第七章 贝叶斯分类器

By Xian2207, 13689903575, wszhangxian@126.com

7.1 贝叶斯理论

原理：基于概率和判断损失来选择最优的类别。公式如下

$P(c|x) = \frac{P(c) \cdot P(x|c)}{P(x)}$ 7.7

其中$P(c)$是"先验"（prior）概率；$P(x|c)$是样本 x 相对于类标记 c 的类条件概率（class-conditional probability）或"似然"（likelihood）; $P(x)$是用来归一化的"证据"（evidence）因子，$P(c|x)$是后验概率（post-prior）。对给定样本 x，证据因子$P(x)$与类标记无关，因此问题转化为，估计$P(c|x)$即如何基于训练数据 D 来估计先验概率$P(c)$和条件概率$P(x|c)$。对条件概率而言，c 指属性，假设样本的 d 个属性都是二值或三值的，如西瓜数据的颜色属性（青绿，乌黑，浅白），硬度（软滑，硬滑），则样本空间 d 有2^d或3^d个可能的取值，由于训练样本数 m 一定小于等于 d，导致$P(x|c)$的取值肯定小于等于 d，就有2^d – d或$3^d - d$个取值没有在训练集中出现，这样估算的概率$P(x|c)$肯定不行，因为"未被观测到"与"出现概率为0"是不同的概念。

7.2 极大似然（条件概率）估计

为解决条件概率取值疏漏问题，假设$P(x|c)$具有确定形式且被参数向量θ_c唯一确定，我们的任务转变为利用训练集 D 估计参数θ_c. 采用统计学派中的频率学派（另一支为贝叶斯学派）的极大似然估计（maximum likelihood estimation, MLE）法，即采用数据采样来估计概率分布参数。令 Dc 表示训练集 D 中第 c 类样本组成的集合，假设样本是独立分布的，则似然满足

$P(D_c|\theta_c) \prod_{x \in D_c}(P(x|\theta_c))$ 7.9

符号$\prod$是指连续乘积，实际操作时容易溢出，通常采用对数似然（log-likelihood）

$log[P(D_c|\theta_c)] = log[\prod P(x|\theta_c)]$ 7.10

$log[P(D_c|\theta_c)] = \sum_{x \in D_c} logP(x|\theta_c)$ 7.11

记极大似然估计为$\hat{\theta}_c$，假设概率密度函数$p(x|c) \sim N(\mu_c, \sigma_c^2)$，则参数$\mu_c$和$\sigma_c^2$的极大似然估计为

$\hat{\mu}_c = \frac{1}{D_c} \sum_{x \in D_c} x$ 7.12

$\hat{\sigma}_c^2 = \frac{1}{D_c} \sum_{x \in D_c} (x - \hat{\mu}_c)^T (x - \hat{\mu}_c)$ 7.13

把$logP(x|\theta_c)$差分开，可理解为 m1+m2+m3,相当于整体为$\sum_{x \in D_c} x$。由于通过极大似然法得到的正太分布恰好可以写成上式，而上式为正态分布的样本均值，方差即为$(x - \hat{\mu}_c)^T(x -$

$\hat{\mu}_c$)的均值。，在离散属性下，根据样本之间概率独立，恰好可以通过此估计最大似然。

7.3 朴素贝叶斯分类器

根据属性独立假设，朴素贝叶斯分类器公式为

$P(c|x) = \frac{P(c) \cdot P(x|c)}{P(x)} = \frac{P(c)}{P(x)} \cdot \prod_{i=1}^{d} P(x_i|c)$　　7.14

对所有类别来说 P(x)是相同的，故贝叶斯判定准则是

$\max\{P(c) \cdot \prod_{i=1}^{d} P(x_i|c)\}$　　7.15

基于贝叶斯分类器，估计$P(c|x)$的主要困难在于估计似然$P(x|c)$上所有属性的联合概率。令D_c表示训练集 D 中第 c 类样本组成的几何，若有充足独立分布的数据样本，则类先验概率为

