数 理 统 计

开课院系：数学与统计学院概率与统计系

教师： 成灵妍 副教授

邮箱： cly@njust.edu.cn

上节回顾

充分统计量

设总体X的分布族为$\{p(x;\theta),\theta\in\Theta\}$，$X_1,X_2,\ldots,X_n$是$F(x;\theta)$的样本，$T(X_1,X_2,\ldots,X_n)$是一个统计量.

若在已知T的条件下，$(X_1,X_2,\ldots,X_n)$的条件分布与θ无关，则称T为参数θ的充分统计量.

因子分解定理

T是参数θ的充分统计量当且仅当$p_\theta(x_1,x_2,\ldots,x_n)$可以分解为

$$p_\theta(x_1,x_2,\ldots,x_n)=g_\theta(T(x_1,x_2,\ldots,x_n))h(x_1,x_2,\ldots,x_n)$$

主观题 1分 设置

习题4：设X_1,X_2是来自$N(0,1)$的样本，试求常数k，使得

$$P\left(\frac{(X_1+X_2)^2}{(X_1-X_2)^2+(X_1+X_2)^2}>k\right)=0.05.$$

作答

单选题 1分 设置

习题5：设总体$X\sim N(\mu,\sigma^2)$，$\overline{X}$为该总体的样本均值，则$P(\overline{X}<\mu)=?$

- A <1/2
- B >1/2
- C =1/2
- D =1/4

提交

主观题 1分 设置

习题6：设总体X服从正态分布$N(0,1)$, $X_1, X_2, \ldots, X_8$是它的简单随机样本，设

$$Y=(X_1+X_2+X_3+X_4)^2+(X_5+X_6+X_7+X_8)^2,$$

试决定常数c,使得cY服从卡方分布.

正常使用主观题需2.0以上版本雨课堂

作答

第七章

参数估计

第七章 参数估计

例子:

某公司要采购一批产品, 每件产品不是合格品就是不合格品, 但是该产品总有一个不合格品率p. 由此, 若从该批产品中随机抽取一件, 用X表示抽出产品的不合格品数, 不难看出X服从二点分布B(1, p), 但分布中的参数p确实不知道的. 显然, p的大小决定了该产品的质量, 它直接影响采购行为的经济效益. 因此, 人们会对p提出一些问题, 比如:

(1)p的大小如何——点估计

(2)p大概落在什么范围内——区间估计

第七章 参数估计

§7.1 参数的点估计

★ **点估计的概念**

定义：设θ为总体**X~F(x; θ)**,θ∈Θ的待估计参数, 若取样本**X_1, ⋯ ,X_n**的一个统计量

$$\hat{\theta}=\hat{\theta}(X_1,\cdots,X_n)$$

来估计θ, 则称 $\hat{\theta}$为θ的一个估计量。称 $\hat{\theta}=\hat{\theta}(x_1,\cdots,x_n)$ 为θ的估计值, 并仍记为$\hat{\theta}$。

§7.1 参数的点估计

★ 点估计的概念

由于估计值 $\hat{\theta}=\hat{\theta}(x_1,\cdots,x_n)$是**R**上的一个点，故称这种估计为点估计。

注：上述定义中，若考虑多个未知参数时，θ也可以看成一个向量，即 $\theta=(\theta_1,\theta_2,\ldots,\theta_k)$，表示有$k$个未知参数，则 $\hat{\theta}_i=\hat{\theta}_i(X_1,X_2,\ldots,X_n)$ 为θ_i的估计量.

§7.1 参数的点估计

★ 点估计的两种方法

1 矩估计法

2 最大似然估计法

§7.1 参数的点估计

★ 矩估计法——皮尔逊（1890）替换原理（矩法）

矩估计法是以样本矩作为相应的总体矩的估计。

样本原点矩$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$，总体原点矩$a_k=E(X^k)$.

即当一个参数可以表达成某些总体矩的函数时，就以样本矩的同一函数作为那个参数的估计。

样本中心矩$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k$，总体中心距$b_k=E((X-EX)^k)$

§7.1 参数的点估计

★ 矩估计的依据原理

回忆:辛钦大数定律

若$\{X_k\}_{k=1,2,\ldots}$为独立同分布随机变量序列，$EX_k=\mu$ 存在，k=1, 2, … 则

$$\overline{X}_n=\frac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{P}\mu$$

§7.1 参数的点估计

矩估计的依据原理

引理：$X_1,\cdots,X_n \overset{iid}{\sim} X$,

$$A_m = \frac{1}{n}\sum_{i=1}^{n} X_i^m \xrightarrow{P} E(X^m),\quad m=1,2,\cdots$$

证明：令 $Y_i = X_i^m$, $i=1,2,\cdots$, **则随机变量序列** $\{Y_i\}_{i=1,2,\cdots}$**独立同分布，且** $E(Y_i)=E(X_i^m), i=1,2,\cdots$

由辛钦大数定理知

$$A_m = \frac{1}{n}\sum_{i=1}^{n} X_i^m \xrightarrow{P} E(X^m),\quad m=1,2,\cdots$$

§7.1 参数的点估计

矩估计的方法与技巧

设$X_1,X_2,\ldots,X_n \sim X, i.i.d.$, X的分布函数为$F(x;\theta_1,\theta_2,\ldots,\theta_k)$,含有$k$个未知参数. 设总体至少存在$k$阶原点矩， i. e. $E(X^k)$存在. 记

$a_l = E(X^l) = g_l(\theta_1,\theta_2,\ldots,\theta_k), l=1,2,\ldots,k.$

若样本的l阶原点矩为$A_l = \frac{1}{n}\sum_{i=1}^{n} X_i^l, l=1,2,\ldots,k,$

§7.1 参数的点估计

矩估计的方法与技巧

则由如下k个方程解出$\theta=(\theta_1,\theta_2,\ldots,\theta_k)$,

$$\begin{cases} a_1 = E(X) = g_1(\theta_1,\theta_2,\ldots,\theta_k) \\ a_2 = E(X^2) = g_2(\theta_1,\theta_2,\ldots,\theta_k) \\ \cdots \\ a_k = E(X^k) = g_k(\theta_1,\theta_2,\ldots,\theta_k) \end{cases}$$

从中解出未知参数$\theta_j, j=1,2,\ldots,k,$

$\theta_j = \theta_j(a_1,a_2,\ldots,a_k)$,再用$A_l$替换$a_l$,得到$\theta$的估计

$\hat\theta_j = \hat\theta_j(A_1,A_2,\ldots,A_k) = \hat\theta_j(X_1,X_2,\ldots,X_n),$

$j=1,2,\ldots,k.$

§7.1 参数的点估计

矩估计的方法与技巧

注意:

(1)有k个未知数就取k个方程从而解出全部未知参数;

(2)未知参数$\theta_1,\cdots,\theta_k$在方程组的右边

(3)解出未知参数后，再用样本矩替换总体矩，从而得到未知参数的估计。

按上述方法求出的$\hat\theta_j, j=1,2,\ldots,k,$

称为未知参数$\theta_j (j=1,2,\ldots,k)$的矩法估计.

§7.1 参数的点估计

★ 矩估计的方法步骤总结（先解方程再替换）

1. 求出总体前k阶原点矩，注意是关于$\theta_1,\theta_2,\ldots,\theta_k$的函数：
$a_l = E(X^l) = g_l(\theta_1,\theta_2,\ldots,\theta_k), l=1,2,\ldots,k;$

2. 从这k个方程中解出$\theta_1,\theta_2,\ldots,\theta_k$：
$\theta_j = \theta_j(a_1,a_2,\ldots,a_k),\ j=1,2,\ldots,k;$

3. 用 A_l替换 a_l，得到 θ的估计
$\hat{\theta}_j = \hat{\theta}_j(A_1,A_2,\ldots,A_k) = \hat{\theta}_j(X_1,X_2,\ldots,X_n),$
$j=1,2,\ldots,k.$

§7.1 参数的点估计

★ 矩估计的方法步骤总结（先替换再解方程）

1. 求出总体前k阶原点矩，注意是关于$\theta_1,\theta_2,\ldots,\theta_k$的函数：
$a_l = E(X^l) = g_l(\theta_1,\theta_2,\ldots,\theta_k), l=1,2,\ldots,k;$

2. 用A_l替换a_l，得到k个方程
$A_l = g_l(\theta_1,\theta_2,\ldots,\theta_k), l=1,2,\ldots,k;$

3. 解出 θ的估计 $\hat{\theta}_j$
$\hat{\theta}_j = \hat{\theta}_j(A_1,A_2,\ldots,A_k) = \hat{\theta}_j(X_1,X_2,\ldots,X_n),$
$j=1,2,\ldots,k.$

§7.1 参数的点估计

★ **例1** 设$X_1,\cdots,X_n \overset{iid}{\sim} U(a,b), a<b$，试求 $\hat{a}_M$ 和 $\hat{b}_M$.

解：$E(X)=\dfrac{a+b}{2}$，　$D(X)=\dfrac{(b-a)^2}{12}$，

$E(X^2)=D(X)+[E(X)]^2=\dfrac{(b-a)^2}{12}+\dfrac{(a+b)^2}{4}$，

用$A_1=\overline{X}$代替$E(X)$，$A_2=\dfrac{1}{n}\sum_{i=1}^n X_i^2$代替$E(X^2)$

得方程 $\begin{cases} A_1=\overline{X}=\dfrac{a+b}{2} \\ A_2=\dfrac{1}{n}\sum_{i=1}^n X_i^2=\dfrac{(b-a)^2}{12}+\dfrac{(a+b)^2}{4} \end{cases}$

解之

$\hat{a}_M=\overline{X}-\sqrt{\dfrac{3}{n}\sum_{i=1}^n (X_i-\overline{X})^2}$，$\hat{b}_M=\overline{X}+\sqrt{\dfrac{3}{n}\sum_{i=1}^n (X_i-\overline{X})^2}$

§7.1 参数的点估计

★ **例2** 设总体**X**的概率密度为 $f(x)=\dfrac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}$

X_1，…，**X_n**为样本，求参数σ^2的矩估计。

解：$E(X)=\int_{-\infty}^{+\infty}\dfrac{x}{2\sigma}e^{-\frac{|x|}{\sigma}}dx=0$

$E(X^2)=\int_{-\infty}^{+\infty}\dfrac{x^2}{2\sigma}e^{-\frac{|x|}{\sigma}}dx=\int_0^{+\infty}\dfrac{x^2}{\sigma}e^{-\frac{x}{\sigma}}dx=2\sigma^2$

令 $A_2=\dfrac{1}{n}\sum_{i=1}^n X_i^2$替换上述方程中的$EX^2$

解得$\hat{\sigma}^2=\dfrac{1}{2n}\sum_{i=1}^n X_i^2$

§7.1 参数的点估计

★ **例3** 设总体X的均值μ，方差σ^2都存在，且$\sigma^2>0$，但μ，σ^2未知，又设$X_1,\cdots,X_n$是一个样本；求：μ,σ^2的矩估计量。

解：因为$EX=\mu$，$EX^2=\sigma^2+\mu^2$

令$A_1=\overline{X}$替换EX，令$A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$替换$EX^2$得

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

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

§7.1 参数的点估计

★ **Remark：1.矩估计法不需要知道总体的分布，只需要知道总体矩就可以。**

2. 二维总体(X,Y)的相关系数$\rho=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$

的矩估计为$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$

3. 频率是概率的矩估计：

因为在伯努利总体$X\sim B(1,p)$中，$E(X)=p$，

而样本均值$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{\mu_n}{n}$就是频率.

