第二章 参数估计

§3 区间估计

理学院

预习提纲

1. 什么是参数的区间估计？它与点估计的区别是什么？

2. 置信区间与置信概率的定义与联系？

- **本次课教学目的：**
  - 掌握区间估计的一些重要定理并会进行区间估计
- **重点难点：**
  - 一个总体参数的区间估计方法
  - 两个总体参数的区间估计方法

§3 区间估计

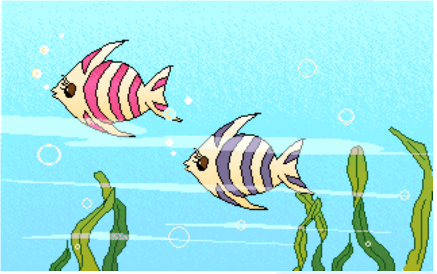
Southwest Petroleum University
西南石油大学

3.1 区间估计概述

为什么需要进行区间估计？

点估计是用样本的估计量直接作为总体参数的估计值，它没有给出估计值接近总体参数程度的信息。区间估计正好弥补了点估计的这个缺陷。

譬如，在估计湖中鱼数的问题中，若我们根据一个实际样本，得到鱼数 N 的估计为1000条.

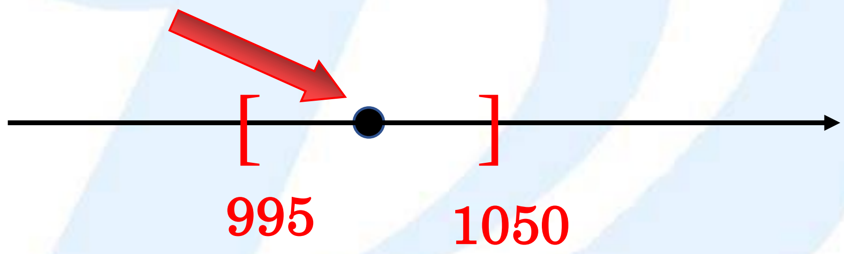


实际上，N 的真值可能大于1000条，也可能小于1000条。

若我们能给出一个区间，在此区间内我们合理地相信 N 的真值位于其中。这样对鱼数的估计就有把握多了。

也就是说，我们希望确定一个区间，使我们能以比较高的可靠程度相信它包含真参数值.

90%可能性包含鱼数
的真实值



这里所说的“可靠程度”是用概率来度量的，称为置信概率或置信水平。

习惯上把置信概率记作 $1-\alpha$，这里 α 是一个很小的正数。置信水平的大小是根据实际选定的，一般地，可取 $1-\alpha=0.95$ 或0.90等。

设母体 X 的分布函数是 $F(x;\theta)$，其中 θ 是未知参数。从母体中抽取子样 $(X_1,X_2,\cdots,X_n)$ ，作统计量 $\theta_1(X_1,X_2,\cdots,X_n)$ 和 $\theta_2(X_1,X_2,\cdots,X_n)$ 使 $P(\theta_1<\theta<\theta_2)=1-\alpha$，称区间 (θ_1,θ_2) 为 θ 的置信概率为 $1-\alpha$ 的置信区间。θ_1 和 θ_2 分别称为置信下限和置信上限。

Southwest Petroleum University 西南石油大學

说明：

- 参数 θ的区间估计的意义可以解释为：随机区间(θ_1,θ_2)包含参数θ 的真值的概率为$1-\alpha$，因此，若认为区间(θ_1,θ_2)包含着参数 θ的真值，则犯错误的概率为 α 。

- 由于 θ不是随机变量，所以不能说参数 θ以 $1-\alpha$ 的概落入随机区间 (θ_1,θ_2)，而只能说区间 (θ_1,θ_2) 以 $1-\alpha$ 的概率包含 θ。

- 置信度$1-\alpha$反映了区间估计的可靠性，而置信区间$\theta_2-\theta_1$的长度则反映了区间估计的精度。一般来说，置信度$1-\alpha$越大越好，置信区间$\theta_2-\theta_1$的长度越小越好。

