STAT 543 :)

Lec 16, F, Feb 28

No assignments until M, Mr 3

Sufficiency and Point Estimation

Minimal Sufficiency

Question: Suppose $\underset{\sim}{S} \equiv (S_1, \dots, S_k)$ is sufficient for $\underset{\sim}{\theta}$ and S_0 is another arbitrary statistic. Is $\underset{\sim}{S}^* \equiv (S_0, S_1, \dots, S_k)$ is sufficient for $\underset{\sim}{\theta}$? Yes!

proof: Since $\underset{\sim}{S}$ is sufficient,

$f(\underset{\sim}{x}|\underset{\sim}{\theta}) = g(\underset{\sim}{S}, \underset{\sim}{\theta})h(\underset{\sim}{x})$ holds by Factorization Theorem

But $\underset{\sim}{S} = (S_1, \dots, S_k) = d(\underset{\sim}{S}^*)$ is a function of $\underset{\sim}{S}^*$

So, $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = g(\underset{\sim}{S}, \underset{\sim}{\theta})h(\underset{\sim}{x})$

$\qquad = g(d(\underset{\sim}{S}^*), \underset{\sim}{\theta})h(\underset{\sim}{x}) = g_1(\underset{\sim}{S}^*, \underset{\sim}{\theta})h(\underset{\sim}{x})$

$\therefore$ $\underset{\sim}{S}^*$ is sufficient Fact. Theorem.

Definition: A vector of statistics $\underset{\sim}{S}$ is called **minimally sufficient** if

1. $\underset{\sim}{S}$ is sufficient for $\underset{\sim}{\theta}$, and
2. for any other vector $\underset{\sim}{T}$ of sufficient statistics for $\underset{\sim}{\theta}$, $\underset{\sim}{S}$ is a function of $\underset{\sim}{T}$.

(Later: we can check "minimally sufficient" using "completeness with sufficiency.")

$X_1, \dots, X_n \xrightarrow{\text{data reduction}}$ $\underset{\sim}{T}$ is sufficient for $\underset{\sim}{\theta}$ (more)

$\Big\downarrow$ data reduction $g(\underset{\sim}{T}) = \underset{\sim}{S}$ ($\underset{\sim}{S}$ is a function of $\underset{\sim}{T}$)

minimally sufficient $\underset{\sim}{S}$ (less)

e.g. In last MVN example, $A = I_{n\times n}$, it turns out that $\underset{\sim}{S} = (\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i)$ is minimally sufficient for (μ, σ^2).

Later: We can show "minimal sufficiency" using "completeness"

Sufficiency and Point Estimation

Remarks on Sufficiency

1. If $X_1, \ldots, X_n$ is a random sample (iid) from pdf/pmf $f(x|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta$, then the order statistics $X_{(1)}, \ldots, X_{(n)}$ are sufficient for $\underset{\sim}{\theta}$.

   proof: By the factorization theorem, $X_{(1)}, \ldots, X_{(n)}$ are sufficient for $\underset{\sim}{\theta}$ because we can write

$$\begin{aligned} \text{the joint pdf/pmf } f(\underset{\sim}{x}|\underset{\sim}{\theta}) &= \prod_{i=1}^n f(x_i|\underset{\sim}{\theta}) = \prod_{i=1}^n f(x_{(i)}|\underset{\sim}{\theta}) \\ &= \underbrace{g(x_{(1)}, \ldots, x_{(n)}, \underset{\sim}{\theta})}_{\prod_{i=1}^n f(x_{(i)}|\underset{\sim}{\theta})} \underbrace{h(\underset{\sim}{x})}_{1}, \quad \text{for all } \underset{\sim}{x}, \underset{\sim}{\theta} \end{aligned}$$

2. If $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ is sufficient for real-valued $\theta \in \Theta \subset \mathbb{R}$, then any Bayes estimator is a function of $\underset{\sim}{S}$.

   $f_{\theta|\underset{\sim}{x}}(\theta) \propto f(\underset{\sim}{x}|\theta)\pi(\theta)$
   $\propto g(\underset{\sim}{s}, \theta)\pi(\theta)$

   Example: From homework, consider $X_1, \ldots, X_n$ iid Bernoulli(θ), $0 < \theta < 1$; loss $L(t, \theta) = \frac{(t-\theta)^2}{\theta(1-\theta)}$; and uniform(0,1) prior $\pi(\theta)$.

