模式识别
第三章 概率密度函数的估计

贝叶斯决策的实际使用

* 计算贝叶斯后验概率进行决策如何实现?

$$P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{\sum_{j=1}^{2} p(x|\omega_j)P(\omega_j)}$$

* 实现中有问题吗?能直接计算吗?

贝叶斯决策的实际使用

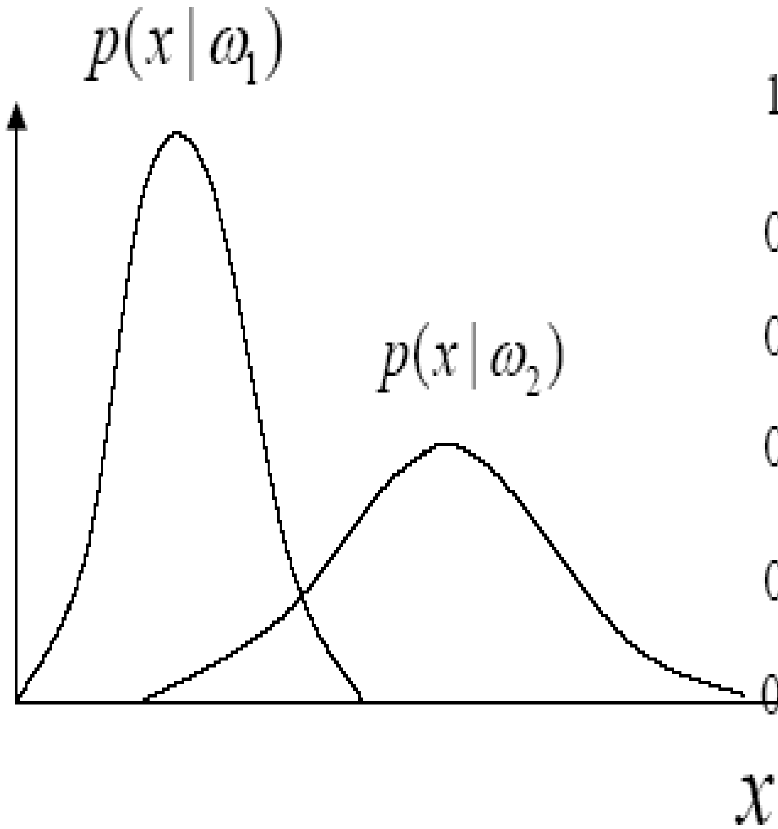
* 以第二章医生诊病为例:

