深度之眼
deepshare.net

数学基础—概率论

导师：Johnson

主要内容

- **概率论**
  - **概率基础**
    - 1.随机试验样本空间和随机事件
    - 2.概率的定义
    - 3.条件概率与乘法公式
    - 4.全概率公式与贝叶斯公式以及应用
    - 5.独立性
  - **随机变量与多维随机变量**
    - 1.离散随机变量
    - 2.连续随机变量
    - 3.多维随机变量
    - 4.边缘分布和条件分布
    - 5.独立性
  - **期望与方差**
    - 1.期望与方差的定义
    - 2.常用随机变量期望与方差的计算
    - 3.期望与方差的性质
    - 4.协方差与协方差矩阵
  - **参数的估计**
    - 1.矩估计
    - 2.极大似然估计
    - 3.最大后验估计

主要内容

- 概率论
  - 概率基础
    - 1.随机试验样本空间和随机事件
    - 2.概率的定义
    - 3.条件概率与乘法公式
    - 4.全概率公式与贝叶斯公式以及应用
    - 5.独立性
  - 随机变量与多维随机变量
    - 1.离散随机变量
    - 2.连续随机变量
    - 3.多维随机变量
    - 4.边缘分布和条件分布
    - 5.独立性
  - 期望与方差
    - 1.期望与方差的定义
    - 2.常用随机变量期望与方差的计算
    - 3.期望与方差的性质
    - 4.协方差与协方差矩阵
  - 参数的估计
    - 1.矩估计
    - 2.极大似然估计
    - 3.最大后验估计

矩估计

深度之眼
deepshare.net

设X为连续型随机变量，其概率密度为$f(x;\theta_1,\theta_2,\cdots,\theta_k)$，或$X$为离散型随机变量，其分布律为$P\{X=x\}=p(x;\theta_1,\theta_2,\cdots,\theta_k)$，其中$\theta_1,\theta_2,\cdots,\theta_k$为待估计参数，$X_1,X_2,\cdots,X_n$是来自$X$的样本. 假设总体$X$的前$k$阶矩

$$\mu_l=E(X^l)=\int_{-\infty}^{\infty}x^l f(x;\theta_1,\theta_2,\cdots,\theta_k)dx(X\text{连续型})$$

$$\mu_l=E(X^l)=\sum_{x\in R_X}x^l p(x;\theta_1,\theta_2,\cdots,\theta_k)(X\text{离散型})$$

$$l=1,2,\cdots,k$$

样本矩 $A_l=\frac{1}{n}\sum_{i=1}^{n}X_i^l$

矩估计

例1 设总体X在$[a,b]$上服从均匀分布，a,b未知，$X_1,X_2,\cdots,X_n$ 是来自X的样本，试求a,b的矩估计量.

解 $\mu_1=E(X)=(a+b)/2,$

$$\mu_2=E(X^2)=D(X)+[E(X)]^2$$

$$=(b-a)^2/12+(a+b)^2/4.$$

$$即\begin{cases}a+b=2\mu_1,\\ b-a=\sqrt{12(\mu_2-\mu_1^2)}\end{cases}.$$

解这一方程组得

$$a=\mu_1-\sqrt{3(\mu_2-\mu_1^2)},b=\mu_1+\sqrt{3(\mu_2-\mu_1^2)}.$$

分别以A_1,A_2代替μ_1，μ_2，得到a,b的矩估计量分别为(注意到

$$\frac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2):$$

$$\hat{a}=A_1-\sqrt{3(A_2-A_1^2)}=\overline{X}-\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2},$$

$$\hat{b}=A_1+\sqrt{3(A_2-A_1^2)}=\overline{X}+\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}.$$

矩估计

深度之眼 deepshare.net

例2 设总体X的均值μ及方差σ^2都存在，且有$\sigma^2 > 0$.但 μ， σ^2均为未知.

又设 $X_1, X_1, \cdots, X_n$ 是来自X的样本. 试求μ， σ^2的矩估计量.

解

$$\begin{cases} \mu_1 = E(X) = \mu, \\ \mu_2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2. \end{cases}$$

$$\text{解得} \begin{cases} \mu = \mu_1, \\ \sigma^2 = \mu_2 - \mu_1^2. \end{cases}$$

分别以A_1， A_2代替μ_1， μ_2，得μ和σ^2的矩估计量分别为

$$\hat{\mu} = A_1 = \overline{X},$$

$$\widehat{\sigma^2} = A_2 - A_1^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

极大似然估计

设 $X_1, \cdots, X_n$ 是来自X的样本

$$L(\theta) = L(x_1, x_2, \cdots x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta)$$

$$L(\theta) = L(x_1, x_2, \cdots x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

$$\frac{d}{d\theta} L(\theta) = 0$$

$$\frac{d}{d\theta} \ln L(\theta) = 0$$

极大似然估计

例1 设 $X \sim b(1,p).X_1,\ \cdots,X_n$ 是来自X的一个样本，试求参数P的最大似然估计量.

解 设$x_1,x_2,\cdots x_n$是相应于样本$X_1,\ \cdots,X_n$的一个样本值.