- **区间估计原则：在保证可靠性（一般α为0.1，0.05，0.01等）的前提下，努力提高精度（即尽量选取长度短的置信区间）。**

3.2 大子样对母体平均数区间估计

1. 假定条件
   - 总体分布未知，且平均数(μ) 和方差(σ^2)均未知
   - 大子样 $(n \geq 50)$

2. 由中心极限定理，当n很大时，$\overline{X}$ 近似地服从正态分布，σ 可用子样标准差S近似，固有

$$U = \frac{\overline{X} - \mu}{S/\sqrt{n}} \overset{\text{近似}}{\sim} N(0,1)$$

3. 给定1-α，可找到$\mu_{\frac{\alpha}{2}}$，使

$$P\left\{|U| < \mu_{\frac{\alpha}{2}}\right\} = P\left\{\frac{|\overline{X} - \mu|}{S/\sqrt{n}} < \mu_{\frac{\alpha}{2}}\right\} \approx 1 - \alpha$$

即有：$P\left\{\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}}<\mu<\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}}\right\}\approx 1-\alpha$

于是可得 μ 的置信区间为：$(\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}},\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}})$

例1： 从一台机床加工的轴中随机地抽取200根，测量其椭圆度。由测量值计算得平均值 $\bar{x}=0.081$ 毫米，标准差 $s=0.025$ 毫米。给定置信概率为95%，求此机床加工轴平均椭圆度的置信区间。

解： 由于n=200,故可认为是大子样。已知 $1-\alpha=0.95$，查附表可得 $\mu_{\frac{\alpha}{2}}=1.96$。于是有：

$$P\left\{|U|<\mu_{\frac{\alpha}{2}}\right\}=P\left\{\frac{|\overline{X}-\mu|}{S/\sqrt{n}}<\mu_{\frac{\alpha}{2}}\right\}\approx 1-\alpha$$

由此得到

置信下限为：$\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}}=0.081-1.96\times\dfrac{0.025}{\sqrt{200}}=0.078$

置信上限为：$\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}}=0.081+1.96\times\dfrac{0.025}{\sqrt{200}}=0.084$

故置信区间为:(0.078,0.084)。

3.3 正态母体平均数区间估计

设母体 X 服从正态分布 $N(\mu,\sigma^2)$

情况一：方差 σ^2 已知，则有 $U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

给定$1-\alpha$，可找到$\mu_{\frac{\alpha}{2}}$，使

$$P\left\{|U|<\mu_{\frac{\alpha}{2}}\right\}=P\left\{\frac{|\overline{X}-\mu|}{\sigma/\sqrt{n}}<\mu_{\frac{\alpha}{2}}\right\}\approx 1-\alpha$$

即有：$P\left\{\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}<\mu<\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}\right\}\approx 1-\alpha$

于是可得 μ 的置信区间为：$(\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}},\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}})$

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

$$P\left\{|U|<\mu_{\frac{\alpha}{2}}\right\}=P\left\{\frac{|\overline{X}-\mu|}{\sigma/\sqrt{n}}<\mu_{\frac{\alpha}{2}}\right\}\approx 1-\alpha$$

例2: 设轴承内环的锻压零件的平均高度X服从正态分布$N(\mu,0.4^2)$，现在从中抽取20只内环，其平均高度为32.3毫米。求内环平均高度的置信度为95%的置信区间。

解: $1-\alpha=0.95$，查表得$\mu_{\alpha/2}=\mu_{0.025}=\Phi(0.975)=1.96$

又$\bar{x}=32.3,\sigma=0.4,n=20$，算得

$$\bar{x}-\mu_{\alpha/2}\frac{\sigma}{\sqrt{n}}=32.3-1.96\times\frac{0.4}{\sqrt{20}}=32.12$$

$$\bar{x}+\mu_{\alpha/2}\frac{\sigma}{\sqrt{n}}=32.3+1.96\times\frac{0.4}{\sqrt{20}}=32.48$$

所以μ的一个置信度为95%的置信区间为(32.12,32.48)

Southwest Petroleum University
西南石油大学

情况二：方差σ^2未知，便无法利用$U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，需要利用其他的函数及其分布。

