第7章 参数估计

•参数的点估计;

•估计量的优良性准则

•区间估计

1

数理统计问题：如何选取样本, 如何根据样本来对总体的种种统计特征作出判断。

参数估计问题：知道随机变量（总体）的分布类型，但确切的形式不知道，根据样本来估计总体的参数，这类问题称为参数估计（paramentric estimation）。

参数估计的类型——点估计、区间估计

2

参数θ的估计量

设总体的分布函数为**F(*x*,θ)**（θ未知），**X_1**，**X_2**，**…**，**X_n** 为样本，构造一个统计量 $\theta=\theta(X_1,X_2,\cdots,X_n)$ 来估计参数θ，则称 $\theta(X_1,X_2,\cdots,X_n)$ 为参数θ的一个估计量。

将样本观测值 $x_1,x_2,\cdots,x_n$ 代入 $\theta(X_1,X_2,\cdots,X_n)$ ，得到的值 $\theta(x_1,x_2,\cdots,x_n)$ 称为参数θ的一个估计值。

3

点估计（**point estimation)**：如果构造一个统计量 $\theta(X_1,X_2,\cdots,X_n)$ 来作为参数θ的估计量，则称为参数θ的点估计。

区间估计（**interval estimation**）：如果构造两个统计量 $\hat{\theta}_1(X_1,X_2,\cdots,X_n),\hat{\theta}_2(X_1,X_2,\cdots,X_n)$ 而用 $(\hat{\theta}_1,\hat{\theta}_2)$ 来作为参数θ可能取值范围的估计，称为参数θ的区间估计。

4

第1节 参数的点估计

•数字特征法

•矩估计法

•最大似然法

数字特征法

样本的数字特征法：以样本的数字特征作为相应总体数字特征的估计量。

以样本均值 $\overline{X}$ 作为总体均值 μ 的点估计量

$$\hat{\mu}=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \qquad \hat{\mu}=\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$ 点估计值

以样本方差 S^2 作为总体方差 σ^2 的点估计量

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

$$\hat{\sigma}^2=s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$$ 点估计值

例1 一批钢件的**20**个样品的屈服点（**t/cm²**）为

4.98 5.11 5.20 5.20 5.11 5.00 5.35

5.61 4.88 5.27 5.38 5.48 5.27 5.23

4.96 5.15 4.77 5.35 5.38 5.54

试估计该批钢件的平均屈服点及其方差。

解 由数字特征法，得屈服点的均值及方差的估计值为

$$\hat{\mu} = \bar{x} = \frac{1}{20}\sum_{i=1}^{20} x_i = 5.21$$

$$\hat{\sigma}^2 = s^2 = \frac{1}{19}\sum_{i=1}^{20}(x_i - 5.21)^2 = 0.049$$

7

矩估计法

定义 设 X 为随机变量，若 $E(|X|^k)$ 存在，则称 $E(X^k)$ 为 X 的 k 阶**原点矩**，记作 $a_k = E(X^k)$；若 $E(|X-EX|^k)$ 存在，则称 $E(X-EX)^k$ 为 X 的 k 阶**中心矩**，记作 $b_k = E(X-EX)^k$

样本的 k 阶原点矩，记作 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^{\,k}$

样本的 k 阶中心矩，记作 $B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$

8

参数的矩估计

矩法估计：用样本的矩作为总体矩的估计量，即

$$a_k = A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \quad \hat{b}_k = B_k = \frac{1}{n}\sum_{n=1}^{n}(X_i - \overline{X})^k$$

若总体X的分布函数中含有m个参数θ$_1$，θ$_2$，…，θ$_m$，总体的k阶矩V$_k$或U$_k$存在，k=1,2,…,m,则

$$a_k(\theta_1,\theta_2,\cdots,\theta_m) = \frac{1}{n}\sum_{i=1}^{n} X_i^k \ (k=1,2,\cdots,m)$$

$$\hat{b}_k(\theta_1,\theta_2,\cdots,\theta_m) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k \ (k=1,2,\cdots,m)$$