$$P(\omega_1), P(\omega_2)$$

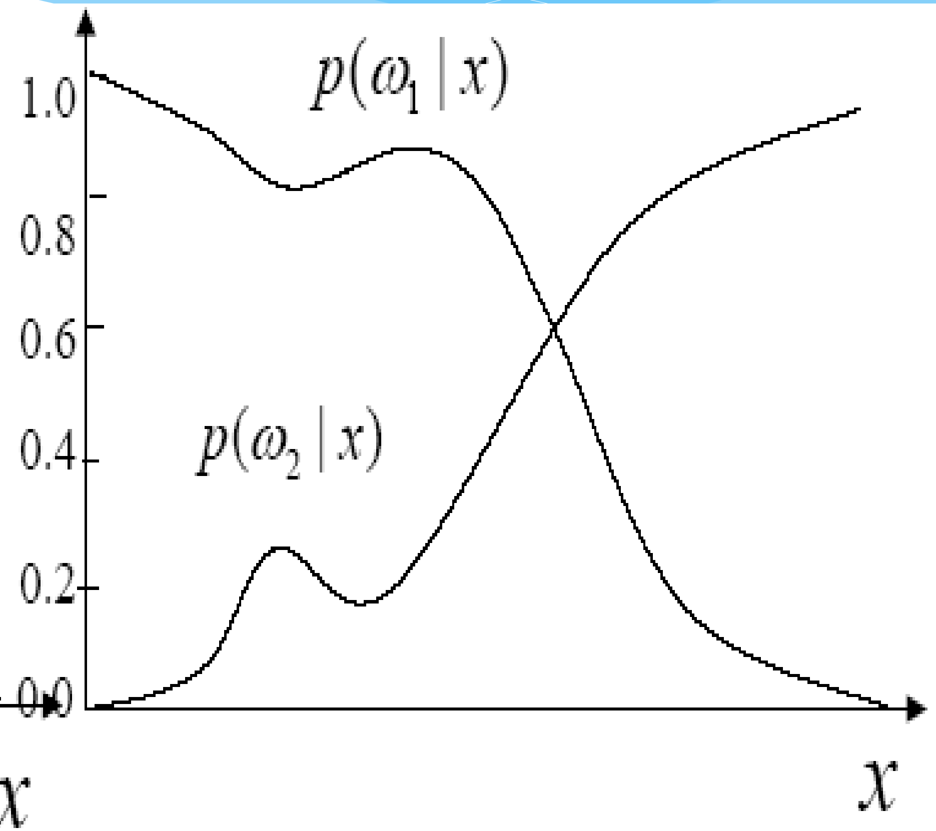
$$P(x|\omega_1), P(x|\omega_2)$$

* 先验概率未知,医生可大致估计
* 类条件概率密度呢?
* 只能通过已知的一些数据去估计

基于样本的两步贝叶斯决策



类条件概率密度函数



后验概率

基于样本的两步贝叶斯决策

* 利用样本集估计$P(\omega_i)$和$\mathrm{p}(x|\omega_i)$
* 得到$\widehat{P}(\omega_i)$和$\widehat{\mathrm{p}}(x|\omega_i)$
* 然后将估计量带入贝叶斯决策规则中
* 得到决策结果

基于样本的两步贝叶斯决策

* 首先通过训练样本估计概率密度函数
  * 先验概率估计－训练样本分布情况/根据领域知识认定
  * 类条件概率密度估计难，本章重点
* 统计决策进行类别判定

基于样本的两步贝叶斯决策

* 训练样本有限，难以涵盖所有情况！
* 当训练样本无限多，基于样本的两步贝叶斯决策可无限趋近第二章中的理论贝叶斯决策

当$\mathbf{N} \rightarrow \infty$

$$\hat{P}(\omega_i) \rightarrow P(\omega_i)$$
$$\hat{\mathrm{p}}(x|\omega_i) \rightarrow \mathrm{p}(x|\omega_i)$$

概率密度估计方法

* 由训练样本集估计总体概率密度的方法可分为：
  * **监督参数估计**：样本所属类别及类条件总体概率密度函数形式已知，表征概率密度函数的某些参数未知。
  * **非监督参数估计**：已知总体概率密度函数形式但未知样本所属类别，要求推断出概率密度函数的某些参数。
  * **非参数估计**：已知样本所属类别,未知总体概率密度函数的形式，直接推断概率密度函数本身。

参数估计的基本概念

* 统计量：样本中包含着总体的信息，针对不同要求构造出样本的某种函数，这种函数在统计学中称统计量。
* 参数空间：假设总体概率密度函数形式已知，未知分布中的参数θ，θ的全部可容许值组成的集合称为参数空间，记为Θ

参数估计的基本概念

* 点估计：点估计问题就是要构造一个统计量$d(x_1, x_2, \ldots, x_N)$作为参数θ的估计$\hat{\theta}$。
* 估计量：$\hat{\theta}$被称为θ的估计量。
* 估计值：如果$x_1^i, x_2^i, \ldots, x_N^i$是属于类别$\omega_i$的样本的观察值，代入统计量d就得到对于第i类的$\hat{\theta}$的具体数值，这一数值被称为$\theta$的估计值。

参数估计的基本概念

* 区间估计: 估计出区间(d_1, d_2)作为θ可能的取值范围，这个区间叫置信区间，这类问题称为区间估计。

概率密度估计的评估

* 如何评估概率密度估计的好坏?
* 单次抽样得到的估计值与真实值的偏差?
* 基于平均和方差进行评估较为公平!
* 常用标准:
  * 无偏性
  * 有效性
  * 一致性

概率密度估计的评估

* 无偏性：$\boldsymbol{\theta}$的估计量$\widehat{\boldsymbol{\theta}}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N)$的数学期望是$\boldsymbol{\theta}$
* 渐进无偏：N趋于无穷时估计具有无偏性
* 有效性：一种估计比另一种的方差小，此种估计更有效
* 对于任意正数ε，有

$$\lim_{n\to\infty} P\left(\left|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right| > \varepsilon\right) = 0$$

