河北师范大学软件学院
Software College of Hebei Normal University

机器学习入门

贝叶斯决策

河北师范大学软件学院

2017.04.16-04.26

张朝晖

河北师范大学软件学院
Software College of Hebei Normal University

监督式学习

分类
- A. 懒惰学习--KNN法(K最近邻法)分类模型
- B. 概率学习 --贝叶斯分类器
- C. 分而治之--分类树

回归
- D. KNN法回归
- E. 分而治之--回归树
- F. 最小二乘回归

集成学习

非监督式学习
- 聚类 -- G. K-Means Clustering
- 特征提取 -- H. PCA

模型评价

本课概要

两种典型的贝叶斯分类模型

(1)最小错误率贝叶斯分类

--朴素贝叶斯分类

(2)最小风险贝叶斯分类

两步贝叶斯决策过程

概率密度函数的估计--最大似然估计

主要内容

1. 引言

2. 贝叶斯决策模型

   2.1 最小错误率贝叶斯决策

   连续情况；离散情况

   2.2 最小风险的贝叶斯决策

   连续情况；离散情况

3. 正态分布的概率密度函数及性质

4. 概率/概率密度函数估计

5. 朴素贝叶斯(Naive Bayes)分类

机器学习中的分类问题----“状态决策”问题

根据待识别对象的观测，将其划分到某类别

统计模式识别：

用概率统计的观点和方法来解决模式分类问题

贝叶斯决策论（统计决策理论）：

是统计模式识别的基本方法和基础;

利用概率的不同，分类决策、或决策代价的折中;

----“最优分类器”

几个基本概念

河北师范大学软件学院
Software College of Hebei Normal University

[1]特征空间(输入空间) 及特征维数

d维特征空间，记为$\Re = \boldsymbol{R}^d$

[2]特征向量(随机向量) $\boldsymbol{X} = [X_1, \ldots, X_d]^T$

其中 $X_1, X_2, \ldots, X_d$分别为随机变量。

$\forall \boldsymbol{x} = [x_1, \ldots, x_d]^T \in \boldsymbol{R}^d$ 或：观察样本$\boldsymbol{x}$为d维

[3]状态空间(输出空间) $\{\omega_1, \ldots, \omega_c\}$

类别状态数$\boldsymbol{c}$，类别状态变量ω(随机变量) $\omega \in \{\omega_1, \ldots, \omega_c\}$

[4](状态)先验概率

预先已知的，或可估计的分类系统位于某一类别的概率。

一般的$\boldsymbol{c}$类问题：各类别ω_i的先验概率$P_\omega(\omega_i), \boldsymbol{i} = 1, \ldots, c$

$$P_\omega(\omega_1) + \ldots + P_\omega(\omega_c) = 1$$

[5]样本分布密度(或：总体概率密度) $p_X(x)$

$$\int_{\Re} p_X(x)dx = 1$$

[6]类条件概率密度函数

$$(class\text{-}conditional\ probability\ density\ function)$$

系统位于某种类别条件下，模式样本x的概率密度分布。同一类别对象的各属性具有一定变化范围，以函数形式表示，记为：$p_{X|\omega}(x|\omega_i), \boldsymbol{i}=1,\ldots,c$

$$\int_{\Re} p_{X|\omega}(x|\omega_i)dx = 1, \qquad \boldsymbol{i}=1,\ldots,c$$

[7]后验概率(*posterior probability*)

给定某具体模式样本的观测x，该样本属于某类别的概率。

记为：$P_{\omega|X}(\omega_i|x), \boldsymbol{i}=1,\ldots,c$

$$\sum_{i=1}^{c} P_{\omega|X}(\omega_i|x) = 1$$

连续情况下，贝叶斯决策条件：

[1] 类别数目c一定，类别状态ω_i， $\boldsymbol{i}=1,\ldots,c$

[2] 给定：

$$\begin{cases} \text{类先验概率} & P_{\omega}(\omega_i), \quad \boldsymbol{i}=1,\ldots,c \\ \text{类条件概率密度} & p_{X|\omega}p(x\mid\omega_i), \quad \boldsymbol{i}=1,\ldots,c \\ \text{损失代价} & \lambda_{ij} \end{cases}$$

(对真实状态为ω_j类的向量$\boldsymbol{x}$采取决策α_i所带来的损失)

