点估计
区间估计

§7.3 正态母体参数的置信区间

前面，我们讨论了参数的点估计. 它是用样本算得的一个值去估计未知参数. 但是，点估计值仅仅是未知参数的一个近似值，它没有反映出这个近似值的误差范围，使用起来把握不大. 区间估计正好弥补了点估计的这个缺陷.

一、 区间估计（Interval estimation）的概念

[,]

区间估计 给出总体未知参数所在的一个随机范围，使得这个范围包含待估参数的概率为一个固定值.

$P(\mu \in (\quad , \quad)) = 95\%$

引例 设 $\xi \sim N(\mu, \sigma_0^2)$,其中 σ_0^2 已知; $\xi_1, \xi_2, \cdots, \xi_n$ 是 ξ 的一个样本,求一区间,使之以95%的把握断定这个区间包含 μ 的真值.

分析

$$\bar{\xi} \sim N\left(\mu, \frac{\sigma_0^2}{n}\right) \qquad U = \frac{\bar{\xi} - \mu}{\sigma_0 / \sqrt{n}} \sim N(0,1)$$

$$P\{|U| < \lambda\} = 0.95$$

$$2\Phi(\lambda) - 1 = 0.95$$

$$\lambda = ? \quad \lambda = 1.96$$

$$0.95 = P\left\{\left|\frac{\bar{\xi} - \mu}{\sigma_0 / \sqrt{n}}\right| < 1.96\right\} = P\left\{-1.96 < \frac{\bar{\xi} - \mu}{\sigma_0 / \sqrt{n}} < 1.96\right\}$$

$$= P\left\{\bar{\xi} - 1.96\frac{\sigma_0}{\sqrt{n}} < \mu < \bar{\xi} + 1.96\frac{\sigma_0}{\sqrt{n}}\right\} = 0.95.$$

区间 $\left(\bar{\xi} - 1.96\frac{\sigma_0}{\sqrt{n}}, \bar{\xi} + 1.96\frac{\sigma_0}{\sqrt{n}}\right)$ 以95%的概率包含真值 μ

定义7.1

设母体具有概率函数$f(x;\theta)$,θ为未知参数,$\xi_1,\xi_2,\cdots,\xi_n$为取自母体ξ的一个样本,若对于事先给定的α,$0<\alpha<1$,存在两个统计量

$$\boxed{\underline{\theta}(\xi_1,\xi_2,\cdots,\xi_n)\text{和}\overline{\theta}(\xi_1,\xi_2,\cdots,\xi_n),\text{满足}P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha}$$

则称随机区间$(\underline{\theta},\overline{\theta})$是$\theta$的$1-\alpha$置信区间(confidence interval),$\underline{\theta}$和$\overline{\theta}$分别称为置信下限和置信上限,$1-\alpha$称为置信度(置信水平).

说明:

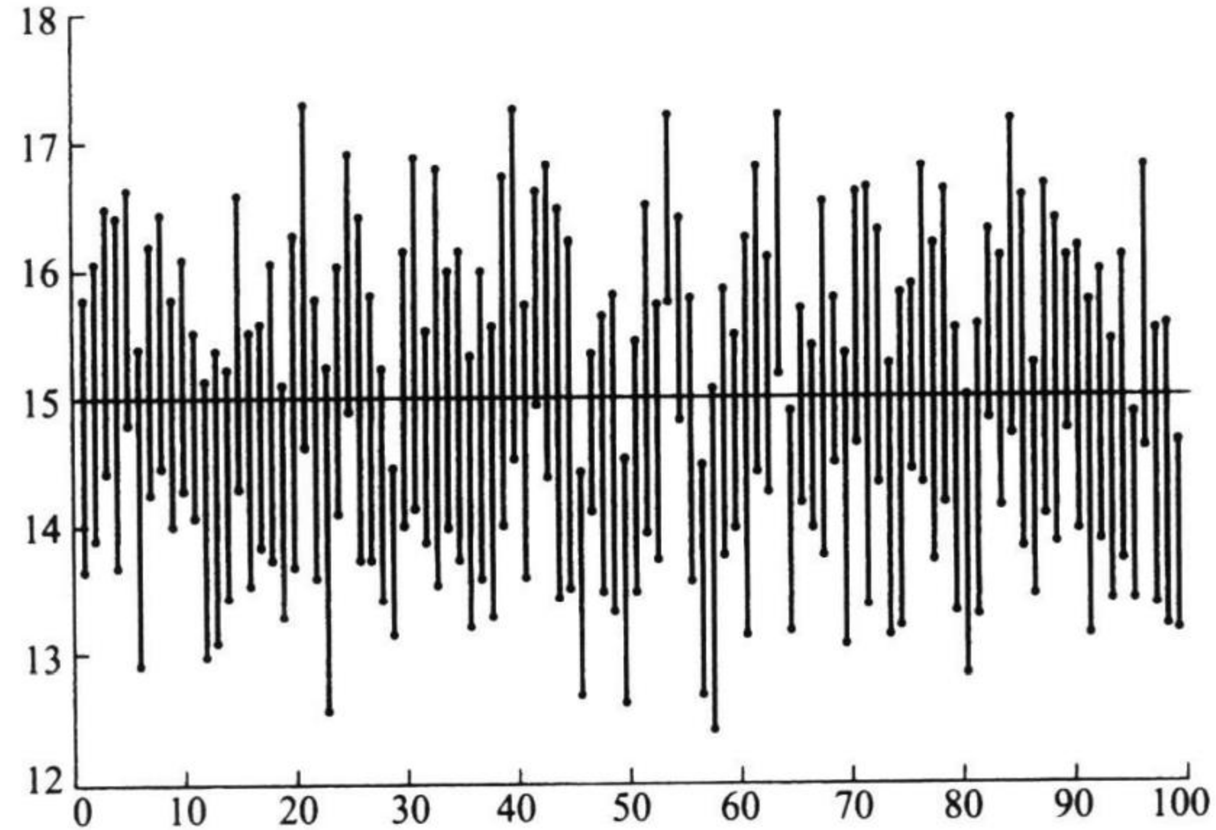
1. 置信区间$(\underline{\theta}(\xi_1,\xi_2,\cdots,\xi_n),\overline{\theta}(\xi_1,\xi_2,\cdots,\xi_n))$是一个随机区间, 且两端点不依赖于未知参数$\theta$.