则$\widehat{\boldsymbol{\theta}}$是$\boldsymbol{\theta}$的一致估计

最大似然估计
基本假设

- **Maximum Likelihood Estimation.**
- 参数θ是确定（非随机）的而未知的量。
- 按类别把样本集分开，$\Re_j$类中的每个样本都是独立地从概率密度为$\mathbf{p}(x|\omega_j)$的总体中独立地抽取出来的－独立同分布。

最大似然估计
基本假设

- 类条件概率密度$\mathrm{p}(x|\omega_j)$为已知分布，参数向量未知

$$\mathrm{p}(x|\omega_j) \rightarrow \mathrm{p}(x|\omega_j, \theta_j)$$

- 假设$\mathfrak{R}_i$中不包含关于$\theta_j(j \neq i)$的信息，即不同类别的参数在函数上是独立的。

最大似然估计
似然函数的定义

- 在一类中独立抽取样本集来估计未知参数

- 假设某类样本集中包含有N个样本

$$\Re = \{x_1, x_2, \ldots, x_N\}$$

- 因样本独立抽取，样本出现在样本集中的概率为

$$P(\Re|\theta) = p(x_1, x_2, \ldots, x_N|\theta) = \prod_{k=1}^{N} p(x_k|\theta)$$

最大似然估计
似然函数的定义

* $P(\Re|\theta)$也被称为相对于样本集$\Re$的θ的似然函数(likelihood function)
* 似然函数的定义

$$l(\theta) = p(x_1, x_2, \ldots, x_N|\theta)$$
$$= p(x_1|\theta)\,p(x_2|\theta)\ldots p(x_N|\theta)$$

最大似然估计
似然函数的定义

* 似然函数示例
* 假设$N=1$，x为一维，均值为6，方差为1的正态分布
* 最可能出现的样本$x_1'=6$
* 此时，似然函数为

$$l(\theta)=p(x_1'|6,1)=\max_x p(x|6,1)$$

最大似然估计
最大似然估计量

* 令$l(\theta)$为样本集$\Re$的似然函数

$$\Re = \{x_1, x_2, \dots, x_N\}$$

* 如果

$$\hat{\theta} = d(\Re) = d(x_1, x_2, \dots, x_N)$$

$\hat{\theta}$是参数空间中能使似然函数最大化的θ

则$\hat{\theta}$是θ的最大似然估计量

最大似然估计
最大似然估计的求解

* 假设第i类样本的类条件概率密度:
  $$p(x^i|\omega_i) = p(x^i|\omega_i, \theta^i) = p(x^i|\theta^i)$$
* 属于i类的训练样本为$X^i = (x_1, x_2, \dots, x_N)$
* 求θ^i的最大似然估计即为将$p(x^i|\theta^i)$看成θ^i的函数，求使它最大时θ^i的值

最大似然估计
最大似然估计的求解

* 由于训练样本独立从总体样本集中抽取的

$$P\left(X^i|\omega_i,\theta^i\right)=P\left(X^i|\theta^i\right)=\prod_{k=1}^{N}p\left(x_k|\theta^i\right)$$

* 对N个训练样本出现概率的乘积取对数：

$$\log\prod_{k=1}^{N}p\left(x_k|\theta^i\right)=\sum_{k=1}^{N}\log p\left(x_k|\theta^i\right)$$

最大似然估计
最大似然估计的求解

* 对θ^i求导，令其等于0

$$\begin{bmatrix} \frac{\partial}{\partial_{\theta_1}} \\ \cdots \\ \frac{\partial}{\partial_{\theta_p}} \end{bmatrix} \sum_{k=1}^{N} logP\left(x_k|\theta^i\right) = 0$$