特殊情况　　0－1代价 $\lambda_{ij}=\begin{cases} 0 & \boldsymbol{i}=\boldsymbol{j} \\ 1 & \boldsymbol{i}\neq\boldsymbol{j} \end{cases}$

问题：如何最为合理地对观测向量$\boldsymbol{x}=\left[x_1,x_2,\ldots,x_d\right]^{\boldsymbol{T}}\in\boldsymbol{R}^d$进行判决？

几个常用决策模型

为方便描述，贝叶斯模型描述作如下简化：

$$p_{\mathbf{X}}(x) \quad \Rightarrow \quad p(x)$$

$$P_{\mathbf{X}}(x) \quad \Rightarrow \quad P(x)$$

$$p_{\mathbf{X}|\omega}(x \mid \omega_i) \quad \Rightarrow \quad p(x \mid \omega_i)$$

$$P_{\mathbf{X}|\omega}(x \mid \omega_i) \quad \Rightarrow \quad P(x \mid \omega_i)$$

$$P_{\omega|\mathbf{X}}(\omega_i \mid x) \quad \Rightarrow \quad P(\omega_i \mid x)$$

$$P_{\omega}(\omega_i) \quad \Rightarrow \quad P(\omega_i)$$

河北师范大学软件学院

Software College of Hebei Normal University

主要内容

1. 引言

2. 贝叶斯决策模型

    2.1 最小错误率贝叶斯决策

        连续情况；离散情况

    2.2 最小风险的贝叶斯决策

        连续情况；离散情况

3. 正态分布的概率密度函数及性质

4. 概率/概率密度函数估计

5. 朴素贝叶斯(Naive Bayes)分类

分析一个两类问题----正常/异常细胞的识别问题

➔ 预测某具体观测细胞x的类别状态。

给定：

(1) 两类的分类问题： c=2

类别状态集合{w1=正常， w2=异常}

(2) 关于该问题，需要借助的一些信息，几种可能：

方式1. 仅仅依靠先验知识：$P(w=w_1)$， $P(w=w_2)$

方式2. 先验知识、该细胞的具体观测x

方式3. 先验知识、观测x、不同决策所带来的损失

➔ 如何决策？

1.仅依靠先验知识(未使用关于该样本x的具体观测信息)

判决规则 $\boxed{\text{若}P(\omega_1)>P(\omega_2),\text{则决策为}\mathbf{x}\in\omega_1;\ \text{反之},\ \mathbf{x}\in\omega_2}$

$P(\omega_1)=0.9 \quad P(\omega_2)=0.1$

➔ 所有待识别细胞全部决策为 ω_1= 正常细胞。

➔ 决策错误率：$P(error)=1-P(\omega_1)=P(\omega_2)$

产生原因：状态先验概率提供的分类信息太少。

$P(\omega 1)$、$P(\omega 2)$ 不等时，仅利用先验概率，只能把未知样本都归于某一类，无法达到正确分类的目的。

因此,先验概率不能作为判决的唯一依据。

河北师范大学软件学院

Software College of Hebei Normal University

2. “先验知识” + “证据(即：该观测样本x)”

贝叶斯公式的实质：通过观测信息x (证据)，将状态先验概率$P(\omega_j)$转化为状态后验概率$P(\omega_j|x)$

$$posterior = \frac{likelihood \times prior}{evidence} \qquad \sum_{i=1}^{c} P(\omega_j \mid x) = 1$$

