概率论与数理统计

庄玮玮
weizh@ustc.edu.cn

安徽 合肥

2020 年 4 月

第 五 章　参数估计

§5.6 参数的区间估计

在独立同分布场合，样本均值$\overline{X}_n$和样本方差S^2分别是总体均值μ和总体方差σ^2的无偏估计和强相合估计，说明样本均值和样本方差都是不错的估计量．它告诉我们，在n比较大的时候，真值μ就在$\overline{X}_n$附近，真值σ^2就在S^2附近．但是到底离真值有多近呢？n多大就够了呢？区间估计可以回答这一问题．

A. 已知σ时, μ的置信区间

设$Z \sim N(0,1)$, 对正数$\alpha \in (0,1)$, 有唯一的z_α使得

$$P(Z \geqslant z_\alpha) = \alpha.$$

这时称z_α为标准正态分布$N(0,1)$的上α分位数.

对于$\alpha = 0.05$和0.025, 查标准正态分布的上α分位数表C2得出(见图5.6.1)

$$z_{0.05} = 1.645, \quad z_{0.025} = 1.96.$$

于是有

$$\begin{aligned} P(Z \geqslant 1.645) &= 0.05, \ P(Z \leqslant 1.645) = 0.95, \\ P(Z \geqslant 1.96) &= 0.025, \ P(Z \leqslant 1.96) = 0.975. \end{aligned}$$

对于$\alpha \in (0,1)$, 利用标准正态概率密度的对称性得到(见图5.6.2)

$$
\begin{aligned}
P(\,|Z| \geqslant z_{\alpha/2}\,) &= P(\,Z \geqslant z_{\alpha/2}) + P(\,Z \leqslant -z_{\alpha/2}) \\
&= P(\,Z \geqslant z_{\alpha/2}) + P(\,-Z \geqslant z_{\alpha/2}) \\
&= \alpha/2 + \alpha/2 = \alpha.
\end{aligned}
$$

于是得到

$$P(\,|Z| \geqslant z_{\alpha/2}\,) = \alpha, \quad P(\,|Z| \leqslant z_{\alpha/2}\,) = 1 - \alpha.$$

取$\alpha = 0.05$和0.025时, 分别得到

$$P(|Z| \leqslant 1.96) = 0.95, \; P(|Z| \leqslant 1.645) = 0.90.$$

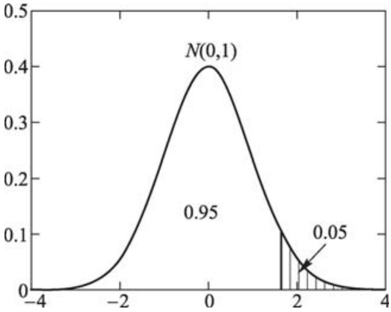

图5.6.1 $\alpha = 0.05$, $z_\alpha = 1.645$

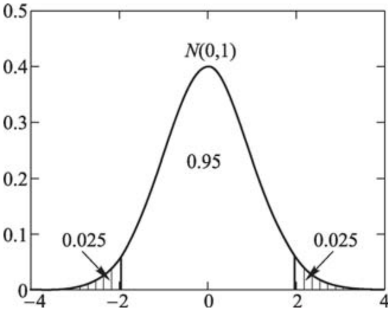

图5.6.2 $\alpha = 0.05$, $z_{\alpha/2} = 1.96$

例1.1 在导电材料中，铜的电阻率仅大于银．但是含杂质的铜的电阻率会增大．现在电缆厂对供应商提交的样品铜的电阻率进行了12次独立重复测量，测得了以下的电阻率(单位:$10^{-6}\Omega\cdot m$):

0.0173	0.0172	0.0145	0.0128	0.0177	0.0161
0.0156	0.0151	0.0138	0.0159	0.0140	0.0146

已知测量仪器的标准差是$\sigma=0.0012$，测量没有系统偏差(也就是说测量值X的数学期望等于样品的真实电阻率)，估计样品铜的电阻率．

解 用X_i表示第i次测量值, 对$n=12$, 容易计算出样本均值

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i=0.015\,4.$$

所以对该样品铜的电阻率的估计值是0.015 4.

问题好像解决了, 但是因为测量带有随机误差, 所以我们还不能确定该样品铜的真实电阻率为0.015 4. 只能确定其电阻率会落在一个含有$\overline{X}_n=0.015\,4$的小区间内, 下面计算这个区间.

解　用X_i表示第i次测量值，对$n=12$，容易计算出样本均值

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = 0.015\,4.$$

所以对该样品铜的电阻率的估计值是0.015 4.

问题好像解决了，但是因为测量带有随机误差，所以我们还不能确定该样品铜的真实电阻率为0.015 4. 只能确定其电阻率会落在一个含有$\overline{X}_n = 0.015\,4$的小区间内，下面计算这个区间.

用μ表示样品铜的实际电阻率. 用X表示测量值, 用ε表示测量误差, 则$X=\mu+\varepsilon$. 因为测量误差ε服从正态分布$N(0,\sigma^2)$, 所以$X\sim N(\mu,\sigma^2)$, 其中$\sigma=0.0012$. 现在$X_1,X_2,\cdots,X_n$是总体$N(\mu,\sigma^2)$的样本. 根据正态分布的性质知道

$$Z=\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\sim N(0,1). \tag{1.1}$$

于是得到

$$P\left(|\overline{X}_n-\mu|\leqslant\frac{1.96\sigma}{\sqrt{n}}\right)=P(|Z|\leqslant 1.96)=0.95,$$

即以0.95的概率保证

$$|\overline{X}_n - \mu| \leqslant \frac{1.96\sigma}{\sqrt{n}}.$$

也就是说我们以0.95的概率保证样品铜的电阻率

$$\mu \in \left[\overline{X}_n - \frac{1.96\sigma}{\sqrt{n}},\ \overline{X}_n + \frac{1.96\sigma}{\sqrt{n}}\right], \qquad (1.2)$$

将$\overline{X}_n = 0.0154$, $\sigma = 0.0012$, $n = 12$代入(1.2), 得到

$$\left[\overline{X}_n - \frac{1.96\sigma}{\sqrt{n}},\ \overline{X}_n + \frac{1.96\sigma}{\sqrt{n}}\right] = [0.0147,\ 0.0161]. \qquad (1.3)$$

所以, 我们以0.95的概率保证样品铜的电阻率 $\mu \in [0.0147, 0.0161]$.

在例1.1中，称[0.0147, 0.0161] 为μ的置信水平为0.95的置信区间．置信区间的长度是

$$0.0161 - 0.0147 = 0.0014.$$

在上面的例子中，称由(1.1)定义的Z为**枢轴量**，这是因为它的分布和未知参数无关．构造μ的置信区间时，本例中的枢轴量Z还起着中心的作用．从例1.1中容易看到以下的结论．

定理1.1 如果$X_1, X_2, \cdots, X_n$是总体$N(\mu, \sigma^2)$的样本，σ已知，则 μ 的置信水平为 $1-\alpha$ 的(双侧)置信区间是

$$\left[\overline{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}}, \ \overline{X}_n + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right]. \tag{1.4}$$

置信区间的长度是

$$2 \cdot \frac{z_{\alpha/2}\sigma}{\sqrt{n}}.$$

在例1.1中，如果要计算样品铜的电阻率μ的置信水平为$1-\alpha=0.90$的置信区间，只要取(6.4)中的 $z_{\alpha/2}=z_{0.05}=1.645$ 即可．这时的置信区间为

$$\left[\overline{X}_n-\frac{z_{\alpha/2}\sigma}{\sqrt{n}},\ \overline{X}_n+\frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right]=[0.014\,8,\ 0.016\,0]. \qquad (1.5)$$

置信区间的长度是

$$2\times 1.645\sigma/\sqrt{n}=0.001\,1.$$

可以看出，当置信水平从0.95降低到0.90，置信区间的长度从0.001 4减少为0.001 1．

可以看出，当置信水平从0.95降低到0.90，置信区间的长度从0.001 4减少为0.001 1.