2. $1-\alpha$ 表示随机区间$(\underline{\theta},\overline{\theta})$包含 θ 真值的可信程度, 一般来说是比较大的, 常取 0.9, 0.95 等.

3. 置信水平为$1-\alpha$的置信区间不是唯一的

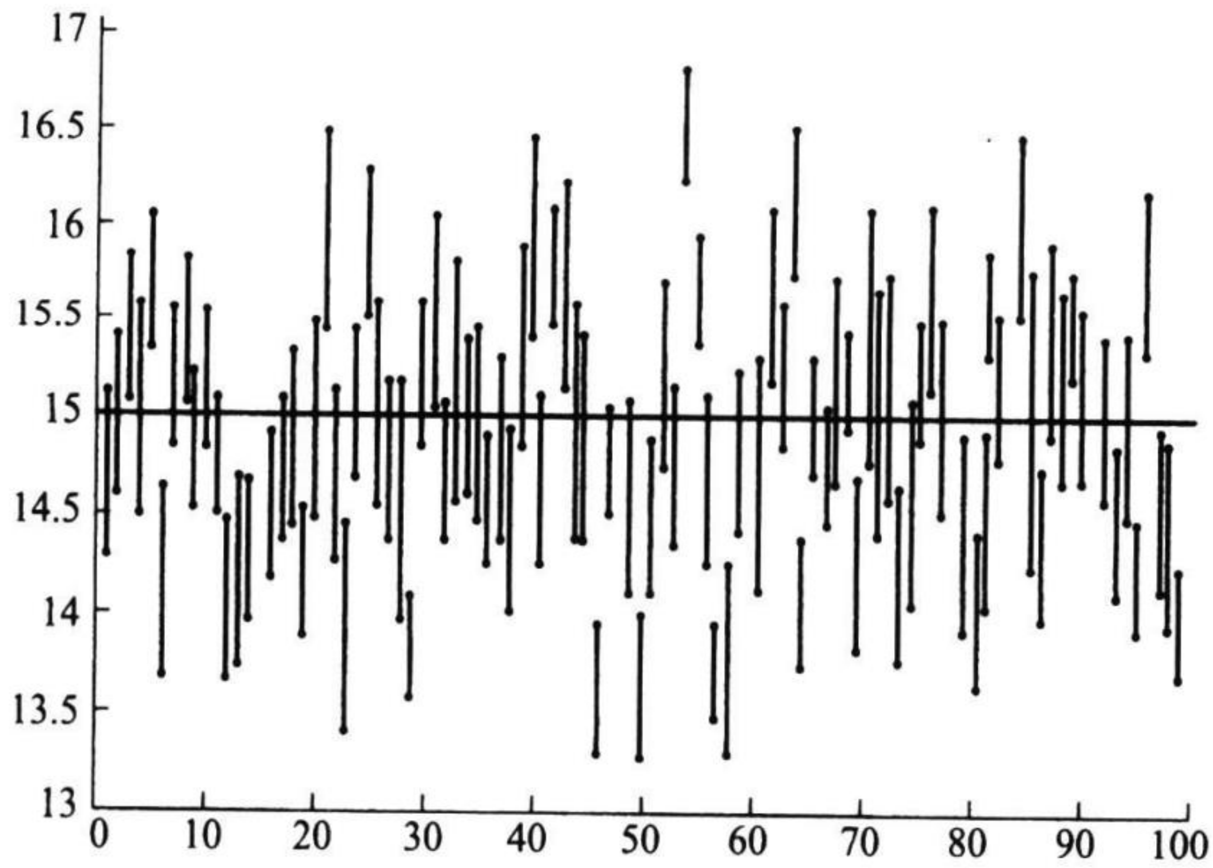
$$P\{|U|<\lambda\}=0.95 \longrightarrow P\{\lambda_1<U<\lambda_2\}=0.95$$

$$(-\lambda,\lambda) \qquad\qquad (\lambda_1,\lambda_2)$$



μ 的置信水平为 0.90 的置信区间

100个区间中有91个包含参数真值15, 另外9个不包含参数真值.



μ 的置信水平为 0.50 的置信区间

100个区间中有50个包含参数真值15, 另外50个不包含参数真值.