$P(c) = \frac{D_c}{D}$　　7.16

对离散属性而言，令D_{c,x_i}表示D_c中在第 i 个属性上取值为x_i的样本组成的集合，则似然为

$P(x_i|c) = \frac{D_{c,x_i}}{D_c}$　　7.17

对连续属性可考虑概率密度函数，假定$p(x_i|c) \sim N(\mu_{c,i}, \sigma_{c,i}^2)$，其中$\mu_{c,i}$和$\sigma_{c,i}^2$分别是第 c 类样本在第 i 个属性上取值的均值和方差，则

$p(x_i|c) = \frac{1}{\sqrt{2\pi\sigma_{c,i}}} \exp(-\frac{(x_i-\mu_{c,i})^2}{2\sigma_{c,i}^2})$　　7.18

若某个属性值在训练集中没有与某个类同时出现，基于公式 7.17 概率估计，再根据 7.15 判断会出现问题，如概率值等于 0 的情况。因此，无论该样本其他属性如何，哪怕其他属性明显是好瓜，分类结果都是“好瓜=否”，这显然不合理。为避免其他属性被训练集中未出现的属性值抹去，在估计概率值时要进行“平滑”smooth，常用“拉普拉斯修正”（Laplacian Correction）。令 N 表示 D 中可能的类别数，Ni 表示第 i 个属性可能的取值，则 7.16 和 7.17 修正为

$\hat{P}(c) = \frac{D_c+1}{D+N}$　　7.19

$\hat{P}(x_i|c) = \frac{D_{c,x_i}+1}{D+N_i}$　　7.20

以上修正过程在训练集变大时，修正引入的误差可忽略不计，使估计值趋向于实际概率值。现实任务中，朴素贝叶斯可将概率值都存储起来，预测时仅需查表判断。若任务频繁更替，新样本增加，则对新增样本属性所涉及的概率值进行计数修正即可实现增量学习。

7.4 半朴素贝叶斯分类器

为降低估计公式 7.8 中的后验概率$P(c|x)$的难度，朴素贝叶斯分类器采用属性条件独立性假设，这在现实任务中很难成立。后人们尝试对属性独立假设进行一定程度放松，产生半朴素贝叶斯分类器（semi-naïve Bayes classifier）。基本思想是适当考虑一部分属性的相互依赖性，但同时仍旧假设大部分属性独立分布。“独依赖估计”（One-Dependent Estimator, ODE）是半

朴素贝叶斯分类器最常用的一种策略。所谓“独依赖”就是假设每个属性在类别之外最多依赖一个其他属性，即

$P(c|x) \propto P(c) \prod_{i=1}^{d} P(x_i|c, pa_i)$ 7.21

其中pa_i为属性x_i所依赖的属性，称为x_i的父属性。此时，对每个属性x_i，若父属性pa_i已知，可采用类似 7.20 的办法计算似然$P(x_i|c, pa_i)$。问题关键从而转化为如何确定每个属性的父属性。常用方法如下

