Tutorial for 11-03 and 11-06

1. Sufficient statistic

- Definition

Definition 2 (One-parameter cases)

Let $\boldsymbol{X} = \{X_i : i = 1, \dots, n\}$ be a r.s. from pdf $f(\cdot\,|\theta)$ or pmf $p(\cdot\,|\theta)$, where $\theta \in \Theta \subset R$. A statistic $S = S(\boldsymbol{X})$ is said to be ***sufficient*** if and only if the conditional distribution of $\boldsymbol{X}$ given $S = s$ does not depend on θ, for all values s of S.

$f(\underline{X} \mid S(X) = s)$

Definition 3 (Multi-parameter cases)

Let $\boldsymbol{X} = \{X_i : i = 1, \dots, n\}$ be a r.s. from pdf $f(\cdot\,|\theta)$ or pmf $p(\cdot\,|\theta)$, where $\theta \in \Theta \subset R^k$. A vector of statistics $S_1 = S_1(\boldsymbol{X}), S_2 = S_2(\boldsymbol{X}), \dots, S_r = S_r(\boldsymbol{X})$ is said to be ***jointly sufficient***, where $r \geq k$, if and only if the conditional distribution of $\boldsymbol{X}$ given $S_1 = s_1, S_2 = s_2, \dots, S_r = s_r$ does not depend on θ, for all values s_1 of S_1, s_2 of S_2, ..., and s_r of S_r.

$f(\underline{X} \mid S_1(X) = S_1, S_2(X) = S_2, \dots)$

the sample are not needed ∵ they can tell us nothing more about θ

- Relation with fisher Information

Lemma 6. Under the regularity conditions, $T(\boldsymbol{X})$ is a sufficient statistic for θ if and only if

$$I_{T(\boldsymbol{X})}(\theta) = I_{\boldsymbol{X}}(\theta).$$

Remark that this lemma tells us that Fisher information can show the fact that a sufficient statistic carries as much information about as the sample.

2. Factorization Theorem

the following theorem to give us an easy way of obtaining the sufficient statistic.

Theorem 3 (Fisher-Neyman Factorization Theorem):

Ronald Fisher (1890 – 1962)

Let $\boldsymbol{X} = \{X_i : i = 1, \dots, n\}$ be a r.s. from pdf $f(\cdot\,|\theta)$ or pmf $p(\cdot\,|\theta)$, where $\theta \in \Theta \subset R^k$. A set of statistics $S_1(\boldsymbol{X}), S_2(\boldsymbol{X}), \dots, S_r(\boldsymbol{X})$ is said to be ***jointly sufficient***, where $r \geq k$, if and only if the joint pdf $f_{X_1,\dots,X_n}(x_1, \dots, x_n|\theta)$ or the joint pmf $p_{X_1,\dots,X_n}(x_1, \dots, x_n|\theta)$ of the r.s. can be factorized in form of

$$g(S_1(x_1, \dots, x_n), \dots, S_r(x_1, \dots, x_n)|\theta)\; h(x_1, \dots, x_n),$$

where $g(S_1(x_1, \dots, x_n), \dots, S_r(x_1, \dots, x_n)|\theta)$ is a non-negative function of $x_1, \dots, x_n$ ONLY through the functions $S_1, \dots, S_r$ and depends on θ, and $h(\cdot)$ is a non-negative function of $x_1, \dots, x_n$ alone, i.e. it does not involve θ.

Jerzy Neyman (1894 – 1981)

The above theorem indeed tells us that all information about θ contained in the rs is ***completely transferred into*** the set of statistics $S_1, \dots, S_r$. ← Take joint pdf as an example

when $n=1$: $\mathcal{X} = \{X_1\} = X_1$

$S(x)$ is sufficient for θ iff $f_\theta(x) = g(S(x), \theta) \cdot h(x)$

$$f_\theta(x) = \prod_{i=1}^{n} f_\theta(x_i)$$

- relation with MLE

Corollary 2: Let $\boldsymbol{X} = \{X_i : i = 1, \dots, n\}$ be a r.s. from a distribution with a pdf $f(\cdot\,|\theta)$ or a pmf $p(\cdot\,|\theta)$, where $\theta \in \Theta \subset R$. If ① a sufficient statistic $S(\boldsymbol{X})$ for θ exists and if ② a MLE $\hat{\theta}(\boldsymbol{X})$ for θ also exists **uniquely**, then $\hat{\theta}(\boldsymbol{X})$ is a function of $S(\boldsymbol{X})$.

- Sufficiency is closed (doesn't change after transformation)

Theorem 4 (1-1 sufficiency): Let $\boldsymbol{X} = \{X_i : i = 1, \dots, n\}$ be a r.s. of size n. If a set of statistics $S_1(\boldsymbol{X}), S_2(\boldsymbol{X}), \dots, S_r(\boldsymbol{X})$ is *jointly sufficient*, where $r \geq k$, then any set of one-to-one function (or transformation) of $S_1(\boldsymbol{X}), S_2(\boldsymbol{X}), \dots, S_r(\boldsymbol{X})$ is also *jointly sufficient*.

f

$\underline{f_1}[S_1(X)], \underline{f_2}[S_2(X)], \dots, \underline{f_r}[S_r(X)]$ ($f_1, f_2, \dots, f_r$ can be different)

e.g. $\sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} X_i^2$ are jointly sufficient ---- 1

$\downarrow f(x) = \frac{x}{n}$ $\qquad \downarrow$ $\qquad$ 1 to 1

$\bar{x}$ and $\sum_{i=1}^{n} (X_i - \bar{x})^2$ are also .... -- 1.

if $f(x) = \left(\frac{x}{n}\right)^2$, the conclusion may not valid.

