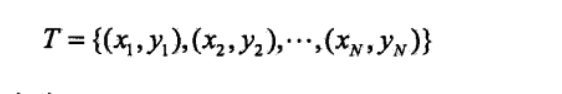
朴素贝叶斯(naive Bayes)法是基于贝叶斯定理和特征条件独立假设的分类方法。对于给定的训练数据集，首先基于特征条件独立假设学习输入/输出的联合概率；然后基于此模型，对给定的输入x，利用贝叶斯定理求出后验概率最大的输出y。

# 1.朴素贝叶斯法的学习与分类

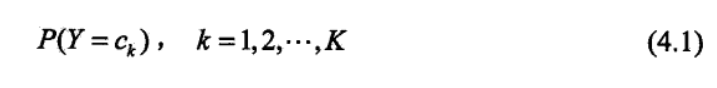
## 1.1 学习

给定输入集合X，输出集合Y，P(X,Y)是输入X和输出Y的联合概率分布。训练数据集：

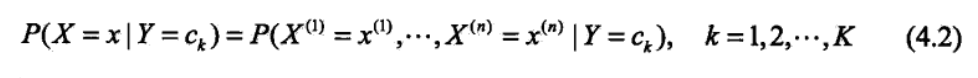


由P(X,Y)独立同分布产生。

朴素贝叶斯法通过训练数据集学习联合概率分布P(X,Y)。具体就是学习先验概率分布:

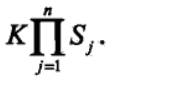


和条件概率分布：

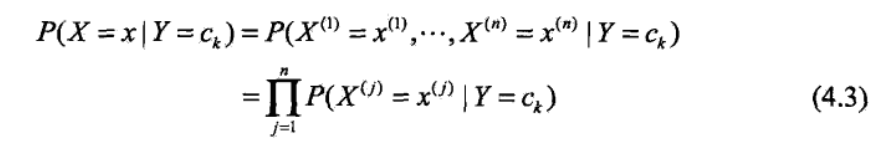


然后通过先验概率分布和条件概率分布学习到联合概率分布P(X,Y)。

条件概率分布P(X=x|Y=Ck)的参数数量是指数量级的，对其进行估计实际上并不可行。假设x(j)可取值有Sj个,j=1,2,...,n,Y可取值有K个，那么参数数量为：



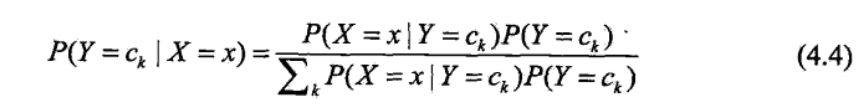
朴素贝叶斯法对条件概率分布作了条件独立性假设，即：



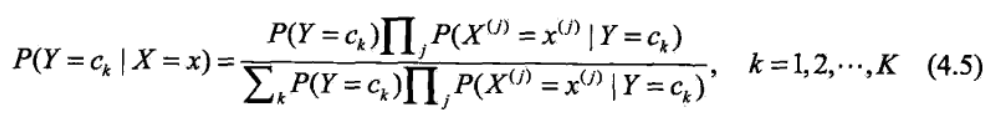
也就是说用于分类的特征再类确定的条件下是条件独立的，这一假设虽然会使得分类变得简单，但会牺牲一定的分类准确率。

## 1.2 分类

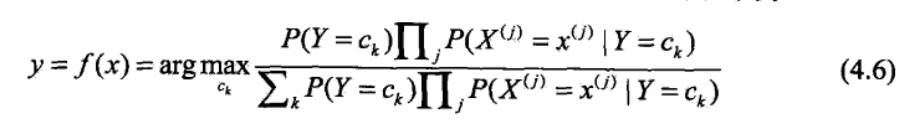
朴素贝叶斯法在分类时，对给定的输入x，通过学习到的模型计算后验概率分布P(Y=Ck|X=x)，将后验概率最大的类作为x的类输出，后验概率通过贝叶斯定理计算得出：