从对于例1.1的分析，可以得到置信区间(1.4)的如下结论：

(1) 置信区间的中心是样本均值$\overline{X}_n$;

(2) 置信水平$1-\alpha$越高，则置信区间越长;

(3) 样本量n越大，则置信区间越短.

很明显，越长的置信区间提供的有用信息就越少．在例1.1中，不用计算就知道μ的置信水平为1的置信区间为$(0,\infty)$，但是这个置信区间没有任何有用的信息.

为了避免置信区间过长带来的不足，同时考虑置信水平不能太低，人们一般使用置信水平为 $1-\alpha=0.95$ 的置信区间．这时的

$$z_{\alpha/2}=z_{0.025}=1.96$$

值得牢记.

例1.2 为了得到鲜牛奶的冰点，对其冰点进行了21次独立重复测量，得到数据如下(单位:°C):

$$
\begin{array}{ccccccc}
-0.541 & -0.545 & -0.543 & -0.554 & -0.547 & -0.543 & -0.538 \\
-0.548 & -0.552 & -0.544 & -0.551 & -0.547 & -0.542 & -0.545 \\
-0.552 & -0.551 & -0.548 & -0.543 & -0.552 & -0.535 & -0.546
\end{array}
$$

已知测量的标准差是$\sigma = 0.004\,8$，测量没有系统偏差，计算鲜牛奶冰点的置信水平为0.95的置信区间，并计算置信区间的长度.

解 用μ表示鲜牛奶的冰点，用X表示测量值，则 $X \sim N(\mu, \sigma^2)$，其中 $\sigma = 0.004\,8$. 对于置信水平 $1-\alpha = 0.95$，有 $\alpha = 0.05$，$z_{\alpha/2} = 1.96$. 对样本量 $n = 21$，容易从测量数据计算出 $\overline{X}_n = -0.546$. 将以上数据代入公式(1.4)得到鲜牛奶冰点 μ 的置信水平为0.95的置信区间

$$\left[\overline{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}},\ \overline{X}_n + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right] = [-0.548\,1,\ -0.544\,0].$$

置信区间的长度为0.004 1.

B.未知σ时, μ的置信区间

在例1.1中, 如果σ 是未知数, 自然想到用样本标准差

$$S=\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_j-\hat{\mu})^2}$$

代替枢轴量(1.5)中的σ, 得到新的枢轴量

$$T_{n-1}=\frac{\overline{X}_n-\mu}{S/\sqrt{n}}. \tag{1.6}$$

T_{n-1} 服从$n-1$个自由度的t分布, 记做$T_{n-1}\sim t(n-1)$.

t分布的概率密度$p_m(t)$是偶函数，其形状和标准正态概率密度的形状相似，见图5.6.3．特别当$m \geqslant 35$时，$p_m(t)$可以用标准正态分布密度$\varphi(t)$近似．

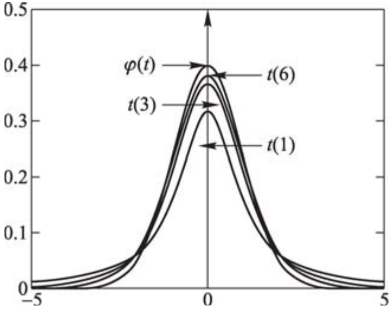

图5.6.3　$p_m(t)$的图形

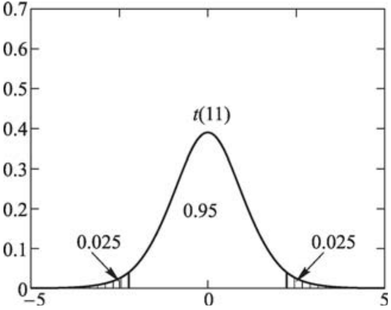

图5.6.4 $\alpha = 0.025$,
$t_{0.025}(11) = 2.201$

设$T_m \sim t(m)$，对正数$\alpha \in (0,1)$，有唯一的$t_\alpha(m)$使得

$$P(T_m \geqslant t_\alpha(m)) = \alpha.$$

这时称$t_\alpha(m)$为t分布的上α分位数.

对于$\alpha = 0.05$和0.025，查t分布的上α分位数表C3得出(参考图5.6.4)

m	11	12	13	14	15
$t_{0.025}(m)$	2.201	2.179	2.160	2.145	2.131
$t_{0.05}(m)$	1.796	1.782	1.771	1.761	1.753

m	16	17	18	19	20
$t_{0.025}(m)$	2.120	2.110	2.101	2.093	2.086
$t_{0.05}(m)$	1.746	1.740	1.734	1.729	1.725

对于$\alpha \in (0,1)$, 利用t分布的对称性得到(参考图5.6.4)

$$P(T_m \geqslant t) = P(T_m \leqslant -t),$$

于是有

$$\begin{aligned} P(|T_m| \geqslant t_{\alpha/2}(m)) &= P(T_m \geqslant t_{\alpha/2}(m)) + P(T_m \leqslant -t_{\alpha/2}(m)) \\ &= \alpha/2 + \alpha/2 = \alpha. \end{aligned}$$

这就得到

$$P(|T_m| \geqslant t_{\alpha/2}(m)) = \alpha, \quad P(|T_m| \leqslant t_{\alpha/2}(m)) = 1 - \alpha. \qquad (1.7)$$

定理1.2 设$X_1, X_2, \cdots, X_n$是正态总体$N(\mu, \sigma^2)$的样本， μ和 σ^2 未知,则μ的置信水平为$1-\alpha$的置信区间为

$$\left[\overline{X}_n - \frac{t_{\alpha/2}(n-1)S}{\sqrt{n}},\ \overline{X}_n + \frac{t_{\alpha/2}(n-1)S}{\sqrt{n}}\right]. \qquad (1.8)$$

证明 设T_{n-1}由(1.6)定义. 对于置信水平$1-\alpha$, 由 $T_{n-1} \sim t(n-1)$ 得到

$$P\left(\frac{|\overline{X}_n - \mu|}{S/\sqrt{n}} \leqslant t_{\alpha/2}(n-1)\right) = P\left(|T_{n-1}| \leqslant t_{\alpha/2}(n-1)\right) = 1-\alpha.$$

由于

$$
\begin{aligned}
& \left\{\frac{|\overline{X}_n-\mu|}{S/\sqrt{n}} \leqslant t_{\alpha/2}(n-1)\right\} \\
= \ & \left\{\overline{X}_n-\frac{t_{\alpha/2}(n-1)S}{\sqrt{n}} \leqslant \mu \leqslant \overline{X}_n+\frac{t_{\alpha/2}(n-1)S}{\sqrt{n}}\right\},
\end{aligned}
$$

所以在置信水平$1-\alpha$下, μ的置信区间是(1.8).

