第三节 参数的区间估计

- 一、区间估计的基本概念
- 二、正态总体均值与方差的区间估计
- 三、内容小结

一、区间估计基本概念

1. 问题的提出

点估计法：找一个统计量$\hat{\theta}$来估计参数θ,

$$\theta \approx \hat{\theta}_0 \quad (\hat{\theta}\text{的观察值})$$

不足之处：未给出估计值量$\hat{\theta}_0$与参数真值θ的误差及估计的可靠程度.

例1 $E(X) = \mu$,

$$\mu \approx \hat{\mu} = \overline{x}$$ —样本均值的观察值.

问： 误差 $|\bar{x}-\mu|$ 多大？

用 $\bar{X}$ 估计μ的可靠程度如何？

即 给定α： $\underline{0<\alpha<<1}$ 很小

使 $P\{|\bar{X}-\mu|<\varepsilon\}=\underline{1-\alpha}$ 较大

中的$\varepsilon=?$

区间估计解决了上述问题，从而克服了点估计的不足之处.

2. 置信区间与置信度

定义6.7

设总体 X 的分布函数 $F(x;\theta)$ 含有一个未知参数 θ, 对于给定值 α $(0<\alpha<1)$, 若由样本 $X_1, X_2, \cdots, X_n$ 确定的两个统计量

$\hat{\theta}_1=\hat{\theta}_1(X_1, X_2, \cdots, X_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(X_1, X_2, \cdots, X_n)$ 满足

$$P\{\hat{\theta}_1<\theta<\hat{\theta}_2\}=1-\alpha,$$

则称随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 是 θ 的置信度为 $1-\alpha$ 的置信区间, $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 分别称为置信度为 $1-\alpha$ 的双侧置信区间的置信下限和置信上限, $1-\alpha$ 为置信度.

关于定义的说明

被估计的参数θ虽然未知, 但它是一个常数, 没有随机性, 而区间$[\hat{\theta}_1, \hat{\theta}_2]$是随机的.

因此定义中下述表达式

$$P\{\hat{\theta}_1 \le \theta \le \hat{\theta}_2\} = 1 - \alpha$$

的本质是:

随机区间$[\hat{\theta}_1, \hat{\theta}_2]$以$1-\alpha$的概率包含着参数$\theta$的真值, 而不能说参数$\theta$以$1-\alpha$的概率落入随机区间$[\hat{\theta}_1, \hat{\theta}_2]$.

由定义可见，

对参数θ作区间估计，就是要设法找出两个只依赖于样本的界限(构造统计量)

$$\hat{\theta}_1 = \hat{\theta}_1(X_1,\ldots X_n)$$
$$\hat{\theta}_2 = \hat{\theta}_2(X_1,\ldots X_n)$$

$$(\hat{\theta}_1 < \hat{\theta}_2)$$

一旦有了样本，就把θ估计在区间$[\hat{\theta}_1, \hat{\theta}_2]$内. 这里有两个要求:

1. 要求 θ 以很大的可能被包含在区间$[\hat{\theta}_1,\hat{\theta}_2]$内，就是说，概率$P\{\hat{\theta}_1 \le \theta \le \hat{\theta}_2\}$要尽可能大. 即要求估计尽量可靠.

2. 估计的精度要尽可能的高. 如要求区间长度 $\hat{\theta}_2 - \hat{\theta}_1$ 尽可能短，或能体现该要求的其它准则.

可靠度与精度是一对矛盾，一般是在保证可靠度的条件下尽可能提高精度.

3. 求置信区间的一般步骤(共3步)

1° 寻求一个样本 $X_1, X_2, \cdots, X_n$ 的函数:

$$Z = Z(X_1, X_2, \cdots, X_n; \theta)$$

其中仅包含待估参数θ, 并且Z的分布已知且不依赖于任何未知参数(包括θ).

