福建师范大学闽南科技学院教案

编号：0701

课时安排： 3 学时	教学课型：理论课 √ 实验课☐ 习题课☐ 实践课☐ 其它☐

题目（教学章、节或主题）：

Ch7 参数估计

§7.1 点估计

§7.2 基于截尾样本的最大似然估计

教学目的要求（分掌握、熟悉、了解三个层次）：

1） 了解参数估计

2） 掌握参数点估计

3） 掌握矩估计法

4） 掌握极大似然估计

教学内容（注明：* 重点 # 难点 ？疑点）：

一. 参数估计 （10分钟）

参数估计问题：一个总体 X 的分布函数，可以用 $F(x,\theta)$ 表示，其中 X 是一个随机变量， 是一个未知的参数，如 $X \sim N(\mu,\sigma^2)$，其中 μ,σ^2 为未知参数。在实际中，总体 $F(x,\theta)$ 的参数通常是未知的，这时可以利用样本提供的信息，对参数有一个基本的估计，这就是参数估计。

例1 设某个厂生产彩电寿命X服从正态分布 $N(\mu,\sigma^2)$，但平均寿命 μ 未知，于是厂家抽查了100台彩电，测得这100台彩电的样本均值 X =4.2万小时，将这个 X =4.2万小时做为该厂彩电平均寿命 μ 的一个估计，这种方法叫做参数估计，参数估计分为点估计和区间估计。

二．点估计

（一）点估计介绍 （5分钟）

定义7.1.1 设总体X的分布函数为 $F(x,\theta)$，其中 θ 为所有未知参数，从总体中 X 抽取样本 $X_1,X_2,\cdots,X_n$，其观察值为 $X_1,X_2,\cdots,X_n$。构造某个统计量 $\hat{\theta}$（$X_1,X_2,\cdots,X_n$），用它的观察值来估计未知参数 θ，则称 $\hat{\theta}$（$X_1,X_2,\cdots,X_n$）为 θ 的估计值，且称此为 θ 的估计量。象这种估计称为 θ 的一个点估计。例如在上述例子里。可以将 $\hat{\mu}=\bar{x}=4.2$ 万小时做为平均寿命一个估量值。

从定义看，任何一个统计量都可以作为未知参数的一个估计量，但是否合理需要验证，下面介绍两种符合一定理论的估计方法。

（二）矩估计法 （45分钟）

设在总体X的分布函数 $F(x,\theta)$ 中，θ 代表 $\theta_1,\theta_2,\ldots\theta_k$ 未知参数（有 k 个参数）。假定总体的矩的 k 个原点矩存在，则这些矩可表示为 θ 的函数，即 $E(X^r)=g_r$（$\theta_1,\theta_2,\ldots\theta_k$）其中 r =1，2，…，k．取r阶样本原点矩A_r 作为总体k阶原点矩的估计量，即有方程组

$$\begin{cases} A_1 = g_1(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k) \\ A_2 = g_2(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k) \\ .... \\ A_k = g_k(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k) \end{cases} \quad (1)$$

其中$A_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r$，r=1，2，。。。，k 解出（1）式得

$$\begin{cases} \hat{\theta}_1 = f_1(A_1, A_2, ..., A_k) \\ \hat{\theta}_2 = f_2(A_1, A_2, ..., A_k) \\ ...... \\ \hat{\theta}_k = f_k(A_1, A_2, ..., A_k) \end{cases}$$

则称$\hat{\theta}_r$叫做θ_r的矩估计量，这种方法称为矩估计法。

例2 设总体X具有均值μ和方差σ^2，但未知，求它们的矩估计量。

解 设从总体X中具一个样本$X_1, X_2, \cdots, X_n$，则样本矩$A_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}$， $A_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

而总体矩$E(X) = \mu$，$E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2$（$\mu, \sigma$均未知）

由矩法估计，用样本矩作为总体矩的估计，得

$$\begin{cases} A_1 = \hat{\mu} \\ A_2 = \hat{\sigma}^2 + \hat{\mu}^2 \end{cases}$$

