§6.4 正态总体均值与方差的区间估计

一. 单个正态总体的情况

设给定置信度为$1-\alpha$, $X_1, X_2, \cdots, X_n$为总体$N(\mu, \sigma^2)$的样本, $\bar{X}, S^2$ 分别是样本均值和样本方差.

1. 均值 μ 的置信区间

(1) σ^2为已知, 由上节可知:

μ双侧置信区间为: $\left(\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$

μ单侧置信区间为: $\left(-\infty, \bar{X} + \frac{\sigma}{\sqrt{n}} z_\alpha\right)$, $\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_\alpha, +\infty\right)$

(2) σ^2为未知,

μ 的置信度为 $1-\alpha$的置信区间 $\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right)$.

因为根据第五章定理2. 5知 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$,

则 $P\left\{-t_{\alpha/2}(n-1) < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\right\} = 1-\alpha$,

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

类似地，可得μ的单侧置信区间为:

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

例1 有一大批糖果, 现从中随机地取16袋, 称得重量分别如下:

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

设袋装糖果的重量近似服从正态分布, 试求总体均值 μ 的置信度为 0.95 的置信区间.

解 已知 $\alpha = 0.05, \quad n - 1 = 15,$

查 $t(n-1)$ 分布表可知: $t_{0.025}(15) = 2.1315,$

计算得 $\bar{x} = 503.75, \quad s = 6.2022,$

得μ的置信度为95%的置信区间

$$\left(503.75 \pm \frac{6.2022}{\sqrt{16}} \times 2.1315\right)$$ 即 (500.4, 507.1).

就是说估计袋装糖果重量的均值在500.4克与507.1克之间, 这个估计的可信程度为95%.

若依此区间内任一值作为μ的近似值,

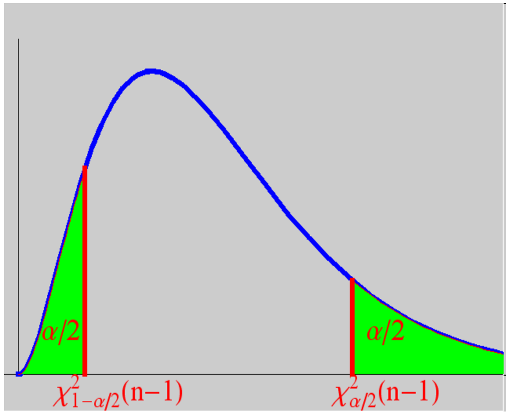
其误差不大于 $\frac{6.2022}{\sqrt{16}} \times 2.1315 \times 2 = 6.61$ (克).

这个误差的可信度为95%.

2. 方差 σ^2 的置信区间

(1) 当 μ 已知时 σ^2 的置信度为 $1-\alpha$ 的置信区间

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



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

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

(2) 当 $\boldsymbol{\mu}$ 未知，方差 $\boldsymbol{\sigma}^2$ 的置信度为 $\mathbf{1-\alpha}$ 的置信区间

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

因为 $\boldsymbol{S}^2$ 是 $\boldsymbol{\sigma}^2$ 的无偏估计，

根据第五章定理2.3知 $\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$

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

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

进一步可得:

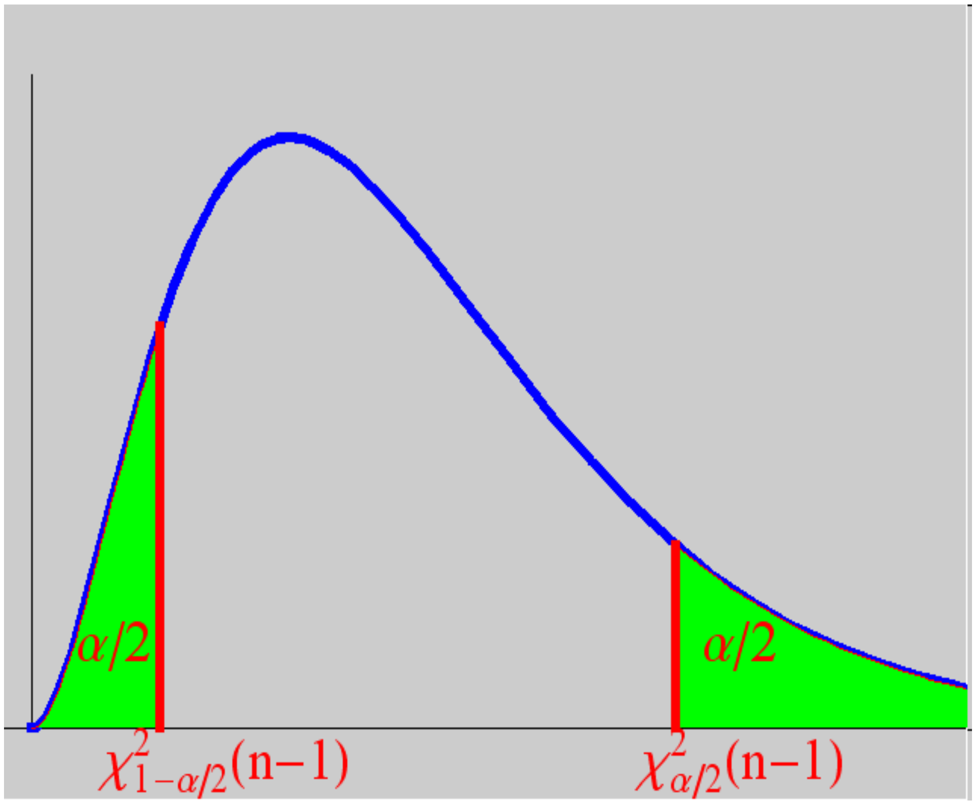
标准差σ的一个置信度为$1-\alpha$的置信区间

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

在应用中一般更关注方差的上限，类似上面的推导，可得方差σ^2的置信度为$1-\alpha$的单侧置信区间为

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

注意: 在密度函数不对称时，如 χ^2 分布和 F分布，习惯上仍取对称的分位点来确定置信区间(如图).