§7.1 参数的点估计

★ **思考：矩估计是不是唯一的？**

作业 P377 1

（作业）总体密度函数为$p(x;\theta)=\frac{2}{\theta^2}(\theta-x),0<x<\theta,\theta>0$，$X_1,X_2,\ldots,X_n$样本，求未知参数$\theta$的矩估计.

（思考题）甲乙两校对员彼此独立的对同一本书的样稿进行校对，校完后，甲发现a个错别字，乙发现b个错别字，其中共同发现的错别字有c个，试用矩估计给出下列两个未知参数的估计.

（1）该书样稿的总错别字个数；

（2）未被发现的错字数.

提示：频率替换概率

主观题 1分 设置

设总体X密度函数为

$$f(x;\theta,\mu)=\frac{1}{\theta}e^{-\frac{x-\mu}{\theta}},x>\mu,\theta>0,$$

其中θ,μ是未知参数，

$X_1,X_2,\ldots,X_n$是X的样本，求θ,μ的矩估计.

正常使用主观题需2.0以上版本雨课堂

作答

§7.1 参数的点估计

★ 最大似然估计小故事

§7.1 参数的点估计

★ 最大似然估计（maximum likelihood estimation简写MLE）

例1：设外形完全相同的两个箱子，甲箱子中有99个白球，一个黑球，乙箱子中有99个黑球，一个白球，今随机抽取一箱并从中随机抽取一球，得到一个白球，问这个球是从哪个箱子取出的？

分析：考虑事件A={取出一球为白球}，那么对于甲箱子，P(A)=0.99，而对于乙箱子，P(A)=0.01。现在经过一次试验A发生了，那么此白球最像是从甲箱子抽出来的。这里的最像就是“最大似然”得意思。

表彰榜


塔娜


梅卜玮


姚树仁


陈昌杰

上节回顾

点估计

矩估计

用样本矩替
代总体矩

大数定律

方法步骤！！！一定要会

§7.1 参数的点估计

★ 最大似然估计（maximum likelihood estimation简写MLE）

***例2: 一射手击中目标的概率可能是**

$$p=\frac{1}{5},\ \frac{8}{15},\ \frac{4}{5}.$$

现让他打三发子弹，在不同的命中目标的次数下，我们应该如何取 p 的估计值 $\hat{p}$?

解: 令 X 表示射手命中目标的次数，则 $X\sim B(3,p)$, 即

$$P\{X=x\}=C_3^x p^x(1-p)^{3-x}=P(x;p).\quad x=0,1,2,3$$

计算结果列表如下：

§7.1 参数的点估计

例2(续)

命中次数 x	0	1	2	3
$P(x;\frac{1}{5})$	$\frac{1728}{3375}$	$\frac{1296}{3375}$	$\frac{324}{3375}$	$\frac{27}{3375}$
$P(x;\frac{8}{15})$	$\frac{343}{3375}$	$\frac{1176}{3375}$	$\frac{1344}{3375}$	$\frac{512}{3375}$
$P(x;\frac{4}{5})$	$\frac{27}{3375}$	$\frac{324}{3375}$	$\frac{1296}{3375}$	$\frac{1728}{3375}$

因为 $P(0;\frac{1}{5})>P(0;\frac{8}{15})>P(0;\frac{4}{5})$,

这表明, 若他打三发子弹均未命中目标时, 取

$\hat{p}=\frac{1}{5}$ 最合理.

§7.1 参数的点估计

例2(续)

命中次数 x	0	1	2	3
$P(x;\frac{1}{5})$	$\frac{1728}{3375}$	$\frac{1296}{3375}$	$\frac{324}{3375}$	$\frac{27}{3375}$
$P(x;\frac{8}{15})$	$\frac{343}{3375}$	$\frac{1176}{3375}$	$\frac{1344}{3375}$	$\frac{512}{3375}$
$P(x;\frac{4}{5})$	$\frac{27}{3375}$	$\frac{324}{3375}$	$\frac{1296}{3375}$	$\frac{1728}{3375}$

由上表可得下面的结论：

打三发命中次数 $x=1$ 时，命中率 p 的合理估计 $\hat{p}=\frac{1}{5}$;

打三发命中次数 $x=2$ 时，命中率 p 的合理估计 $\hat{p}=\frac{8}{15}$;

打三发命中次数 $x=3$ 时，命中率 p 的合理估计 $\hat{p}=\frac{4}{5}$.

§7.1 参数的点估计

★ 最大似然估计（maximum likelihood estimation简写MLE）

一般说，事件**A**发生的概率与参数θ∈Θ有关，θ取值不同，则**P(A)**也不同。因而应记事件**A**发生的概率为**P(A;θ)**. 若**A**发生了，则认为此时的θ值应是在Θ中使**P(A;θ)** 达到最大的那一个。这就是最大似然估计思想。

§7.1 参数的点估计

★ 最大似然估计（maximum likelihood estimation简写MLE）

用样本观点解释

设$X_1, X_2, \ldots, X_n$是总体X的简单随机样本, X分布为$P(x;\theta_1,\theta_2,\ldots,\theta_k)$, 事件$A=\{X_1=x_1, X_2=x_2,\ldots,X_n=x_n\}$, $P(A;\theta_1,\theta_2,\ldots,\theta_k)$即观测到样本值为$(x_1,x_2,\ldots,x_n)$的概率, 我们需要求$(\theta_1,\theta_2,\ldots,\theta_k)$使得$P(A;\theta_1,\theta_2,\ldots,\theta_k)$最大.

§7.1 参数的点估计

★ 最大似然估计（离散时）

$\mathrm{P}\{\mathrm{X}=\mathrm{x};\theta\}=\mathrm{p}(\mathrm{x};\theta), \theta\in\Theta$, 则

$$P(X_1=x_1, X_2=x_2,\ldots,X_n=x_n;\theta)$$

$$=\prod_{i=1}^{n}P(X_i=x_i;\theta)=\prod_{i=1}^{n}p(x_i;\theta)$$

$$\arg\max_{\theta\in\Theta}\left\{\prod_{i=1}^{n}p(x_i;\theta)\right\}=\arg\max_{\theta\in\Theta}L(\theta)=\hat{\theta}_{MLE}$$

§7.1 参数的点估计

★ 最大似然估计（连续时）

由于连续随机变量在一点的取值概率为0，因此我们考虑样本$(X_1,X_2,\ldots,X_n)$落在$(x_1,x_2,\ldots,x_n)$小邻域内的概率

$$P(x_1\le X_1<x_1+\Delta x_1,\ldots,x_n\le X_n<x_n+\Delta x_n;\theta)$$

$$=\prod_{i=1}^{n}P(x_i\le X_i<x_i+\Delta x_i;\theta)$$

$$\approx\prod_{i=1}^{n}f(x_i;\theta)\Delta x_i$$

$$\arg\max_{\theta\in\Theta}\left\{\prod_{i=1}^{n}f(x_i;\theta)\Delta x_i\right\}=\arg\max_{\theta\in\Theta}\left\{\prod_{i=1}^{n}f(x_i;\theta)\right\}$$

$$=\arg\max_{\theta\in\Theta}L(\theta)=\hat{\theta}_{MLE}$$

§7.1 参数的点估计

★ 似然函数

$$设X_1,\cdots,X_n\overset{iid}{\sim}\begin{cases}f(x;\theta),(连续)\\ \mathrm{P}\{\mathrm{X}=\mathrm{x};\theta\}=\mathrm{p}(\mathrm{x};\theta),(离散)\end{cases},\quad \theta\in\Theta,\quad 则称$$

$$L(\theta)=L(x_1,\cdots,x_n;\theta)=\begin{cases}\prod_{i=1}^{n}f(x_i;\theta), & （连续）\\ \prod_{i=1}^{n}p(x_i;\theta), & （离散）\end{cases}$$

为该总体的似然函数。

似然函数是样本的联合密度函数或分布律。

§7.1 参数的点估计

最大似然估计的定义

当选取$\hat{\theta}$作为θ的估计时，若有$\hat{\theta}(x_1,\cdots,x_n)\in\Theta$,

使得$L(\hat{\theta})=\max_{\theta\in\Theta} L(\theta)$

即$\hat{\theta}$是使得$L(\theta)$达到最大值的θ的取值，

称$\hat{\theta}$为θ的最大似然估计值，

称$\hat{\theta}(X_1,\cdots,X_n)$为$\theta$的极大似然估计量，都记为$\hat{\theta}_{MLE}$.

§7.1 参数的点估计

求最大似然估计方法与技巧

问题的提法

$P\{X=x;\theta\}=p(x;\theta),\theta\in\Theta$，则

设$X_1,X_2,\ldots,X_n$为X的一个样本，其观测值为

$x_1,x_2,\ldots,x_n$，求$\hat{\theta}_{\text{MLE}}$.（离散时，$X\sim p(x;\theta)$；连续时，

$X\sim f(x;\theta)$.）

§7.1 参数的点估计

求最大似然估计方法与技巧

(1)写出似然函数：

$$L(\theta)=L(x_1,x_2,\ldots,x_n;\theta)=\begin{cases}\prod_{i=1}^{n} p(x_i;\theta),(\text{离散})\\ \prod_{i=1}^{n} f(x_i;\theta),(\text{连续})\end{cases},$$

(2)求θ使得$L(\theta)$达到最大值.

§7.1 参数的点估计

如何求最大似然函数的最大值

注：通常，最大似然函数是可微的

方法：求导！

$$\frac{dL(\theta)}{d\theta}=0 \longrightarrow$$

观察最大似然函数的形式，当样本量比较大时，计算量大且复杂

此时考虑函数$\ln L(\theta)$，由于$\ln x$关于x单调增，所以$L(\theta)$与$\ln L(\theta)$在相同的点处取最大值.