则总体均值μ可用样本均值来估计，总体方差σ^2可用样本二阶中心矩来估计。

例3 在某校初三毕业班期末语文成绩中随机取9人，分数分别为78，75，85，71，89，65，55，63，94。试估计该校初三毕业班语文成绩的平均分和方差。

解 $\mu = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{9}\sum_{i=1}^{9} x_i = 75$

$$\hat{\sigma}_2 = B_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \approx 132.67$$

例4 总体$X \sim U[a,b]$，其中a, b未知，试求a, b的矩估计量。

解 E（X）=$\frac{a+b}{2}$，D（X）=$\frac{(a-b)^2}{12}$，则E（X^2）=D（X）+E（X）^{2}=$\frac{1}{12}(a-b)^2 + \frac{1}{4}(a+b)^2$

设$\overline{X}$为样本均值，B_2样本二阶中心矩，根据矩估计法

$$\begin{cases} \overline{X} = E(X) = \dfrac{\hat{a}+\hat{b}}{2} \\ B_2 + (\overline{X})^2 = \dfrac{1}{12}(\hat{a}-\hat{b})^2 + \dfrac{1}{4}(\hat{a}+\hat{b})^2 \end{cases}$$

得 $\hat{a} = \overline{X} - \sqrt{3}B_2 \qquad \hat{b} = \overline{X} + \sqrt{3}B_2$

（三）极大似然法估计 （15分钟）

极大似然法估计是急于概率最大事件最可能出现思想来估计参数。下面结合例子介绍极大似然估计法的思想和方法。

设一袋中装有红与白两球，设P是从袋中随机摸到白球的概率，p是未知的，现估计p，做放回抽样，摸球10次，其结果是3次白球7次红球，则发生这种事件的概率L（P）=P^3（1-P）^{7}

在随机抽取中有3次白球7次红球，说明发生这种事件的概率很大，即L（P）很大，问什么时候是最大，则发生该事件组最有可能？

$$\frac{dL(p)}{dp} = 0,\quad 3p^2(1-p)^7 > p^3(1-p)^6 = 0$$，得p=0.3

即p=0.3时，L（P）最大，发生该事件概率最大，根据这种思想则可以求出参数p。

1）似然函数（10分钟）

设总体X是离散型随机变量，分布律$P(X=x)=p(x,\theta)$，其中θ是未知参数，当样本$X_1, X_2, \cdots, X_n$得到一组观察值$x_1, x_2, \cdots, x_n$时，由样本的独立同分布，记$L(\theta) = P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2)\cdots P(X_n = x_n) = \prod_{i=1}^{n} p(x_i, \theta)$ 称$L(\theta)$为似然函数。同理可以定义连续函数的似然函数$L(\theta)=\prod_{i=1}^{n} f(x_i, \theta)$.

2）极大似然估计法（25分钟）

根据极大似然估计法的思想，就是估计参数θ，使极大似然函数值最大。

A 若只有一个参数，则可解$\dfrac{dL(\theta)}{d\theta}$=0 ，得$\theta = \hat{\theta}$做为参数估计值

B 若多于一个以上 参数，则解方程组 $\begin{cases} \dfrac{\partial L(\theta)}{\partial \theta_1} = 0 \\ \dfrac{\partial L(\theta)}{\partial \theta_2} = 0 \\ \ldots \\ \dfrac{\partial L(\theta)}{\partial \theta_n} = 0 \end{cases}$ 解得$\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$，将它们做为参数$\theta_1, \theta_2, \ldots, \theta_n$的

估计值。特别有$\begin{cases}\dfrac{\partial\ln(\theta)}{\partial\theta_1}=0\\ \dfrac{\partial\ln(\theta)}{\partial\theta_2}=0\\ \cdots\\ \dfrac{\partial\ln(\theta)}{\partial\theta_n}=0\end{cases}$

例5 设总体$X\sim B(m,p)$，$X_1,X_2,\cdots,X_n$为样本观察值，求p的极大似然估计值和极大似然估计量。（15分钟）

解：总体X的分布律为$P(X=x)=p(x,p)=C_m^x p^x(1-p)^{m-x}(x=0,1,2,\cdots,m)$.

