STAT 5430
Lecture 12, M, Feb 17

After HW3, NO new homework this week →

- Homework 3 is assigned in Canvas (due by Monday, Feb 17 by midnight)
  (practice on CRLB/UMVUE/Bayes)
- Solutions to Homeworks 1-2 posted.

- Exam 1 is scheduled for W, Feb 26 (two weeks)
  6:15-8:15 PM (Snedecor seminar room 3105)
  - No regular class on W, Feb 26
  - See Canvas for study guide, practice exams
  - can bring 1 page formula sheet (front/back) with anything on it
  - see Canvas for a "canned" sheet
  - I'll provide table with STAT 5420 distributions (see Canvas)

Large Sample Properties of Estimators

Asymptotic

Terminology

$n \equiv$ "sample size"

Definition: Let $\{T_n\}$ be a sequence of estimators of a parametric function $\gamma(\theta)$.

1. $\{T_n\}$ is called **consistent** for $\gamma(\theta)$ if, for any $\epsilon > 0$, "little distance"

   (a) $\lim_{n\to\infty} P_\theta\left(|T_n - \gamma(\theta)| < \epsilon\right) = 1$, for all $\theta \in \Theta$,

   (b) or, equivalently, if $\lim_{n\to\infty} P_\theta\left(|T_n - \gamma(\theta)| \geq \epsilon\right) = 0$, for all $\theta \in \Theta$.

   (c) or, equivalently, if "T_n converges in probability to $\gamma(\theta)$ as $n \to \infty$" i.e, $T_n \xrightarrow{p} \gamma(\theta)$ as $n \to \infty$.

2. $\{T_n\}$ is called **mean squared error consistent (MSEC)** for $\gamma(\theta)$ if

$$\lim_{n\to\infty} \text{MSE}_\theta(T_n) \equiv \lim_{n\to\infty} \text{E}_\theta\left(\left[T_n - \gamma(\theta)\right]^2\right) = 0, \quad \forall\theta \in \Theta.$$

3. $\{T_n\}$ is called **asymptotically unbiased** for $\gamma(\theta)$ if

$$\lim_{n\to\infty} \text{E}_\theta(T_n) = \gamma(\theta), \quad \forall\theta \in \Theta.$$

$$\lim_{n\to\infty}\left[\underbrace{E_\theta(T_n) - \gamma(\theta)}_{\equiv \text{bias } b_\theta[T_n]}\right] = 0$$

Definitions 1-3 represent different senses in which T_n can be "close" to $\gamma(\theta)$ as sample size $n \to \infty$

Large Sample Properties of Estimators

Terminology, cont'd

Example. Let $X_1, \ldots, X_n$ be iid with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$.

parameters $\theta = (\mu, \sigma^2)$

- Consider the sequence of estimators $T_n = \bar{X}_n$, $n \geq 1$, of μ.

It turns out $\bar{X}_n \xrightarrow{P} EX_1 = \mu$ as $n \to \infty$ by WLLN (later)

or pick $\epsilon > 0$,

$$P_\theta(|\bar{X}_n - \mu| \geq \epsilon) = P_\theta(|\bar{X}_n - \mu|^2 \geq \epsilon^2)$$

Markov inequality

$$\leq \frac{E_\theta(|\bar{X}_n - \mu|^2)}{\epsilon^2} = \frac{Var_\theta(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \to 0 \text{ as } n \to \infty \text{ for any } \theta = (\mu, \sigma^2)$$

$\therefore \bar{X}_n$ is consistent for μ!

- $\text{MSE}_\theta(\bar{X}_n) = \text{E}_\theta\left(\left[\bar{X}_n - \mu\right]^2\right)$

$= Var_\theta(\bar{X}_n) = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$ for any $\theta = (\mu, \sigma^2)$

$\therefore \bar{X}_n$ is MSEC for μ too!

← MLE/MOM of σ^2 when data are normal

- Consider the sequence of estimators $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2 = \frac{n-1}{n}S^2$ of σ^2, where $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2$, $n \geq 1$.

$$E_\theta(\hat{\sigma}_n^2) = \frac{n-1}{n} E_\theta(S^2) = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

So $\hat{\sigma}_n^2$ is biased, but $\lim_{n \to \infty} E_\theta(\hat{\sigma}_n^2) = \sigma^2$.