可对$\ln L(\theta)$求导，令$\dfrac{d\ln L(\theta)}{d\theta}=0$.

第七章 参数估计

§7.1 参数的点估计

★ **最大似然估计的通常方法步骤总结**

1 写出似然函数：$L(\theta)=L(x_1,x_2,\dots,x_n;\theta)=\begin{cases}\prod_{i=1}^{n}p(x_i;\theta),(\text{离散})\\ \prod_{i=1}^{n}f(x_i;\theta),(\text{连续})\end{cases}$,

2 做对数似然函数：$\ln L(\theta)=\ln L(x_1,\cdots,x_n;\theta)=\begin{cases}\sum_{i=1}^{n}\ln p(x_i;\theta)\\ \sum_{i=1}^{n}\ln f(x_i;\theta)\end{cases}$

3 求导(列似然方程)

$\frac{d\ln L(\theta)}{d\theta}=0$, 一般若有解，其解就是 $\hat{\theta}_{MLE}$.

第七章 参数估计

§7.1 参数的点估计

★ **最大似然估计的通常方法步骤总结**

注：

(1)当似然函数不可微，或者似然方程无解或者解不唯一时，则需根据具体情况求似然函数的最大值;

(2)当似然函数相对简单时，也可以直接对其求导求出取最大值的点。

第七章 参数估计

§7.1 参数的点估计

★ **例1** 设 $X\sim f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$

$X_1,\cdots,X_n\overset{iid}{\sim}f(x;\lambda)$, 求$\hat{\lambda}_{MLE}$.

解： $L(\lambda)=\prod_{i=1}^{n}f(x_i;\lambda)=\prod_{i=1}^{n}(\lambda e^{-\lambda x_i})=\lambda^n e^{-\lambda\sum_{i=1}^{n}x_i}$

$\ln L(\lambda)=n\ln\lambda-\lambda\sum_{i=1}^{n}x_i$ 令 $\frac{d\ln L(\lambda)}{d\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i=0$

$\therefore\hat{\lambda}=\frac{n}{\sum_{i=1}^{n}X_i}=\frac{1}{\overline{X}}$

第七章 参数估计

§7.1 参数的点估计

★ **例2** 设$X_1,\cdots,X_n\overset{iid}{\sim}B(m,p),0<p<1$，试求$\hat{p}_{MLE}$。

解:$\because L(p)=\prod_{i=1}^{n}\binom{m}{x_i}p^{x_i}(1-p)^{m-x_i}=[\prod_{i=1}^{n}\binom{m}{x_i}]p^{\sum_{i=1}^{n}x_i}(1-p)^{n(m-\bar{x})}$

$\ln L=\ln[\prod_{i=1}^{n}\binom{m}{x_i}]+\sum_{i=1}^{n}x_i\ln p+n(m-\bar{x})\ln(1-p)$

令 $\frac{d\ln L}{dp}=\frac{1}{p}\sum_{i=1}^{n}x_i-\frac{n(m-\bar{x})}{1-p}=0$

解得 $\hat{p}_{MLE}=\frac{1}{m}\overline{X}$

§7.1 参数的点估计

★ **注：**

(1)若总体分布中含有多个未知参数，

$$X_1,\cdots,X_n \overset{iid}{\sim} \begin{cases} p(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k) \\ f(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k) \end{cases}, \quad (\theta_1,\cdots,\theta_k)\in\Theta$$

则可解方程组

$$\frac{\partial L}{\partial \theta_j}=0,\text{或}\ \frac{\partial \ln L}{\partial \theta_j}=0,\ j=1,2,\dots,k,$$

得出 θ_j的$MLE\ \hat{\theta}_j,\ j=1,2,\dots,k.$

§7.1 参数的点估计

★ **注：**

(2)同样，方程组有唯一解时，$\hat{\theta}_j$ 就是 θ_j的MLE.

例3：设X_1，…，X_n为取自$N(\mu,\sigma^2)$总体的样，求参数μ, σ^2 的最大似然估计。

解 $\because L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$

§7.1 参数的点估计

★ **例3(续)** $\ln L=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}$

$$\begin{cases} \dfrac{\partial \ln L}{\partial \mu}=\dfrac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2}=0 \\ \dfrac{\partial \ln L}{\partial \sigma^2}=-\dfrac{n}{2\sigma^2}+\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4}=0 \end{cases}$$

解得：$\hat{\mu}=\bar{X}$, $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2.$

回顾 矩估计的例子

★ **例3** 设总体X的均值μ，方差 σ^2 都存在，且$\sigma^2>0$，但μ，σ^2未知，又设$X_1,\cdots,X_n$是一个样本；求：μ,σ^2的矩估计量。

解: 因为$EX=\mu$，$EX^2=\sigma^2+\mu^2$

令$A_1=\bar{X}$替换EX，令$A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$替换$EX^2$得

$$\hat{\mu}=\bar{X}$$

$$\hat{\sigma}^2=A_2-\hat{\mu}^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

§7.1 参数的点估计

☆ **注(3)**

若 $\hat{\theta}$ 是未知参数 θ 的极大似然估计，

$g(\theta)$ 具有单值反函数，则 $g(\theta)$ 的极大似然估计为 $g(\hat{\theta})$.

例4：设X1，…，Xn为取自参数为λ的指数分布总体的样本，a>0为一给定实数。求p=P{X<a}的最大似然估计。

解 $\because p = P\{X < a\} = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}$

关于 λ 严单，$\therefore$ 若 $\hat{\lambda}$ 是 λ 的最大似然估计，则 $\hat{p} = 1 - e^{-\hat{\lambda} a}$

其中 $\hat{\lambda} = \dfrac{1}{\bar{X}}$.

§7.1 参数的点估计

☆ **例5** 设X_1，…，X_n为取自 U[a,b] 的样本，a<b 均未知，求参数a, b的最大似然估计。

解: 若按通常的做法, 得

$$L(a,b)=\begin{cases}\prod_{i=1}^{n}\dfrac{1}{b-a}=\dfrac{1}{(b-a)^n} & a\le x_i\le b,\ i=1,2,\cdots,n\\ 0 & else\end{cases}$$

只考虑 $a \le x_i \le b$，$i = 1, 2, \cdots, n$时，

$\ln L(a,b) = -n\ln(b-a)$

$$\begin{cases}\dfrac{\partial \ln L}{\partial a} = \dfrac{n}{b-a} \ne 0\\ \dfrac{\partial \ln L}{\partial b} = -\dfrac{n}{b-a} \ne 0\end{cases}$$

§7.1 参数的点估计

☆ **例5(续)**

用最大似然估计思想求

$$L(a,b)=\begin{cases}\prod_{i=1}^{n}\dfrac{1}{b-a}=\dfrac{1}{(b-a)^n} & a\le x_i\le b,\ i=1,2,\cdots,n\\ 0 & else\end{cases}$$

$$=\begin{cases}\dfrac{1}{(b-a)^n} & a\le \min\{x_1,\cdots,x_n\}\le \max\{x_1,\cdots,x_n\}\le b\\ 0 & else\end{cases}$$

对满足条件 $a \le x_{(1)} \le x_{(n)} \le b$ **的任意a,b有**

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

因此取 $\begin{cases}\hat{a} = X_{(1)}\\ \hat{b} = X_{(n)}\end{cases}$

§7.1 参数的点估计

☆ **例6** 在一次试验中事件$A_1, A_2, \ldots, A_r$发生的概率分别为 $p_1, p_2, \ldots, p_r (p_i > 0, i = 1, 2, \ldots, r)$，且$p_1 + p_2 + \ldots + p_r = 1$.设在$n$次独立试验中，事件$A_1, A_2, \ldots, A_r$分别发生$n_1, n_2, \ldots, n_r$次，$n_1 + n_2 + \ldots + n_r = n$，试求参数$p_1, p_2, \ldots, p_r$的最大似然估计.

解：由于常数系数不影响取最大值的θ，可取似然函数为

$L(p_1, p_2, \ldots, p_r) = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$.

$\ln L(p_1, p_2, \ldots, p_r) = n_1 \ln p_1 + n_2 \ln p_2 + \cdots n_r \ln p_r$.

如何求似然函数的最大值？——拉格朗日数乘法

§7.1 参数的点估计

★ 例6 回忆拉格朗日数乘法

求$f(x,y)$在给定条件$g(x,y)=0$时的极值：

步骤

1.令$F(x,y,\lambda)=f(x,y)+\lambda g(x,y)$

2.求导

$$\frac{\partial F(x,y,\lambda)}{\partial x}=0$$

$$\frac{\partial F(x,y,\lambda)}{\partial y}=0$$

$$\frac{\partial F(x,y,\lambda)}{\partial \lambda}=g(x,y)=0$$

3. 解出(x,y).

§7.1 参数的点估计

★ 例6 回忆拉格朗日数乘法

取$g(p_1,p_2,\ldots,p_r)=p_1+p_2+\ldots+p_r-1$,则

$f(p_1,p_2,\ldots,p_r,\lambda)=n_1\ln p_1+n_2\ln p_1+\cdots n_r\ln p_r-\lambda g(p_1,p_2,\ldots,p_r)$

求导得：

$\frac{n_i}{p_i}-\lambda_i=0\Rightarrow p_i=\frac{n_i}{\lambda}$,

又$p_1+p_2+\ldots+p_r$=1,

所以$\lambda=n_1+n_2+\ldots+n_r=n$,

$\hat{p}_i=\frac{n_i}{n}$.

主观题 1分 设置

设总体X服从几何分布$P(X=x)=(1-p)^{x-1}p$,

$x=1,2,\ldots.$,其中p为未知参数,$0<p<1$.

设$X_1,X_2,\ldots,X_n$为X的一个样本，

求参数p的矩估计和最大似然估计.

正常使用主观题需2.0以上版本雨课堂

作答

主观题 10分 设置

设X的分布律为

X	0	1	2	3
p	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中$0<\theta<0.5$,现观测到样本观测值为

3,1,3,0,3,1,2,3, 求θ的极大似然估计值.

作答

§7.1 参数的点估计

★ §7.1.2 估计量的优良性

1 无偏性

2 有效性

3 相合性（一致性）

§7.1 参数的点估计

★ §7.1.2 估计量的优良性

1 无偏性

设$\hat{\theta}$是参数θ的估计量，如果$\forall \theta \in \Theta$，

$E_\theta(\hat{\theta}) = \theta$，则称$\hat{\theta}$是$\theta$的无偏估计.