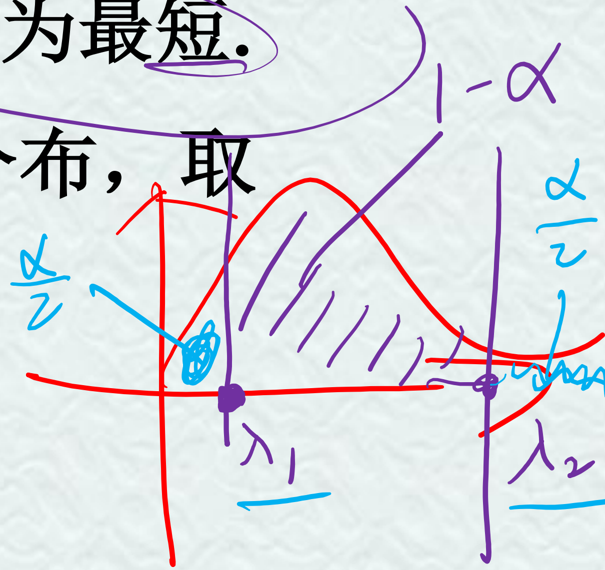
$\chi^2 \sim \chi^2(n)$　　$P(|U|<\lambda)=1-\alpha$

4. 密度函数为对称时 (如 $N(0,1)$, $t(n)$ 分布)，在置信度一定时，取对称的 λ 的置信区间为最短.

密度函数不对称时，例如 $\chi^2(n)$ 分布，取

$$P\{\lambda_1<\chi^2(\cancel{n})<\lambda_2\}=1-\alpha$$

$$\boxed{P\{\chi^2(\cancel{n})\le\lambda_1\}=P\{\chi^2(\cancel{n})\ge\lambda_2\}=\frac{\alpha}{2}}$$



等尾

$1-\alpha=0.95$ $U=\frac{\bar{\xi}-\mu}{\sigma/\sqrt{n}}$ $1-\alpha=0.95$

$u_{0.975}=1.96$

一般步骤:

(1) 首先, 寻求样本的一个函数 $u(\xi_1,\cdots,\xi_n;\theta)$, 它只含待估参数 θ 而不含其他未知参数, 而且该样本函数的分布是已知的 (可以直接查附表).

$P(U>u_{0.975})=0.025$ $1-\frac{\alpha}{2}$

一般的, 称 $u(\xi_1,\xi_2,\cdots,\xi_n;\theta)$ 为枢轴量 (*pivotal*). $\frac{\alpha}{2}$

→ 从无偏的点估计量出发

(2)对于给定的置信度 $1-\alpha$, 通过 u 求出数集 A, 使得 $P\{u(\xi_1,\cdots,\xi_n;\theta)\in A\}=1-\alpha$,即确定分位数.

$\lambda=1.96$

(3)利用不等式的变形, 从$u\in A$中得到$\underline{\theta}<\theta<\overline{\theta}$.

二、 正态母体均值的区间估计

设母体ξ服从$N(\mu,\sigma^2)$，$\xi_1,\xi_2,\cdots\xi_n$是来自总体的一个样本.

1. σ^2已知，求μ的置信区间

$\lambda = u_{1-\frac{\alpha}{2}}$

由引例知：$\bar{\xi} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ $\quad U = \dfrac{\bar{\xi}-\mu}{\sigma/\sqrt{n}} \sim N(0,1).$

$P(|U|<\lambda)$

对于给定的$\alpha(0<\alpha<1)$，$1-\alpha$为置信度

查正态分布表 (附表3) 得 λ：$P\left\{\left|\dfrac{\bar{\xi}-\mu}{\sigma/\sqrt{n}}\right|<\lambda\right\}=1-\alpha$

$\frac{\alpha}{2}$

$$P\left\{\bar{\xi}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{\xi}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right\}=1-\alpha$$

母体均值μ的$1-\alpha$置信区间为：$\left(\bar{\xi}-u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \bar{\xi}+u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right)$

$1-\alpha=0.9$

例1: 设母体 $\xi \sim N(\mu, 0.2^2), \xi_1, \xi_2, \cdots \xi_9$ 是来自母体 ξ 的容量为9的样本,计算 $\bar{x}=12.35$,试求母体期望 μ 的90%的置信区间.

解:已知 $\bar{x}=12.35, \sigma=0.2, \sqrt{9}=3, 1-\alpha=0.9$

$$u_{1-0.05}=u_{0.95}=1.64$$

μ 的 $1-\alpha$ …… 的

$$\left(\bar{\xi}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{\xi}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

μ 的置信度为90%的置信区间为:

$$\left(12.35-1.64\frac{0.2}{3}, 12.35+1.64\frac{0.2}{3}\right)$$

$u_{0.95}=1.645$

1.64

$u_{0.975}=1.96$

即 (12.24,12.46)

2.σ^2未知,求μ的置信区间

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2$$

$$U = \frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}} \sim N(0,1).$$

$$T = \frac{\bar{\xi} - \mu}{S_n^* / \sqrt{n}} \sim ?\quad t(n-1)$$

$$S_n^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2$$

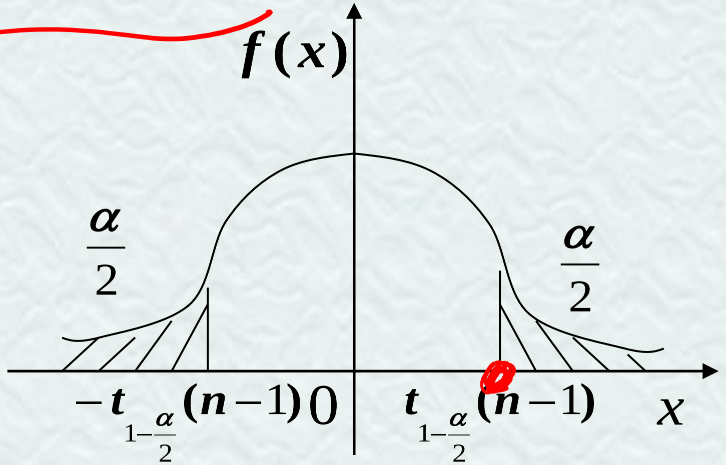
$$nS_n^2 = (n-1)S_n^{*2} = \sum_{i=1}^{n}(\xi_i - \bar{\xi})^2$$

