Sufficient Statistic

A statistic $T(\mathbf{X})$ is a *sufficient statistic for* θ if the conditional distribution of the sample $\mathbf{X}$ given the value of $T(\mathbf{X})$ does not depend on θ.

Minimal Sufficient Statistic

A sufficient statistic $T(\mathbf{X})$ is called a *minimal sufficient statistic* if, for any other sufficient statistic $T'(\mathbf{X})$, $T(\mathbf{x})$ is a function of $T'(\mathbf{x})$.

Ancillary Statistic

A statistic $S(\mathbf{X})$ whose distribution does not depend on the parameter θ is called an *ancillary statistic*.

Ratio Theorem: Sufficient Statistic

$T(\mathbf{X})$ is a sufficient statistic for θ if, for every $\mathbf{x}$ in the sample space, the ratio

$$\frac{p(\mathbf{x}|\theta)}{q(T(\mathbf{x})|\theta)}$$

is constant as a function of θ.

Ratio Theorem: Minimal Sufficient

$T(\mathbf{X})$ is a minimal sufficient statistic for θ if, for every $\mathbf{x}$ and $\mathbf{y}$, the ratio

$$f(\mathbf{x}|\theta)/f(\mathbf{y}|\theta)$$

is constant as a function of θ iff $T(\mathbf{x}) = T(\mathbf{y})$.

Basu's Theorem

If $T(\mathbf{X})$ is a complete and minimal sufficient statistic, then,

$T(\mathbf{X})$ is independent of every ancillary statistic.

Factorization Theorem

$T(\mathbf{X})$ is a sufficient statistic for θ if and only if there exist functions $g(t|\theta)$ and $h(\mathbf{x})$ such that, for all sample points $\mathbf{x}$ and all parameter points θ,

$$f(\mathbf{x}|\theta) = g(T(\mathbf{x})|\theta)h(\mathbf{x})$$

Complete and Minimal Sufficient

If a minimal sufficient statistic exists, then,

any complete statistic is also a minimal sufficient statistic.

Complete Statistic

The family $f(t|\theta)$ of pdfs or pmfs for a statistic $T(\mathbf{X})$ is called *complete* if

$$\mathrm{E}_\theta g(T) = 0 \text{ for all } \theta$$

$$\text{implies } P_\theta(g(T) = 0) = 1 \text{ for all } \theta$$

Exponential Sufficient Statistic

If $f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})e^{\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)}$ then

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

is a sufficient statistic for $\boldsymbol{\theta}$

Exponential Complete Statistic

$$T(\mathbf{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \ldots, \sum_{i=1}^{n} t_k(X_i)\right)$$

is complete if $\{(w_1(\boldsymbol{\theta}), \ldots, w_k(\boldsymbol{\theta})) : \boldsymbol{\theta} \in \Theta\}$ contains an open set in $\mathbb{R}^k$.