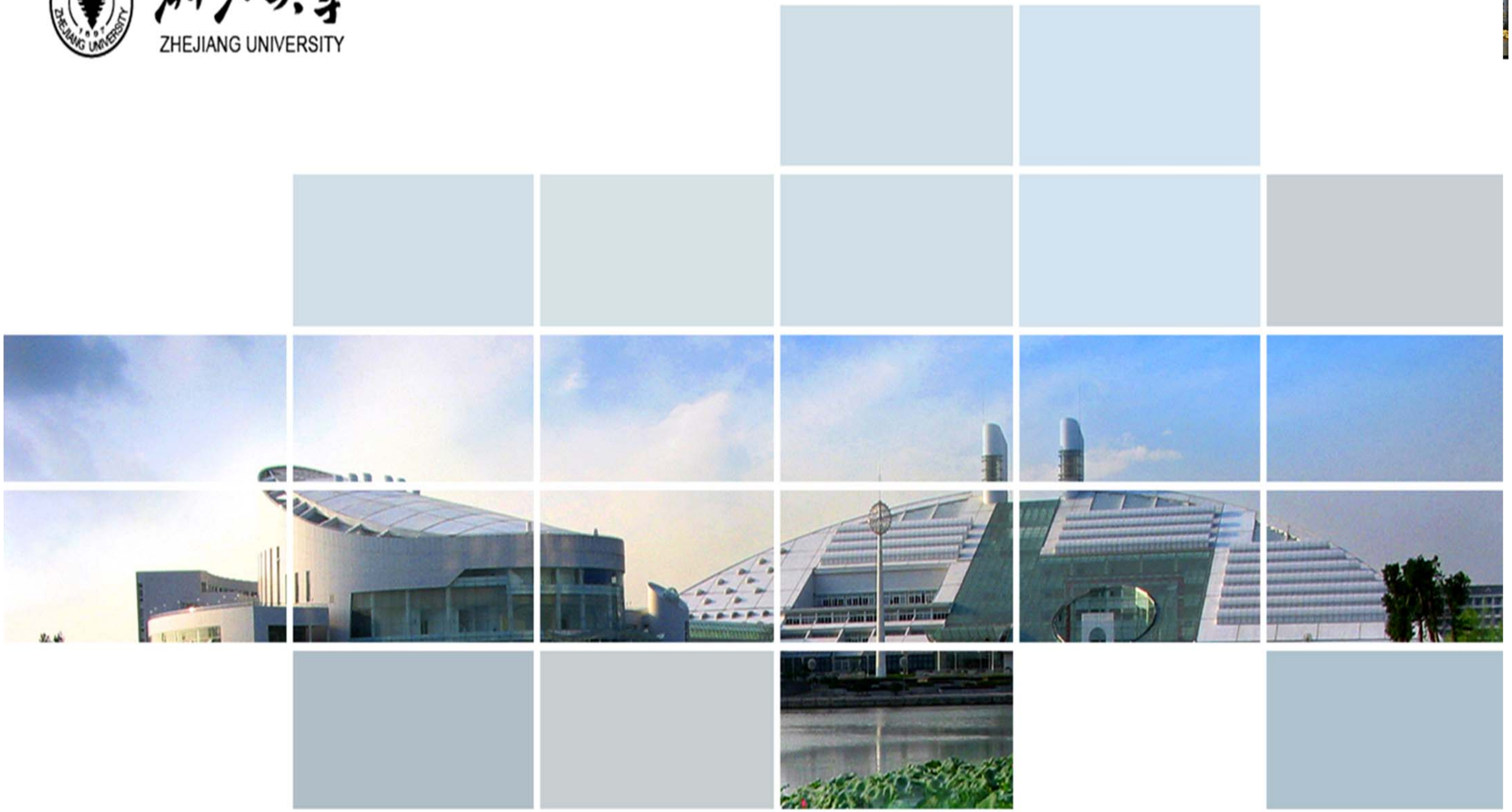
浙江大学
ZHEJIANG UNIVERSITY



第51讲 单个正态总体均值的区间估计

1. 正态总体均值μ的置信区间

设总体$X \sim N\left(\mu,\sigma^2\right)$，$X_1,X_2,\cdots,X_n$为样本.