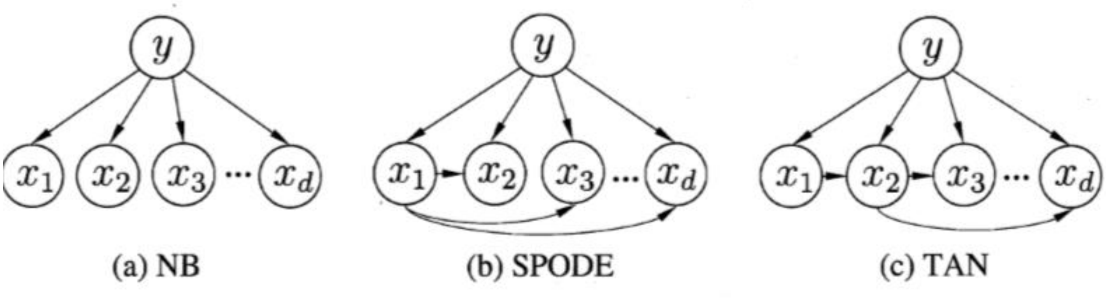

图 7.1　朴素贝叶斯与两种半朴素贝叶斯分类器所考虑的属性依赖关系

上图第一个是：Naïve Bayes, NB；第二个是 Super-Parent One-Dependent Estimator, SPODE, Tree Augmented Naïve Bayes, TAN。其他衍生的还有 Averaged One-Dpendent Estimator, AODE。

7.5 贝叶斯网

详情参考周志华书 P157 即可。主要涉及贝叶斯网的结构和学习。

7.5.1 结构

7.5.2 学习

7.6 EM 算法

以上均假设属性没有缺失，如果缺失了，则需使用“Expectation-Maixumization”EM 算法来做贝叶斯分类。将缺失变量即因数据未观测到但本应存在的变量称为“隐变量”（latent variable）。令 X 表示已观测变量集，Z 表示隐变量集，M 为模型参数。若对 M 做最大似然估计，用下列公式

$$LL(\Theta \mid \mathbf{X}, \mathbf{Z}) = \ln P(\mathbf{X}, \mathbf{Z} \mid \Theta) . \qquad (7.34)$$

$$LL(\Theta \mid \mathbf{X}) = \ln P(\mathbf{X} \mid \Theta) = \ln \sum\nolimits_{\mathbf{Z}} P(\mathbf{X}, \mathbf{Z} \mid \Theta) . \qquad (7.35)$$

细节略，非重点。

7.7 实例

假设西瓜数据 3.0 训练集如下，根据此训练集求编号为 1 的瓜是好瓜还是坏瓜。

编号,	色泽,	根蒂,	敲声,	纹理,	脐部,	触感,	密度,	含糖率,	好瓜
1,	青绿,	蜷缩,	浊响,	清晰,	凹陷,	硬滑,	0.697,	0.46,	是
2,	乌黑,	蜷缩,	沉闷,	清晰,	凹陷,	硬滑,	0.774,	0.376,	是
3,	乌黑,	蜷缩,	浊响,	清晰,	凹陷,	硬滑,	0.634,	0.264,	是
4,	青绿,	蜷缩,	沉闷,	清晰,	凹陷,	硬滑,	0.608,	0.318,	是
5,	浅白,	蜷缩,	浊响,	清晰,	凹陷,	硬滑,	0.556,	0.215,	是
6,	青绿,	稍蜷,	浊响,	清晰,	稍凹,	软粘,	0.403,	0.237,	是
7,	乌黑,	稍蜷,	浊响,	稍糊,	稍凹,	软粘,	0.481,	0.149,	是
8,	乌黑,	稍蜷,	浊响,	清晰,	稍凹,	硬滑,	0.437,	0.211,	是
9,	乌黑,	稍蜷,	沉闷,	稍糊,	稍凹,	硬滑,	0.666,	0.091,	否
10,	青绿,	硬挺,	清脆,	清晰,	平坦,	软粘,	0.243,	0.267,	否
11,	浅白,	硬挺,	清脆,	模糊,	平坦,	硬滑,	0.245,	0.057,	否
12,	浅白,	蜷缩,	浊响,	模糊,	平坦,	软粘,	0.343,	0.099,	否
13,	青绿,	稍蜷,	浊响,	稍糊,	凹陷,	硬滑,	0.639,	0.161,	否
14,	浅白,	稍蜷,	沉闷,	稍糊,	凹陷,	硬滑,	0.657,	0.198,	否
15,	乌黑,	稍蜷,	浊响,	清晰,	稍凹,	软粘,	0.36,	0.37,	否
16,	浅白,	蜷缩,	浊响,	模糊,	平坦,	硬滑,	0.593,	0.042,	否
17,	青绿,	蜷缩,	沉闷,	稍糊,	稍凹,	硬滑,	0.719,	0.103,	否