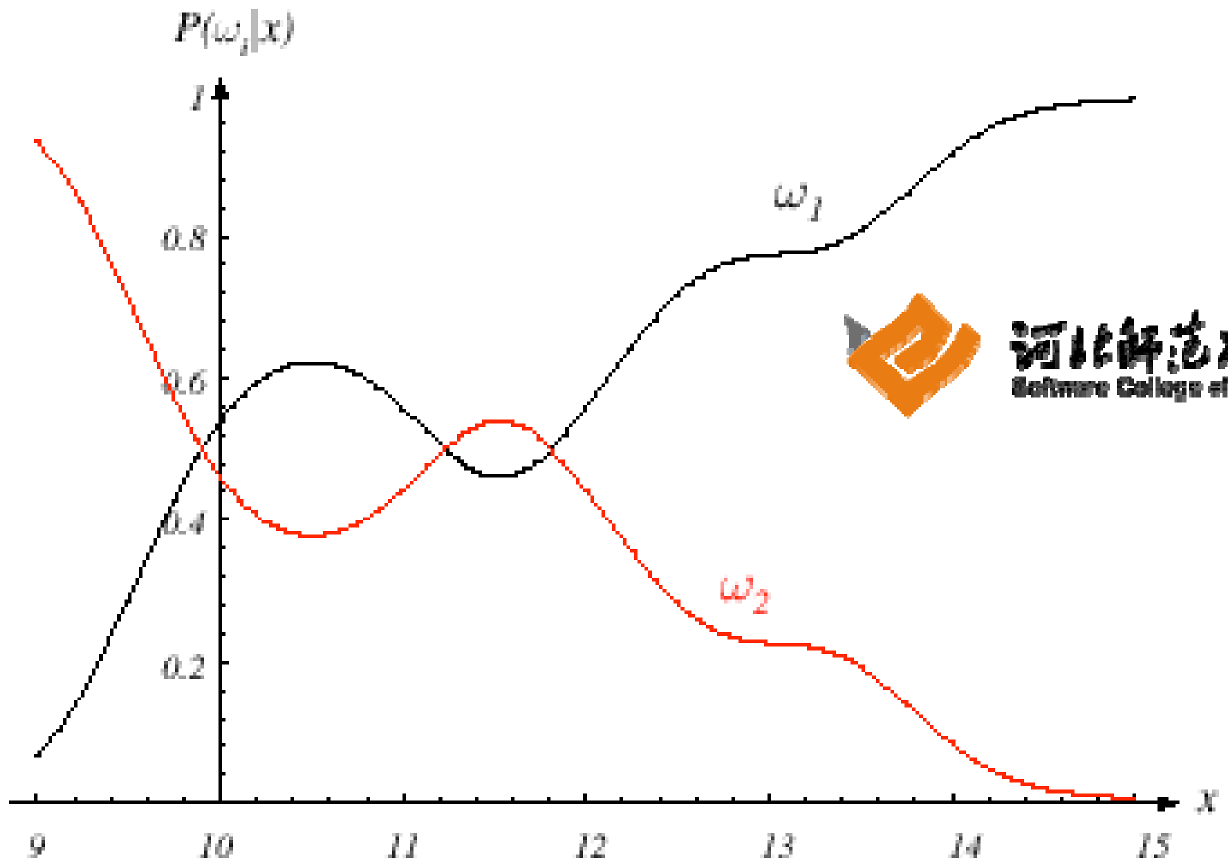
观测空间为连续特征空间

$$P(\omega_j \mid x) = \frac{p(x,\omega_j)}{p(x)} = \frac{p(x \mid \omega_j)P(\omega_j)}{p(x)} = \frac{p(x \mid \omega_j)P(\omega_j)}{\sum_{i=1}^{c} p(x \mid \omega_i)P(\omega_i)}$$

观测空间为离散特征空间

$$P(\omega_j \mid x) = \frac{P(x,\omega_j)}{P(x)} = \frac{P(x \mid \omega_j)P(\omega_j)}{P(x)} = \frac{P(x \mid \omega_j)P(\omega_j)}{\sum_{i=1}^{c} P(x \mid \omega_i)P(\omega_i)}$$

例：一维连续特征情况下，某两类问题的**状态后验概率**



似然函数： 其它条件相等情况下，对应较大取值$p(x|\omega_j)$的ω_j更有可能为x的真实类别.因此，称$p(x|\omega_j)$为ω_j关于x的似然函数。

(1) 连续特征空间，两类别分类问题：决策规则

对于观测样本$\boldsymbol{x}$，计算$P(\omega_j \mid \boldsymbol{x})$:

$$P(\omega_j \mid \boldsymbol{x}) = \frac{p(\boldsymbol{x} \mid \omega_j)P(\omega_j)}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{x} \mid \omega_j)P(\omega_j)}{\sum_{i=1}^{2} p(\boldsymbol{x} \mid \omega_i)P(\omega_i)} \qquad j = 1, 2$$

决策规则 $\begin{cases} 若 P(\omega_1 \mid \boldsymbol{x}) > P(\omega_2 \mid \boldsymbol{x}), 则决策为 \boldsymbol{x} \in \omega_1 \\ 若 P(\omega_2 \mid \boldsymbol{x}) > P(\omega_1 \mid \boldsymbol{x}), 则决策为 \boldsymbol{x} \in \omega_2 \end{cases}$