9

参数的矩法估计

矩估计：用样本的矩作为总体矩的估计量，即

$$a_k(\theta_1,\theta_2,\cdots,\theta_m) = \frac{1}{n}\sum_{i=1}^{n} X_i^k \ (k=1,2,\cdots,m)$$

$$\hat{b}_k(\theta_1,\theta_2,\cdots,\theta_m) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k \ (k=1,2,\cdots,m)$$

得m个方程构成方程组，解得的 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m$ 即为参数 $\theta_1,\theta_2,\cdots,\theta_m$ 的矩估计量，代入样本观测值，即得参数的矩估计值。

10

例2 设某总体X的数学期望为EX=μ，方差DX=σ²，X_1，X_2，…，X_n为样本，试求μ和σ²的矩估计量。

解 总体的一阶二阶阶原点矩为 $a_1 = \mu$

$$a_2 = E(X^2) = DX + \left(EX\right)^2 = \sigma^2 + \mu^2$$

样本的一阶二阶阶原点矩为 $A_1 = \overline{X}$ $\quad A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

由矩法估计，应有 $\overline{X} = \mu$ $\quad \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \sigma^2 + \mu^2$

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

11

结论：不管总体X服从何种分布，总体期望和方差的矩估计量分别为样本均值、样本的二阶中心矩，即

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

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

估计值为 $\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ $\quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$

12

例3 设X_1，X_2，…，X_n为总体X的样本，试求下列总体分布参数的矩估计量。

(1) $X \sim N\left(\mu,\sigma^2\right)$ (2)$X \sim B\left(N,p\right)$(N已知）（3） $X \sim P(\lambda)$

解 （1）由于 $EX=\mu \quad DX=\sigma^2$

所以参数μ和σ^2的矩估计量为

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

（2）由于 $EX=Np$

所以 $Np=\overline{X}$

得参数p的矩估计量为 $\hat{p}=\frac{1}{N}\overline{X}$

例3 设X_1，X_2，…，X_n为总体X的样本，试求下列总体分布参数的矩估计量。

(1) $X \sim N\left(\mu,\sigma^2\right)$ (2)$X \sim B\left(N,p\right)$(N已知）（3） $X \sim P(\lambda)$

解 （3）由于 $EX=DX=\lambda$

所以参数λ的矩估计量为

$$\lambda=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \qquad \text{或} \qquad \lambda=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

一阶矩　　　　　　　二阶中心矩

可见：同一个参数的矩估计量可以不同，所以估计量存在“优、劣”之分。

例4 设总体X服从[θ_1， θ_2]上的均匀分布， $\theta_1<\theta_2$，求θ_1， θ_2的矩估计量， X_1，X_2，…，X_n为X的一个样本。

解 由于 $EX=\dfrac{\theta_1+\theta_2}{2}, DX=\dfrac{(\theta_2-\theta_1)^2}{12}$

所以由矩法估计，得

$$\bar{X}=\frac{\theta_1+\theta_2}{2} \qquad S_n^2=\frac{(\theta_2-\theta_1)^2}{12}$$