3. Minimal Sufficient statistic

- Definition:

Given any other sufficient statistic $S(x)$, $T(x)$ is a function of $S(x)$

$\Rightarrow$ $T(x)$ is minimal sufficient $[\ T(x) = L(S(x))\ ]$

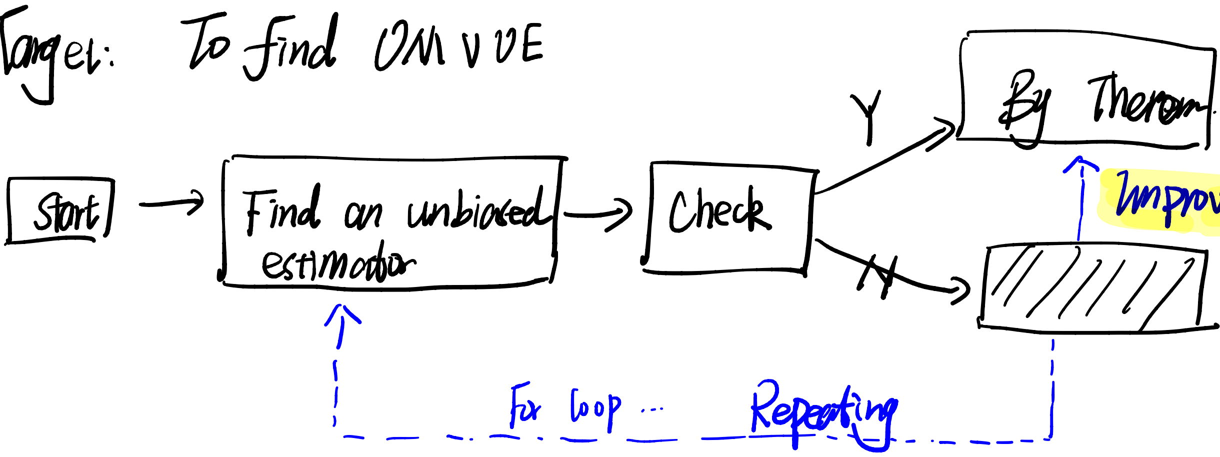
In general, if $T(x) = L[S(x)]$

$T(x)$ is (minimal) sufficient. $S(x)$ is also (minimal) sufficient

4.

- Why we need sufficient statistic ???

Target: To find UMVUE



Rao and Black shows that

The UMVUE must be a function of jointly sufficient statistic

Theorem 5 (Rao-Blackwell Theorem):

Let $\boldsymbol{X} = \{X_i : i = 1, \dots, n\}$ be a r.s. from pdf $f(\cdot\,|\theta)$ or pmf $p(\cdot\,|\theta)$, where $\theta \in \Theta \subset R^k$, and a set of statistics $S_1(\boldsymbol{X}), S_2(\boldsymbol{X}), \dots, S_r(\boldsymbol{X})$ be *jointly sufficient*, where $r \geq k$. Suppose that a statistic $T = T(\boldsymbol{X})$ is an unbiased estimator for $g(\theta)$. Define T' by $E(T|S_1, \dots, S_r)$. Then,

1. T' is a statistic, and it is a function of the jointly sufficient statistics.
2. T' is also unbiased for $g(\theta)$.
3. $Var(T') \leq Var(T)$, for all $\theta \in \Theta$.

Example:

let $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$, find a UMVUE for

$$g(\lambda) = P(X_1 = k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for any } k = 0, 1, 2, \ldots$$

Solution:

$$S(X) = I\{X_1 = k\}$$

$$E\,S(X) = g(\lambda)$$ - - - - - unbiased

Then, we can find a complete and sufficient statistic (C-S)

will learn later.

$$T(X) = \sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$$

$$T' = E[\,S(X) \mid T(X) = t\,]$$

① unbiased estimate

② C-S

③ add Rao-Black Them

$\Rightarrow$ UMVUE

$$= P(X_1 = k \mid T = t)$$

$$= \frac{P(X_1 = k,\ T = t)}{P(T = t)}$$

$$= \frac{P(X_1 = k,\ \sum_{i=2}^{n} X_i = t - k)}{P(\sum_{i=1}^{n} X_i = t)}$$

$$= \frac{\frac{e^{-\lambda}\lambda^k}{k!} \cdot \frac{e^{-(n-1)\lambda}[(n-1)\lambda]^{t-k}}{(t-k)!}}{e^{-n\lambda}[n\lambda]^t / t!}$$

$$\Rightarrow \binom{t}{k}\left(\frac{1}{n}\right)^k\left(1 - \frac{1}{n}\right)^{t-k}.$$