

理解贝叶斯先验与后验

名词“先验”和“后验”来自于西方中世纪逻辑区分的两种论证，从原因到结果的论证称为“先验的”，而从结果到原因的论证称为“后验的”。

先验概率是指根据以往的经验、资料或者常识分析得到的概率，它往往作为“由因求果”问题中的“因”出现；**后验概率是指在得到“结果”的信息后重新修正的概率**，即基于新的观测信息，修正原来的先验概率后获得的更接近实际情况的概率估计，往往作为“执果寻因”问题中的“因”出现。

在概率论中，先验概率和后验概率是和贝叶斯公式联系在一起的。在贝叶斯公式

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

中， $P(A)$ 就是先验 (prior) 概率， $P(A|B)$ 是后验 (posterior) 概率， $P(B|A)$ 被称为类条件概率 (class-conditional probability) 或似然 (likelihood) 。

通俗的说，先验概率就是统计概率，而后验概率是条件概率。

用“瓜熟蒂落”这个词语举个例子：

- **先验概率**：在没有任何条件的情况下，根据经验和常识得出的瓜熟的概率；
- **后验概率**：在观察到瓜蒂脱落（结果）的情况下，瓜熟（原因）的概率。

注意上面的因果关系，是因为瓜成熟（原因）了才导致瓜蒂脱落（结果），后验概率就是根据瓜蒂脱落这个结果去推断瓜成熟的概率。

用贝叶斯公式来表达，就是：

$$P(\text{瓜熟} | \text{已知蒂落}) = P(\text{瓜熟}) \times P(\text{蒂落} | \text{瓜熟}) / P(\text{蒂落})$$

这其中，**似然函数的意义就是根据已知结果去推测由某种原因（或者固有性质）导致的可能性**，它代表的是对事物固有性质的拟合程度，而没有具体的概率意义。在上面的例子中，似然函数关心的不是瓜熟的概率，而是瓜熟与蒂落之间的关系。

贝叶斯方法在机器学习中十分重要，因为很多问题都需要计算机在已知条件下进行最佳决策，它实际上是在模仿人脑的决策过程。贝叶斯公式就是人脑根据以往的经验，结合已知条件，做出直觉判断的过程的一种数学表示。