根根据贝叶斯分类公式

$$P(c|x)=\frac{P(c)\cdot P(x|c)}{P(x)}=\frac{P(c)}{P(x)}\cdot\prod_{i=1}^{d}P(x_i|c)$$

7.7.1 先验概率 prior $P(c)$:

P（好瓜=是）=8/17= 0.471；
P（坏瓜=否）=9/17-0.529；

7.7.2 证据因子（evidence）$P(x)$:

$P(x)=1/17$

7.7.3 编号 1 每个属性的似然估计$P(x_i|c)$:

7.7.3.1 离散属性

$$P(x_i|c)=\frac{D_{c,x_i}}{D_c} \qquad 7.17$$

P（青绿|是）=P（色泽=亲绿|好瓜=是）=3/8=0.375；
P（青绿|否）=P（色泽=亲绿|好瓜=否）=3/9=0.333；
P（蜷缩|是）=P（根蒂=蜷缩|好瓜=是）=5/8=0.625；

P（蜷缩|否）=P（根蒂=蜷缩|好瓜=否）=3/9=0.333；
P（浊响|是）=P（敲声=浊响|好瓜=是）=6/8=0.750；
P（浊响|否）=P（敲声=浊响|好瓜=否）=4/9=0.444；
P（清晰|是）=P（纹理=清晰|好瓜=是）=7/8=0.875；
P（清晰|否）=P（纹理=清晰|好瓜=否）=2/9=0.222；
P（凹陷|是）=P（脐部=凹陷|好瓜=是）=6/8=0.750；
P（凹陷|否）=P（脐部=凹陷|好瓜=否）=2/9=0.222；
P（硬滑|是）=P（触感=硬滑|好瓜=是）=6/8=0.750；
P（硬滑|否）=P（触感=硬滑|好瓜=否）=6/9=0.667；

7.7.3.2 连续属性

（1）朴素贝叶斯分类器：

$p(x_i|c)=\frac{1}{\sqrt{2\pi\sigma_{c,i}}}\exp(-\frac{(x_i-\mu_{c,i})^2}{2\sigma_{c,i}^2})$ 7.18

（1）P_密度：0.697|是=P（密度=0.697|好瓜=是）

均值$\mu_{c,i}$

$\mu_{c,i}$=0.697+0.774+0.634+0.608+0.556+0.403+0.481+0.437/8=0.574;

方差$\sigma_{c,i}$

$\sigma_{c,i}$=
[(0.697-0.574)^2+(0.774-0.574)^2+(0.634-0.574)^2+(0.608-0.574)^2+(0.556-0.574)^2+(0.403-0.574)^2+(0.481-0.574)^2+(0.437-0.574)^2]/7=0.129

（注意这里的 7，是 n-1，而非 n，故不能是 8）

$\sigma_{c,i}^2$=0.0146；

密度属性极大似然估计

$$p(0.697|是)=\frac{1}{\sqrt{2\pi*0.129}}\exp\left(-\frac{(0.694-0.574)^2}{2*0.129^{\wedge}2}\right)=1.959$$

（2）P_密度：0.697|否=P（密度=0.697|好瓜=否）

均值$\mu_{c,i}$

假设$\mu_{c,i}=0.496$

方差

假设$\sigma_{c,i}=0.195$

密度属性极大似然估计

$$p(0.697|否)=\frac{1}{\sqrt{2\pi*0.195}}\exp\left(-\frac{(0.694-0.496)^2}{2*0.195^{\wedge}2}\right)=1.203$$

