第二章　贝叶斯决策与分类

基于概率统计知识的分类方法

- 客观现象或事物可以分为两类：一类是确定性的，此类事物在一定条件下必然要发生或不发生；另一类是随机性的，此类事物有若干可能的结果，在实验或实现前不能预知会出现哪种结果，但是其有统计规律，这种规律可用概率分布（密度）函数或数字特征来刻划。

- 贝叶斯决策方法是一个有效的应对具有随机性的模式的分类方法

“概率论”有关知识

■ **划分**：设S为实验E的样本空间，$B_1, B_2, \ldots B_n$为E一组事件，若 $B_i \cap B_j = \Phi,\ i \neq j,\ i, j = 1, 2, \ldots, n$；$B_1 \cup B_2 \cdots \cup B_n = S$ 则称，$B_1, B_2, \ldots B_n$为S一个划分。对每次实验 $B_1, B_2, \ldots B_n$中必有一个且仅有一个发生。

$$P(B_i) > 0\ (i = 1, 2, \cdots, n)$$

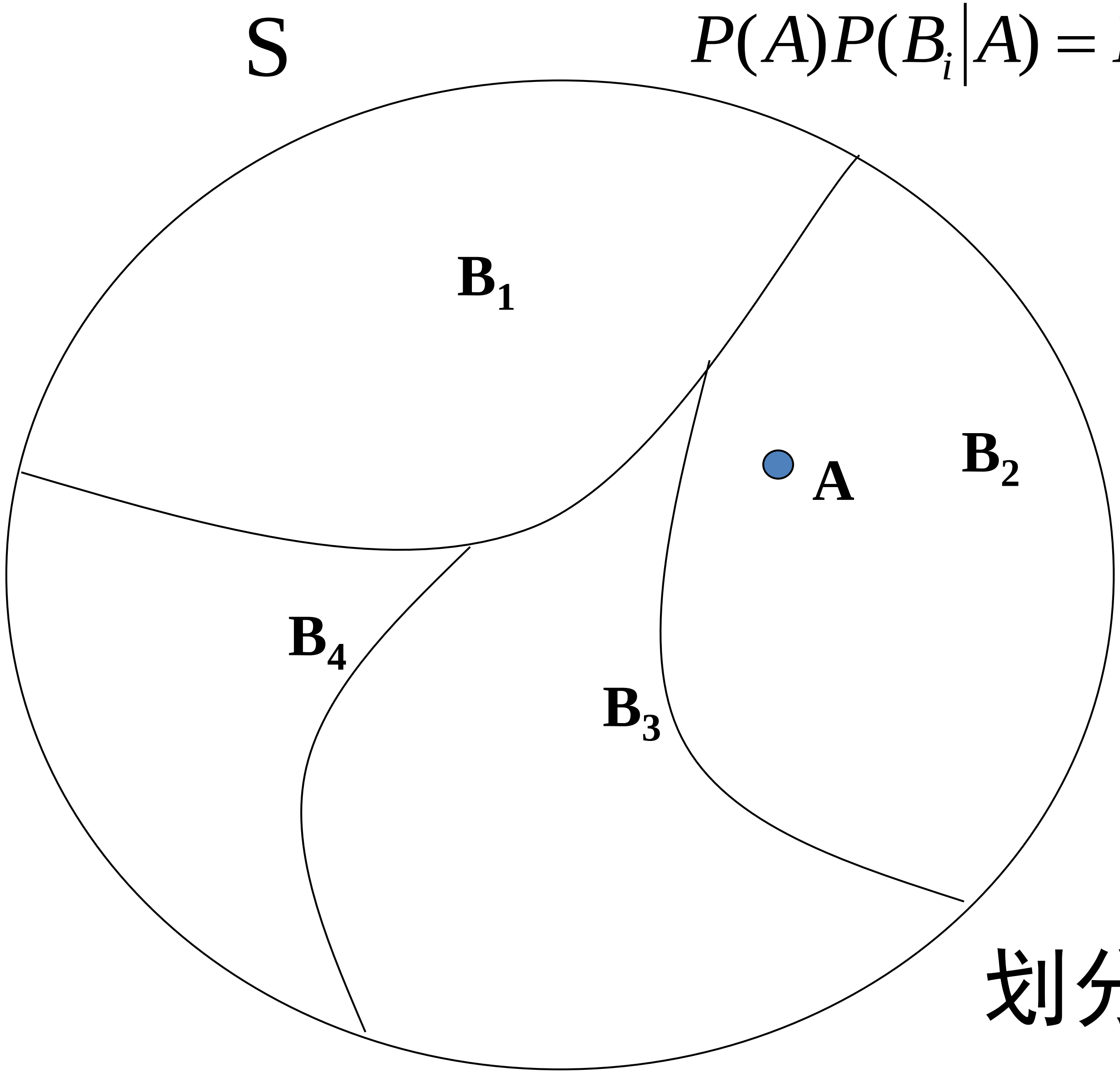
“概率论”有关知识

■ **全概率公式**：设实验E的样本空间为S，A为E的事件， $B_1, B_2, \ldots B_n$为S的一个划分，且，$P(B_i) > 0\ (i = 1, 2, \cdots, n)$ 则

$$\boxed{P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n)}$$

“概率论”有关概念复习

$$P(A)P(B_i|A) = P(B_i)P(A|B_i)$$



划分示意图

“概率论”有关概念复习

Bayes公式： 设实验E的样本空间为S，A为E的事件，$B_1, B_2, \ldots, B_n$为S的一个划分，且$P(A)>0$，$P(B_i)>0$，（i=1, 2, ..., n），则:

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A)} = \frac{P(A \mid B_i)P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j)P(B_j)}$$