$$\frac{\bar{\xi}-\mu}{\sigma/\sqrt{n}} \longrightarrow T=\frac{\bar{\xi}-\mu}{\frac{S_n^*}{\sqrt{n}}}\sim t(n-1)$$

$\sigma \to S_n^*$

$$P\left\{\left|\frac{\bar{\xi}-\mu}{\frac{S_n^*}{\sqrt{n}}}\right|<\lambda\right\}=1-\alpha$$

$t_{1-\frac{\alpha}{2}}(n-1)$



$$P\left\{\bar{\xi}-t_{1-\alpha/2}(n-1)\frac{S_n^*}{\sqrt{n}}<\mu<\bar{\xi}+t_{1-\alpha/2}(n-1)\frac{S_n^*}{\sqrt{n}}\right\}=1-\alpha$$

$t_{1-\frac{\alpha}{2}}(n-1)$

σ^2 未知时, μ 的 $1-\alpha$ 置信区间为:

$$\left(\bar{\xi}-t_{1-\alpha/2}(n-1)\cdot\frac{S_n^*}{\sqrt{n}},\ \bar{\xi}+t_{1-\alpha/2}(n-1)\cdot\frac{S_n^*}{\sqrt{n}}\right)$$

$$\left(\overline{\xi}-t_{1-\alpha/2}(n-1)\cdot\frac{S_n^*}{\sqrt{n}},\overline{\xi}+t_{1-\alpha/2}(n-1)\cdot\frac{S_n^*}{\sqrt{n}}\right)$$

例2： 假设人的身高服从正态分布. 今从高三毕业班中随机抽查十名女生, 测其身高如下：162, 159.5, 168, 160, 157, 162, 163.4, 158.5, 170.3, 166 (单位：厘米), 求高三女生平均身高的0.95的置信区间.

$\alpha=0.05$

解：用ξ表示女生身高，记$\mu=E\xi$. $\overline{x}=162.67$, $\alpha=0.05$,

$$S_{10}^{*2}=\frac{1}{10-1}\sum_{i=1}^{10}(x_i-\overline{x})^2=\frac{1}{9}\left(\sum_{i=1}^{10}x_i^2-10\overline{x}^2\right)=18.43$$

查表得：$\lambda=t_{1-\frac{\alpha}{2}}(n-1)=t_{0.975}(9)=2.2622$

μ的置信度为0.95的置信区间(159.6,165.74)

思考：

1、 如果从样本方差出发，可得怎样的置信区间？

2、为何在方差已知的情况下不用后一种方法？

三、 正态总体方差的区间估计

1. $\mu = \mu_0$ 已知，求 σ^2 的置信区间

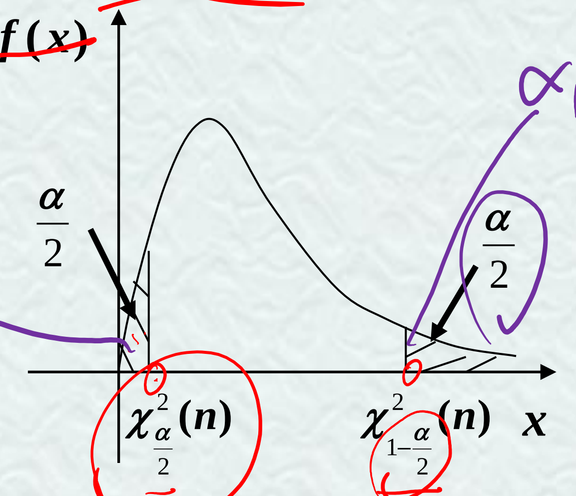
χ^2

$$\because \xi_i \sim N(\mu_0, \sigma^2) \therefore \frac{\xi_i - \mu_0}{\sigma} \sim N(0,1) \quad \sum_{i=1}^{n}\left(\frac{\xi_i - \mu_0}{\sigma}\right)^2 \sim \chi^2(n) \ (\mu_0\text{已知})$$

$$P\left\{ \lambda_1 < \frac{\sum_{i=1}^{n}(\xi_i - \mu_0)^2}{\sigma^2} < \lambda_2 \right\} = 1-\alpha$$

$$\Rightarrow P\left\{ \chi^2_{\frac{\alpha}{2}}(n) < \frac{\sum_{i=1}^{n}(\xi_i - \mu_0)^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}(n) \right\} = 1-\alpha$$

$$\Rightarrow P\left\{ \frac{\sum_{i=1}^{n}(\xi_i - \mu_0)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)} < \sigma^2 < \frac{\sum_{i=1}^{n}(\xi_i - \mu_0)^2}{\chi^2_{\frac{\alpha}{2}}(n)} \right\} = 1-\alpha$$



