典 型 例 题 讲 解

例 1.1 设$X_1, X_2, \ldots, X_6$是来自正态总体$N(\mu, 4)$的样本, 记

$$Y = a(X_1 + X_2 - 2\mu)^2 + b(2X_3 + 3X_4 - 4X_5 - X_6)^2$$

试确定a, b 使Y服从 $\chi^2(k)$ 分布, 并求出k 值.

解 $\because\ X_1 + X_2 \sim N(2\mu, 8)$, $\therefore\ Y_1 = \dfrac{X_1 + X_2 - 2\mu}{\sqrt{8}} \sim N(0,1)$.

$\because\ 2X_3 + 3X_4 - 4X_5 - X_6 \sim N(0, 120)$, $\therefore\ Y_2 = \dfrac{2X_3 + 3X_4 - 4X_5 - X_6}{\sqrt{120}} \sim N(0,1)$.

由样本的独立性可知，Y_1, Y_2相互独立.

由卡方分布的定义可知 $Y = Y_1^2 + Y_2^2 \sim \chi^2(2)$.

所以 $a = 1/8$, $b = 1/120$ ，$k = 2$.

例1.2 设 $X_1, X_2, \ldots, X_7$ 是来自总体$N(\mu, \sigma^2)$ 的样本,

$$\overline{X}=\frac{1}{6}\sum_{k=1}^{6}X_k,\quad S^2=\frac{1}{5}\sum_{k=1}^{6}(X_k-\overline{X})^2,\quad Y=\sqrt{\frac{6}{7}}\frac{X_7-\overline{X}}{S}.$$

求统计量Y的概率分布.

解 $X_7-\overline{X}\sim N(0,\frac{7}{6}\sigma^2)$, $\quad \dfrac{X_7-\overline{X}}{\sqrt{7/6}\,\sigma}\sim N(0,1).\quad \dfrac{5S^2}{\sigma^2}\sim\chi^2(5).$

$\dfrac{X_7-\overline{X}}{\sqrt{7/6}\,\sigma}$ 与 $\dfrac{5S^2}{\sigma^2}$ 独立. 由 t 分布的定义可知

$$\frac{(X_7-\overline{X})/(\sqrt{7/6}\,\sigma)}{\sqrt{\dfrac{5S^2/\sigma^2}{5}}}=\sqrt{\frac{6}{7}}\frac{X_7-\overline{X}}{S}=Y\sim t(5).$$

例1.3 设总体 $X \sim N(20, 5^2)$，总体 $Y \sim N(10, 2^2)$，X和Y独立，从两总体中分别抽取容量为 10 和 8 的样本，求：

1） $P(\overline{X} - \overline{Y} > 6)$， 2） $P(S_1^2 / S_2^2 < 23)$.

解 1)

$$U = \frac{(\bar{X} - \bar{Y}) - (20 - 10)}{\sqrt{\frac{5^2}{10} + \frac{2^2}{8}}} = \frac{\bar{X} - \bar{Y} - 10}{\sqrt{3}} \sim N(0,1),$$

$$P(\bar{X} - \bar{Y} > 6) = P\left(\frac{\bar{X} - \bar{Y} - 10}{\sqrt{3}} > \frac{6 - 10}{\sqrt{3}}\right) = P(U > -2.31).$$

$$= \quad 1 - P(U \le -2.31) = 1 - \Phi(-2.31) = \Phi(2.31) = 0.9896.$$

例1.4 设总体 $X \sim N(20, 5^2)$, 总体 $Y \sim N(10, 2^2)$, X和Y独立，从两总体中分别抽取容量为 10 和 8 的样本，求

1） $P(\overline{X}-\overline{Y}>6)$, 2） $P(S_1^2/S_2^2<23)$.

解 2) $F=\dfrac{S_1^2/S_2^2}{5^2/2^2}\sim F(9,7)$.

$$P(S_1^2/S_2^2<23)=P\left(\frac{S_1^2/S_2^2}{5^2/2^2}<\frac{23}{5^2/2^2}\right)=P(F<3.68) = 1-P(F\geq 3.68)$$

查表得: $F_{0.05}(9,7)=3.68$

所以 $P(S_1^2/S_2^2<23)=1-0.05=0.95.$

例2.1 设总体 X 的分布函数为 $F(x,a,b)=\begin{cases}1-(\frac{a}{x})^b, & x\ge a;\\ 0, & x<a.\end{cases}$

其中参数 $a>0, b>1$, $X_1, X_2, \cdots X_n$为样本.