$$\boxed{P(A)P(B_i \mid A) = P(B_i)P(A \mid B_i)}$$

“概率论”有关概念复习

$$P(A)P(B_i|A) = P(B_i)P(A|B_i)$$

$$\boxed{p(\vec{x})P(\omega_i|\vec{x}) = P(\omega_i)p(\vec{x}|\omega_i)}$$

先验概率： $P(\omega_i)$表示类ω_i出现的先验概率，简称类ω_i的概率。

条件概率： 事件A发生的条件下，事件B发生的概率P(B|A)。

“概率论”有关概念复习

后验概率：$P(\omega_i|\mathrm{x})$表示x出现条件下类ω_i出现的概率, 称其为类别的后验概率，对于模式识别来讲可理解为（“观察”到的）样本x来自（属于）类ω_i的概率。

类概密：$\boldsymbol{p}(\mathrm{x}|\omega_i)$表示在类$\omega_i$条件下的概率密度，即类$\omega_i$样本 x 的概率分布密度，简称为类概密，也称为第i类的似然函数。

“概率论”有关概念复习

后验概率：$P(Male|x=170)$??

物理意义：已知一个人身高为170cm，其为男性的概率

在类*Male*条件下的概率密度:$P(x=170|Male)$??

最直观的解释：男性中身高为170cm的比例

实例：1300个学生中男生与女生的人数分别为1000与300。其中身高为175cm的男生有200个，身高为175cm的女生有15个

ω_1: *male*

ω_2: *female*

χ ：身高为175cm的事件

$$p(x\,|\,\omega_1)=\frac{200}{1000},\ p(x\,|\,\omega_2)=\frac{15}{300}$$

$$p(x,\omega_1)=\frac{200}{1300},\ p(x,\omega_2)=\frac{15}{1300}$$

$$p(x)=\frac{200+15}{1300}$$

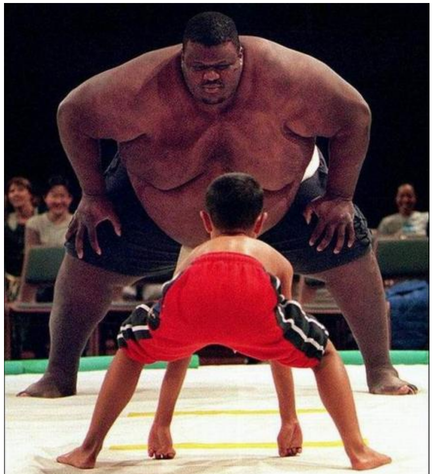
进行观察之前(似然未知或相等)

- **问题**

给定所有可能类别的先验概率，在不进行观察的前提下，预测下一个可能出现的模式的类别

最简单的但不精确的决策规则:

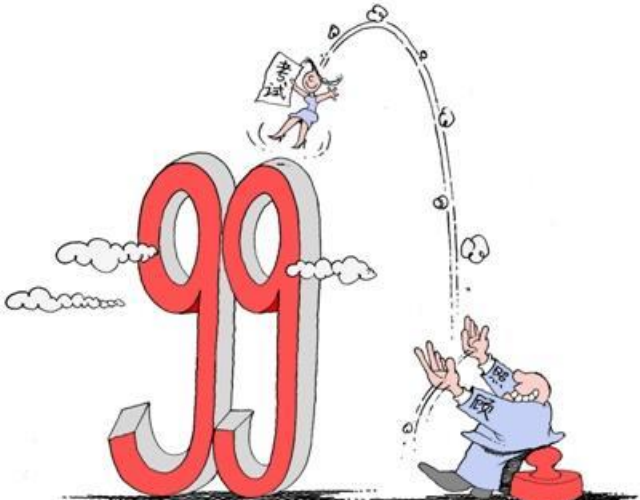
如果 $P(\omega_j) \geq P(\omega_i), \forall i \neq, j$ 则预测下一个模式为 ω_j



Bayes法则—最小错误贝叶斯分类

对于两类ω_1， ω_2问题，直观地，可以根据后验概率做判决：

$$\text{若}\quad p(\omega_1|\vec{x}) > p(\omega_2|\vec{x}) \quad \text{则}\quad \vec{x} \in \omega_1$$
$$\text{若}\quad p(\omega_1|\vec{x}) < p(\omega_2|\vec{x}) \quad \text{则}\quad \vec{x} \in \omega_2$$



相似问题：基于Bayes法则的因果分析

$$P(reason|result)P(result) = P(reason)P(result|reason)$$

百度为您找到相关结果约386,000个　　　　搜索工具

[Bayesian Causal Analysis用贝叶斯网络进行因果分析_百度学术]

王双成，林士敏 - 《计算机科学》 - 2000

The Bayesian causal analysis includes two techniques, one of which takes advantage of Bayesian network structure learning under the Causal Markov assumptio...

xueshu.baidu.com

[贝叶斯 因果分析_相关论文(共734篇)_百度学术]

最高相关	最新发表	
用贝叶斯网络进行因果分析	计算机科学	被引:22
用于因果分析的混合贝叶斯网络结构学习	《智能系统学报》	被引:13
用于离散变量因果分析的贝叶斯网络学习	系统工程学报	被引:6

基于Bayes法则的因果分析

1763年，英国的长老会牧师贝叶斯发表了一篇论文“论有关机遇问题的求解”，提出了解决的框架：那就是用不断增加的信息和经验，可以逐步逼近未知的真相或理解未知。并给出了算法（其实贝叶斯由于是一个牧师，他关心的原始问题本来的表述是：人能不能根据凡人世界的经验和现实世界的证据，证明上帝的存在，因为宗教人士的逻辑是机遇就是上帝存在的主要证据，能够认识机遇的规律，几乎等同于证明上帝存在）。

后来经拉格朗日等数学家进一步努力，获得了大突破，贝叶斯理论成为现代统计学两大支柱之一。

基于Bayes法则的因果分析

人类思考问题有两个方向，一个是正向，也即知道结果找原因（例如现在我们经常讨论的明朝灭亡的原因）；一个是倒向，也即根据一些现象判断结果

例如如果我们事先并不知道黑箱里面黑白球的比例，而是闭着眼睛摸出一个或好几个球，观察这些取出来的球的颜色之后，就此对黑箱里面的黑白球的比例进行推测。现实需要大量的倒向计算，例如现在某些现象出现，企业会不会破产，现在应该怎么办等等。

Bayes法则—最小错误贝叶斯分类

根据Bayes公式，后验概率 $p(\omega_i / \vec{x})$ 可由类ω_i的先验概率$P(\omega_i)$和条件概率密度 $p(\vec{x} / \omega_i)$ 来表示，即

$$p(\omega_i \mid \vec{x}) = \frac{p(\vec{x} \mid \omega_i)P(\omega_i)}{p(\vec{x})} = \frac{p(\vec{x} \mid \omega_i)P(\omega_i)}{\sum_{i=1}^{n} p(\vec{x} \mid \omega_i)P(\omega_i)}$$

式中，$p(x|\omega_i)$又称似然函数(likelihood function of class ω_i)，可由已知样本求得。

解释

贝叶斯概率公式：

最小错误贝叶斯分类重要公式：

$$p(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x} \mid \omega_i)p(\omega_i)}{p(\mathbf{x})}$$

