Theorem 6.2.6. *$T(\vec{X})$ is a sufficient statistic for θ if and only if there are functions $g(t|\theta)$ and $h(\vec{x})$ such that:*

$$f_{\vec{X}}(\vec{x}|\theta) = g(T(\vec{x})|\theta) \cdot h(\vec{x}).$$

Remark. We will present the proof for the discrete case. The continuous case requires some additional assumptions (which are usually met in practice) and the argument requires attention to some details that do not arise in the discrete case. The central ideas are the same, though.

Proof. Note that

$$p_{\vec{X}}(\vec{x}|\theta) = P_\theta(\vec{X} = \vec{x}). \tag{$**$}$$

Also, using conditional probability as in $(*)$ in the last lecture:

$$P_\theta(\vec{X} = \vec{x}) = P_\theta(T(\vec{X}) = T(\vec{x})) \cdot P_\theta(\vec{X} = \vec{x} \mid T(\vec{X}) = T(\vec{x})). \tag{$***$}$$

($\Rightarrow$) Suppose T is sufficient for θ. We need to show that $P_\theta(\vec{X} = \vec{x})$ factors in the stated way. Let $g(t|\theta) := P_\theta(T(\vec{X}) = t)$ and let $h(\vec{x}) := P_\theta(\vec{X} = \vec{x} \mid T(\vec{X}) = T(\vec{x}))$. By the sufficiency assumption, $h(\vec{x})$ does not depend on θ. The factorization in the theorem is just a restatement of $(***)$.

($\Leftarrow$) Suppose the factorization $P_\theta(\vec{X} = \vec{x}) = g(T(\vec{x})|\theta) \cdot h(\vec{x})$ can be made. We need to show that T is sufficient for θ. Observe that

$$\begin{aligned} P_\theta(T(\vec{X}) = t) &= \sum_{T(\vec{y})=t} P_\theta(\vec{X} = \vec{y}) \\ &= \sum_{T(\vec{y})=t} g(T(\vec{y})|\theta) \cdot h(\vec{y}) \\ &= g(t|\theta) \sum_{T(\vec{y})=t} h(\vec{y}). \end{aligned}$$

Hence, using $(***)$:

$$\begin{aligned} P_\theta(\vec{X} = \vec{x} \mid T(\vec{X}) = T(\vec{x})) &= \frac{P_\theta(\vec{X} = \vec{x})}{P_\theta(T(\vec{X}) = T(\vec{x}))} \\ &= \frac{h(\vec{x})}{\sum_{T(\vec{y})=T(\vec{x})} h(\vec{y})}, \end{aligned}$$

and this does not vary with changes in θ. ////

Example. Suppose $X_1, \ldots X_n$ are iid discrete on $\{1, 2, 3, \ldots\}$, each with *pmf*:

$$f_{X_i}(x_i|\theta) = \begin{cases} 1/\theta, & \text{if } x_i \in \{1, \ldots, \theta\}; \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$f_{\vec{X}}(\vec{x}|\theta) = \begin{cases} \theta^{-n}, & \text{if } \vec{x} \in \{1, \ldots, \theta\}^n; \\ 0, & \text{otherwise.} \end{cases}$$

Let $T(\vec{x}) = \max(x_1, \ldots, x_n)$, let

$$g(t|\theta) = \begin{cases} \theta^{-n}, & \text{if } t \leq \theta; \\ 0, & \text{otherwise,} \end{cases}$$

and let $h(\vec{x}) = 1$ if all the components of $\vec{x}$ are positive integers, $h(\vec{x}) = 0$ otherwise. Then

$$f_X(\vec{x}|\theta) = g(T(\vec{x})|\theta) \cdot h(\vec{x}),$$

so by the theorem $T(\vec{x}) = \max(x_1, \ldots, x_n)$ is a sufficient statistic for θ.

Bonus Problem. Sufficiency says that anything we can determine about θ from a sample $\vec{X}$, we can determine from $T(\vec{X})$. As a function of n, θ and k, what is the probability that $\theta \leq k \cdot T(\vec{X})$, where $k \geq 1$ and $\vec{X}$ is an n-element sample from the distribution above with parameter θ.

Homework.

1. Use the factorization theorem to find a sufficient statistic for the exponential distribution $f_X(x|\lambda) = \lambda e^{-\lambda x}$, $x \geq 0$.
2. Prove Theorem 6.2.10.
3. 6.6, page 300.