Ⅱ Note: (from 6-1)

Sufficiency Principle

Sufficiency Principle (6-2)

Def: $\underline{X} \sim P(\underline{x}, \theta)$, $T(\underline{X})$ is a statistic of $\underline{X}$.

We say $T(\underline{X})$ is <u>sufficient</u> for θ, if $P(\underline{x} \mid T(\underline{x}))$ does not depend on θ.

Thm: (Factorization theorem)

Given a distribution family $P(\underline{x}, \theta)$, $\theta \in \Theta$, $T(\underline{X})$ is a statistic of $\underline{X}$, then

$$T(\underline{X}) \text{ is sufficient for } \theta \iff P(\underline{x}, \theta) = h(\underline{x})\, g(T(\underline{x}), \theta), \quad \forall \theta, \forall \underline{x}.$$

Note:
1° Parameters θ influence the pr. distri. of a random sample $\underline{X}$; but it only manipulates the pr. distr. on a coarse-grained space of r.s., numbered by a sufficient stat. $T(\underline{X})$.
2° $T(\underline{X}) = t$ represents a subspace of the space of r.s., $\mathbb{R}^n$, within which the relative pr. distr. is not influenced by θ. The subspace can also be named as a coarse grain.

Proof: "$\Rightarrow$"

$$\begin{aligned} P(\underline{x}, \theta) &= P_\theta(\underline{x}) \\ &= P_\theta(\underline{x}, T(\underline{X}) = T(\underline{x})) \\ &= P_\theta(\underline{x} \mid T(\underline{X}) = T(\underline{x})) \cdot P_\theta(T(\underline{X}) = T(\underline{x})) \\ &= P(\underline{x} \mid T(\underline{X}) = T(\underline{x})) \cdot P_\theta(T(\underline{X}) = T(\underline{x})) \\ &= h(\underline{x}) \cdot g(T(\underline{x}), \theta) \end{aligned}$$

"$\Leftarrow$" let $t = T(\underline{x})$, then

$$\begin{aligned} P_\theta(\underline{x} \mid T(\underline{X}) = t) &= \frac{P_\theta(\underline{x}, T(\underline{X}) = t)}{P_\theta(T(\underline{X}) = t)} \\ &= \frac{P_\theta(\underline{x})}{P_\theta(T(\underline{X}) = t)} \\ &= \frac{h(\underline{x})\, g(t, \theta)}{\sum_{\{\underline{y} : T(\underline{y}) = t\}} h(\underline{y})\, g(t, \theta)} \\ &= \frac{h(\underline{x})}{\sum_{\{\underline{y} : T(\underline{y}) = t\}} h(\underline{y})} \quad \text{doesn't depend on } \theta. \end{aligned}$$

Note:
$f_{T(\underline{X})}(t) \propto g(T(\underline{x}), \theta)$
The pdf/pmf of the sufficient stat. $T(\underline{X})$ is proportional to $g(T(\underline{x}), \theta)$, except for a factor $c(T(\underline{x}))$.
Or, we can choose $g(T(\underline{x}), \theta)$ to be the pdf/pmf of $T(\underline{X})$!

$\square$.

Thm: If $T(\underline{X})$ is sufficient for θ, and $T(\underline{X}) = f(S(\underline{X}))$, then $S(\underline{X})$ is also sufficient for θ.

Note: 1° $T(\underline{X}) = f(S(\underline{X}))$ is a renumbering of grains (subspaces), which either maintains the structure of the coarse grained space (of r.s.), or further coarsens the space.

2° If the parameter θ cannot manipulate the relative pr distr. with a coarse grain, it also cannot manipulate within a finer grain!

Page

E.g.: 1° Exponential family

$$P(\underline{x},\theta) = h(\underline{x})\, c(\theta) \exp\left\{ \sum_{i=1}^{k} w_i(\theta)\, t_i(\underline{x}) \right\}$$

Statistic $T(\underline{X}) = (t_1(\underline{X}), \cdots, t_k(\underline{X}))$ is sufficient for θ.

2° $X_1, \cdots, X_n$ iid., and are distributed as an exponential family,

then $P(\underline{x},\theta) = \left[\prod_{j=1}^{n} h(x_j)\right] \cdot c^n(\theta) \cdot \exp\left\{ \sum_{i=1}^{k} w_i(\theta) \sum_{j=1}^{n} t_i(x_j) \right\}$.