2° 对于给定的置信度 $1-\alpha$, 求出两个常数 a, b, 使

$$P\{a \le Z(X_1, X_2, \cdots, X_n; \theta) \le b\} = 1-\alpha.$$

3° 解出 θ

$$a \le Z(X_1, X_2, \cdots, X_n; \theta) \le b$$

$$\Leftrightarrow \quad \hat{\theta}_1 \le \theta \le \hat{\theta}_2,$$

其中$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \cdots, X_n)$, $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \cdots, X_n)$

那么$[\hat{\theta}_1, \hat{\theta}_2]$就是$\theta$的一个置信度为$1-\alpha$的置信区间.

二、正态总体均值与方差的区间估计

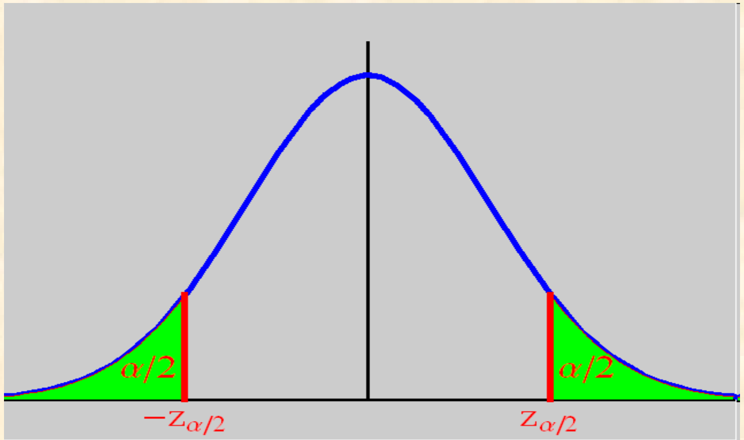
I. 单个总体 $N(\mu,\sigma^2)$ 的情况

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

1. 均值 μ 的置信区间

(1) σ^2为已知, 求 μ 的置信区间

1° 取 $U=\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$,



2° 给定α，要求：

$$P\{\,|U|\le u_{\alpha/2}\,\}=1-\alpha,$$

$$\Phi(u_{\frac{\alpha}{2}})-\Phi(-u_{\frac{\alpha}{2}})=1-\alpha,$$

$$2\Phi(u_{\frac{\alpha}{2}})-1=1-\alpha,\quad \Phi(u_{\frac{\alpha}{2}})=1-\frac{\alpha}{2},$$

3° 查表可确定 $u_{\frac{\alpha}{2}}$.

4° 作等价变形 解出 μ

$$|U| \le u_{\frac{\alpha}{2}} \Leftrightarrow \left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right| \le u_{\frac{\alpha}{2}} \Leftrightarrow |\overline{X}-\mu| \le \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}$$

$$\Leftrightarrow \overline{X} - \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}} \le \mu \le \overline{X} + \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}$$

$\therefore$ μ的一个置信度为$1-\alpha$的置信区间

$[\overline{X} - \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}},\ \overline{X} + \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}]$, 简写成 $\left(\overline{X} \pm \frac{\sigma}{\sqrt{n}} u_{\alpha/2}\right)$.

其置信区间的长度为 $2 \times \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$.

例2 包糖机某日开工包了12 包糖,称得重量(单位:克)分别为506, 500, 495, 488, 504, 486, 505, 513, 521, 520, 512, 485. 假设重量服从正态分布,且标准差为 $\sigma=10$, 试求糖包的平均重量 μ 的 $1-\alpha$ 置信区间 (取 $\alpha=0.10$).

解 $\sigma=10,\quad n=12,$

计算得 $\bar{x}=502.92,$

(1) 当 $\alpha=0.10$ 时, $\quad 1-\dfrac{\alpha}{2}=0.95,$

查表得 $u_{\alpha/2}=u_{0.05}=1.645,$

$$\bar{x}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}=502.92-\frac{10}{\sqrt{12}}\times 1.645=498.17,$$

$$\bar{x}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}=502.92+\frac{10}{\sqrt{12}}\times 1.645=507.67,$$