解得 $\hat{\theta}_1=\bar{X}-\sqrt{3}S_n \quad \hat{\theta}_2=\bar{X}+\sqrt{3}S_n$

区间长度的矩估计量为 $\hat{\theta}_2-\hat{\theta}_1=2\sqrt{3}S_n$

15

例5 对容量为n的子样，求下列密度函数中参数 a的矩估计量。

$$f(x)=\begin{cases}\dfrac{2}{a^2}(a-x), & (0<x<a)\\ 0, & \text{其它}\end{cases}$$

解 由于 $EX=\int_0^a x\cdot\dfrac{2}{a^2}(a-x)dx=\dfrac{a}{3}$

所以由矩法估计，得 $\bar{X}=\dfrac{a}{3}$

解得 $a=3\bar{X}=\dfrac{3}{n}\sum_{i=1}^{n}X_i$

所以，参数a的矩估计量为 $a=\dfrac{3}{n}\sum_{i=1}^{n}X_i$

16

最大似然法

思想：设总体**X**的密度函数为**f(x,θ)**，**θ**为未知参数，则样本（$X_1,X_2,\ldots,X_n$）的联合密度函数为

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

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

参数θ的估计量 $\hat{\theta}$，使得样本（$X_1,X_2,\ldots,X_n$）落在观测值 $(x_1,x_2,\cdots,x_n)$ 的邻域内的概率**L(θ)**达到最大，即

$$L(x_1,x_2,\cdots,x_n,\hat{\theta})=\max L(x_1,x_2,\cdots,x_n,\theta)$$

则称 $\hat{\theta}$ 为参数θ的最大似然估计值。

17

参数的最大似然估计

求解方法：

（1）构造似然函数 $L(\theta)=f(x_1,x_2,\cdots,x_n,\theta)=\prod_{i=1}^{n}f(x_i,\theta)$

（2）取自然对数 $\ln L(x_1,x_2,\cdots,x_n,\theta)=\sum_{i=1}^{n}\ln f(x_i,\theta)$

（3）令 $\dfrac{d\ln L}{d\theta}=0$

其解 $\hat{\theta}$ 即为参数θ的最大似然估计。

若总体的密度函数中有多个参数θ_1，θ_2，…，θ_n，则将第（3）步改为 $\dfrac{\partial \ln L}{\partial \theta_i}=0,(i=1,2,\cdots,n)$ 解方程组即可。

对于离散总体，将密度函数换成分布律即可。

18

例6 假设（X_1，X_2，…，X_n）是取自正态总体$N(\mu,\sigma^2)$的样本，求μ和σ^2的最大似然估计量。

解 构造似然函数

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

取对数 $\ln L=\sum_{i=1}^{n}\ln\left(\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}\right)$

$$=\sum_{i=1}^{n}\left(-\frac{(x_i-\mu)^2}{2\sigma^2}-\ln\sqrt{2\pi}-\ln\sigma\right)$$

19

求偏导数，并令其为0

$$\frac{\partial\ln L}{\partial\mu}=\sum_{i=1}^{n}\left(-\frac{2(x_i-\mu)(-1)}{2\sigma^2}\right)=\frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2}=0$$

$$\frac{\partial\ln L}{\partial\sigma^2}=\sum_{i=1}^{n}\left(\frac{(x_i-\mu)^2}{2\left(\sigma^2\right)^2}-\frac{1}{2}\cdot\frac{1}{\sigma^2}\right)=0$$

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

所以μ，σ^2的最大似然估计量为

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

与矩估计量相同

20

第2节 估计量的优良性准则

——无偏性、有效性

无偏估计量：设 $\hat{\theta}$ 是 θ 的估计量，如果 $E(\hat{\theta})=\theta$，则称 $\hat{\theta}$ 是 θ 的无偏估计量（**unbiased estimation**）

无偏性的意义在于，用无偏估计去估计未知参数时，虽然存在随机性，但是其平均值等于未知参数.

21

例题 设总体的数学期望**EX**和方差**DX**都存在，证明：样本均值 $\bar{X}$ 、样本方差 $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 分别是**EX**、**DX**的无偏估计。

证明 $E(S^2)=\dfrac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]$

$$=\frac{n}{n-1}E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right]=\frac{n}{n-1}E\left[\frac{1}{n}\sum_{i=1}^{n}X_i^2-\left(\bar{X}\right)^2\right]$$

$$=\frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)-E\left(\bar{X}\right)^2\right]$$

$$=\frac{n}{n-1}\left[\frac{n[DX+(EX)^2]}{n}-D\left(\bar{X}\right)-(E\bar{X})^2\right]$$

