MA 3140: Statistical Inference

Dr. Sameen Naqvi
Department of Mathematics, IIT Hyderabad
Email id: sameen@math.iith.ac.in

A method to find UEs

Solve directly the equation

$$E[T(X)] = g(\theta), \forall\ \theta \in \Theta.$$

- **Example:**

  Let X be a truncated Poisson r.v. with zero missing

$$P(X = x) = \frac{1}{e^{\lambda} - 1} \frac{\lambda^x}{x!}, \quad x = 1, 2, \ldots$$

What will an UE of λ?

A method to find UEs cont'd

$$
\begin{aligned}
E[T(X)] &= \lambda,\ \forall\ \lambda > 0 \\
\sum_{x=1}^{\infty} T(x) \frac{1}{e^{\lambda} - 1}\ \frac{\lambda^x}{x!} &= \lambda,\ \forall\ \lambda > 0 \\
T(1)\lambda + \frac{T(2)}{2!}\lambda^2 + \ldots &= \lambda(e^{\lambda} - 1) \\
&= \lambda\Big[\lambda + \frac{\lambda^2}{2!} + \ldots\Big],\ \forall\ \lambda > 0
\end{aligned}
$$

Since the two power series can be identical on an open interval iff all their coefficients match, we get

A method to find UEs cont'd

$$
\begin{aligned}
T(1) &= 0, \\
T(2) &= 2! = 2, \\
T(3) &= 3!/2! = 3, \\
&\vdots \\
T(r) &= r!/(r-1)! = r
\end{aligned}
$$

So the UE is

$$
T(X) = \begin{cases} 0, \text{ if } X = 1 \\ X, \text{ if } X = 2, 3, \dots \end{cases}
$$

Remark

(i) UEs may not exist.

Example: Let $X \sim Bin(n, p)$ and $g(p) = p^{n+1}$.

$$
\begin{aligned}
E[T(X)] &= p^{n+1}, \ \forall\ 0 \le p \le 1 \\
\Longrightarrow \sum_{x=0}^{n} T(x) \binom{n}{x} p^x (1-p)^{n-x} &= p^{n+1}, \ \forall\ 0 \le p \le 1
\end{aligned}
$$

Since the LHS is a polynomial of degree atmost n, it cannot be equal to the RHS which has a power $n+1$. Thus the above equation has no solution and UE of p^{n+1} does not exist.

Remark cont'd

(ii) UEs may not be reasonable.

Example 1:

Consider Example 4 where we were interested to find the probability of no occurrence and $g(p) = e^{-\lambda} \in (0, 1)$. However, the UE we obtained was:

$$I(X) = \begin{cases} 1, & \text{if } X = 0 \\ 0, & \text{if } X \neq 0. \end{cases}$$

So, this is not a very proper estimator.

Remark cont'd

Example 2:

Consider $X \sim P(\lambda)$ and $g(\lambda) = e^{-3\lambda}$.

$$E[T(X)] = e^{-3\lambda}, \ \forall \ \lambda > 0$$

$$\Longrightarrow \sum_{x=0}^{\infty} T(x) \frac{e^{-\lambda}\lambda^x}{x!} = e^{-3\lambda}, \ \forall \ \lambda > 0$$

$$\Longrightarrow \sum_{x=0}^{\infty} T(x) \frac{\lambda^x}{x!} = \frac{e^{-3\lambda}}{e^{-\lambda}} = e^{-2\lambda}$$

$$\Longrightarrow \frac{T(0)\lambda^0}{0!} + \frac{T(1)\lambda^1}{1!} + \frac{T(2)\lambda^2}{2!} + \ldots = 1 - 2\lambda + \frac{(2\lambda)^2}{2!} - \frac{(2\lambda)^3}{3!}$$

Remark cont'd

Comparing the coefficients, we get

$$\begin{aligned} T(0) &= 1, \\ T(1) &= -2, \\ T(2) &= 4, \\ T(3) &= -8, \\ &\vdots \end{aligned}$$

So the UE is $T(X) = (-2)^X$.

Here, it can be noted that the estimator is taking absurdly different values from that of the parametric function $g(\lambda) = e^{-3\lambda}$.

Remark cont'd

(iii) MME and MLE may not be unbiased

Example:

Let $X \sim \mathrm{U}(0, \theta)$, where $\theta \in \Theta = (0, \infty)$ is unknown, and let the estimand be $g(\theta) = \sqrt{\theta}$. It can be shown that MME and MLE of $g(\theta)$ is $\sqrt{2X}$ and $\sqrt{X}$, respectively. Check whether these estimators are unbiased.

$$E[\sqrt{X}] = \int_0^\theta \frac{\sqrt{x}}{\theta} dx = \frac{2}{3}\sqrt{\theta} \neq \theta.$$

$$E[\sqrt{2X}] = \sqrt{2}E[\sqrt{X}] = \frac{2\sqrt{2}}{3}\sqrt{\theta} \neq \theta.$$

Thus, none of these estimators is unbiased.

Remark cont'd

(iv) Typically, there are many UEs for a given estimand.

Example:

Let $X_1, \ldots, X_n$ be a r.s. from a $N(\theta, 1)$ distribution, where $\theta \in \Theta = (-\infty, \infty)$ is unknown, and let the estimand be $g(\theta) = \theta$.