§7.1 参数的点估计

★ §7.1.2 估计量的优良性

注：

(1) 若$\hat{\theta}$是θ的无偏估计量，则

$g(\hat{\theta}) = a\hat{\theta} + b$是$g(\theta) = a\theta + b$的无偏估计量。

(2) 对一般函数$g(\theta)$，

$g(\hat{\theta})$一般不是$g(\theta)$的无偏估计量。

§7.1 参数的点估计

★ 反例 $g(\theta) = \theta^2$

若$\hat{\theta}$是θ的无偏估计量，且$D(\hat{\theta}) > 0$，

$E(\hat{\theta}^2) = D(\hat{\theta}) + (E(\hat{\theta}))^2 > \theta^2$.

§7.1 参数的点估计

例1

$\overline{X}$是总体均值μ的无偏估计吗？

S^2是总体方差σ^2的无偏估计吗？

$$E(\overline{X}) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} EX_i = \mu$$

$$ES^2 = E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2] = \sigma^2 \text{ (已证结论)}$$

样本k阶原点矩是总体k阶原点矩的无偏估计吗？

第七章 参数估计

§7.1 参数的点估计

例1

注：

若总体k阶矩存在时，则样本原点矩是总体原点矩的无偏估计，但是样本中心距则不一定，考虑二阶中心距

$$EV^2 = E[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2] = \frac{n-1}{n}\sigma^2 \neq \sigma^2,$$

但 $\lim_{n\to\infty} EV^2 = \sigma^2$，

所以它是渐近无偏的，即当样本量较大时，二阶样本中心距可以近似看作是总体方差的无偏估计。

第七章 参数估计

表彰榜

陈静波

白煜涵

第七章 参数估计

上节回顾

最大似然估计

思想：观测到的事件就是发生的概率最大的

基本方法和步骤：
1.写出似然函数
2.写出对数似然函数（若似然函数易求最大值点可省去）
3.列似然方程（求导），求出达到最大值的点
注：若求估计量就写关于样本的函数，若求估计值就把样本观测值带入求出结果

求最大值点的常用方法：
1.求导
2.判断函数单调性
3.拉格朗日数乘法
4.具体情况具体分析O(∩_∩)O

§7.1 参数的点估计

渐近无偏估计

设$\hat{\theta}$是参数θ的估计量，如果$\forall\theta\in\Theta$，

$E_\theta(\hat{\theta})\xrightarrow{n\to+\infty}\theta$，则称$\hat{\theta}$是$\theta$的渐近无偏估计.

§7.1 参数的点估计

例2 设$X_1,\cdots,X_n\overset{iid}{\sim}U(0,\theta)$，试讨论$\theta$的矩估计$\hat{\theta}_M$与最大似然估计$\hat{\theta}_{MLE}$的无偏性.

解 已知$\hat{\theta}_M=2\overline{X}$，$\hat{\theta}_{MLE}=X_{(n)}$

$\because E\hat{\theta}_M=E(2\overline{X})=2EX=2(\frac{\theta}{2})=\theta$，$\therefore\hat{\theta}_M$是$\theta$的无偏估计.

$$F_{X_{(n)}}(x)=(F_X(x))^n=\begin{cases}0 & x<0\\ x^n/\theta^n & 0\le x<\theta\\ 1 & x\ge\theta\end{cases}$$

极大值的分布

§7.1 参数的点估计

例2(续)

$$\therefore X_{(n)}\sim f_n(x)=\begin{cases}\dfrac{nx^{n-1}}{\theta^n}, & 0<x<\theta\\ 0, & 其它\end{cases}$$

$$E\hat{\theta}_{MLE}=EX_{(n)}=\int_0^\theta xn\frac{x^{n-1}}{\theta^n}dx=\frac{n}{n+1}\theta\ne\theta$$

$\therefore\hat{\theta}_{MLE}$不是$\theta$的无偏估计.

§7.1 参数的点估计

§7.1.2 估计量的优良性

2 有效性

设$\hat{\theta}_1$和$\hat{\theta}_2$是参数θ的两个无偏估计，

如果对$\forall\theta\in\Theta$，$D_\theta(\hat{\theta}_1)\le D_\theta(\hat{\theta}_2)$，

且至少对某个$\theta\in\Theta$上式中的不等号成立，

则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

单选题 1分 设置

例2(补) 考虑$\hat{\theta}_{MLE}$修正$\hat{\theta}=\frac{n+1}{n}X_{(n)}$,证明它为$\theta$的无偏估计,并比较$\hat{\theta}$与$\hat{\theta}_M$的有效性.

A 矩估计更有效

B 修正之后更有效

C 一样有效

D 不知道

提交

解: $\hat{\theta}_M=2\bar{X}$ 与 $\hat{\theta}=\frac{n+1}{n}X_{(n)}$ 都是θ的无偏估计

$X_{(n)}=\max(X_1,X_2,\cdots,X_n)$ 的概率密度函数为 $f_n(x)=\begin{cases}\frac{nx^{n-1}}{\theta^n}, & 0<x<\theta,\\ 0, & 其它,\end{cases}$

$$D(2\bar{X})=4D(\bar{X})=\frac{4}{n}D(X)=\frac{4}{n}\cdot\frac{\theta^2}{12}=\frac{\theta^2}{3n}$$

$$D(X_{(n)})=\int_0^\theta x^2\frac{n}{\theta^n}x^{n-1}dx-\frac{n^2}{(n+1)^2}\theta^2=\frac{n}{(n+1)^2(n+2)}\theta^2$$

$$D\left(\frac{n+1}{n}X_{(n)}\right)=\frac{(n+1)^2}{n^2}\cdot\frac{n}{(n+1)^2(n+2)}\theta^2=\frac{1}{n(n+2)}\theta^2$$

由于 $D\left(\frac{n+1}{n}X_{(n)}\right)\le D(2\bar{X})$ 且当 $n>1$ 时，不等号成立，故 $\frac{n+1}{n}X_{(n)}$ 比 $2\bar{X}$ 有效.

第七章 参数估计

§7.1 参数的点估计

★ §7.1.2 估计量的优良性

3 相合性（一致性）

设$\hat{\theta}_n$为样本容量为n时未知参数θ的估计量,如果

$\lim_{n\to+\infty}P(|\hat{\theta}_n-\theta|\ge\varepsilon)=0$,对一切$\varepsilon$成立,

则称$\hat{\theta}_n$为θ的相合估计或一致估计.

第七章 参数估计

§7.1 参数的点估计

★ §7.1.2 估计量的优良性

3 相合性（一致性）

样本k阶原点矩是总体k阶原点矩的相合估计吗？

§7.1 参数的点估计

★ §7.1.2 估计量的优良性

3 相合性（一致性）

定理：若$\hat{\theta}_n$是θ的相合估计，$g(\theta)$是θ的连续函数，则$g(\hat{\theta}_n)$是$g(\theta)$的相合估计.

注：该定理可以推广到θ是向量的情形，即多个参数情形.
见茆诗松书p77页定理2.1.2

§7.1 参数的点估计

★ 例3 设总体 $X \sim N(\mu,\ \sigma^2)$，其中 μ 已知，而 $\sigma^2>0$ 为未知参数，$X_1,\cdots,X_n$ 是从该总体中抽取的一个样本. 证明σ^2的最大似然估计是无偏估计

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$$

由于 $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$

$$E(\hat{\sigma}^2)=E\left(\frac{\sigma^2}{n}\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\right)=\frac{\sigma^2}{n}\times n=\sigma^2$$

主观题 1分 设置

★ 例3(续) 比较S^2与$\hat{\sigma}^2$的有效性

判断S^2与$\hat{\sigma}^2$的是否是σ^2的相合估计？

正常使用主观题需2.0以上版本雨课堂

作答

§7.1 参数的点估计

★ 渐近正态性

相合性：定性的性质

收敛的速度 ← 渐近正态性

§7.1 参数的点估计

★ 渐近正态性

定义：

称$\hat{\theta}_n$是θ的渐近正态估计或$\hat{\theta}_n$具有渐近正态性，
如果存在趋于0的非负常数序列$\sigma_n(\theta)$，使得
$\dfrac{\hat{\theta}_n-\theta}{\sigma_n(\theta)}$依分布收敛于标准正态分布.
这时也称$\hat{\theta}_n$服从渐近正态分布$N(\theta,\sigma_n^2(\theta))$,
记为$\hat{\theta}_n \sim AN(\theta,\sigma_n^2(\theta))$.

§7.1 参数的点估计

★ 渐近正态性

注：

（1）$\sigma_n^2(\theta)$称为$\hat{\theta}_n$的渐近方差

（2）$\sigma_n(\theta)$表示着估计量$\hat{\theta}_n$依概率收敛到θ的速度

（3）$\sigma_n(\theta)$不唯一

（4）渐近正态性$\Rightarrow$相合性

§7.1 参数的点估计

★ 渐近正态性

例1: 若总体的均值为μ，方差为σ^2，证明
样本均值$\overline{X}$是μ的渐近正态估计.

§7.1 参数的点估计

★ 渐近正态性

例2: 若总体$X \sim N(\mu,\sigma^2)$,证明样本方差
S^2是σ^2的渐近正态估计.

表彰榜

丁子缘

上节回顾

无偏性 $E\hat{\theta}=\theta$

有效性

设$\hat{\theta}_1$和$\hat{\theta}_2$是参数θ的两个无偏估计，

如果对$\forall\theta\in\Theta$，$D_\theta(\hat{\theta}_1)\le D_\theta(\hat{\theta}_2)$，

且至少对某个$\theta\in\Theta$上式中的不等号成立，

则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

相合性（一致性）

设$\hat{\theta}_n$为样本容量为n时未知参数θ的估计量，如果

$\lim_{n\to+\infty}P(|\hat{\theta}_n-\theta|\ge\varepsilon)=0$，对一切$\varepsilon$成立，

则称$\hat{\theta}_n$为θ的相合估计或一致估计.

上节回顾

渐近正态性

称$\hat{\theta}_n$是θ的渐近正态估计：

$\dfrac{\hat{\theta}_n-\theta}{\sigma_n(\theta)}$依分布收敛于标准正态分布

$\sigma_n(\theta)$：趋于0的非负常数序列

$\hat{\theta}_n\sim AN(\theta,\sigma_n^2(\theta))$.

§7.1.2 一致最小方差无偏估计（UMVUE）

判断估计量的好坏

大样本场合 $\longrightarrow$ **相合性**

小样本场合 $\longrightarrow$ **如何判断?**

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 均方误差

判断一个点估计的好坏指标，我们最自然的想法考虑的是估计量与真实参数之间的距离，即 $|\hat{\theta}-\theta|$ 。通常我们最常用的函数是距离的平方，又由于估计量的随机性，所以我们考虑

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

这个函数即为均方误差，记为 $MSE(\hat{\theta})=E(\hat{\theta}-\theta)^2$.