上述决策对应的错误概率$P(e \mid \boldsymbol{x})$:

$$P(e \mid \boldsymbol{x}) = \begin{cases} 1 - P(\omega_1 \mid \boldsymbol{x}) = P(\omega_2 \mid \boldsymbol{x}) < P(\omega_1 \mid \boldsymbol{x}) & 若决策 \boldsymbol{x} \in \omega_1 \\ 1 - P(\omega_2 \mid \boldsymbol{x}) = P(\omega_1 \mid \boldsymbol{x}) < P(\omega_2 \mid \boldsymbol{x}) & 若决策 \boldsymbol{x} \in \omega_2 \end{cases}$$

$$= 1 - \max\{P(\omega_1 \mid \boldsymbol{x}), P(\omega_2 \mid \boldsymbol{x})\} = \min\{P(\omega_1 \mid \boldsymbol{x}), P(\omega_2 \mid \boldsymbol{x})\}$$

基于**最大后验概率**的决策 $\Leftrightarrow$ 基于**最小条件错误概率**$P(e \mid \boldsymbol{x})$的决策

连续特征空间，错误率(平均错误率)

错误率(平均错误率)是错误概率$P(e|x)$的期望，记为$P(e)$

对于连续随机变量x，有

$$P(e)=\int_{\Re} P(e|x)p(x)dx \qquad P(e|x)\geq 0, p(x)\geq 0$$

最大后验概率贝叶斯决策就是基于最小错误率贝叶斯决策

$$P(e)=\int_{\Re} P(e|x)p(x)dx$$

每个x的判决，都对应一个条件错分概率$P(e|x)$，

对于所有x，若能保证决策时关于x的后验概率最大，

则可保证$P(e|x)$最小，则有$P(e)$最小。

河北师范大学软件学院
Software College of Hebei Normal University

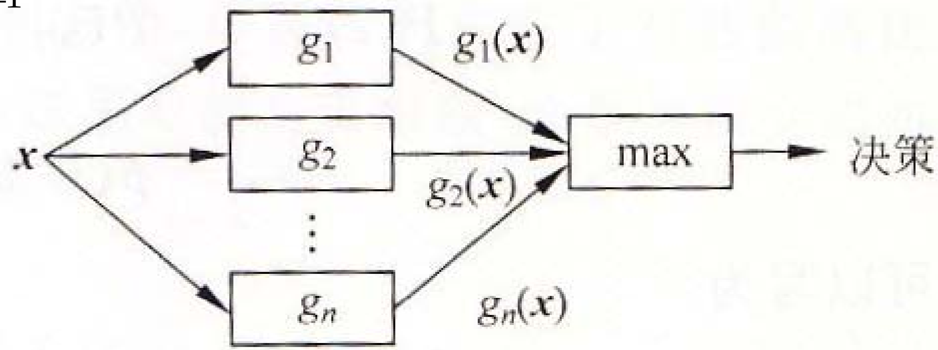
(2).多类问题情况下，基于最小错误率的贝叶斯决策

后验概率

$$P(\omega_j \mid x) = \frac{p(x \mid \omega_j)P(\omega_j)}{p(x)} = \frac{p(x \mid \omega_j)P(\omega_j)}{\sum_{i=1}^{c} p(x \mid \omega_i)P(\omega_i)} \qquad j = 1, 2, \ldots, c$$

样本x的错分概率

$$P(e \mid x) = 1 - \max_{j=1,2,\ldots,c} P(\omega_j \mid x)$$



基于最小错误率的决策规则：

若$P(\omega_i \mid x) = \max_{j=1,2,\ldots,c} P(\omega_j \mid x)$ 则 $x \in \omega_i$

或

若$p(x \mid \omega_i)P(\omega_i) = \max_{j=1,2,\ldots,c} p(x \mid \omega_j)P(\omega_j)$ 则 $x \in \omega_i$