α_1

$\alpha - \alpha_1$

$\chi^2_{\alpha-\alpha_1}(n)$

$\chi^2_{1-\alpha_1}(n)$

$$\Rightarrow P\left\{\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}<\sigma^2<\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right\}=1-\alpha$$

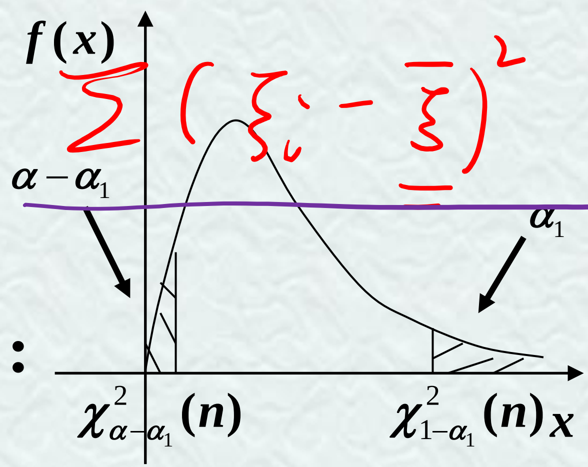
未知 μ^2

$\sum(\xi_i-\cancel{\mu_0})$ ×

$\mu=\mu_0$已知,σ^2的$1-\alpha$置信区间为:

$$\left(\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)},\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right)$$

$\sum(\xi_i-\bar{\xi})^2$



若不是等区间($\frac{\alpha}{2}$),设为α_1,则置信区间为:

$$\left(\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{1-\alpha_1}(n)},\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{\alpha-\alpha_1}(n)}\right)$$

2. μ 未知，求 σ^2 的置信区间

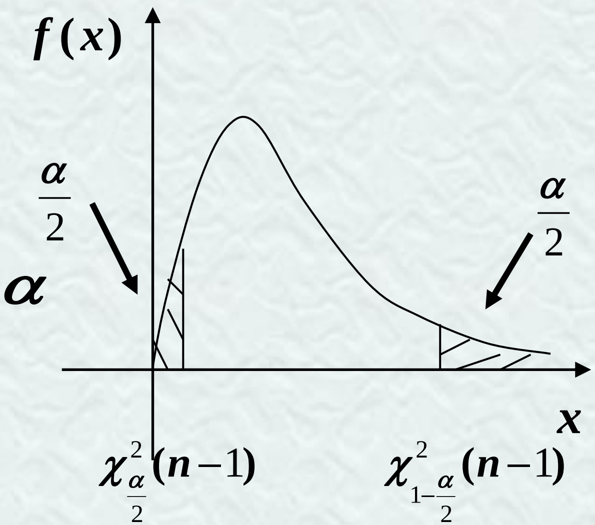
(寻找合适的枢轴量)

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

由定理5.4： $\Rightarrow \dfrac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$

$$P\left\{\lambda_1 < \frac{(n-1)S_n^{*2}}{\sigma^2} < \lambda_2\right\} = 1-\alpha$$

$$P\left\{\chi^2_{\frac{\alpha}{2}}(n-1) < \frac{(n-1)S_n^{*2}}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}(n-1)\right\} = 1-\alpha$$



$$\Rightarrow P\left\{\frac{(n-1)S_n^{*2}}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)S_n^{*2}}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right\} = 1-\alpha$$

μ未知时,σ^2的$1-\alpha$置信区间为:

$$\left(\frac{(n-1)S_n^{*2}}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} , \frac{(n-1)S_n^{*2}}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right)$$

若不是等区间($\frac{\alpha}{2}$),设为α_1,则置信区间为:

$$\left(\frac{(n-1)S_n^{*2}}{\chi^2_{1-\alpha_1}(n-1)} , \frac{(n-1)S_n^{*2}}{\chi^2_{\alpha-\alpha_1}(n-1)} \right)$$

标准差σ的一个置信度为$1-\alpha$的置信区间

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

$\chi^2_{0.975}(15)$　　$\chi^2_{0.025}(15)$

例3 有一大批糖果,现从中随机地取16袋, 称得重量(克)如下:

506	508	499	503	504	510	497	512
514	505	493	496	506	502	509	496

设袋装糖果的重量服从正态分布, 试求总体方差 σ^2 和标准差 σ 的置信度为0.95的置信区间. $\alpha = 0.05.$

解 $\dfrac{\alpha}{2} = 0.025,\quad 1-\dfrac{\alpha}{2} = 0.975,\quad n-1=15,$

查 $\chi^2(n-1)$ 分布表可知:

$$\chi^2_{1-0.025}(15) = 27.488,\qquad \chi^2_{0.025}(15) = 6.262,$$

计算得 $s_n^* = 6.2022$,

S_n^{*2}

方差 σ^2 的置信度为 $1-\alpha = 95\%$ 的置信区间

$$\left(\frac{(n-1)S_n^{*2}}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)S_n^{*2}}{\chi^2_{\alpha/2}(n-1)} \right).$$

$$\left(\frac{15 \times 6.2022^2}{27.488}, \frac{15 \times 6.2022^2}{6.262} \right)$$ 即 $(20.99, 92.14)$

标准差的置信区间 $\left(\sqrt{20.99}, \sqrt{92.14}\right) = (4.58,\ 9.60).$

四、双正态母体（相互独立）均值差的置信区间

(寻找合适的枢轴量)

