2012—2013年 数理统计 DATE: / / /

1. X_1, X_2, X_3 是来自总体的样本，服从 $N(0,4)$，$Y_1 = X_1-2X_2+X_3$，$Y_2=X_1+X_2+X_3$，$Y_3 = X_1-X_3$，求 $P\{Y_1^2 < 37Y_2^2 + 55.5Y_3^2\}$

解：由已知得
$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

由于 X_1, X_2, X_3 相互独立 $\Rightarrow$ Y_1, Y_2, Y_3 相互独立

$Y_1 \sim N(0,24)$；$Y_2 \sim N(0,12)$，$Y_3 \sim N(0,8)$

得：$\frac{Y_1^2}{24} \sim \chi^2(1)$，$\frac{Y_2^2}{12} \sim \chi^2(1)$，$\frac{Y_3^2}{8} \sim \chi^2(1)$

$$\therefore \frac{Y_1^2/24}{\left(\frac{Y_2^2}{12}+\frac{Y_3^2}{8}\right)/2} \sim F(1,2)$$

$$\therefore P\{Y_1^2 < 37Y_2^2 + 55.5Y_3^2\} = P\left\{\frac{Y_1^2/24}{\left(\frac{Y_2^2}{12}+\frac{Y_3^2}{8}\right)/2} < \frac{444}{12}\right\}$$

$$= P\{F(1,2) < 37\} = 0.95$$

2. 总体X服从均匀分布 $U(0,\theta)$，θ的先验分布服从于 $U(0,1)$，求θ的贝叶斯估计。

解：由 $\theta \sim U(0,1)$，得 $\pi(\theta) = \begin{cases} 1, & 0<\theta<1 \\ 0, & 其它 \end{cases}$

由 $X \sim U(0,\theta)$，得 $f(x) = \begin{cases} \frac{1}{\theta} & 0<x<\theta \\ 0, & 其它 \end{cases}$

$\therefore h(y \mid x_1, \cdots, x_n) \propto \pi(y) f(x_1, \cdots, x_n; y) = \left(\frac{1}{y}\right)^n = y^{(-n+1)-1} \cdot (1-y)^{1-1}$

$\therefore h(y|x) \sim \beta(1-n, 1)$

$\therefore E(\theta|x) = \bar{\theta} = \frac{1-n}{2n}$

3. $X \sim N(0,\delta^2)$，样本为 $X_1, X_2, \cdots, X_n$

(1). 求 δ^2 极大似然估计。

(2) 求均方误差是多少？

解：(1) $L(\delta^2) = \left(\frac{1}{\sqrt{2\pi}\delta}\right)^n e^{-\frac{\sum_{i=1}^{n} x_i^2}{2\delta^2}} = (2\pi\delta^2)^{-\frac{n}{2}} \cdot e^{-\frac{\sum_{i=1}^{n} x_i^2}{2\delta^2}}$

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

$\frac{d\ln L(\delta^2)}{d\delta^2} = -\frac{n}{2}\cdot\frac{2\pi}{2\pi\delta^2} + \frac{\sum_{i=1}^{n} x_i^2}{2\delta^4}$

令 $\frac{d\ln L(\delta^2)}{d\delta^2} = 0$，得：$\hat{\delta}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$　　$D(X) = E(X^2) - [E(X)]^2$

(2). $MSE = E(\hat{\delta}^2 - \delta^2)^2$　　$MSE(\varphi) = E[\varphi - g(\theta)]^2$

$= D(\hat{\delta}^2 - \delta^2) + [E(\hat{\delta}^2 - \delta^2)]^2$　　$= E[\varphi - E\varphi + (E\varphi - g(\theta))]^2$

$= D(\hat{\delta}^2) + [E(\hat{\delta}^2 - \delta^2)]^2$　　$= D(\varphi) + (E\varphi - g(\theta))^2$

