$T(\underline{X}) = \underline{X}$ is sufficient

Let $T(\underline{X})$ be sufficient and define $T^{\circ}(\underline{X}) = r(T(\underline{X}))$, $\forall X$

Then $f(\underline{x}|\theta) = g(T(\underline{x})|\theta) = g(r^{-1}(T^{\circ}(\underline{x}))|\theta)\, h(\underline{x})$

$\Rightarrow T^{\circ}(\underline{x})$ is also sufficient.

Theorem 6.2.10

$X_1, \dots, X_n$ iid from a pmf/pdf $f(x|\underline{\theta})$

where $f(x|\underline{\theta}) = h(x)\, c(\underline{\theta})\, e^{\sum_{i=1}^{k} w_i(\underline{\theta}) t_i(x)}$

and $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$, $d \le k$. Then

$T(\underline{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \dots, \sum_{j=1}^{n} t_k(X_j)\right)$ is a sufficient statistic for $\underline{\theta}$

Proof.

$$f(\underline{x}|\underline{\theta}) = \prod_{j=1}^{n} h(x_j)\, c(\underline{\theta})\, e^{\sum_{i=1}^{k} w_i(\underline{\theta}) t_i(x_j)}$$

$$= \underbrace{\prod_{j=1}^{n} h(x_j)}_{h(x)} \underbrace{\left((c(\underline{\theta}))^{n}\, e^{\sum_{i=1}^{k} w_i(\underline{\theta}) \sum_{j=1}^{n} t_i(x_j)} \right)}_{g(T_1(\underline{x}), \dots, T_k(x)|\theta)}$$

Example

$$f(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma}\, e^{-\frac{\mu^2}{2\sigma^2} - \frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}}$$

$t_1(x) = x$, $w_1(\mu, \sigma^2) = \frac{\mu}{\sigma^2}$, $t_2(x) = x^2$, $w_2(\mu, \sigma^2) = -\frac{1}{2\sigma^2}$

$\Rightarrow$ for $X_1, \dots, X_n$ that.

$\sum_{j=1}^{n} X_j$ and $\sum_{j=1}^{n} X_j^2$ are sufficient for (μ, σ^2).

Example

$X_1, X_2, X_3 \sim$ Bernoulli(θ) and let $T(\underline{X}) = X_1 + X_2 + X_3$

$T'(\underline{X}) = \{X_1, X_2 + X_3\}$

X - space

(0,0,0)	(0,0,1)
(1,0,0)	(1,0,1)
(0,1,0)	(0,1,1)
(1,1,0)	(1,1,1)

T' - space

(0,0)	(1,0)
(0,1)	(1,1)
(0,2)	(1,2)

T - space

0
1
2
3

$$B_{T'(\underline{x})} = \{\underline{y} : T'(\underline{y}) = T'(\underline{x}) = \underline{t}\} = B_t'$$

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

$B_{(0,0)}$	(0,0,0)	(1,0,0)	$B_{(1,0)}$
$B_{(0,1)}$	(0,1,0), (0,0,1)	(1,0,1), (1,0,1)	$B_{(1,1)}$
$B_{(0,2)}$	(0,1,1)	(1,1,1)	$B_{(1,2)}$

