STAT 5430
Lecture 06, M, Feb 3

practice on point estimation

- Homework 1 is assigned in Canvas
  (submit/due by Monday, Feb 3, by midnight)
- Homework 2 is assigned in Canvas
  (due by next Monday, Feb 10, by midnight)

Office hours Mine: FM, 12-1 PM & by appointment
TA (Min-Yi): WR 11-12 in Snedecor 2404

Criteria for Evaluating Point Estimators

Bias, cont'd

2. It is NOT always possible to find an U.E. of $\gamma(\theta)$

   $\leftarrow$ data

   Example: Let X be Binomial(n,p), $0 < p < 1$. Show there is no U.E. of $\gamma(p) = 1/p$.

Solution: If possible, suppose $h(X)$ is U.E of $1/p$

$$\Rightarrow E_p h(X) = \sum_{x=0}^{n} h(x)\binom{n}{x}p^x(1-p)^{n-x} = 1/p, \ \forall\, 0<p<1$$

multiply both sides by p & let $p \downarrow 0$

$$\underbrace{\sum_{x=0}^{n} h(x)\binom{n}{x}p^{x+1}(1-p)^{n-x}}_{\downarrow 0 \text{ as } p \downarrow 0} = 1, \ \forall\, 0<p<1$$

a contradiction!

Note: $\frac{X}{n}$ is estimator of p $\Rightarrow$ $\frac{n}{X}$ is estimator of $1/p$
(U.E of p) but not U.E.

3. Unbiasedness, while good, is not everything! $\leftarrow$ X_i is $\begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } (1-p) \end{cases}$

   Example: Let $X_1, \ldots, X_n$ be iid Bernoulli(p), $0 < p < 1$, and consider two estimators of p given by

$$T_1 = X_1, \qquad T_2 = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Note: $E_p(T_1) = E_p(X_1) = p$, $E_p(T_2) = E_p(\bar{X}_n) = E_p(X_1) = p$, $\forall\, 0<p<1$

Both T_1 & T_2 are U.E of p.

Note:

$$E_p(T_1 - p)^2 = E_p(X_1 - E_p(X_1))^2 = Var_p(X_1) = p(1-p)$$

$$E_p(T_2 - p)^2 = E_p(\bar{X}_n - E_p(\bar{X}_n))^2 = Var_p(\bar{X}_n) = \frac{Var_p(X_1)}{n} = \frac{p(1-p)}{n}$$

$\bar{X}_n$ is better (smaller variance) "smaller expected squared distance from center p"

Criteria for Evaluating Point Estimators

Uniform Minimum Variance Unbiased Estimator (UMVUE)

"best U.E"

Definition: Let $f(x_1, \ldots, x_n|\theta)$ be the joint pdf/pmf of $X_1, \ldots, X_n$. Then, an estimator T of a real-valued parametric function $\gamma(\theta)$ is called the **Uniform Minimum Variance Unbiased Estimator (UMVUE)** of $\gamma(\theta)$ if

1. T is an U.E. of $\gamma(\theta)$, that is, $\mathrm{E}_\theta(T) = \gamma(\theta)$, $\forall \theta \in \Theta$

2. $\mathrm{Var}_\theta(T) < \infty$, $\forall \theta \in \Theta$

3. Given any other U.E. of $\gamma(\theta)$, say T_1, it holds that

$$\mathrm{Var}_\theta(T) \leq \mathrm{Var}_\theta(T_1), \quad \forall \theta \in \Theta.$$

T has the smallest (best) variance compared to any other U.E T_1 of $\gamma(\theta)$

Finding a UMVUE

There are two general strategies for finding a UMVUE:

- Use CRLB (this doesn't always work)
  next
- Use "sufficiency" + "completeness" (later)

Criteria for Evaluating Point Estimators

Cramèr-Rao Lower Bound (CRLB)

Motivation for the CRLB:

← given T is U.E of $\gamma(\theta)$

- Suppose T is an *unbiased* estimator of a *real-valued* parametric function $\gamma(\theta)$ (i.e., $\mathrm{E}_\theta(T) = \gamma(\theta)$, any $\theta \in \Theta$) & wish to know if T is the UMVUE for $\gamma(\theta)$.