最大似然估计
最大似然估计的求解

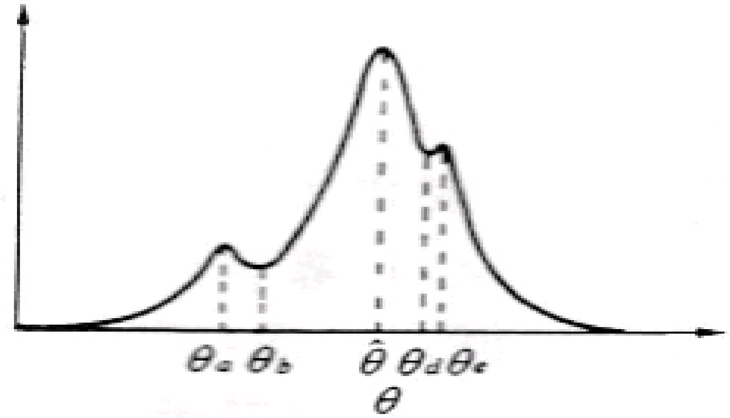
$$\sum_{k=1}^{N}\frac{\partial}{\partial_{\theta_1}}logP\left(x_k|\theta^i\right)=0$$

$$\cdots$$

$$\sum_{k=1}^{N}\frac{\partial}{\partial_{\theta_p}}logP\left(x_k|\theta^i\right)=0$$

最大似然估计
最大似然估计的求解

* 利用上式可以求出θ^i的估值$\hat{\theta}$
* 则 $\theta^i = \hat{\theta}$
* 注意：有时上式是多解的
* 例： 此图有5个解,只有一个解最大



最大似然估计
多维正态分布的最大似然估计

* 情况1：∑已知，μ未知，估计μ
* $P\left(X^i|\theta^i\right)$服从正态分布
* 待估计参数为$\theta^i = \theta_1 = \mu$
* 求解方程

$$\sum_{k=1}^{N} \frac{\partial}{\partial \mu} \log p(x_k|\mu) = 0$$

最大似然估计
多维正态分布的最大似然估计

* 正态分布时

$$\boldsymbol{p}(\boldsymbol{x}_k|\boldsymbol{\mu}) = -\frac{1}{2}\log[(2\boldsymbol{\pi})^n|\boldsymbol{\Sigma}|] - \frac{1}{2}(\boldsymbol{x}_k - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_k - \boldsymbol{\mu})$$

* 求偏导

最大似然估计
多维正态分布的最大似然估计

* 所以

$$\Sigma^{-1}\sum_{k=1}^{N}(x_k - N\mu) = 0$$

$$\mu = \hat{\mu} = \frac{1}{N}\sum_{k=1}^{N} x_k$$

* 说明未知均值的最大似然估计正好是训练样本的算术平均

最大似然估计
多维正态分布的最大似然估计

- 情况2：∑，μ均未知
- 一维时，n=1对于每个训练样本只有一个特征的简单情况：

$$\theta_1 = \mu_1, \theta_2 = \sigma_1^2$$

则：

$$logP\left(x_k|\theta^i\right) = -\frac{1}{2}log2\pi\theta_2 - \frac{1}{2\theta_2}(x_k - \theta_1)^2$$

最大似然估计
多维正态分布的最大似然估计

* 代入$\sum_{k=1}^{N}\frac{\partial}{\partial_{\theta_1}}logP\left(x_k|\theta^i\right)=\sum_{k=1}^{N}\frac{1}{\theta_2}(x_k-\theta_1)=0$

$$\sum_{k=1}^{N}\frac{\partial}{\partial_{\theta_2}}logP\left(x_k|\theta^i\right)=\sum_{k=1}^{N}[-\frac{1}{2\theta_2}+\frac{(x_k-\theta_1)^2}{2\theta_2^2}]=0$$

所以：$\hat{\theta}_1=\widehat{\mu_1}=\frac{1}{N}\sum_{k=1}^{N}x_k$ 训练样本的算术平均

$$\hat{\theta}_2=\widehat{\sigma_1^2}=\frac{1}{N}\sum_{k=1}^{N}(x_k-\hat{\mu})^2$$ 样本方差

最大似然估计
多维正态分布的最大似然估计

* 一维的结论：
  * 正态总体均值的最大似然估计即为训练样本的算术平均
  * 正态总体方差的最大似然估计与样本的方差不同(一个是算术平均，一个是期望值），当N较大的时候，二者的差别不大。
* 注意，这两个也正是人们经常使用的对于均值和方差的估计！

最大似然估计
多维正态分布的最大似然估计

* 多维情况：N个特征
* 与一维时的情况类似
* 估计值：

$$\hat{\theta}_1 = \hat{\mu} = \frac{1}{N}\sum_{k=1}^{N} x_k$$

$$\hat{\theta}_2 = \hat{\Sigma} = \frac{1}{N}\sum_{k=1}^{N} (x_k - \hat{\mu})(x_k - \hat{\mu})^T$$

最大似然估计
多维正态分布的最大似然估计

* 结论：
* μ的估计即为训练样本的算术平均
* 估计的协方差矩阵是如下矩阵的算数平均

$$(x_k - \hat{\mu})(x_k - \hat{\mu})^T$$

最大似然估计
多维正态分布的最大似然估计

* 这两个估计是否为无偏的估计量?

