Chapter 6: Principles of Data Reduction

January 1, 2022

Exercise 6.1

Yes

$$\frac{p(x|0,\sigma)}{q(|X||0,\sigma)} = \frac{p(x|0,\sigma)}{p(x|0,\sigma)+p(-x|0,\sigma)} = \frac{1}{2}$$

Does not depend on the paramters.

Exercise 6.2

The pdf for X_i is $f_{X_i}(x|\theta) = \exp(i\theta - x)\mathbb{1}_{x\geq i\theta}$. Then

$$\begin{aligned}
f_X(x_i|\theta) &= \prod_{i=1}^{n} f_{X_i}(x_i|\theta) \\
&= \exp\left(\sum_{i=1}^{n} i\theta - x_i\right)\prod_{i=1}^{n}\mathbb{1}_{x_i\geq i\theta} \\
&= \exp\left(\frac{n(n+1)\theta}{2}\right)\exp\left(\sum_i x_i\right)\prod_{i=1}^{n}\mathbb{1}_{\frac{x_i}{i}\geq\theta} \\
&= \exp\left(\frac{n(n+1)\theta}{2}\right)\mathbb{1}_{\min\frac{x_i}{i}\geq\theta}\exp\left(\sum_i x_i\right)
\end{aligned}$$

$g(T(x)|\theta) = g(\min\frac{x_i}{i}|\theta) = \exp\left(\frac{n(n+1)\theta}{2}\right)\mathbb{1}_{\min\frac{x_i}{i}\geq\theta}$ and $h(x) = \exp\left(\sum_i x_i\right)$. By factorization theorem, it is sufficient statistic.

Exercise 6.3

Given the pdf

$$f(x|\mu,\sigma) = \frac{1}{\sigma}e^{-(x-\mu)/\sigma}, \quad u < x < \infty, 0 < \sigma < \infty$$

We need to get rid of the dependency on u in the range of x. With indicator, we can rewrite it as

$$f(x|\mu,\sigma) = \frac{1}{\sigma}e^{-(x-\mu)/\sigma}\mathbb{1}_{\mu<x}, \quad -\infty < x < \infty, 0 < \sigma < \infty$$

Then

$$\begin{aligned} f(\mathbf{x}|\mu,\sigma) &= \prod_{i=0}^{n} f(x_i|\mu,\sigma) \\ &= \frac{1}{\sigma^n}\exp\left(-\frac{\sum_i x_i - n\mu}{\sigma}\right)\prod_{i=0}^{n}\mathbb{1}_{\mu<x_i} \\ &= \frac{1}{\sigma^n}\exp\left(-\frac{\sum_i x_i - n\mu}{\sigma}\right)\mathbb{1}_{\mu<x_{\min}} \end{aligned}$$

Define $T((x)) = (t_1, t_2) = (\sum_i x_i, x_{\min})$ and $h(\mathbf{x}) = 1$. By factorization theorem, it is a sufficient statistics.

Exercise 6.4

The pdf is

$$\begin{aligned} f(x|\theta) &= \left[\prod_{i}^{n} h(x_i)\right] c^n(\theta)\exp\left(\sum_{i=1}^{n}\sum_{j=1}^{k} w_j(\theta)t_j(x_i)\right) \\ &= H(x)C(\theta)\exp\left(\sum_{j=1}^{k} w_j(\theta)\sum_{i=1}^{n} t_j(x_i)\right) \\ &= H(x)C(\theta)\exp\left(\sum_{j=1}^{k} w_j(\theta)T_j(x)\right) \\ &= H(x)C(\theta)g(T(x)|\theta) \end{aligned}$$

Therefore by Factorization theorem, $T(X) = [T_j(X)] = (\sum_{i=1}^{n} t_j(x_i))$ is sufficient statistics.

Exercise 6.8

The pdf of sample X is $f(x|\theta) = \prod_{i=1}^{n} f(x_i - \theta)$. By theorem 6.2.13, we take the ratio of $f(x|\theta)$ and $f(y|\theta)$,

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

Note that the above expression is a constant of θ only when the terms cancel out which implies there exists an ordering T such that $f(T(x_i) - \theta) = f(T(y_i) - \theta)$. Order statistics is such an ordering. Therefore it is the minimal statistics.