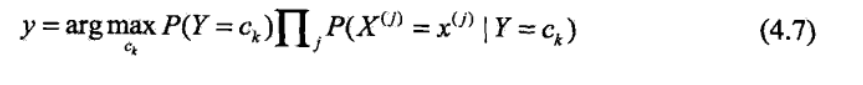
将4.3式代入4.4有：



于是朴素贝叶斯分类器可表示为：

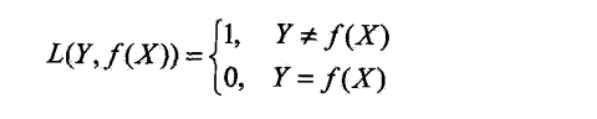


上式中，分母对所有Ck都是相同的，所以：

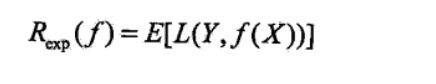


## 1.3 后验概率最大化

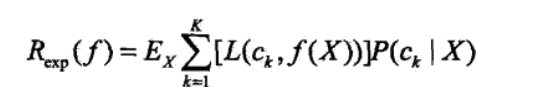
朴素贝叶斯法将实例分到后验概率最大的类中，等价于期望风险最小化。假设选择0-1损失函数：



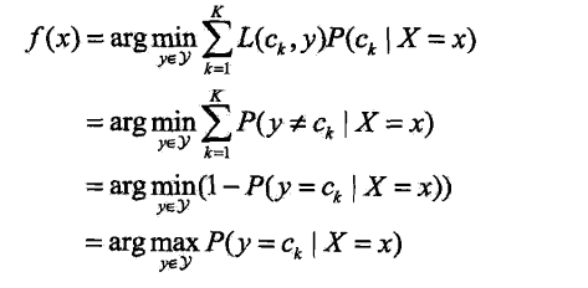
式中f(X)是分类决策函数。这时，期望风险函数为：



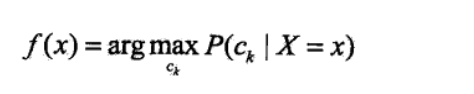
该期望是对联合分布P(X,Y)取的，由此取条件期望：



对X=x逐个极小化可以使得期望风险最小化：



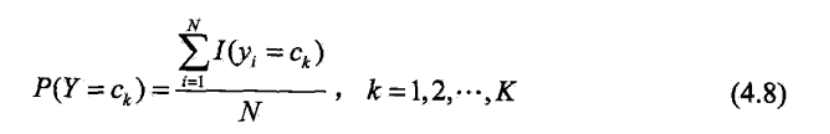
从上式可知，根据期望风险最小化准则可得后验概率最大化准则：



# 2.朴素贝叶斯法的参数估计

## 2.1 极大似然估计

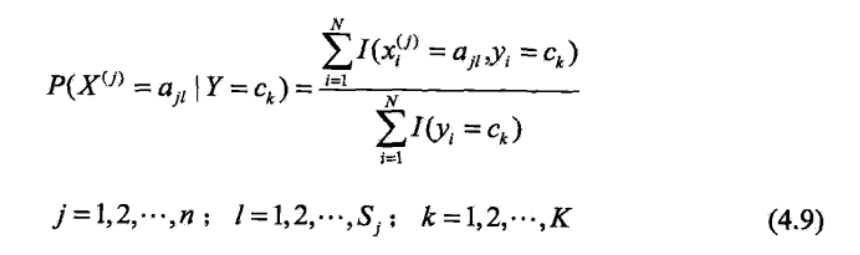
朴素贝叶斯法就是估计P(Y=Ck)和P(X(j)|Y=Ck)，从而求得联合概率分布P(X,Y),然后计算后验概率。先验概率P(Y=Ck)的极大似然估计是：



设第j个特征x(j)可能取值的集合为



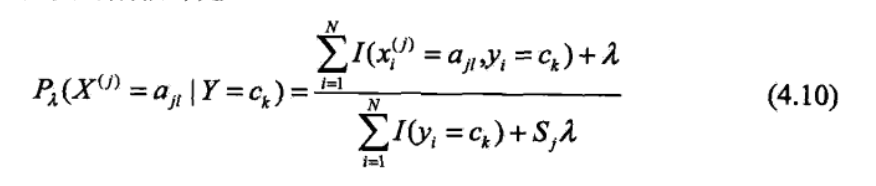
则条件概率P(X(j)|Y=Ck)的极大似然估计是：



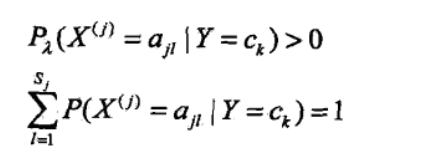
式中，x(j)i是第i个样本的第j个特征；aij是第j个特征可能取得第l个值；I为指示函数。

## 2.2 贝叶斯估计

条件概率得贝叶斯估计是：



式中λ≥0，等价于在特征得各个取值得频数上赋予一个整数λ>0,λ=0时就是极大似然估计，λ=1时称为拉普拉斯平滑。对于任何l=1,2,...,Sj,k=1,2,...K,有：



同样先验概率得贝叶斯估计为：