解释： $\mathbf{x}$ 属于第 i 类的概率（即后验概率）等于先验概率与似然的乘积除以 $\mathbf{x}$ 的全概率。

解释

贝叶斯概率公式：

$$p(\omega_i|\mathrm{x}) = \frac{p(\mathrm{x}\,|\,\omega_i)\,p(\omega_i)}{p(\mathrm{x})}$$

理解： (1) 先验概率越大，后验概率就越大

(2) 似然越大，后验概也越大

(1)的形象例子： 北方的蛇基本上都是无毒蛇（无毒蛇这个类别的先验概率很大），所以在北方被蛇咬上不用担心会致命（在北方见碰到一条蛇时，它属于无毒蛇这个类别的概率--后验概率大）

(2)的形象例子： 因为男艺术家中留长发的居多（似然大），所以我们看到留长发的男人就猜测他可能是艺术家（后验概率大）

例题1： 鱼类加工厂对鱼进行自动分类，ω_1:鲈鱼；ω_2:鲑鱼。模式特征x=x(长度)。

已知：（统计结果）

先验概率： $P(\omega_1)=1/3$（鲈鱼出现的概率）

$P(\omega_2)=1-P(\omega_1)=2/3$ （鲑鱼出现的概率）

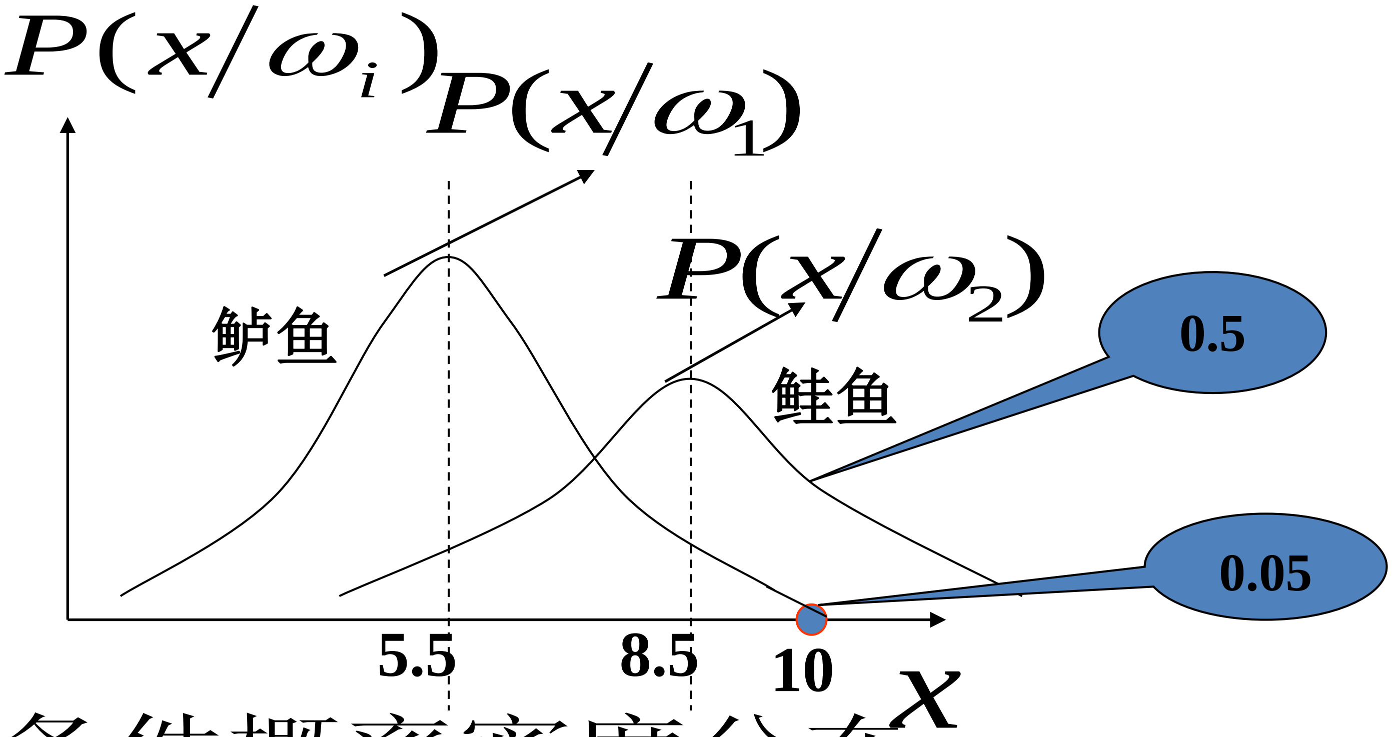
例题1： 鱼类加工厂对鱼进行自动分类，ω_1:鲈鱼；ω_2:鲑鱼。模式特征x=x(长度)。

