Sufficiency

Equivalent procedure based on a sufficient statistic

Used framework - loss function
X observed
T(θ) a sufficient statistic of θ

Let S be any decision procedure.

Note that conditionally given T(θ) = x, S(x)
some distribution not depending on θ.

Draw a sample S*(A) from the conditional distribution of S(A) given T(θ) = x
(Note: S* is a randomized procedure)

E[θ(θ, S*(θ)) | T(θ) = x] = E[θ[θ(θ, S(θ)) | T(θ) = x]]

By double expectation theorem (E[E[Y | Z]] = E[E])
we see that

E[θ[θ(θ, S*(θ)) | T(θ) = x]] = E[θ[θ(θ, S(θ)) | T(θ) = x]]

E[θ(θ, S*(θ)) | T(θ) = x] = E[θ(θ, S(θ))]

All θ

Therefore exists a randomized procedure whose performance
is same as that of S & its distribution
depends only on the sufficient statistic T(θ).
Convex Loss Function & Improvement by Conditioning

General decision theory framework as above, but take \( \mathbf{A} \in \mathbb{R}^k \) bounded, convex and assume

\[
l(\theta, a) : \mathbf{A} \rightarrow [0, 100]
\]

is a convex function for all \( \theta \in \Theta \).

Example:

\[
l(\theta, a) = ||\theta - a||^2
\]

Let \( T(X) \) be a sufficient statistic of \( \theta \) and let \( \delta(X) \) be any decision procedure.

Define

\[
\delta^*(X) = \mathbb{E}_\theta[\delta(X) \mid T(X)]
\]

This defines a decision procedure. Note: \( \delta^*(X) \) is bounded so expectation exists.

Define

\[
\delta^*(X) = \mathbb{E}_\theta[\delta(X) \mid T(X)]
\]

This defines a decision procedure. Note: \( \delta^*(X) \) is bounded so expectation exists.

Jensen's inequality:

\[
E[\delta(Y)] \geq \Psi(EY)
\]

is \( \psi \) convex

\[
E[\psi(Y)] \geq \psi(EY)
\]

\( R(\theta, \delta^*(X)) \geq R(\theta, \delta(X)) \) all \( \theta \)

\( \delta^*(X) \) does no worse than \( \delta \)

depends only on \( T \)
Example: $X_1, \ldots, X_n$ i.i.d. Uniform $(0, \theta)$

pdf of $X_i$:
\[ f(x; \theta) = \frac{1}{\theta} I_{(0, \theta)}(x) \]

cdf of $X_i$:
\[ F(x; \theta) = \frac{x}{\theta} I_{(0, \theta)}(x) \]

Factorization Theorem:
\[ p(x_1, \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \frac{1}{\theta} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
\[ = \frac{1}{\theta^n} I_{(0, \theta)}(\max x_i) \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
\[ g(\theta; \max x_i) \underbrace{g(x_1, \theta) \ldots g(x_n, \theta)}_{n \times n} \]

$T_n = \max x_i$ is a sufficient statistic for $\theta$

What is the conditional distribution of $(X_1, \ldots, X_n)$ given $\max x_i$?

Exercise B.2.10

$X_1, \ldots, X_n \mid X_n$ has pdf of the form
\[ f(x_1, \ldots, x_n; \theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
\[ = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
\[ \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
\[ \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
\[ \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]

\[ (x_{(n)} \mid \theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0, \theta)}(x_i) \]
Take as estimator of $\theta$ using squared error loss $2\widehat{X}$

\[ E(2\widehat{X}) = 2 \frac{\theta}{\sigma^2} = \theta \text{ unbiased} \]

\[ \text{Var}(2\widehat{X}) = \frac{4}{n} \text{Var} X_i = \frac{4\theta^2}{12} = \frac{\theta^2}{3} \]

Above theorem says we improve on $\bar{X}$ by taking

\[ E[2\widehat{X} | X(n)] = \frac{2}{n} \sum_{i=1}^{n} E[X(i) | X(n)] \]

\[ = \frac{2}{n} \sum_{i=1}^{n} E[X(i)] \]

(Continued)

If $X_1, \ldots, X_n$ are iid Uniform (0, $\theta$) then conditionally given $X(n)$, $X_{(1)}, \ldots, X_{(n-1)}$ are distributed as the order statistics for any iid sample $X'_1, \ldots, X'_n$.

\[ E[X_{(i)} | X(n)] = \frac{X_{(i)} - E[X_{(i)} | X(n)]}{X_{(i)}} \]

\[ \Rightarrow \quad E[X_{(i)} | X(n)] = \frac{X_{(i)}}{X(n)} \]

For $X_{(i)} = \frac{X(n)}{X(n) - X_{(i-1)}}$

so

\[ X_{(i)} \sim \text{Beta}(i, n-i) \]

\[ 2 \ E[X | X(n)] = \frac{2}{n} \sum_{i=1}^{n} X_{(i)} \frac{i}{n} = \frac{2X(n)}{n} \frac{n(n+1)}{2} \]

\[ E[\frac{n+1}{n}X(n)] \]

\[ \hat{\theta} = \frac{n+1}{n} X(n) \]

\[ E_{\hat{\theta}} = \frac{n+1}{n} E[X(n)] = \frac{n+1}{n} \theta = \theta \]

\[ \text{Var} \hat{\theta} = (\frac{n+1}{n})^2 \text{Var} X(n) = (\frac{n+1}{n})^2 \text{Var}(\theta X(n)) \]
\[
\left( \frac{\sigma^2}{n} \right) \theta \left( B(n) \right) \\
= \left( \frac{n \theta}{n^2} \right) \theta \left( \frac{n}{n(n+2)} \right) = \frac{\theta^2}{n(n+2)}
\]

Compare to \( \text{Var} \bar{X} = \frac{\theta^2}{3n} \)