正态分布中μ的区间估计

例　设$X \sim N(\mu, \sigma^2)$，$\sigma^2 = \sigma_0^2$，求参数μ的置信度为$1-\alpha$的置信区间.

分析：

1）要估计参数，就涉及统计量；而选取统计量应根据优良性质准则来选.

这里μ的优良估计是：　$\bar{X}$

它是无偏、有效、相合估计.

2）考察统计量所服从的分布，这里：

$$\bar{X} \sim N\left(\mu, \frac{\sigma_0^2}{n}\right)$$

3）将统计量化为常用分布，再通过临界值确定区间，这里：

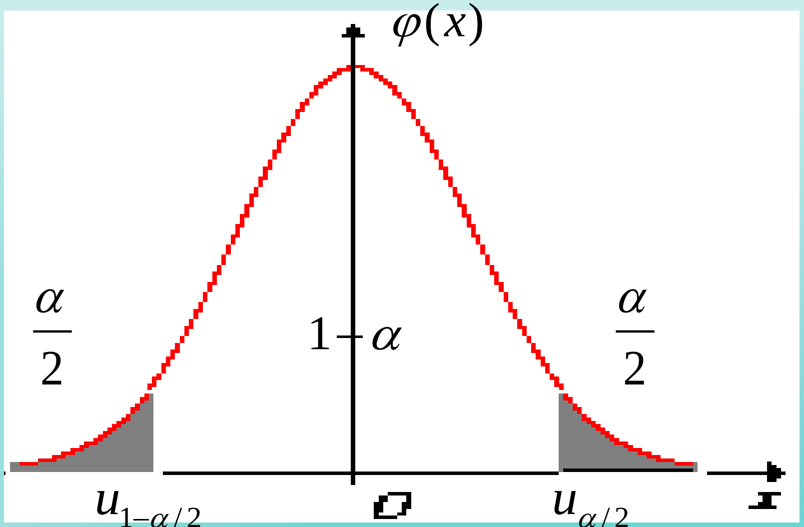
$$\frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}} \sim N(0,1)$$

解： $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是 μ 的优良估计，且

$$\overline{X}\sim N(\mu,\frac{\sigma_0^2}{n})$$

从而 $U=\frac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1)$ 令

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



由标准正态分布的对称性可知

$$u_{1-\alpha/2}=-u_{\alpha/2}$$

从而，前式可化为：

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

$$P\left\{-u_{\frac{\alpha}{2}}\le\frac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}\le u_{\frac{\alpha}{2}}\right\}=1-\alpha$$

从而 $$P\left\{\overline{X}-\frac{\sigma_0}{\sqrt{n}}u_{\frac{\alpha}{2}}\le\mu\le\overline{X}+\frac{\sigma_0}{\sqrt{n}}u_{\frac{\alpha}{2}}\right\}=1-\alpha$$

由此可得， μ的置信度为$1-\alpha$的置信区间为：

$$\left[\overline{X}-\frac{\sigma_0}{\sqrt{n}}u_{\frac{\alpha}{2}},\ \overline{X}+\frac{\sigma_0}{\sqrt{n}}u_{\frac{\alpha}{2}}\right]$$

特别，当 $\sigma_0=1$ ， $\alpha=0.05$ ，样本观测值为 ：

5.1	5.1	4.8	5.0	4.7	5.0	5.2	5.1	5.0

$u_{\alpha/2}=1.96$ ，μ的置信区间为：[4.35，5.65]

估计量的选取

例 设X~N(μ,σ^2)，求参数σ^2的置信度为1-α的置信区间.

分析:

1) 当μ未知时，应选统计量为: S^2

要化至常用分布，由抽样分布定理可知:

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

2) 当μ已知时，应选统计量为: $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$

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

原因：它是最简单的无偏、有效、相合估计量.

未知参数的替换

例　设X~N(μ,σ^2)，σ^2 未知，求参数μ的置信度为1-α 的置信区间.

分析：1. $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 是μ的优良估计，且 $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$

思考：是否仍选统计量 $U = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

令 $P\{u_{1-\alpha/2} \le U \le u_{\alpha/2}\} = 1-\alpha$ 求得置信区间？

不可

因为σ^2 未知，故 U 不是统计量

2. 选一个统计量去替代σ^2：S^2 、 M_2 选哪一个较好？

选S^2

因为它是σ^2的无偏、有效、相合估计

化至常用分布，应为 t 分布，据抽样分布定理有：

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

3. 得 T 的置信区间：$P\left\{t_{1-\alpha/2}(n-1)\le T\le t_{\alpha/2}(n-1)\right\}=1-\alpha$

由分布的对称性，即 $t_{1-\alpha/2}(n-1)=-t_{\alpha/2}(n-1)$ 可化为：

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

4. 代换后可得 μ 的置信区间：

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

比较：

$\sigma^2=\sigma_0^2$ 时， μ的置信区间为

$$\left[\bar{X}-\frac{\sigma_0}{\sqrt{n}}\boxed{u_{\frac{\alpha}{2}}},\ \bar{X}+\frac{\sigma_0}{\sqrt{n}}u_{\frac{\alpha}{2}}\right]$$

σ^2未知 时， μ的置信区间为

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

零件长度的方差

例　从自动机床加工的同类零件中任取16件测得长度值为(单位：mm)

12.15	12.12	12.01	12.28	12.09	12.16	12.03	12.01
12.06	12.13	12.07	12.11	12.08	12.01	12.03	12.06

求方差的估计值和置信区间(α=0.05).

