Minimal sufficient statistic

* Among all sufficient statistic, a minimal sufficient statistic reduces the data as much as possible while retaining sufficiency.

* Minimal sufficient statistics are not unique but are equivalant up to a one-to-one transformation.

* The partition of the sample space induced by a minimum sufficient statistic is the coarsest possible among sufficient statistics.

$T(X) = X$ is always a sufficient statistic, with $h(x) = 1$

$$\Rightarrow \quad f(x|\theta) = g(T(x)|\theta)\, h(x)$$

$$= g(x|\theta)\, 1$$

$$= f(x|\theta)$$

and every one-to-one function of sufficient statistic is also a sufficient statistic

let $T(x)$ is a sufficient statistic

and η is a one-to-one function

$\Rightarrow \eta^{-1}$ exists.

Then $T^*(x) = \eta(T(x)) \ \forall x$ is also sufficient statistic.

$$f(x|\theta) = g(T(x)|\theta)\, h(x)$$

$$= g(\eta^{-1}(T^*(x))|\theta)\, h(x)$$

Define $g^*(t.|\theta)$

$$= g(\eta^{-1}(t)|\theta)$$

$\Rightarrow$ $f(x|\theta) = g^*(T^*(x)|\theta)\, h(x)$

$\Rightarrow$ $T^*(x)$ is also a sufficient statistic

$\Rightarrow$ There are many sufficient statistic.

$\Rightarrow$ The purpose of a sufficient statistic is to achieve data-reduction without loss of information about the parameter θ.

$\Rightarrow$ Thus, a statistic that achieves the most data reduction while still retaining all the information about θ might be considered preferable.

Def: 6.2.11 :-

A sufficient statistic $T(X)$ is called a minimal sufficient statistic if, for any other sufficient statistic $T'(X)$ $T(X)$ is a function of $T'(X)$

$T(x)$ must be derivable from $T'(x)$

* This makes $T(x)$ a "minimal" way of summarizing the data, and no other sufficient statistic is more consice than $T(x)$

$\Rightarrow$ The minimal sufficient statistic is the most efficient summary of the data, meaning it cannot be reduced further without losing information about θ.

we know that

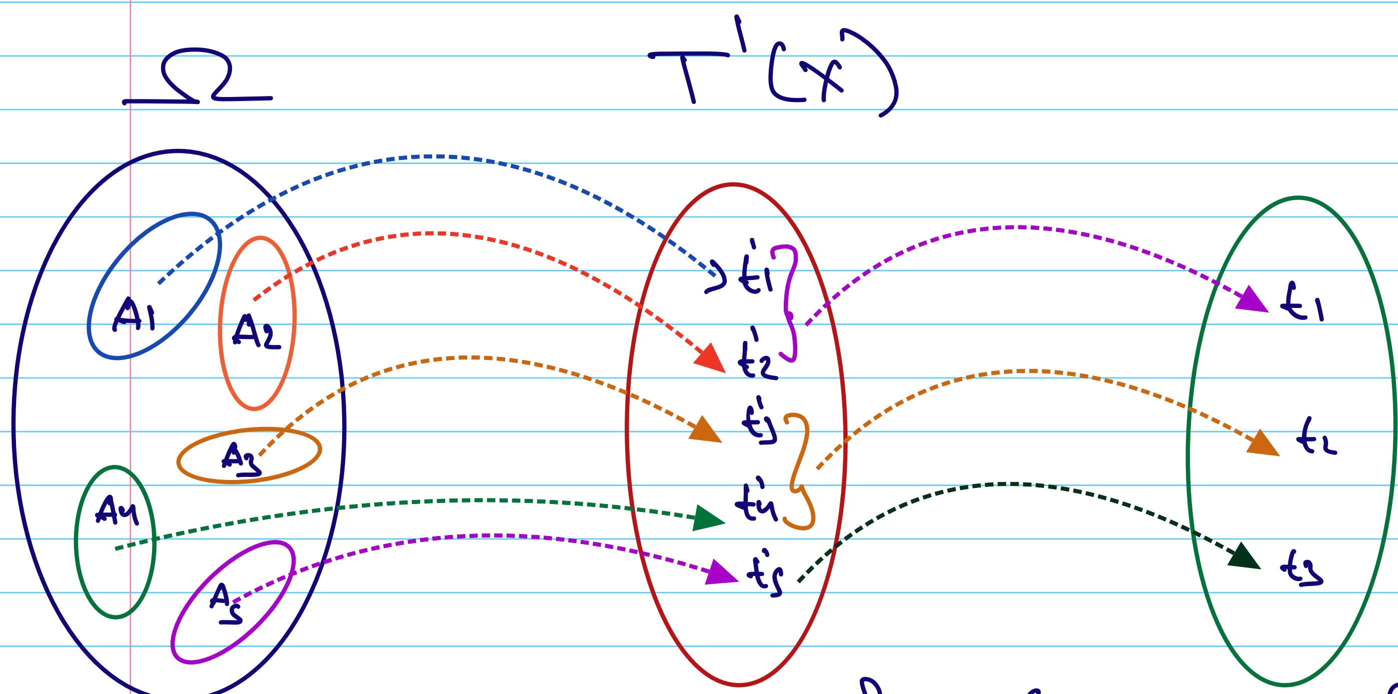
$$A_t = \{x : T(x) = t\}$$

To say $T(x)$ is a function of $T'(x)$ simply mean's if $T'(x) = T'(y) \implies T(x) = T(y)$

<u>Understanding:</u>

$$T(x) = c\left(T'(x)\right)$$

where the function c does not need to be one-to-one. (can be many-to-one)



$\{B_{t'} : t' \in T'\}$ $\{A_t : t \in T\}$

$B_{t'}$ is some subset of some A_t.