1) 当 $a=1$时，求参数b的极大似然估计量;

1) 当 $b=2$时，求参数a的极大似然估计量.

解 1）当 $a=1$时， X 的概率密度为 $f(x,1,b)=\begin{cases}\dfrac{b}{x^{b+1}}, & x\ge 1;\\ 0, & x<1.\end{cases}$

当$x_i \ge 1(i-1,2,\cdots,n)$时， 参数$b$的似然函数为

$$L(b)=\prod_{i=1}^{n} f(x_i,1,b)=\prod_{i=1}^{n}\frac{b}{x_i^{b+1}}=b^n(\prod_{i=1}^{n} x_i)^{-b-1}$$

$$\ln L(b) = n\ln b - (b+1)\sum_{i=1}^{n}\ln x_i$$

对b求导，并令其等于零

$$\frac{d\ln L(b)}{db} = \frac{n}{b} - \sum_{i=1}^{n}\ln x_i = 0$$

解得 $\hat{b} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i$

所以参数b的极大似然估计量为 $\hat{b} = \frac{1}{n}\sum_{i=1}^{n}\ln X_i$

解 2）当 $b=2$ 时， X 的概率密度为 $f(x,a,2)=\begin{cases}\dfrac{2a^2}{x^3}, & x\geq a;\\ 0, & x<a.\end{cases}$

当$x_i\geq a(i=1,2,\cdots,n)$时，参数$a$的似然函数为

$$L(a)=\prod_{i=1}^{n}f(x_i,a,1)=\prod_{i=1}^{n}\frac{2a^2}{x_i^3}=\frac{2^n a^{2n}}{(x_1x_2\cdots x_n)^3}$$

a越大，$L(a)$越大， 由于 $a\leq x_i(i=1,2,\cdots,n)$

所以a的极大似然估计值为 $\hat{a}=\min\{x_1,x_2,\cdots,x_n\}$

由此可得a的极大似然估计量为 $\hat{a}=\min\{X_1,X_2,\cdots,X_n\}$

例2. 2 设总体 X 的分布律为

X	1	2	3
p	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $\theta\in(0,1)$ 为未知参数. 已知取得了样本值： $x_1=1,\ x_2=2,\ x_3=1$

求的矩估计值和极大似然估计值.

解 1）$E(X)=3-2\theta$. 令 $3-2\theta=A_1=\bar{X}$,

得θ的矩估计量为 $\hat{\theta}=\dfrac{3-\bar{X}}{2}$

所以θ的矩估计值为

$$\hat{\theta}=\frac{3-\bar{x}}{2}=\frac{3-4/3}{2}=\frac{5}{6}.$$

（续）设总体 X 的分布律为

X	1	2	3
p	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $\theta\in(0,1)$ 为未知参数. 已知取得了样本值： $x_1=1,\ x_2=2,\ x_3=1$

求的矩估计值和极大似然估计值.

解 2）似然函数 $L(\theta)=\prod_{k=1}^{3}p(x_k,\theta)$

$$=p(1,\theta)p(2,\theta)p(1,\theta)=2\theta^5(1-\theta)$$

$\ln L(\theta)=\ln 2+5\ln\theta+\ln(1-\theta)$, 令 $\dfrac{d\ln L(\theta)}{d\theta}=\dfrac{5}{\theta}-\dfrac{1}{1-\theta}=0$

解得θ的极大似然估计值为 $\hat{\theta}=\dfrac{5}{6}$.

例2.3 从一批产品中放回抽样，依次抽取60件样品，发现其中有3件次品，用极大似然法估计这批产品的次品率.

解 设这批产品的次品率为 p ，随机变量 X 表示任一次抽样时取得次品的件数，则 X 服从两点分布 B(1, p).