同理计算糖度

$$p(0.460|是)=0.788$$
$$p(0.460|否)=0.066$$

对编号 1 分类预测

$$\frac{P(c)}{P(x)}\cdot\prod_{i=1}^{d}P(x_i|c)$$

因为$P(x)$是固定的，所以这里省去，故只需要看$P(c)\cdot\prod_{i=1}^{d}P(x_i|c)$

$$P(c)\cdot\prod_{i=1}^{d}P(x_i|c)$$

$$=P(好瓜|是)*\mathrm{P}(青绿|是)*\mathrm{P}(蜷缩|是)*\mathrm{P}(浊响|是)*\mathrm{P}(清晰|是)*\mathrm{P}(凹陷|是)*\mathrm{P}(硬滑|是)*p(0.697|是)*p(0.460|是)$$
$$=0.471*0.375*0.325*0.750*0.875*0.750*0.750*1.959*0.788$$
$$=0.037$$

$$\prod_{i=1}^{d}P(x_i|c)$$

$$=P(好瓜|否)*\ \mathrm{P}(青绿|否)*\mathrm{P}(蜷缩|否)*\mathrm{P}(浊响|否)*\mathrm{P}(清晰|否)*\mathrm{P}(凹陷|否)*\mathrm{P}(硬滑|否)*p(0.697|否)*p(0.460|否)$$

= 0.529 *0.333*0.333*0.444*0.222*0.222*0.667*1.203*0.066=6.7978e-5

由于 0.038>6.7978e-5，因此贝叶斯分类器将编号 1 判别为“好瓜|是”。以上数据无缺失值，若样本不充分，观测不仔细，有缺失值怎么办？下面举例说明。

（2）带缺失值的朴素贝叶斯分类器

编号,	色泽,	根蒂,	敲声,	纹理,	脐部,	触感,	密度,	含糖率,	好瓜
1,	青绿,	蜷缩,	清脆,	清晰,	凹陷,	硬滑,	0.697,	0.46,	---
2,	乌黑,	蜷缩,	沉闷,	清晰,	凹陷,	硬滑,	0.774,	0.376,	是
3,	乌黑,	蜷缩,	浊响,	清晰,	凹陷,	硬滑,	0.634,	0.264,	是
4,	青绿,	蜷缩,	沉闷,	清晰,	凹陷,	硬滑,	0.608,	0.318,	是
5,	浅白,	蜷缩,	浊响,	清晰,	凹陷,	硬滑,	0.556,	0.215,	是
6,	青绿,	稍蜷,	浊响,	清晰,	稍凹,	软粘,	0.403,	0.237,	是
7,	乌黑,	稍蜷,	浊响,	稍糊,	稍凹,	软粘,	0.481,	0.149,	是
8,	乌黑,	稍蜷,	浊响,	清晰,	稍凹,	硬滑,	0.437,	0.211,	是
9,	乌黑,	稍蜷,	沉闷,	稍糊,	稍凹,	硬滑,	0.666,	0.091,	否
10,	青绿,	硬挺,	清脆,	清晰,	平坦,	软粘,	0.243,	0.267,	否
11,	浅白,	硬挺,	清脆,	模糊,	平坦,	硬滑,	0.245,	0.057,	否
12,	浅白,	蜷缩,	浊响,	模糊,	平坦,	软粘,	0.343,	0.099,	否
13,	青绿,	稍蜷,	浊响,	稍糊,	凹陷,	硬滑,	0.639,	0.161,	否
14,	浅白,	稍蜷,	沉闷,	稍糊,	凹陷,	硬滑,	0.657,	0.198,	否
15,	乌黑,	稍蜷,	浊响,	清晰,	稍凹,	软粘,	0.36,	0.37,	否
16,	浅白,	蜷缩,	浊响,	模糊,	平坦,	硬滑,	0.593,	0.042,	否
17,	青绿,	蜷缩,	沉闷,	稍糊,	稍凹,	硬滑,	0.719,	0.103,	否