例：离散特征空间基于最小错误率的贝叶斯决策—例题分析

例1：为了对某种疾病进行诊断，对一批人进行一次普查，对每个人各打试验针，观察反应，然后进行统计，规律如下：

(1)这一批人中，每1000个人中有5个关于该疾病的病人；

(2)这一批人中，每100个正常人中有一个试验呈阳性反应；

(3)这一批人中，每100个病人中有95人试验呈阳性反应。

问：若某人（甲）呈阳性反应，甲是否正常？

分析：类别状态空间$\Omega = \{\omega_1 = \text{正常}, \omega_2 = \text{异常}\}$

$$\begin{cases} (1) \Rightarrow \text{先验概率} \\ (2) \Rightarrow \text{类条件概率} \boldsymbol{P}(\boldsymbol{x} = \text{阳性} \mid \omega_1) \\ (3) \Rightarrow \text{类条件概率} \boldsymbol{P}(\boldsymbol{x} = \text{阳性} \mid \omega_2) \end{cases}$$

确定$\boldsymbol{P}(\omega_i \mid \boldsymbol{x} = \text{阳性}), \boldsymbol{i} = 1, 2$

河北师范大学软件学院
Software College of Hebei Normal University

解：观测x=阳性。

(1) 类别状态ω有两种：$\omega=\omega_1$ 正常；$\omega=\omega_2$ 某疾病患者

(2) 根据已知条件，计算关于类别状态的后验概率。

由条件1，状态先验概率： $P(\omega_1)=0.995, P(\omega_2)=0.005$

由条件2,3，类条件概率：$\begin{cases} P(x=\text{阳性}\mid w_1)=0.01 \\ P(x=\text{阳性}\mid w_2)=0.95 \end{cases}$

(3) 决策过程：

$$P(\omega_1\mid x=\text{阳性})=\frac{P(\omega_1)P(x=\text{阳性}\mid w_1)}{P(x=\text{阳性})}$$

$$=\frac{P(\omega_1)P(x=\text{阳性}\mid w_1)}{P(\omega_1)P(x=\text{阳性}\mid w_1)+P(\omega_2)P(x=\text{阳性}\mid w_2)}=\frac{0.995\times 0.01}{0.995\times 0.01+0.005\times 0.95}=0.677$$

$P(\omega_2\mid x=\text{阳性})=1-0.677=0.323$

$P(\omega_1\mid x=\text{阳性})>P(\omega_2\mid x=\text{阳性})$

由最小错误率决策规则，将"阳性反应的甲"判决为ω_1（正常人），

判决错误率：0. 323

由于$P(\omega_1)>P(\omega_2)$，判决时，先验概率起了明显的作用。

最小错误率贝叶斯决策—小结

两类情况下，基于最小错误率的判决规则，各等价形式：

1，状态后验概率
$$\begin{cases} P(\omega_1 \mid x) > P(\omega_2 \mid x), 则 x \in \omega_1 类 \\ P(\omega_1 \mid x) < P(\omega_2 \mid x), 则 x \in \omega_2 类 \end{cases}$$

2，似然值×先验概率
$$\begin{cases} P(\omega_1) p(x \mid \omega_1) > P(\omega_2) p(x \mid \omega_2), 则 x \in \omega_1 类 \\ P(\omega_1) p(x \mid \omega_1) < P(\omega_2) p(x \mid \omega_2), 则 x \in \omega_2 类 \end{cases}$$

3，似然比
$$\begin{cases} 若\ l(x) = \dfrac{p(x \mid \omega_1)}{p(x \mid \omega_2)} > \dfrac{P(\omega_2)}{P(\omega_1)}, 则\ x \in \omega_1 类 \\ 若\ l(x) = \dfrac{p(x \mid \omega_1)}{p(x \mid \omega_2)} < \dfrac{P(\omega_2)}{P(\omega_1)}, 则\ x \in \omega_2 类 \end{cases}$$

最小错误率贝叶斯决策—小结

多类情况下，基于最小错误率的判决规则，各等价形式

(1)后验概率：

若 $P(\omega_i \mid \boldsymbol{x}) = \max\limits_{j=1,2,..,c} P(\omega_j \mid \boldsymbol{x})$

则 $\boldsymbol{x} \in \omega_i$类

(2)似然值×先验概率：

若 $p(\boldsymbol{x} \mid \omega_i)P(\omega_i) = \max\limits_{j=1,2,...,c} p(\boldsymbol{x} \mid \omega_j)P(\omega_j)$

则 $\boldsymbol{x} \in \omega_i$类

河北师范大学软件学院
Software College of Hebei Normal University

主要内容

1. 引言
2. 贝叶斯决策模型
   2.1 最小错误率贝叶斯决策
      连续情况；离散情况
   2.2 最小风险的贝叶斯决策
      连续情况；离散情况
   “先验知识”＋“证据观测X”＋“不同决策所带来的损失
3. 正态分布的概率密度函数及性质
4. 概率/概率密度函数估计
5. 朴素贝叶斯(Naive Bayes)分类

