

贝叶斯认为参数是随机变量，而样本 X 是固定的，由于样本是固定的，所以重点研究的是参数的分布。

我们在抛硬币的时候，结果往往只有正面，背面两种结果，符合0-1分布，那么这种就被称为先验分布，而后验分布往往是给定样本 X 的情况下 θ 的条件分布，使得后验分布的 θ 值最大的被称为最大后验估计，类似经典统计学中的极大似然估计。

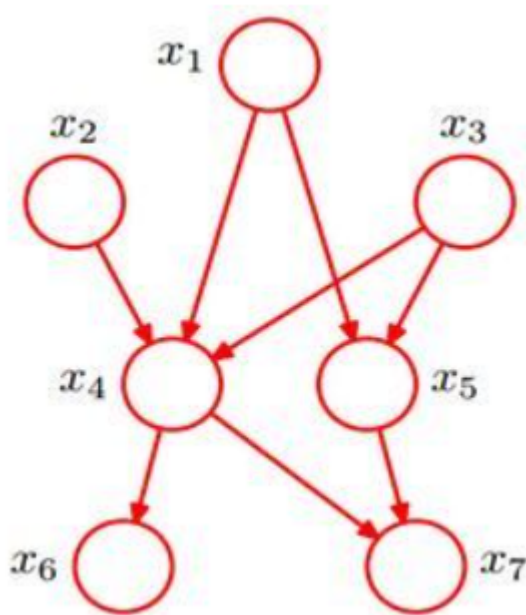
综合起来看，则好比是人类刚开始时对大自然只有少得可怜的先验知识，但随着不断观察、实验获得更多的样本、结果，使得人们对自然界的规律摸得越来越透彻。所以，贝叶斯方法既符合人们日常生活的思考方式，也符合人们认识自然的规律，经过不断的发展，最终占据统计学领域的半壁江山，与经典统计学分庭抗礼。

贝叶斯定理：

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

贝叶斯网络：

贝叶斯网络(Bayesian network)，又称信念网络(Belief Network)，或有向无环图模型(directed acyclic graphical model)，是一种概率图模型，于1985年由Judea Pearl首先提出。它是一种模拟人类推理过程中因果关系的不确定性处理模型，其网络拓朴结构是一个有向无环图(DAG)。



上述的每一个节点都可以以贝叶斯公式的形式写出。

$$P(x_1, \dots, x_k) = P(x_k | x_1, \dots, x_{k-1}) \dots P(x_2 | x_1) P(x_1)$$

特殊情况：



只与前一项有关，又被称为马尔可夫链(Markov chain)

朴素贝叶斯:

朴素贝叶斯(Naive Bayesian)是经典的机器学习算法之一，也是为数不多的基于概率论的分类算法。

简单来讲就是一个生成模型，利用已知的分布，来估计新数据的概率，哪个概率大就对应哪个特征。

朴素贝叶斯缺点:

理论上，朴素贝叶斯模型与其他分类方法相比具有最小的误差率。但是实际上并非总是如此，这是因为朴素贝叶斯模型假设属性之间相互独立，这个假设在实际应用中往往是不成立的，在属性个数比较多或者属性之间相关性较大时，分类效果不好。