例1.3 在例1.1中, 假设标准差σ未知, 计算均值μ的置信水平为0.95的置信区间.

解 从例1.1中的数据可以计算出样本标准差$S = 0.0015.$, 由$1-\alpha = 0.95$得$\alpha/2 = 0.025$, 查表得到$t_{0.025}(11) = 2.201$. 将这些数和$n = 12$, $\overline{X}_n = 0.0154$代入置信区间(1.8), 得到所要的置信区间为 $[0.0144, 0.0164]$.

例1.3 在例1.1中, 假设标准差σ未知, 计算均值μ的置信水平为0.95的置信区间.

解 从例1.1中的数据可以计算出样本标准差$S = 0.0015.$, 由$1-\alpha = 0.95$得$\alpha/2 = 0.025$, 查表得到$t_{0.025}(11) = 2.201$. 将这些数和$n = 12$, $\overline{X}_n = 0.0154$代入置信区间(1.8), 得到所要的置信区间为 $[0.0144, 0.0164]$.

例1.4 在例1.2中，如果标准差σ未知，在置信水平0.95下，计算冰点μ的置信区间及置信区间的长度.

解 从例1.2中的数据可以计算出样本标准差$S = 0.005$，由$1-\alpha = 0.95$得$\alpha/2 = 0.025$，查表得到$t_{0.025}(20) = 2.086$. 将这些数和$n = 21$, $\overline{X}_n = -0.546$代入置信区间(1.8)，得到所要的置信区间为 $[-0.548\,3, -0.543\,7]$. 置信区间的长度是 $0.004\,6$.

应当注意置信区间(1.4)和(1.8)的特点：已知标准差σ时使用(1.4)；未知σ时用样本标准差S代替σ，将$z_{\alpha/2}$换成$t_{\alpha/2}(n-1)$ 即得到(1.8).

例1.4　在例1.2中，如果标准差σ未知，在置信水平0.95下，计算冰点μ的置信区间及置信区间的长度.

解　从例1.2中的数据可以计算出样本标准差$S = 0.005$，由$1 - \alpha = 0.95$得$\alpha/2 = 0.025$，查表得到$t_{0.025}(20) = 2.086$. 将这些数和$n = 21$, $\overline{X}_n = -0.546$代入置信区间(1.8)，得到所要的置信区间为 $[-0.548\,3, -0.543\,7]$. 置信区间的长度是 $0.004\,6$.

应当注意置信区间(1.4)和(1.8)的特点：已知标准差σ时使用(1.4)；未知σ时用样本标准差S代替σ，将$z_{\alpha/2}$换成$t_{\alpha/2}(n-1)$ 即得到(1.8).

C. 方差$\boldsymbol{\sigma}^2$的置信区间

设$X_1, X_2, \cdots, X_n$是正态总体$N(\mu, \sigma^2)$的样本,

$$S^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2$$

是样本方差. 枢轴量

$$\chi^2_{n-1} \equiv \frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 \qquad (1.9)$$

服从$n-1$个自由度的χ^2分布, 记做$\chi^2_{n-1} \sim \chi^2(n-1)$.

χ^2 分布的概率密度 $p_n(u)$的形状见图5.6.5.

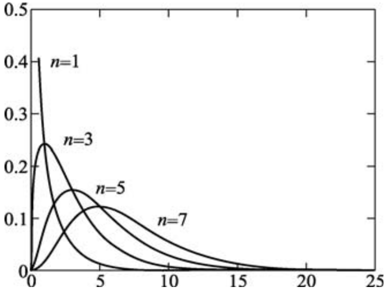

图5.6.5 $\chi^2(n)$的概率密度, $n = 1, 3, 5, 7$

设$\chi_m^2 \sim \chi^2(m)$, 对正数$\alpha \in (0,1)$, 有唯一的$\chi_\alpha^2(m)$使得(见图5.6.6)

$$P\big(\chi_m^2 \geqslant \chi_\alpha^2(m)\big) = \alpha.$$

这时称$\chi_\alpha^2(m)$ 为$\chi^2(m)$ 分布的上α分位数. 于是有

$$P\big(\chi_m^2 \leqslant \chi_\alpha^2(m)\big) = 1-\alpha,$$

及

$$P\big(\chi_m^2 \geqslant \chi_{\alpha/2}^2(m)\big) = \frac{\alpha}{2}, \quad P\big(\chi_m^2 \leqslant \chi_{\alpha/2}^2(m)\big) = 1-\frac{\alpha}{2}.$$

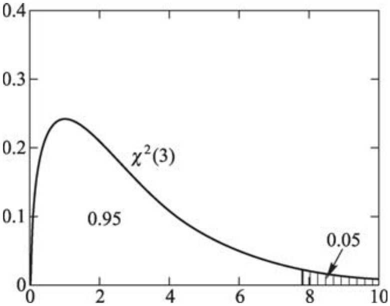

图5.6.6 $\chi^2_{0.05}(3) = 7.815$

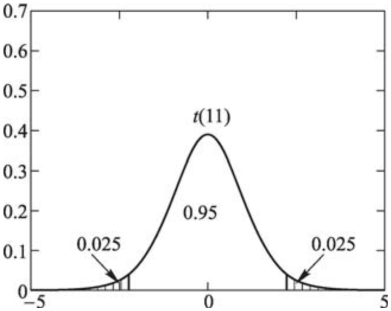

图5.6.7 $\chi^2_{0.975}(18) = 8.231$, $\chi^2_{0.025}(18) = 31.53$

对于$\alpha = 0.05$和0.025，查χ^2分布的上α分位数表得到(参考图5.6.7):

m	11	12	13	14	15
$\chi^2_{0.025}(m)$	21.92	23.34	24.74	26.12	27.49
$\chi^2_{0.975}(m)$	3.816	4.404	5.009	5.629	6.262

m	16	17	18	19
$\chi^2_{0.025}(m)$	28.85	30.19	31.53	32.85
$\chi^2_{0.975}(m)$	6.908	7.564	8.231	8.907

定理1.3 设$X_1, X_2, \cdots, X_n$是正态总体$N(\mu, \sigma^2)$的样本，μ和σ^2未知.则σ^2的置信水平为$1-\alpha$的置信区间为

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \quad \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right]. \tag{1.10}$$

证明 设χ^2_{n-1}由(1.9)定义. 直接计算得到

$$
\begin{aligned}
&P\Big(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\Big)\\
&= P\Big(\chi^2_{1-\alpha/2}(n-1) \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \chi^2_{\alpha/2}(n-1)\Big)\\
&= P\Big(\chi^2_{1-\alpha/2}(n-1) \leqslant \chi^2_{n-1} \leqslant \chi^2_{\alpha/2}(n-1)\Big)\\
&= P\Big(\chi^2_{n-1} \geqslant \chi^2_{1-\alpha/2}(n-1)\Big) - P\Big(\chi^2_{n-1} > \chi^2_{\alpha/2}(n-1)\Big)\\
&= 1 - \frac{\alpha}{2} - \frac{\alpha}{2}\\
&= 1 - \alpha.
\end{aligned}
$$

例1.5 地球生物的演变经历了漫长的岁月，只有化石为这一演变进行了记录. 现在科学家们利用物质的放射性衰变来研究生物的演变规律. 几乎所有的矿物质含有K(钾)元素及其同位素^{40}K(钾40). ^{40}K并不稳定，它可以缓慢地衰变成 ^{40}Ar(氩40)和 ^{40}Ca(钙40). 于是知道了^{40}K的衰变速率，就可以通过测量化石中的^{40}K和^{40}Ar的比例(钾氩比)估计化石的形成年代. 下面是根据钾氩比估算出的德国黑森林中发掘的19个化石样品的形成年龄(单位：百万年):

249 254 243 268 253 269 287 241 273

306 303 280 260 256 278 344 304 283 310

假设每个样品的估算年代都服从正态分布 $N(\mu, \sigma^2)$. 为评价钾氩比方法的估算精度，试完成以下工作.