X的分布律为 $P\{X=x\}=p^x(1-p)^{1-x},x=0,1.$

故似然函数为

$$L(p)=\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}=p^{\sum\limits_{i=1}^{n} x_i}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}$$

而

$$\ln L(p)=(\sum_{i=1}^{n} x_i)\ln p+(n-\sum_{i=1}^{n} x_i)\ln(1-p),$$

令

$$\frac{d}{dp}\ln L(p)=\frac{\sum\limits_{i=1}^{n} x_i}{p}-\frac{n-\sum\limits_{i=1}^{n} x_i}{1-p}=0.$$

解得 p 的最大似然估计值

$$\hat{p}=\frac{1}{n}\sum_{i=1}^{n} x_i=\bar{x}.$$

极大似然估计

例2 设 $x\sim N(\mu,\sigma^2)$，μ,σ^2 为未知参数，$x_1,x_2,\cdots x_n$ 是来自X的一个样本值.
求μ，σ^2的最大似然估计量.

解 X的概率密度为 $f(x;\mu,\sigma^2)=\dfrac{1}{\sqrt{2\pi\sigma}}\exp[-\dfrac{1}{2\sigma^2}(x-\mu)^2]$,

似然函数为 $L(\mu,\sigma^2)=\prod_{i=1}^{n}\dfrac{1}{\sqrt{2\pi\sigma}}\exp[-\dfrac{1}{2\sigma^2}(x_i-\mu)^2]$

$$=(2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2]$$

而 $\ln L=-\dfrac{n}{2}\ln(2\pi)-\dfrac{n}{2}\ln\sigma^2-\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$

令
$$\begin{cases}\dfrac{\partial}{\partial\mu}\ln L=\dfrac{1}{\sigma^2}(\sum_{i=1}^{n}x_i)-n\mu=0,\\[2ex] \dfrac{\partial}{\partial\sigma^2}\ln L=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2=0.\end{cases}$$

由前一式解得

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$$

代入后一式得

$$\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

极大似然估计

例3 设总体X在$[a,b]$上服从均匀分布，a,b未知，$x_1,x_2,\cdots,x_n$是一个样本值，试求a,b的最大似然估计量.

解 记 $x_{(1)}=\min\{x_1,x_2,\cdots,x_n\},x_{(n)}=\max\{x_1,x_2,\cdots,x_n\}.$

X的概率密度是

$$f(x;a,b)=\begin{cases}\dfrac{1}{b-a},a\le x\le b,\\ 0,\text{其他}.\end{cases}$$

似然函数为 $L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n},a\le x_1,x_2,\cdots,x_n\le b,\\ 0,\text{其他}.\end{cases}$

由于$a\le x_1$，x_2，…，$x_n\le b$,等价于$a\le x_{(1)}$，$x_{(n)}\le b$,似然函数可写成

$$L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a\le x_{(1)}, \quad b\ge x_{(1)}\\ 0, & \text{其他}\end{cases}$$

于是对于满足条件$a\le x_{(1)},b\ge x_{(n)}$的任意$a,b$有

$$L(a,b)=\frac{1}{(b-a)^n}\le\frac{1}{\left(x_{(n)}-x_{(1)}\right)^n}$$

即$L(a,b)$ 在$a=x_{(1)},b=x_{(n)}$时取到最大值$\left(x_{(n)}-x_{(1)}\right)^{-n}$.

故a,b的最大似然估计值为$\hat{a}=x_{(i)}=\min\limits_{1\le i\le n}x_i$，$\hat{b}=x_{(n)}=\max\limits_{1\le i\le n}x_i$

a,b的最大似然估计量为$\hat{a}=\min\limits_{1\le i\le n}X_i$，$\hat{b}=\max\limits_{1\le i\le n}X_i$

最大后验估计

最大似然估计认为使似然函数最大的参数θ即为最好的θ，此时最大似然估计是将θ看作固定的值，只是其值未知；

最大后验概率分布认为θ是一个随机变量，即θ具有某种概率分布，称为先验分布，求解时除了要考虑似然函数$P(X|\theta)$之外，还要考虑θ的先验分布$P(\theta)$，因此其认为使$P(X|\theta)P(\theta)$取最大值的θ就是最好的θ。

用数学公式表达：$\underset{\theta}{argmax}\, P(\theta|X) = \underset{\theta}{argmax}\, P(X|\theta)P(\theta)$

频率派为极大似然估计：$\underset{\theta}{argmax}\, P(X|\theta)$

最大后验估计

深度之眼
deepshare.net

西瓜价格问题估计的两个门派的区别

频率派：

假设去不同超市调研西瓜的价格得到5组数据(2.0，1.8， 2.2，1.9，2.1)

价格价格X~高斯分布，则可以用极大似然估计出西瓜价格均值为2.0元

贝叶斯派

假设去不同超市调研西瓜的价格得到5组数据(2.0，1.8， 2.2，1.9，2.1) （2019年）

同时我还会去统计2019年之前的西瓜价格

2018	1.5
2017	1.3
2016	1.2
2015	1.1
2014	1.2

$$P(u) = \frac{1}{\sqrt{2\pi}0.1} e^{-\frac{(u-1.26)^2}{0.02}}$$

$$P(u|x) = P(x|u)P(u)$$

$$= \left\{ \prod_{i=1}^{5} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[- \frac{(x_i - u)^2}{2\sigma^2} \right] \right\} \frac{1}{\sqrt{2\pi}0.1} e^{-\frac{(u-1.26)^2}{0.02}}$$

深度之眼
deepshare.net

联系我们:

电话：18001992849

邮箱：service@deepshare.net

Q Q： 2677693114

公众号

客服微信