6.2.4 Sufficient, Ancillary, and Complete statists.

$\Rightarrow$ A minimal sufficient statistic is a statistic that has achieved the maximal amount of data reduction possible while still retaining all the information about the parameter's θ.

$\Rightarrow$ distribution of ancillary statistic does not depend on θ.

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

then $(X_{(1)}, X_{(n)})$ is a minimal sufficient statistic

$\Rightarrow$ $\left(X_{(n)} - X_{(1)}, \frac{X_{(n)} + X_{(1)}}{2}\right)$

is also a minimal sufficient statisl

$\Rightarrow$ $X_{(n)} - X_{(1)}$ is a ancillary statistic is a minimal sufficient statistic.

Ex:- 6.2.20 (Ancillary Precision)

$X_1, X_2 \overset{iid}{\sim} X$

$$P_\theta(X=\theta) = P_\theta(X=\theta+1) = P_\theta(X=\theta+2) = \frac{1}{3}$$

$$(R, M) = \left(X_{(n)} - X_{(1)},\ \frac{X_{(1)} + X_{(2)}}{2}\right)$$

is the minimal sufficient statistic.

& R is ancillary statistic.

+ The knowledge of the value of R alone would give us no information about θ, but it increased our knowledge about θ along with M

For many important situation's, however, our intuition that a minimal sufficient statistic is independent of any ancillary statistic is correct.

Definition 6.2.21 :-

Let $f(t \mid \theta)$ be a family of pdf's or pmf's for a statistic $T(X)$. The family of probability distribution's is called complete if

if
$$E_\theta[g(T)] = 0 \quad \forall \theta$$
$$\Downarrow$$
$$\mathbb{P}_\theta(g(T) = 0) = 1 \quad \forall \theta$$

Equivalently $T(X)$ is called a complete statistic.

Linear algebra Analogy: Completeness

(1) in linear algebra, a set of vector's $\{v_1, v_2, \ldots, v_n\}$ is complete if it span's the entire vector space.

$\Rightarrow$ This mean's any vector ω in the space can be written as a linear combination of $\{v_1, v_2, \ldots v_n\}$

$\Rightarrow$ if ω is orthogonal to all vector's in $\{v_1, v_2, \ldots, v_n\}$, then $\omega = 0$ (the 0 vector)

(2) For Completeness in statistics

$\Rightarrow$ The family of distribution's $\{f(t|\theta)\}$ is analogous to the set of vector's $\{v_1, v_2, \ldots, v_n\}$

$\Rightarrow$ A function $g(T)$ is analogous to a vector ω

$\Rightarrow$ Completeness mean's the family $\{f(t|\theta)\}$ span's the entire "function space" of T, leaving no direction's (function's $g(T)$) orthogonal to it except for the zero function.

③ if $g(T) \neq 0$ and $E_\theta[g(T)] = 0$ for all θ, it implies.

$\Rightarrow$ The family $\{f(t|\theta)\}$ does not span the whole function space.

$\Rightarrow$ Completeness is violated, as $g(T)$ represents a "hidden direction" orthogonal to all $f(t|\theta)$

Complete Statistics: Linear Algebra Analogy

1. **Family of Distributions as Vectors**:
   - The family of distributions $f(t|\theta)$ can be thought of as a set of vectors in a vector space of functions of T.
2. **Completeness Definition in Linear Algebra Terms**:
   - A statistic T is **complete** if:

$$E_\theta[g(T)] = 0 \text{ for all } \theta \implies g(T) = 0 \text{ for all } T.$$

   - This means there is no function $g(T) \neq 0$ that is orthogonal to all $f(t|\theta)$.
3. **Orthogonality and the Inner Product**:
   - The expectation $E_\theta[g(T)] = 0$ is analogous to the inner product of two vectors being 0 (orthogonality):

$$\langle g(T), f(t|\theta) \rangle = \int g(T) f(t|\theta) dt = 0.$$

4. **Spanning the Space**:
   - Completeness implies that the family $f(t|\theta)$ spans the entire function space of T.
   - Any function $g(T)$ in this space can be written as a "linear combination" of $f(t|\theta)$.
5. **No Orthogonal Directions Left**:
   - If $g(T) \neq 0$ exists and is orthogonal to all $f(t|\theta)$, it lies outside the span of $f(t|\theta)$.
   - Completeness ensures there are no such $g(T)$; the family $f(t|\theta)$ is sufficient to describe all functions of T.

Key Points on Completeness

1. **Why Completeness Matters**:
   - Completeness ensures that the statistic T fully captures all the information about θ.
   - No "extra" functions $g(T)$ are left unexplained by $f(t|\theta)$.
2. **Connection to Sufficiency**:
   - Completeness often arises in the context of **minimal sufficient statistics**.
   - If a statistic T is sufficient and complete, it's often the best summary of the data for inference.
3. **What if $g(T) \neq 0$**:
   - If $g(T) \neq 0$ and is orthogonal to $f(t|\theta)$, it represents a "direction" in the function space that $f(t|\theta)$ does not capture.
   - This would mean T is **not complete**, as the family $f(t|\theta)$ is not rich enough to span the space.
4. **Analogy to Linear Independence**:
   - Completeness is similar to a set of vectors being **linearly independent** and spanning the space.
   - If the family $f(t|\theta)$ is complete, it forms a "basis" for the function space of T.