设$\xi \sim N(\mu_1,\sigma_1^2),\eta \sim N(\mu_2,\sigma_2^2)$ 则$\bar{\xi} \sim N(\mu_1,\dfrac{\sigma_1^2}{n_1}),\bar{\eta} \sim N(\mu_2,\dfrac{\sigma_2^2}{n_2})$

$\bar{\xi}$与$\bar{\eta}$的独立性,$\bar{\xi}-\bar{\eta} \sim N(\mu_1-\mu_2,\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2})$

1.若σ_1^2和σ_2^2已知

$$U=\frac{\bar{\xi}-\bar{\eta}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

$$\left(\bar{\xi}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\bar{\xi}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) \longrightarrow$$

$$\left(\bar{\xi}-\bar{\eta}-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\bar{\xi}-\bar{\eta}+u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

2.若 $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知

类似的, $\bar{\xi}-\bar{\eta}\sim N(\mu_1-\mu_2,(\frac{1}{n_1}+\frac{1}{n_2})\sigma^2)$

且 $\frac{(n_1-1)S_{n_1}^{*2}}{\sigma^2}\sim\chi^2(n_1-1),\frac{(n_2-1)S_{n_2}^{*2}}{\sigma^2}\sim\chi^2(n_2-1)$

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$$

由独立性及 χ^2 分布的可加性:

$$\frac{(n_1-1)S_{n_1}^{*2}}{\sigma^2}+\frac{(n_2-1)S_{n_2}^{*2}}{\sigma^2}\sim\chi^2(n_1+n_2-2)$$

由 t 分布的构造,则 $T=\dfrac{(\bar{\xi}-\bar{\eta})-(\mu_1-\mu_2)}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$

其中 $S_W^2=\dfrac{(n_1-1)S_{n_1}^{*2}+(n_2-1)S_{n_2}^{*2}}{n_1+n_2-2}$ (两个样本的联合方差)

$$\left(\bar{\xi}-t_{1-\frac{\alpha}{2}}(n-1)\frac{S_n^*}{\sqrt{n},},\bar{\xi}+t_{1-\frac{\alpha}{2}}(n-1)\frac{S_n^*}{\sqrt{n}}\right)\longrightarrow$$

$$\left(\bar{\xi}-\bar{\eta}-t_{1-\frac{\alpha}{2}}(n_1+n_2-2)S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\bar{\xi}-\bar{\eta}+t_{1-\frac{\alpha}{2}}(n_1+n_2-2)S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

其中$\boldsymbol{S_w^2}=\dfrac{(\boldsymbol{n}_1-1)\boldsymbol{S}_{\boldsymbol{n}_1}^{*2}+(\boldsymbol{n}_2-1)\boldsymbol{S}_{\boldsymbol{n}_2}^{*2}}{\boldsymbol{n}_1+\boldsymbol{n}_2-2}$

(两个样本的联合方差)

例4 为比较I, II两种型号步枪子弹的枪口速度, 随机地取I型子弹10发, 得到枪口速度的平均值为 $\overline{x}_1 = 500(\mathrm{m/s})$, 标准差 $s_1^* = 1.10(\mathrm{m/s})$, 随机地取II型子弹20发, 得枪口速度平均值为 $\overline{x}_2 = 496(\mathrm{m/s})$, 标准差 $s_2^* = 1.20(\mathrm{m/s})$, 假设两总体都可认为近似地服从正态分布,且由生产过程可认为它们的方差相等, 求两总体均值差 $\mu_1 - \mu_2$的置信度为 0.95的置信区间.

解 由题意, 两总体样本独立且方差相等(但未知),

$$\frac{\alpha}{2}=0.025,\quad n_1=10,\quad n_2=20,\quad n_1+n_2-2=28,$$

查 $t(n-1)$ 分布表可知： $t_{0.975}(28)=2.0484,$

$$s_w^{\ 2}=\frac{9\times1.10^2+19\times1.20^2}{28},\quad s_w=\sqrt{s_w^{\ 2}}=1.1688,$$

于是得 $\mu_1-\mu_2$的一个置信度为 0.95的置信区间

$$\left(\bar{x}_1-\bar{x}_2\pm S_w\times t_{0.975}(28)\sqrt{\frac{1}{10}+\frac{1}{20}}\right)=(4\pm0.93),$$

即所求置信区间为(3.07, 4.93).

例5 为提高某一化学生产过程的得率，试图采用一种新的催化剂，为慎重起见，在试验工厂先进行试验．设采用原来的催化剂进行了$n_1=8$次试验，得到得率的平均值 $\bar{x}_1=91.73$．样本方差 $s_1^{*2}=3.89$，又采用新的催化剂进行 $n_2=8$次试验，得到得率的平均值 $\bar{x}_2=93.75$，样本方差 $s_2^{*2}=4.02$，假设两总体都可认为近似地服从正态分布，且方差相等，求两总体均值差 $\mu_1-\mu_2$ 的置信水平为0.95的置信区间.

解 因为

$$s_w^{\ 2} = \frac{(n_1-1)s_1^{*2}+(n_2-1)s_2^{*2}}{n_1+n_2-2} = 3.96,$$

于是得 $\mu_1-\mu_2$的一个置信水平为0.95的置信区间

$$\left(\bar{x}_1-\bar{x}_2 \pm s_w \times t_{1-0.025}(14)\sqrt{\frac{1}{8}+\frac{1}{8}}\right) = (-2.02 \pm 2.13),$$

即 所求置信区间为$(-4.15,\ 0.11)$.