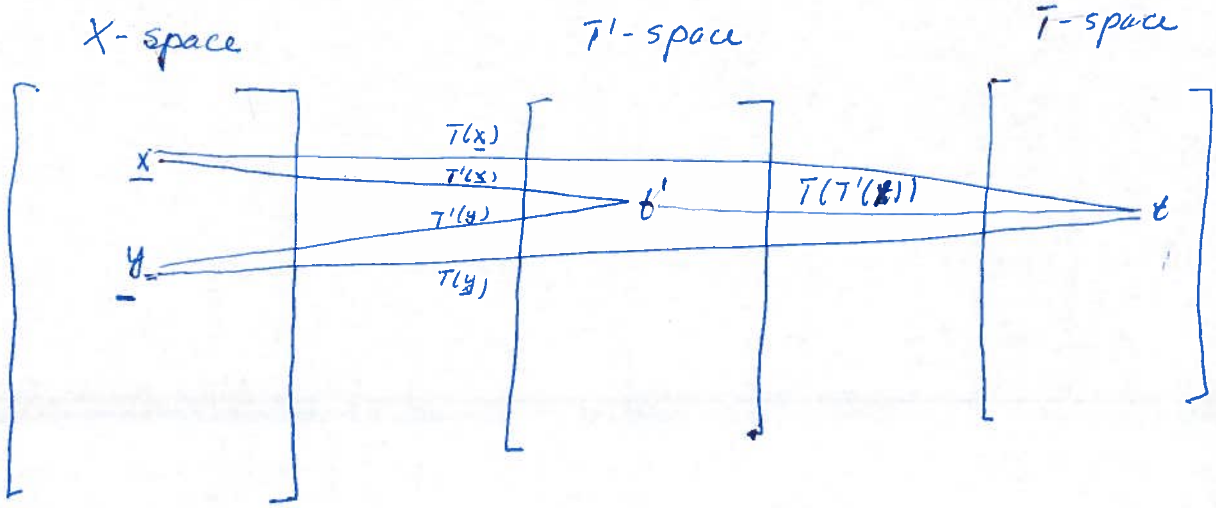
A_0	(0,0,0)	(0,1,1), (1,0,1), (1,1,0)	A_2
A_1	(1,0,0), (0,1,0), (0,0,1)	(1,1,1)	A_3

Both $T'(\underline{x})$ and $T(\underline{x})$ are sufficient but clearly $T(\underline{x})$ achieves a greater data reduction. Note $T(\underline{x}) = f(T'(\underline{x}))$ which means that if $T'(\underline{x}) = T'(\underline{y}) \Rightarrow T(\underline{x}) = T(\underline{y})$

Definition 6.2.11

A sufficient statistic $T(\underline{x})$ is called a minimal sufficient statistic if for any other sufficient statistic $T'(\underline{x})$, $T(\underline{x})$ is a function of $T'(\underline{x})$. which means that $T'(\underline{x}) = T'(\underline{y}) \Rightarrow T(\underline{x}) = T(\underline{y})$

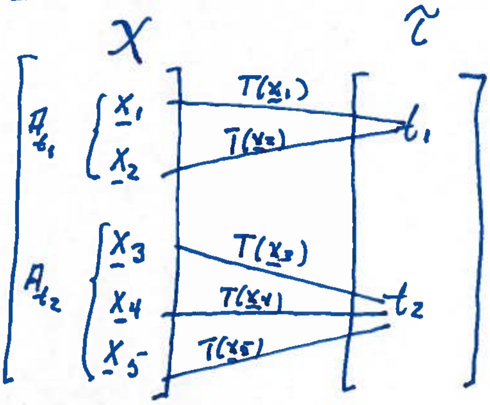
$T(\underline{X})$ minimal suffiaint, $T'(\underline{X})$ sufficient



Theorem 6.2.13

Let $f(\underline{x}|\theta)$ be the pmf/pdf of a sample $\underline{x}$. Suppose there exists a statistic $T(\underline{x})$ such that for every two sample points $\underline{x}$ and $\underline{y}$, we have $\frac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)}$ is a constant as a function of $\theta \Leftrightarrow T(\underline{x}) = T(\underline{y})$.

Then $T(\underline{x})$ is a minimal sufficient statistic for θ.



Proof

$T(\underline{x})$ sufficient. Let $A_t = \{\underline{x} : T(\underline{x}) = t\}$. For each A_t fix an $\underline{x}_t \in A_t$ and let for any $\underline{x} \in \mathcal{X}$, $\underline{x}_{T(\underline{x})}$ be the fixed element in the same partition set as $\underline{x}$. Define a function on $\mathcal{T}$ as

$g(t|\theta) = f(\underline{x}_t|\theta)$.