若编号 1 变为上述，训练集也是上述，重复（1）的计算，会发现

先验概率：

P（好瓜=是）=7/17= 0.412；

P（坏瓜=否）=9/17-0.529；

每个属性的极大似然估计

P（青绿|是）=P（色泽=亲绿|好瓜=是）=2/8=0.250；

P（青绿|否）=P（色泽=亲绿|好瓜=否）=3/9=0.333；

P（蜷缩|是）=P（根蒂=蜷缩|好瓜=是）=4/8=0.500；

P（蜷缩|否）=P（根蒂=蜷缩|好瓜=否）=3/9=0.333；

P（清脆|是）=P（敲声=清脆|好瓜=是）=0/8=0.000；
P（清脆|否）=P（敲声=清脆|好瓜=否）=2/9=0.222；
P（清晰|是）=P（纹理=清晰|好瓜=是）=6/8=0.750；
P（清晰|否）=P（纹理=清晰|好瓜=否）=2/9=0.222；
P（凹陷|是）=P（脐部=凹陷|好瓜=是）=5/8=0.625；
P（凹陷|否）=P（脐部=凹陷|好瓜=否）=2/9=0.222；
P（硬滑|是）=P（触感=硬滑|好瓜=是）=5/8=0.625；
P（硬滑|否）=P（触感=硬滑|好瓜=否）=6/9=0.667；
假设 0.697 是

$p(0.697|是) = 1.959$

$p(0.697|否) = 1.203$

$p(0.460|是) = 0.788$

$p(0.460|否) = 0.066$

假设 0.697 否

$p(0.697|是) = 0$

$p(0.697|否) = \mathrm{var2}$

$p(0.460|是) = var3$

$p(0.460|否) = var4$

假设 0.460 是

$p(0.697|是) = 1.959$

$p(0.697|否) = 1.203$

$p(0.460|是) = 0.788$

$p(0.460|否) = 0.066$

假设 0.460 否

$p(0.697|是) = 1.959$

$p(0.697|否) = 1.203$

$p(0.460|是) = 0$

$p(0.460|否) = var6$

...

相当于假设糖度和密度分别是或否，进行排列组合

对编号 1 分类预测

假设 0.697 是

$$P(c) \cdot \prod_{i=1}^{d} P(x_i|c)$$

$$= P(好瓜|是) * \mathrm{P}(青绿|是) * \mathrm{P}(蜷缩|是) * \mathrm{P}(浊响|是) * \mathrm{P}(清晰|是) * \mathrm{P}(凹陷|是) * \mathrm{P}(硬滑|是) * p(0.697|是) * p(0.460|是)$$

$$= 0.412 * 0.25 * 0.5 * 0.00 * 0.750 * 0.625 * 0.625 * 1.959 * 0.788 = 0.038$$

$$P(c)\cdot\prod_{i=1}^{d}P(x_i|c)$$

$$= P(好瓜|否) * \mathrm{P}(青绿|否) * \mathrm{P}(蜷缩|否) * \mathrm{P}(浊响|否) * \mathrm{P}(清晰|否) * \mathrm{P}(凹陷|否) * \mathrm{P}(硬滑|否) * p(0.697|否) * p(0.460|否)$$
$$= 0.412 * 0.25 * 0.5 * 0.00 * 0.750 * 0.625 * 0.625 * 1.959 * 0.788$$
$$= 6.8e-5$$

假设 0.697 否

$$P(c)\cdot\prod_{i=1}^{d}P(x_i|c)$$

$$= P(好瓜|是) * \mathrm{P}(青绿|是) * \mathrm{P}(蜷缩|是) * \mathrm{P}(浊响|是) * \mathrm{P}(清晰|是) * \mathrm{P}(凹陷|是) * \mathrm{P}(硬滑|是) * p(0.697|是) * p(0.460|是)$$
$$= 0.412 * 0.25 * 0.5 * 0.00 * 0.750 * 0.625 * 0.625 * 1.959 * 0.788 = 0$$

$$P(c)\cdot\prod_{i=1}^{d}P(x_i|c)$$

$$= P(好瓜|否) * \mathrm{P}(青绿|否) * \mathrm{P}(蜷缩|否) * \mathrm{P}(浊响|否) * \mathrm{P}(清晰|否) * \mathrm{P}(凹陷|否) * \mathrm{P}(硬滑|否) * p(0.697|否) * p(0.460|否)$$
$$= 0.412 * 0.25 * 0.5 * 0.00 * 0.750 * 0.625 * 0.625 * 1.959 * 0.788 = restult$$

