§ 8.3　区间估计

引例： 已知 $X \sim N(\mu, \sigma^2)$，　σ^2已知，　μ未知

$x_1, x_2, \ldots, x_n$ 是一组样本值。

估计未知参数μ

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{\sqrt{\sigma^2 / n}} \sim N(0, 1)$$

则 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right|>z_{1-\alpha/2}\right)=\alpha$$

故 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right|\le z_{1-\alpha/2}\right)=1-\alpha$$

即 $$P\left(\bar{X}-z_{1-\alpha/2}\sqrt{\sigma^2/n}\le\mu\le\bar{X}+z_{1-\alpha/2}\sqrt{\sigma^2/n}\right)=1-\alpha$$

称随机区间 $$\left(\bar{X}-z_{1-\alpha/2}\sqrt{\sigma^2/n},\ \bar{X}+z_{1-\alpha/2}\sqrt{\sigma^2/n}\right)$$

为未知参数 μ 的置信度为1-α的置信区间.

置信区间的意义

$$\left(\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n},\quad \bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n} \right)$$

反复抽取容量为n的样本，都可得到一个区间,这个区间可能包含未知参数 μ 的真值,也可能不包含未知参数的真值, 包含真值的区间占1-α.

$\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n}$ —— μ 的置信下限

$\bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n}$ —— μ 的置信上限

$1-\alpha$ —— 置信度

置信区间的定义

设 θ 是一个待估计的参数, α 是一给定的数, $(0 < \alpha < 1)$. 若能找到两个统计量

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

使得 $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$

则称随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 为参数 θ 的置信度为 $1 - \alpha$ 的置信区间, 分别称 $\hat{\theta}_1, \hat{\theta}_2$ 为置信下限与置信上限, $1 - \alpha$ 称为置信水平或置信度.

几点说明 $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$

- 置信区间的长度 $\hat{\theta}_2 - \hat{\theta}_1$ 反映了估计的精度

  $\hat{\theta}_2 - \hat{\theta}_1$ 越小，估计的精度越高.

例：$(\bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n}) - (\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n}) = 2 \cdot z_{1-\alpha/2}\sqrt{\sigma^2/n}$

- α 反映了估计的可靠程度，α 越小, 越可靠.

  α 越小，1- α 越大, 估计的可靠程度越高,但

  这时，$\hat{\theta}_2 - \hat{\theta}_1$ 往往增大, 因而估计的精度降低.

- α 确定后，置信区间 的选取方法不唯一，常选长度最小的一个.

求置信区间的步骤

- 寻找一个样本的函数

$$g(X_1, X_2, \cdots, X_n, \theta)$$ —— 称为**枢轴量**

例如 $$\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} = g(X_1, X_2, \cdots, X_n, \mu) \sim N(0,\ 1)$$

它含有待估参数，不含其它未知参数，它的分布已知，且分布不依赖于待估参数 (常由 θ 的点估计出发考虑).

- 给定置信度 $1-\alpha$，定出两个常数 a , b ,使得

$$P(a < g(X_1, X_2, X_n, \theta) < b) = 1-\alpha$$

例如

$$P\left(\left| \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \right| \le z_{1-\alpha/2} \right) = 1-\alpha$$

- 由 $a < g(X_1, X_2, X_n, \theta) < b$ 解出

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

得置信区间 $(\underline{\theta}, \ \bar{\theta})$

例如 $\left(\bar{X} - z_{1-\alpha/2}\sqrt{\sigma^2/n}, \bar{X} + z_{1-\alpha/2}\sqrt{\sigma^2/n} \right)$

置信区间常用公式

(一) 一个正态总体$X \sim N(\mu, \sigma^2)$的情形

(1) 方差σ^2已知, μ的置信区间

$$\left(\bar{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) \quad \cdots\cdots\cdots \quad (1)$$

推导 由 $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ 选取枢轴量

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

由 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right| > z_{1-\alpha/2}\right) = \alpha$$

得 $$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right| \le z_{1-\alpha/2}\right) = 1-\alpha$$

即 $$P\left(\overline{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

故 μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\quad \overline{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

(2) 方差σ^2未知，μ的置信区间

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

推导　选取枢轴量 $T=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim T(n-1)$

由 $P\left(\left|\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\right|\geq t_{1-\frac{\alpha}{2}}(n-1)\right)=\alpha$

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=S^2$$

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

$$P\left(\bar{X}-t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}<\mu<\bar{X}+t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)=1-\alpha$$

故 μ 的置信区间为

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

用到：设总体$X \sim N(\mu，\sigma^2)$，$(X_1，X_2，\dots\dots X_n)$是总体 X 的样本，S^2为样本方差，证明：

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

证明：

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

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

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}\Big/ n-1}} \sim t(n-1)$$

(3) 当 μ 已知时，方差 σ^2 的 置信区间

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

由概率

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

得 σ^2 的置信度为 $1-\alpha$ 置信区间为

$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)},\ \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right) \quad \cdots\cdots\cdots \quad (3)$$