22

$$E(S^2)=\frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]$$

$$=\frac{n}{n-1}\left[\frac{n[DX+(EX)^2]}{n}-D\left(\bar{X}\right)-(E\bar{X})^2\right]$$

$$=\frac{n}{n-1}\left[\frac{n[DX+(EX)^2]}{n}-\frac{DX}{n}-(EX)^2\right]\quad = DX$$

$$E\left\{\frac{1}{n}\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]\right\}=\frac{n-1}{n}DX$$

有 效 性

设θ_1和θ_2是θ的两个无偏估计量，如果

$$D(\hat{\theta}_1)\le D(\hat{\theta}_2)$$

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

若在θ 的所有无偏估计类中，θ 的方差达到最小，则称 θ 是 θ 的一致最小方差无偏估计.

例如 $\overline{X}$ 及 $\sum_{i=1}^{n} a_i X_i$（其中$\sum_{i=1}^{n} a_i = 1$）都是EX的无偏估计，但 $\overline{X}$ 比 $\sum_{i=1}^{n} a_i X_i$ 有效。

因为 $E\left(\overline{X} - EX\right)^2 = E\left(\overline{X} - E\overline{X}\right)^2 = D\overline{X} = \dfrac{DX}{n}$

$$E\left[\sum_{i=1}^{n} a_i X_i - EX\right]^2 = E\left[\sum_{i=1}^{n} a_i X_i - E\left(\sum_{i=1}^{n} a_i X_i\right)\right]^2$$

$$= D\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 D(X_i) = D(X)\sum_{i=1}^{n} a_i^2$$

$$= \frac{1}{n} D(X) \cdot n\sum_{i=1}^{n} a_i^2 \geq \frac{1}{n} D(X) \cdot \left(\sum_{i=1}^{n} a_i\right)^2 = \frac{1}{n} D(X)$$

25

小结 参数估计的点估计方法

数字特征法：以样本均值、方差作为总体期望、方差的估计量。

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

矩估计法：以样本k阶矩作为总体k阶矩的估计量。

$$\hat{V}_k(\theta_1, \theta_2, \cdots, \theta_m) = \frac{1}{n}\sum_{i=1}^{n} X_i^k \ (k = 1, 2, \cdots, m)$$

$$\hat{U}_k(\theta_1, \theta_2, \cdots, \theta_m) = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^k \ (k = 1, 2, \cdots, m)$$

最大似然估计

26

课后练习

习题7

1--7

作业

3，6

27

第3节 区间估计

本节主要研究正态总体的均值和方差的区间估计.

- 方差已知,对均值的区间估计
- 方差未知,对均值的区间估计
- 均值已知,对方差的区间估计
- 均值未知,对方差的区间估计

28

区间估计的思想

点估计总是有误差的，但没有衡量偏差程度的量，区间估计则是按一定的可靠性程度对待估参数给出一个区间范围。

引例　设某厂生产的灯泡使用寿命X~N（μ，100^2），现随机抽取5只，测量其寿命如下：**1455**，**1502**，**1370**，**1610**，**1430**，则该厂灯泡的平均使用寿命的点估计值为

$$\bar{x}=\frac{1}{5}(1455+1502+1370+1610+1430)=1473.4$$

29

可以认为该种灯泡的使用寿命在**1473.4**个单位时间左右，但范围有多大呢？又有多大的可能性在这“左右”呢？

如果要求有**95%**的把握断言μ在**1473.4**左右，则由**U**统计量可知　$U=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

由　$P\left\{\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right|<\varepsilon\right\}=0.95 \longrightarrow \Phi(\varepsilon)=0.975$

查表得　$\varepsilon=1.96$

$$\longrightarrow \bar{X}-1.96\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+1.96\frac{\sigma}{\sqrt{n}}$$