假设 0.460 是

…

假设 0.460 否

…

所以排列组合

…

这就出现了 7.3 节所说的，属性值同类的标号没有同时出现导致概率为 0 的不合理情况。欲解决此问题，需借助拉普拉斯修正法。令 N 表示 D 中可能的类别数，Ni 表示第 i 个属性可能的取值，则修正公式为

$$\hat{P}(c) = \frac{D_c+1}{D+N}$$

$$\hat{P}(x_i|c) = \frac{D_{c,x_i}+1}{D+N_i}$$

对离散属性

A：先验概率

P（好瓜|是）=(8+1)/(17+2)=0.474 （因为好瓜的属性值只有 2 个，分子 1 代表缺 1 个类别）

P（好瓜|否）=(9+1)/(17+2)=0.526 （因为好瓜的属性值只有 2 个，分子 1 代表缺 1 个类别）

估计每个属性的似然

P（青绿|是）=P（色泽=亲绿|好瓜=是）=(2+1)/(8+3)=0.364；（注意属性值有 3 个）

P（青绿|否）=P（色泽=亲绿|好瓜=否）=(3+1)/(9+3)=0.333；

P（蜷缩|是）=P（根蒂=蜷缩|好瓜=是）=(4+1)/(8+3)=0.455；

P（蜷缩|否）=P（根蒂=蜷缩|好瓜=否）=(3+1)/(9+3)=0.417；

P（清脆|是）=P（敲声=清脆|好瓜=是）=(0+1)/(8+3)=0.091；

P（清脆|否）=P（敲声=清脆|好瓜=否）=(2+1)/(9+1)=0.3；

P（清晰|是）=P（纹理=清晰|好瓜=是）=(6+1)/(8+3)=0.636；
P（清晰|否）=P（纹理=清晰|好瓜=否）=(2+1)/(9+3)=0.25；
P（凹陷|是）=P（脐部=凹陷|好瓜=是）=(5+1)/(8+3)=0.545；
P（凹陷|否）=P（脐部=凹陷|好瓜=否）=(2+1)/(9+3)=0.25；
P（硬滑|是）=P（触感=硬滑|好瓜=是）=(5+1)/(8+3)=0.545；
P（硬滑|否）=P（触感=硬滑|好瓜=否）=(6+1)/(9+2)=0.545；（注意触感的属性值仅 2 个）

对连续属性

$p(0.697|是)$

均值：(0.774+0.634+0.608+0.556+0.403+0.481+0.437+1)/(8+1)=0.544; 设连续属性值只有 1 个

方差：((0.744-0.544)^2+(0.634-0.544)^2+(0.608-0.544)^2+(0.556-0.544)^2+(0.403-0.544)^2+(0.481-0.544)^2+(0.437-0.544)^2+(1-0.544)^2)/(9-1)=0.192

$p(0.697|否)$ = var1; 假设一个数，省略计算

同理糖度

$p(0.46|是) = \text{var2}$

$p(0.46|否) = \text{var3}$

预测分类

$$P(c) \cdot \prod_{i=1}^{d} P(x_i|c)$$

用上述公式即可。总之，需要核实连续属性怎么赋值，即以下两个公式怎么确定 N，这个要问问别人或查资料。

$\hat{P}(c) = \frac{D_c+1}{D+N}$

$\hat{P}(x_i|c) = \frac{D_{c,x_i}+1}{D+N_i}$