Common Theorems

I'll be making heavy use of a few theorems in this assignment, so I'll include them at the start for reference.

Theorem 6.2.6 (Factorization Theorem)

Let $f(\vec{x} \mid \theta)$ denote the joint pdf or pmf of a sample $\vec{X}$. A statistic $T(\vec{X})$ is a sufficient statistic for θ if and only if there exist functions $g(t \mid \theta)$ and $h(x)$ such that, for all sample points $\vec{x}$ and all parameter points θ,

$$f(\vec{x} \mid \theta) = g(T(\vec{x}))h(\vec{x})$$

Theorem 6.2.13

Let $f(\vec{x} \mid \theta)$ be the pmf or pdf of a sample $\vec{X}$. Suppose there exists a function $T(\vec{x})$ such that, for every two sample points x and y, the ratio $f(x \mid \theta)/f(y \mid \theta)$ is constant as a function of θ if and only if $T(\vec{x}) = T(\vec{y})$. Then $T(\vec{x})$ is a minimal sufficient statistic for θ.

6.1

Let X be one observation from a $N(0, \sigma^2)$ population. Is $|X|$ a sufficient statistic?

$$\begin{aligned} f(x \mid 0, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-0)^2\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \\ &\text{Note: } x^2 = |x|^2 \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|x|^2}{2\sigma^2}\right) \cdot I_{(-\infty,\infty)}(x) \end{aligned}$$

By the factorization theorem we have,

$$\begin{aligned} g(T(x) \mid \theta) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|x|^2}{2\sigma^2}\right) \\ h(x) &= I_{(-\infty,\infty)}(x) \end{aligned}$$

Thus, $T(x) = |x|$ is a sufficient statistic for theta.

6.2

Let $X_i, ..., X_n$ be a random sample from the pdf

$$f_{x_i}(x|\theta) = \begin{cases} e^{i\theta - x} & x \geq i\theta \\ 0 & x < i\theta \end{cases}$$

Prove that $T = \min_i(X_i/i)$ is a sufficient statistic for θ

First, let us rewrite the pdf using indicator functions and some rearrangement of the inequality.

$$\begin{aligned} x &\geq i\theta \\ \frac{x}{i} &\geq \theta \end{aligned}$$

This allows us the use of the indicator function $I_{(\theta \leq x/i)}(x)$.

$$\begin{aligned} f_{x_i}(x \mid \theta) &= e^{i\theta - x} I_{(\theta \leq x/i)}(x) \\ f(\vec{x} \mid \theta) &= \prod_{i=1}^{n} e^{i\theta - x_i} I_{(\theta \leq x_i/i)}(x_i) \\ &= \exp\left(\sum_{i=1}^{n} i\theta\right) \cdot \exp\left(-\sum_{i=1}^{n} x_i\right) \cdot \prod_{i=1}^{n} I_{(\theta \leq x_i/i)}(x_i) \end{aligned}$$

Now we return to the indicator function but within the context of a product now. This product requires that, for all i, that $\theta \leq x_i/i$. This is important because if any of the values of the sample violate this the entire product evaluates to 0. So we can look at the minimum value that x_i/i holds for the entire sample and create an indicator function around that. That is, we use $min_i(X_i/i)$. In other words, $\prod I_{(\theta \leq x_i/i)}(x_i) = I_{(\theta \leq min_i(x_i/i))}$. So,

$$f(\vec{x} \mid \theta) = \exp\left(-\sum_{i=1}^{n} x_i\right) \cdot \exp\left(\sum_{i=1}^{n} i\theta\right) \cdot I_{(\theta \leq min_i(x_i/i))}$$

Giving us

$$\begin{aligned} h(\vec{x}) &= \exp\left(-\sum_{i=1}^{n} x_i\right) \\ g(T(\vec{x}) \mid \theta) &= \exp\left(\sum_{i=1}^{n} i\theta\right) \cdot I_{(\theta \leq min_i(x_i/i))} \end{aligned}$$

Thus, $T(\vec{x}) = min_i(X_i/i)$ is a sufficient statistic for θ.

6.4

Prove Theorem 6.2.10

Theorem 6.2.10

Let $X_1, \cdots, X_n$ be iid observations from a pdf or pmf $f(x \mid \theta)$, which belongs to an exponential family given by:

$$f(x \mid \theta) = h(x)c(\theta)\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right),$$

where $\theta = (\theta_1, \cdots, \theta_d), d \leq k$. Then

$$T(\vec{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \cdots, \sum_{j=1}^{n} t_k(X_j)\right)$$

is a sufficient statistic for θ.

We'll be leveraging the factorization theorem for this proof.

We begin by writing out the joint pdf of the exponential family.

$$\begin{aligned}
f(\vec{x} \mid \theta) &= \prod_{i=1}^{n} h(x_i)c(\theta)\exp\left(\sum_{j=1}^{k} w_j(\theta)t_j(x_i)\right) \\
&= c(\theta)^n \left(\prod_{i=1}^{n} h(x_i)\right) \cdot \exp\left(\sum_{j=1}^{k} w_j(\theta)\sum_{i=1}^{n} t_j(x_i)\right) && \text{(Distribute product)} \\
&= \left(\prod_{i=1}^{n} h(x_i)\right) \cdot c(\theta)^n \cdot \exp\left(\sum_{j=1}^{k} w_j(\theta)\sum_{i=1}^{n} t_j(x_i)\right) && \text{(Rearrange)}
\end{aligned}$$

Giving us

$$g(T(\vec{x}) \mid \theta) = c(\theta)^n \cdot \exp\left(\sum_{j=1}^{k} w_j(\theta)\sum_{i=1}^{n} t_j(x_i)\right)$$

$$h(\vec{x}) = \prod_{i=1}^{n} h(x_i)$$

Therefore,

$$T(\vec{x}) = \left(\sum_{i=1}^{n} t_1(x_i), \cdots, \sum_{i=1}^{n} t_k(x_i) \right)$$

is a sufficient statistic for θ by the factorization theorem.