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 均方误差

注：

(1)均方误差是评价点估计的最一般的标准。

(2)均方误差与方差的关系：

$MSE(\hat{\theta})=D(\hat{\theta})+B^2$,

其中$B=E(\hat{\theta}-\theta)=E(\hat{\theta})-\theta$称为偏差。

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差无偏估计

定义: 若 $\hat{\theta}_0$是参数 θ的一个无偏估计量，而 $\hat{\theta}$是参数 θ的任意一个无偏估计量，即

$E_\theta(\hat{\theta}_0)=E_\theta(\hat{\theta})=\theta$,

如果有

$D_\theta(\hat{\theta}_0)\le D_\theta(\hat{\theta}), \forall\theta\in\Theta$

成立, 则称 $\hat{\theta}_0$是 θ的一致最小方差无偏估计，记为UMVUE（Uniformly Minimum Variance Unbiased Estimator）.

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差无偏估计判定准则

定理: 设$X_1,X_2,\ldots,X_n$为来自某总体的一个样本，

$\hat{\theta}(X_1,X_2,\ldots,X_n)$是参数 θ的一个无偏估计，

$D_\theta(\hat{\theta}(X_1,X_2,\ldots,X_n))<+\infty$,则 $\hat{\theta}$是 θ的UMVUE

$\Leftrightarrow$

对任意满足

$E_\theta(T(X_1,X_2,\ldots,X_n))=0$且

$D_\theta(T(X_1,X_2,\ldots,X_n))<+\infty$

的$T(X_1,X_2,\ldots,X_n)$,都有$Cov_\theta(\hat{\theta},T)=0, \forall\theta\in\Theta$.

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差无偏估计判定方法

注：上述定理表明，UMVUE与任意一个均值为0，方差有限的估计量都不相关.

第七章 参数估计

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差无偏估计判定方法

推论： 对于参数θ而言，最多存在一个UMVUE.（概率意义下）

第七章 参数估计

§7.1.2 一致最小方差无偏估计（UMVUE）

★ **例** 设总体$X \sim N(\mu,\sigma^2)$，其中μ,σ^2未知. 试证μ,σ^2的一致最小方差无偏估计量分别为$\overline{X},S^2$.（S^2是σ^2的UMVUE(作业)）.

主观题 10分

设置

证明：若$\hat{\theta}_1$，$\hat{\theta}_2$分别是θ_1，θ_2的UMVUE，则对任意的非零常数a，b，$\widehat{a\theta_1} + \widehat{b\theta_2}$是$a\theta_1 + b\theta_2$的UMVUE.

作答

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差线性无偏估计

定义: 设$\hat{\theta}$是参数θ的估计量，若满足

(1)$\hat{\theta}(X_1,X_2,\dots X_n)=\sum_{i=1}^{n}C_iX_i$,其中$C_i$为常数;

(2)$E_\theta(\hat{\theta})=\theta$;

(3)对满足上述(1)(2)的估计量$\hat{\theta}'$,都有

$D_\theta(\hat{\theta})\le D_\theta(\hat{\theta}'),\forall\theta\in\Theta$,

则称$\hat{\theta}$是样本$X_1,X_2,\dots X_n$线性组合类当中

θ的一致最小方差线性无偏估计.

第七章 参数估计

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差线性无偏估计

例 设$X_1,X_2,\dots,X_n$是为总体X的一个样本，记μ为总体均值，$a_i(i=1,2,\dots n)$为常数，且

$\sum_{i=1}^{n}a_i=1$.

(1)证明$\sum_{i=1}^{n}a_iX_i$是μ的无偏估计量.

(2)证明在μ的所有形如$\sum_{i=1}^{n}a_iX_i$的线性无偏估计中，以$\overline{X}$最有效.

第七章 参数估计

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 一致最小方差线性无偏估计

定理: 设$\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_m$是参数$\theta$的$m$个相互独立的线性无偏估计量，且有相同的方差

$D_\theta(\hat{\theta}_j)=\sigma^2<+\infty, j=1,2,\dots m$,

则统计量$\overline{\theta}=\frac{1}{m}\sum_{i=1}^{m}\hat{\theta}_i$是$\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_m$线性组合类中

θ的一致最小方差线性无偏估计量, 且$D_\theta(\overline{\theta})=\frac{\sigma^2}{m}$.

第七章 参数估计

表彰榜

张源钧

§7.1.2 一致最小方差无偏估计（UMVUE）

★ 充分性原则

定理:

设总体的分布为$p(x;\theta)$，$\theta\in\Theta$，$X_1,X_2,\dots,X_n$是其样本，

$T(X_1,X_2,\dots,X_n)$是θ的充分统计量，

设$\varphi(X_1,X_2,\dots,X_n)$是$g(\theta)$的一个无偏估计，则

$\hat{g}(T)=E(\varphi(X_1,X_2,\dots,X_n)\,|\,T)$

也是$g(\theta)$的无偏估计，并且

$D_\theta(\hat{g}(T))\le D_\theta(\varphi(X_1,X_2,\dots,X_n)),\forall\theta\in\Theta,$

且等号成立的充要条件是

$P(\varphi(X_1,X_2,\dots,X_n)=\hat{g}(T))=1.$

§7.1.2 一致最小方差无偏估计（UMVUE）

★ Cramer——Rao不等式

定义（费希尔信息量Fisher information）：

设总体概率函数为$f(x;\theta),\theta\in\Theta$满足下列条件

(1)参数空间Θ是直线上的一个开区间；

(2)支撑$S=\{x:f(x;\theta)>0\}$与θ无关；

(3)导数$\frac{\partial}{\partial\theta}f(x;\theta)$对$\theta\in\Theta$都存在；

(4)对$f(x;\theta)$,积分和微分运算可以交换顺序；

(5)期望$E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2$存在，

则称$I(\theta)=E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2$为总体分布的费希尔信息量.

§7.1.2 一致最小方差无偏估计（UMVUE）

★ Cramer——Rao不等式

定理（C-R不等式）：

设$X_1,X_2,\dots,X_n$是为总体X的一个样本，$X\sim f(x;\theta)$,

$f(x;\theta)$满足上述定义中的条件，设

$g(\theta)$是Θ的可微函数，

$\hat{g}(X_1,X_2,\dots,X_n)$是$g(\theta)$的无偏估计量，且对$\Theta$中的一切$\theta$，

$$g(\theta)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\hat{g}(X_1,X_2,\dots,X_n)\prod_{i=1}^{n}f(x_i;\theta)dx_1\dots dx_n$$

的微分可在积分号下进行，即

$$g'(\theta)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\hat{g}(X_1,X_2,\dots,X_n)\left[\frac{\partial}{\partial\theta}\prod_{i=1}^{n}f(x_i;\theta)\right]dx_1\dots dx_n,$$

则有$D_\theta(\hat{g}(X_1,X_2,\dots,X_n))\ge\frac{[g'(\theta)]^2}{nI(\theta)}$，称为C-R不等式.

§7.1.2 一致最小方差无偏估计（UMVUE）

★ Cramer——Rao不等式

注:

(1)C-R不等式表明在样本容量n给定时，待估参数的无偏估计量的方差不可能无限小，它有一个下界:

$\frac{[g'(\theta)]^2}{nI(\theta)}$ 称为C-R下界.

(2)如果有一个待估参数的无偏估计量达到了C-R下界，那么该估计量就是UMVUE.

§7.1.2 一致最小方差无偏估计（UMVUE）

★ **例** 证明Fisher信息量的另一个计算公式

$$I(\theta)=E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2=-E\left[\frac{\partial^2\ln f(X;\theta)}{\partial\theta^2}\right].$$

Pf: $E\left(\frac{\partial}{\partial\theta}\ln f(X;\theta)\right)=0$

$$0=\frac{\partial}{\partial\theta}E\left(\frac{\partial}{\partial\theta}\ln f(X;\theta)\right)$$

$$=\frac{\partial}{\partial\theta}\int\frac{\partial}{\partial\theta}\ln f(x;\theta)\,f(x;\theta)\,dx$$

积分求导可交换. $\int\frac{\partial^2\ln f(x;\theta)}{(\partial\theta)^2}\cdot f(x;\theta)\,dx$

$$+\int\frac{\partial\ln f(x;\theta)}{\partial\theta}\cdot\frac{\partial f(x;\theta)}{\partial\theta}\,dx$$

$$=E\left(\frac{\partial^2\ln f(X;\theta)}{(\partial\theta)^2}\right)+\int\frac{\partial\ln f(x;\theta)}{\partial\theta}\cdot\frac{\partial\ln f(x;\theta)}{\partial\theta}\cdot f(x;\theta)\,dx$$

$$=E\left(\frac{\partial^2(\ln f(X;\theta))}{(\partial\theta)^2}\right)+\underbrace{E\left[\left(\frac{\partial\ln f(X;\theta)}{\partial\theta}\right)^2\right]}_{=I(\theta)}$$

上节回顾

C-R不等式：

$$\boxed{D_\theta(\hat g(X_1,X_2,\dots,X_n))\ge\frac{[g'(\theta)]^2}{nI(\theta)}}$$

$\hat g(X_1,X_2,\dots,X_n)$是$g(\theta)$的无偏估计量

$I(\theta)=E\left[\frac{\partial}{\partial\theta}\ln f(x;\theta)\right]^2$ 为总体分布的费希尔信息量

上节回顾

Fisher信息量的计算公式：

$$I(\theta)=E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2=-E\left[\frac{\partial^2\ln f(X;\theta)}{\partial\theta^2}\right].$$

§7.1.2 一致最小方差无偏估计（UMVUE）

例 设总体$X\sim N(\mu,\sigma_0^2)$，σ_0^2已知，$X_1,X_2,\dots,X_n$为X的一个样本：
(1)求均值μ的矩估计，并证明其无偏性；
(2)证明该估计量是均值μ的UMVUE.

(1)解：
因为$E(X)=\mu$，用$\overline{X}$替换$E(X)$，得μ的矩估计量为$\overline{X}$，
又因为$E(\overline{X})=\mu$，所以$\overline{X}$是μ的无偏估计量.

§7.1.2 一致最小方差无偏估计（UMVUE）

例(续) (2)证明：

$$f(x;\mu)=\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(x-\mu)^2}{2\sigma_0^2}},$$

$$\ln f(x;\mu)=-\ln(\sqrt{2\pi}\sigma_0)-\frac{(x-\mu)^2}{2\sigma_0^2},$$

$$\frac{\partial\ln f(x;\mu)}{\partial\mu}=\frac{x-\mu}{\sigma_0^2},$$

$$I(\mu)=E\left[\left(\frac{X-\mu}{\sigma_0^2}\right)^2\right]=\frac{E\left[(X-\mu)^2\right]}{\sigma_0^4}=\frac{1}{\sigma_0^2},$$

所以C-R下界为$\dfrac{(\mu')^2}{nI(\mu)}=\dfrac{\sigma_0^2}{n}$，

又$D(\overline{X})=\dfrac{\sigma_0^2}{n}$，所以$\overline{X}$是$\mu$的UMVUE.