3.两个总体方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间

仅讨论总体均值 μ_1, μ_2 为未知的情况.

$\dfrac{\sigma_1^2}{\sigma_2^2}$ 的一个置信度为 $1-\alpha$ 的置信区间

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

推导过程如下:

由于 $\dfrac{(n_1-1)S_1^{*2}}{\sigma_1^2} \sim \chi^2(n_1-1)$, $\dfrac{(n_2-1)S_2^{*2}}{\sigma_2^2} \sim \chi^2(n_2-1)$,

且由假设知 $\dfrac{(n_1-1)S_1^{*2}}{\sigma_1^2}$ 与 $\dfrac{(n_2-1)S_2^{*2}}{\sigma_2^2}$ 相互独立,

根据F分布的定义, 知

$$\frac{S_1^{*2}\Big/\sigma_1^2}{S_2^{*2}\Big/\sigma_2^2}=\frac{\dfrac{(n_1-1)S_1^{*2}}{\sigma_1^2}\Bigg/(n_1-1)}{\dfrac{(n_2-1)S_2^{*2}}{\sigma_2^2}\Bigg/(n_2-1)}\sim F(n_1-1,n_2-1),$$

$$P\left\{F_{\alpha/2}(n_1-1,n_2-1)<\frac{S_1^{*2}\big/\sigma_1^2}{S_2^{*2}\big/\sigma_2^2}<F_{1-\alpha/2}(n_1-1,n_2-1)\right\}=1-\alpha,$$

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

于是得 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的一个置信度为 $1-\alpha$ 的置信区间

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

例6 研究由机器 A 和机器 B 生产的钢管内径, 随机抽取机器 A 生产的管子 18 只, 测得样本方差为 $s_1^{*2}=0.34(\text{mm}^2)$; 抽取机器$B$生产的管子 13 只, 测得样本方差为 $s_2^{*2}=0.29(\text{mm}^2)$. 设两样本相互独立,且设由机器 A 和机器 B 生产的钢管内径分别服从正态分布 $N(\mu_1,\sigma_1^2), N(\mu_2,\sigma_2^2)$, $\mu_i,\sigma_i^2 (i=1,2)$ 均未知, 求方差比 σ_1^2/σ_2^2 的置信度为 0.90 的置信区间.

解 $n_1=18,\quad n_2=13,\quad \alpha=0.10,$

$$s_1^{*2}=0.34(\text{mm}^2),\quad s_2^{*2}=0.29(\text{mm}^2),$$

$$F_{1-\alpha/2}(n_1-1,n_2-1)=F_{0.95}(17,12)=2.59,$$

$$F_{\alpha/2}(17,12)=F_{0.05}(17,12)=\frac{1}{F_{0.95}(12,17)}=\frac{1}{2.38},$$

于是得 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的一个置信度为0.90的置信区间

$$\left(\frac{0.34}{0.29}\times\frac{1}{2.59},\ \frac{0.34}{0.29}\times 2.38\right)=(0.45,\ 2.79).$$

补充：单侧置信区间

一、问题的引入

二、基本概念

三、典型例题

四、小结

一、问题的引入

在以上各节的讨论中,对于未知参数θ,我们给出两个统计量$\underline{\theta}$,$\overline{\theta}$,得到θ的双侧置信区间$(\underline{\theta},\overline{\theta})$.

但在某些实际问题中,例如,对于设备、元件的寿命来说,平均寿命长是我们希望的,我们关心的是平均寿命θ的“下限”;与之相反,在考虑产品的废品率p时,我们常关心参数p的“上限”,这就引出了单侧置信区间的概念.

二、基本概念

1. 单侧置信区间的定义

对于给定值 $\alpha\ (0<\alpha<1)$, 若由样本 $X_1, X_2, \cdots, X_n$ 确定的统计量 $\underline{\theta}=\underline{\theta}(X_1, X_2, \cdots, X_n)$, 对于任意 $\theta\in\Theta$ 满足

$$P\{\theta>\underline{\theta}\}\geq 1-\alpha,$$

则称随机区间 $(\underline{\theta}, +\infty)$ 是 θ 的置信水平为 $1-\alpha$ 的单侧置信区间, $\underline{\theta}$ 称为 θ 的置信水平为 $1-\alpha$ 的单侧置信下限.

又如果统计量 $\overline{\theta}=\overline{\theta}(X_1,X_2,\cdots,X_n)$,对于任意$\theta\in\Theta$满足

$$P\{\theta<\overline{\theta}\}\geq 1-\alpha,$$

则称随机区间$(-\infty,\overline{\theta})$是$\theta$的置信水平为$1-\alpha$的单侧置信区间,$\overline{\theta}$称为$\theta$的置信水平为$1-\alpha$的单侧置信上限.

