§7.4　区间估计

前面我们研究了参数的点估计。也就是回答了问题“参数等于多少？”

现在的新问题：参数在什么范围？

答案的形式：给出一个由统计量定义的区间

（$\underline{\theta}(X_1, X_2, \cdots, X_n)$，$\overline{\theta}(X_1, X_2, \cdots, X_n)$）

第二个问题：参数真的在这个区间中吗？

定义

设总体X的分布函数$F(x;\theta)$ 含有未知参数θ,

对于给定值$\alpha(0<\alpha<1)$，若统计量

$\underline{\theta}=\underline{\theta}(X_1,X_2,\cdots,X_n)$ 和 $\overline{\theta}=\overline{\theta}(X_1,X_2,\cdots,X_n)$

$(\underline{\theta}<\overline{\theta})$ 使得，对于任意 $\theta\in\Theta$

$$P\{\underline{\theta}(X_1,X_2,\cdots,X_n)<\theta<\overline{\theta}(X_1,X_2,\cdots,X_n)\}\geq 1-\alpha$$

则称随机区间 $(\underline{\theta},\ \overline{\theta})$ 是 θ 的置信水平（置信度）为 $1-\alpha$ 的置信区间。

例　设总体$X \sim N(\mu,1)$，μ为未知参数。设$X_1, X_2, \cdots, X_n$是来自X的样本。求μ的置信水平为$1-\alpha$的置信区间.

解：$\overline{X}$是μ 的无偏估计。且有

$\overline{X} - \mu \sim N(0,1/n)$，所以 $\dfrac{\overline{X} - \mu}{1/\sqrt{n}} \sim N(0,1)$

按标准正态分布的上α 分位点的定义,

$$P\left\{-z_{0.025} < \frac{\overline{X}-\mu}{1/\sqrt{n}} < z_{0.025}\right\} = 0.95$$

$$P\left\{ \overline{X} - \frac{1}{\sqrt{n}} z_{0.025} < \mu < \overline{X} + \frac{1}{\sqrt{n}} z_{0.025} \right\} = 0.95$$

（$\overline{X} - \frac{1}{\sqrt{n}} z_{0.025}$， $\overline{X} + \frac{1}{\sqrt{n}} z_{0.025}$）是$\mu$的一个置信度为0.95的置信区间。

注意：置信水平为$1-\alpha$的置信区间并不唯一。

以上述例子来说，对于给定的$\alpha = 0.05$，也有

$$P\left\{-z_{0.04} < \frac{\overline{X}-\mu}{1/\sqrt{n}} < z_{0.01}\right\} = 0.95$$

即$P\left\{\overline{X} - \frac{1}{\sqrt{n}} z_{0.01} < \mu < \overline{X} + \frac{1}{\sqrt{n}} z_{0.04}\right\} = 0.95$

原则：对于给定的置信度，尽量选择短的那个