所以似然函数为

$$L(p)=\prod_{i1=}^{n}p(x_i,p)=\prod_{i=1}^{n}C_m^{x_i}p^{x_i}(1-p)^{m-x_i}=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{\sum\limits_{i=1}^{n}(m-x_i)}\prod_{i=1}^{n}C_m^{x_i}$$

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

$$\frac{d\ln L(p)}{dp}=\frac{1}{p}\sum_{i=1}^{n}x_i-\frac{1}{1-p}(\sum_{i=1}^{n}m-x_i)=0$$ 解得 $\hat{p}=p=\dfrac{\frac{1}{n}\sum\limits_{i=1}^{n}x_i}{m}=\dfrac{\bar{x}}{m}$ 即用$\dfrac{\bar{x}}{m}$做为极大似然估计值,估计量为$\widehat{p}=\dfrac{\bar{x}}{m}$.

例6 该总体$X\sim N(\mu,\sigma^2)$，其中μ,σ^2都是未知参数，$X_1,X_2,\cdots,X_n$为样本观察值。（15分钟）

解 似然函数L（μ,σ^2）$=\prod\limits_{i=1}^{n}f(x_i,\mu,\sigma^2)=\prod\limits_{i=1}^{n}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$＝（2$\pi)^{-\frac{n}{2}}$（$(\sigma^2)^{-\frac{n}{2}}e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2}$

lnL（μ,σ^2）$=-\dfrac{n}{2}$ln（2$\pi)-\dfrac{n}{2}\ln(\sigma^2)-\dfrac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2$ （σ^2作为一个整体）

$$\begin{cases}\dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\mu}=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)=0\\ \dfrac{\partial\ln L(\mu,\sigma^2)}{\partial\sigma^2}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum\limits_{i=1}^{n}(x_i-n)^2=0\end{cases}$$ 得$\begin{cases}\widehat{\mu}=\bar{x}\\ \widehat{\sigma}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2=B_2\end{cases}$ （同矩估计值一样）

教学方式、手段：讲授、示教

教学媒介：教科书、板书、多媒体课件

板书设计：

定义，定理，性质等一些概念性的内容	图形，例题，习题的演算

思考题、作业：

思考题：1。矩估计与极大似然估计得到的结果是否会一样的？

作业：习题7、 1、2、4

参考书目：

1.《概率论》，复旦大学编，北京：高等教育出版社，1979。

2.《概率论与数理统计》，浙江大学数学系高等数学教研组编，人民教育出版社，1979。

3.《概率论与数理统计》，王松桂，程维虎，高旅端编，北京：科学出版社，2000

教师姓名：傅金波　　　　职称：讲师　　　　年　月　日

福建师范大学闽南科技学院教案

编号：0702

课时安排：　3 学时	教学课型：理论课 √　实验课□　习题课□　实践课□　其它□

题目（教学章、节或主题）：

CH7 参数估计

§7.3 估计量的评价标准

§7.4 区间估计

教学目的要求（分掌握、熟悉、了解三个层次）：

1) 掌握参数估计中三种常用评价标准及它们的使用，
2) 确定置信区间以及正态总体的置信区间

教学内容（注明：* 重点　# 难点　? 疑点）：

一．估计量的评价标准

（一） 无偏性（20分钟）

定义：设$\hat{\theta}=\hat{\theta}(X_1,X_2,\dots,X_n)$的数学期望等于参数$\theta$，即E（$\hat{\theta}$）$=\theta$，则称$\hat{\theta}(X_1,X_2,\dots,X_n)$是参数$\theta$的无偏估计量，反之称为有偏估计量。

例1 设总体X服从任意分布，且E（X）$=\mu, D(X)=\sigma^2$，$X_1,X_2,\cdots,X_n$是样本。证明样本均值$\overline{X}$和样本方差S^2分别是μ和σ^2的无偏估计。（与矩法估计的$\overline{X}$和B_2比较）

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

$$E(S^2)=E(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-X)^2)=\frac{1}{n-1}E(\sum_{i=1}^{n}(X_i-\overline{X})^2)=\frac{1}{n-1}(\sum_{i=1}^{n}EX_i^{\ 2}-nE(\overline{X})^2)$$

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

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

$$=\sigma^2$$

所以$E(S^2)=\sigma^2$为无偏估计

例2 设X~N（$\mu,1^2$），X_1,X_2为其随机样本，考虑μ的三个统计量

$$\begin{cases}\hat{\mu}_1=1/3X_1+2/3X_2\\ \hat{\mu}_2=1/4X_1+3/4X_2\\ \hat{\mu}_3=1/2X_1+1/2X_2\end{cases}$$