Hence, $\hat{\sigma}_n^2$ is asym. unbiased.

Large Sample Properties of Estimators

Tools for Showing Consistency

Question: Is $\hat{\sigma}_n^2$ consistent for σ^2?

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X}_n)^2$$

notice 2 types of sample means (for X_i^2 & X_i)

Two useful results for establishing consistency $W_i \equiv \begin{pmatrix} W_{i,1} \\ \vdots \\ W_{i,k} \end{pmatrix}$

1. (WLLN): If $W_1, W_2, \ldots$ are iid k-dimensional random vectors with $\mathrm{E}|W_1| \equiv \mathrm{E}|W_{1,1}| + \cdots + \mathrm{E}|W_{1,k}| < \infty$ where $W_1 = (W_{1,1}, \ldots, W_{1,k})'$, then

weak law of large numbers

$$\bar{W}_n = \frac{1}{n}\sum_{i=1}^{n} W_i \xrightarrow{p} \mathrm{E}(W_1) \equiv \begin{pmatrix} \mathrm{E}(W_{1,1}) \\ \vdots \\ \mathrm{E}(W_{1,k}) \end{pmatrix} \quad \text{as } n \to \infty.$$

← each component of W_1 has mean

k sample averages

$|W_i| = |W_{i,1}| + \cdots + |W_{i,k}|$

2. (Continuous Mapping Theorem): Suppose $Y_n \xrightarrow{p} Y$, where $\{Y_n\}, Y$ are k-dimensional random vectors, $k \geq 1$. Then,

   (a) for any continuous function $g : \mathbb{R}^k \to \mathbb{R}^p$ $(p \geq 1)$, we have that

$$g(Y_n) \xrightarrow{p} g(Y) \quad \text{as } n \to \infty.$$

special case of (a) $\quad Y_n \xrightarrow{P} c$ as $n \to \infty$

   (b) if $P(Y = c) = 1$ for some $c \in \mathbb{R}^k$ (i.e., Y is essentially a constant c) and $g : \mathbb{R}^k \to \mathbb{R}^p$ is *continuous at* $c \in \mathbb{R}^k$, then

$$g(Y_n) \xrightarrow{p} g(c) \quad \text{as } n \to \infty.$$

Large Sample Properties of Estimators

Showing Consistency

Example. Let $X_1, \ldots, X_n$ be iid with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. Show $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X}_n)^2$ is consistent for σ^2.

Let $W_i = \begin{pmatrix} X_i^2 \\ X_i \end{pmatrix}$, for $i \geq 1$.

Then $W_1, \ldots, W_n$ are iid random vectors

with $E|W_1| = E|X_1^2| + E|X_1| < \infty$.

By WLLN, $\bar{W}_n = \frac{1}{n}\sum_{i=1}^{n} W_i = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} X_i^2 \\ \frac{1}{n}\sum_{i=1}^{n} X_i \end{pmatrix} \xrightarrow{P} \begin{pmatrix} EX_1^2 \\ EX_1 \end{pmatrix}$

as $n \to \infty$.

Note $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X}_n)^2$

$= g\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2, \frac{1}{n}\sum_{i=1}^{n} X_i\right)$, where

$g(x,y) = x - y^2$

Since $g(x,y)$ is continuous on $\mathbb{R}^2$,

by continuous mapping theorem,

$\hat{\sigma}_n^2 = g\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2, \frac{1}{n}\sum_{i=1}^{n} X_i\right) \xrightarrow{P} g(EX_1^2, EX_1)$

$= EX_1^2 - (EX_1)^2$

$= \sigma^2$.

$\therefore$ $\hat{\sigma}_n^2$ is consistent for σ^2 //

Large Sample Properties of Estimators

Remarks & Relationships on Consistency

1. MSEC $\Rightarrow$ Consistency

   Pick $\epsilon > 0$ & use Chebychev,

   Markov's inequality — definition of MSE

   $$P_\theta(|T_n - \gamma(\theta)| > \epsilon) = P_\theta(|T_n - \gamma(\theta)|^2 > \epsilon^2) \le \frac{\mathrm{E}_\theta(|T_n - \gamma(\theta)|^2)}{\epsilon^2} = \frac{\mathrm{MSE}_\theta(T_n)}{\epsilon^2} \to 0 \text{ as } n \to \infty$$

   $\therefore$ T_n is consistent for $\gamma(\theta)$.