30

置信水平、置信区间

设总体X的分布函数 $F(x;\theta)$ 中含有一个未知参数θ，$X_1, X_2, \cdots, X_n$ 为来自X的样本。若能构造两个统计量 $\theta_1(X_1, X_2, \cdots, X_n)$ 和 $\theta_2(X_1, X_2, \cdots, X_n)$ 使得对给定的足够小的α值（0< α<1,通常取α＝0.05，0.01，0.1 ），满足 $P\{\theta_1 \le \theta \le \theta_2\} = 1-\alpha$

则称区间[θ_1，θ_2]为参数θ的置信度（或置信水平）为 1- α的置信区间， α 为显著性水平。

θ_1——置信下限 θ_2——置信上限

31

几点说明

1、参数θ的置信水平为1-α的置信区间（ θ_1， θ_2） 表示该区间有100（1-α）%的可能性包含总体参数θ的真值。

2、不同的置信水平，参数θ的置信区间不同。

3、置信区间越小，估计越精确，但置信水平会降低；相反，置信水平越大，估计越可靠，但精确度会降低，置信区间会较长。一般：对于固定的样本容量，不能同时做到精确度高（置信区间小），可靠程度也高（1- α大）。如果不降低可靠性，而要缩小估计范围，则必须增大样本容量n，增加抽样成本。

32

一. 正态总体均值的区间估计

1. 正态总体方差已知，对均值的区间估计

如果总体X~N（μ，σ²），其中σ²已知， μ未知，

则构造 $U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ ，对μ做区间估计。

对给定的置信水平1-α，由 $P\left\{|U|\le u_{\frac{\alpha}{2}}\right\}=1-\alpha$

确定临界值（X的双侧α分位数）

33

$$|U|\le u_{\frac{\alpha}{2}} \Rightarrow \left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right|\le u_{\frac{\alpha}{2}}$$

$$\Rightarrow -u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\le \overline{X}-\mu\le u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \overline{X}-u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\le \mu\le \overline{X}+u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

得μ的置信区间

$$\left[\overline{X}-u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{X}+u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

将观测值 $(x_1,x_2,\cdots,x_n)$ 代入，则可得具体的区间。

34

例1 某车间生产滚珠，从长期实践中知道，滚珠直径X可以认为服从正态分布，从某天的产品中随机抽取6个，测得直径为（单位：cm）

14.6，15.1，14.9，14.8，15.2，15.1

（1）试求该天产品的平均直径EX的点估计；

（2）若已知方差为0.06，试求该天平均直径EX的置信区间：α=0.05；α=0.01。

解 （1）由矩法估计得EX的点估计值为

$$EX=\bar{x}=\frac{1}{6}(14.6+15.1+14.9+14.8+15.2+15.1)=14.95$$

35

（2）由题设知X~N（μ，0.06）

构造 U，得EX的置信区间为

$$\left[\bar{X}-u_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{X}+u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

而 $\bar{x}=14.95,\quad \frac{\sigma}{\sqrt{n}}=\frac{\sqrt{0.06}}{\sqrt{6}}=0.1$

当α=0.05时，$u_{0.025}=1.96$

所以，EX的置信区间为[14.754，15.146]

当α=0.01时，$u_{0.005}=2.575$

所以，EX的置信区间为[14.693，15.208]

置信水平提高，置信区间扩大，估计精确度降低。36

例2 假定某地一旅游者的消费额X服从正态分布N（μ，σ²），且标准差σ=12元，今要对该地旅游者的平均消费额EX加以估计，为了能以95%的置信度相信这种估计误差小于2元，问至少要调查多少人？

解 由题意知：消费额X~N（μ，12²），设要调查n人。

由 $1-\alpha=0.95$ 得 $\alpha=0.05$ 查表得 $u_{\alpha/2}=1.96$

即 $P\left\{\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right|\le 1.96\right\}=0.95$