定理1：设随机变量 $X_1,X_2,\cdots,X_n$独立同分布，且每个随机变量服从正态分布$N(\mu,\sigma^2)$，则$(n-1)\dfrac{S^{*2}}{\sigma^2}=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$服从自由度为$n-1$的$\chi^2$分布，且$\overline{X}$与$S^{*2}$相互独立。

定理1.3.3(P15)

根据定理1和t分布的定义，可得：

$$T=\frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^{*2}/\sigma^2}{n-1}}}=\frac{\overline{X}-\mu}{\dfrac{S^*}{\sqrt{n}}}$$

服从自由度为$n-1$的t分布。 定理1.3.4(P16)

对于给定$1-\alpha$，存在$t_{\frac{\alpha}{2}}(n-1)$ 使

$$P\left\{|T| < t_{\frac{\alpha}{2}}(n-1)\right\} = P\left\{-t_{\frac{\alpha}{2}}(n-1) < \frac{\overline{X}-\mu}{S^*/\sqrt{n}} < t_{\frac{\alpha}{2}}(n-1)\right\} \approx 1-\alpha$$

即有：$P\left\{\overline{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S^*}{\sqrt{n}} < \mu < \overline{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S^*}{\sqrt{n}}\right\} \approx 1-\alpha$

于是可得μ的置信区间为:

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

例3：已知某种灯泡的寿命服从正态分布，现从一批灯泡中随机抽取16只，测得其使用寿命(小时)如下。求该批灯泡平均使用寿命95%的置信区间。

16只灯泡使用寿命的数据			
1510	1520	1480	1500
1450	1480	1510	1520
1480	1490	1530	1510
1460	1460	1470	1470

解：已知 $X \sim N(\mu, \sigma^2)$，$n=16$, $1-\alpha = 95\%$，$t_{\alpha/2}(15)=2.131$

根据样本数据计算得：$\bar{x}=1490$ ，$s=24.77$

总体均值μ在$1-\alpha$置信水平下的置信区间为

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = 1490 \pm 2.131 \times \frac{24.77}{\sqrt{16}}$$

$$= 1490 \pm 13.2 = (1476.8, 1503.2)$$

该种灯泡平均使用寿命的置信区间为1476.8小时～1503.2小时

$$P\left\{|T| < t_{\frac{\alpha}{2}}(n-1)\right\} = P\left\{-t_{\frac{\alpha}{2}}(n-1) < \frac{\bar{X}-\mu}{S^*/\sqrt{n}} < t_{\frac{\alpha}{2}}(n-1)\right\} \approx 1-\alpha$$

小结—总体均值(μ)的区间估计

对母体（或子样）要求	所用函数及其分布	置信区间
大子样	$U=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\overset{\text{近似}}{\sim}N(0,1)$	$(\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}},\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{S}{\sqrt{n}})$
正态母体方差σ_0^2已知	$U=\dfrac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1)$	$(\overline{X}-\mu_{\frac{\alpha}{2}}\dfrac{\sigma_0}{\sqrt{n}},\overline{X}+\mu_{\frac{\alpha}{2}}\dfrac{\sigma_0}{\sqrt{n}})$
正态母体方差σ_0^2未知	$T=\dfrac{\overline{X}-\mu}{S^*/\sqrt{n}}\sim t(n-1)$	$(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\dfrac{S^*}{\sqrt{n}},\overline{X}+t_{\frac{\alpha}{2}}(n-1)\dfrac{S^*}{\sqrt{n}})$

3.4 大子样对两个母体平均数之差区间估计

1. 假定条件

- 两个总体分布未知，且平均数 μ_1，μ_2，方差 σ_1，σ_2 均未知
- 两个样本均为大样本 $(n_1 \geq 50, \; n_2 \geq 50)$
- 两个样本是独立的随机样本

2. 由于子样的独立性可知 $\bar{X}_1$ 和 $\bar{X}_2$ 是独立的，因而有

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2, D(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

由中心极限定理，当 n_1, n_2 很大时，$\bar{X}_1$ 和 $\bar{X}_2$ 都近似地服从正态分布，σ_1^2 和 σ_2^2 可分别用 S_1^2 和 S_2^2 替代，固有：

$$U=\frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\overset{近似}{\sim}N(0,1)$$