Theorem 6.2.13 :-

Let $f(x|\theta)$ be the pmf or pdf of a sample X. Suppose there exists a function $T(x)$ such that, for every two sample points x and y, the ratio

$$\frac{f(x|\theta)}{f(y|\theta)}$$ is constant as a function of $\theta \iff T(x) = T(y)$. Then $T(x)$ is a minimal sufficient statistic for θ.

Proof: The sample space $\Omega = \mathcal{X}$

$x \in \mathcal{X}$.

Let $\mathcal{T} = \{t : t = T(x) \text{ for some } x \in \mathcal{X}\}$

$\Rightarrow T(x) : \mathcal{X} \longrightarrow \mathcal{T}$

The partition sets induced by $T(x)$ on $\mathcal{X}$ are $A_t = \{x : T(x) = t\}$

let's take $x \in A_t \subseteq \mathcal{X}$

let's take one element for each partition.

$x_{T(x)}$ is the fixed element for each set.

$\Rightarrow$ for any $x \in \mathcal{X}$ choosen randomly $x_{T(x)}$ is the fixed element in that partition

$$\Rightarrow \quad T(x) = T(x_{T(x)})$$

$$f(x|\theta) = g(T(x)|\theta)\, h(x)$$

$$\Rightarrow \frac{f(x|\theta)}{f(x_{T(x)}|\theta)} = \frac{g(T(\cancel{x})|\theta)\, h(x)}{g(T(\cancel{x}_{T(x)})|\theta)\, h(x_{T(x)})}$$

$$= \frac{h(x)}{h(x_{T(x)})}$$

independent of θ.

* define a function on $\mathcal{X}$ by

$$h(x) = \frac{f(x|\theta)}{f(x_{T(x)}|\theta)}$$

(independent of θ)

(this is not similar to previous $h(x)$)

* define a function on $\mathcal{T}$ by

$$g(t|\theta) = f(x_t|\theta)$$

$$\Rightarrow f(x|\theta) = \frac{f(x_{T(x)}|\theta)}{f(x_{T(x)}|\theta)} \cdot f(x|\theta)$$

$$= f(x_{T(x)}|\theta) \cdot h(x)$$

$$= g(T(x)|\theta) \cdot h(x)$$

$\Rightarrow$ $T(x)$ is sufficient statistic

Now we need to show $T(x)$ is minimal.

Let's take $T'(x)$ is another sufficient statistic

$$\Rightarrow \qquad f(x|\theta) = g(T(x)|\theta)\, h(x)$$
$$= g'(T'(x)|\theta)\, h'(x)$$

Let x and y be any two sample points with $T'(x) = T'(y)$. Then

$$\frac{f(x|\theta)}{f(y|\theta)} = \frac{g'(T'(x)|\theta)\, h'(x)}{g'(T'(y|\theta)\ \ h'(y)}$$

$$= \frac{h'(x)}{h'(y)}$$

Ex: 6.2.14 (Normal minimal sufficient statistic)

$X_1, X_2, \ldots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Both μ, σ^2 are unknown.

Let x, y are two data points, and

let $(\bar{x}, s_x^2), (\bar{y}, s_y^2)$ be the

sample mean, sample variances.

$$\frac{f(x|\mu,\sigma^2)}{f(y|\mu,\sigma^2)} = \frac{\frac{1}{\cancel{(2\pi\sigma^2)^{n/2}}} \exp\left(-\frac{1}{2\sigma^2}\left(n(\bar{x}-\mu)^2 + (n-1)s_x^2\right)\right)}{\frac{1}{\cancel{(2\pi\sigma^2)^{n/2}}} \exp\left[-\frac{1}{2\sigma^2}\left(n(\bar{y}-\mu)^2 + (n-1)s_y^2\right)\right)}$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left(n\left[(\bar{x}-\mu)^2 - (\bar{y}-\mu)^2\right] + (n-1)(s_x^2 - s_y^2)\right)\right)$$

if $\bar{x} = \bar{y}$ & $S_x^2 = S_y^2$

$$= \exp\left(-\frac{1}{2\sigma^2}\left(n \cdot 0 + (n-1) \cdot 0\right)\right)$$

$= 1$

$\Rightarrow \dfrac{f(x|\mu,\sigma^2)}{f(y|\mu,\sigma^2)}$ is constant if $\bar{x} = \bar{y}$, $S_x^2 = S_y^2$

$\Rightarrow (\bar{X}, S^2)$ is minimal sufficient statistic.

Example 6.2.15 (Uniform minimal sufficient statistic)

$X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{unif}(\theta, \theta+1)$

then $f(x|\theta) = \begin{cases} 1 & \theta < x_i < \theta+1, \; i = 1, 2, \ldots n \\ 0 & \text{o.w} \end{cases}$

we can write

$$f(x|\theta) = \begin{cases} 1 & \max_i x_i - 1 < \theta < \min_i x_i \\ 0 & \text{o.w} \end{cases}$$

$\Rightarrow$ for two random samples x, y

$$\frac{f(x|\theta)}{f(y|\theta)}$$