而 $|\bar{X}-\mu|<2$ $\longrightarrow$ $1.96\times\frac{\sigma}{\sqrt{n}}=2$

解得 $n=\left(\frac{1.96\times 12}{2}\right)^2\approx 138.3$ 至少要调查139人

37

2. 正态总体方差未知，对均值的区间估计

如果总体X~N（μ，σ²），其中μ，σ均未知

由 $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$ 构造随机变量T，$T=\frac{\bar{X}-\mu}{S/\sqrt{n}}$

当置信水平为1-α时，由 $P\{|T|\le t_{\alpha/2}(n-1)\}=1-\alpha$

查t-分布表确定 $t_{\alpha/2}(n-1)$

从而得μ的置信水平为1-α的置信区间为

$$\left[\bar{X}-\frac{S}{\sqrt{n}}\cdot t_{\alpha/2}(n-1),\quad \bar{X}+\frac{S}{\sqrt{n}}\cdot t_{\alpha/2}(n-1)\right]$$

38

例3 某厂生产的一种塑料口杯的重量X被认为服从正态分布，今随机抽取9个，测得其重量为（单位：克）：21.1，21.3，21.4，21.5，21.3，21.7，21.4，21.3，21.6。试用95%的置信度估计全部口杯的平均重量。

解 由题设可知：口杯的重量X~N（μ，σ²）

由抽取的9个样本，可得 $s=0.18$ $\bar{x}=21.4$ $n=9$

由 $1-\alpha=0.95$ 得 $\alpha=0.05$ 查表得 $t_{0.025}(8)=2.306$

$$t_{\alpha/2}(8)\times\frac{S}{\sqrt{n}}=2.306\cdot\frac{0.18}{\sqrt{9}}=0.1384$$

全部口杯的平均重量的置信区间为[21.26，21.54]

39

例4 旅行社随机访问了25名旅游者,得知平均消费额 $\bar{x}=80$ 样本标准差 $s=12$.已知旅游者消费额服从正态分布,那么该地旅游者平均消费额 μ 的95%置信区间是什么?

解 由题设可知：平均消费额X~N（μ，σ²）

$$S=12 \quad \bar{x}=80 \quad n=25$$

40

由 $1-\alpha=0.95$ 得 $\alpha=0.05$ 查表得 $t_{0.025}(24)=2.064$

$$t_{\alpha/2}(24)\times\frac{S}{\sqrt{n}}=2.064\times\frac{12}{\sqrt{25}}=4.9536$$

平均消费额的置信区间为[75.0464，84.9536]

估计误差为 $2\times4.9536=9.9072>2\times2$

与例**2**结果相比

精确度降低——原因：**1.**样本容量减少 **,2.**用S^2代替σ^2

实际应用中，方差未知的均值的区间估计较有应用价值。

41

例**5** 假设某片居民每月对某种商品的需求量**X**服从正态分布，经调查**100**家住户，得出每户每月平均需求量为**10**公斤，方差为**9**，如果某商店供应**10000**户，试就居民对该种商品的平均需求量进行区间估计（α=**0.01**），并依此考虑最少要准备多少这种商品才能以**99%**的概率满足需求？

解 由题设可知：平均需求量**X~N**（μ，σ^2）

$$S^2=9 \quad \bar{x}=10 \quad n=100$$

由 $\alpha=0.01$ 查表得 $t_{0.005}(99)\approx u_{0.005}=2.575$

$$t_{\alpha/2}(99)\times\frac{S}{\sqrt{n}}=2.575\times\frac{3}{\sqrt{100}}=0.773$$

平均需求量的**99%**的置信区间为[9.227，10.773]