3. 给定1−α，可找到$\mu_{\frac{\alpha}{2}}$，使

$$P\left\{|U|<\mu_{\frac{\alpha}{2}}\right\}=P\left\{-\mu_{\frac{\alpha}{2}}<\frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}<\mu_{\frac{\alpha}{2}}\right\}\approx 1-\alpha$$

于是可得$\mu_1-\mu_2$的置信区间为:

$$\left(\overline{X}_1-\overline{X}_2-\mu_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}},\right.$$

$$\left.\overline{X}_1-\overline{X}_2+\mu_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)$$

例4：某地区教育委员会想估计两所中学的学生中考时的英语平均分数之差，为此在两所中学独立抽取两个随机样本，有关数据如右表。建立两所中学中考英语平均分数之差95%的置信区间

两个样本的有关数据	
中学1	中学2
$n_1 = 64$	$n_2 = 52$
$\bar{x}_1 = 86$	$\bar{x}_2 = 78$
$S_1 = 5.8$	$S_2 = 7.2$

解：两个总体均值之差在1-α置信水平下的置信区间为

$$(\bar{x}_1-\bar{x}_2)\pm\mu_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$=(86-78)\pm1.96\times\sqrt{\frac{5.8^2}{64}+\frac{7.2^2}{52}}$$

$$=8\pm2.42=(5.58,10.42)$$

两所中学中考英语平均分数之差的置信区间为(5.58, 10.42)

3.5 两个正态母体平均数之差区间估计

1. 假定条件
   - 两个总体都服从正态分布
   - 两个总体方差相等但未知：$\sigma_1{}^2=\sigma_2{}^2$
   - 两个样本独立
2. 根据3.3节定理1，可知

定理1.3.3(P15)

$$\frac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{\sigma^2}\sim\chi^2(n_1+n_2-2)$$

再根据t分布的定义，有

$$T=\frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\Big/\sqrt{\frac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{n_1+n_2-2}}\sim t(n_1+n_2-2)$$

定理1.3.6(P17)

令 $S^{*2}=\dfrac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{n_1+n_2-2}$

对于给定1−α，存在 $t_{\frac{\alpha}{2}}(n_1+n_2-2)$ 使

$$P\left\{-t_{\frac{\alpha}{2}}(n_1+n_2-2)<\frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S^*}<t_{\frac{\alpha}{2}}(n_1+n_2-2)\right\}=1-\alpha$$

于是可得$\mu_1-\mu_2$的置信区间为：

$$(\overline{X}_1-\overline{X}_2-t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S^*,$$

$$\overline{X}_1-\overline{X}_2+t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S^*)$$

Southwest Petroleum University 西南石油大學

例5： 随机地从甲、乙两厂生产的蓄电池中抽取一些样本，测得蓄电池的电容量如下

甲厂：144，141，138，142，141，143，138，137

乙厂：142，143，139，140，138，141，140，138，142，136

设两厂生产的蓄电池电容量分别服从正态总体 $N(\mu_1,\sigma_1^2)$ 和 $N(\mu_2,\sigma_2^2)$。两样本独立，若已知 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知。求 $\mu_1-\mu_2$ 的置信度为0.95的置信区间。

解： 由题意得：$n_1=8,\quad n_2=10$

通过一系列的计算，可得：

$$\bar{x}=140.5,\quad s_1^{*2}=\frac{\sum_{i=1}^{n_1}(x_i-\bar{x})^2}{n_1-1}=6.57$$

$$\bar{y}=139.9,\quad s_2^{*2}=\frac{\sum_{i=1}^{n_2}(y_i-\bar{y})^2}{n_2-1}=4.77$$

$$S^*=\sqrt{\frac{7s_1^{\ 2}+9s_2^{\ 2}}{16}}=2.36,\quad \alpha=0.05,\quad t_{0.025}(16)=2.1199$$

于是可得 $\mu_1-\mu_2$ 置信度为0.95的区间为：

$$(\bar{X}_1-\bar{X}_2-t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S^*,\bar{X}_1-\bar{X}_2+t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S^*)=(-1.77,2.97)$$