* θ_1为无偏估计量
* θ_2并不是无偏的

贝叶斯估计和贝叶斯学习

* Bayesian Estimation
* 贝叶斯估计实质：贝叶斯决策来决策参数的取值
* 与最大似然估计同为概率密度估计中的主要参数估计方法
* 结果多数情况下与最大似然估计相同
* 区别：
  * 最大似然估计把待估计的参数当作未知但固定的量
  * 贝叶斯估计把待估计的参数也看为随机变量

贝叶斯估计和贝叶斯学习

- **Bayesian Learning**
- 把贝叶斯估计的原理用于直接从数据对概率密度函数进行迭代估计

贝叶斯估计和贝叶斯学习
最小风险贝叶斯决策回顾

* 回顾最小风险贝叶斯决策

$$R(a_i|x) = E\left[\lambda(a_i, \omega_j)\right]$$

$$= \sum_{j=1}^{c} \lambda(a_i, \omega_j) P(\omega_j|x), i = 1, 2, \dots, a$$

$$R(a_k|x) = min_{i=1,\dots,a} R(a_i|x)$$

* 则a_k就是最小风险贝叶斯决策

贝叶斯估计和贝叶斯学习
贝叶斯估计量

- $R(\hat{\theta}|x)$为给定x条件下估计量$\hat{\theta}$的期望损失，称为条件风险
- 定义：如果θ的估计量$\hat{\theta}$使得条件风险最小，则称$\hat{\theta}$是θ的贝叶斯估计量

贝叶斯估计和贝叶斯学习
损失函数

* 决策分类时我们需要事先定义决策风险表即损失表
* 估计连续随机变量时我们需要定义损失函数
* 损失函数有许多种，最常见的损失函数为平方误差损失函数

$$\lambda(\hat{\theta},\theta)=(\hat{\theta}-\theta)^2$$

贝叶斯估计和贝叶斯学习
贝叶斯估计

* 定理:
* 如果损失函数为二次函数，即

$$\lambda(\hat{\theta},\theta)=(\hat{\theta}-\theta)^2$$

* 则θ的贝叶斯估计量$\hat{\theta}$为在给定x时θ的条件期望，如下

$$\hat{\theta}=E[\theta|x]=\int_{\Theta}\theta p(\theta|x)d\theta$$

贝叶斯估计和贝叶斯学习
参数估计问题

- 假设有一个样本集$\mathfrak{R}$，要求我们找出估计量$\hat{\theta}$用来估计$\mathfrak{R}$所属的总体分布的某个真实参数θ使带来的贝叶斯风险最小，这就是贝叶斯估计

贝叶斯估计和贝叶斯学习
贝叶斯估计

- 最大似然估计是把待估的参数看作固定的未知量，而贝叶斯估计则是把待估的参数作为具有某种先验分布的随机变量
- 通过对第i类训练样本X^i的观察，使得概率密度分布$P\left(X^i|\theta\right)$转化为后验概率$P\left(\theta|X^i\right)$，再求贝叶斯估计

贝叶斯估计和贝叶斯学习
贝叶斯估计具体步骤

* 确定θ的先验分布$P(\theta)$,待估参数为随机变量
* 用第i类样本X^i=(X_1, X_2,.... X_N)求出样本的联合概率密度分布$P\left(X^i|\theta\right)$，它是θ的函数
* 利用贝叶斯公式,求θ的后验概率

$$P\left(\theta|X^i\right) = \frac{P\left(X^i|\theta\right)P(\theta)}{\int P(X^i|\theta)P(\theta)d\theta}$$

* 最后求贝叶斯估计

$$\hat{\theta} = \int_{\theta} \theta P\left(\theta|X^i\right)d\theta$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* 一维正态分布:
* 已知σ^2,估计μ

