**Definition.** Let $\{f_T(t|\theta): \theta \in \Theta\}$ be a family of pdfs or pmfs for a statistic $T(\underset{\sim}{X})$.

We say this family is complete if

$$E_\theta[g(T)] = 0 \quad \text{for all } \theta \in \Theta$$

$$\Rightarrow \quad P_\theta[g(T)=0] = 1 \quad \forall \theta \in \Theta$$

we call $T = T(\underset{\sim}{X})$ a complete statistic

$X_1, \cdots, X_n \overset{iid}{\sim} f(x|\theta)$

$\theta \in \Theta$

The sampling distribution of $T = T(\underset{\sim}{X})$ depends on $\theta$

---

**Example.** (Binomial complete sufficient statistic) $\Theta = (0,1)$

Suppose that $X_1, \cdots, X_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$, $0<\theta<1$

show that $T = T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ is a complete sufficient statistic for $\theta$

**Ans.** Sufficiency ✓

Completeness: $T(\underset{\sim}{X}) = \sum_{i=1}^n X_i \sim \text{Binomial}(n, \theta)$ $0<\theta<1$

$$\{f_T(t|\theta): \theta \in (0,1)\} = \{\underset{\text{pmf}}{\text{Binomial}}(n,\theta): \theta \in (0,1)\}$$

we will show if $E_\theta[g(T)] = 0 \quad \forall \theta \in (0,1)$

Then $P_\theta[g(T)=0] = 1 \quad \forall \theta \in (0,1)$

$$0 = E_\theta[g(T)] = \sum_{t=0}^{n} g(t) f_T(t|\theta)$$

$$= \sum_{t=0}^{n} g(t) \binom{n}{t} \theta^t (1-\theta)^{n-t} \qquad \forall \theta \in (0,1)$$

$$= \underbrace{(1-\theta)^n}_{>0} \sum_{t=0}^{n} g(t) \binom{n}{t} \Big(\underset{\overset{\shortparallel}{r}}{\frac{\theta}{1-\theta}}\Big)^t$$

$$r = \frac{\theta}{1-\theta} \in (0, +\infty)$$

$$0 = \sum_{t=0}^{n} g(t) \binom{n}{t} r^t \quad \forall\, r \in (0, +\infty)$$

$$= g(0)\binom{n}{0} \times 1 + g(1)\binom{n}{1} \times r$$

$$+ g(2)\binom{n}{2} \times r^2 + \cdots + g(n)\binom{n}{n} r^n$$

$$= a_0 + a_1 r + a_2 r^2 + a_3 r^3 + \cdots + a_n r^n \quad \text{for all } r \in (0, +\infty)$$

$$\Rightarrow \quad a_0 = a_1 = \cdots = a_n = 0$$

$$\Rightarrow \quad g(0) = g(1) = \cdots = g(n) = 0$$

$$\Rightarrow \quad P_\theta[g(T) = 0] = 1 \quad \text{for all } \theta$$

---

Example. $X_1, \cdots, X_n \sim U(\theta, \theta+1)$ $\quad \theta \in \mathbb{R}$

$$T(\underset{\sim}{X}) = \begin{pmatrix} X_{(1)} \\ \\ X_{(n)} \end{pmatrix} \text{ is } \underset{\text{minimal}}{\text{sufficient}} \text{ for } \theta$$

Is $T(\underset{\sim}{X})$ complete? No

$$\frac{E[X_{(1)}]}{E[X_{(n)}]}$$

If $X_1, \cdots, X_n \sim U(\theta, \theta+1)$

$$E[X_{(1)}] = \frac{1}{1+n} + \theta$$

$$E[X_{(n)}] = \frac{n}{1+n} + \theta$$

$\Leftarrow$

$X_1, \cdots, X_n \sim U(0,1)$

$X_{(1)} \sim Beta(1, n)$ $\quad E[X_{(1)}] = \frac{1}{1+n}$

$X_{(n)} \sim Beta(n, 1)$ $\quad E[X_{(n)}] = \frac{n}{n+1}$

$$P(X_{(1)} > a) = \prod P(X_i > a)$$

$$= (1-a)^n$$

$$f_{X_{(1)}}(a) = n(1-a)^{n-1} \quad 0 < a < 1$$

$$E[X_{(1)}] - E[X_{(n)}] = \left(\frac{1}{1+n} + \theta\right) - \left(\frac{n}{1+n} + \theta\right)$$

$$= \frac{1-n}{1+n}$$

$$g(T) = g\left(\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}\right) = t_1 - t_2 - \frac{1-n}{1+n}$$ is non-zero

$$E_\theta[\,g(T(\underset{\sim}{X}))] = E_\theta\left[g\begin{pmatrix} X_{(1)} \\ X_{(n)} \end{pmatrix}\right] = E_\theta\left[X_{(1)} - X_{(n)} - \frac{1-n}{1+n}\right]$$

$$= \frac{1-n}{1+n} - \frac{1-n}{1+n} = 0$$

it implies $T(\underset{\sim}{X}) = \begin{pmatrix} X_{(1)} \\ X_{(n)} \end{pmatrix}$ is not complete.

Example. $X_1, \cdots, X_n \overset{iid}{\sim} Unif(0, \theta)$ $\theta > 0$

we know $T(\underset{\sim}{X}) = X_{(n)}$ is sufficient and complete.

Ancillary: $\frac{X_1}{X_2}$, $\frac{X_{(1)}}{X_{(n)}}$, $\frac{X_{(n)}}{X_{(1)}}$, ...