小结—平均数之差$(\mu_1-\mu_2)$的区间估计

对母体（或子样）要求	所用函数及其分布	置信区间
大子样	$U=\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\overset{\text{近似}}{\sim}N(0,1)$	$(\bar{X}_1-\bar{X}_2-\mu_{\frac{\alpha}{2}}\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}},$ $\bar{X}_1-\bar{X}_2+\mu_{\frac{\alpha}{2}}\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}})$
两个正态母体，方差相等	$T=\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}S^*}\sim t(n_1+n_2-2)$	$(\bar{X}_1-\bar{X}_2-t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}S^*,$ $\bar{X}_1-\bar{X}_2+t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}S^*)$

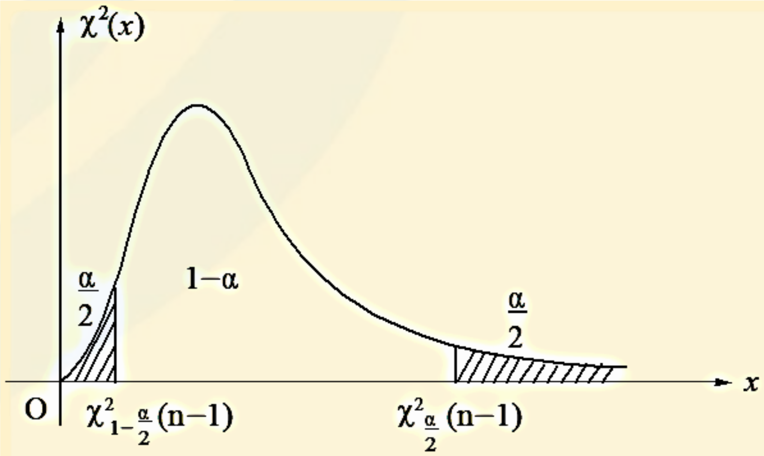
3.6 正态母体方差区间估计

1. 假定条件

- 总体服从正态分布$N(\mu, \sigma^2)$
- 平均数 μ和方差σ^2均未知

2. 根据3.3定理1，又由于样本方差S^{*2}是总体方差σ^2的无偏估计量，故有：$\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$　　定理1.3.3(P15)

3. 对于给定的置信度1−α和自由度n−1，查分布分位数表，可得到两个临界值 $\chi^2_{\alpha/2}(n-1)$和$\chi^2_{1-\alpha/2}(n-1)$，使得$P\{\chi^2_{1-\alpha/2}<\chi^2<\chi^2_{\alpha/2}\}=1-\alpha$



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

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

故总体方差σ^2置信度为1-α的置信区间为：

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

例6： 设高速公路上汽车的速度服从正态分布，现对汽车的速度独立地作了5次测试，求得这5次测试值的方差 $s^2=0.09$。求汽车速度的方差σ^2的置信度为0.9的置信区间。

解： 由题意得：$1-\alpha=0.9,\quad \alpha=0.1\quad ,n-1=4$

查表可得：$\chi^2_{\alpha/2}(4)=\chi^2_{0.05}(4)=9.448$

$$\chi^2_{1-\alpha/2}(4)=\chi^2_{0.95}(4)=0.711$$

于是得到：$\dfrac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}=\dfrac{4\times0.09}{9.448}=0.038$

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}=\frac{4\times0.09}{0.711}=0.506$$

所求置信区间为:**(0.038,0.506)**

3.7 两个正态母体方差比的区间估计

1. 假定条件

- 两个总体均服从正态分布
- 总体平均数 μ_1，μ_2和方差σ_1^2，σ_2^2 均未知

2. 根据3.3定理1有：$\dfrac{(n_1-1)S_1^{*2}}{\sigma_1^2} \sim \chi(n_1-1), \dfrac{(n_1-1)S_2^{*2}}{\sigma_2^2} \sim \chi(n_2-1)$

再由F分布的定义知：

定理1.2.4(P13)

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

3. 对于给定的置信度1-α ，有

$$P\left\{F_{1-\alpha/2}(n_2-1,n_1-1)<F<F_{\alpha/2}(n_2-1,n_1-1)\right\}=1-\alpha$$