(1) 计算σ^2的置信水平为0.95的置信区间;

(2) 计算标准差σ的置信水平为0.95的置信区间.

解　本例中$n-1=18$, $\alpha/2=0.025$, $1-\alpha/2=0.975$. 可以计算出

$$\overline{X}_n=276.9,\quad S^2=733.4.$$

查表得到

$$\chi^2_{0.025}(18)=31.53,\quad \chi^2_{0.975}(18)=8.231.$$

(1) 将上述数据代入(1.9), 得到σ^2的置信水平为0.95的置信区间

$$\left[\frac{18\times 733.4}{31.53},\ \frac{18\times 733.4}{8.231}\right]=[418.7,\ 1\,603.8].$$

(2) 由于 $\sigma\in[\sqrt{418.7},\ \sqrt{1\,603.8}]$ 和 $\sigma^2\in[418.7,\ 1\,603.8]$ 等价, 所以σ 的置信水平为 0.95 的置信区间是

$$[\sqrt{418.7},\ \sqrt{1\,603.8}]=[20.46,\ 40.05].$$

D. 单侧置信限

在例1.2中，已经得到鲜牛奶冰点的21次独立重复测量值的样本均值$\overline{X}_n = -0.546$. 已知测量的标准差为$\sigma = 0.0048$时，因为枢轴量

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

所以得到

$$P\left(\mu \geqslant \overline{X}_n - \frac{z_{0.05}\sigma}{\sqrt{n}}\right) = P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant z_{0.05}\right) = 0.95.$$

因为$z_{0.05}=1.645$, $n=21$, $\overline{X}_n-z_{0.05}\sigma/\sqrt{n}=-0.5477$, 所以我们以0.95的概率保证所测鲜牛奶的真实冰点$\mu\geqslant-0.5477$. 这时称$[-0.5477,\infty)$为μ的置信水平为0.95的单侧置信区间, 或称-0.5477为μ的置信水平为0.95的单侧置信下限.

因为

$$P\left(\mu\leqslant\overline{X}_n+\frac{z_{0.05}\sigma}{\sqrt{n}}\right)=P\left(\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\geqslant-z_{0.05}\right)=0.95,$$

$\overline{X}_n+z_{0.05}\sigma/\sqrt{n}=-0.5443$, 所以我们以0.95的概率保证所测鲜牛奶的真实冰点 $\mu\leqslant-0.5443$. 这时称$(-\infty,-0.5443]$为μ的置信水平为0.95的单侧置信区间, 或称-0.5443为μ的置信水平为0.95的单侧置信上限.

在例1.2中，如果标准差 σ 是未知的，则使用枢轴量

$$T_{20} = \frac{\overline{X}_n - \mu}{S/\sqrt{n}} \sim t(20)$$

得到

$$P\left(\mu \geqslant \overline{X}_n - \frac{t_{0.05}(20)S}{\sqrt{n}}\right) = P\left(\frac{\overline{X}_n - \mu}{S/\sqrt{n}} \leqslant t_{0.05}(20)\right) = 0.95.$$

因为$t_{0.05}(20)=1.725$, $S=0.005$, $\overline{X}_n-t_{0.05}(20)S/\sqrt{n}=-0.5479$, 所以我们以0.95的概率保证所测鲜牛奶的真实冰点$\mu\geqslant-0.5479$. 这时称$[-0.5479,\infty)$为μ的置信水平为0.95的单侧置信区间, 或称 -0.5479 为 μ的置信水平为 0.95 的单侧置信下限.

因为

$$P\left(\mu\leqslant\overline{X}_n+\frac{t_{0.05}(20)S}{\sqrt{n}}\right)=P\left(\frac{\overline{X}_n-\mu}{S/\sqrt{n}}\geqslant-t_{0.05}(20)\right)=0.95,$$

所以$\overline{X}_n+t_{0.05}(20)\sigma/\sqrt{n}=-0.5441$, 所以我们以0.95的概率保证所测鲜牛奶的真实冰点$\mu\leqslant-0.5441$. 这时称$(-\infty,-0.5441]$ 为μ的置信水平为0.95的单侧置信区间, 或称-0.5441为μ的置信水平为0.95的单侧置信上限.

在定理1.3中，因为枢轴量

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

所以

$$\begin{aligned} & P\left(\sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}\right) \\ = \quad & P\left(\frac{(n-1)S^2}{\sigma^2} \geqslant \chi^2_{1-\alpha}(n-1)\right) \\ = \quad & 1-\alpha. \end{aligned}$$

也就是说，我们以$1-\alpha$的概率保证

$$\sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}.$$

于是称

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

为σ^2的置信水平为$1-\alpha$的单侧置信区间，或称 $\dfrac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$ 为 σ^2 的置信水平为 $1-\alpha$ 的单侧置信上限.

相同的方法可以得到σ^2的置信水平为$1-\alpha$的单侧置信区间

$$\Big[\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)},\ \ \infty\Big),$$

从而得到σ^2的置信水平为$1-\alpha$的单侧置信下限$\dfrac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}$.

例1.6　对于例1.5中根据钾氩比估算出的德国黑森林中发掘的19个化石样品的形成年龄，计算测量方差σ^2的置信水平为0.95的单侧置信上限和单侧置信下限.

解　已经算得$S^2 = 733.4$，查表得到$\chi^2_{0.95}(18) = 9.390$，$\chi^2_{0.05}(18) = 28.869$. 分别计算出$\sigma^2$的单侧置信上限为

$$\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)} = 1405.9,$$

单侧置信下限为

$$\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)} = 457.28.$$

例1.6 对于例1.5中根据钾氩比估算出的德国黑森林中发掘的19个化石样品的形成年龄，计算测量方差σ^2的置信水平为0.95的单侧置信上限和单侧置信下限.