* 假设概率密度服从正态分布
* $P(X|\mu) = N(\mu, \sigma^2)$
* $P(\mu) = N(\mu_0, \sigma_0^2)$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* 第i类样本$X^i=(X_1, X_2,\dots X_N)$

* 后验概率：

$$P(\mu|X^i) = \frac{P(X^i|\mu)P(\mu)}{\int P(X^i|\mu)P(\mu)d\mu}$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* 因为N个样本是独立抽取的：

$$P(\mu|X^i) = a\prod_{k=1}^{N} P(x_k|\mu)P(\mu)$$

其中a为比例因子，只与x有关，与μ无关

$$a = \frac{1}{\int P(X^i|\mu)P(\mu)d\mu}$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* 由

$$P(X|\mu)=N(\mu,\sigma^2)$$
$$P(\mu)=N(\mu_0,\sigma_0^2)$$

* 则

$$P(\mu|X^i)=a\prod_{k=1}^{N}\frac{1}{\sqrt{2\pi}\sigma}exp\left\{-\frac{1}{2}\left(\frac{x_k-\mu}{\sigma}\right)^2\frac{1}{\sqrt{2\pi}\sigma}exp[-\frac{1}{2}(\frac{\mu-\mu_0}{\sigma_0})^2]\right\}$$

$$=a'exp\left\{-\frac{1}{2}\left[\sum_{k=1}^{N}\left(\frac{x_k-\mu}{\sigma}\right)^2+\left(\frac{\mu-\mu_0}{\sigma_0}\right)^2\right]\right\}$$

$$=a''exp\{-\frac{1}{2}[\left(\frac{N}{\sigma^2}+\frac{1}{\sigma_0^2}\right)\mu^2-2(\frac{1}{\sigma^2}\sum_{k=1}^{N}x_k+\frac{\mu_0}{\sigma_0^2})\mu]\}$$

其中a’,a”包含了所有与μ无关的因子

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* $P(\mu|X^i)$是μ的二次函数的指数函数
* $P(\mu|X^i)$仍然是一个正态函数, $P(\mu|X^i) = N(\mu_N, \sigma_N^2)$
* 另外后验概率可以直接写成正态形式:

$$P(\mu|X^i) = \frac{1}{\sqrt{2\pi}\sigma_N} exp[-\frac{1}{2}(\frac{\mu-\mu_N}{\sigma_N})^2]$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* 比较以上两个式子,对应的系数应该相等

$$\frac{1}{\sigma_N^2} = \frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}$$

$$\frac{\mu_N}{\sigma_N^2} = \frac{1}{\sigma^2}\sum_{k=1}^{N} x_k + \frac{\mu_0}{\sigma_0^2}$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

* 求解，可得:

$$\mu_N = \frac{\sigma_0^2}{N\sigma_0^2 + \sigma^2} \sum_{k=1}^{N} x_k + \frac{\sigma^2}{N\sigma_0^2 + \sigma^2} \mu_0$$

$$\sigma_N^2 = \frac{\sigma_0^2 \sigma^2}{N\sigma_0^2 + \sigma^2}$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

- 将μ_N和σ_N^2代入$P(\mu|X^i)$可以得到后验概率，再用$\hat{\theta}=\int_\theta \theta P(\theta|X^i)d\theta$求$\mu$的估计

$$\hat{\mu}=\int \mu P(\mu|X^i)d\mu=\mu_N \text{ (因为正态分布)}$$

$$\widehat{\mu_N}=\mu_N=\frac{\sigma_0^2}{N\sigma_0^2+\sigma^2}\sum_{k=1}^{N}x_k+\frac{\sigma^2}{N\sigma_0^2+\sigma^2}\mu_0$$

贝叶斯估计和贝叶斯学习
正态分布的均值估计

$$P(\mu_N) = N(\mu_0, \sigma_0) = N(0, 1)$$

所以：

$$\widehat{\mu_N} = \frac{1}{N+1}\sum_{k=1}^{N} x_k$$

* 与最大似然估计结果相似，分母不同

贝叶斯估计和贝叶斯学习
贝叶斯学习

* 前面学习了两种参数估计的方法．最终目的是估计总体分布

$$P(x|\mathfrak{R}) \quad \mathfrak{N} \to X^{i}$$

贝叶斯估计和贝叶斯学习
贝叶斯估计的另一计算方式

* 确定θ的先验分布P(θ),待估参数为随机变量
* 用第i类样本$X^i=(X_1, X_2, \ldots X_N)$求出样本的联合概率密度分布$P(X^i|\theta)$，它是θ的函数
* 利用贝叶斯公式,求θ的后验概率

$$P\left(\theta|X^i\right) = \frac{P\left(X^i|\theta\right)P(\theta)}{\int_\theta P(X^i|\theta)P(\theta)d\theta}$$