$= \left(\frac{n-1}{n}\right)^2 D(S^2) + \left[\frac{n-1}{n}ES^2 - \delta^2\right]^2$　　$E(S^2) = \sigma^2$

$= \left(\frac{n-1}{n}\right)^2 \frac{2\delta^4}{n-1} + \frac{1}{n^2}\delta^4$　　$D(S^2) = \frac{2\sigma^4}{n-1}$

$= \frac{2n-1}{n^2}\delta^4$

4. $X_1, X_2, \ldots, X_n$ 是一个总体样本，服从 $N(\theta, \theta)$，$\theta > 0$ 未知，求：θ 在置信水平 $1-\alpha$ 下区间估计。

(1). 当把问题看成方差未知，对均值 θ 进行估计的问题时，

由 $\bar{X} \sim N(\theta, \frac{\theta}{n})$ 有：$\frac{\bar{X}-\theta}{\sqrt{\theta/n}} \sim N(0,1)$，$\frac{(n-1)S^2}{\theta} \sim \chi^2(n-1)$

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

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

所求区间为：$\left[\bar{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right]$

② 当把~~问题看~~成均值未知，对方差θ进行估计时.

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

$P\left\{\chi_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)S^2}{\theta} < \chi^2_{\frac{\alpha}{2}}(n-1)\right\} = 1-\alpha$，~~得~~

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

5. 两台机器服从正态分布，且方差已知 δ_1^2, δ_2^2。μ_1, μ_2 为两台机器的平均产量。$X_1, \cdots, X_{n_1}$ 为甲机器样本 $Y_1 \cdots Y_{n_2}$ 为乙机器产量样本。$H_0: 2\mu_1 > \mu_2$，$H_1: 2\mu_1 \le \mu_2$，

求：在 $1-\alpha$ 水平下的 H_0 拒绝域

解：由已知得，$\bar{X} \sim N(\mu_1, \frac{\delta_1^2}{n_1})$，$\bar{Y} \sim N(\mu_2, \frac{\delta_2^2}{n_2})$

$2\bar{X} \sim N(2\mu_1, \frac{4\delta_1^2}{n_1})$，$2\bar{X}-\bar{Y} \sim N(2\mu_1-\mu_2, \frac{4\delta_1^2}{n_1}+\frac{\delta_2^2}{n_2})$

由 $P\left\{\frac{2\bar{X}-\bar{Y}-(2\mu_1-\mu_2)}{\sqrt{\frac{4\delta_1^2}{n_1}+\frac{\delta_2^2}{n_2}}} < -U_{1-\alpha}\right\} = \alpha$，得所求拒绝域为

$$\left\{\frac{2\bar{X}-\bar{Y}}{\sqrt{\frac{4\delta_1^2}{n_1}+\frac{\delta_2^2}{n_2}}} < -U_{1-\alpha}\right\}.$$

6. 秩和检验，在 $n_1>10$，$n_2>10$，$H_0: F_1(x)=F_2(x)$，$H_1: F_1(x)\neq F_2(x)$

$R_1 \sim N\left(\frac{n_1(n_1+n_2+1)}{2}, \frac{n_1 n_2(n_1+n_2+1)}{12}\right)$，求在 $1-\alpha$ 水平下的 H_0 的拒绝域.

解：由 $R_1 \sim N\left(\frac{n_1(n_1+n_2+1)}{2}, \frac{n_1 n_2(n_1+n_2+1)}{12}\right)$

得：$$\frac{R_1-\frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}}} \sim N(0,1)$$

由 $$P\left\{\left|\frac{R_1-\frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}}}\right| > U_{\frac{\alpha}{2}}\right\}=\alpha.$$

得拒绝域为：$R_1 > U_{\frac{\alpha}{2}}\sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}} + \frac{n_1(n_1+n_2+1)}{2}$

或 $R_1 < \frac{n_1(n_1+n_2+1)}{2} - U_{\frac{\alpha}{2}}\sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}}$.