证明它们都是μ的无偏估计

证明：E（$\hat{\mu}_1$）=1/3 $E(X_1)$ +2/3 $E(X_2)$ =1/3E（X）+2/3E（X）=E（X）= μ

同理可E（$\hat{\mu}_2$）=E（$\hat{\mu}_2$）= μ

（二） 有效性（10分钟）

定义 设$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, ..., X_n)$和$\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, ..., X_n)$是$\theta$的两个无偏估计量，若D（$\hat{\theta}_1$) <= $D(\hat{\theta}_2)$

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

当θ的两个无偏估计量$\hat{\theta}_1$和$\hat{\theta}_2$的取值都在θ周围波动，但若$\hat{\theta}_1$取值比$\hat{\theta}_2$取值更集中在θ的附近，便认为以$\hat{\theta}_1$来估计$\hat{\theta}_2$更好些。

例3 设总体$X \sim N(\mu, 1)$，X_1, X_2为其样本，记

$\hat{\mu}_1 = 1/3X_1 + 2/3X_2, \hat{\mu}_2 = 1/4X_1 + 3/4X_2, \hat{\mu}_3 = 1/2X_1 + 1/2X_2, \hat{\mu}_4 = 2/5X_1 + 3/5X_2$，这四个无偏估计量中，最有效的是$\mu_3$。

解： D（$\hat{\mu}_1$）=5/9; D（$\hat{\mu}_2$ ）=10/16; D（$\hat{\mu}_3$）=1/2; D（$\hat{\mu}_4$）=13/25.

故最有效的量是$\hat{\mu}_3$

（三） 一致性（10 分钟）

定义 若对任意$\varepsilon > 0,$有 $\lim\limits_{n\to\infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1$, 则称$\hat{\theta}_n$是$\theta$的一致估计量。此性质在样本容量很大时才起作用

例4 样本均值$\overline{X}$是总体数学期望μ的一致估计量.

证明 设E（X）= μ，D（X）= σ^2，$X_1, X_2, \cdots, X_n$是其样本，则E（$\overline{X}$）= μ，D（$\overline{X}$ ）= $\frac{\sigma^2}{n}$

利用切比雪夫不等式P{$\left|\overline{X} - \mu\right| \geq \xi$}=P{$\left|\overline{X} - E(\overline{X}) \geq \xi\right|$}$\leq \frac{D(\overline{X})}{\xi^2} = \frac{\sigma^2}{n\xi^2}$

$$P\{\left|\overline{X} - \mu\right| \prec \xi\} = P\{\left|\hat{\mu} - \mu\right| \prec \xi\} \geq 1 - \frac{\sigma^2}{n\xi^2}$$

$$\lim_{n\to\infty} P\{|\hat{\mu} - \mu| < \xi| = 1, \text{故 } \hat{\mu} = \overline{X} \text{ 是 } \mu \text{ 的一致估计量.}$$

二．正态总体参数的区间估计 （20分钟）

例5 设总体$X \sim N(\mu, \sigma^2)$，其中$\sigma^2 = 1$，μ未知。对给定的样本$X_1, X_2, ..., X_n$，由极大似然估计知μ的点估计为$\hat{\mu} = \overline{X}$，现在来找区间估计。

定义 设θ是总体X的未知参数，若有由样本$X_1,X_2,...,X_n$确定的两个统计量$\underline{\theta}=\underline{\theta}(X_1,X_2,...,X_n)$及$\overline{\theta}=\overline{\theta}(X_1,X_2,...,X_n)$，对给定的$\alpha$ $(0<\alpha<1)$，满足$P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha$，

便称随机区间$(\underline{\theta},\overline{\theta})$为未知参数$\theta$的置信度为$1-\alpha$的置信区间，并分别称$\underline{\theta}$及$\overline{\theta}$是$\theta$的置信下限及置信上限。而确定$\theta$的置信区间的过程，称为$\theta$的区间估计。在不致混淆的情况下，置信区间及其观察区间统称为置信区间，在具体抽样后所求得的置信区间均指相应的观察区间。