  ← want to check if T is UMVUE

- Suppose further that we know of a function of θ, say $c(\theta)$, where it holds that

$$\mathrm{Var}_\theta(T_1) \geq c(\theta) \quad \text{for } \textit{any} \text{ unbiased estimator } T_1 \text{ of } \gamma(\theta) \text{ and for any } \theta \in \Theta.$$

← suppose some lower bound (LB) on variance can be found

- If you find $\mathrm{Var}_\theta(T) = c(\theta)$ for all $\theta \in \Theta$, then you'd know T is the UMVUE.
- Sometimes you can explicitly find such a lower bound $c(\theta)$ by the **Cramèr-Rao Inequality** or **Cramèr-Rao Lower Bound**

Theorem (Cramèr-Rao Inequality): *Let $f(x_1, x_2, \ldots, x_n|\theta)$, $\theta \in \Theta$, be the joint pdf/pmf of $X_1, X_2, \ldots, X_n$. Assume that*

e.g. rules out UNIF(0,θ)

1. *Θ is an open subset of $\mathbb{R}$*
2. *$A \equiv \{(x_1, x_2, \ldots, x_n) : f(x_1, x_2, \ldots, x_n|\theta) > 0\}$ does not depend on θ*
3. *$\dfrac{d\, f(x_1, x_2, \ldots, x_n|\theta)}{d\theta}$ exists on Θ, for all $(x_1, x_2, \ldots, x_n) \in A$*
4. *For any estimator $T^* = T^*(X_1, X_2, \ldots, X_n)$ with $\mathrm{E}(T^*)^2 < \infty$ for all θ, it holds that*

smoothness conditions on pdf/pmf of data

$$\frac{d\,\mathrm{E}_\theta(T^*)}{d\theta} = \begin{cases} \displaystyle\int_A T^*(x_1, x_2, \ldots, x_n) \frac{d\, f(x_1, x_2, \ldots, x_n|\theta)}{d\theta} dx_1 dx_2 \ldots dx_n & \text{if } X_i\text{'s continuous} \\ \displaystyle\sum_{(x_1, x_2, \ldots, x_n) \in A} T^*(x_1, x_2, \ldots, x_n) \frac{d\, f(x_1, x_2, \ldots, x_n|\theta)}{d\theta} & \text{if } X_i\text{'s discrete} \end{cases}$$

5. *For all $\theta \in \Theta$,* $\quad 0 < \underbrace{I_n(\theta)}_{\text{sample size } n} \equiv \mathrm{E}_\theta\left[\left(\dfrac{d \log f(X_1, X_2, \ldots, X_n|\theta)}{d\theta}\right)^2\right] < \infty$

"information number"

Then, for any unbiased estimator T of $\gamma(\theta)$, it holds that

$$\mathrm{Var}_\theta(T) \geq \underbrace{\frac{(\gamma'(\theta))^2}{I_n(\theta)}}_{\text{CRLB } c(\theta)} \quad \text{for all } \theta \in \Theta \tag{1}$$

where $\gamma'(\theta) \equiv d\,\gamma(\theta)/d\theta$ is assumed to exist on Θ.

Criteria for Evaluating Point Estimators

Cramèr-Rao Lower Bound (CRLB)

Remarks:

- The right-hand side of (1) on page 23 is called the **Cramèr-Rao Lower Bound**.

- Conditions 1 - 5 in the Theorem are called the "Cramèr-Rao Regularity Conditions." These are satisfied if $X_1, X_2, \ldots, X_n$ are a random sample from the 1-parameter exponential family. Eg., Binomial(n, θ), Poisson(θ), Geometric(θ), $\mathrm{N}(\theta, \sigma^2)$, $\mathrm{N}(\mu, \theta)$, gamma(α, θ), gamma(θ, β)

- $I_n(\theta)$ is called the **Fisher Information number** (for size n sample)

- If $X_1, X_2, \ldots, X_n$ are iid with common pdf/pmf $f(x|\theta)$, then

to compute info numbers

$$I_n(\theta) = nI_1(\theta), \quad \text{where } I_1(\theta) = \mathrm{E}_\theta\left[\left(\frac{d \log f(X_1|\theta)}{d\theta}\right)^2\right] \tag{2}$$

and $I_1(\theta)$ represents the Fisher information for one observation.

- If $\dfrac{d^2 f(x_1, x_2, \ldots, x_n|\theta)}{d\theta^2}$ exits on Θ, for all $(x_1, x_2, \ldots, x_n) \in A$, then

$$I_n(\theta) = \mathrm{E}_\theta\left[\left(\frac{d \log f(X_1, X_2, \ldots, X_n|\theta)}{d\theta}\right)^2\right] = -\mathrm{E}_\theta\left(\frac{d^2 \log f(X_1, X_2, \ldots, X_n|\theta)}{d\theta^2}\right).$$

If, in addition, $X_1, X_2, \ldots, X_n$ are iid with common pdf/pmf $f(x|\theta)$, then we have

two ways to compute $I_1(\theta)$

$$I_n(\theta) = nI_1(\theta) \text{ where } I_1(\theta) = \mathrm{E}_\theta\left[\left(\frac{d \log f(X_1|\theta)}{d\theta}\right)^2\right] = -\mathrm{E}_\theta\left(\frac{d^2 \log f(X_1|\theta)}{d\theta^2}\right)$$

Criteria for Evaluating Point Estimators

ASIDE

Cramèr-Rao Lower Bound (CRLB)