最小错误率的贝叶斯决策，仅仅考虑判决错误。

实际决策过程中，不同类别的错误决策，还存在不同程度的损失或风险。

例：疾病的诊断

正常诊断为疾病，带来精神负担，但可进一步检查；

疾病诊断为正常，延误治疗、危及生命，损失严重。

因此，决策时可考虑不同类别的决策错误所引起的风险（损失）。

最小风险贝叶斯决策

1.问题描述

[1]观测$\boldsymbol{x}$

$\boldsymbol{x}$为$\boldsymbol{d}$维随机向量，$\boldsymbol{x}=\left[x_1,x_2,\ldots,x_d\right]^T$

$x_1,x_2,\ldots,x_d$分别为随机变量。

[2]状态空间Ω

$\boldsymbol{c}$个自然状态（$\boldsymbol{c}$类），$\Omega=\{\omega_1,\ldots,\omega_{\boldsymbol{c}}\}$

[3]决策空间(或：行动空间)$\mathcal{A}$

对观测$\boldsymbol{x}$可能采取的k个决策 $\mathcal{A}=\{\alpha_1,\ldots,\alpha_k\}$

注：$\boldsymbol{c}$与k可能不同

[4]损失函数$\lambda(\alpha_i, \omega_j)$

对真实类别状态为ω_j的观察$\boldsymbol{x}$，采取决策α_i

所带来的损失(或：风险)，简称λ_{ij}。

$\boldsymbol{i} = 1, 2, \ldots, \boldsymbol{k} \qquad \boldsymbol{j} = 1, 2, \ldots, \boldsymbol{c}$

以"决策表格"形式提供λ_{ij}

[5]**特征空间连续**--类条件概率密度$p(\boldsymbol{x} \mid \omega_i)$

$$\int_{\Re} p(x \mid \omega_i)dx = 1, \qquad \boldsymbol{i} = 1, \ldots, c$$

特征空间离散--类条件概率$P(\boldsymbol{x} \mid \omega_i)$

$$\sum_{\boldsymbol{x} \in} P(\boldsymbol{x} \mid \omega_i) = 1, \qquad \boldsymbol{i} = 1, \ldots, c$$

[6]状态先验概率$P(\omega_i)$ $\qquad \sum_{\boldsymbol{i}=1}^{\boldsymbol{c}} P(\omega_i) = 1 \qquad \boldsymbol{i} = 1, \ldots, \boldsymbol{c}$

$\Rightarrow$目标：对所有的$\boldsymbol{x}$进行决策，使损失最小。

2.与“风险”“损失”有关的几个名词

[1]决策表(损失函数)

对于具体观测 $\boldsymbol{x}$，所有可能采取的决策或行动都会带来一定风险，以决策表表示：

决策	自然状态					
	ω_1	ω_2	$\cdots$	ω_j	$\cdots$	ω_c
α_1	$\lambda(\alpha_1,\omega_1)$	$\lambda(\alpha_1,\omega_2)$	$\cdots$	$\lambda(\alpha_1,\omega_j)$	$\cdots$	$\lambda(\alpha_1,\omega_c)$
α_2	$\lambda(\alpha_2,\omega_1)$	$\lambda(\alpha_2,\omega_2)$	$\cdots$	$\lambda(\alpha_2,\omega_j)$	$\cdots$	$\lambda(\alpha_2,\omega_c)$
$\vdots$						$\vdots$
α_i	$\lambda(\alpha_i,\omega_1)$	$\lambda(\alpha_i,\omega_2)$	$\cdots$	$\lambda(\alpha_i,\omega_j)$	$\cdots$	$\lambda(\alpha_i,\omega_c)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
α_k	$\lambda(\alpha_k,\omega_1)$	$\lambda(\alpha_k,\omega_2)$	$\cdots$	$\lambda(\alpha_k,\omega_j)$	$\cdots$	$\lambda(\alpha_k,\omega_c)$

$$\text{损失函数}\ \lambda(\alpha_i,\omega_j)=\lambda_{ij}$$

[2]观测x的**条件期望损失**（或：x的条件风险）$R(\alpha(x)|x)$

分析：

$$\boxed{\text{观测}x} \rightarrow \boxed{\begin{array}{l}\text{真实类别状态}\omega=?\\ \omega_j,\quad j=1,2,\ldots,c\end{array}} \xrightarrow[f(\omega)]{\lambda(\alpha_i,\omega_j)} \boxed{\begin{array}{l}\text{决策行为}\alpha(x)\\ \alpha_i,\quad i=1,2,\ldots,k\end{array}}$$

