STA 522/Solutions to Homework 3

Problem 6.1

Use the Factorization theorem on the pdf of X:

$$f(x \mid \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \ \exp\left(-\frac{x^2}{2\sigma^2}\right) = \underbrace{\left\{\frac{1}{\sigma} \ \exp\left(-\frac{|x|^2}{2\sigma^2}\right)\right\}}_{=g(T(x)|\sigma)} \underbrace{\left(\frac{1}{2\pi}\right)}_{=h(x)}$$

where $T(x) = |x|$. Therefore by the Factorization theorem, $T(X) = |X|$ is a sufficient statistic for σ.

Problem 6.4

If $X_1, X_2, \ldots, X_n$ is a random sample from a pdf/pmf from the exponential family, then their joint pdf/pmf is given by:

$$\begin{aligned} f(x_1, \ldots, x_n \mid \underline{\theta}) &= \prod_{j=1}^{n} \left\{ h(x_j) \ c(\underline{\theta}) \ \exp\left(\sum_{i=1}^{k} w_i(\underline{\theta}) \ t_i(x_j)\right)\right\} \\ &= \underbrace{\left\{ c(\underline{\theta})^n \exp\left(\sum_{i=1}^{k} w_i(\underline{\theta}) \ \sum_{j=1}^{n} t_i(x_j)\right)\right\}}_{=g(\underline{T}(\underline{x})|\underline{\theta})} \underbrace{\left(\prod_{i=1}^{n} h(x_j)\right)}_{=h(\underline{x})} \end{aligned}$$

where $\underline{T}(\underline{x}) = \left(\sum_{j=1}^{n} t_1(x_j), \ldots, \sum_{j=1}^{n} t_k(x_j)\right)$. Hence, by the Factorization theorem it follows that $\underline{T}(\underline{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \ldots, \sum_{j=1}^{n} t_k(X_j)\right)$ is sufficient for $\underline{\theta}$.

Problem 6.6

The joint pdf of $\underline{X} = (X_1, X_2, \ldots, X_n)$ is given by:

$$f(\underline{x} \mid \alpha, \beta) = \prod_{i=1}^{n} \left\{ \frac{1}{\Gamma(\alpha)\beta^\alpha} \ x_i^{\alpha-1} \ e^{-x_i/\beta} \right\} = \underbrace{\left(\frac{1}{\Gamma(\alpha)\beta^\alpha}\right)^n \exp\left[(\alpha-1)\sum_{i=1}^{n} \log x_i - \frac{1}{\beta}\sum_{i=1}^{n} x_i\right]}_{g(T_1(\underline{x}),T_2(\underline{x})|\alpha,\beta)} \ \underbrace{1}_{=h(\underline{x})}$$

Therefore, from the Factorization theorem, $(\sum_{i=1}^{n} \log X_i, \sum_{i=1}^{n} X_i)$ is a sufficient statistic for (α, β).

Problem 6.9

Part (a): Consider two sample points $\underline{x}$ and $\underline{y}$ from the density. Then

$$\begin{aligned}
\frac{f(\underline{x} \mid \theta)}{f(\underline{y} \mid \theta)} &= \frac{(2\pi)^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i - \theta)^2\right)}{(2\pi)^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^n (y_i - \theta)^2\right)} \\
&= \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^n x_i^2 + n\overline{x}\theta - \frac{1}{2}n\theta^2\right)}{\exp\left(-\frac{1}{2}\sum_{i=1}^n y_i^2 + n\overline{y}\theta - \frac{1}{2}n\theta^2\right)} \\
&= \exp\left[\frac{1}{2}\sum_{i=1}^n (y_i^2 - x_i^2) - n\theta(\overline{y} - \overline{x})\right]
\end{aligned}$$

is constant as a function of θ if and only if $\overline{y} = \overline{x}$. Hence, $\overline{X}$ is a minimal sufficient statistic for θ.

Part (b): Consider two sample points $\underline{x}$ and $\underline{y}$ from the density. Then

$$\begin{aligned}
\frac{f(\underline{x} \mid \theta)}{f(\underline{y} \mid \theta)} &= \frac{\exp\left(-\sum_{i=1}^n (x_i - \theta)\right)\ \prod_{i=1}^n I(x_i > \theta > 0)}{\exp\left(-\sum_{i=1}^n (y_i - \theta)\right)\ \prod_{i=1}^n I(y_i > \theta > 0)} \\
&= \exp\left(\sum_{i=1} (y_i - x_i)\right) \frac{I\left(0 < \theta < x_{(1)}\right)}{I\left(0 < \theta < y_{(1)}\right)}.
\end{aligned}$$

This is constant as a function of θ if and only if the ratio of indicators is 1 for all θ, i.e., $x_{(1)} = y_{(1)}$. This means $X_{(1)} = \min_i X_i$ is a minimal sufficient statistic for θ.