§7.1.2 一致最小方差无偏估计（UMVUE）

例 设总体$X\sim B(1,p)$，$X_1,X_2,\dots,X_n$为X的一个样本，
证明$\overline{X}$是p的UMVUE.

证明：
$E(\overline{X})=E(X)=p$，所以$\overline{X}$是p的无偏估计量，

$$D(\overline{X})=\frac{D(X)}{n}=\frac{p(1-p)}{n}.$$

下证C-R下界为$\dfrac{p(1-p)}{n}$.

§7.1.2 一致最小方差无偏估计（UMVUE）

例(续) 下证C-R下界为$\dfrac{p(1-p)}{n}$.

$p(x;p)=p^x(1-p)^{1-x}$，$x=0$或1，

$\ln p(x;p)=x\ln p+(1-x)\ln(1-p)$，

$$\frac{\partial\ln p(x;p)}{\partial p}=\frac{x}{p}-\frac{1-x}{1-p}=\frac{x-p}{p(1-p)},$$

$$I(p)=E\left[\left(\frac{X-p}{p(1-p)}\right)^2\right]=\frac{E\left[(X-p)^2\right]}{p^2(1-p)^2}=\frac{1}{p(1-p)},$$

所以C-R下界为$\dfrac{1}{nI(p)}=\dfrac{p(1-p)}{n}=D(\overline{X})$，

所以$\overline{X}$是p的UMVUE.

§7.1.2 一致最小方差无偏估计（UMVUE)

★ **总结：用C-R不等式证明UMVUE的方法步骤**

设总体$X \sim f(x;\theta), X_1, X_2, \dots, X_n$为$X$的一个样本，

证明$\hat{\theta}$是θ的UMVUE.

1 证明$\hat{\theta}$是θ的无偏估计量，即验证$E(\hat{\theta})=\theta$;

2 计算$D(\hat{\theta})$;

3 计算C-R下界：

(1)计算Fisher信息量$I(\theta)=E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2=-E\left[\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}\right]$.

(2)C-R下界为$\frac{1}{nI(\theta)}$

4 验证C-R下界与$D(\hat{\theta})$是否相等.

§7.1.2 一致最小方差无偏估计（UMVUE)

★ **总结：用C-R不等式证明UMVUE的方法步骤**

设总体$X \sim f(x;\theta), X_1, X_2, \dots, X_n$为$X$的一个样本，

证明$\hat{g}(X_1, X_2, \dots, X_n)$是$g(\theta)$的UMVUE.

1 证明$\hat{g}(X_1, X_2, \dots, X_n)$是$g(\theta)$的无偏估计量，即验证$E(\hat{g})=g(\theta)$;

2 计算$D(\hat{g}(X_1, X_2, \dots, X_n))$;

3 计算C-R下界：

(1)计算Fisher信息量$I(\theta)=E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2=-E\left[\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}\right]$.

(2)C-R下界为$\frac{[g'(\theta)]^2}{nI(\theta)}$

4 验证C-R下界与$D(\hat{g}(X_1, X_2, \dots, X_n))$是否相等.

§7.1.2 一致最小方差无偏估计（UMVUE)

★ **思考**

1.如果没有达到C-R下界，则UMVUE一定不存在吗？

有效估计：

若定理中等号成立，即$D_\theta(\hat{g}(X_1, X_2, \dots, X_n))=\frac{[g'(\theta)]^2}{nI(\theta)}$,

则称$\hat{g}(X_1, X_2, \dots, X_n)$是$g(\theta)$的有效估计.

C-R定理表明有效估计一定是UMVUE，反之不一定.

单选题 1分 设置

关于UMVUE下面说法错误的是

A UMVUE是无偏的

B UMVUE与零均值且方差存在的统计量都不相关

C 至多存在一个UMVUE（概率意义下）

D 在样本量给定时，UMVUE的方差可以无限小

提交

§7.1 参数的点估计

★ 渐近正态性

条件：只考虑θ为一维情况

设总体X的密度函数为$f(x;\theta),\theta\in\Theta,\Theta$为非退化区间，假定

（1）对任意的x，偏导数

$\frac{\partial f(x;\theta)}{\partial\theta},\frac{\partial^2 f(x;\theta)}{\partial\theta^2},\frac{\partial^3 f(x;\theta)}{\partial\theta^3}$对$\forall\theta\in\Theta$都存在

（2）对$\forall\theta\in\Theta$有

$$\left|\frac{\partial f(x;\theta)}{\partial\theta}\right|\le F_1(x),\left|\frac{\partial^2 f(x;\theta)}{\partial\theta^2}\right|\le F_2(x),\left|\frac{\partial^3 \ln f(x;\theta)}{\partial\theta^3}\right|\le H(x)\text{成立},$$

其中

$$\int_{-\infty}^{+\infty}F_i(x)dx<+\infty,i=1,2.\qquad \int_{-\infty}^{+\infty}H(x)f(x;\theta)dx<+\infty.$$

（3）对$\forall\theta\in\Theta,0<I(\theta)<+\infty$

§7.1 参数的点估计

★ 渐近正态性

定理：

设$X_1,X_2,\dots,X_n$是满足上述条件总体的简单随机样本，且设对数似然方程有唯一的解$\hat\theta$，则$\hat\theta$是θ的相合渐近正态估计（consistent Asymptotic Normal）（CAN估计）且

$$\sqrt{n}(\hat\theta-\theta)\xrightarrow{d}Y\sim N(0,\frac{1}{I(\theta)})$$

即$\hat\theta\sim AN(\theta,\frac{1}{nI(\theta)})$.

主观题 10分 设置

计算正态总体$N(\mu,\sigma^2)$中σ^2的fisher信息量以及C-R下界，μ已知.

作答

§7.2 区间估计

★ 区间估计

前面介绍了估计总体未知参数θ的两种方法，这两种方法都是要估计出θ的一个值．从几何角度来看表示一个点，所以这两种估计方法有称为点估计．

但有的时候，我们还希望估计出包含未知参数θ的一个范围，并且还希望知道这个范围包含θ的可靠程度，这就是区间估计．

$$\theta\begin{cases}\text{点估计}\begin{cases}\text{矩法估计}\\\text{最大拟然估计}\end{cases}\\\text{区间估计}\end{cases}$$

§7.2 区间估计

定义(置信区间)

设总体X的分布函数为$F(x;\theta)$，θ为未知参数，$\theta \in \Theta$，对于给定值$\alpha(0<\alpha<1)$，若有两个统计量

$\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\ldots X_n)$和$\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\ldots X_n)$使得

$$P\left\{\hat{\theta}_1<\theta<\hat{\theta}_2\right\}=1-\alpha$$

对一切$\theta \in \Theta$成立，则称随机区间$(\hat{\theta}_1,\hat{\theta}_2)$是参数$\theta$的置信度为$1-\alpha$的置信区间，$\hat{\theta}_1,\hat{\theta}_2$分别称为置信下限和置信上限.

§7.2 区间估计

注：

(1)参数θ虽然未知，但是却是常数，而$\hat{\theta}_1,\hat{\theta}_2$是随机变量，所以置信区间是一个随机区间.

(2)置信区间的意思是$(\hat{\theta}_1,\hat{\theta}_2)$以$1-\alpha$的概率包含参数$\theta$.

一般我们说θ的区间估计，就是在给定置信度$1-\alpha$的情况下，给出其置信区间.

§7.2 区间估计

定义(单侧置信区间)

设总体X的分布函数为$F(x;\theta)$，θ为未知参数，$\theta \in \Theta$，对于给定值$\alpha(0<\alpha<1)$，若存在统计量

$\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\ldots X_n)$使得

$P\left\{\theta>\hat{\theta}_1\right\}=1-\alpha$对一切$\theta \in \Theta$成立，

则称随机区间$(\hat{\theta}_1,+\infty)$是参数θ的置信度为$1-\alpha$的单侧置信区间，$\hat{\theta}_1$称为$1-\alpha$的单侧置信下限.

§7.2 区间估计

定义(续)(单侧置信区间)

又若存在统计量

$\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\ldots X_n)$使得

$P\left\{\theta<\hat{\theta}_2\right\}=1-\alpha$对一切$\theta \in \Theta$成立，

则称随机区间$(-\infty,\hat{\theta}_2)$是参数θ的置信度为$1-\alpha$的单侧置信区间，$\hat{\theta}_2$称为$1-\alpha$的单侧置信上限.

§7.2 区间估计

★ 注

1. **估计精度：** 置信区间的长度就是区间估计的估计精度，区间长度越短，精度越高.
2. **置信度：** 置信度的大小反映了这个区间估计的可靠程度，即随机区间包含参数的概率大小
3. **精度与置信度的关系：** 在样本容量一定的情况下，精度和置信度是此消彼长的关系.

§7.2 区间估计

★ 枢轴量法

枢轴量法是构造未知参数置信区间的最常用的方法.

设$G(X_1,X_2,\dots,X_n,\theta)$是关于样本$X_1,X_2,\dots,X_n$和未知参数$\theta$的函数，不包含其他未知参数，$G(X_1,X_2,\dots,X_n,\theta)$分布已知且不依赖于$\theta$.
一般称具有这种性质的G为枢轴量.

§7.2 区间估计

★ 枢轴量法的具体步骤

(i)设法构造一个枢轴量$G(X_1,X_2,\dots,X_n,\theta)$;

(ii)对于给定的置信度$1-\alpha$,利用$G(X_1,X_2,\dots,X_n,\theta)$的已知分布，确定两个常数$a,b$,使得

$$P\{a<G(X_1,X_2,\dots,X_n,\theta)<b\}=1-\alpha.$$

(iii)利用不等式运算，将$a<G(X_1,X_2,\dots,X_n,\theta)<b$进行等价变换，若能得到$\hat\theta_1<\theta<\hat\theta_2$的形式，则$(\hat\theta_1,\hat\theta_2)$是$\theta$的$1-\alpha$的置信区间.

§7.2 区间估计

★ 注：

(1)构造枢轴量一般从θ的点估计$\hat\theta$出发，构造函数$G(\hat\theta,\theta)$,此处的$\hat\theta$常用最大似然估计.

(2)满足(iii)中条件的a,b有很多，选择的目的是希望置信区间的平均长度$E_\theta(\hat\theta_2-\hat\theta_1)$尽可能短.