$\Rightarrow$ **条件期望损失**：$R(\alpha(x)|x)$

对**观测x**在各种可能类别下，采取**某具体决策**$\alpha(x)$所造成的平均损失。

$$R(\alpha(x)|x)=E[f(\omega)|x]=\sum_{j=1}^{c}f(\omega=\omega_j)P(\omega=\omega_j|x)$$

河北师范大学软件学院
Software College of Hebei Normal University

[2]观测x的**条件期望损失**(或：x的条件风险)$R(\alpha(x)|x)$

$$R(\alpha(x)|x)=E[f(\omega)|x]=E[\lambda(\alpha(x),\omega)|x]$$

$$=\sum_{j=1}^{c} f(\omega=\omega_j)P(\omega=\omega_j|x)$$

$$=\sum_{j=1}^{c}\lambda(\alpha(x),\omega=\omega_j)P(\omega=\omega_j|x)=\sum_{j=1}^{c}\lambda(\alpha(x),\omega_j)P(\omega_j|x)$$

当$\alpha(x)=\alpha_i$时，

$$R(\alpha_i|x)=\sum_{j=1}^{c}\lambda(\alpha_i,\omega_j)P(\omega_j|x) \qquad i=1,2,\ldots,k$$

$\Rightarrow$ $\boxed{\text{对于观测}x\text{，可选择行为}\alpha_i\text{，使}R(\alpha_i|x)=\min_{j=1,2,\ldots,k}R(\alpha_j|x)}$

河北师范大学软件学院
Software College of Hebei Normal University

[3] **期望风险(平均风险)** $R(\alpha)$

$R(\alpha)$是条件期望风险$R(\alpha(x)|x)$关于观测x的数学期望。

$$R(\alpha)=\int R(\alpha(x)|x)p(x)dx$$

意义：对特征空间的所有x，采取相应决策$\alpha(x)$所带来的平均风险。

$\Rightarrow$ $\boxed{\text{对于每个}x\text{，若能选择}\alpha(x)\text{，使条件风险}R(\alpha(x)|x)\text{最小，则能保证期望风险}R=R(\alpha)\text{最小}}$

[4] **贝叶斯风险**：最小的期望风险 $R^*=\min\limits_{\alpha} R(\alpha)$

河北师范大学软件学院
Software College of Hebei Normal University

3.最小风险贝叶斯决策规则

若　对于$\boldsymbol{x}$，$\boldsymbol{R}(\alpha_j \mid \boldsymbol{x}) = \min_{\boldsymbol{i}=1,2,\ldots,\boldsymbol{k}} \boldsymbol{R}(\alpha_{\boldsymbol{i}} \mid \boldsymbol{x})$

则　应对$\boldsymbol{x}$采取行为$\alpha = \alpha_j$

4.最小风险贝叶斯决策的计算步骤

对于观测$\boldsymbol{x}$

(1) 计算后验概率：

$$P(\omega_j|\boldsymbol{x})=\frac{P(\omega_j)p(\boldsymbol{x}|\omega_j)}{p(\boldsymbol{x})}=\frac{P(\omega_j)p(\boldsymbol{x}|\omega_j)}{\sum_{i=1}^{c}P(\omega_i)p(\boldsymbol{x}|\omega_i)}$$

$$j=1,2,\ldots,c$$

(2) 计算$\boldsymbol{x}$的条件风险： $R(\alpha_i|\boldsymbol{x})=\sum_{j=1}^{c}\lambda\left(\alpha_i,\omega_j\right)P(\omega_j|\boldsymbol{x})$

$$i=1,2,\ldots,k$$

(3) 决策： 选择关于$\boldsymbol{x}$的条件风险最小的决策α，即

$$\alpha=\arg\min_{i=1,2,\ldots,k}R(\alpha_i|\boldsymbol{x})$$

5.最小风险贝叶斯决策举例--两类别问题

例2：在例1的异常诊断问题中，所有化验结果对应两种判决。试验人为两类；并且，由于决策会产生不同程度风险。对应决策表如下。

决策 α_i $i=1,2$	自然状态 ω_j $j=1,2$	
	ω_1 =正常	ω_2 =异常
α_1 =判决为 ω_1	$\lambda(\alpha_1,\omega_1)=0.5$	$\lambda(\alpha_1,\omega_2)=6$
α_2 =判决为 ω_2	$\lambda(\alpha_2,\omega_1)=2$	$\lambda(\alpha_2,\omega_2)=0.5$

