6.9 Let $X_1, \cdots, X_n$ be a random sample. Find a minimal sufficient statistic for θ.

By Theorem 6.2.13, the ratio $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is constant as a function of θ iff $T(\mathbf{x}) = T(\mathbf{y})$. Then $T(x)$ is a minimal sufficient statistic for θ.

(a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}$, $x \in (-\infty, \infty)$ and $\theta \in (-\infty, \infty)$ (normal)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\cancel{(2\pi)^{-n/2}}\, e^{-\sum_{i=1}^{n}(x_i-\theta)^2/2}}{\cancel{(2\pi)^{-n/2}}\, e^{-\sum_{i=1}^{n}(y_i-\theta)^2/2}} = \exp\left\{-\frac{1}{2}\underbrace{\sum_{i=1}^{n}\left[(x_i-\theta)^2-(y_i-\theta)^2\right]}\right\}$$

$$\sum_{i=1}^{n}\left[(x_i-\theta)^2-(y_i-\theta)^2\right] = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2 + 2\theta\left(\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2 + 2\theta n(\bar{y}-\bar{x})$$

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

Since $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is constant as a function of θ iff $\bar{y}=\bar{x}$, so $\bar{X}$ is a minimal sufficient statistic for θ.

(b) $f(x|\theta) = e^{-(x-\theta)}$, $x \in (\theta, \infty)$, $\theta \in (-\infty, \infty)$ (location exponential)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{e^{-\sum_{i=1}^{n}(x_i-\theta)}\prod_{i=1}^{n} I_{(\theta,\infty)}(x_i)}{e^{-\sum_{i=1}^{n}(y_i-\theta)}\prod_{i=1}^{n} I_{(\theta,\infty)}(y_i)} = \frac{e^{-\sum x_i}\,\cancel{e^{n\theta}}\, I_{(\theta,\infty)}(\min x_i)}{e^{-\sum y_i}\,\cancel{e^{n\theta}}\, I_{(\theta,\infty)}(\min y_i)}$$

Since $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is constant as a function of θ iff $\min\{x_1, \cdots, x_n\} = \min\{y_1, \cdots, y_n\}$, so $T(\mathbf{X}) = \min\{X_1, X_2, \cdots, X_n\}$ is a minimal sufficient statistic for θ.

6.9 Continue ...

(c) $f(x|\theta) = \dfrac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}$, $x \in (-\infty, \infty)$, $\theta \in (-\infty, \infty)$ (logistic)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{e^{-\sum_{i=1}^{n}(x_i-\theta)}}{\prod_{i=1}^{n}(1+e^{-(x_i-\theta)})^2} \cdot \frac{\prod_{i=1}^{n}(1+e^{-(y_i-\theta)})^2}{e^{-\sum_{i=1}^{n}(y_i-\theta)}}$$

$$= e^{-\sum_{i=1}^{n}(x_i-y_i)} \left(\frac{\prod_{i=1}^{n}(1+e^{-(y_i-\theta)})}{\prod_{i=1}^{n}(1+e^{-(x_i-\theta)})} \right)^2$$

Since $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is constant as a function of θ iff $\mathbf{x}$ and $\mathbf{y}$ have the same order statistics, therefore the order statistics are minimal sufficient for θ.

(d) $f(x|\theta) = \dfrac{1}{\pi[1+(x-\theta)^2]}$, $x \in (-\infty, \infty)$, $\theta \in (-\infty, \infty)$ (Cauchy)

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{i=1}^{n}\pi[1+(y_i-\theta)^2]}{\prod_{i=1}^{n}\pi[1+(x_i-\theta)^2]} = \frac{\cancel{\pi^n}\prod_{i=1}^{n}(1+(y_i-\theta)^2)}{\cancel{\pi^n}\prod_{i=1}^{n}(1+(x_i-\theta)^2)}$$

Since $\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)}$ is constant as a function of θ iff $\mathbf{x}$ and $\mathbf{y}$ have the same order statistics, thus the order statistics are minimal sufficient for θ.

#6.9 continue...

(e) $f(x|\theta) = \frac{1}{2} e^{-|x-\theta|}$, $x \in (-\infty, \infty)$, $\theta \in (-\infty, \infty)$ (double exponential)

$$\frac{f(x|\theta)}{f(y|\theta)} = \frac{\cancel{2^n} e^{-\sum_{i=1}^n |x_i - \theta|}}{\cancel{2^n} e^{-\sum_{i=1}^n |y_i - \theta|}} = \exp\left\{ -\sum_{i=1}^n \left(|x_i - \theta| - |y_i - \theta| \right) \right\}.$$

Define a set $A = \{i : x_i \le \theta\}$ and $n(A)$ = the numbe of elements in A.
Similarly define a set $B = \{i : y_i \le \theta\}$ and $n(B)$ = the number of elements in B.

$$\sum_{i=1}^n |x_i - \theta| - \sum_{i=1} |y_i - \theta| = \underbrace{\left\{ \sum_{i \in A} (\theta - x_i) + \sum_{i \in A^c} (x_i - \theta) \right\}}_{\text{part}(A)} - \underbrace{\left\{ \sum_{i \in B} (\theta - y_i) + \sum_{i \in B^c} (y_i - \theta) \right\}}_{\text{par}(B)}$$

For part(A),

$$\sum_{i \in A} (\theta - x_i) + \sum_{i \in A^c} (x_i - \theta) = n(A)\theta - \sum_{i \in A} x_i + \sum_{i \in A^c} x_i - (n - n(A)\theta)$$
$$= \sum_{i \in A^c} x_i - \sum_{i \in A} x_i + 2\theta n(A) - n$$

For part (B),

$$\sum_{i \in B} (\theta - y_i) + \sum_{i \in B^c} (y_i - \theta) = \sum_{i \in B^c} y_i - \sum_{i \in B} y_i + 2\theta n(B) - n$$

Then, we have

$$\sum_{i=1}^n |x_i - \theta| - \sum_{i=1}^n |y_i - \theta| = 2\theta\left(n(A) - n(B) \right) + \left[\left(\sum_{i \in A^c} x_i - \sum_{i \in B^c} y_i \right) - \left(\sum_{i \in A} x_i - \sum_{i \in B} y_i \right) \right]$$

$\frac{f(x|\theta)}{f(y|\theta)}$ is constant as a function of θ iff $\sum_{i \in A^c} x_i = \sum_{i \in B^c} y_i$ and $\sum_{i \in A} x_i = \sum_{i \in B} y_i$.

In otherwords, this is valid iff x and y have the same order statics.

Thus, the order statistics for X are a minimal sufficient statistic.