   Then the Bayes estimator is $T_0 = \bar{X}_n$, which is sufficient for θ (by factorization theorem).

3. If $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ is sufficient for $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ and $\hat{\theta}$ is the unique MLE of $\underset{\sim}{\theta}$, then $\hat{\theta}$ is a function of $\underset{\sim}{S}$.

   $f(\underset{\sim}{x}|\theta) = g(\underset{\sim}{s}, \theta)h(\underset{\sim}{x})$ by Fact theorem

Sufficiency and Point Estimation

Rao-Blackwell Theorem & Sufficiency

Rao-Blackwell Theorem. Let $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = f(x_1, \ldots, x_n|\underset{\sim}{\theta})$ be the joint pdf/pmf of $(X_1, \ldots, X_n)$ and $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ be sufficient for $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p) \in \Theta \subset \mathbb{R}^p$. Also let T be any UE of a real-valued $\gamma(\underset{\sim}{\theta})$ and $T^* = \mathrm{E}(T|\underset{\sim}{S})$ *(this conditional expectation does not depend on $\underset{\sim}{\theta}$, since $\underset{\sim}{S}$ is sufficient, and so is a statistic).*

new estimator (just depend on $\underset{\sim}{S}$ through data & NOT on $\underset{\sim}{\theta}$)

Then,

1. T^* is a function of $\underset{\sim}{S}$ and an UE of $\gamma(\underset{\sim}{\theta})$.

2. $\mathrm{Var}_{\underset{\sim}{\theta}}(T^*) \leq \mathrm{Var}_{\underset{\sim}{\theta}}(T)$, for all $\underset{\sim}{\theta} \in \Theta$.

3. If $\mathrm{Var}_{\underset{\sim}{\theta}_0}(T^*) = \mathrm{Var}_{\underset{\sim}{\theta}_0}(T)$ holds for some $\underset{\sim}{\theta}_0 \in \Theta$, then $P_{\underset{\sim}{\theta}_0}(T = T^*) = 1$.

Idea: T is UE of $\gamma(\underset{\sim}{\theta})$ $\xrightarrow{\text{condition on sufficient } \underset{\sim}{S}}$ new $T^* = E(T|\underset{\sim}{S})$

"Rao-Blackwellization" or we say we "Rao-Blackwellize" T using $\underset{\sim}{S}$

Remarks

- Given an UE T of $\gamma(\underset{\sim}{\theta})$, the theorem shows how to obtain an UE T^* that is at least as good as T in terms of variance (in fact, better than T unless $T = T^*$ with probability 1 for all $\underset{\sim}{\theta}$). That is, you can "Rao-Blackwellize" an UE T by conditioning on a sufficient statistic $\underset{\sim}{S}$.

$T^* = E(T|\underset{\sim}{S})$

(ASIDE: $T = g(\underset{\sim}{S}) \Rightarrow T^* = E(T|\underset{\sim}{S}) = E(g(\underset{\sim}{S})|\underset{\sim}{S}) = g(\underset{\sim}{S})$)

- For finding an UMVUE of $\gamma(\underset{\sim}{\theta})$ we may restrict attention to the class of estimators that are functions of a sufficient statistic.

$T^* = E(T|\underset{\sim}{S})$ function of $\underset{\sim}{S}$

Sufficiency and Point Estimation

Rao-Blackwell Theorem: Illustration

$n=2$

Example: Suppose X_1, X_2 are iid Exponential(θ). Note $T = X_1$ is an UE of θ and $\text{Var}_\theta(T) = \text{Var}_\theta(X_1) = \theta^2$.

$E_\theta T = E_\theta X_1 = \theta$

Also note that $S = X_1 + X_2$ is sufficient for θ by factorization theorem & S is GAMMA(2,θ)-distributed.