解 已经算得$S^2 = 733.4$，查表得到 $\chi^2_{0.95}(18) = 9.390$, $\chi^2_{0.05}(18) = 28.869$. 分别计算出$\sigma^2$的单侧置信上限为

$$\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)} = 1405.9,$$

单侧置信下限为

$$\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)} = 457.28.$$

现在把本节介绍的有关置信区间的内容作一小结:
设$X_1, X_2, \cdots, X_n$是总体$N(\mu, \sigma^2)$的样本.
(1) 当σ已知时, 在置信水平$1-\alpha$下, μ的(双侧)置信区间是

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

单侧置信上限为

$$\overline{X}_n + \frac{z_{\alpha}\sigma}{\sqrt{n}},$$

单侧置信下限为

$$\overline{X}_n - \frac{z_{\alpha}\sigma}{\sqrt{n}}.$$

(2) 当σ未知时, 在置信水平$1-\alpha$下, μ的(双侧)置信区间是

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

单侧置信上限为

$$\overline{X}_n + \frac{t_{\alpha}(n-1)S}{\sqrt{n}},$$

单侧置信下限为

$$\overline{X}_n - \frac{t_{\alpha}(n-1)S}{\sqrt{n}}.$$

(3) 在置信水平$1-\alpha$下, σ^2的(双侧)置信区间为

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

单侧置信上限为

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

单侧置信下限为

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

A. 均值差$\mu_1-\mu_2$的置信区间

称总体$X\sim N(\mu_1,\sigma_1^2)$ 和总体$Y\sim N(\mu_2,\sigma_2^2)$独立，意指这两个总体的样本独立. 也就是说，如果$X_1,X_2,\cdots,X_n$是总体$X$的样本，$Y_1,Y_2,\cdots,Y_m$是总体$Y$的样本，则

$$X_1,X_2,\cdots,X_n,Y_1,Y_2,\cdots,Y_m$$

相互独立. 对于上述样本，下面构造 $\mu_1-\mu_2$的 置信区间.

用$\overline{X}_n$ 和$\overline{Y}_m$分别表示$\{X_i\}$和$\{Y_j\}$的样本均值，用S_1^2 和S_2^2分别表示$\{X_i\}$和$\{Y_j\}$的样本方差，则

$$\overline{X}_n \sim N(\mu_1, \sigma_1^2/n), \quad \overline{Y}_m \sim N(\mu_2, \sigma_2^2/m),$$

从而

$$\overline{X}_n - \overline{Y}_m \sim N(\mu_1 - \mu_2, \sigma_1^2/n + \sigma_2^2/m).$$

于是得到

$$Z = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}} \sim N(0, 1). \tag{2.1}$$

(1) 已知σ_1^2, σ_2^2时, 对置信水平$1-\alpha$, 利用枢轴量(2.1)构造出$\mu_1-\mu_2$的置信区间

$$\left[(\overline{X}_n-\overline{Y}_m)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}},\ (\overline{X}_n-\overline{Y}_m)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\right]. \tag{2.2}$$

(2) 已知$\sigma_1^2=\sigma_2^2$, 但不知道σ_1^2,σ_2^2的具体值时, 利用$\mathrm{E}\,S_1^2=\mathrm{E}\,S_2^2=\sigma_2^2$, 可以验证

$$S_W^2=\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2} \tag{2.3}$$

是σ_1^2和σ_2^2的无偏估计: $\mathrm{E}\,S_W^2=\sigma_1^2=\sigma_2^2$.

用S_W^2代替(2.1)中的σ_1^2和σ_2^2, 得到新的枢轴量及其分布如下:

$$T = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{S_W\sqrt{1/n + 1/m}} \sim t(n+m-2). \tag{2.4}$$

利用(2.4)可以构造出$\mu_1 - \mu_2$的置信水平为$1-\alpha$的置信区间

$$\left[(\overline{X}_n - \overline{Y}_m) - t_{\alpha/2} S_W \sqrt{\frac{1}{n} + \frac{1}{m}},\ (\overline{X}_n - \overline{Y}_m) + t_{\alpha/2} S_W \sqrt{\frac{1}{n} + \frac{1}{m}}\right], \tag{2.5}$$

其中$t_{\alpha/2} = t_{\alpha/2}(n+m-2)$.

用S_W^2代替(2.1)中的σ_1^2和σ_2^2, 得到新的枢轴量及其分布如下:

$$T=\frac{(\overline{X}_n-\overline{Y}_m)-(\mu_1-\mu_2)}{S_W\sqrt{1/n+1/m}}\ \sim\ t(n+m-2). \tag{2.4}$$

利用(2.4)可以构造出$\mu_1-\mu_2$的置信水平为$1-\alpha$的置信区间

$$\left[(\overline{X}_n-\overline{Y}_m)-t_{\alpha/2}S_W\sqrt{\frac{1}{n}+\frac{1}{m}},\ (\overline{X}_n-\overline{Y}_m)+t_{\alpha/2}S_W\sqrt{\frac{1}{n}+\frac{1}{m}}\right], \tag{2.5}$$

其中$t_{\alpha/2}=t_{\alpha/2}(n+m-2)$.

例2.1 X, Y 两个渔场在春季放养相同的鲫鱼苗，但是使用不同的饵料饲养．三个月后，从X渔场打捞出16条鲫鱼，从Y渔场打捞出 14!条鲫鱼．分别秤出他们的平均质量和样本标准差如下(单位：kg)：

$$\overline{X}_n = 0.181,\ S_1 = 0.021,\ \ \overline{Y}_m = 0.185,\ S_2 = 0.020. \qquad (2.6)$$

假设X和Y渔场的鲫鱼质量分别服从正态分布$N(\mu_1, \sigma_1^2)$和$N(\mu_2, \sigma_2^2)$,且 $\sigma_1^2 = \sigma_2^2$，在置信水平0.95下，求$\mu_1 - \mu_2$的置信区间．

解 用X_i和Y_j分别表示X和Y渔场的第i条鱼和第j条鱼的质量，则$X_1, X_2, \cdots, X_{16}$ 和 $Y_1, Y_2, \cdots, Y_{14}$ 分别是总体$N(\mu_1, \sigma_1^2)$和 $N(\mu_2, \sigma_2^2)$ 的样本，且这两个总体独立．从置信水平$1-\alpha = 0.95$ 得 $\alpha/2 = 0.025$．对 $n = 16$，$m = 14$，$n + m - 2 = 28$，查表得到$t_{0.025}(28) = 2.048$．按公式(2.3)计算出

$$S_W = 0.020\,5.$$

将这些数连同(2.6)代入(2.5)，得到$\mu_1 - \mu_2$的置信水平为0.95 的置信区间

$$[-0.019\,4,\ 0.011\,4].$$

在实际问题中，还会遇到已知方差比例σ_1^2/σ_2^2，但是未知具体的方差值的情况，这时可以使用下面的置信区间(2.10)．

(3) 已知方差比$\sigma_1^2/\sigma_2^2 = b^2$, 但不知道$\sigma_1^2, \sigma_2^2$的具体值时, 从(2.1)得到

$$Z = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{b^2\sigma_2^2/n + \sigma_2^2/m}} \sim N(0,1). \qquad (2.7)$$

利用$\mathrm{E}\, S_1^2 = \sigma_1^2 = b^2\sigma_2^2$, $\mathrm{E}\, S_2^2 = \sigma_2^2$, 可以验证

$$S_b^2 = \frac{(n-1)S_1^2/b^2 + (m-1)S_2^2}{n+m-2} \qquad (2.8)$$

是σ_2^2的无偏估计: $\mathrm{E}\, S_b^2 = \sigma_2^2$. 用$S_b^2$代替(2.7)中的$\sigma_2^2$, 得到的枢轴量

$$T = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{S_b\sqrt{b^2/n + 1/m}} \sim t(n+m-2). \qquad (2.9)$$

利用(2.9)可以构造出$\mu_1-\mu_2$的置信水平为$1-\alpha$的置信区间

$$\left[(\overline{X}_n-\overline{Y}_m)-t_{\alpha/2}S_b\sqrt{\frac{b^2}{n}+\frac{1}{m}},\ (\overline{X}_n-\overline{Y}_m)+t_{\alpha/2}S_b\sqrt{\frac{b^2}{n}+\frac{1}{m}}\right], \tag{2.10}$$

其中$t_{\alpha/2}=t_{\alpha/2}(n+m-2)$.

