Sufficient Statistic

Practical application (Reduction of dimension).

Theoretical application (Rao - Blackwellization / Lehman - Sheffe theorem).

Definition:- $X_1, \dots, X_n \sim f(x \mid \theta)$

$\theta \longrightarrow$ unknown parameter.

$T_n = T_n(X_1, \dots, X_n)$ is a sufficient statistic if it contains all information about θ.

Mathematically speaking, the conditional dist$^{\underline{n}}$ of $(X_1, \dots, X_n)$ given $T_n = T_n(X_1, \dots, X_n)$ is indep. of θ.

Example:-

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Ber}(1, \theta)$.

Claim:- $T_n(\underset{\sim}{X}) = T_n(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i$ is sufficient statistic for θ.

Verification:-

Case 1:- $P[X_1 = x_1, \ldots, X_n = x_n \mid T_n = t] = 0$ if $\sum_{i=1}^{n} x_i \neq t$

Case 2:- If $\sum_{i=1}^{n} x_i = t$, Then

$$P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n \mid T_n = t]$$

$$= \frac{P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n \cap T_n = t]}{P[T_n = t]}$$

$$= \frac{P[X_1 = x_1, \ldots, X_n = x_n]}{P[T_n(\sum_{i=1}^{n} X_i) = t]} = \frac{P[X_1 = x_1] \times \cdots \times P[X_n = x_n]}{P[T_n = t]} = \frac{\theta^{x_1}(1-\theta)^{1-x_1} \times \cdots \times \theta^{x_n}(1-\theta)^{1-x_n}}{\binom{n}{t}\theta^{t}(1-\theta)^{n-t}}$$

$\to T_n \sim \text{Bin}(n, \theta)$

$$= \frac{\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}}{\binom{n}{t} \theta^t (1-\theta)^{n-t}} \quad \text{since } \sum_{i=1}^{n} x_i = t \quad = \frac{\theta^t (1-\theta)^{n-t}}{\binom{n}{t} \theta^t (1-\theta)^{n-t}} = \frac{1}{\underbrace{\binom{n}{t}}_{\text{is indep. of } \theta}}$$

Hence, $\sum_{i=1}^{n} X_i$ is sufficient statistic for θ.

<u>Remark:-</u> It is clear from the defn., whole Random sample $(X_1, \dots, X_n)$ is sufficient. Also, we have seen that $\sum_{i=1}^{n} X_i$ is sufficient. Hence, it is ~~easy~~ direct to conclude that sufficient statistic is NOT unique.

A Theoretical result:-

Neyman- Fisher Factorization Theorem:- (NFT)

Convince yourself about the motivation

Let $X_1, \dots, X_n \overset{i.i.d.}{\sim} f(x|\theta)$, $\theta \in \Theta$.

unknown parameter

$\hookrightarrow$ Parameter space.

A Statistic $T_n = T_n(X_1, \dots, X_n)$ is sufficient for θ if and only if

$$\underbrace{f_{X_1 \dots X_n}(x_1, \dots, x_n; \theta)}_{\text{Joint density}} = g(\theta; T_n(x_1, \dots, x_n))\, h(x_1, \dots, x_n).$$

Joint density

Here $g(\theta; T_n(x_1, \dots, x_n))$ is a $f^{\underline{n}}$ of θ and $x_1, \dots, x_n$ only through $T_n(x_1, \dots, x_n)$ & $h(x_1, \dots, x_n)$ is indep. of θ.

Remark:- If $h(x_1, \ldots, x_n) = 1$, it is clear from NFT that the whole random sample $(X_1, \ldots, X_n)$ is suffiint statistic for θ.

Remark (Result):- Sufficient statistic is NOT unique.

Justification:- Let T_n be a sufficient statistic.

Define $T_n^* = \Psi(T_n)$, where Ψ^{-1} exists.