Verify that

1. $T^* = \text{E}_\theta(T|S) = \text{E}_\theta(X_1|S)$ is a function of S;
2. T^* doesn't depend on θ;
3. T^* is unbiased for θ;
4. and compare $\text{Var}_\theta(T)$ and $\text{Var}_\theta(T^*)$

Solution: Given $S = s > 0$, first find the conditional pdf of $X_1|S = s$ as

$$f(x_1|S=s) = \frac{f_{X_1,S}(x_1, s|\theta)}{f_S(s|\theta)} = \frac{f_{X_1,X_2}(x_1, x_2 = s - x_1|\theta)}{f_S(s|\theta)}$$

joint pdf of (X_1, S)

$$= \begin{cases} \dfrac{\theta^{-2}e^{-x_1/\theta}e^{-(s-x_1)/\theta}}{\theta^{-2}se^{-s/\theta}} = s^{-1} & \text{if } 0 < x_1 < s \\ 0 & \text{otherwise} \end{cases}$$

Gamma pdf

So, given $S = s > 0$, the conditional distribution of X_1 is UNIF$(0, S)$

Hence, the conditional expectation is $\text{E}_\theta(X_1|S = s) = \frac{0+S}{2} = S/2$

Now, treating S as a random variable, we have $T^* = \text{E}_\theta(X_1|S) = \frac{S}{2} = \frac{X_1+X_2}{2} = \bar{X}_2$

1. T^* is function of S
2. T^* doesn't depend on θ
3. T^* is UE of θ $(E_\theta(\bar{X}_2) = \theta)$
4. $Var_\theta(\bar{X}_2) = \theta^2/2 < Var_\theta(X_1) = \theta^2$

Sufficiency and Point Estimation

Completeness

side note : to be used with sufficiency

- "Completeness" is a statistical property for use in conjunction with sufficiency.

- "Completeness" has a rather technical definition: a statistic $\underset{\sim}{T}$ is complete if the only function $u(\underset{\sim}{T})$ of $\underset{\sim}{T}$ that can be an UE of zero is $u(\underset{\sim}{T}) = 0$ wp1

Note: $u(\underset{\sim}{T}) \equiv 0$ (always zero) has expectation zero (i.e. $E_\theta u(\underset{\sim}{T}) = 0$).

It is true that $u(T) \equiv 0$ is the ONLY function of $\underset{\sim}{T}$ that is U.E of zero? If so, $\underset{\sim}{T}$ is complete

"Completeness" Definition: Let $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = f(x_1, \ldots, x_n|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$, be the *joint* pdf/pmf of $(X_1, \ldots, X_n)$ and let $f_{\underset{\sim}{T}}(\underset{\sim}{t}|\underset{\sim}{\theta})$ denote the pdf/pmf of a vector of statistics $\underset{\sim}{T}$. Then,

sampling distribution of $\underset{\sim}{T}$

(i) $\underset{\sim}{T}$ (and/or the family $\mathcal{F}_{\underset{\sim}{T}} \equiv \{f_{\underset{\sim}{T}}(\underset{\sim}{t}|\underset{\sim}{\theta}) : \underset{\sim}{\theta} \in \Theta\}$) is called **complete** if, for any real-valued function $u(\underset{\sim}{T})$,

$$\text{whenever } \underbrace{\mathrm{E}_{\underset{\sim}{\theta}}\, u(\underset{\sim}{T}) = 0 \text{ holds for all } \underset{\sim}{\theta} \in \Theta}_{\text{says "}u(\underset{\sim}{T})\text{ is U.E of zero"}} \text{ then } P_{\underset{\sim}{\theta}}\Big[u(\underset{\sim}{T}) = 0\Big] = 1 \text{ for } \underset{\sim}{\theta} \in \Theta \quad (3)$$

then, $u(\underset{\sim}{T})$ must zero w.p. 1

(ii) $\underset{\sim}{T}$ is called **bounded complete** if (4) holds for all *bounded* functions $u(\cdot)$.

completeness of $\underset{\sim}{T}$ depends on $f_{\underset{\sim}{T}}(t|\theta)$ & "richness/size of parameter space Θ"

The more parameter $\underset{\sim}{\theta}$ in Θ, the more constraints are on $u(\underset{\sim}{T})$ in $E_\theta(u(\underset{\sim}{T}))=0$, $\forall \underset{\sim}{\theta} \in \Theta$

e.g. $T \equiv X_1 \sim N(\theta, 1)$, $\theta \in \Theta$

① $\Theta = \{1\}$

Note: $u(X_1) = X_1 - 1$

$E_\theta u(X_1) = E_\theta (X_1 - 1) = 0$, $\forall \theta \in \Theta = \{1\}$

But, $u(X_1)$ is NOT zero wp.1 $\leftarrow N(0,1) \sim X_1 - 1$

($P_\theta(X_1 - 1 = 0) = P(Z = 0) = 0$)

So, $T \equiv X_1$ is NOT complete.