例2.2 在例2.1中，假定$\sigma_1/\sigma_2 = 21/20$. 在置信水平0.95下，计算$\mu_1 - \mu_2$的置信区间.

解 这时有$b = 21/20$, 利用(2.8)计算出

$$S_b = 0.02.$$

将$\overline{X}_n = 0.181$, $\overline{Y}_m = 0.185$, $n = 16$, $m = 14$, $n + m - 2 = 28$, $t_{0.025}(28) = 2.048$, $S_b = 0.02$ 代入(2.10)，得到$\mu_1 - \mu_2$的置信水平为0.95的置信区间 $[-0.019, 0.011]$.

例2.2 在例2.1中，假定$\sigma_1/\sigma_2 = 21/20$. 在置信水平0.95下，计算$\mu_1 - \mu_2$的置信区间.

解 这时有$b = 21/20$, 利用(2.8)计算出

$$S_b = 0.02.$$

将$\overline{X}_n = 0.181$, $\overline{Y}_m = 0.185$, $n = 16$, $m = 14$, $n + m - 2 = 28$, $t_{0.025}(28) = 2.048$, $S_b = 0.02$ 代入(2.10), 得到$\mu_1 - \mu_2$的置信水平为0.95的置信区间 $[-0.019,\ 0.011]$.

B. 方差比σ_1^2/σ_2^2的置信区间

设总体$X \sim N(\mu_1, \sigma_1^2)$ 和总体$Y \sim N(\mu_2, \sigma_2^2)$独立， $X_1, X_2, \cdots,$ X_n是总体X的样本，$Y_1, Y_2, \cdots, Y_m$ 是总体Y的样本. 可以计算出枢轴量

$$F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}. \qquad (2.11)$$

F服从自由度为$n-1$和$m-1$的F分布，记做 $F \sim F(n-1, m-1)$.

F分布概率密度的形状见图5.6.8. 设$F \sim F(n-1, m-1)$, 对正数$\alpha \in (0,1)$, 有唯一的$F_\alpha(n-1, m-1)$使得

$$P(F > F_\alpha(n-1, m-1)) = \alpha,$$

这时称$F_\alpha(n-1, m-1)$为$F(n-1, m-1)$分布的上α分位数. 查表时, 对于 $\alpha > 0.5$, 需要用下面的公式进行换算:

$$F_\alpha(n, m) = \frac{1}{F_{1-\alpha}(m, n)}, \quad n, m \geqslant 1. \qquad (2.12)$$

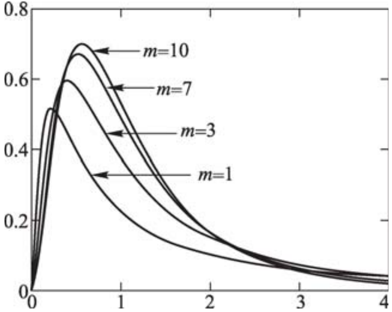

图5.6.8 $F(6, m)$的概率密度, $m = 1, 3, 7, 10$

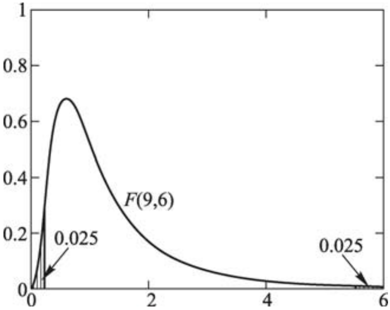

图5.6.9 $F_{0.975}(9, 6) = 0.23$, $F_{0.025}(9, 6) = 5.52$

利用枢轴量(2.11)可以得到σ_1^2/σ_2^2的置信水平为$1-\alpha$的置信区间

$$\left[\frac{S_1^2/S_2^2}{F_{\alpha/2}}, \frac{S_1^2/S_2^2}{F_{1-\alpha/2}}\right]. \qquad (2.13)$$

这是因为有

$$\begin{aligned} & P\left(\frac{S_1^2/S_2^2}{F_{\alpha/2}} \leqslant \sigma_1^2/\sigma_2^2 \leqslant \frac{S_1^2/S_2^2}{F_{1-\alpha/2}}\right) \\ = \quad & P\left(F_{1-\alpha/2} \leqslant \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \leqslant F_{\alpha/2}\right) \\ = \quad & P(F \geqslant F_{1-\alpha/2}) - P(F > F_{\alpha/2}) \\ = \quad & 1-\alpha/2-\alpha/2 = 1-\alpha. \end{aligned}$$

例2.3 在例2.1中，计算方差比σ_1^2/σ_2^2和标准差之比σ_1/σ_2的置信水平为0.95的置信区间.

解 对$n=16$, $m=14$, 查表得到$F_{0.025}(15,13)=3.05$,

$$F_{0.975}(15,13)=1/F_{0.025}(13,15)=1/2.92=0.34,$$

连同$S_1=0.021$, $S_2=0.020$ 代入(2.13), 得到σ_1^2/σ_2^2的置信水平为0.95的置信区间[0.361, 3.243].

因为$\sigma_1^2/\sigma_2^2\in[0.361,\ 3.243]$ 和$\sigma_1/\sigma_2\in[\sqrt{0.361},\ \sqrt{3.243}]$ 等价, 所以σ_1/σ_2的置信水平为0.95的置信区间是

$$[\sqrt{0.361},\ \sqrt{3.243}]=[0.601,\ 1.801].$$

例2.3 在例2.1中，计算方差比σ_1^2/σ_2^2和标准差之比σ_1/σ_2的置信水平为0.95的置信区间.

解 对$n = 16$, $m = 14$, 查表得到$F_{0.025}(15, 13) = 3.05$,

$$F_{0.975}(15, 13) = 1/F_{0.025}(13, 15) = 1/2.92 = 0.34,$$

连同$S_1 = 0.021$, $S_2 = 0.020$ 代入(2.13)，得到σ_1^2/σ_2^2的置信水平为0.95的置信区间[0.361, 3.243].

因为$\sigma_1^2/\sigma_2^2 \in [0.361,\ 3.243]$ 和$\sigma_1/\sigma_2 \in [\sqrt{0.361},\ \sqrt{3.243}]$ 等价，所以σ_1/σ_2的置信水平为0.95的置信区间是

$$[\sqrt{0.361},\ \sqrt{3.243}] = [0.601,\ 1.801].$$

A. 正态逼近法

在实际问题中，常常遇到总体分布不是正态分布的情况，这时也需要对均值和方差计算置信区间. 设$X_1, X_2, \cdots, X_n$是总体X的样本，$\mu = EX$，$\sigma^2 = \mathrm{Var}(X)$分别是总体均值和总体方差. 根据中心极限定理，对较大的样本量n，

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$$

近似服从标准正态分布$N(0,1)$. 于是对较大的n，有

$$P\left(|\overline{X}_n - \mu| \leqslant \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right) \approx 1 - \alpha. \qquad (3.1)$$

于是，已知标准差σ时，在置信水平$1-\alpha$下，总体均值μ的近似置信区间仍然是

$$\left[\overline{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}},\ \overline{X}_n + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right].$$

当σ未知时，对较大的n，S是σ的强相合估计，可用S代替σ，有

$$\frac{\overline{X}_n - \mu}{S/\sqrt{n}}$$

近似服从标准正态分布$N(0,1)$，则未知标准差σ时，均值μ的置信水平为$1-\alpha$的近似置信区间是

$$\left[\overline{X}_n - \frac{z_{\alpha/2}S}{\sqrt{n}},\ \overline{X}_n + \frac{z_{\alpha/2}S}{\sqrt{n}}\right]. \tag{3.2}$$

需要注意的是，n越大，近似的程度越好．以上方法被称为正态逼近法．使用正态逼近法时，一般的要求是$n \geqslant 30$.

