Lec-23. 正态总体的区间估计

主讲教师: 吴利苏 (**wulisu@sdust.edu.cn**)
主　　页: wulisu.cn

本次课内容

单个正态总体参数的区间估计

- σ^2 已知, μ 的置信区间
- σ^2 未知, μ 的置信区间
- 其他总体 μ 的置信区间
- μ 未知, σ^2 的置信区间

两个正态总体参数的区间估计

- σ_1^2, σ_2^2 已知, $\mu_1 - \mu_2$ 的置信区间
- $\sigma_1^2 = \sigma_1^2 = \sigma^2$ 未知, $\mu_1 - \mu_2$ 的置信区间
- μ_1, μ_2 未知, σ_1^2/σ_2^2 的置信区间

正态总体均值 μ 的置信区间 (σ^2 已知时)

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 $1-\alpha$.

- σ^2 已知时, 取枢轴量 $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, 则 μ 的一个置信水平为 $1-\alpha$ 的区间估计为

$$\left(\overline{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right).$$

σ^2 已知时, $\overline{X}$ 是 μ 的最大似然估计, 枢轴量

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

设常数 $a<b$ 满足:

$$P\{a<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<b\} \geq 1-\alpha$$

等价于

$$P\left\{\overline{X}-\frac{\sigma}{\sqrt{n}}b<\mu<\overline{X}-\frac{\sigma}{\sqrt{n}}a\right\} \geq 1-\alpha$$

此时区间的长度为 $L=(b-a)\frac{\sigma}{\sqrt{n}}$.

由正态分布的对称性知, 当

$$-a = b = z_{\alpha/2}$$

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

固定 n, L 变大, $z_{\alpha/2}$ 增大, 则 $(1-\alpha)$ 增大, 置信水平提高, 精确度降低; 反之亦然.

所以, μ 的

- 双侧置信区间为:

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

- 单侧置信下限为 $\overline{X} - \frac{\sigma}{\sqrt{n}}z_{\alpha}$,
- 单侧置信上限为 $\overline{X} + \frac{\sigma}{\sqrt{n}}z_{\alpha}$.

正态总体均值 μ 的置信区间 (σ^2 未知时)

设总体 $X \sim N\left(\mu, \sigma^2\right)$, $X_1, X_2, \cdots, X_n$ 为样本.
$\bar{X}$ 和 S^2 分别为样本均值和样本方差.

- σ^2 未知时, 取枢轴量 $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$, 则 μ 的一个置信水平为 $1-\alpha$ 的区间估计为

$$\left(\overline{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right).$$

σ^2 未知时, S^2 是 σ^2 的无偏估计, 用 S 替换 σ, 得枢轴量

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

由

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

解得,

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

所以 μ 的

- 置信区间为:

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

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

例

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 16 件, 称得重量为:

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

$(\overline{x} = 503.75, s = 6.2022,)$ 求在

(1) $\sigma = 3$;

(2) σ 未知

两种情况下 μ 的置信水平为 95% 的双侧置信区间.

解: $n=16, n-1=15, \alpha/2=0.025$. 计算得
$\overline{x}=503.75, s=6.2022$.
(1) $\sigma=3$, 查表得 $z_{0.025}=1.96$ 所以, μ 的置信水平为 95% 的置信区间为
$$(\overline{x}-\frac{3}{\sqrt{16}}z_{0.025}, \overline{x}+\frac{3}{\sqrt{16}}z_{0.025})=(502.28, 505.22).$$

(2) σ 未知, 查表得 $t_{0.025}(15)=2.1315$ 此时, μ 置信水平为 95% 的置信区间为

$$(\overline{x}-\frac{S}{\sqrt{16}}t_{0.025}(15), \overline{x}+\frac{S}{\sqrt{10}}t_{0.025}(15))=(500.4, 507.1)$$

□

实际中 σ^2 未知的情况更多.

例

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

求 μ 的置信水平为 95% 的双侧置信区间.
$(\overline{x} = 3200, s = 665.48)$

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:

$$(\overline{x}-\frac{S}{\sqrt{16}}t_{0.025}(15), \overline{x}+\frac{S}{\sqrt{16}}t_{0.025}(15))=(2845.4, 3554.6).$$

$\square$

其他总体均值的区间估计

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,

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

设 $\bar{X}$ 和 S^2 分别为样本均值和样本方差. μ 的置信水平为 $1 - \alpha$ 的置信区间

- σ^2 已知时, 置信区间近似为 $\left(\overline{X} \pm z_{\alpha/2}\sigma/\sqrt{n}\right)$.
- σ^2 未知时, 置信区间近似为 $\left(\overline{X} \pm z_{\alpha/2}S/\sqrt{n}\right)$.