即μ的置信度为90%的置信区间为

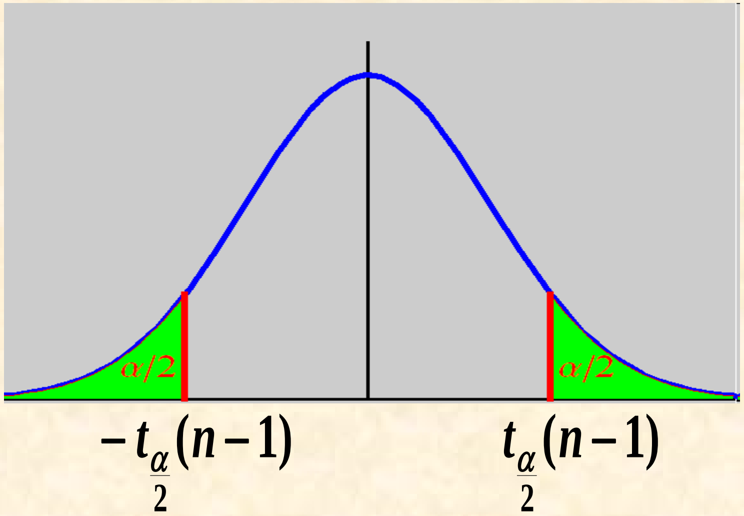
$$[498.17,\ 507.67].$$

(2) σ^2为未知, 求 μ 的置信区间

1° 取 $T = \dfrac{\bar{X} - \mu}{S_n^* / \sqrt{n}} \sim t(n-1),$

2° 给定α，要求：

$$P\left\{ |T| \le t_{\frac{\alpha}{2}}(n-1) \right\} = 1 - \alpha,$$



3° 查表可确定 $t_{\frac{\alpha}{2}}(n-1)$.

4° 作等价变形（解出 μ）

$$|T| \le t_{\frac{\alpha}{2}}(n-1) \Leftrightarrow \left| \frac{\bar{X} - \mu}{S_n^* / \sqrt{n}} \right| \le t_{\frac{\alpha}{2}}(n-1)$$

$$\Leftrightarrow \quad \left|\frac{\overline{X}-\mu}{S_n^*/\sqrt{n}}\right| \le t_{\frac{\alpha}{2}}(n-1)$$

$$\Leftrightarrow \quad \overline{X}-\frac{S_n^*}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1) \le \mu \le \overline{X}+\frac{S_n^*}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)$$

μ的置信度为$1-\alpha$的置信区间

$$[\overline{X}-\frac{S_n^*}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\ \overline{X}+\frac{S_n^*}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)],$$

简写成 $$\left(\overline{X}\pm\frac{S_n^*}{\sqrt{n}}t_{\alpha/2}(n-1)\right).$$

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, 试求总体均值 μ 的置信度为 0.95 的置信区间.

解 $\alpha = 0.05,\ \ n-1=15,$

查 $t(n-1)$ 分布表可知: $t_{\frac{\alpha}{2}}(15) = t_{0.025}(15) = 2.1315,$

计算得 $\bar{x} = 503.75,\quad s_n^* = 6.2022,$

得μ的置信度为95%的置信区间

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

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 $\dfrac{6.2022}{\sqrt{16}} \times 2.1315 \times 2 = 6.61$(克).

这个误差的可信度为95%.

2. 方差σ^2的置信区间

根据实际需要, 只介绍μ未知的情况.

方差σ^2的置信度为$1-\alpha$的置信区间

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

推导过程如下:

因为S_n^{*2}是σ^2的无偏估计,

根据第五章第三节定理知 $\dfrac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$, a　　　　　　　　　　　　b

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

解σ^2

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

于是得方差σ^2的置信度为$1-\alpha$的置信区间

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

进一步可得:

标准差σ的一个置信度为$1-\alpha$的置信区间

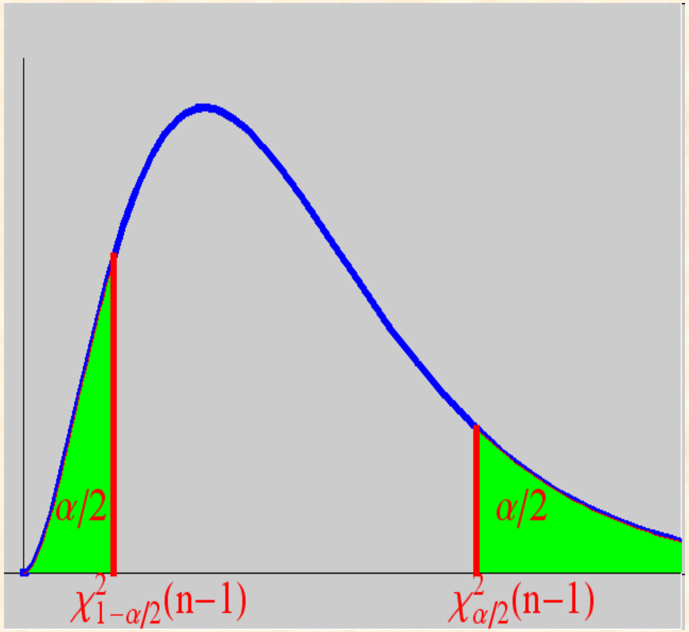
$$\left(\frac{\sqrt{n-1}S_n^*}{\sqrt{\chi^2_{\alpha/2}(n-1)}},\ \frac{\sqrt{n-1}S_n^*}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right).$$

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5 (续例3) 求例3中总体标准差σ的置信度为0.95 的置信区间.

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:

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

计算得 $s_n^*=6.2022$, $\left(\dfrac{\sqrt{n-1}S_n^*}{\sqrt{\chi^2_{\alpha/2}(n-1)}},\ \dfrac{\sqrt{n-1}S_n^*}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right).$

代入公式得标准差的置信区间

$[4.58,\ 9.60].$

Ⅱ.两个总体 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ 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、二个总体的修正样本方差.

讨论两个总体均值差和方差比的估计问题.

1. 两个总体均值差 $\mu_1 - \mu_2$ 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

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

推导过程如下:

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

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

由 $\bar{X}, \bar{Y}$ 的独立性及

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

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

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

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

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

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

取 $$T=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2),$$

$\mu_1-\mu_2$的一个置信度为$1-\alpha$的置信区间

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

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

例6 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 $\bar{x}_1 = 500(m/s)$, 修正标准差 $s_1^* = 1.10(m/s)$, 随机地取II型子弹20发, 得枪口速度平均值为 $\bar{x}_2 = 496(m/s)$, 修正标准差 $s_2^* = 1.20(m/s)$, 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：$t_{0.025}(28)=2.0484,$

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

于是得 $\mu_1-\mu_2$ 的一个置信度为 0.95 的置信区间

$$\left(\overline{x}_1-\overline{x}_2\pm S_w\times t_{0.025}(28)\sqrt{\frac{1}{10}+\frac{1}{20}}\right)=(4\pm0.93),$$

即所求置信区间为[3.07，4.93].

例7 为提高某以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了$n_2=8$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),

且 $S_w^{\ 2}=\dfrac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{n_1+n_2-2}=3.96,$

于是得 $\mu_1-\mu_2$的一个置信度为 0.95的置信区间

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

即所求置信区间为$[-4.15,\ 0.11]$.

说明: 由于所得置信区间包含零, 在实际中我们认为采用这两种催化剂所得的得率的均值没有显著差别.

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

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

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

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

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

$$=1-\alpha,$$

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

$$=1-\alpha,$$

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

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

例8 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 $s_1^{*2}=0.34(mm^2)$; 抽取机器B生产的管子13只,测得样本方差为 $s_2^{*2}=0.29(mm^2)$. 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 $n_1=18,\quad n_2=13,\quad \alpha=0.10,$

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

$$F_{\alpha/2}(n_1-1, n_2-1) = F_{0.05}(17, 12) = 2.583,$$

$$F_{1-\alpha/2}(17, 12) = F_{0.95}(17, 12) = \frac{1}{F_{0.05}(12, 17)} = \frac{1}{2.381},$$

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

$$\left(\left[\frac{0.34}{0.29}\times\frac{1}{2.583},\ \frac{0.34}{0.29}\times 2.381\right]\right) = ([0.45,\ 2.79]).$$

由于在$\dfrac{\sigma_1^2}{\sigma_2^2}$的置信区间中包含1, 在实际中我们认为$\sigma_1^2$和$\sigma_2^2$两者没有显著差别.