对于给定的置信度$1-\alpha$和来自总体X的容量为n的样本$X_1,X_2,...,X_n$，可按以下步骤确定未知参数θ的置信区间：

（1） 取一个未知参数θ的较优的点估计$\hat{\theta}(X_1,X_2,...,X_n)$，最好是无偏的；

（2） 从$\hat{\theta}(X_1,X_2,...,X_n)$出发，找一个样本函数$W=W(X_1,X_2,...,X_n;\ \theta)$，其分布已知，且只含惟一一个未知参数$\theta$；

（3） 确定适当的常数a和b，使$P\{a<W<b\}=1-\alpha$；

（4） 由$a<W<b$解出θ，得出其等价形式$\underline{\theta}(X_1,X_2,...,X_n)<\theta<\overline{\theta}(X_1,X_2,...,X_n)$。于是，$(\underline{\theta},\overline{\theta})$是$\theta$的置信度为$1-\alpha$的置信区间，这时有$P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha$。

（一）．单个正态总体参数的区间估计

1. 设总体$X\sim N(\mu,\sigma^2)$，σ^2为已知，又$X_1,X_2,...,X_n$为总体样本，求μ的置信度为$1-\alpha$的置信区间。（20 分钟）

2. 设总体$X\sim N(\mu,\sigma^2)$，σ^2未知。又$X_1,X_2,...,X_n$为总体样本，求μ的置信度为$1-\alpha$的置信区间。（15 分钟）

例 6 从刚生产出来的一大堆钢珠中随机抽出 7 个，测量它们的直径（单位：mm）为

5.52，5.41，5.18，5.32，5.64，5.22，5.76

若钢珠直径$X\sim N(\mu,\sigma^2)$，试求总体均值μ的置信度为 95%的置信区间：（1）若已知$\sigma=0.16$；（2）若σ未知。

3. 设总体$X\sim N(\mu,\sigma^2)$，μ已知，又$X_1,X_2,...,X_n$为总体样本，试求σ^2的置信度为$1-\alpha$的置信区间。（20 分钟）

4. 设总体$X\sim N(\mu,\sigma^2)$，μ未知，又$X_1,X_2,...,X_n$为总体样本，试求σ^2及σ的置信度为$1-\alpha$的置信区间。（20 分钟）

例 7 设某自动车床加工的一种零件的长度$X\sim N(\mu,\sigma^2)$。现从产品中随机抽取 16 件，测得它们的长度（单位：mm）为：

12.15， 12.12， 12.01， 12.28， 12.08， 12.16， 12.03， 12.06

12.01， 12.13， 12.07， 12.11， 12.08， 12.01， 12.03， 12.06

试求长度 X 标准差 σ 的置信度为 0.95 的置信区间。

教学方式、手段：讲授、示教、练习

教学媒介：教科书、板书

板书设计：

定义，定理，性质等一些概念性的内容	图形，例题，习题的演算

讨论、思考题、作业：

作业：已知 $X_1, X_2, \cdots, X_n$ 为指数分布 f（x）= $\begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, x \succ 0 \\ 0, x \le 0 \end{cases}$ 的一个样本。

证明 $\theta_1 = \overline{X} = 1/n \sum X_i$ 是 θ 的一个极大似然无偏估计量。

参考书目：

1.《概率论》，复旦大学编，北京：高等教育出版社，1979。

2.《概率论与数理统计》，浙江大学数学系高等数学教研组编，人民教育出版社，1979。

3.《概率论与数理统计》，王松桂，程维虎，高旅端编，北京：科学出版社，2000

教师姓名：傅金波　　　职称：讲师　　　年　月　日

福建师范大学闽南科技学院教案

编号：0703

课时安排：　2 学时	教学课型：理论课 √　实验课□　习题课□　实践课□　其它□

题目（教学章、节或主题）：

GH7 参数估计

§7.5 正态总体参数的区间估计

教学目的要求（分掌握、熟悉、了解三个层次）：

1）掌握双正态总体参数的区间估计

教学内容（注明：* 重点　# 难点　？疑点）：

一．两个正态总体参数的区间估计情形

1. 设正态总体 $X\sim N(\mu_1,\sigma_1^2)$ 与正态总体 $Y\sim N(\mu_2,\sigma_2^2)$ 相互独立，$X_1,X_2,\cdots,X_{n_1}$，$Y_1,Y_2,\cdots,Y_{n_2}$ 分别为总体 X，Y 的独立样本，且方差 σ_1^2 和 σ_2^2 均为已知，求 μ_1-μ_2 的置信度为1-α 的置信区间。（20分钟）