② $\Theta \equiv \mathbb{R}$. Then, $T \equiv X_1$ is complete (later)

Example. Suppose $X_1, \ldots, X_n$ are iid Poisson(θ), $\theta > 0$. Show that $T \equiv \sum_{i=1}^n X_i$ is complete.

Solution: $T \sim \text{Poisson}(n\theta)$, $\theta > 0$.

Now for some $u(\cdot)$, suppose it holds that $E_\theta u(T) = 0$, $\forall \theta > 0$.

$$\Leftrightarrow \sum_{t=0}^{\infty} u(t) e^{-n\theta} \frac{(n\theta)^t}{t!} = 0, \quad \forall \theta > 0$$

$$\Leftrightarrow \sum_{t=0}^{\infty} u(t) \frac{(n\theta)^t}{t!} = 0, \quad \forall \theta > 0.$$

$$\Leftrightarrow \left[u(0) + u(1)(n\theta) + u(2)\frac{(n\theta)^2}{2!} + \cdots \right] = 0, \forall \theta > 0$$

let $\theta \downarrow 0$, get $u(0) = 0$.

$$\Leftrightarrow \left[u(1) + u(2)\frac{(n\theta)}{2!} + \cdots \right] = 0, \forall \theta > 0$$

let $\theta \downarrow 0$, get $u(1) = 0$.

So, get for any $t \geq 0$, $u(t) = 0$

$\Rightarrow P_\theta(u(T) = 0) = 1$ for all $\theta > 0$

$\therefore$ T is complete!

(Later, we'll see easier ways to check completeness)

Sufficiency and Point Estimation

Remarks on Completeness

1. If T is complete, then T is boundedly complete; the converse is false.

connection between sufficiency & completeness

2. If T is sufficient and boundedly complete, then T is minimal sufficient.

   So by Remark 1 above, if T is sufficient and complete, then T is minimal sufficient.

3. Suppose T is complete and $h_1(T), h_2(T)$ are two estimators of $\gamma(\underset{\sim}{\theta})$

$$\begin{aligned}
&\text{if } \mathrm{E}_{\underset{\sim}{\theta}} h_1(T) = \gamma(\underset{\sim}{\theta}) = \mathrm{E}_{\underset{\sim}{\theta}} h_2(T), \ \text{ for all} \underset{\sim}{\theta} \in \Theta \\
\Rightarrow \ & \mathrm{E}_{\underset{\sim}{\theta}} u(T) = 0, \ \text{ for all} \underset{\sim}{\theta} \in \Theta, \quad \text{where } u(T) = h_1(T) - h_2(T) \\
\Rightarrow \ & P_{\underset{\sim}{\theta}}\big(u(T) = 0\big) = 1, \ \text{ for all} \underset{\sim}{\theta} \in \Theta \\
\Rightarrow \ & P_{\underset{\sim}{\theta}}\big(h_1(T) = h_2(T)\big) = 1, \text{for all} \underset{\sim}{\theta} \in \Theta
\end{aligned}$$

   Hence, there can be at most one (i.e., unique) UE of a parametric function $\gamma(\underset{\sim}{\theta})$ that is a function of a complete statistic.

4. Let $T \equiv h(X_1, \ldots, X_n)$ be an UE of $\gamma(\theta)$ & suppose $\underset{\sim}{S}$ is sufficient.

Recall:

$T \xrightarrow[\text{Rao-Blackwellize}]{\text{sufficient } \underset{\sim}{S}} T^* = E(T|\underset{\sim}{S})$ is U.E of $\gamma(\theta)$ & $Var_\theta(T^*) \le Var_\theta(T)$

$T \xrightarrow{\text{sufficient \& complete } \underset{\sim}{S}} T^* = E(T|\underset{\sim}{S})$ is UMVUE of $\gamma(\theta)$!