条件概率： $p(x|\omega_1)$ 见图示（鲈鱼的长度特征分布概率）

$p(x|\omega_2)$见图示（鲑鱼的长度特征分布概率）

求：后验概率： $P(\omega|x=10)=?$

（如果一条鱼x＝10，将其判定为什么类别？）



条件概率密度分布

[例题1图示]

解：

利用**Bayes**公式

$$P(\omega_1 \mid x=10) = \frac{p(x=10 \mid \omega_1)P(\omega_1)}{p(x=10)}$$

$$= \frac{p(x=10 \mid \omega_1)P(\omega_1)}{p(x=10 \mid \omega_1)P(\omega_1) + p(x=10 \mid \omega_2)P(\omega_2)}$$

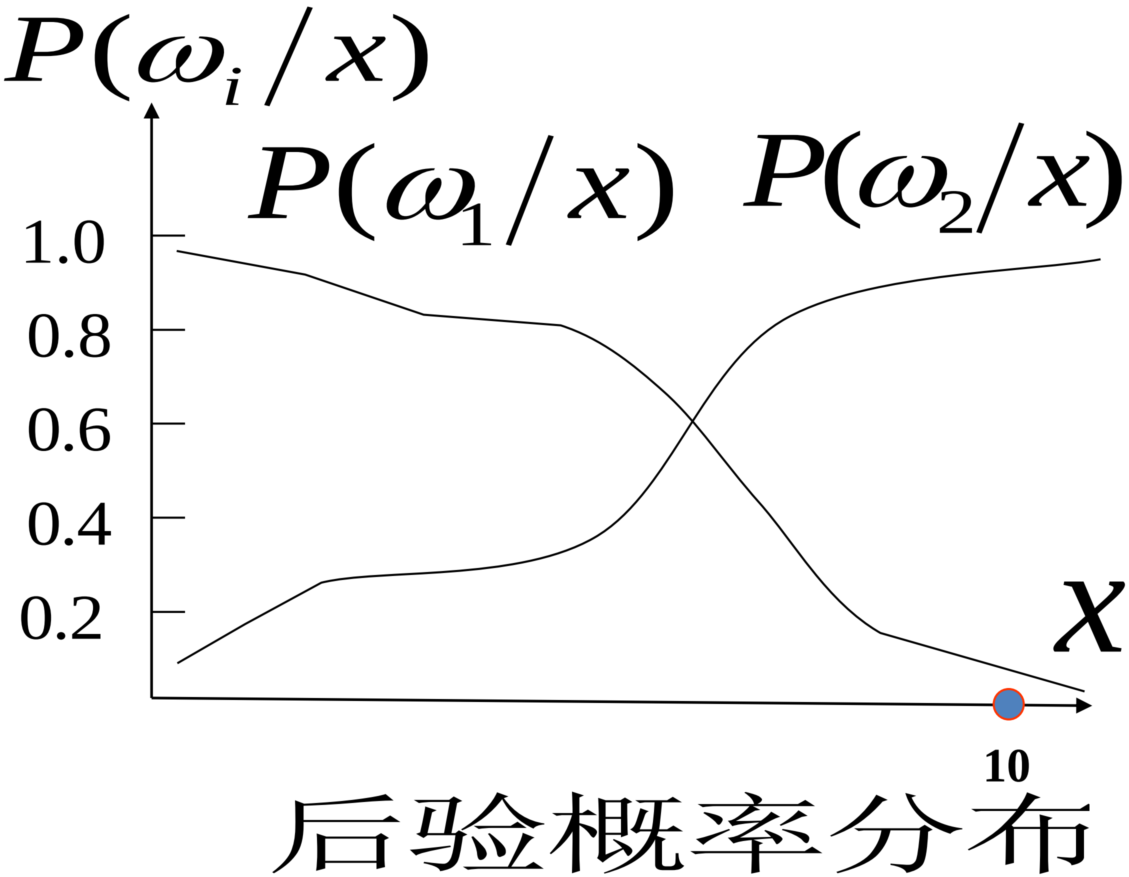
$$= \frac{0.05 \times 1/3}{0.05 \times 1/3 + 0.50 \times 2/3} = 0.048$$