引用服从标准正态分布的统计量 $\dfrac{(X-Y)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$，　因为

$$P\{\left|\frac{(\bar{X}-\bar{Y}-(\mu_1-\mu_2\ ))}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\right|\prec U_{\alpha/2}\}=1-\alpha$$

于是有 μ_1-μ_2 的置信度为1-α 的置信区间为

$$\left(\bar{X}-\bar{Y}-U_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ \bar{X}-\bar{Y}+U_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

2. 设正态总体 $X\sim N(\mu_1,\sigma_1^2)$ 与正态分布 $Y\sim N(\mu_2,\sigma_2^2)$ 相互独立，$X_1,X_2,\cdots,X_n$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别为总体 X，Y 的样本，且方差 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知，求 $\mu_1-\mu_2$ 的置信度为1-α 的置信区间。（30分钟）

此时，利用第六章th6.4中统计量

$$t=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t（n_1+n_2-2），其中S_w=\sqrt{S_w^2}=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

不难由此求得 $\mu_1-\mu_2$ 的置信度为1-α 的置信区间为

$$\left(\bar{X}-\bar{Y}-t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{X}-\bar{Y}-t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

例1　为了估计灌溉制度对水稻产量的影响，现选20块条件大致相同的土地。10块采用A种灌溉制度，另外10块采用B种灌溉制度，得亩产量（单位：kg）

A种亩产量：326，332，335，340，355，367，363，368，374，366

B种亩产量：264，270，275，276，299，312，290，306，262，286

设两种灌溉制度下亩产量都服从正态分布，且方差相同，求两种灌溉制度下平均亩产量之差 $\mu_1-\mu_2$ 的置信度为0.90的置信区间。

解　本题 $n_A=10, n_B=10, \alpha=0.10, t_{\alpha/2}(n_A+n_B-2)=t_{0.05}(18)=2.5524,$

经计算得：$\bar{x}_A=352.6, s_A^2=310.7, \bar{x}_B=284, s_B^2=306.4$

从而有

$$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}=308.55,$$

S_w=17.5656

故所求置信区间为（48.54，88.66）

3. 设正态总体 $X\sim N(\mu_1,\sigma_1^2)$ 与正态总体 $Y\sim N(\mu_2,\sigma_2^2)$ 相互独立，$\overline{X}, S_1^2$ 是总体 X 的容量为 n_1 的样本均值和样本方差 $\overline{Y}, S_2^2$ 是总体 Y 的容量为 n_2 的样本均值和样本方差。$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均为未知，求 σ_1^2/σ_2^2 的置信度为1-α 的置信区间。（30分钟）

依条件，统计量

$$(n_1-1)S_1^2/\sigma_1^2\sim\chi^2(n_1-1)$$

$$(n_2-1)S_2^2/\sigma_2^2\sim\chi^2(n_2-1)$$

且相互独立，故由F分布定义

$$\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}=\frac{(n_1-1)S_1^2/[\sigma_1^2(n_1-1)}{(n_2-1)S_2^2/[\sigma_2^2(n_2-1)]}\sim F(n_1-1,n_2-1)$$

上述F统计量不含除 σ_1^2,σ_2^2 之外其他未知参数，所以由此可得

$$P\{F_{1-\alpha/2}(n_1-1,n_2-1)\prec\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}\prec F_{\alpha/2}(n_1-1,n_2-1)\}=1-\alpha$$

得到 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信度为1-α 的置信区间为

$$\left(\frac{S_1^2}{S_2^2F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2F_{1-\alpha/2}(n_1-1,n_2-1)}\right)$$

例2　两正态总体N（μ_1,σ_1^2），N（μ_2,σ_2^2）的参数均为未知。依次取容量为13，10的两独立样本，测得样本方差 $s_1^2=8.41, s_2^2=5.29$，求两总体方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信度为90%的置信区间。（10分钟）