$\overline{X}$和S^2分别为样本均值和样本方差.

置信水平为$1-\alpha$.

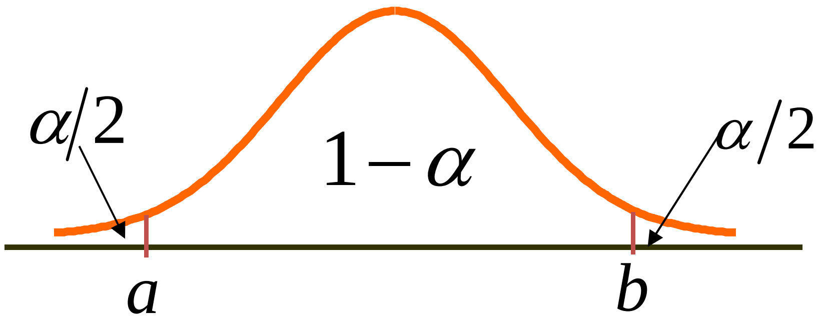
(1) σ^2已知时

$\bar{X}$是μ的极大似然估计，取枢轴量$G=\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

设常数$a<b$满足：$P\{a<G<b\}\geq 1-\alpha$

等价于 $P\left\{\bar{X}-\dfrac{\sigma}{\sqrt{n}}b<\mu<\bar{X}-\dfrac{\sigma}{\sqrt{n}}a\right\}\geq 1-\alpha$

此时区间的长度为 $(b-a)\sigma/\sqrt{n}$



由正态分布的对称性知，$a=-b=-z_{\alpha/2}$

时，区间的长度L达到最短. $L=2z_{\alpha/2}\sigma/\sqrt{n}$. n固定.

置信水平提高，即$(1-\alpha)$增大，则$z_{\alpha/2}$增大，

所以L变大，精确度降低；反之亦然.

所以μ的双侧置信区间为：

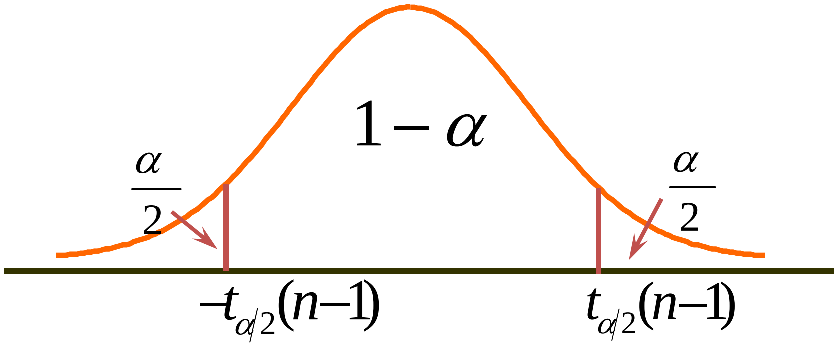
$$\left(\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2}, \bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$$

单侧置信下限为$\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha}$

单侧置信上限为$\bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha}$

(2) σ^2未知时

以S^2估计σ^2，得 枢轴量$G=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$



由 $-t_{\alpha/2}(n-1)<G<t_{\alpha/2}(n-1)$解得，

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

所以μ的置信区间为：

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

单侧置信下限为$\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha}(n-1)$,

单侧置信上限为$\bar{X}+\frac{S}{\sqrt{n}}t_{\alpha}(n-1)$.

正态总体均值μ的置信区间模拟见实验13.

例 1. 某袋装食品重量(单位：克) $X \sim N(\mu,\sigma^2)$. 现从一大批该产品中随机抽取10件, 称得重量为：

101.3 96.6 100.4 98.8 94.6

103.1 102.3 97.5 105.4 100.2

(1) $\sigma = 3$, (2) σ未知，求(1),(2)两种情况下 μ的置信水平为95%的双侧置信区间.