例9 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 $n_1=9,\qquad n_2=6,\qquad \alpha=0.02,$

$$F_{1-\alpha/2}(n_1-1,n_2-1)=F_{0.99}(8,5)=1/6.63,$$

$$F_{\alpha/2}(8,5)=F_{0.01}(8,5)=10.3,$$

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

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

$$=\left([\sqrt{\frac{0.245}{0.357\times 10.3}},\ \sqrt{\frac{0.245\times 6.63}{0.357}}]\right)=([0.258,\ 2.133]).$$

三、小结

点估计不能反映估计的精度, 故而本节引入了区间估计.

置信区间是一个随机区间$[\hat{\theta}_1, \hat{\theta}_2]$, 它覆盖未知参数具有预先给定的高概率(置信水平), 即对于任意的$\theta \in \Theta$, 有 $P\{\hat{\theta}_1 \le \theta \le \hat{\theta}_2\} \ge 1-\alpha$.

求置信区间的一般步骤(分三步).

求置信区间的一般步骤(共3步)

1° 寻求一个样本$X_1, X_2, \cdots, X_n$的函数:

$$Z = Z(X_1, X_2, \cdots, X_n; \theta)$$

其中仅包含待估参数θ,并且Z的分布已知且不依赖于任何未知参数(包括θ).

2° 对于给定的置信度$1-\alpha$,定出两个常数a, b, 使

$$P\{a \le Z(X_1, X_2, \cdots, X_n; \theta) \le b\} = 1-\alpha.$$

3° 作等价变形

$$a \le Z(X_1, X_2, \cdots, X_n; \theta) \le b$$

$$\Leftrightarrow \quad \hat{\theta}_1 \le \theta \le \hat{\theta}_2,$$

其中$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \cdots, X_n)$, $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \cdots, X_n)$

那么$[\hat{\theta}_1, \hat{\theta}_2]$就是$\theta$的一个置信度为$1-\alpha$的置信区间.

正态总体均值与方差的区间估计

1. 单个总体均值 μ 的置信区间

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

2. 单个总体方差 σ^2 的置信区间

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

3. 两个总体均值差 $\mu_1-\mu_2$ 的置信区间

σ_1^2 和 σ_2^2 均为已知, $\left(\overline{X}-\overline{Y}\pm u_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}\right)$.

σ_1^2 和 σ_2^2 均为未知, $\left(\overline{X}-\overline{Y}\pm u_{\alpha/2}\sqrt{\dfrac{S_1^{*2}}{n_1}+\dfrac{S_2^{*2}}{n_2}}\right)$.

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

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

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

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

备份题

例5 设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, \sigma^2)$的样本,其中σ^2和μ为未知参数,设随机变量L是关于μ的置信度为$1-\alpha$的置信区间的长度,求$E(L^2)$.

解 当σ^2未知时,

μ的置信度为$1-\alpha$的置信区间为

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

置信区间长度$L = \dfrac{2S_n^*}{\sqrt{n}} t_{\alpha/2}(n-1)$,

$$L^2 = \frac{4S_n^{*2}}{n}[t_{\alpha/2}(n-1)]^2,$$

又 $E(S_n^{*2}) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right]$

$$= E\left\{\frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^{\ 2} - n\overline{X}^{\,2}\right]\right\}$$

$$= \frac{1}{n-1}\left\{\sum_{i=1}^{n} E(X_i^{\ 2}) - nE(\overline{X}^{\,2})\right\}$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}[D(X_i)+E(X_i)^2]-n[D(\overline{X})+E^2(\overline{X})]\right\}$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}[\sigma^2+\mu^2]-n\left[\frac{\sigma^2}{n}+\mu^2\right]\right\}=\sigma^2,$$

于是 $E(L^2)=E\left(\dfrac{4S_n^{*2}}{n}[t_{\alpha/2}(n-1)]^2\right)$

$$=\frac{4}{n}[t_{\alpha/2}(n-1)]^2E(S_n^{*2})=\frac{4}{n}[t_{\alpha/2}(n-1)]^2\sigma^2.$$