Then if $T(\underline{x}) = T(\underline{x}_{T(\underline{x})}) \Leftrightarrow \frac{f(\underline{x}|\theta)}{f(\underline{x}_{T(\underline{x})}|\theta)} = h(\underline{x})$ (a constant as a function of θ), we have

$$f(\underline{x}|\theta) = f(\underline{x}_{T(\underline{x})}|\theta) \cdot \frac{f(\underline{x}|\theta)}{f(\underline{x}_{T(\underline{x})}|\theta)} = g(T(\underline{x})|\theta)\, h(\underline{x})$$

$\Leftrightarrow T(\underline{x})$ is sufficient (factorization theorem).

$T(\underline{x})$ minimal sufficient

Let $T'(\underline{x})$ be any other sufficient statistic and let $\underline{x}$ and $\underline{y}$ be any two sample points such that $T'(\underline{x}) = T'(\underline{y})$

By the factorization theorem there exists g' and h' such that $\dfrac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)} = \dfrac{g'(T'(\underline{x})|\theta)\cdot h'(\underline{x})}{g'(T'(\underline{y})|\theta)\; h'(\underline{y})} = \dfrac{h'(\underline{x})}{h'(\underline{y})}$ is a constant as a function of θ $\Rightarrow$ $T(\underline{x}) = T(\underline{y})$ and $T(\underline{x})$ is a function of $T'(\underline{x})$.

Example. $X_1, \ldots, X_n$ iid $\sim N(\mu, \sigma^2)$, μ and σ^2 unknown.

Let $\underline{x}$ and $\underline{y}$ be two sample points. Then

$$\frac{f(\underline{x}/\mu,\sigma^2)}{f(\underline{y}/\mu,\sigma^2)} = \frac{(2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n(\bar{x}-\mu)^2}{2\sigma^2} - \frac{\sum(x_i-\bar{x})^2}{2\sigma^2}}}{(2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n(\bar{y}-\mu)^2}{2\sigma^2} - \frac{\sum(y_i-\bar{y})^2}{2\sigma^2}}}$$

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

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

is a constant as a function of μ and $\sigma^2 \Leftrightarrow \bar{x} = \bar{y}$ and $s_x^2 = s_y^2 \Leftrightarrow (\bar{x}, s^2)$ is minimal sufficient for (μ, σ^2)

Suppose $f(x/\underline{\theta}) = h(x)\, c(\underline{\theta})\, e^{\sum_{i=1}^{n} w_i(\underline{\theta})\, t_i(x)}$

$\Rightarrow f(\underline{x}/\underline{\theta}) = c(\underline{\theta})^n \prod_{j=1}^{n} h(x_j)\, e^{\sum_{i=1}^{n} w_i(\underline{\theta}) \sum_{j=1}^{n} t_i(x_j)}$

Then $$\frac{f(\underline{x}/\underline{\theta})}{f(\underline{y}/\underline{\theta})} = \frac{c(\underline{\theta})^n}{c(\underline{\theta})^n} \frac{\prod_{j=1}^{n} h(x_j)}{\prod_{j=1}^{n} h(y_j)}\, e^{\sum_{i=1}^{k} w_i(\underline{\theta})\left[\sum_{j=1}^{n} t_i(x_j) - \sum_{j=1}^{n} t_i(y_j)\right]}$$

Constant as a function of $\theta \Longleftrightarrow \sum_{j=1}^{n} t_i(x_j) = \sum_{j=1}^{n} t_i(y_j)$, $i = 1, 2, \ldots, k$

$\Rightarrow \underline{T}(\underline{X}) = \left[\sum_{j=1}^{n} t_1(X_j), \ldots \sum_{j=1}^{n} t_k(X_j)\right]$ is minimal sufficient.

under some restrictions on $w_i(\theta)$, $i = 1, 2, \ldots k$.

They must be linearly independent.