X 的分布律为 $p(x,p)=P(X=x)=p^{x}(1-p)^{1-x}$， $x=0,1.$

似然函数

$$L(p)=\prod_{i=1}^{n}p(x_i,p)=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}.$$

两边取对数得 $$\ln L(p)=(\sum_{i=1}^{n}x_i)\ln p+(n-\sum_{i=1}^{n}x_i)\ln(1-p).$$

令 $$\frac{d\ln L(p)}{dp}=\frac{\sum_{i=1}^{n}x_i}{p}+\frac{n-\sum_{i=1}^{n}x_i}{1-p}=0.$$

解得 p 的极大似然估计值为 $$\hat{p}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}.$$

由已知 $n=60,\quad \sum_{i=1}^{n}x_i=3.$

所以这批产品的次品率 p 的极大似然估计值为

$$\hat{p}=\frac{3}{60}=5\%.$$

例2.4 设总体X的服从 $N(0,\sigma^2)$分布，$X_1, X_2, \cdots X_n$ 为其样本，求σ^2的极大似然估计量，并判断它是否为σ^2的最小方差无偏估计量.

解 1）X的概率密度为 $f(x,\sigma^2)=\dfrac{1}{\sqrt{2\pi\sigma}}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}$

似然函数

$$L(\sigma^2)=\prod_{i=1}^{n} f(x_i,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma}}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}=(2\pi\sigma^2)^{-\frac{n}{2}}\mathrm{e}^{-\sum\limits_{i=1}^{n}\frac{x_i^2}{2\sigma^2}}$$

$$\ln L(\sigma^2)=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2$$

令 $$\frac{\ln L(\sigma^2)}{d\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n} x_i^2 = 0$$

解得θ 的极大似然估计值为 $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

所以θ 的极大似然估计量为 $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

2) 由于 $$E(\hat{\sigma}^2) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2) = E(X^2)$$

$$= D(X) + E^2(X) = D(X) = \sigma^2$$

所以$\hat{\sigma}^2$是σ^2的无偏估计量.

$$D(\hat{\sigma}^2)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i^2)=\frac{1}{n}D(X^2)$$

$$=\frac{1}{n}[E(X^4)-E^2(X^2)]=\frac{1}{n}(3\sigma^4-\sigma^4)=\frac{2\sigma^4}{n}$$

$$\ln f(X,\sigma^2)=-\ln(\sqrt{2\pi})-\frac{1}{2}\ln\sigma^2-\frac{X^2}{2\sigma^2}$$

$$I(\sigma^2)=-E\left[\frac{\partial^2\ln f(X;\sigma^2)}{\partial(\sigma^2)^2}\right]=-E\left(\frac{1}{2\sigma^4}-\frac{X^2}{\sigma^6}\right)=\frac{1}{2\sigma^4}$$

估计的效率 $e_n=\dfrac{1}{D(\hat{\sigma}^2)nI(\sigma^2)}=1$

所以 $\hat{\sigma}^2$ 是 σ^2 的最小方差无偏估计量.

例2.5 设总体X服从几何分布：

$$P(X=x)=p(1-p)^{x-1}\ (x=1,2,\ldots)$$

$X_1,X_2,\cdots X_n$ 为其样本，$\theta=\dfrac{1}{p}$，求θ的矩估计量，并判断它是否为θ的最小方差无偏估计量.

解 $$E(X)=\sum_{x=1}^{+\infty}xp(1-p)^{x-1}=p\sum_{k=1}^{+\infty}x(1-p)^{x-1}=\frac{1}{p}=\theta$$

所以 θ 的矩估计量为 $\hat{\theta}=\bar{X}$，显然$\hat{\theta}$是θ的无偏估计量.

$$\boxed{e_n=\frac{1}{D(\hat{\theta})nI(\theta)}}$$

$\hat{\theta}$ 是θ的最小方差无偏估计量，证明如下：

$$D(\hat{\theta})=D(\bar{X})=\frac{1}{n}D(X)=\frac{1-p}{np^2}=\frac{\theta^2-\theta}{n}$$

$$p(X,\theta)=\frac{1}{\theta}(1-\frac{1}{\theta})^{X-1}$$

$$\ln p(X,\theta)=-\ln\theta+(X-1)[\ln(\theta-1)-\ln\theta]$$

$$I(\theta)=-E[\frac{\partial^2}{\partial\theta^2}\ln p(X;\theta)]=-E[\frac{1}{\theta^2}-\frac{2\theta-1}{(\theta^2-\theta)^2}(X-1)]$$

$$=-\frac{1}{\theta^2}+\frac{2\theta-1}{(\theta^2-\theta)^2}\bullet(\theta-1)=-\frac{1}{\theta^2}+\frac{2\theta-1}{\theta^2(\theta-1)}=\frac{1}{\theta^2-\theta}$$

估计的效率 $e_n=\dfrac{1}{D(\hat{\theta})nI(\theta)}=\dfrac{1}{\frac{\theta^2-\theta}{n}n\frac{1}{\theta^2-\theta}}=1$

所以 $\hat{\theta}$ 是θ的最小方差无偏估计量.