本例的Excel计算见实验14.

解：$n=10,\ \alpha=0.05$

计算得 $\overline{x}=100.1, s=2.54,$

（1）$\sigma=3$, 查表得 $z_{0.025}=1.96$

所以, μ的置信水平为95%的双侧置信区间为

$$\left(\overline{x}-\frac{3}{\sqrt{10}}z_{0.025}, \overline{x}+\frac{3}{\sqrt{10}}z_{0.025}\right)=(98.24,\ 101.96)$$

(2) σ未知，查表得$t_{0.025}(9)=2.26$

此时，μ置信水平为95%的双侧置信区间为

$$(\overline{x}-\frac{s}{\sqrt{10}}t_{0.025}(9),\overline{x}+\frac{s}{\sqrt{10}}t_{0.025}(8))=(98.29,101.91)$$

实际中σ^2未知的情况更多.

例 2. 设新生儿体重(单位：克) $X \sim N(\mu, \sigma^2)$, μ, σ^2未知.

现从某妇产医院随机抽查16名新生儿, 称得重量为：

3200	**3050**	**2600**	**3530**	**3840**	**4450**	**2900**	**4180**
2150	**2650**	**2750**	**3450**	**2830**	**3730**	**3620**	**2270**

求μ的置信水平为95%的双侧置信区间.

解：$n=16, \alpha=0.05$, σ未知

计算得 $\overline{x}=3200, s=665.48$

查表得$t_{0.025}(15)=2.1315$

所以μ的置信水平为95%的双侧置信区间为：

$$(\overline{x}-\frac{s}{\sqrt{16}}t_{0.025}(15), \overline{x}+\frac{s}{\sqrt{16}}t_{0.025}(15))$$

$$=(2845.4, \quad 3554.6)$$

例 3. 某种产品的寿命(单位：千小时) $X \sim N(\mu,\sigma^2)$, μ,σ^2未知. 现随机抽查10件产品进行寿命试验，测得：样本均值$\overline{x}=5.78$, 样本标准差$s=0.92$. 求μ的置信水平为95%的单侧置信下限.

解: 查表得$t_{0.05}(9)=1.8331$

所以μ的置信水平为95%的单侧置信下限为:

$$\overline{x}-\frac{s}{\sqrt{10}}t_{0.05}(9)=5.78-\frac{0.92}{\sqrt{10}}\times 1.8331=5.25$$

2.其他总体均值的区间估计

设总体X的均值为μ,方差为σ^2,非正态分布或不知分布形式.样本为$X_1,\ldots,X_n$.

当n充分大（一般$n>30$）时，由中心极限定理知,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{\text{近似}}{\sim} N(0,1).$$

当σ^2已知时，μ的置信水平为$1-\alpha$的近似置信区间为

$$\left(\bar{X} - z_{\alpha/2}\sigma/\sqrt{n},\quad \bar{X} + z_{\alpha/2}\sigma/\sqrt{n}\right)$$

当σ^2未知时，以样本方差S^2代入，得近似置信区间为

$$\left(\bar{X} - z_{\alpha/2}S/\sqrt{n},\quad \bar{X} + z_{\alpha/2}S/\sqrt{n}\right)$$

例4: 某市随机抽取1500个家庭，调查知道其中有375家拥有私家车. 试根据此调查结果，求该市拥有私家车比例 p 的置信水平为95%近似置信区间.

解：$\hat{p}=\bar{x}=\dfrac{375}{1500}=0.25,\ s^2\approx\hat{p}(1-\hat{p})=0.1875$

代入近似置信区间

$$\left(\bar{X}-z_{0.025}S/\sqrt{n},\quad \bar{X}+z_{0.025}S/\sqrt{n}\right)$$

得近似置信区间为 $(0.228,\quad 0.272)$.