Statistic $T(\underline{X}) = \left(\sum_{j=1}^{n} t_1(x_j), \cdots, \sum_{j=1}^{n} t_k(x_j)\right)$ is sufficient for θ.

3° $X_1, \cdots, X_n$ iid $U[0,\theta]$,

Statistic $X_{(n)}$ is sufficient for θ,

$$f(\underline{x}|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} 1(0<x_i<\theta) = \frac{1}{\theta^n} 1(x_1,\cdots,x_n>0,\ x_{(n)}<\theta) = \frac{1}{\theta^n} 1(x_1,\cdots,x_n>0)\, 1(x_{(n)}<\theta)$$

4° Sufficient order statistics - (P275)

5° Minimal sufficient stat.

Note: P278

Def: $T(\underline{X})$ is a sufficient statistic of $\underline{X}$, we say $T(\underline{X})$ is a <u>minimal sufficient statistic</u>, if $\forall$ sufficient statistic $S(\underline{X})$, $\exists\, g(\cdot)$, s.t.

$$T(\underline{X}) = g(S(\underline{X}))$$

Intermediate page → Note: P280, P284(1)

Def: A statistic $T(\underline{X})$ is <u>complete</u> if

$$E_\theta\, g(T(\underline{X})) = 0 \quad \forall \theta \in \Theta \Rightarrow g(t) = 0 \text{ on the support of } T(\underline{X}),\ \forall \theta \in \Theta.$$

Note: 1° S_n is a complete sufficient statistic for Bernoulli random sample. ($S_n \equiv \sum_{i=1}^{n} X_i$)

2° S_n is a complete sufficient statistic for Poisson random sample.

3° $X_{(n)}$ is a complete sufficient statistic for Uniform$(0,\theta)$ random sample.

4° S_n is not a complete sufficient statistic for $N(\theta,\theta^2)$ random sample.

Thm: $\underline{X} \sim f(\underline{x}|\theta)$, If $\exists$ statistic $T(\underline{X})$, s.t. $\forall \underline{x}, \underline{y}$,

$$\frac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)} \text{ doesn't depend on } \theta \iff T(\underline{x}) = T(\underline{y}),$$

Question: P281

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

(Note: for proof, see textbook P281).

Eg: 1° $X \sim N(\mu, \sigma^2)$, then $(\bar{X}, S^2)$ is a minimal sufficient statistic for (μ, σ^2).

2° $X \sim U(\theta, \theta+1)$, then $(X_{(1)}, X_{(n)})$ is a minimal sufficient statistic for θ. The dimensions the minimal sufficient statistic and the parameter do not match.

Note: Any 1-1 function of a minimal sufficient statistic is also a minimal sufficient statistic.

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

E.g.: 1° $X \sim U(\theta, \theta+1)$, then the range statistic $R = X_{(n)} - X_{(1)}$ is ancillary. And $R \sim Beta(n-1, 2)$

2° (location family) $X \sim F(x-\theta)$, where $F(x)$ is a cdf. Then the range $R = X_{(n)} - X_{(1)}$ is ancillary.

3° (Scale family) $X \sim F(\frac{x}{\sigma})$, $\sigma > 0$. Then $\forall$ statistic T, s.t. $T = T(\frac{X_1}{X_n}, \dots, \frac{X_{n-1}}{X_n})$, is ancillary.

($\because$ the joint cdf of $(\frac{X_1}{X_n}, \dots, \frac{X_{n-1}}{X_n})$ doesn't depend on σ.)

Note: P284(2)

Thm: (Basu's theorem)

If $T(\underline{X})$ is a complete and (minimal) sufficient statistic, then $T(\underline{X})$ is independent of every ancillary statistic.

Note: P289

Thm: ~~$X_1, \ldots, X_n$ iid f~~

(Complete statistics in exponential family)

$X_1, \cdots, X_n$ iid $f(x|\underline{\theta}) = h(x)c(\underline{\theta})\exp\{\sum_{j=1}^{k} w_j(\underline{\theta})t_j(x)\}$

with $\underline{\theta} = (\theta_1, \cdots, \theta_k)$.

Then statistic $T(\underline{X}) = (\sum_{i=1}^{n} t_1(X_i), \cdots, \sum_{i=1}^{n} t_k(X_i))$

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

(natural parameter space)

Thm: (Bahadur's thm) If a minimal sufficient statistic exists, then any complete sufficient statistic is also a minimal sufficient statistic.