例2 (续例1) 求例1中总体标准差 σ的置信度为0.95的置信区间.

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

查 $\chi^2(n-1)$ 分布表可知:

$\chi^2_{0.025}(15) = 27.488,\quad \chi^2_{0.975}(15) = 6.262,$

计算得 $s = 6.2022,$

代入公式得标准差的置信区间 $(4.58,\ 9.60)$.

二. 两个正态总体的情况

设给定置信度为$1-\alpha$, 并设 $X_1, X_2, \cdots, X_{n_1}$ 为第一个总体 $N(\mu_1, \sigma_1^2)$的样本, $Y_1, Y_2, \cdots, Y_{n_2}$ 为第二个总体 $N(\mu_2, \sigma_2^2)$的样本, $\bar{X}, \bar{Y}$分别是第一、二个总体的样本均值, S_1^2, S_2^2分别是第一、二个总体的样本方差.

讨论两个整体总体均值差和方差比的估计问题.

1. 两个正态总体均值差$\mu_1-\mu_2$的置信区间

(1) σ_1^2和σ_2^2均为已知

$\mu_1-\mu_2$的一个置信度为$1-\alpha$的置信区间

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

推导过程如下:

因为$\bar{X}$, $\bar{Y}$分别是μ_1, μ_2的无偏估计,

所以$\bar{X}-\bar{Y}$是$\mu_1-\mu_2$的无偏估计,

P101定理2.1

由 $\overline{X}, \overline{Y}$ 的独立性及

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

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

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

于是得 $\mu_1-\mu_2$ 的一个置信度为 $1-\alpha$ 的置信区间

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

(2) $\boldsymbol{\sigma_1^{\,2}}$和$\boldsymbol{\sigma_2^{\,2}}$均为未知

只要$\boldsymbol{n_1}$和$\boldsymbol{n_2}$都很大(实用上 >50 即可), 则有

$\mu_1-\mu_2$的置信度为 $1-\alpha$ 的近似置信区间

$$\left(\bar{X}-\bar{Y}\pm z_{\alpha/2}\sqrt{\frac{S_1^{\,2}}{n_1}+\frac{S_2^{\,2}}{n_2}}\right).$$

P105定理2.6

(3) $\boldsymbol{\sigma_1^{\ 2} = \sigma_2^{\ 2} = \sigma^2}$**, 但** $\boldsymbol{\sigma^2}$ **为未知,**

$\mu_1 - \mu_2$的一个置信度为$1-\alpha$的置信区间为

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

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

例3 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 $\bar{x}_1 = 500(\mathrm{m/s})$,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 $\bar{x}_2 = 496(\mathrm{m/s})$, 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：$t_{0.025}(28)=2.0484,$

$$s_w^{\ 2}=\frac{9\times 1.10^2+19\times 1.20^2}{28}, \quad s_w=\sqrt{S_w^{\ 2}}=1.1688,$$

于是得 $\mu_1-\mu_2$ 的一个置信度为 0.95的置信区间

$$\left(\bar{x}_1-\bar{x}_2 \pm S_w \times t_{0.025}(28)\sqrt{\frac{1}{10}+\frac{1}{20}}\right)=(4\pm 0.93),$$

即所求置信区间为 (3.07, 4.93).

2. 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

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

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

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

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

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

$$=1-\alpha,$$

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

$$=1-\alpha,$$

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

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

进而得到$\dfrac{\sigma_1}{\sigma_2}$的一个置信度为$1-\alpha$的置信区间

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

例4 甲、乙两台机床加工同一种零件, 在机床甲加工的零件中抽取9个样品, 在机床乙加工的零件中抽取6个样品,并分别测得它们的长度(单位:mm), 由所给数据算得 $s_1^2 = 0.245$, $s_2^2 = 0.357$,在置信度 0.98下, 试求这两台机床加工精度之比 σ_1/σ_2 的置信区间. 假定测量值都服从正态分布, 方差分别为 σ_1^2, σ_2^2.

解 $n_1 = 9, \quad n_2 = 6, \quad \alpha = 0.02,$

$$F_{1-\alpha/2}(n_1 - 1, n_2 - 1) = F_{0.99}(8, 5) = 10.3,$$

$$F_{\alpha/2}(8,5)=F_{0.01}(8,5)=\frac{1}{F_{0.99}(5,8)}=\frac{1}{6.63},$$

于是得 $\dfrac{\sigma_1}{\sigma_2}$ 的一个置信度为 0.98的置信区间

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

$$=\left(\sqrt{\frac{0.245}{0.357\times10.3}},\ \sqrt{\frac{0.245\times6.63}{0.357}}\right)=(0.258,\ 2.133).$$

小　结

1. 单个总体均值 μ 的置信区间

$$\begin{cases} (1)\ \sigma^2\text{为已知}, \left(\bar{X} \pm \dfrac{\sigma}{\sqrt{n}} z_{\alpha/2}\right). \\ (2)\ \sigma^2\text{为未知}, \left(\bar{X} \pm \dfrac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right). \end{cases}$$

2. 单个总体方差 σ^2 的置信区间

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

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

$\sigma_1^{\,2}$和$\sigma_2^{\,2}$均为未知, $\left(\bar{X}-\bar{Y}\pm z_{\alpha/2}\sqrt{\dfrac{S_1^{\,2}}{n_1}+\dfrac{S_2^{\,2}}{n_2}}\right).$

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

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

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

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