2. $\{T_n\}$ is MSEC $\Leftrightarrow$ $\lim_{n\to\infty} \mathrm{Var}_\theta(T_n) = 0$ & $\{T_n\}$ is asymptotically unbiased.

   $\lim_{n\to\infty} b_\theta(T_n) = 0$

   This follows because

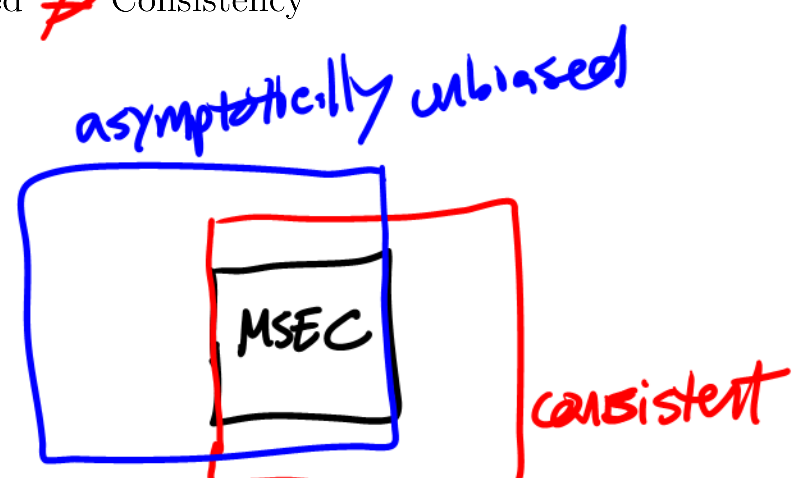
   $$\underbrace{\mathrm{MSE}_\theta(T_n)}_{\ge 0} = \underbrace{\mathrm{Var}_\theta(T_n)}_{\ge 0} + \underbrace{[b_\theta(T_n)]^2}_{\ge 0}.$$

   $\therefore$ MSEC $\Rightarrow$ asymp. unbiased

3. Consistency $\not\Rightarrow$ Asymptotically Unbiased

   Asymptotically Unbiased $\not\Rightarrow$ Consistency

picture



Large Sample Properties of Estimators

Remarks & Relationships on Consistency, cont'd

$\swarrow$ mean $E_\theta X_1 = \theta$

Example: Consider $X_1, X_2, \ldots$ iid $N(\theta, 1)$, $\theta \in \mathbb{R}$. Consider $T_n = \exp[(\bar{X}_n)^4]$ to estimate $\gamma(\theta) \equiv \exp(\theta^4)$

- Note that $\bar{X}_n \xrightarrow{p} \theta \Rightarrow$

by WLLN

$$T_n = \exp\left[(\bar{X}_n)^4\right] \xrightarrow{p} \exp[\theta^4] \equiv \gamma(\theta).$$

$\uparrow$ by continuous mapping theorem $g(\bar{X}_n) \xrightarrow{p} g(\theta)$

T_n is consistent for $\gamma(\theta)$

But, $\mathrm{E}_\theta(T_n) = +\infty$ for all n, so T_n is not asymptotically unbiased for $\gamma(\theta)$.

T_n has no finite mean $\qquad +\infty = \lim_{n\to\infty} E_\theta(T_n) \neq \gamma(\theta) < \infty$

- Let $T_n = \dfrac{1}{2}\left(X_1 + \bar{X}_n\right)$ be an estimator of $\mathrm{E}_\theta(X_1) = \theta$.

Note $\mathrm{E}_\theta(T_n) = \theta$, for all n, θ, so that $\{T_n\}$ is asymptotically unbiased for θ.