7. $Y=X\beta+\varepsilon$，$\varepsilon\sim N(0,\sigma^2 I_n)$，$X$ 是一个 15×7 满秩阵

$H_0: \beta_1=\beta_2=\cdots=\beta_6=0$.

(1) 利用 X 求 $Cov(\hat{\beta}, Y)$

(2) 填表，并对 H_0 作 $1-\alpha$ 水平的参数检验

来源	平方和	自由度	均方误差	F值
组间	$\boxed{55960}$	5	11192	10.03.
组内	10040	9	1115.56	
总平方和	$\boxed{66000}$	14.		

解：(1) $Cov(\hat{\beta}, Y) = Cov[(X'X)^{-1}\cdot X'\cdot Y, Y) = Cov(L^{-1}X'Y, Y)$

$= L^{-1}X' Cov(Y, Y) = L^{-1}X'\sigma^2 I_n$ （其中 $L^{-1}=(X'X)^{-1}$

② 填表

因为 X 是 15×7 的满秩阵，则 $n=15$，$k+1=7$，$\Rightarrow n=15$，$k=6$

∴ 自由度分别为：$n-k=9$；$k-1=5$.

$$Y=\begin{bmatrix} y_1 \\ \vdots \\ y_{15} \end{bmatrix},\quad X=\begin{bmatrix} 1 & x_{11} & \cdots & x_{16} \\ 1 & x_{21} & \cdots & x_{26} \\ \vdots & \vdots & & \vdots \\ 1 & x_{15,1} & \cdots & x_{15,6} \end{bmatrix},\quad \beta=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_6 \end{bmatrix},\quad \varepsilon=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{15} \end{bmatrix}$$

满足关系式：$y_t=\beta_0+\beta_1 x_{t1}+\cdots+\beta_6 x_{t6}+\varepsilon_t$，$t=1,\cdots,15$.

$Y=X\beta+\varepsilon$（矩阵形式），其中 X 为已知的 $n\times(k+1)$ 阶矩阵.

8. $y=a\cdot e^{-\beta x}\cdot\varepsilon$，$\ln\varepsilon\sim N(0,\sigma^2)$，样本 $(x_1,y_1),\cdots,(x_n,y_n)$

给出 a 的最小二乘估计。

解. $\ln y=-\beta x+\ln a+\ln\varepsilon$ …①.

令 $z=\ln y$，$\beta_0=\ln a$，从而①式变为：$z=-\beta x+\beta_0+\ln\varepsilon$.

$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，$\bar{z}=\frac{1}{n}\sum_{i=1}^{n}\ln y_i$，$L_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2$，$L_{zz}=\sum_{i=1}^{n}\left(\ln y_i-\frac{1}{n}\sum_{i=1}^{n}\ln y_i\right)^2$

$L_{xz}=\sum_{i=1}^{n}(x_i-\bar{x})(z_i-\bar{z})=\sum_{i=1}^{n}(x_i-\bar{x})\left(\ln y_i-\frac{1}{n}\sum_{i=1}^{n}\ln y_i\right)$

∴ $\hat{\beta}_1=\frac{L_{xz}}{L_{xx}}$

$\hat{\beta}_0=\ln\hat{a}=\bar{z}+\hat{\beta}_1\bar{x}=\bar{z}+\frac{L_{xz}}{L_{xx}}\cdot\bar{x}$

则 $\hat{a}=e^{\hat{\beta}_0}=\exp\left\{\bar{z}+\frac{L_{xz}}{L_{xx}}\cdot\bar{x}\right\}$

$$=\exp\left\{\frac{1}{n}\sum_{i=1}^{n}\ln y_i+\frac{\sum_{i=1}^{n}(x_i-\bar{x})\left(\ln y_i-\frac{1}{n}\sum_{i=1}^{n}\ln y_i\right)}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\right\}$$