例2.6 已知 $X_1, X_2, \ldots, X_n$是取来自总体 $N(1, \sigma^2)$ 的样本，记

$$Y = X_1 + X_2 + X_3 - 3, Z = X_2 - X_3, T = \frac{Y}{|Z|}$$

（1）证明Y与Z相互独立；（2）求常数c, 使得 $cT \sim t(n)$, 并求n的值.

解 （1）证明：由于（X_2, X_3 ）服从二维正态分布，根据二维正态分布的性质可知（$X_2 + X_3$, $X_2 - X_3$ ）也服从二维正态分布

由 $Cov(X_2 + X_3, X_2 - X_3) = D(X_2) - D(X_3) = 0$

可知与 $X_2 - X_3$不相关， 所以$X_2 + X_3$ 与 $X_2 - X_3$相互独立，

由X_1 与 $X_2 - X_3$相互独立，所以 $X_1 + X_2 + X_3 - 3$与 $X_2 - X_3$相互独立，

即Y与Z相互独立.

（2）由独立随机变量的性质可知 $Y \sim N(0,3\sigma^2), Z \sim N(0,2\sigma^2)$,

所以

$$\frac{Y}{\sqrt{3}\sigma} \sim N(0,1), \frac{Z}{\sqrt{2}\sigma} \sim N(0,1), \left(\frac{Z}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1)$$

由 t 分布的定义可知

$$\frac{Y/\sqrt{3}\sigma}{\sqrt{\left(Z/\sqrt{2}\sigma\right)^2/1}} \sim t(1)$$

即 $\sqrt{\dfrac{2}{3}}\dfrac{Y}{|Z|} = \sqrt{\dfrac{2}{3}}T \sim t(1)$

所以 $c = \sqrt{\dfrac{2}{3}}, n = 1.$

例2.7 设 $X_1, X_2, \ldots X_n$为总体 X 的一个样本，$E(X)=\mu, D(X)=\sigma^2$，

$\hat{\theta}=c\sum_{k=1}^{n-1}(X_{k+1}-X_k)^2$ 为σ^2的无偏估计， 则 $c=$ ________.

解 $$E\left[\sum_{k=1}^{n-1}(X_{k+1}-X_k)^2\right]=\sum_{k=1}^{n-1}E(X_{k+1}-X_k)^2$$

$$=\sum_{k=1}^{n-1}\left[D(X_{k+1}-X_k)+E^2(X_{k+1}-X_k)\right]=\sum_{k=1}^{n-1}\left[2\sigma^2+0^2\right]=2(n-1)\sigma^2.$$

令 $E(\hat{\theta})=\sigma^2$，即 $c\cdot 2(n-1)\sigma^2=\sigma^2$， $\therefore\ c=\dfrac{1}{2(n-1)}$.

解　方法2

$$E\left[\sum_{k=1}^{n-1}(X_{k+1}-X_k)^2\right]=\sum_{k=1}^{n-1}E\,(X_{k+1}-X_k)^2$$

$$=\sum_{k=1}^{n-1}E\left[(X_{k+1}-\mu)-(X_k-\mu)\right]^2$$

$$=\sum_{k=1}^{n-1}\{\,E(X_{k+1}-\mu)^2-2E[\,(X_k-\mu)(X_{k+1}-\mu)\,]+E(X_k-\mu)^2\}$$

$$=\sum_{k=1}^{n-1}\left[\sigma^2+2\times 0+\sigma^2\right]=2(n-1)\sigma^2. \qquad \therefore \quad c=\frac{1}{2(n-1)}.$$

例2.8 某厂生产的一批金属材料，其抗弯强度服从正态分布，从这批金属材料中抽取11件测试，得抗弯强度为

42.5，42.7，43.0，42.3，43.4，44.8，44.0，43.8，44.1，43.9，43.7.