$X_{(n)}$ is independent of $\frac{X_{(1)}}{X_{(n)}}$

$$T(\underset{\sim}{X}) = X_{(n)}$$

is $T(\underset{\sim}{X}) = X_{(n)}$ complete?

$$\{ f_T(t|\theta) : \theta > 0 \}$$

$$P(X_{(n)} < a) = \prod_{i=1}^{n} P(X_i < a)$$

$$= \prod_{i=1}^{n} \left(\frac{a}{\theta}\right) = \frac{a^n}{\theta^n} \qquad a < \theta$$

$$f_T(t|\theta) = \begin{cases} n\, t^{n-1} \theta^{-n}, & 0 < t < \theta \\ 0, & o.w. \end{cases}$$

to verify completeness

if $E_\theta[g(T)] = 0$ $\forall \theta$ then $P_\theta[g(T) = 0] = 1$ for all $\theta$

$$0 = E_\theta[g(T)] = \int_0^\theta g(t)\, n t^{n-1} \theta^{-n} dt \qquad \text{holds for all } \theta$$

because $\theta > 0$ $\quad 0 = \int_0^\theta g(t)\, n t^{n-1} dt$

take $\frac{d}{d\theta}$:

$$0 = g(\theta)\, n\, \theta^{n-1} \quad \text{holds for all } \theta > 0$$

$$\Rightarrow \quad 0 = g(\theta) \text{ for all } \theta > 0 \Rightarrow P_\theta[g(T) = 0] = 1$$

Comment: we only considered Riemann-Integrable functions.
but it is good enough!

Basu's theorem. (Version I) Suppose $T(\underset{\sim}{X})$ is a sufficient statistic (for $\theta$)
If $T(\underset{\sim}{X})$ is also complete. Then $T(\underset{\sim}{X})$ is independent of every ancillary statistic $S(\underset{\sim}{X})$

---

(Version II) Suppose $T(\underset{\sim}{X})$ is sufficient. $S(\underset{\sim}{X})$ is ancillary
If $T(\underset{\sim}{X})$ and $S(\underset{\sim}{X})$ are not independent
Then $T(\underset{\sim}{X})$ is not complete.

---

Example $X_1, \cdots, X_n \overset{iid}{\sim} U(0, \theta)$

Find $E\left[\frac{X_{(1)}}{X_{(n)}}\right]$

we know $X_{(n)}$ and $\frac{X_{(1)}}{X_{(n)}}$ are independent (Basu's thm)

$$E[X_{(1)}] = E\left[X_{(n)} \times \frac{X_{(1)}}{X_{(n)}}\right] = E[X_{(n)}] \times E\left[\frac{X_{(1)}}{X_{(n)}}\right]$$

$$E\left[\frac{X_{(1)}}{X_{(n)}}\right] = \frac{E[X_{(1)}]}{E[X_{(n)}]}$$

# Exponential Family.

Let $X_1, \dots, X_n \overset{iid}{\sim} f(x|\underset{\sim}{\theta}) = h(x)\, c(\underset{\sim}{\theta}) \exp\left(\sum_{j=1}^{k} w_j(\underset{\sim}{\theta}) t_j(x)\right)$

where $\underset{\sim}{\theta} = (\theta_1, \dots, \theta_d)$, $d \le k$.

From previous ... : know that $T = \begin{pmatrix} \sum_{i=1}^{n} t_1(X_i) \\ \vdots \\ \sum_{i=1}^{n} t_k(X_i) \end{pmatrix}$ is sufficient for $\underset{\sim}{\theta}$

- If $d = k$, then $T$ is complete
- If $d < k$, then $T$ is not complete.

---

Example. $X_1, \dots, X_n \overset{iid}{\sim} \text{gamma}\left(\alpha, \frac{1}{\alpha^2}\right)$ $\theta = \alpha$, $d = 1$

$$f_X(x|\alpha) = \frac{1}{\Gamma(\alpha)\left(\frac{1}{\alpha^2}\right)^{\alpha}} x^{\alpha-1} e^{-x/\left(\frac{1}{\alpha^2}\right)} I(x>0)$$

$$= \frac{I(x>0)}{x} \frac{1}{\Gamma(\alpha)\left(\frac{1}{\alpha^2}\right)^{\alpha}} \exp\left\{-\alpha^2 x + \alpha \times \ln x\right\}$$

$$T = \begin{pmatrix} \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} \ln(X_i) \end{pmatrix}$$ is sufficient for $\alpha$

$k = 2 > d = 1$

so $T$ is not complete.!

Example $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ $-\infty < \mu < \infty$, $\sigma > 0$

prove $\bar{X}_n \perp\!\!\!\perp S_n^2$ using Basu's theorem.

$\underset{\text{Sample mean}}{\uparrow}$ $\underset{\text{Sample variance}}{\uparrow}$

Proof: we focus on $N(\mu, \sigma_0^2)$ where $\sigma_0^2$ is any value.

we can find that $\bar{X}_n$ is sufficient and complete

By Basu's theorem, $\bar{X}_n \perp\!\!\!\perp$ any ancillary statistic

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \text{ is ancillary}$$

$$\Rightarrow \bar{X}_n \perp\!\!\!\perp S_n^2 \text{ for } N(\mu, \sigma_0^2)$$

Because we pick $\sigma_0^2$ arbitrarily!!!

$\bar{X}_n \perp\!\!\!\perp S_n^2$ should hold for all $\sigma_0^2 > 0$

Hence $\bar{X}_n \perp\!\!\!\perp S_n^2$ holds for any $N(\mu, \sigma^2)$ !

Theorem 6.2.28. Suppose $T(\underset{\sim}{X})$ is sufficient.

If $T(\underset{\sim}{X})$ is complete

Then $T(\underset{\sim}{X})$ is minimal sufficient !

(the other way does not hold : minimal sufficiency $\Rightarrow$ completeness)