§7.2 区间估计

★ 注：在$G(X_1,X_2,\dots,X_n,\theta)$为单峰分布时，常用如下两种方法确定$a,b$:

(1)当$G(X_1,X_2,\dots,X_n,\theta)$的分布对称时，如正态分布，t分布，由对称性可取$b$,使得

$$P\{-b<G(X_1,X_2,\dots,X_n,\theta)<b\}=1-\alpha,$$

这时$a=-b$,b为$G(X_1,X_2,\dots,X_n,\theta)$的分布的$\frac{\alpha}{2}$上侧分位数.

§7.2 区间估计

★ 注：

(2)当$G(X_1,X_2,\dots,X_n,\theta)$的分布不对称时，如$\chi^2$分布，F分布，可取$a,b$使得

$$P\{G(X_1,X_2,\dots,X_n,\theta)\le a\}=\frac{\alpha}{2},$$

$$P\{G(X_1,X_2,\dots,X_n,\theta)\ge b\}=\frac{\alpha}{2}$$

这时a为$G(X_1,X_2,\dots,X_n,\theta)$的分布的$\frac{\alpha}{2}$分位数，

b为$G(X_1,X_2,\dots,X_n,\theta)$的分布的$\frac{\alpha}{2}$上侧分位数.

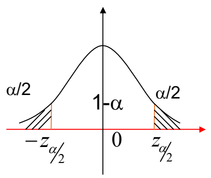
§7.2 区间估计

★ 单个正态总体参数的置信区间

1、设$X_1,\cdots,X_n \overset{iid}{\sim} N(\mu, \sigma^2)$，给定$\alpha$，求$\mu$的置信区间.

(1)、σ^2已知

$$\boxed{\because\ Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)}$$



对给定的α，有

$$P\{-z_{\alpha/2}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_{\alpha/2}\}=1-\alpha$$

其中 $\boxed{\Phi(z_{\alpha/2})=1-\alpha/2}$

§7.2 区间估计

★ 单个正态总体参数的置信区间

等价变形得 $P\{\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}=1-\alpha$

所以，μ的置信度为1−α的置信区间为

$$\boxed{(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}})=(\overline{X}\pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}})}$$

单选题 1分 设置

例1: 设X~N(μ, 0.09), 测得X的4个观测值为12.6, 13.4,12.8, 13.2, 求μ的95%置信区间.

x	5	6
1.6	0.9505	0.9515
1.9	0.9744	0.9750

A (12.71，13.29)

B (12.75，13.25)

C (12,14)

D (12.65，13.35)

提交

§7.2 区间估计

★ 单个正态总体参数的置信区间

例1: 设X~N(μ, 0.09), 测得X的4个观测值为12.6, 13.4,12.8, 13.2, 求μ的95%置信区间.

解: $\alpha=0.05$, 查表得$z_{\alpha/2}=z_{0.025}=1.96$

所以μ的95%置信区间为

$$(\bar{X}\pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}})=(13\pm1.96\times0.3/\sqrt{4})=(12.71,\ 13.29)$$

§6.2 正态总体的抽样分布

回顾

★ 单个正态总体场合

设$X\sim N(\mu,\sigma^2),X_1,X_2,\ldots,X_n$为它的简单随机样本，则有

1 $Z=\frac{\bar{X}-\mu}{\sigma}\sqrt{n}\sim N(0,1);$

2 $T=\frac{\bar{X}-\mu}{S}\sqrt{n}\sim t(n-1);$

3 在μ已知时，$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n);$

4 在μ未知时，$\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2}\sim\chi^2(n-1).$

§7.2 区间估计

★ 单个正态总体参数的置信区间

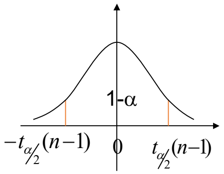
(2)、σ^2未知

选取枢轴量 $T=\frac{\bar{X}-\mu}{S}\sqrt{n}\sim t(n-1);$

§7.2 区间估计

★ 单个正态总体参数的置信区间

(2)、σ^2未知



对给定的α，有

$$P\{-t_{\alpha/2}(n-1) < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\} = 1-\alpha$$

得 $P\{\bar{X}-t_{\alpha/2}(n-1)\,S/\sqrt{n} < \mu < \bar{X}+t_{\alpha/2}(n-1)\,S/\sqrt{n}\}=1-\alpha$

所以，μ的置信度为$1-\alpha$的置信区间

$$(\bar{x}-t_{\alpha/2}(n-1)\,s/\sqrt{n},\ \bar{x}+t_{\alpha/2}(n-1)\,s/\sqrt{n}) = (\bar{x}\pm t_{\alpha/2}(n-1)\,s/\sqrt{n})$$

§7.2 区间估计

★ 单个正态总体参数的置信区间

2、设$X_1,\cdots,X_n \overset{iid}{\sim} N(\mu,\ \sigma^2)$，给定置信度$1-\alpha$，求$\sigma^2$(或$\sigma$) 的置信区间。

(1)μ已知

$$\boxed{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}} \sim \chi^2(n);$$

$$P\left\{\chi^2_{1-\alpha/2}(n) < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} < \chi^2_{\alpha/2}(n)\right\} = 1-\alpha$$

§7.2 区间估计

★ 单个正态总体参数的置信区间

得 $$P\left\{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)} < \sigma^2 < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right\} = 1-\alpha$$

得 $$P\left\{\sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}} < \sigma < \sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}}\right\} = 1-\alpha$$

§7.2 区间估计

★ 单个正态总体参数的置信区间

所以，σ^2与σ的置信度为**1-** α的置信区间分别为

$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right),$$

$$\left(\sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}}, \sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}}\right).$$

主观题 1分 设置

2、设$X_1,\cdots,X_n \overset{iid}{\sim} N(\mu,\ \sigma^2)$，给定置信度$1-\alpha$，求$\sigma^2$(或$\sigma$) 的置信区间。($\mu$未知时)

正常使用主观题需2.0以上版本雨课堂

作答

§7.2 区间估计

★ 单个正态总体参数的置信区间

2、设$X_1,\cdots,X_n \overset{iid}{\sim} N(\mu,\ \sigma^2)$，给定置信度$1-\alpha$，求$\sigma^2$(或$\sigma$) 的置信区间。

(2)μ未知

$$\because\ \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$$P\{\chi^2_{1-\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}(n-1)\} = 1-\alpha$$

§7.2 区间估计

★ 单个正态总体参数的置信区间

得 $$P\{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\} = 1-\alpha$$

$$P\{\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{\alpha/2}(n-1)}} < \sigma < \frac{\sqrt{n-1}S}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\} = 1-\alpha$$

所以，σ^2与σ的置信度为$1-\alpha$的置信区间分别为

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)\quad \left(\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{\alpha/2}(n-1)}}, \frac{\sqrt{n-1}S}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right)$$

表彰榜

池浩闻

上节回顾

★ 总结：用C-R不等式证明UMVUE的方法步骤

设总体$X \sim f(x;\theta)$, $X_1, X_2, \ldots, X_n$为X的一个样本，
证明$\hat{g}(X_1, X_2, \ldots, X_n)$是$g(\theta)$的UMVUE.

1. 证明$\hat{g}(X_1, X_2, \ldots, X_n)$是$g(\theta)$的无偏估计量，即验证$E(\hat{g}) = g(\theta)$;
2. 计算$D(\hat{g}(X_1, X_2, \ldots, X_n))$;
3. 计算C-R下界：
   (1)计算Fisher信息量$I(\theta) = E\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2 = -E\left[\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}\right]$.
   (2)C-R下界为$\frac{[g'(\theta)]^2}{nI(\theta)}$
4. 验证C-R下界与$D(\hat{g}(X_1, X_2, \ldots, X_n))$是否相等.

上节回顾

区间估计　　置信区间：　随机区间$(\hat{\theta}_1, \hat{\theta}_2)$

枢轴量法：(i)构造一个枢轴量G;
(ii)对于给定的置信度$1-\alpha$,确定两个常数a, b,使得
$P\{a < G(X_1, X_2, \ldots, X_n, \theta) < b\} = 1-\alpha$;
(iii)对$a < G(X_1, X_2, \ldots, X_n, \theta) < b$进行等价变换，若可以解出
$\hat{\theta}_1 < \theta < \hat{\theta}_2$的形式，则$(\hat{\theta}_1, \hat{\theta}_2)$是$\theta$的$1-\alpha$的置信区间.

单个正态总体区间估计：置信度1-α

§7.2 区间估计

★ 两个正态总体参数的置信区间

设$X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$,且相互独立，
$X_1, X_2, \ldots, X_{n_1}$，$Y_1, Y_2, \ldots, Y_{n_2}$为抽自这两个总体的样本，记

$$\overline{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i, \overline{Y} = \frac{1}{n_2}\sum_{j=1}^{n_2} Y_j,$$

$$S_1^2 = \frac{\sum_{i=1}^{n_1}(X_i - \overline{X})^2}{n_1 - 1}, S_2^2 = \frac{\sum_{j=1}^{n_2}(Y_j - \overline{Y})^2}{n_2 - 1},\text{则}$$

$$S_w^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}\text{称为混合样本方差.}$$

待估参数		枢轴量	随机变量的分布	置信区间的上、下限
μ	σ^2已知	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$N(0,\ 1)$	$\overline{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
	σ^2未知	$\frac{\overline{X}-\mu}{S/\sqrt{n}}$	$t(n-1)$	$\overline{X} \pm t_{\frac{\alpha}{2}}(n-1) \cdot \frac{S}{\sqrt{n}}$
σ^2	μ已知	$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$	$\chi^2(n)$	$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}$　$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}$
	μ未知	$\frac{(n-1)S^2}{\sigma^2}$	$\chi^2(n-1)$	$\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}$　$\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$

§6.2 正态总体的抽样分布

回顾

两个正态总体场合

1 在σ_1,σ_2已知时，$\overline{X}\pm\overline{Y}\sim N(\mu_1\pm\mu_2,\frac{\sigma_1^{\ 2}}{n_1}+\frac{\sigma_2^{\ 2}}{n_2})$，且有

$$\frac{\overline{X}\pm\overline{Y}-(\mu_1\pm\mu_2)}{\sqrt{\frac{\sigma_1^{\ 2}}{n_1}+\frac{\sigma_2^{\ 2}}{n_2}}}\sim N(0,1);$$

第六章 抽样分布

§6.2 正态总体的抽样分布

回顾

两个正态总体场合

2 虽然σ_1,σ_2未知，但是当$\sigma_1^{\ 2}=\sigma_2^{\ 2}$时，则

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{(n_1-1)S_1^2+(n_2-1)S_2^2}}\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}$$

$$=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2),$$

其中$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$称为混合样本方差；

第六章 抽样分布

§6.2 正态总体的抽样分布

回顾

两个正态总体场合

3 若μ_1,μ_2已知，则

$$F=\frac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2/n_1\sigma_1^2}{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2/n_2\sigma_2^2}\sim F(n_1,n_2);$$

第六章 抽样分布

§6.2 正态总体的抽样分布

回顾

两个正态总体场合

4 若μ_1,μ_2未知，则

$$F=\frac{\sum_{i=1}^{n_1}(X_i-\overline{X})^2/(n_1-1)\sigma_1^2}{\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2/(n_2-1)\sigma_2^2}\sim F(n_1-1,n_2-1).$$

§7.2 区间估计

★ 两个正态总体参数的置信区间

1、给定α，求$\mu_1-\mu_2$的区间估计.