因为 $n_1 = 12, n_2 = 10, \alpha = 0.10$ 查附表得

$$F_{\alpha/2}(n_1-1, n_2-1) = F_{0.05}(12,9) = 3.07$$

$$F_{1-\alpha/2}(n_1-1, n_2-1) = F_{0.95}(12,9) = \frac{1}{F_{0.05}(9,12)} = \frac{1}{2.80}$$

而 $\frac{s_1^2}{s_2^2} = \frac{8.41}{5.29} = 1.59$ 所以 $\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间为

（1.59/3.07，1.59/2.80）=（0.52，4.45）

教学方式、手段：讲授、示教

教学媒介：教科书、板书、多媒体课件

板书设计：

定义，定理，性质等一些概念性的内容	图形，例题，习题的演算

思考题、作业：

讨论：置信区间长度取值与准确值之间如何调整？

作业：习题7 13、16、19

参考书目：

1.《概率论》，复旦大学编，北京：高等教育出版社，1979。

2.《概率论与数理统计》，浙江大学数学系高等数学教研组编，人民教育出版社，1979。

3.《概率论与数理统计》，王松桂，程维虎，高旅端编，北京：科学出版社，2000

教师姓名：傅金波　　　　职称：讲师　　　　年　月　日

福建师范大学闽南科技学院教案

编号：0704

课时安排：　2 学时	教学课型：理论课√　实验课□　习题课□　实践课□　其它□

题目（教学章、节或主题）：

CH7 参数估计

§7.6 （0－1）分布参数的区间估计

§7.7 单侧置信区间

教学目的要求（分掌握、熟悉、了解三个层次）：

1) 总体分布已知时，对未知参数用大样本作区间估计.
2) 在总体分布未知下利用大样本进行总体分布参数区间估计的问题.

教学内容（注明：* 重点　# 难点　? 疑点）：

1.（0-1）分布参数的区间估计（30分钟）

设总体X服从（0-1）分布，分布律为　f（x；p）=p^x（1-p）$^{1-x}$ x=0，1，其中参数p（$0 \prec p \prec 1$）未知，欲求其置信度为1-α的置信区间。就要利用某些有关统计量的极限分布对参数作近似的区间估计。利用极限分布进行参数估计，要求样本容量比较大，这样的样本称为大样本。根据经验，样本容量n≥50为好。

现设$X_1, X_2, \cdots, X_n$为总体X的大样本，因此$\sum X_i$服从二项分布$B(n,p)$，其均值与方差分别是np和np（1-p）。根据中心极限定理，当n很大时，下述标准化随机变量

Z=$\dfrac{\sum X_i - np}{\sqrt{np(1-p)}} = \dfrac{n\overline{X} - np}{\sqrt{np(1-p)}}$渐近服从于N（0，1）从标准正态分布表中查得双侧$\alpha$分位点$U_{\alpha/2}$，

使P{$\left|\dfrac{n\overline{X} - np}{\sqrt{np(1-p)}}\right| \prec U_{\alpha/2}$} $\approx 1-\alpha$

括号内的不等式等价于（n$\overline{X}$-np）$^2 \prec U_{\alpha/2}^2$np（1-p）

经整理得　（n+$U_{\alpha/2}^2$）P^2-（2n$\overline{X}$+$U_{\alpha/2}^2$）P+n$\overline{X}^2 \prec 0$，改写该二次不等式为 ap^2+bp+c$\prec 0$

解得参数p的置信度近似为1-α的置信区间

（p_1, p_2）=（$\dfrac{-b-\sqrt{b^2-4ac}}{2a}, \dfrac{-b+\sqrt{b^2-4ac}}{2a}$）

其中，a=n+$U_{\alpha/2}^2$，b=-（2n$\bar{X}$+$U_{\alpha/2}^2$），C=n$\overline{X}^2$

例1　从一批同类产品中抽取的100个样品中含60个一级品，试求该批产品一级品率p的置信度为95%的置信区间。

解　以X记随机抽取的一件产品中含一级品的数目，则X服从参数p的（0-1）分布，样本容量n=100，属于大样本，样本均值$\bar{x}$=60/100=0.6。又因$\alpha = 0.05, U_{0.05/2} = 1.96, U_{0.05/2}^2 = 3.84$