例

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: $\hat{p} = \bar{x} = \frac{375}{1500} = 0.25, s^2 \approx \hat{p}(1 - \hat{p}) = 0.1875$
代入近似置信区间

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

得近似置信区间为 $(0.228, 0.272)$. $\square$

正态总体方差 σ^2 的置信区间 (μ 未知时)

设总体 $X \sim N\left(\mu, \sigma^2\right)$, $\quad X_1, X_2, \cdots, X_n$ 为样本. $\bar{X}$ 和 S^2 分别为样本均值和样本方差.

- μ 未知, 取枢轴量 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 则 μ 的一个置信水平为 $1-\alpha$ 的区间估计为

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

S^2 为 σ^2 的无偏估计, 故取枢轴量

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

则

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

等价的,

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

正态总体标准差 σ 的置信区间 (μ 未知时)

设总体 $X \sim N\left(\mu, \sigma^2\right)$, $X_1, X_2, \cdots, X_n$ 为样本. $\bar{X}$ 和 S^2 分别为样本均值和样本方差.

- μ 未知, 取枢轴量 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, 则 μ 的一个置信水平为 $1-\alpha$ 的区间估计为

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

例

某袋装食品重量 (单位: 克) $X \sim N(\mu, \sigma^2)$, μ 未知. 现从一大批该产品中随机抽取 16 件, 称得重量为:

$$
\begin{array}{cccccccc}
506 & 508 & 499 & 503 & 504 & 510 & 497 & 512 \\
514 & 505 & 493 & 496 & 506 & 502 & 509 & 496
\end{array}
$$

求标准差 σ 的置信度为 95% 置信区间.

例

一个园艺科学家正在培养一个新品种的苹果，这种苹果除了口感好和颜色鲜艳以外，另一个重要特征是单个重量差异不大．为了评估新苹果，她随机挑选了 25 个测试重量（单位：克），其样本方差为 $s^2 = 4.25$．试求 σ^2 的置信水平为 95% 置信区间．

解: $n = 25, s^2 = 4.25, \alpha = 0.05$ 查表得:
$\chi^2_{0.025}(24) = 39.4, \quad \chi^2_{0.975}(24) = 12.4;$

$$\chi^2_{0.95}(24) = 13.85,$$

σ^2 的双侧置信区间为

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} \right) = (2.59, 8.23).$$

$\square$

两个正态总体均值差 $\mu_1 - \mu_2$ 的置信区间

设总体 $X \sim N\left(\mu_1, \sigma_1^2\right)$, $Y \sim N\left(\mu_2, \sigma_2^2\right)$, σ_1^2, σ_2^2 已知.
$X_1, \cdots, X_{n_1}$ 和 $Y_1, \cdots, Y_{n_2}$ 分别为来自总体 X, Y 的样本,
这两个样本相互独立. $\bar{X}, \bar{Y}$ 分别为 X, Y 的样本均值,
S_1^2, S_2^2 分别为 X, Y 的样本方差. 取枢轴量

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

- 则 $\mu_1 - \mu_2$ 的一个置信水平为 $1 - \alpha$ 的区间估计为

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

$\bar{X}$ 和 $\bar{Y}$ 分别为 μ_1, μ_2 的无偏估计，故 $\bar{X} - \bar{Y}$ 是 $\mu_1 - \mu_2$ 的无偏估计. $\bar{X}, \bar{Y}$ 相互独立,
$\bar{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$, $\bar{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$.

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right),$$

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

$$P\{-z_{\alpha/2} < G < z_{\alpha/2}\} = 1 - \alpha$$

解得置信区间为 $\left(\bar{X} - \bar{Y} \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$.

两个正态总体均值差 $\mu_1 - \mu_2$ 的置信区间

设总体 $X \sim N\left(\mu_1, \sigma_1^2\right)$, $Y \sim N\left(\mu_2, \sigma_2^2\right)$, $\sigma_1^2 = \sigma_1^2 = \sigma^2$ 未知. 取枢轴量

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2),$$

- 则 $\mu_1 - \mu_2$ 的一个置信水平为 $1 - \alpha$ 的区间估计为

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

其中 $S_W = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$, $S_W = \sqrt{S_W^2}$.

例

两个正态总体中, 方差未知且相等. 设样本独立且

$n_1 = 10,\ \bar{x}_1 = 500,\ s_1 = 1.10;$

$n_2 = 20,\ \bar{x}_2 = 496,\ s_2 = 1.20.$

求 $\mu_1 - \mu_2$ 的一个置信水平为 0.95 的置信区间.

例

两个正态总体中，方差未知且相等. 设样本独立且

$n_1 = 10$, $\bar{x}_1 = 500$, $s_1 = 1.10$;

$n_2 = 20$, $\bar{x}_2 = 496$, $s_2 = 1.20$.

