Ai)=P(T|Ai)P(Ai)P(T)

基于以上所有信息，如果观察到一个穿裤子的学生，并且是女生的概率是 

P(G|T)=P(T|G)P(G)P(T)=0.5×0.4÷0.8=0.25.