* 最后求贝叶斯估计

$$\hat{\theta} = \int_\theta \theta P\left(\theta|X^i\right)d\theta$$

贝叶斯估计和贝叶斯学习
贝叶斯估计的另一计算方式

* 我们在第三步后可以直接通过联合密度求类条件概率密度

$$p(x|X^i)\int p(x,\theta|X^i)\ d\theta = \int p(x|\theta)p(\theta|X^i)d\theta$$

贝叶斯估计和贝叶斯学习
非监督参数估计

* 在未知样本类别的条件下的参数估计称为非监督参数估计.
* 几个基本假设:
  * 样本来自类数为c的各类中，但不知道每个样本究竟来自哪一类
  * 每类的先验概率$P(\omega_j), j = 1, 2, \dots, c$已知
  * 类条件概率密度的形式$\mathrm{p}(x|\omega_j, \theta_j)$已知
  * 未知的只是c个参数向量$\theta_1, \theta_2, \dots, \theta_c$的值

贝叶斯估计和贝叶斯学习
非监督参数估计

* 似然函数：为样本集的联合密度

$$\boldsymbol{l(\theta) = p(\mathfrak{R}|\theta)}$$

* 假设每个样本的类别未知，但样本从混合密度为

$$\boldsymbol{p(x|\theta) = \sum_{j=1}^{c} p(x|\omega_j, \theta_j)P(\omega_j)}$$

* 的总体中独立被抽出来，被观察函数的似然函数定义为

$$\boldsymbol{l(\theta) = p(\mathfrak{R}|\theta) = \prod_{k=1}^{N} p(x_k|\theta)}$$

贝叶斯估计和贝叶斯学习
非监督参数估计

* 对数似然函数为：

$$H(\theta) = \ln[\, l(\theta)] = \sum_{k=1}^{N} \ln p\,(x_k|\theta)$$

* 最大似然估计$\hat{\boldsymbol{\theta}}$就是使似然函数最大的$\boldsymbol{\theta}$满足

$$\boldsymbol{l(\hat{\theta}) = \max_{\theta \in \Theta} \prod_{k=1}^{N} p(x_k|\theta)}$$

$$\boldsymbol{H(\hat{\theta}) = \underset{\theta \in \Theta}{max} \sum_{k=1}^{N} ln[\,p(x_k|\theta)]}$$

贝叶斯估计和贝叶斯学习
非监督参数估计

* 如何求得估计量?

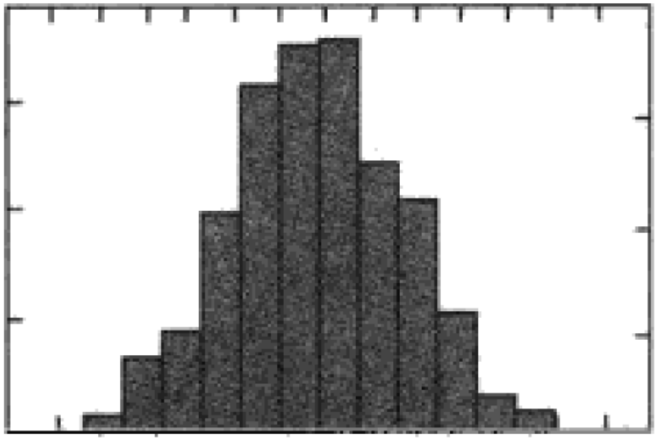
$$\nabla_{\theta_i} H(\theta) = \sum_{k=1}^{N} \frac{1}{p(x_k|\theta)} \nabla_{\theta_i} [\sum_{j=1}^{c} p(x_k|\omega_j, \theta_j) P(\omega_j)] = 0$$