Proof of (2), page 24/continuous case. For any sample size n,

$$
\begin{aligned}
& \mathrm{E}_\theta\left(\frac{d \log f(X_1, X_2, \ldots, X_n|\theta)}{d\theta}\right) \\
&= \int_A \frac{d \log f(x_1, x_2, \ldots, x_n|\theta)}{d\theta} f(x_1, x_2, \ldots, x_n|\theta) dx_1, \ldots, dx_n \\
&= \int_A \frac{d\, f(x_1, x_2, \ldots, x_n|\theta)}{d\theta} \frac{f(x_1, x_2, \ldots, x_n|\theta)}{f(x_1, x_2, \ldots, x_n|\theta)} dx_1, \ldots, dx_n \quad \text{derivative of log} \\
&= \int_A 1 \cdot \frac{d\, f(x_1, x_2, \ldots, x_n|\theta)}{d\theta} dx_1, \ldots, dx_n \\
&= \frac{d\,\mathrm{E}_\theta(1)}{d\theta} \quad \text{by condition 4. of Theorem with } T^* = 1 \\
&= \frac{d}{d\theta} 1 \\
&= 0
\end{aligned}
$$

so that

$$
\begin{aligned}
I_n(\theta) &= \mathrm{E}_\theta\left[\left(\frac{d \log f(X_1, X_2, \ldots, X_n|\theta)}{d\theta}\right)^2\right] \\
&= \mathrm{Var}_\theta\left(\frac{d \log f(X_1, X_2, \ldots, X_n|\theta)}{d\theta}\right) \\
&= \mathrm{Var}_\theta\left(\frac{d}{d\theta} \sum_{i=1}^n \log f(X_i|\theta)\right) \quad \text{since } f(x_1, x_2, \ldots, x_n|\theta) = \textstyle\prod_{i=1}^n f(x_i|\theta) \\
&= \mathrm{Var}_\theta\left(\sum_{i=1}^n \frac{d \log f(X_i|\theta)}{d\theta}\right) \\
&= \sum_{i=1}^n \mathrm{Var}_\theta\left(\frac{d \log f(X_i|\theta)}{d\theta}\right) \quad \text{sum of independent variables} \\
&= n\mathrm{Var}_\theta\left(\frac{d \log f(X_1|\theta)}{d\theta}\right) \quad \text{iid variables} \\
&= n\mathrm{E}_\theta\left[\left(\frac{d \log f(X_1|\theta)}{d\theta}\right)^2\right] \\
&= nI_1(\theta)
\end{aligned}
$$

← $I_n(\theta)$ is a type of variance

(optional)

Criteria for Evaluating Point Estimators

Cramèr-Rao Lower Bound (CRLB)

Example: Suppose $X_1, \ldots, X_n$ are iid Exponential(θ), $\theta > 0$ with density

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & x > 0 \\ 0 & \text{otherwise.} \end{cases} \quad \Rightarrow E_\theta(X_1) = \theta, \quad Var_\theta(X_1) = \theta^2$$

Find the CRLB for estimating $\gamma(\theta) = \theta$ and the UMVUE of θ.

Solution: $I_n(\theta) = n I_1(\theta)$

$$\frac{d \log f(X_1|\theta)}{d\theta} = \frac{d\left[-\log\theta - X_1/\theta\right]}{d\theta} = -1/\theta + \frac{X_1}{\theta^2}$$

$$I_1(\theta) = E_\theta\left[\left(\frac{-1}{\theta} + \frac{X_1}{\theta^2}\right)^2\right] = \frac{1}{\theta^4} \underbrace{E_\theta\left[(X_1-\theta)^2\right]}_{Var_\theta(X_1) = \theta^2} = \frac{\theta^2}{\theta^4} = \frac{1}{\theta^2}$$

$$I_1(\theta) = -E_\theta\left[\frac{d^2 \log f(X_1|\theta)}{d\theta^2}\right] = -E_\theta\left[1/\theta^2 - \frac{2X_1}{\theta^3}\right]$$

$$= -\left(1/\theta^2 - \frac{2E_\theta X_1}{\theta^3}\right)$$

$$= -\left(1/\theta^2 - \frac{2\theta}{\theta^3}\right) = 1/\theta^2$$

$\gamma(\theta) = \theta$

$\gamma'(\theta) = 1$ $\Rightarrow$ CRLB is $\frac{[\gamma'(\theta)]^2}{I_n(\theta)} = \frac{[1]^2}{n \cdot 1/\theta^2} = \frac{\theta^2}{n}$

check: $\bar{X}_n \Rightarrow E_\theta(\bar{X}_n) = E_\theta(X_1) = \theta, \ \forall \theta > 0$

($\bar{X}_n$ is UE of θ)

& $Var_\theta(\bar{X}_n) = Var_\theta(X_1)/n = \theta^2/n \equiv$ CRLB

Hence, $\bar{X}_n$ is UMVUE of θ!