6.6

Let $X_1, \cdots, X_n \overset{\text{iid}}{\sim} Gamma(\alpha, \beta)$
Find a two dimension sufficient statistic for (α, β).

$$
\begin{aligned}
f(x \mid \alpha, \beta) &= \left(\Gamma(\alpha)\beta^{\alpha}\right)^{-1} x^{\alpha-1} e^{-x/\beta} I_{x\geq 0}(x) \\
f(\vec{x} \mid \alpha, \beta) &= \prod_{i=1}^{n} \left(\Gamma(\alpha)\beta^{\alpha}\right)^{-1} x_i^{\alpha-1} e^{-x_i/\beta} I_{x\geq 0}(x_i) \\
&= \left(\Gamma(\alpha)\beta^{\alpha}\right)^{-n} e^{-\frac{n}{\beta}\sum x_i} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \prod_{i=1}^{n} I_{x\geq 0}(x_i)
\end{aligned}
$$

Here we have

$$
\begin{aligned}
g(T(\vec{x}) \mid \theta) &= \left(\Gamma(\alpha)\beta^{\alpha}\right)^{-n} e^{-\frac{n}{\beta}\sum x_i} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \\
h(\vec{x}) &= \prod_{i=1}^{n} I_{x\geq 0}(x_i)
\end{aligned}
$$

Therefore, by the factorization theorem,

$$
T(\vec{x}) = \left(\prod_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i\right)
$$

is a sufficient statistic for (α, β).

6.9 (b,c)

For each of the following distributions let $X_1, \cdots, X_n$ be a random sample. Find a minimal sufficient statistic for θ.

B

Our goal here will be to set up the joint pdf and then utilize theorem 6.2.13 to find the minimal sufficient statistic for θ.

$$\begin{aligned} f(x \mid \theta) &= e^{-(x-\theta)} I_{(\theta,\infty)}(x) \\ f(\vec{x} \mid \theta) &= \prod_{i=1}^{n} e^{-(x_i-\theta)} I_{(\theta,\infty)}(x_i) \\ &= e^{-\sum x_i + n\theta} \prod_{i=1}^{n} I_{(\theta,\infty)}(x_i) \end{aligned}$$

Let us examine the indicator function for a moment. It states that, for all i, that $x_i > \theta$. What this means is that if even a single x_i violates this that the entire product evaluates to 0. This means we can look at the minimum value of our sample. I'll write this as the first order statistic, $x_{(1)}$. If the minimum of our sample is greater than θ then we're in the clear. So we can rewrite our indicator function using that fact.

Using that knowledge,

$$f(\vec{x} \mid \theta) = e^{-\sum x_i + n\theta} I_{(\theta,\infty)}(x_{(1)})$$

Now we use theorem 6.2.13.

$$\begin{aligned} \frac{f(\vec{x} \mid \theta)}{f(\vec{y} \mid \theta)} &= \frac{e^{-\sum x_i + n\theta} I_{(\theta,\infty)}(x_{(1)})}{e^{-\sum y_i + n\theta} I_{(\theta,\infty)}(y_{(1)})} \\ &= e^{\sum y_i - \sum x_i} \cdot \frac{I_{(\theta,\infty)}(x_{(1)})}{I_{(\theta,\infty)}(y_{(1)})} \end{aligned}$$

This is only a constant with respect to θ if $x_{(1)} = y_{(1)}$, thus $T(\vec{x}) = x_{(1)}$ is a minimal sufficient statistic for θ.

C

$$
\begin{aligned}
f(x \mid \theta) &= e^{-(x-\theta)} \cdot (1+e^{-(x-\theta)})^{-2} \\
f(\vec{x} \mid \theta) &= \prod_{i=1}^{n} e^{-(x_i-\theta)} \cdot (1+e^{-(x_i-\theta)})^{-2} \\
&= e^{-\sum x_i + n\theta} \prod_{i=1}^{n} (1+e^{-(x_i-\theta)})^{-2} \\
\frac{f(\vec{x} \mid \theta)}{f(\vec{y} \mid \theta)} &= \frac{e^{-\sum x_i + n\theta} \prod_{i=1}^{n} (1+e^{-(x_i-\theta)})^{-2}}{e^{-\sum y_i + n\theta} \prod_{i=1}^{n} (1+e^{-(y_i-\theta)})^{-2}} \\
&= e^{\sum y_i - \sum x_i} \left(\frac{\prod_{i=1}^{n} (1+e^{-(x_i-\theta)})}{\prod_{i=1}^{n} (1+e^{-(y_i-\theta)})} \right)^2
\end{aligned}
$$

From here we can leverage the results from example 6.2.5 that state that the order statistics are sufficient statistics for θ. So we get,

$$
\frac{f(\vec{x} \mid \theta)}{f(\vec{y} \mid \theta)} = e^{\sum y_i - \sum x_i} \left(\frac{\prod_{i=1}^{n} (1+e^{-(x_{(i)}-\theta)})}{\prod_{i=1}^{n} (1+e^{-(y_{(i)}-\theta)})} \right)^2
$$

This is only constant with respect to θ if $x_{(i)} = y_{(i)} \forall i$, therefore the order statistics are minimal sufficient statistics for θ.