贝叶斯估计和贝叶斯学习
非参数设计

* 多数情况下我们对样本的分布没有充分的了解，无法给出密度函数形式，部分样本情况难以用简单的函数来描述。
* 此时，非参数设计－不对密度函数进行假设，直接利用样本估计整个函数，登场
* 只有数值形式，无封闭函数形式。
* 从所有可能的函数中进行一种选择

贝叶斯估计和贝叶斯学习
非参数设计

* 几种常见的非参数设计
  * 直方图方法
  * K近邻估计方法
  * Parzen窗法

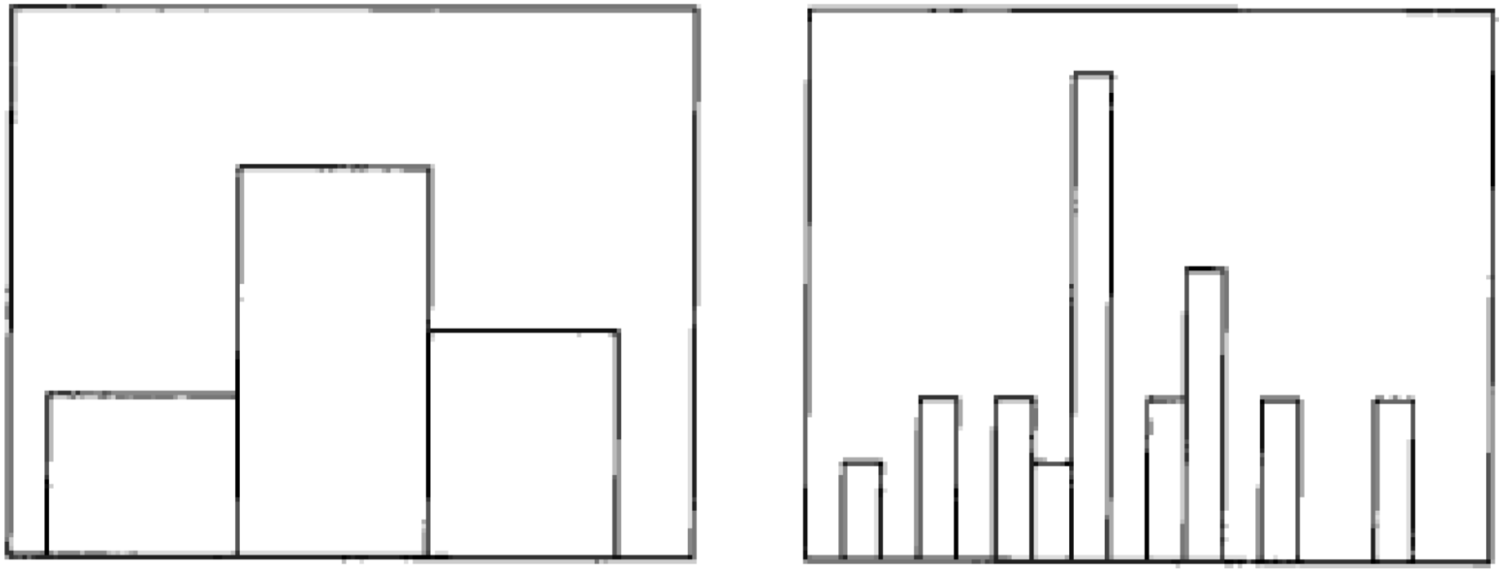


贝叶斯估计和贝叶斯学习
非参数设计－直方图估计

* 直方图方法为日常人们常用之方法
* 过程：
  * 将样本x的每个分量在其取值范围内分割成k个等间隔小窗。如果x是d维向量，则这种分割会得到k^d个小舱，每个小舱体积为V
  * 统计落入每个小舱内的样本数目q_i
  * 把每个小舱的概率密度看为常数，并用$q_i/(NV)$作为其估计值，N为样本总数。

贝叶斯估计和贝叶斯学习
非参数设计－直方图估计

* 小舱的选择与估计效果息息相关



(a) 小窗过宽　　(b) 小窗过窄

大作业1

* 1．学习总体分布的非参数估计
* 2．阅读文章
  * Density Estimation based on Parzen Window
  * K-nearest Algorithm
* 3．实现PARZEN窗估计和K近邻算法(非分类)
* 4．以PPT形式口头报告，包括两种算法的原理实现

谢谢