(1)在σ_1,σ_2已知时，选取枢轴量

$$\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}}\sim N(0,1);$$

§7.2 区间估计

★ 两个正态总体参数的置信区间

1、给定α，求$\mu_1-\mu_2$的区间估计.

（2）虽然σ_1,σ_2未知，但是当$\sigma_1{}^2=\sigma_2{}^2=\sigma^2$时，选取枢轴量

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{(n_1-1)S_1^2+(n_2-1)S_2^2}}\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}$$

$$=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2),$$

§7.2 区间估计

★ 两个正态总体参数的置信区间

2、给定α，求$\frac{\sigma_1^2}{\sigma_2^2}$的区间估计.

(1)若μ_1,μ_2已知，则选取枢轴量

$$F=\frac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2/n_1\sigma_1^2}{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2/n_2\sigma_2^2}\sim F(n_1,n_2);$$

§7.2 区间估计

★ 两个正态总体参数的置信区间

2、给定α，求$\frac{\sigma_1^2}{\sigma_2^2}$的区间估计.

(2)若μ_1,μ_2未知，则

$$F=\frac{\sum_{i=1}^{n_1}(X_i-\overline{X})^2/(n_1-1)\sigma_1^2}{\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2/(n_2-1)\sigma_2^2}\sim F(n_1-1,n_2-1).$$

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

对一般总体X,若其均值为μ,方差为σ^2,只要样本容量n足够大,根据中心极限定理,可以认为

$\overline{X}$近似服从正态分布,

即$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$近似服从标准正态分布.

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

1.比例数(或概率)的近似置信区间估计

我们可以把比例看成是一个服从伯努利分布总体X的参数

$$X=\begin{cases}1,\text{表示抽到具有某种属性A的个体},\\0,\text{表示抽到不具有某种属性A的个体},\end{cases}$$

则$X\sim B(1,p), E(X)=p, D(X)=p(1-p)$.

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

1.比例数(或概率)的近似置信区间估计

设$X_1,X_2,\ldots,X_n$是X的一个样本,则$\overline{X}$即是在n次观察中A出现的频率$\dfrac{\mu_n}{n}$,其中$\mu_n=X_1+X_2+\cdots+X_n$,

$\overline{X}$作为p的估计,显然有

$E(\overline{X})=p, D(\overline{X})=\dfrac{p(1-p)}{n}$,当$n$足够大时,近似的有

$$\frac{\overline{X}-p}{\sqrt{p(1-p)/n}}\sim N(0,1).$$

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

1.比例数(或概率)的近似置信区间估计

选取枢轴量$\dfrac{\overline{X}-p}{\sqrt{p(1-p)/n}}$,

$$P\left\{-z_{\alpha/2}<\frac{\overline{X}-p}{\sqrt{p(1-p)/n}}<z_{\alpha/2}\right\}=1-\alpha.$$

求$\left|\dfrac{\overline{X}-p}{\sqrt{p(1-p)/n}}\right|<z_{\alpha/2}\Rightarrow(\overline{X}-p)^2<z_{\alpha/2}^2p(1-p)/n,$

$$\Rightarrow\left(1+z_{\alpha/2}^2/n\right)p^2-\left(2\overline{X}+z_{\alpha/2}^2/n\right)p+\overline{X}^2<0,$$

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

1.比例数(或概率)的近似置信区间估计

求解p的二次方程得

$$\hat{p}=\frac{n}{n+z_{\alpha/2}^2}\left[\overline{X}+z_{\alpha/2}^2/2n\pm z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}+\frac{z_{\alpha/2}^2}{4n^2}}\right],$$

当n足够大时，略去$z_{\alpha/2}^2/n$，得p得置信度为

$1-\alpha$得置信区间为

$$\left(\overline{X}-z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}},\overline{X}+z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\right).$$

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

2.泊松分布中参数得近似置信区间：

设总体X服从参数为λ的泊松分布$P(\lambda)$,

$X_1,X_2,\ldots,X_n$是X的一个样本,则$\overline{X}$是λ的点估计.

由$E(\overline{X})=\lambda, D(\overline{X})=\dfrac{\lambda}{n}$,当$n$足够大时，近似的有

$$\frac{\overline{X}-\lambda}{\sqrt{\lambda/n}}\sim N(0,1).$$

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

2.泊松分布中参数得近似置信区间：

选取枢轴量$\dfrac{\overline{X}-\lambda}{\sqrt{\lambda/n}}$,

$$P\left\{-z_{\alpha/2}<\frac{\overline{X}-\lambda}{\sqrt{\lambda/n}}<z_{\alpha/2}\right\}=1-\alpha.$$

求$\left|\dfrac{\overline{X}-\lambda}{\sqrt{\lambda/n}}\right|<z_{\alpha/2}\Rightarrow(\overline{X}-\lambda)^2<z_{\alpha/2}^2\lambda/n$

$$\Rightarrow\lambda^2-\left(2\overline{X}+z_{\alpha/2}^2/n\right)\lambda+\overline{X}^2<0$$

§7.2 区间估计

★ 非正态总体均值的区间估计（大样本场合）

2.泊松分布中参数得近似置信区间：

求解λ的二次方程得

$$\hat{\lambda}=\left[\overline{X}+z_{\alpha/2}^2/2n\pm z_{\alpha/2}\sqrt{\frac{\overline{X}}{n}+\frac{z_{\alpha/2}^2}{4n^2}}\right],$$

当n足够大时，略去$z_{\alpha/2}^2/n$，得λ得置信度为

$1-\alpha$得置信区间为

$$\left(\overline{X}-z_{\alpha/2}\sqrt{\frac{\overline{X}}{n}},\overline{X}+z_{\alpha/2}\sqrt{\frac{\overline{X}}{n}}\right).$$

主观题 1分 设置

例：在某综艺节目收视率的调查中抽取1000户家庭，其中有401户收看，试求该收视率p的置信度为95%的置信区间.

正常使用主观题需2.0以上版本雨课堂

作答

主观题 10分 设置

复习题1：

已知某种零件尺寸服从方差 1.21的正态分布，对一批这类零件检查 6 件，得尺寸数据（单位：mm）32.56 29.66 31.64 30.00 31.37 31.03

问：n等于多少才能使由样本$X_1, X_2, \dots, X_2$构成的参数的 95%置信区间的长度小于 1?

作答

主观题 10分 设置

复习题2：

设总体X服从 $f(x;\theta)=\frac{2\theta}{x^3}e^{-\theta/x^2}$，$x>0$，$f(x;\theta)=0$,其他，参数$\theta>0$未知，

$X_1, X_2, \dots, X_n$ 为它的样本，求 $\frac{1}{\theta}$ 的MLE，

并考察其是否是$\frac{1}{\theta}$的UMVUE，若是请给出证明，若不是请说明理由.

作答

解：

似然函数为$L(\theta)=\prod_{i=1}^{n}\frac{2\theta}{x_i^3}e^{-\frac{\theta}{x_i^2}}, x_i>0, i=1,2,\dots,n.$

$\ln L(\theta)=\sum_{i=1}^{n}(\ln 2\theta-3\ln x_i)-\sum_{i=1}^{n}\frac{\theta}{x_i^2},$

$\frac{\partial \ln L(\theta)}{\partial\theta}=\frac{n}{\theta}-\sum_{i=1}^{n}\frac{1}{x_i^2}=0\Rightarrow\theta=\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i^2}}$，所以$\hat{\theta}_{MLE}=\frac{n}{\sum_{i=1}^{n}\frac{1}{X_i^2}}.$

由不变性知：$\frac{\hat{1}}{\theta_{MLE}}=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i^2}$

$$E\left(\frac{\hat{1}}{\theta_{MLE}}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i^2}\right)=E\left(\frac{1}{X^2}\right)=\int_0^{+\infty}\frac{1}{x^2}\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}dx$$

$=\frac{1}{x^2}e^{-\frac{\theta}{x^2}}\Big|_0^{+\infty}-\int_0^{+\infty}e^{-\frac{\theta}{x^2}}d\frac{1}{x^2}=\frac{1}{\theta}$,所以$\frac{\hat{1}}{\theta_{MLE}}$是$\frac{1}{\theta}$的无偏估计.

$$D\left(\frac{\hat{1}}{\theta_{MLE}}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i^2}\right)=\frac{1}{n}\left(E\left(\frac{1}{X^4}\right)-\frac{1}{\theta^2}\right),$$

$$E\left(\frac{1}{X^4}\right)=\int_0^{+\infty}\frac{1}{x^4}\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}dx=\frac{1}{x^4}e^{-\frac{\theta}{x^2}}\Big|_0^{+\infty}-\int_0^{+\infty}e^{-\frac{\theta}{x^2}}d\frac{1}{x^4}$$

$$=\frac{2}{\theta}\int_0^{+\infty}\frac{1}{x^2}\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}dx=\frac{2}{\theta^2},$$

所以$D\left(\frac{\hat{1}}{\theta_{MLE}}\right)=\frac{1}{n\theta^2}$

$$\ln f(x;\theta)=\ln 2+\ln\theta-3\ln x-\frac{\theta}{x^2}$$

$$\frac{\partial\ln f(x;\theta)}{\partial\theta}=\frac{1}{\theta}-\frac{1}{x^2},\frac{\partial^2\ln f(x;\theta)}{(\partial\theta)^2}=-\frac{1}{\theta^2},$$

$$I(\theta)=-E\left(\frac{\partial^2\ln f(X;\theta)}{(\partial\theta)^2}\right)=\frac{1}{\theta^2},$$

因此C-R下界为$\frac{((1/\theta)')^2}{nI(\theta)}=\frac{1}{n\theta^2}=D\left(\frac{\hat{1}}{\theta_{MLE}}\right)$,

所以$\frac{\hat{1}}{\theta_{MLE}}$是$\frac{1}{\theta}$的UMVUE.