例3.1 人们一直在研究年龄和血液中的各种成份之间的关系. 现在通过随机抽样调查了30个30岁健康公民的血小板数. 数据如下(单位: 万/mm^3):

26	19	18	16	26	17	20	20	19	22	19	12	29	15	22
19	27	25	28	24	35	28	19	23	31	30	23	30	17	22

用μ表示30岁健康公民的血小板数的总体均值. 对于置信水平$1-\alpha=0.95$, 计算μ的置信区间.

解 可以认为被选到的个体的血小板数是独立同分布的. 经过计算得到$\overline{X}_n = 22.7$, $S = 5.45$. 代入(3.2)得到置信水平0.95下, μ的(近似)置信区间

$$\begin{aligned} & \left[22.7 - 1.96 \times 5.45/\sqrt{30},\ 22.7 + 1.96 \times 5.45/\sqrt{30}\right] \\ = \quad & [20.75,\ 24.65]. \end{aligned}$$

B. 比例p的置信区间

例3.2 设$X_1, X_2, \cdots, X_n$是两点分布$\mathcal{B}(1,p)$的样本, $\hat{p} = \overline{X}_n$是 p 的最大似然估计. 对置信水平 $1-\alpha$, 当 n 较大(至少使得$5 \leqslant n\hat{p} \leqslant n-5$), p的(近似)置信区间是

$$\left[\frac{b-\sqrt{b^2-4ac}}{2a}, \quad \frac{b+\sqrt{b^2-4ac}}{2a}\right], \tag{3.3}$$

其中

$$a = 1 + \frac{z_{\alpha/2}^2}{n}, \quad b = 2\overline{X}_n + \frac{z_{\alpha/2}^2}{n}, \quad c = \overline{X}_n^2. \tag{3.4}$$

*证明 设总体$X \sim \mathcal{B}(1, p)$, $p \in (0, 1)$, 则$p = \mathrm{E}X$, $\sigma^2 = \mathrm{Var}\,(X) = p(1-p)$, $\mathrm{E}\overline{X}_n = p$, $\mathrm{Var}\,(\overline{X}_n) = p(1-p)/n$. 对较大的样本量$n$, 由中心极限定理知道

$$\frac{\overline{X}_n - p}{\sqrt{p(1-p)/n}}$$

近似服从标准正态分布$N(0, 1)$. 于是

$$P\left(\left|\frac{\overline{X}_n - p}{\sqrt{p(1-p)/n}}\right| \leqslant z_{\alpha/2}\right) \approx 1 - \alpha.$$

将

$$\left|\frac{\overline{X}_n - p}{\sqrt{p(1-p)/n}}\right| \leqslant z_{\alpha/2} \qquad (3.5)$$

两边平方，整理后得到

$$\begin{aligned} g(p) &\equiv \left(1 + z_{\alpha/2}^2/n\right)p^2 - \left(2\overline{X}_n + z_{\alpha/2}^2/n\right)p + \overline{X}_n^2 \\ &= ap^2 - bp + c \\ &\leqslant 0. \qquad (3.6) \end{aligned}$$

$y = g(p)$是一条开口向上的抛物线，和横坐标的交点$\hat{p}_1, \hat{p}_2$分别是

$$\hat{p}_1 = \frac{b - \sqrt{b^2 - 4ac}}{2a}, \quad \hat{p}_2 = \frac{b + \sqrt{b^2 - 4ac}}{2a}. \qquad (3.7)$$

从

$$\{p \in [\hat{p}_1, \hat{p}_2]\} = \{g(p) \leqslant 0\} = \left\{ \left| \frac{\overline{X}_n - p}{\sqrt{p(1-p)/n}} \right| \leqslant z_{\alpha/2} \right\}$$

得到

$$P(\hat{p}_1 \leqslant p \leqslant \hat{p}_2) \approx 1 - \alpha.$$

所以置信水平为$1-\alpha$时, p的(近似)置信区间是(3.3).

因为在实际问题中经常遇到计算比例p的置信区间的问题，所以有必要通过随机模拟对置信区间(3.3)加深理解. 给定比例p的具体值后，用计算机产生$n = 200$个服从两点分布$\mathcal{B}(1, p)$的随机数，然后取置信水平0.95，计算置信区间(3.3)及其长度L. 将以上试验独立重复10 000次，用cp表示真值p落在置信区间(3.3)中的频率，用$\hat{L}$表示上述10 000个置信区间(3.3)的平均长度. 计算结果如下:

p	0.5	0.3	0.15	0.1
cp	0.9449	0.9480	0.9424	0.9537
$\hat{L}$	0.1369	0.1257	0.0984	0.0832

p	0.08	0.06	0.05	0.04
cp	0.9474	0.9496	0.9563	0.9536
$\hat{L}$	0.0756	0.0666	0.0614	0.0557

通常把cp称为置信区间(3.3)的模拟覆盖率. 从以上计算结果看出, 总体上讲, $n = 200$ 时的模拟覆盖率和理论覆盖率0.95差别不大. 而当p由0.5减少到0.04时, 置信区间的平均长度$\hat{L}$也随之减少.

下面将样本量减少为$n=80$, 重复上述试验. 对于使得$np \geqslant 5$的p, 得到:

p	0.5	0.3	0.15	0.1	0.08
cp	0.942 9	0.961 6	0.958 5	0.962 7	0.960 6
$\hat{L}$	0.212 8	0.195 7	0.154 7	0.131 7	0.120 3

这次的模拟覆盖率也和理论覆盖率差别不大, 但是置信区间的平均长度比$n=200$时的平均长度明显增加.

C. 样本量的确定

在许多实际问题中，人们往往需要事先了解n取多大比较合适．我们知道$\overline{X}_n$是p的估计，n越大，$\mathrm{Var}\,(\overline{X}_n) = \mathrm{E}(\overline{X}_n - p)^2 = p(1-p)/n$ 越小，估计越精确．但是增加 n，往往会增加抽样的成本，于是给定置信水平$1-\alpha$和置信区间的长度d后，需要找到较小的n使得置信区间的长度小于等于d．下面的例3.3回答了这个问题．注意d越小，置信区间越精确．

例3.3 给定置信水平$1-\alpha$，要使得置信区间(3.3)的长度不超过d，只要取样本量

$$n \geqslant \left(\frac{z_{\alpha/2}}{d}\right)^2. \qquad (3.8)$$

C. 样本量的确定

在许多实际问题中，人们往往需要事先了解n取多大比较合适. 我们知道$\overline{X}_n$是p的估计，n越大，$\mathrm{Var}\,(\overline{X}_n) = \mathrm{E}(\overline{X}_n - p)^2 = p(1-p)/n$ 越小，估计越精确. 但是增加 n，往往会增加抽样的成本，于是给定置信水平$1-\alpha$和置信区间的长度d后，需要找到较小的n使得置信区间的长度小于等于d. 下面的例3.3回答了这个问题. 注意d越小，置信区间越精确.