求 $\mu_1 - \mu_2$ 的一个置信水平为 0.95 的置信区间.

得到的置信区间的下限大于零，则推断 $\mu_1 > \mu_2$.

例

两个正态总体中，方差未知且相等. 设样本独立且

$n_1 = 8,\ \bar{x}_1 = 91.73,\ s_1^2 = 3.89;$

$n_2 = 8,\ \bar{x}_2 = 93.75,\ s_2^2 = 4.02.$

求 $\mu_1 - \mu_2$ 的一个置信水平为 0.95 的置信区间.

例

两个正态总体中，方差未知且相等. 设样本独立且

$n_1 = 8,\ \bar{x}_1 = 91.73,\ s_1^2 = 3.89;$

$n_2 = 8,\ \bar{x}_2 = 93.75,\ s_2^2 = 4.02.$

求 $\mu_1 - \mu_2$ 的一个置信水平为 0.95 的置信区间.

得到的置信区间的包含零，则推断 μ_1 和 μ_2 没有显著差别.

设两个正态总体均值差 $\mu_1 - \mu_2$ 的置信区间为 $(\underline{\theta}, \overline{\theta})$,

- 若 $0 \in (\underline{\theta}, \overline{\theta})$, 则推断 $\mu_1 = \mu_2$;

$$\underline{\theta} < 0 < \overline{\theta} \xrightarrow{\text{推断}} \mu_1 = \mu_2$$

- 若 $(\underline{\theta}, \overline{\theta})$ 在 0 的右侧, 则推断 $\mu_1 > \mu_2$;

$$0 < \underline{\theta} \xrightarrow{\text{推断}} \mu_1 > \mu_2$$

- 若 $(\underline{\theta}, \overline{\theta})$ 在 0 的左侧, 则推断 $\mu_1 < \mu_2$.

$$\overline{\theta} < 0 \xrightarrow{\text{推断}} \mu_1 < \mu_2.$$

两个正态总体方差 σ_1^2/σ_2^2 的置信区间

设总体 $X \sim N\left(\mu_1, \sigma_1^2\right)$, $Y \sim N\left(\mu_2, \sigma_2^2\right)$, μ_1, μ_2 未知. 取枢轴量

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

- 则 σ_1^2/σ_2^2 的一个置信水平为 $1-\alpha$ 的区间估计为

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

例

两个正态总体中，均值方差均未知. 设样本独立且 $n_1 = 18,\ s_1^2 = 0.34;\ n_2 = 13,\ s_2^2 = 0.29.$
求 σ_1^2/σ_2^2 的一个置信水平为 0.90 的置信区间.

例

两个正态总体中，均值方差均未知．设样本独立且 $n_1 = 18,\ s_1^2 = 0.34;\ n_2 = 13,\ s_2^2 = 0.29.$
求 σ_1^2/σ_2^2 的一个置信水平为 0.90 的置信区间．

得到的置信区间的包含 1，则推断 σ_1^2 和 σ_2^2 没有显著差别．

设两个正态总体方差商 σ_1^2/σ_2^2 的置信区间为 $(\underline{\theta}, \overline{\theta})$,

- 若 $1 \in (\underline{\theta}, \overline{\theta})$, 则推断 $\underline{\theta} = \overline{\theta}$;

$$\underline{\theta} < 1 < \overline{\theta} \xrightarrow{\text{推断}} \sigma_1^2 = \sigma_2^2$$

- 若 $(\underline{\theta}, \overline{\theta})$ 在 1 的右侧, 则推断 $\sigma_1^2 > \sigma_2^2$;

$$\underline{\theta} > 1 \xrightarrow{\text{推断}} \sigma_1^2 > \sigma_2^2$$

- 若 $(\underline{\theta}, \overline{\theta})$ 在 1 的左侧, 则推断 $\sigma_1^2 < \sigma_2^2$.

$$0 < \overline{\theta} < 1 \xrightarrow{\text{推断}} \sigma_1^2 < \sigma_2^2$$

单个正态总体参数的区间估计

待估参数	其他参数	枢轴量	置信区间
μ	σ^2 已知	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(\overline{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$
μ	σ^2 未知	$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\overline{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right)$
σ^2	μ 未知	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$

两个正态总体参数的区间估计

待估参数	其他参数	枢轴量	置信区间
$\mu_1-\mu_2$	σ_1^2,σ_2^2 已知	$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$	$\left(\bar{X}-\bar{Y}\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$
$\mu_1-\mu_2$	$\sigma_1^2=\sigma_2^2=\sigma^2$ 未知	$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$	$\left(\bar{X}-\bar{Y}\pm t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$
σ_1^2/σ_2^2	μ_1,μ_2 未知	$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$	$\left(\frac{S_1^2}{S_2^2}\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right)$