42

要以**99%**的概率满足**10000**户居民对该种商品的需求，则最少要准备的量为

$$10.773 \times 10000 = 107730 \quad （公斤）$$

二. 正态总体方差的区间估计

1. 正态总体均值已知，对方差的区间估计

如果总体**X~N**（μ，σ²），其中μ已知，σ²未知

由 $\dfrac{X_i-\mu}{\sigma} \sim N(0,1)$ 构造随机变量 χ^2

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

对给定的置信水平1-α，由

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

查χ^2- 分布表，确定双侧分位数

$$\chi^2_{1-\alpha/2}(n), \chi^2_{\alpha/2}(n)$$

从而得σ^2的置信水平为1-α的置信区间为

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

45

例6 已知某种果树产量服从N（218，σ^2），随机抽取6棵计算其产量为（单位：公斤）

221，191，202，205，256，236

试以95%的置信水平估计产量的方差。

解 计算 $\sum_{i=1}^{6}(x_i-\mu)^2 = 2931$

查表 $\chi^2_{1-0.05/2}(6) = 1.237, \chi^2_{0.05/2}(6) = 14.449$

果树方差的置信区间为

$$\left[\frac{2931}{14.449}, \frac{2931}{1.237}\right] = [202.851, 2369.442]$$

46

2. 正态总体均值未知，对方差的区间估计

如果总体X~N（μ，σ²），其中σ²未知

由 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ 构造随机变量 χ^2

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

当置信水平为1-α时，

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

47

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

查χ²- 分布表，确定双侧分位数

$$\chi^2_{1-\alpha/2}(n-1)\ ,\ \chi^2_{\alpha/2}(n-1)$$

从而得σ²的置信水平为1-α的置信区间为

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

48

例7 设某灯泡的寿命X~N（μ，σ²），μ，σ²未知，现从中任取5个灯泡进行寿命试验，得数据10.5，11.0，11.2，12.5，12.8（单位：千小时），求置信水平为90%的σ²的区间估计。

解 样本方差及均值分别为 $S^2 = 0.995 \qquad \bar{x} = 11.6$

由 $1-\alpha = 0.9$ 得 $\alpha = 0.1$ 查表得

$$\chi^2_{1-0.05}(4) = 0.711 \quad \chi^2_{0.05}(4) = 9.488$$

$$\frac{(n-1)S^2}{\chi^2_{1-0.05}} = \frac{4\times 0.995}{0.711} \approx 5.598 \qquad \frac{(n-1)S^2}{\chi^2_{0.05}} \approx 0.419$$

σ²的置信区间为[0.419，5.598]

49

小　结

总体服从正态分布的均值或方差的区间估计

假设置信水平为1-α

（1）方差已知，对均值的区间估计

构造随机变量U，反查标准正态分布表，

确定U的双侧分位数 $u_{\alpha/2}$

得EX的区间估计为

$$\left[\bar{X} - u_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \quad \bar{X} + u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

50

总体服从正态分布的均值或方差的区间估计

假设置信水平为1-α

（2）方差未知，对均值的区间估计

构造随机变量T，查t-分布临界值表，

确定T的双侧分位数 $t_{\alpha/2}(n-1)$

得EX的区间估计为

$$\left[\bar{X}-\frac{S}{\sqrt{n}}\cdot t_{\alpha/2}(n-1),\quad \bar{X}+\frac{S}{\sqrt{n}}\cdot t_{\alpha/2}(n-1)\right]$$

51

总体服从正态分布的均值或方差的区间估计

假设置信水平为1-α

（3）均值已知，对方差的区间估计

构造随机变量χ^2，查χ^2-分布临界值表，

确定χ^2的双侧分位数 $\chi^2_{1-\alpha/2}(n), \chi^2_{\alpha/2}(n)$

得σ^2的区间估计为

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

52

总体服从正态分布的均值或方差的区间估计

假设置信水平为1-α

（4）均值未知，对方差的区间估计

构造随机变量 χ^2，查χ^2-分布临界值表，

确定χ^2的双侧分位数 $\chi^2_{1-\alpha/2}(n-1), \chi^2_{\alpha/2}(n-1)$

得σ^2的区间估计为

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

53

课后练习

习题7

8--13

作业

9, 11

54

本章作业

3, 6, 9, 11

下节课交作业，请同学们做好作业，下次课不要忘了带过来