由此解得$\dfrac{\sigma_1^2}{\sigma_2^2}$的置信度为1-$\alpha$ 的置信区间是

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

例7：随机地从甲、乙两厂生产的蓄电池中抽取一些样本，测得蓄电池的电容量如下

甲厂：144，141，138，142，141，143，138，137

乙厂：142，143，139，140，138，141，140，138，142，136

设两厂生产的蓄电池电容量分别服从正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$。两样本独立，若两个总体的方差 $\sigma_1^2 \neq \sigma_2^2$，现求 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信度为0.95的置信区间。

解： 由题意得：$n_1 = 8,\quad n_2 = 10$

通过一系列的计算，可得：

$$\bar{x} = 140.5,\quad s_1^{*2} = \frac{\sum_{i=1}^{n_1}(x_i - \bar{x})^2}{n_1 - 1} = 6.57$$

$$\bar{y} = 139.9,\quad s_2^{*2} = \frac{\sum_{i=1}^{n_2}(y_i - \bar{y})^2}{n_2 - 1} = 4.77$$

查表得 $F_{0.025}(9,7) = 4.82$, 又有 $F_{0.975}(9,7) = \frac{1}{F_{0.025}(7,9)} = \frac{1}{4.20} = 0.24$

于是可得 $\frac{\sigma_1^2}{\sigma_2^2}$ 置信度为0.95的区间为：

$$\left(F_{1-\alpha/2}(n_2-1, n_1-1)\frac{S_1^{*2}}{S_2^{*2}}, F_{\alpha/2}(n_2-1, n_1-1)\frac{S_1^{*2}}{S_2^{*2}}\right) = (0.33, 6.64)$$

小结—方差的区间估计

估计对象	对母体（或子样）要求	所用函数及其分布	置信区间
方差σ^2	正态母体	$\chi^2=(n-1)\dfrac{S^{*2}}{\sigma^2}\sim\chi^2(n-1)$	$\left(\dfrac{(n-1)S^{*2}}{\chi^2_{\frac{\alpha}{2}}(n-1)},\dfrac{(n-1)S^{*2}}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$
方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$	两个正态母体	$F=\dfrac{S_2^*/\sigma_2^*}{S_1^*/\sigma_1^2}\sim F(n_2-1,n_1-1)$	$(F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)\dfrac{S_1^*}{S_2^*}$, $F_{\frac{\alpha}{2}}(n_2-1,n_1-1)\dfrac{S_1^*}{S_2^*})$

3.8 单侧置信区间

在上述讨论中，对于未知参数θ，我们给出两个统计量θ_1和θ_2得到 的双侧置信区间(θ_1,θ_2)，但在一些实际问题中，例如，对于设备，元件的寿命来说，平均寿命长是我们所希望得到的，我们关心的是平均寿命θ的"下限"，与此相反，在考虑化学药品中杂质含量的均值EX时，我们常关心参数EX的"上限"。这就引出了单侧置信区间的概念。

对于给定值 $\alpha(0<\alpha<1)$，若由样本 $(X_1,X_2,\cdots,X_n)$ 确定的统计量 $\theta_1=\theta_1(X_1,X_2,\cdots,X_n)$ 对于任意$\theta\in\Theta$ 满足：$P\{\theta\geq\theta_1\}\geq 1-\alpha$

称随机区间(θ_1,∞)是θ的置信水平为$1-\alpha$的单侧置信区间，

称 θ_1为θ的置信水平为$1-\alpha$ 的单侧置信下限。

对于给定值 $\alpha(0<\alpha<1)$，若由样本 $(X_1,X_2,\cdots,X_n)$ 确定的统计量 $\theta_2=\theta_2(X_1,X_2,\cdots,X_n)$ 对于任意$\theta\in\Theta$ 满足：$P\{\theta\leq\theta_2\}\geq 1-\alpha$

称随机区间 $(-\infty,\theta_2)$是θ的置信水平为$1-\alpha$的单侧置信区间，

称θ_2为θ的置信水平为$1-\alpha$ 的单侧置信上限。

例2.3.2(P39)

