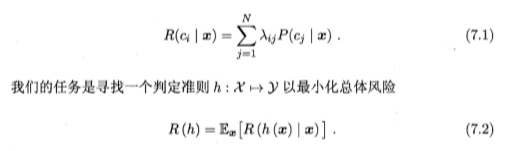
联合概率：联合概率表示为包含多个条件并且所有的条件都同时成立的概率，记作P(X=a,Y=b)P(X=a,Y=b)P(X=a,Y=b) 或 P(a,b)，有的书上也习惯记作P(ab)

大数定理：大数定律(law of large numbers)，是一种描述当试验次数很大时所呈现的概率性质的定律。但是注意到，大数定律并不是经验规律，而是在一些附加条件上经严格证明了的定理，它是一种自然规律因而通常不叫定理而是大数“定律”。而我们说的大数定理通常是经数学家证明并以数学家名字命名的大数定理，如伯努利大数定理。

1.贝叶斯决策论

贝叶斯决策论是概率背景下的实施决策的方法，在所有相关概率已知情况下，贝叶斯决策论考虑基于这些概率的误判损失来选择最优分类标记

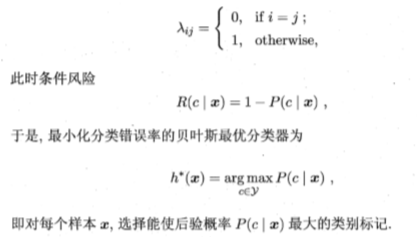
贝叶斯决策用一个叫条件风险的东西作为一个类损失函数，任务就是找到一个判定法则，这个东西判断出来的结果可以使得总体的条件风险最小，那么贝叶斯判断准则中心思想就是使每个样本上的判定风险最小，从而使得整体判定风险最小



那么这个时候我们将会训练出贝叶斯最优分类器



具体来说若目标为最小化分类错误率，则误判损失可写为



不难看出要用贝叶斯决策使得决策风险最小，就要知道后验概率，在现实中这很难获得，从这个角度我们可以把机器学习看成基于有限样本来估计后验概率，主要有两者策略：对于给定x直接建模P（c|x）来预测c,这样是判别式模型；

也可对联合概率建模，再由此得到P（c|x）这是生成式模型，

对于前面的决策树，BP神经网络，支持向量机都是判别式模型

对于生成式模型：