例3.3 给定置信水平$1-\alpha$，要使得置信区间(3.3)的长度不超过d，只要取样本量

$$n \geqslant \left(\frac{z_{\alpha/2}}{d}\right)^2. \qquad (3.8)$$

*证明 当(3.8)成立，有

$$\frac{z_{\alpha/2}^2}{n} \leqslant d^2. \qquad (3.9)$$

因为函数$h(x)=4(x-x^2)$在区间$[0,1]$中的最大值是1，所以从$\overline{X}_n \in [0,1]$得到

$$h(\overline{X}_n) = 4(\overline{X}_n - \overline{X}_n^2) \leqslant 1. \qquad (3.10)$$

利用

$$a = 1 + \frac{z_{\alpha/2}^2}{n}, \quad b = 2\overline{X}_n + \frac{z_{\alpha/2}^2}{n}, \quad c = \overline{X}_n^2,$$

得到

$$
\begin{aligned}
b^2-4ac &= 4\overline{X}_n^2+4\overline{X}_n\frac{z_{\alpha/2}^2}{n}+\frac{z_{\alpha/2}^4}{n^2}-4\left(\overline{X}_n^2+\overline{X}_n^2\frac{z_{\alpha/2}^2}{n}\right)\\
&= 4(\overline{X}_n-\overline{X}_n^2)\frac{z_{\alpha/2}^2}{n}+\frac{z_{\alpha/2}^4}{n^2}\\
&\leqslant \frac{z_{\alpha/2}^2}{n}+\frac{z_{\alpha/2}^4}{n^2} \qquad [\text{用}(3.10)]\\
&= \left(1+\frac{z_{\alpha/2}^2}{n}\right)\frac{z_{\alpha/2}^2}{n}\\
&= a\frac{z_{\alpha/2}^2}{n}\\
&\leqslant ad^2. \qquad [\text{用}(3.9)]
\end{aligned}
$$

于是从$a>1$知道置信区间(3.3)的长度

$$\frac{\sqrt{b^2-4ac}}{a} \leqslant \frac{\sqrt{ad^2}}{a} = \frac{\sqrt{d^2}}{\sqrt{a}} < d.$$

取$1-\alpha=0.95$时，可以列出和d对应的n如下.

d	0.14	0.12	0.10	0.08	0.06	0.04	0.02	0.01
n	196	267	385	601	1 068	2 401	9 604	38 416

(3.11)

例3.4 2009年3月, 有政协委员建议逐步恢复繁体字的提案, 引发了广泛关注和争议. 为广泛了解民意, 需要对于该提案的总体支持率p进行估计. 试解决以下问题:

(a) 为得到p的置信水平为0.95的置信区间, 且置信区间长度不超过0.01, 应当随机抽样调查多少人?

(b) 如果随机抽样调查的4万人中有5 600人支持该提案, 计算p的置信水平为0.95的置信区间;

(c) 计算(b)中置信区间的长度.

解 (a) 本例中$d = 0.01$, $\alpha = 0.05$, $z_{\alpha/2} = 1.96$. 从(3.8)计算出

$$n \geqslant \frac{1.96^2}{0.01^2} = 38\,416,$$

所以至少应当调查38 416个人.

例3.4 2009年3月，有政协委员建议逐步恢复繁体字的提案，引发了广泛关注和争议. 为广泛了解民意，需要对于该提案的总体支持率p进行估计. 试解决以下问题:

(a) 为得到p的置信水平为0.95的置信区间，且置信区间长度不超过0.01，应当随机抽样调查多少人?

(b) 如果随机抽样调查的4万人中有5 600人支持该提案，计算p的置信水平为0.95的置信区间;

(c) 计算(b)中置信区间的长度.

解 (a) 本例中$d = 0.01$, $\alpha = 0.05$, $z_{\alpha/2} = 1.96$. 从(3.8)计算出

$$n \geqslant \frac{1.96^2}{0.01^2} = 38\,416,$$

所以至少应当调查38 416个人.

(b) 当4万个人中有5 600个人同意支持该提案时，可以计算出

$$\overline{X}_n = \frac{5\,600}{40\,000} = 0.14, \quad n = 40\,000.$$

因为$5 < n\overline{X}_n < 40\,000 - 5$，所以可以用正态逼近法. 将

$$a = 1 + \frac{1.96^2}{40\,000}, \quad b = 2 \times 0.14 + \frac{1.96^2}{40\,000}, \quad c = 0.14^2$$

代入置信区间(3.3)，得到总体支持率p的置信区间为

$$[0.136\,6,\ 0.143\,4].$$

即以95%的把握保证，对该提案的支持率在$13.66\% \sim 14.34\%$之间.

(c) 置信区间的长度为$0.143\,4 - 0.136\,6 = 0.006\,8$.

根据前面的讨论，可以给出一般情况下未知参数θ的置信区间的定义.

定义4.1 设$X_1, X_2, \cdots, X_n$是总体X的样本，$\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$，$\theta$是未知参数，$\hat{\theta}_1 = \hat{\theta}_1(\boldsymbol{X})$，$\hat{\theta}_2 = \hat{\theta}_2(\boldsymbol{X})$是两个统计量. 对于给定的$\alpha \in (0,1)$，如果有

$$P(\,\hat{\theta}_1 \leqslant \theta \leqslant \hat{\theta}_2\,) \geqslant 1 - \alpha, \tag{4.1}$$

则称$[\hat{\theta}_1, \hat{\theta}_2]$为参数$\theta$的置信水平为$1-\alpha$的置信区间(confidence interval).

在定义4.1中，置信水平又称为置信度，置信区间的右端点$\hat{\theta}_2$又称为置信上限，置信区间的左端点$\hat{\theta}_1$又称为置信下限．由于$\hat{\theta}_1=\hat{\theta}_1(\boldsymbol{X})$和$\hat{\theta}_2=\hat{\theta}_2(\boldsymbol{X})$都是随机变量的函数，因而是随机变量．但是给定样本观测值$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$，就得到了一个具体的闭区间$[\hat{\theta}_1(\boldsymbol{x}),\hat{\theta}_2(\boldsymbol{x})]$，我们以$1-\alpha$的概率保证未知参数$\theta\in[\hat{\theta}_1(\boldsymbol{x}),\hat{\theta}_2(\boldsymbol{x})]$．很明显，在相同的置信水平下，置信区间的长度越小越好．

定义4.2 设$X_1, X_2, \cdots, X_n$是总体X的样本, $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$, θ是未知参数, $\overline{\theta} = \overline{\theta}(\boldsymbol{X})$, $\underline{\theta} = \underline{\theta}(\boldsymbol{X})$ 是两个统计量. 对于给定的 $\alpha \in (0, 1)$.

(1) 如果

$$P(\theta \leqslant \overline{\theta}) \geqslant 1 - \alpha, \tag{4.2}$$

则称$\overline{\theta}$为参数θ的置信水平为$1-\alpha$的**单侧置信上限**;

(2) 如果

$$P(\theta \geqslant \underline{\theta}) \geqslant 1 - \alpha, \tag{4.3}$$

则称$\underline{\theta}$为参数θ的置信水平为$1-\alpha$的**单侧置信下限**.

很明显,(4.2), (4.3)中的右边不等号$\geqslant$越接近等号, 相应的置信区间越好.