解：设零件长度为$X\sim N\left(\mu,\sigma^2\right)$，用$S^2$作为方差的估计

$$\bar{x}=\frac{1}{16}\sum_{i=1}^{16}x_i=12.08,\quad \sum_{i=1}^{16}(x_i-\bar{x})^2=0.0761$$

故方差的估计值为$\hat{\sigma}^2=\dfrac{1}{15}\sum_{i=1}^{16}(x_i-\bar{x})^2=0.005$

下面计算方差的置信区间：

由于μ未知， S^2是σ^2的优良估计，相应的常用分布为：

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

相应的置信区间为：

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

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

查χ^2分布表可得：$\chi^2_{\alpha/2}(n-1)=\chi^2_{0.025}(15)=\ 27.488$

$$\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.975}(15)=6.262$$

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

$$\chi^2_{\alpha/2}(n-1)=\chi^2_{0.025}(15)=\quad 27.488$$

$$\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.975}(15)=6.262$$

σ^2的置信度为0.95的置信区间为：

$$\left[\frac{0.0761}{27.488},\frac{0.0761}{6.26}\right]\quad 即\quad [0.002768,0.012]$$

比较：σ^2的估计值为 0.005

婴儿体重的估计

例、假定初生婴儿的体重服从正态分布，随机抽取12 名婴儿，测得体重为：（单位：克）

3100, 2520, 3000, 3000, 3600, 3160,
3560, 3320, 2880, 2600, 3400, 2540

试以 95% 的置信度估计初生婴儿的平均体重以及方差.

解：设初生婴儿体重为X克，则 $X \sim N(\mu, \sigma^2)$

(1) 需估计μ,而σ^2 未知：　　α= 0.05，n= 12，

取 $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$，有 $t_{0.025}(11)$= 2.201，

$$\because \quad \overline{X} \approx 3057, \qquad S \approx 375.3$$

μ的置信区间为：$\left[\overline{X} - t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}}\right]$

婴儿体重的估计

代入数据得：

$$\left[3057-\frac{375.3}{\sqrt{12}}\times 2.201, 3057+\frac{375.3}{\sqrt{12}}\times 2.201\right]$$

即 $[\,2818, 3296\,]$

(2) 需估计σ^2 ,而μ未知： 取 $\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$

有 $\chi^2_{0.025}(11)=$ 21.92， $\chi^2_{0.975}(11)=$ 3.816 ,

σ^2的置信区间为： $\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]$

$\because 11\times S^2=1549000$

$\therefore \sigma^2$的置信区间为:$[\frac{1549000}{21.92},\frac{1549000}{3.816}]$

即 $[\,70666, 405922.4\,]$

两稻种产量的期望差的置信区间

例　甲、乙两种稻种分别种在10块试验田中，每块田中甲、乙稻种各种一半. 假设两种稻种产量X、Y 服从正态分布，且方差相等.

10块田中的产量如下表 (单位：公斤)，求两稻种产量的期望差 $\mu_1 - \mu_2$ 的置信区间(α=0.05).

甲	140	137	136	140	145	148	140	135	144	141
乙	135	118	115	140	128	131	130	115	121	125

解：设$X \sim N(\mu_1,\sigma_1^2)$ ，$Y \sim N(\mu_2,\sigma_2^2)$ ，　$\sigma_1^2 = \sigma_2^2 = \sigma^2$ ，要估计$\mu_1 - \mu_2$，取统计量

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

其中，$S_w = \sqrt{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{j=1}^{n_2}(Y_j - \overline{Y})^2} = \sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$

由样本表可计算得：

$\bar{x} = 140.6 \qquad s_1^2 = 16.933 \qquad n_1 = 10$

$\bar{y} = 126.8 \qquad s_2^2 = 71.956 \qquad n_2 = 10$

从而，$S_w = \sqrt{\dfrac{9 \times 16.933 + 9 \times 71.956}{18}} = 6.667$

查t 分布表得： $t_{0.025}(18) =$ 2.1009

可得两稻种产量期望差的置信度为95%的置信区间为：

$$\left[\overline{X} - \overline{Y} - t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)S_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}},\ \overline{X} - \overline{Y} + t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)S_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$$

$$\left[140.6 - 126.8 - 2.10096.667\sqrt{\frac{2}{10}},\ 140.6 - 126.8 - 2.10096.667\sqrt{\frac{2}{10}}\right]$$

即[7.536 , 20.064]

常见的区间估计

$X \sim N(\mu, \sigma^2)$

$P\{w_{1-\alpha/2} \le W \le w_{\alpha/2}\} = 1-\alpha$

被估参数	条件	统计量(枢轴变量)	置信区间
μ	σ^2已知	$U = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left[\bar{X}-\dfrac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}, \bar{X}+\dfrac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\right]$
μ	σ^2未知	$T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\left[\bar{X}-\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1), \bar{X}+\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right]$
σ^2	μ已知	$\chi^2 = \dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$	$\left[\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right]$
σ^2	μ未知	$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left[\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$

$X \sim N(\mu_1, \sigma_1^2)$ $Y \sim N(\mu_2, \sigma_2^2)$ $\boxed{P\{w_{1-\alpha/2} \le W \le w_{\alpha/2}\} = 1-\alpha}$

被估参数	条件	统计量(枢轴变量)
$\mu_1 - \mu_2$	σ_1^2 已知 σ_2^2 已知	$U = \dfrac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$
$\mu_1 - \mu_2$	σ_1^2未知 σ_2^2未知	$T = \dfrac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$ S_w见书
$\dfrac{\sigma_1^2}{\sigma_2^2}$	μ_1未知 μ_2未知	$F = \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \dfrac{\sigma_2^2}{\sigma_1^2}\dfrac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$