第一周作业

PB18000029 舒文炫

目录

1.11 题

设 T=T(X) 是充分统计量，S(X)=G(T(X)), 且函数 S=G(T) 是一对一的 (即 $T_1 \neq T_2 \implies G(T_1) \neq G(T_2)$), 证明：S 也是充分统计量。

证明: 因为 T 是充分统计量，由因子分解定理, 对 X 的分布族 $\{f(x,\theta),\theta \in \Theta\}$,$\Theta$ 为参数空间，我们有

$$f(x,\theta) = g(T(x),\theta)h(X)$$

. 要证 S 是充分统计量，根据因子分解定理，只需要将 $f(x,\theta)$ 做类似分解即可。注意到,$S = G(T)$ 是一对一的，即存在函数 G^{-1}, 使得 $G^{-1}(S) = T$, 那么我们将此式代入上式即得

$$f(x,\theta) = g(G^{-1}(S),\theta)h(x)$$

, 从而 S 是充分统计量。

1.16 题

设 $X_1, ..., X_n$ 是从总体 X 中抽取的 i.i.d 样本，X 的密度函数为

$$f_\theta(x) = \frac{1}{2\theta}e^{-\frac{|x|}{\theta}} \quad (-\infty < x < +\infty, \theta > 0)$$

证明：$T = \sum_{i=1}^n |X_i|$ 是 θ 的充分完全统计量.

证明: 由题 $X_1, ..., X_n$ 的联合密度分布为

$$\begin{aligned} f_\theta(x_1, ..., x_n) &= \frac{1}{2\theta}e^{\sum_{i=1}^n -\frac{|x_i|}{\theta}} \\ &= \frac{1}{2\theta}e^{-\frac{\sum_{i=1}^n |x_i|}{\theta}} \end{aligned}$$

若记 $T = \sum_{i=1}^n |x_i|$, 上式化为

$$f_\theta(x_1, ..., x_n) = \frac{1}{2\theta}e^{-\frac{T}{\theta}}$$

即，我们将式子化成了

$$f(x, \theta) = g(T(x), \theta)h(X)$$

这样的形式，其中 $h(X) = 1$ 根据因子分解定理 $T = \sum_{i=1}^n |x_i|$ 为 θ 的充分统计量

1.23 题

设 $X_1, ..., X_m, i.i.d \sim N(\mu_1, \sigma^2)$,$Y_1, ..., Y_n \sim N(\mu_2, \sigma^2)$ 且两组样本独立，求下列检验问题的检验水平为 α 的似然比检验.

$$H_0 : \mu_2 - \mu_1 = 0 \leftrightarrow H_1 : \mu_2 - \mu_1 \neq 0$$

解: 先写出似然函数

$$f(X, Y, \mu_1, \mu_2, \sigma) = (2\pi\sigma^2)^{-\frac{m+n}{2}} e^{-\frac{\sum_{i=1}^m (x_i - \mu_1)^2}{2\sigma^2} - \frac{\sum_{i=1}^n (y_i - \mu_2)^2}{2\sigma^2}}$$

参数空间为 $\Theta = \{\theta = (\mu_1, \mu_2, \sigma^2) - \infty < \mu_1, \mu_2 < +\infty, \sigma^2 > 0\}$ 零假设对应 Θ 子集为 $\Theta_0 = \{\theta = (\mu_1, \sigma^2) - \infty < \mu_1 < +\infty, \sigma^2 > 0\}$, 因为这里 $\mu_1 = \mu_2$

在 Θ 上，μ_1, μ_2, σ^2 极大似然估计分别为

$$\hat{\mu_1} = \overline{X}$$
$$\hat{\mu_2} = \overline{Y}$$
$$\hat{\sigma^2} = \frac{1}{m+n}[\sum_{i=1}^{m}(X_i - \overline{X}) + \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})]$$

对于 Θ_0, 可将似然函数化成如下形式

$$f(X, Y, \mu_1, \mu_2, \sigma) = (2\pi\sigma^2)^{\frac{m+n}{2}} e^{-\frac{\sum_{i=1}^{m}(x_i-\mu_1)^2}{2\sigma^2} - \frac{\sum_{i=1}^{n}(y_i-\mu_1)^2}{2\sigma^2}}$$

从而 μ_1, σ^2 的极大似然估计为

$$\hat{\mu_1} = \frac{m\overline{X} + n\overline{Y}}{m+n}$$
$$\hat{\sigma^2} = \frac{1}{m+n}[\sum_{i=1}^{m}(X_i - \hat{\mu_1}) + \sum_{i=1}^{n}(Y_i - \hat{\mu_1})]$$

将这些结果代入似然比的公式得

$$\begin{aligned}\sup_{\theta\in\Theta} f(x, y, \theta) &= f(x, y, \hat{\mu_1}, \hat{\mu_2}, \hat{\sigma^2}) \\ &= \{\frac{2\pi e}{m+n}[\sum_{i=1}^{m}(x_i - \overline{x}) + \sum_{i=1}^{n}(y_i - \overline{y})\}^{-\frac{m+n}{2}} \\ \sup_{\theta\in\Theta_0} f(x, y, \theta) &= f(x, y, \hat{\mu_1}, \hat{\sigma^2}) \\ &= \{\frac{2\pi e}{m+n}[\sum_{i=1}^{m}(x_i - \hat{\mu_1}) + \sum_{i=1}^{n}(y_i - \hat{\mu_1})\}^{-\frac{m+n}{2}}\end{aligned}$$

似然比

$$\begin{aligned}\lambda(x, y) &= \frac{\sup_{\theta\in\Theta_0} f(x, y, \theta)}{\sup_{\theta\in\Theta} f(x, y, \theta)} \\ &= \frac{[\sum_{i=1}^{m}(x_i - \overline{x}) + \sum_{i=1}^{n}(y_i - \overline{y})]^{\frac{m+n}{2}}}{[\sum_{i=1}^{m}(x_i - \hat{\mu_1}) + \sum_{i=1}^{n}(y_i - \hat{\mu_1})]^{\frac{m+n}{2}}} \\ &= (1 + \frac{mn(\overline{x} - \overline{y})^2}{(m+n)[(m-1)S_x^2 + (n-1)S_y^2]})^{-\frac{m+n}{2}}\end{aligned}$$

$$S_x^2 = \frac{\sum_{i=1}^{m}(x_i - \overline{x})^2}{m-1}$$
$$S_y^2 = \frac{\sum_{i=1}^{n}(y_i - \overline{y})^2}{n-1}$$

令

$$T(x,y)^2 = \frac{(m+n-2)mn(\overline{x}-\overline{y})^2}{(m+n)[(m-1)S_x^2+(n-1)S_y^2]}$$

λ 为 $|T|$ 的严格减函数,

$$T = \sqrt{\frac{mn}{m+nn}}\frac{(\overline{x}-\overline{y})}{S_w}$$

$$S_w^2 = [(m-1)S_x^2+(n-1)S_y^2]/(m+n-2)$$

且 H_0 成立时，$T \sim t_{m+n-2}$ 检验的否定域 $D = \{X, Y : \lambda(x,y) < c'\} = \{X, Y : T(x,y) > c\}$ 其中 c 满足 $P(\{X, Y : |T(x,y)| > c\}) = \alpha$ 从而求得 $c = t_{m+n-2}(\alpha/2)$ 从而当 $|T| > t_{m+n-2}(\alpha/2)$ 时拒绝原假设，是一个检验水平为 α 的似然比检验.