以下均设$S^2=S_n^{*2}$

2. 正态总体均值与方差的单侧置信区间

设正态总体 X 的均值是μ, 方差是σ^2 (均为未知),

$X_1, X_2, \cdots, X_n$ 是一个样本，由 $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$,

有 $P\left\{\dfrac{\bar{X}-\mu}{S/\sqrt{n}} < t_{1-\alpha}(n-1)\right\} = 1-\alpha$,

即 $P\left\{\mu > \bar{X} - \dfrac{S}{\sqrt{n}} t_{1-\alpha}(n-1)\right\} = 1-\alpha$,

于是得 μ 的一个置信水平为 $1-\alpha$ 的单侧置信区间

$$\left(\bar{X}-\frac{S}{\sqrt{n}}t_{1-\alpha}(n-1),+\infty\right),$$

μ 的置信水平为 $1-\alpha$ 的置信下限 $\underline{\mu}=\bar{X}-\frac{S}{\sqrt{n}}t_{1-\alpha}(n-1).$

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

有 $P\left\{\frac{(n-1)S^2}{\sigma^2}>\chi_{\alpha}^2(n-1)\right\}=1-\alpha,$

即 $$P\left\{\sigma^2 < \frac{(n-1)S^2}{\chi_\alpha^2(n-1)}\right\} = 1-\alpha,$$

于是得 σ^2 的一个置信水平为 $1-\alpha$ 的单侧置信区间

$$\left(0,\ \frac{(n-1)S^2}{\chi_\alpha^2(n-1)}\right),$$

σ^2 的置信水平为 $1-\alpha$ 的单侧置信上限

$$\overline{\sigma^2} = \frac{(n-1)S^2}{\chi_\alpha^2(n-1)}.$$

三、典型例题

例7 设从一批灯泡中, 随机地取5只作寿命试验, 测得寿命(以小时计)为 1050, 1100, 1120, 1250, 1280, 设灯泡寿命服从正态分布, 求灯泡寿命平均值的置信水平为 0.95 的单侧置信下限.

解 $1-\alpha=0.95,\quad n=5,\quad \bar{x}=1160,\quad s^2=9950,$

$$t_{1-\alpha}(n-1)=t_{1-0.05}(4)=2.1318,$$

μ 的置信水平为 0.95 的置信下限

$$\underline{\mu}=\bar{x}-\frac{s}{\sqrt{n}}t_{1-\alpha}(n-1)=1065.$$

补例 设总体 X 在 $[0,\theta]$ 上服从均匀分布，其中 θ $(\theta>0)$ 未知，$(X_1,X_2,\cdots,X_n)$ 是来自总体X的样本，给定 α，求 θ 的置信水平为 $1-\alpha$ 的置信区间.

解 令 $X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}$,

$X_{(n)}$的概率密度为 $$g(y)=\begin{cases}\dfrac{ny^{n-1}}{\theta^n}, & 0\le y\le\theta,\\ 0, & \text{其他}.\end{cases}$$

取枢轴量 $Z=\dfrac{X_{(n)}}{\theta}$,

$Z=\dfrac{X_{(n)}}{\theta}$ 的概率密度为 $g(z)=\begin{cases} nz^{n-1}, & 0\le z\le 1, \\ 0, & \text{其他}. \end{cases}$

对于给定的α,可定出两个常数$a,b(0<a<b\le 1)$,

满足条件 $P\left\{a<\dfrac{X_{(n)}}{\theta}<b\right\}=1-\alpha,$

即 $1-\alpha=\int_a^b nz^{n-1}\mathrm{d}z=b^n-a^n,$

$\Rightarrow P\left\{\dfrac{X_{(n)}}{b}<\theta<\dfrac{X_{(n)}}{a}\right\}=1-\alpha,$ $\left(\dfrac{X_{(n)}}{b},\dfrac{X_{(n)}}{a}\right)$为置信区间.

拓展：

1、 大样本置信区间

2、样本量的确定

作业： p.377 7.16，7.19，7.21

2. 单个总体方差 σ^2 的置信区间

1. $\mu=\mu_0$ 已知

$$\left(\frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}, \frac{\sum_{i=1}^{n}(\xi_i-\mu_0)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right)$$

2. μ 未知

$$\left(\frac{(n-1)S_n^{*2}}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S_n^{*2}}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)$$

3. 两个总体均值差 $\mu_1-\mu_2$ 的置信区间

$\sigma_1^{\ 2}$ 和 $\sigma_2^{\ 2}$ 均为已知，

$$\left(\bar{\xi}-\bar{\eta}-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\bar{\xi}-\bar{\eta}+u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

$\sigma_1^{\ 2}=\sigma_2^{\ 2}=\sigma^2$，但 σ^2 为未知，

$$\left(\bar{\xi}-\bar{\eta}-t_{1-\frac{\alpha}{2}}(n_1+n_2-2)S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\bar{\xi}-\bar{\eta}+t_{1-\frac{\alpha}{2}}(n_1+n_2-2)S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

其中 $S_w^2=\dfrac{(n_1-1)S_{n_1}^{*2}+(n_2-1)S_{n_2}^{*2}}{n_1+n_2-2}$

(两个样本的联合方差)

4. 两个总体方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间

总体均值 μ_1, μ_2 为未知

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

四、小结

正态总体均值 μ 的置信水平为 $1-\alpha$ 的单侧置信区间

$$\left(-\infty, \boxed{\bar{X}+\frac{S}{\sqrt{n}}t_{1-\alpha}(n-1)}\right),$$

单侧置信上限 $\overline{\mu}$

$$\left(\boxed{\bar{X}-\frac{S}{\sqrt{n}}t_{1-\alpha}(n-1)}, +\infty\right),$$

单侧置信下限 $\underline{\mu}$

正态总体方差 σ^2 的置信水平为 $1-\alpha$ 的单侧置信区间

$$\left(0, \boxed{\frac{(n-1)S^2}{\chi_{\alpha}^2(n-1)}}\right).$$

单侧置信上限 $\overline{\sigma^2}$