To Remember

- **Completeness ensures no hidden dimensions in the function space.**
- **The family $f(t|\theta)$ must span the entire space, leaving no orthogonal directions unexplained.**
- It is a critical property in statistical theory, especially when working with sufficient statistics.

Example 6.2.22 (Binomial complete sufficient statistic)

$$T(X) \sim \text{binomial}(n, p)$$

where $0 < p < 1$

$$E_p(g(T)) = \sum_{t=0}^{n} g(t) \binom{n}{t} p^t (1-p)^{n-t} = 0$$

$$\Rightarrow \quad (1-p)^n \sum_{t=0}^{n} g(t) \binom{n}{t} \left(\frac{p}{1-p}\right)^t = 0$$

$$\Rightarrow \quad g(t) = 0 \quad \forall t \in \mathcal{T}$$

$\Rightarrow$ $T(X)$ is a complete statistic

Theorem 6.2.24 (Basu's theorem)

if $T(x)$ is a complete and minimal sufficient statistic, then $T(x)$ is independent of every ancillary statistic.

Proof:

① understanding the setup:

we are given

(i) $S(x)$ is ancillary, meaning

$\mathbb{P}(S(x) = s)$ does not depend on θ

(ii) $T(x)$ is sufficient, meaning all info about θ in the data X is captured by $T(x)$

The goal is to show that $S(x)$ and $T(x)$ are independent, which means:

$$P(S(x)=s \mid T(x)=t) = \mathbb{P}(S(x)=s)$$
$$\forall t \in \mathcal{T}$$

② Law of total Probability

Using the law of total Probability,

we can write the marginal Probability of $S(x)=s$ as

$$\mathbb{P}(S(x)=s) = \sum_{t \in \mathcal{T}} P(S(x)=s \mid T(x)=t)\, P_\theta(T(x)=t),$$

③ Using completeness of $T(x)$

Assume independence of $S(x)$ & $T(x)$

if they are independent, we can writ.

$$\mathbb{P}(S(x)=s) = \sum_{t \in \mathcal{T}} \mathbb{P}(S(x)=s \mid T(x)=t) \cdot P_\theta(T(x)=t)$$

$$= \sum_{t \in \mathcal{T}} \mathbb{P}(S(x)=s) \cdot P_\theta(T(x)=t)$$

$$= \mathbb{P}(S(x)=s) \sum_{t \in \mathcal{T}} P_\theta(T(x)=t)$$

$$= \mathbb{P}(S(x)=s)$$

4 define $g(t)$

$$g(t) = \mathbb{P}(S(x)=s \mid T(x)=t) - P(S(x)=s)$$

$$\Rightarrow E_\theta(g(T)) = \sum_{t \in \mathcal{T}} g(T)\, \mathbb{P}(T(x)=t)$$

$$= \sum_{t \in \mathcal{T}} \Big[\mathbb{P}(S(x)=s \mid T(x)=t) - \mathbb{P}(S(x)=s) \Big] \mathbb{P}(T(x)=t)$$

$$\Rightarrow \sum_{t \in \mathcal{T}} \mathbb{P}(S(x)=s \mid T(x)=t) \cdot \mathbb{P}(T(x)=t) - \sum_{t \in \mathcal{T}} \mathbb{P}(S(x)=s) \cdot \mathbb{P}(T(x)=t)$$

$$\Rightarrow \quad \mathbb{P}(S(X)=s) - \sum_{t \in \mathcal{T}} \mathbb{P}(S(X)=s) \cdot \mathbb{P}(T(X)=t)$$

$$\Rightarrow \quad \mathbb{P}(S(X)=s) - \mathbb{P}(S(X)=s) = 0$$

therefore $\mathbb{E}_\theta [g(T)] = 0$

$\Rightarrow$ according to completeness of $T(X)$

$$g(T) = 0 \quad \forall t \in \mathcal{T}$$

$$\Rightarrow \mathbb{P}(S(X)=s \mid T(X)=t) - \mathbb{P}(S(X)=s) = 0$$

$$\Rightarrow \quad \mathbb{P}(S(X)=s \mid T(X)=t) = P(S(X)=s$$

$\Rightarrow$ Both $S(X)$, $T(X)$ are independent.

Example 6.2.26

$X_1, X_2, \ldots, X_n \sim \exp(\theta)$

exponential distributions.

$$f(x|\theta) = \frac{1}{\theta^n} e^{-\frac{\sum x_i}{\theta}}$$

$\Rightarrow \sum x_i$ is sufficient statistic

$$\frac{f(x|\theta)}{f(y|\theta)} = e^{-\frac{1}{\theta}\left(\sum x_i - \sum y_i\right)}$$

$\Rightarrow T(x) = T(y)$ to make

$\frac{f(x|\theta)}{f(y|\theta)}$ to be independent of θ

$\Rightarrow T(x) = \sum x_i$ is minimal sufficient statistic.

Let $g(x) = \dfrac{x_n}{x_1 + \cdots + x_n}$

Theorem 6.2.28 :- if a minimal sufficient statistic exists, then any complete statistic is also a minimal sufficient statistic.