2013012266

数学作业纸

（科目：　　）

班级：基科32　姓名：曾栩文　编号：　第　页

6.5

$$f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{2i\theta} I(-i(\theta-1) < x_i < i(\theta+1))$$

$$= \frac{1}{2n!\theta^n} \prod_{i=1}^{n} I(-(\theta-1) < \frac{x_i}{i} < \theta+1)$$

$$= \frac{1}{2n!\theta^n} I(\min_i \frac{x_i}{i} > -(\theta-1)) I(\max_i \frac{x_i}{i} < \theta+1)$$

$h(x) = \frac{1}{2n!}$

由因子分解定理．

$T(x) = (\min_i \frac{x_i}{i}, \max_i \frac{x_i}{i})$ 为一个二维充分统计量．

6.6．

$$f(x|\theta) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x_i/\beta}$$

$$= \left(\frac{1}{\Gamma(\alpha)\beta^\alpha}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} e^{-\sum_{i=1}^{n} x_i/\beta}$$

$h(x) = 1$

由因子分解定理．

$T(x) = (\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i)$ 为一个二维充分统计量．

6.8

$$f(x|\theta) = \prod_{i=1}^{n} f(x_i - \theta)$$

$$= \prod_{i=1}^{n} f(x_{(i)} - \theta)$$

由因子分解定理

$T(x) = (X_{(1)}, \ldots, X_{(n)})$ 是一个充分统计量.

在不知道 f 的具体形式时，无法进一步对 T 进行压缩.

例如对 Cauchy 分布 $f(x|\theta) = \frac{1}{\pi(x_i-\theta)^2}$.

若 $\frac{f(x|\theta)}{f(y|\theta)} = \prod_{i=1}^{n} \left(\frac{y_i-\theta}{x_i-\theta}\right)^2 = \prod_{i=1}^{n} \left(\frac{y_{(i)}-\theta}{x_{(i)}-\theta}\right)^2$ 对于 θ 是常数.

则必有 $x_{(i)} = y_{(i)}$.

因此顺序统计量是一个极小充分统计量，因此不可能进一步压缩.

Theorem 6.2.6

若 $T(X)$ 是一个充分统计量

则由定义

$h(x) = P(X=x \mid T(X)=T(x))$ 与 θ 无关

记 $g(t|\theta) = P_\theta(T(X)=t)$

故 $P_\theta(X=x) = P_\theta(X=x, T(X)=T(x))$

$= P_\theta(X=x \mid T(X)=T(x))\, P_\theta(T(X)=T(x))$

$= h(x)\, g(t|\theta)$

另一方面，若 x 的 pmf 可写为

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

则 ~~$P_\theta(T(X)=T(x))$~~

$P_\theta(T(X)=T(x)) = \sum_{T(y)=T(x)} f(y|\theta)$

记 $A_t = \{x \mid T(x)=t\}$

则 $P_\theta(T(X)=T(x)) = \sum_{y\in A_{T(x)}} f(y|\theta)$

$= \sum_{y\in A_{T(x)}} g(T(y)|\theta)\, h(y)$

$= g(T(x)|\theta) \sum_{y\in A_{T(x)}} h(y)$

数 学 作 业 纸 （科目：　　）

班级：　　姓名：　　编号：　　第　　页

故 $P_\theta(X=x \mid T(X)=T(x))$

$$= \frac{P_\theta(X=x)}{P_\theta(T(X)=T(x))}$$

$$= \frac{g(T(x)\mid\theta)\,h(x)}{g(T(x)\mid\theta)\sum_{y\in A_{T(x)}} h(y)} = \frac{h(x)}{\sum_{y\in A_{T(x)}} h(y)}.$$

与 θ 无关

即 $T(X)$ 为一个充分统计量.

2. $f_{X_i}(x_i) = \frac{e^{-\lambda_i}\lambda_i^{x_i}}{x_i!} = \frac{e^{-c_i\lambda}(c_i\lambda)^{x_i}}{x_i!}.$

故 $f(x\mid\theta) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{e^{-c_i\lambda}(c_i\lambda)^{x_i}}{x_i!}$

$$= \frac{e^{-(\sum_{i=1}^{n} c_i)\lambda}\cdot\left(\prod_{i=1}^{n} c_i^{x_i}\right)\cdot\lambda^{\left(\sum_{i=1}^{n} x_i\right)}}{\prod_{i=1}^{n} x_i!}.$$

由因子分解定理

$T(X) = \left(\prod_{i=1}^{n} c_i^{x_i},\ \sum_{i=1}^{n} x_i,\ \prod_{i=1}^{n} x_i!\right)$ 为一个充分统计量.

证明：

记 $A_t = \{x \mid T(x)=t\}$. $n_t = |A_t|$

数 学 作 业 纸 （科目：　　）

班级：　　姓名：　　编号：　　第　　页

$$P_\lambda(T(X)=T(x)) = \sum_{y\in A_{T(x)}} f(y|\theta) = n_t f(x|\theta)$$

故 ~~$P_\lambda(T(x))=$~~

$$P_\lambda(X=x \mid T(X)=T(x))$$

$$= \frac{P_\lambda(X=x)}{P_\lambda(T(X)=T(x))} = \frac{1}{n_t}$$ 与 x 无关.

3.

$$f(x|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2}$$

(a) $T_1(X) = X$

显然 $f(x|\theta) = f(T(x)|\theta)$

故 $T_1(X)$ 为一个充分统计量.

(b). $T_2(X) = (X_1^2, \cdots, X_n^2)$;

有 $f(x|\theta) = g_2(T_2(X)|\theta)$

其中 $g(t|\theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} t_i}$

故 $T_2(X)$ 为一个充分统计量.

数 学 作 业 纸 （科目： ）

班级： 姓名： 编号： 第 页

(c) $f(x|\theta) = g_3(T_3(x)|\theta)$.

其中 $g_3(t|\theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{t_1+t_2}{2\sigma^2}}$.

故 $T_3(X) = (X_1^2+\dots X_m^2,\ X_{m+1}^2+\dots X_n^2)$ 为一个充分统计量.

(d) $T_4(X) = \sum_{i=1}^{n} X_i^2$.

~~故~~有 $f(x|\theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{T_4(x)}{2\sigma^2}}$

故 $T_4(X)$ 为一个充分统计量.

从数据压缩的角度看.

T_4 最好. 而且可以证明，T_4 是一个极小充分统计量.