为了对疾病进行诊断，对一批人进行一次普查，对每个人各打试验针，观察反应，然后进行统计，规律如下：

（1）这一批人中，每1000个人中有5个病人；

（2）这一批人中，每100个正常人中有一个试验呈阳性反应；

（3）这一批人中，每100个病人中有95人试验呈阳性反应。

问：若某人（甲）呈阳性反应，甲是否正常？

分析：

$$\begin{cases}\text{状态空间}\Omega=\{\omega_1=\text{正常},\omega_2=\text{异常}\}\\ \text{决策空间}\boldsymbol{\mathcal{A}}=\{\alpha_1=\text{判决为}\omega_1,\alpha_2=\text{判决为}\omega_2\}\\ \text{有关概率信息}\begin{cases}(1)\Rightarrow\text{先验概率}\\(2)\Rightarrow\text{类条件概率}\boldsymbol{P}(\boldsymbol{x}=\text{阳性}\mid\omega_1)\\(3)\Rightarrow\text{类条件概率}\boldsymbol{P}(\boldsymbol{x}=\text{阳性}\mid\omega_2)\end{cases}\end{cases}$$

确定：$\boldsymbol{R}(\alpha_i\mid\boldsymbol{x}=\text{阳性}),\boldsymbol{i}=1,2$

判决结果与例1相反，影响判决结果的主导因素为“损失”。

解：设$\boldsymbol{x}$ =“实验反应为阳性”。

(1)类别状态ω有两种：$\omega=\omega_1$正常；$\omega=\omega_2$病患

(2)根据已知，计算后验概率。

由条件1，状态先验概率：$\boldsymbol{P}(\omega_1)=0.995, \boldsymbol{P}(\omega_2)=0.005$

由条件2，条件概率：$\mathrm{P}(\boldsymbol{x}\,|\,\omega_1)=0.01, \boldsymbol{P}(\boldsymbol{x}\,|\,\omega_2)=0.95$

后验概率：$\boldsymbol{P}(\omega_1\,|\,\boldsymbol{x})=\dfrac{\boldsymbol{P}(\omega_1)\boldsymbol{P}(\boldsymbol{x}\,|\,\omega_1)}{\boldsymbol{P}(\omega_1)\boldsymbol{P}(\boldsymbol{x}\,|\,\omega_1)+\boldsymbol{P}(\omega_2)\boldsymbol{P}(\boldsymbol{x}\,|\,\omega_2)}=0.677$

$$\boldsymbol{P}(\omega_2\,|\,\boldsymbol{x})=1-0.677=0.323$$

Software College of Hebei Normal University

(3)计算条件风险：

$$\boldsymbol{R}(\alpha_1\,|\,\boldsymbol{x})=\boldsymbol{P}(\omega_1\,|\,\boldsymbol{x})\lambda_{\mathbf{11}}+\boldsymbol{P}(\omega_2\,|\,\boldsymbol{x})\lambda_{\mathbf{12}}=0.677\times0.5+0.323\times6=2.2765$$

$$\boldsymbol{R}(\alpha_2\,|\,\boldsymbol{x})=\boldsymbol{P}(\omega_1\,|\,\boldsymbol{x})\lambda_{\mathbf{21}}+\boldsymbol{P}(\omega_2\,|\,\boldsymbol{x})\lambda_{\mathbf{22}}=0.677\times2+0.323\times0.5=1.5155$$

$$\boldsymbol{R}(\alpha_2\,|\,\boldsymbol{x})<\boldsymbol{R}(\alpha_1\,|\,\boldsymbol{x})$$

由最小风险判决，″阳性反应的甲″判决为ω_2（病患），判决风险：1.5155

最小风险贝叶斯决策-小结

- 最小风险贝叶斯决策意义
- 最小风险贝叶斯决策规则
- 基于最小风险贝叶斯决策规则，对观测X的分类步骤