例8 :从一批灯泡中随机地取5只做寿命测试，测得寿命（以小时计）为 1050，1100，1120，1250，1280。

设灯泡寿命服从正态分布，求灯泡寿命平均值的置信度为0.95的单侧置信下限。 一个正态总体均值的区间估计(方差未知)

解： 此时有：$t=\dfrac{\overline{X}-\mu}{s/\sqrt{n}}\sim t(n-1)$，于是 $P\left\{\dfrac{\overline{X}-\mu}{s/\sqrt{n}}<t_{\alpha}(n-1)\right\}=1-\alpha$

$P\left\{\mu>\overline{X}-\dfrac{s}{\sqrt{n}}t_{\alpha}(n-1)\right\}=1-\alpha$ 又由于

$$1-\alpha=0.95,\quad n=5,\quad \overline{x}=1160,\quad s^2=9950,\quad t_{0.05}(4)=2.1318$$

所以其单侧置信下限为：

$$\overline{x}-\frac{s}{\sqrt{n}}t_{\alpha}(n-1)=1160-\frac{\sqrt{9950}}{\sqrt{5}}\times 2.1318=1065$$

思考题

1. 在测量反应时间时，一个心理学家估计的标准差是0.05秒，为了能以95%的置信度使得他对平均反应时间的估计误差不超过0.01秒，问应该抽取多大的样本容量n。

解：选取 $U=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim N(0,1)$

由条件 $P\{|\overline{X}-\mu|\le 0.01\}=0.95$，可得

$P\{\left|\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\right|\le\dfrac{\sqrt{n}}{5}\}=0.95$ 于是有：$2\Phi_0(\dfrac{\sqrt{n}}{5})-1=0.95$

$\Phi_0(\dfrac{\sqrt{n}}{5})=0.975$ $\quad\dfrac{\sqrt{n}}{5}=1.96\Rightarrow n\approx 96.04$

所以抽取的样本量为n>96。

思考题

2. 以商店销售某种商品来自甲、乙两厂家，为检测商品的性能差异，现从甲、乙两厂产品中分别抽取8件和9件产品，检得性能指标X数据如下表：

甲厂X_1	0.30	0.12	0.18	0.25	0.27	0.08	0.19	0.13	
乙厂X_2	0.28	0.30	0.11	0.14	0.35	0.26	0.14	0.31	0.20

设两厂的产品性能指标服从$x_1 \sim N(\mu_1,\sigma_1^2)$ 和 $x_1 \sim N(\mu_1,\sigma_1^2)$

（1） 若σ_1^2=0. 0064，σ_2^2=0. 0081，
求$\mu_1-\mu_2$置信度为0. 9置信区间。

(2)求$\dfrac{\sigma_1^2}{\sigma_2^2}$置信度为0.9置信区间。

思考题答案

（1）解：$n_1=8, n_2=9, \overline{X}_1=0.19, \overline{X}_2=0.232$

σ_1^2、σ_2^2已知 $\mu_{0.05}=1.64$

因此置信度为0.9的$\mu_1-\mu_2$置信区间为：

$$\left(\overline{X}_1-\overline{X}_2-\mu_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ ,\ \overline{X}_1-\overline{X}_2+\mu_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

$$=(-0.042-1.64\times0.041\ ,\ -0.042+1.64\times0.041)$$

$$=(-0.109, 0.025)$$

方差比 $\frac{\sigma_1^2}{\sigma_2^2}$	两个正态母体	$F=\frac{S_2^*/\sigma_2^*}{S_1^*/\sigma_1^2}\sim F(n_2-1,n_1-1)$	$(F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)\frac{S_1^*}{S_2^*},$ $F_{\frac{\alpha}{2}}(n_2-1,n_1-1)\frac{S_1^*}{S_2^*})$

（2）解：由于α=0.1，查表可得

$$F_{0.05}(7,8)=3.50 \qquad F_{0.05}(8,7)=3.73$$

于是有 $F_{0.95}(8,7)=\dfrac{1}{F_{0.05}(7,8)}=\dfrac{1}{3.50}=0.29$

因此置信度为0.9的 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 置信区间为

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

$$=\left(0.29\times\frac{0.0061}{0.0081},3.73\times\frac{0.0061}{0.0081}\right)=(0.22,2.81)$$

Thank you!