所以$a=100+3.84=103.84$，$b=-(2\times 100\times 0.6+3.84)=-123.84$，

$c=100\times 0.6\times 0.6=36$。

$$(p_1,p_2)=(0.50,0.69)$$

即大约有95%的可靠性说一级品率p介于区间（0.50，0.69）。

2　大样本下总体均值的区间估计（30分钟）

设总体X具有任意的分布，且其均值　$\mu=E(X)$和方差σ^2=D（X）都存在，但均未知。试用样本$X_1,X_2,\cdots,X_n$对总体的均值μ作区间估计。

考虑μ的点估计$\overline{X}$，有中心极限定理知，当n很大时，$\overline{X}$近似服从正态分布，又E（$\overline{X}$）=μ，D（$\overline{X}$）=$\frac{\sigma^2}{n}$，所以$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$渐进服从N（0，1）。

给定α，查标准正态分布的上侧$\alpha/2$分位点$U_{\alpha/2}$，p{$|Z|\prec U_{\alpha/2}$}=P{$\left|\frac{\overline{X}-\mu}{S/\sqrt{n}}\right|\prec U_{\alpha/2}$}$\approx 1-\alpha$

则区间（$\overline{X}-\frac{S}{\sqrt{n}}U_{\alpha/2},\overline{X}+\frac{S}{\sqrt{n}}U_{\alpha/2}$）是$\mu$的置信度等于1-$\alpha$的置信区间。

例2　从轮胎制造厂生产的一大批同类轮胎中，随机抽取100只进行测试，求得平均寿命$\bar{x}$=32000公里，样本的标准差s=4000公里，试求该批轮胎平均寿命μ的95%的置信区间。

解　此题寿命总体X的分布类型。均值及方差均未知。样本的容量n=100比较大，可以考虑用大样本对轮胎的平均寿命μ作区间估计。

已知$\bar{x}$=32000s=4000，查出$U_{0.05/2}=1.96$，求得

（$\bar{x}-\frac{s}{\sqrt{n}}U_{\alpha/2},\bar{x}+\frac{s}{\sqrt{n}}U_{\alpha/2}$）=（31216，32784）是轮胎平均寿命$\mu$的置信度等于95%的置信区间。

3　大样本下两总体均值差的区间估计（30分钟）

设两总体X与Y的分布是任意的，且均值与方差都存在，分别记$\mu_1=E(X),\sigma_1^2=D(X),\mu_2=E(Y),\sigma_2^2=D(Y)$，它们都是未知的，今独立地分别从两总体中各取一容量为n_1与n_2的样本，算得样本均值分别为$\overline{X}$与$\overline{Y}$，样本方差分别为S_1^2和S_2^2。现欲对两总体的均值差$\mu_1-\mu_2$作区间估计。

解　由中心极限定理知，当n_1，n_2都大时，$\overline{X}$和$\overline{Y}$分别近似于正态分布N（$\mu_1,\frac{\sigma_1^2}{n_1}$)和N（$\mu_2,\frac{\sigma_2^2}{n_2}$）z。再由样本的独立性，容易推出$\overline{X}-\overline{Y}$近似于正态分布N（$\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$)。

其标准化随机变量 $\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ 渐进服从标准正态分布N（0，1）未知参数用无偏估计样本方差 S_1^2 和 S_2^2 代替，则给定置信度1-α，求得 $\mu_1-\mu_2$ 的置信度为1-α 的置信区间（$\bar{X}-\bar{Y}$ -U$_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$，$\bar{X}-\bar{Y}+U_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$）。

教学方式、手段：讲授、示教

教学媒介：教科书、板书、多媒体课件

板书设计：

定义，定理，性质等一些概念性的内容	图形，例题，习题的演算

思考题、作业：

作业：习题7 21、23、25

参考书目：

1.《概率论》，复旦大学编，北京：高等教育出版社，1979。

2.《概率论与数理统计》，浙江大学数学系高等数学教研组编，人民教育出版社，1979。

3.《概率论与数理统计》，王松桂，程维虎，高旅端编，北京：科学出版社，2000

教师姓名：傅金波　　　　职称：讲师　　　　年　　月　　日