Since T_n is sufficient, we have (by NFT),

$$f(x_1, \ldots, x_n; \theta) = g(\theta; T_n(x_1, \ldots, x_n))\, h(x_1, \ldots, x_n)$$

$$= g(\theta; \Psi^{-1}(T_n^*(x_1, \ldots, x_n)))\, h(x_1, \ldots, x_n) \quad \text{since } \Psi^{-1} \text{ exists.}$$

Hence, by NFT, T_n^* is also sufficient statistic for θ.

Remark:- NFT holds for more than one dimensional parameter as well

$$\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p).$$

It may happen that $(T_1, \ldots, T_q)$ is "jointly" sufficient for $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p)$, where $p \neq q$.

Further, even if $p = q$, we cannot claim T_i is sufficient for θ_i, $i = 1, \ldots, p$.

$\boxed{\sum_{i=1}^{n} x_i \quad \left\{\left(\sum_{i=1}^{n} x_i\right)^2\right\}^{1/2}}$

Examples of sufficient statistic (by NFT) :-

Example 1 :- Let $X_1, \ldots, X_n \overset{i.i.d}{\sim} N(\mu, 1)$, $\mu \in \mathbb{R}$.

$\hookrightarrow$ Want to find suff. stat. for μ.

$$f(x_1, \ldots, x_n ; \mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}, \quad x_i \in \mathbb{R}, \mu \in \mathbb{R}.$$

check

$$= \underbrace{e^{-\left(-\frac{n}{2}\mu^2 + \mu \sum_{i=1}^{n} x_i\right)}}_{g(\mu, T_n(x_1, \ldots, x_n))} \times \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2}}_{h(x_1, \ldots, x_n)},$$

Hence, $\sum_{i=1}^{n} X_i$ is sufficient ~~for~~ statistic for μ.

Example 2:– $X_1, \ldots, X_n \overset{i.i.d}{\sim} \text{Unif}(0, \theta)$, $\theta > 0$.

Want to find suff. stat. for θ.

$$f(x_1, \ldots, x_n; \theta) \begin{cases} = \dfrac{1}{\theta^n} & \text{if } 0 < x_1 \ldots x_n < \theta \\ = 0 & \text{o.w.} \end{cases}$$

$$\begin{cases} = \dfrac{1}{\theta^n} & \text{if } 0 < x_{(1)} \ldots < x_{(n)} < \theta \\ = 0 & \text{o.w.} \end{cases}$$

$$f(x_1, \ldots, x_n; \theta) = \frac{1}{\theta^n} \mathbb{1}_{(0, x_{(1)})} \mathbb{1}_{(x_{(n)}, \theta)}$$, where $\mathbb{1}_{(a,b)} = 1$ if $a < b$

$= 0$, o.w.

$$= \left\{ \frac{1}{\theta^n} \mathbb{1}_{(x_{(n)}, \theta)} \right\} \times \mathbb{1}_{(0, x_{(1)})}$$

$\{\ldots\} \longrightarrow g(\theta; T_n(x_1 \ldots x_n))$

$\mathbb{1}_{(0, x_{(1)})} \longrightarrow$ f^n of $(x_1, \ldots, x_n)$ $\longrightarrow h(x_1, \ldots, x_n)$.

Hence, by NFT, $X_{(n)}$ is sufficient stat for θ.

<u>Example:—</u> $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$.

(μ, σ^2): parameter

$\Theta = \mathbb{R} \times \mathbb{R}^+$

$$f(x_1, \ldots, x_n; (\mu, \sigma^2)) \overset{\text{Check}}{=} \frac{1}{\sigma^n} e^{\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i\right)} \times \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{n\mu^2}{2\sigma^2}}$$

$h(x_1, \ldots, x_n)$

$g((\mu, \sigma^2); T_n(x_1, \ldots, x_n))$

Here $T_n = \left(\sum_{i=1}^{n} X_i^2, \sum_{i=1}^{n} X_i\right)$ is "jointly" sufficient for (μ, σ^2).