T_n is UE of $\theta \Rightarrow \lim_{n\to\infty} E_\theta(T_n) = \lim_{n\to\infty} \theta = \theta$

Also, the sampling distribution of T_n is normal with mean θ and variance

$$\mathrm{Var}_\theta(T_n) = \frac{1}{4}\mathrm{Var}_\theta\left(\frac{n+1}{n}X_1 + \frac{1}{n}\sum_{i=2}^{n} X_i\right) = \frac{1}{4}\frac{(n+1)^2 + n - 1}{n^2}.$$

check $\lim_{n\to\infty} \mathrm{Var}_\theta(T_n) = 1/4$

Pick any $\epsilon > 0$, let $Z \sim N(0,1)$, and note

$$\begin{aligned} P_\theta\left(|T_n - \theta| > \epsilon\right) &= P_\theta\left(\left|\frac{T_n - \theta}{\sqrt{\mathrm{Var}_\theta(T_n)}}\right| > \frac{\epsilon}{\sqrt{\mathrm{Var}_\theta(T_n)}}\right) \\ &= P\left(|Z| > \frac{\epsilon}{\sqrt{\mathrm{Var}_\theta(T_n)}}\right) \\ &\to P\left(|Z| > 2\epsilon\right) \neq 0. \end{aligned}$$

standard normal

Hence, T_n is not consistent for θ.

Large Sample Properties of Estimators

Asymptotic ~~Efficiency~~ Efficiency

Recall: For two *unbiased* estimators T and T^* of $\gamma(\theta)$, we compare the *variances* of the estimators to judge their *relative efficiency*, i.e. $RE(T, T^*, \theta) = \text{Var}_\theta(T^*)/\text{Var}_\theta(T)$. ie. choose UE with smaller variance

Similarly, we compare large-sample variances of **asymptotically unbiased** estimators.

Definitions: Let $\{T_n^*\}$ and $\{T_n\}$ be asymptotically unbiased for $\gamma(\theta)$. Then, define

$$\lim_{n\to\infty} E_\theta(T_n^*) = \gamma(\theta) = \lim_{n\to\infty} E_\theta(T_n)$$

1. The **asymptotic relative efficiency** of $\{T_n\}$ with respect to $\{T_n^*\}$ at θ is defined as

$$ARE(T_n, T_n^*, \theta) = \lim_{n\to\infty} \frac{\text{Var}_\theta(T_n^*)}{\text{Var}_\theta(T_n)}, \quad \theta \in \Theta$$

compare by limit of variance "ratio"

2. $\{T_n^*\}$ is called **asymptotically efficient** if

$$ARE(T_n, T_n^*, \theta) \leq 1$$

holds $\forall \theta \in \Theta$ for any other $\{T_n\}$ that is asymptotically unbiased for $\gamma(\theta)$.

as $n\to\infty$, T_n^* has smallest variance out of asymp. U.E.

3. The **asymptotic efficiency** of $\{T_n\}$ is defined as

$$AE(T_n, \theta) \equiv ARE(T_n, T_n^*, \theta)$$

if $\{T_n^*\}$ is asymptotically efficient.

Large Sample Properties of Estimators

Asymptotic Effiency

Previous example. Let $X_1, X_2, \ldots$ be iid uniform$(0, \theta)$, $\theta > 0$.

(from earlier work with MSE, etc)

Recall

- T_n = MME of θ based on $X_1, \ldots, X_n = 2\bar{X}_n$

  (Aside: $\bar{X}_n \xrightarrow{P} E_\theta X_1 = \theta/2$ & $2\bar{X}_n \xrightarrow{P} 2 \cdot \theta/2 = \theta$)

- T_n^* = MLE of θ based on $X_1, \ldots, X_n = \max_{1 \le i \le n} X_i = X_{(n)}$

- $\mathrm{E}_\theta T_n = \theta \qquad \mathrm{Var}_\theta(T_n) = \dfrac{3\theta^2}{n}$ ← looks like $\frac{1}{n}$

  MME is UE of θ & so is asymp. unbiased too.

- $\mathrm{E}_\theta T_n^* = \dfrac{n}{n+1}\theta \qquad \mathrm{Var}_\theta(T_n^*) = \dfrac{n\theta^2}{(n+1)^2(n+2)}$ ← looks like $\frac{1}{n^2}$

  The MLE is also asymp. unbiased: $\lim_{n\to\infty} E_\theta(T_n^*) = \theta$.

$$ARE(T_n, T_n^*, \theta) = \lim_{n\to\infty} \frac{Var_\theta(T_n^*)}{Var_\theta(T_n)} = 0.$$

MME MLE