1）求平均抗弯强度μ 的置信水平为0.95 的置信区间；

2）求抗弯强度标准差σ的置信水平为0.90 的置信区间.

解 1）σ^2未知，μ 的置信水平为$1-\alpha$ 的置信区间为 $\left(\bar{x} \pm t_{\alpha/2}(n-1)\cdot s/\sqrt{n}\right)$

其中 $n=11, \alpha/2 = 0.025$, 计算得： $\bar{x} = 43.4, s^2 = 0.523, t_{0.025}(10) = 2.228.$

区间半径为 $t_{\alpha/2}(n-1)\cdot s/\sqrt{n} = 2.228 \times \sqrt{0.523/11} = 0.49.$

所以μ 的置信水平为0.95 的置信区间为 $(42.91, 43.89)$

2）σ的置信水平为1-α 的置信区间为

$$\left(\sqrt{\frac{n-1}{\chi^2_{\alpha/2}(n-1)}}\,s,\quad \sqrt{\frac{n-1}{\chi^2_{1-\alpha/2}(n-1)}}\,s\right).$$

其中 $n=11, \alpha/2=0.05$，查表得：

$$\chi^2_{0.05}(10)=18.307,\ \chi^2_{0.95}(10)=3.940.$$

$$\sqrt{\frac{n-1}{\chi^2_{\alpha/2}(n-1)}}\,s=0.53,\quad \sqrt{\frac{n-1}{\chi^2_{1-\alpha/2}(n-1)}}\,s=1.15.$$

所以 σ的置信水平为 0.90 的置信区间为（0.53，1.15）

例3. 1一个手机生产厂家在其宣传广告中声称他们生产的某种品牌的手机待机时间的平均值至少为71. 5小时，一质检部门抽查了该厂生产的此品牌的手机6部，得到的待机时间为：69，68，72， 70，66，75，设手机的待机时间 $X \sim N(\mu, \sigma^2)$, 由这些数据能否说明其广告有欺骗消费者的嫌疑（$\alpha = 0.05$）？

解 σ^2 未知，检验 H_0: $\mu \geq \mu_0$，H_1: $\mu < \mu_0$, 其中 $\mu_0 = 71.5$.

检验统计量 $T = \dfrac{\overline{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1)$. 拒绝域为 $t = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}} < -t_\alpha(n-1)$.

由题设可得 $\overline{x} = 70, s^2 = 10, n = 6, \alpha = 0.05$. 查表得 $t_{0.05}(5) = 2.015$.

由于 $t = \dfrac{70 - 71.5}{\sqrt{10} / \sqrt{6}} = -1.16 > -t_{0.05}(5) = -2.015$.

所以接受假设 H_0 ，即不能认为其广告有欺骗消费者的嫌疑.

例3.2 有一种新安眠剂，据说在一定剂量下能比某种旧安眠剂平均增加睡眠时间3小时，为了检验新安眠剂的这种说法是否正确，收集到一组使用新安眠剂的睡眠时间(单位:小时):

26.7， 22.0， 24.1， 21.0， 27.2， 25.0， 23.4.

根据资料用某种旧安眠剂时平均睡眠时间为20.8小时，假设用安眠剂后睡眠时间服从正态分布，试问这组数据能否说明新安眠剂的疗效?

解 检验假设 H_0: $u \le 23.8$, H_1: $u > 23.8$

这是正态总体方差未知，关于均值的检验假设.

检验统计量为 $T = \dfrac{\bar{X} - u_0}{s/\sqrt{n}}$ 拒绝域为 $t = \dfrac{\bar{x} - u}{s/\sqrt{n}} > t_\alpha(n-1)$

$$\bar{x}=\frac{1}{7}\times(26.7+22+24.1+21.0+27.2+25+23.4)=24.2$$

$$s^2=\frac{1}{n\text{-}1}\sum_{i=1}^{n}(x_i-\bar{x})^2=5.27,$$

由于 $t=\dfrac{\bar{x}-23.8}{s/\sqrt{n}}=0.46<t_{0.05}(6)=1.943,$

所以接受H_0原假设，

即不能认为这组数据能说明新安眠剂的疗效.