二、两个总体 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ 的情况

设已给定置信水平为 $1-\alpha$，并设 $X_1,X_2\ldots,X_{n_1}$ 是来自第一个总体的样本，$Y_1,Y_2\ldots,Y_{n_2}$ 是来自第二个总体的样本，这两个样本相互独立．且设 $\bar{X},\bar{Y}$ 分别为第一、二个总体的样本均值，S_1^2,S_2^2 为第一、二个总体的样本方差．

1. 两个总体均值差 $\mu_1-\mu_2$ 的置信区间

未知参数

1° σ_1^2,σ_2^2 为已知

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

因为 X,Y 相互独立，所以 $\bar{X},\bar{Y}$ 相互独立.

故

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

μ_1-μ_2的点估计量

或

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

于是得到 $\mu_1-\mu_2$ 的置信水平为 $1-\alpha$ 的置信区间为

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

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

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

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

定理 4 (两总体样本均值差、样本方差比的分布)

设$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$，且X与Y独立，$X_1,X_2,\ldots,X_{n_1}$是来自X的样本，$Y_1,Y_2,\ldots,Y_{n_2}$是取自Y的样本，$\bar{X}$和$\bar{Y}$分别是这两个样本的样本均值，S_1^2和S_2^2 分别是这两个样本的样本方差，则有

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

2、当$\sigma_1^2=\sigma_2^2=\sigma^2$时，$\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$

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

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

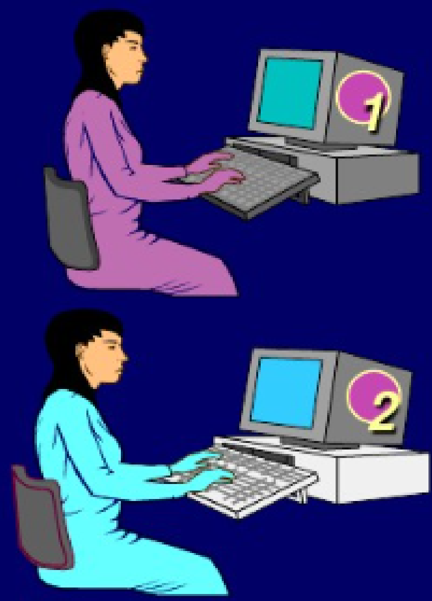
其中 $S_\omega=\sqrt{S_\omega^2}$, $S_\omega^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$.

于是得到 $\mu_1-\mu_2$ 的置信水平为 $1-\alpha$ 的置信区间为

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

【例6】为估计两种方法组装产品所需时间的差异，分别对两种不同的组装方法各随机安排12名工人，每个工人组装一件产品所需的时间(单位：min)下如表。假定两种方法组装产品的时间服从正态分布，且方差相等。试以95%的置信水平建立两种方法组装产品所需平均时间差值的置信区间。

两个方法组装产品所需的时间			
方法1		方法2	
28.3	36.0	27.6	31.7
30.1	37.2	22.2	26.0
29.0	38.5	31.0	32.0
37.6	34.4	33.8	31.2
32.1	28.0	20.0	33.4
28.8	30.0	30.2	26.5



解: 根据样本数据计算得

$$\bar{x}_1 = 32.5 \quad s_1^2 = 15.996 \quad \bar{x}_2 = 28.8 \quad s_2^2 = 19.358$$

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

合并估计量为

$$s_\omega^2 = \frac{(12-1)\times 15.996 + (12-1)\times 19.358}{12+12-2} = 17.677$$

$$S_\omega^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}.$$

$t_{\alpha/2}(n_1+n_2-2)= t_{0.25}(22)=2.0739$

$$(32.5-28.8) \pm 2.0739 \times \sqrt{17.677\times\left(\frac{1}{12}+\frac{1}{12}\right)} = 3.7 \pm 3.56$$

两种方法组装产品所需平均时间之差的置信区间为

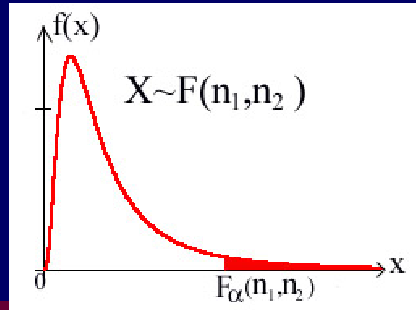
0.14min~7.26min。

置信区间的下限大于0，在实际中认为μ_1比μ_2大。

若置信区间的包含0，则认为μ_1与μ_2的差异不显著。

2. 两个总体方差比 σ_1^2/σ_2^2 的置信区间

$(\mu_1,\mu_2$ 为未知 $)$



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

$$F_{1-\alpha/2}(n_1,n_2)=\frac{1}{F_{\alpha/2}(n_2,n_1)}$$

σ^2/σ^2

$P\{F_{1-\alpha/2}($ $=1-\alpha$

定理 4 (两总体样本均值差、样本方差比的分布)

设 $X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$，且X与Y独立，

$X_1,X_2,\dots,X_{n_1}$是来自X的样本，$Y_1,Y_2,\dots,Y_{n_2}$是取自Y的样本，

$\overline{X}$和$\overline{Y}$分别是这两个样本的样本均值，S_1^2和S_2^2 分别是这两个样本的样本方差，则有

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

即

$P\{\dfrac{S_1^2}{S_2^2}\dfrac{}{F_{\alpha/}}$ $)\}=1-\alpha$

【例7】为了研究男女学生在生活费支出(单位：元)上的差异，在某大学各随机抽取25名男学生和25名女学生，得到下面的结果

男学生： $\bar{x}_1 = 520$　　$s_1^2 = 260$

女学生： $\bar{x}_2 = 480$　　$s_2^2 = 280$

试以90%置信水平估计男女学生生活费支出方差比的置信区间。

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

解：根据自由度 n_1=25-1=24 ， n_2=25-1=24， 查得

$F_{\alpha/2}(24,24)=1.98$， $F_{1-\alpha/2}(24,24)=1/1.98=0.505$

σ_1^2/σ_2^2 置信度为90%的置信区间为

$$\frac{260/280}{1.98}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{260/280}{0.505}$$

男女学生生活费支出方差比的置信区间为

0.47~1.84 。

置信区间包含1，在实际中认为 σ_1^2与σ_2^2 两者没有显著差异。

四、小结

在本节中，我们学习了单个正态总体均值、方差的置信区间，两个正态总体均值差、方差比的置信区间.

- 重点：P172 表7-1