因为， $P(\omega_2 | x=10) = 1 - P(\omega_1 | x=10) = 1 - 0.048 = 0.952$

$$P(\omega_1 | x=10) < P(\omega_2 | x=10)$$

故判决： $(x=10) \in \omega_2$ **，即是鲑鱼。**

结论：长度为10的这条鱼应该判别为鲑鱼！！



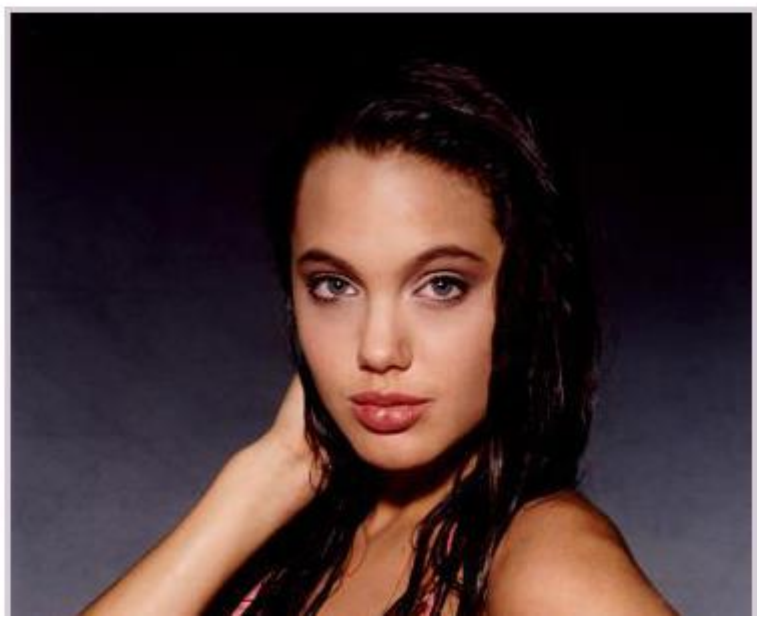
例题1图示

Bayes法则—最小错误贝叶斯分类(两类问题)

$$\text{错误率}=1-\max\left(P(\omega_1/x),\ P(\omega_2/x)\right)$$

1 性感女性安吉丽娜·朱莉做了乳腺切除手术

好莱坞知名女星安吉丽娜·朱莉14日在《纽约时报》上发表了“我的医疗选择”一文，她表示自己已经接受预防性的双侧乳腺切除手术，以降低罹癌风险。而之所以进行这项手术，是因为她有基因缺陷，罹患乳癌和卵巢癌风险恐较高。



52%

例：对一批人进行癌症普查，患癌症者定为属ω_1类，正常者定为属ω_2类。统计资料表明人们患癌的概率$P(\omega_1)=0.005$，从而$P(\omega_2)=0.995$。设有一种诊断此病的试验，其结果有阳性反应和阴性反应之分，依其作诊断。化验结果是一维离散模式特征。统计资料表明：癌症者有阳性反映的概率为**0.95**即$P(x=\text{阳}|\omega_1)=0.95$，从而可知$P(x=\text{阴}|\omega_1)=0.05$，正常人阳性反映的概率为**0.01**即$P(x=\text{阳}|\omega_2)=0.01$，可知$P(x=\text{阴}|\omega_2)=0.99$。

问有阳性反映的人患癌症的概率有多大？

解：

$$P(\omega_1|x=\text{阳})=\frac{P(x=\text{阳}|\omega_1)P(\omega_1)}{P(x=\text{阳})}$$

$$=\frac{P(x=\text{阳}|\omega_1)P(\omega_1)}{P(x=\text{阳}|\omega_1)P(\omega_1)+P(x=\text{阳}|\omega_2)P(\omega_2)}$$

$$=\frac{0.95\times0.005}{0.95\times0.005+0.01\times0.995}$$

$$=0.323$$

说明有阳性反应的人其患癌的概率有32.3%

思考

1. 在多类分类问题中最小错误贝叶斯分类的决策公式是什么?

2. 在多类分类问题中最小错误贝叶斯分类的错误率公式可得到吗?

3. 你认为最小错误贝叶斯分类在实际应用中有什么不便之处（缺点）？