(4) 当 μ 未知时，方差 σ^2 的置信区间

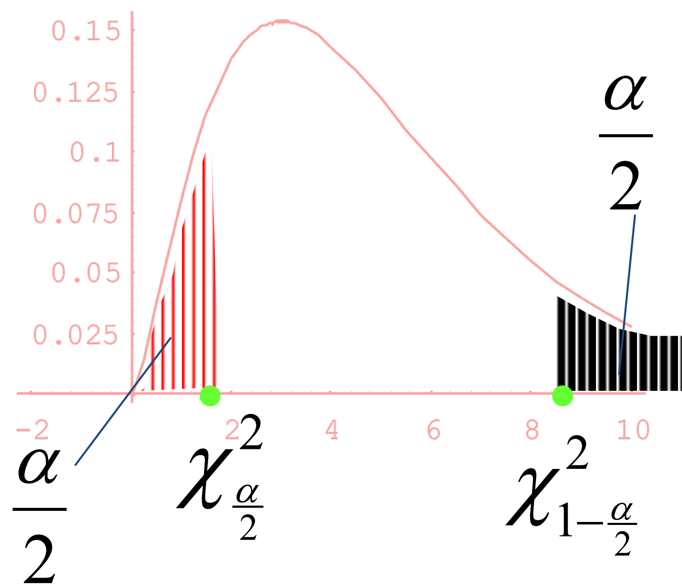
P177 Th3

选取 $K=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ 则由

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

得 σ^2 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\ \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right) \cdots\cdots\cdots \quad (4)$$



例1 某工厂生产一批滚珠，其直径 X 服从正态分布$N(\mu,\sigma^2)$，现从某天的产品中随机抽取6件，测得直径为

15.1，14.8，15.2，14.9，14.6，15.1

(1) 若$\sigma^2=0.06$，求 μ 的置信度为95%的 置信区间;

(2) 若σ^2未知，求 μ 的置信度为95%的置信区间;

(3) 求方差σ^2的置信度为95%的置信区间.

解 (1)若$\sigma^2=0.06$，求 μ 的置信度为95%的置信区间;

$$\left(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\quad \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

$$\sigma^2=0.06 \qquad n=6 \qquad 1\text{-}\alpha=0.95$$

$$z_{1-\frac{\alpha}{2}}=z_{1-0.025}=1.96$$

由给定数据算得 $\bar{x}=\frac{1}{6}\sum_{i=1}^{6}x_i=14.95$

由公式(1) 得 μ 的置信区间为

$$\left(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\quad \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

$$=(14.75,\quad 15.15)$$

(2) 若σ^2未知,求 μ 的置信度为95%的置信区间

$$\left(\bar{x}-\frac{s}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1),\quad \bar{x}+\frac{s}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)\right)$$

查表得 $t_{1-0.025}(5)=2.5706$

由给定数据算得 $\bar{x}=14.95$

$$s^2=\frac{1}{5}\left(\sum_{i=1}^{6}x_i^2-6\bar{x}^2\right)=0.051. \quad s=0.226$$

由公式 (2) 得 μ 的置信区间为

$$\left(\bar{x}-\frac{s}{\sqrt{6}}t_{0.025}(5),\quad \bar{x}+\frac{s}{\sqrt{6}}t_{0.025}(5)\right)=(14.71,\quad 15.187)$$

(3) 求方差σ^2的置信度为95%的置信区间

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

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = 0.051.$$

查表得 $\chi^2_{0.025}(5) = 0.831,\quad \chi^2_{0.975}(5) = 12.833$

由公式 (4) 得 μ的置信区间为

$$\left(\frac{5s^2}{\chi^2_{0.975}(5)}, \quad \frac{5s^2}{\chi^2_{0.025}(5)} \right) = (0.0199, \quad 0.3069)$$

(二) 单侧置信区间

定义 对于给定的 $\alpha\,(0<\alpha<1)$，θ 是待估参数

$(X_1, X_2, \cdots, X_n)$ 是总体 X 的样本,

若能确定一个统计量

$$\underline{\theta}=\underline{\theta}\,(X_1, X_2, \cdots, X_n) \quad (\text{或 } \overline{\theta}=\overline{\theta}\,(X_1, X_2, \cdots, X_n))$$

使得 $P(\theta>\underline{\theta})=1-\alpha \quad (\text{或 } P(\theta<\overline{\theta})=1-\alpha)$

则称 $(\underline{\theta}, +\infty) \quad (\text{或 } (-\infty, \overline{\theta}))$

为置信度为1 - α 的单侧置信区间.

$\underline{\theta}$ —— 单侧置信下限 $\overline{\theta}$ —— 单侧置信下限

例3 已知灯泡的寿命X服从正态分布，从灯泡中随机地抽取 5 只作寿命试验，测得寿命(以小时记) 为

$$1050,\quad 1100,\quad 1120,\quad 1250,\quad 1280$$

求灯泡寿命均值的置信度为0.95的单侧置信下限与灯泡寿命方差的置信度为0.95的单侧置信上限

解 $X \sim N(\mu,\sigma^2)$， μ,σ^2未知

(1) 选取枢轴量 $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

$$故 P\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{1-\alpha}(n-1)\right)=1-\alpha$$

$$即 P\left(\mu \ge \bar{X}-t_{1-\alpha}(n-1)\cdot S/\sqrt{n}\right)=1-\alpha$$

$n=5$, $\bar{x}=1160$, $s^2=\dfrac{1}{4}(\sum_{i=1}^{5} x_i^2-5\bar{x})=9950$

$$t_{1-\alpha}(n-1)=t_{1-0.05}(4)=2.1318$$

$$\underline{\mu}=\bar{x}-t_{1-0.05}(4)\times\frac{s}{\sqrt{5}}=1064.9$$

(2) 选取枢轴量 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

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

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

$$n = 5, \quad s^2 = \frac{1}{4}\left(\sum_{i=1}^{5} x_i^2 - 5\overline{x}\right) = 9950$$

$$\chi^2_{0.05}(4) = 0.711$$

$$\overline{\sigma^2} = \frac{4s^2}{\chi^2_{0.05}(4)} = 55977$$