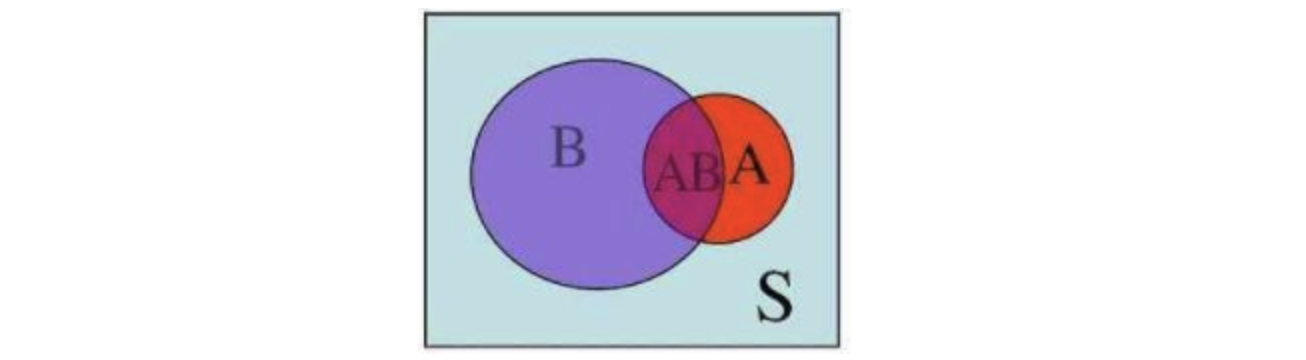


贝叶斯



**1、概率**

**2、全概率**

**3、贝叶斯推断**

称为**先验概率(Prior probability),**即在B时间发生之前我们对A事件概率的一个判断；

称为**后验概率(Posterior probability),** 即在B事件发生之后，我们对A事件概率的重新判断；

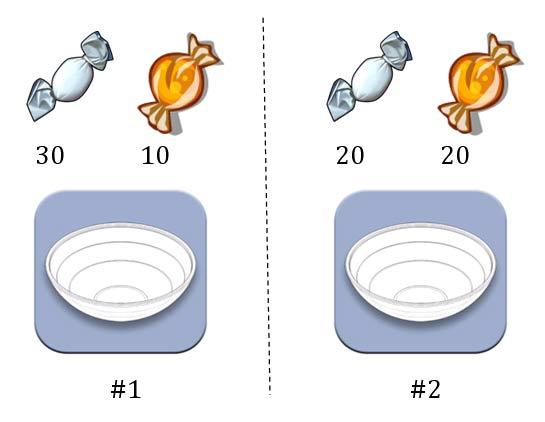
称为**“可能性函数”（Likelyhood），**这是一个调整因子，使得预估概率更加接近真实概率。

条件概率可以写成

**后验概率 = 先验概率 \* 调整因子**

这就是贝叶斯推断的含义。我们先预估一个"先验概率"，然后加入实验结果，看这个实验到底是增强还是削弱了"先验概率"，由此得到更接近事实的"后验概率"。

在这里，如果"可能性函数"P(B|A)/P(B)>1，意味着"先验概率"被增强，事件A的发生的可能性变大；如果"可能性函数"=1，意味着B事件无助于判断事件A的可能性；如果"可能性函数"<1，意味着"先验概率"被削弱，事件A的可能性变小。



两个一模一样的碗，一号碗有30颗水果糖和10颗巧克力糖，二号碗有水果糖和巧克力糖各20颗。现在随机选择一个碗，从中摸出一颗糖，发现是水果糖。请问这颗水果糖来自一号碗的概率有多大？

我们假定，表示一号碗，表示二号碗。由于这两个碗是一样的，所以，也就是说，在取出水果糖之前，这两个碗被选中的概率相同。因此，，我们把这个概率就叫做"先验概率"，即没有做实验之前，来自一号碗的概率是0.5。

再假定，E表示水果糖，所以问题就变成了在已知E的情况下，来自一号碗的概率有多大，即求。我们把这个概率叫做"后验概率"，即在E事件发生之后，对的修正。

根据条件概率公式，得到

已知，等于0.5，为一号碗中取出水果糖的概率，等于30÷(30+10)=0.75，那么求出就可以得到答案。根据全概率公式，

所以，

将数字代入原方程，得到

这表明，来自一号碗的概率是0.6。也就是说，取出水果糖之后，事件的可能性得到了增强。

同时再思考一个问题，在使用该算法的时候，如果不需要知道具体的类别概率，即上面，只需要知道所属类别，即来自一号碗，我们有必要计算这个全概率吗？要知道我们只需要比较 和的大小，找到那个最大的概率就可以。既然如此，两者的分母都是相同的，那我们只需要比较分子即可。即比较和的大小，所以为了减少计算量，**全概率公式在实际编程中可以不使用**。

**4、朴素贝叶斯**

朴素贝叶斯（naive Bayes classifiers）是一种分类器，在机器学习中有着广泛的应用。相信很多人知道贝叶斯定理,即

所以，当我们有一组事件， 我们想通过这组事件去估计一个事件发生的概率，比如我们想估计水果的种类，如果有一组事件分别 黄，长，弯...那我们就可以判断这是一个香蕉。 尽管黄，长，弯几个事件之间可能会相互依赖，但是在朴素贝叶斯模型中，我们假设它们相互独立，这就是他的朴素之处。值得注意的是，这里的朴素对应英语单词naive，单纯的意思，所以也可以理解为很天真单纯的估计（把数据中的每个特征看作独立分布）。

OK，现在我们具体看一下朴素贝叶斯的概率模型，其实就是一个条件概率的模型，在发生的条件下去估计事件C的概率，即计算

根据贝叶斯定理，我们有

我们可以看到, 如果每个特征的取值至少有两个，然后有100个特征，那给定一组数据后，想要计算是几乎不可能的，因为你的联合分布中有种可能，这样计算机是无法扫描完所有的概率空间的，即使可以，在会出现大量的0（即不存在的某种组合）。为了解决这个问题，我们假设所有特征F之间相互独立，这样一来，我们的等式，即可写成

这样即可算出我们想要的结果。

**5、朴素贝叶斯改进之拉普拉斯平滑**

朴素贝叶斯在落实的时候常常需要计算多个条件概率的乘积,由于训练样本量的限制，往往会出现其中之一的概率值为0，这就导致后验概率为0，最后生成的概率为0无法进行比较。详细过程见程序。