Then, $\overline{X}$, X_i, $\frac{X_i + X_j}{2}$, $X_i + X_j - X_k$, and so on, are all unbiased for estimating $g(\theta)$.

UE based on MLE

Remark:

Given any UE T, which is not based on the MLE, there exists an UE based on the MLE, say T_M, and that estimator is a better UE.

Therefore, in finding a sensible UE for an estimand $g(\theta)$, we typically start with the MLE of $g(\theta)$.

If it is unbiased, then we have found the estimator we want. If it is not unbiased, we modify it to make it unbiased.

Example 1: Let $X_1, \ldots, X_n$ be a r.s. from a $Poisson(\theta)$ distribution, where $\theta \in \Theta = (-\infty, \infty)$ is unknown, and let the estimand be $g(\theta) = P_\theta(X = 0) = e^{-\theta}$. Find the UE of $g(\theta)$.

Solution: We know that MLE of $g(\theta)$ is $e^{-\overline{X}}$.

Let $Z = \sum_{i=1}^{n} X_i$ so that $Z \sim Poisson(n\theta)$ and $\overline{X} = \frac{Z}{n}$.

We want the estimator $T(\overline{X}) = T(\frac{Z}{n})$ s.t.

$$E_\theta(T(\overline{X})) = e^{-\theta}, \ \forall \ \theta > 0$$

$$\Leftrightarrow \sum_{j=0}^{\infty} T\left(\frac{j}{n}\right) \frac{e^{-n\theta}(n\theta)^j}{j!} = e^{-\theta}, \ \forall \ \theta > 0$$

UE based on MLE

$$
\begin{aligned}
&\Leftrightarrow \sum_{j=0}^{\infty} T\left(\frac{j}{n}\right)\frac{n^j}{j!}\theta^j = e^{(n-1)\theta},\ \forall\ \theta > 0 \\
&\Leftrightarrow \sum_{j=0}^{\infty} T\left(\frac{j}{n}\right)\frac{n^j}{j!}\theta^j = \sum_{j=0}^{\infty}\frac{(n-1)^j}{j!}\theta^j,\ \forall\ \theta > 0 \\
&\Leftrightarrow \qquad\quad T\left(\frac{j}{n}\right) = \left(1-\frac{1}{n}\right)^j,\ j = 0, 1, \ldots .
\end{aligned}
$$

Thus, it follows that the unbiased estimator based on the MLE is $\left(1-\frac{1}{n}\right)^{n\overline{X}}$.

UE based on MLE

Example 2: Let $X_1, \dots, X_n$ be a r.s. from a $U(0, \theta)$ distribution, where $\theta \in \Theta = (0, \infty)$ is unknown, and let the estimand be $g(\theta) = \theta^r$, for some positive integer r. Find the the UE of $g(\theta)$ based on the MLE.

Solution:

As we know that MLE of θ is $X_{(n)}$, MLE of $g(\theta)$ is $X_{(n)}^r$.
The distribution function of $X_{(n)}$ is

$$
\begin{aligned}
F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) \\
&= P(X_1 \leq x, \dots, X_n \leq x) \\
&= [P(X_1 \leq x)]^n = \begin{cases} 0, & x < 0 \\ (\frac{x}{\theta})^n, & 0 \leq x \leq \theta \\ 1, & x > \theta. \end{cases}
\end{aligned}
$$

UE based on MLE

So the pdf of $X_{(n)}$ is:

$$\because f_{X_{(n)}}(x) = \begin{cases} \frac{nx^{n-1}}{\theta^n}, & 0 \leq x \leq \theta \\ 0, & \text{otherwise.} \end{cases}$$

Now,

$$E(X_{(n)}^r) = \int_0^\theta x^r \; f_{X_{(n)}}(x)dx = \int_0^\theta x^r \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+r}\theta^r$$

$$\implies E\left[\frac{n+r}{n} X_{(n)}^r\right] = \theta^r$$

Thus, $\frac{n+r}{n} \; X_{(n)}^r$ is an UE of $g(\theta) = \theta^r$.

Consistency

An estimator $T_n = T(X_1, \ldots, X_n)$ is said to be consistent for estimating $g(\theta)$ if for each $\epsilon > 0$

$$P(|T_n - g(\theta)| > \epsilon) \longrightarrow 0,$$

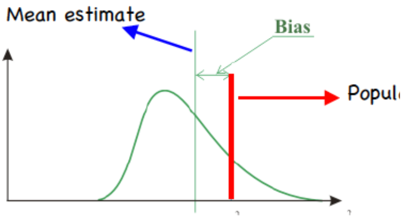
as $n \longrightarrow \infty$, $\forall\ \theta \in \Theta$.

Equivalently,

$$T_n \xrightarrow{p} g(\theta).$$

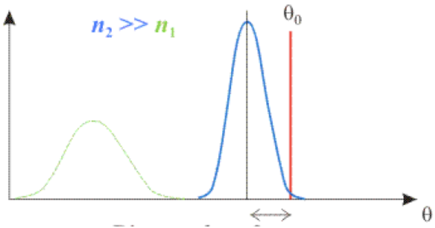
Unbiasedness vs Consistency

Unbiasedness



- It is a finite-sample argument which says that the average (expected) result of an estimator is equal to the true parameter.

Consistency



- Consistency says that an estimator will converge to the true parameter as $n \longrightarrow \infty$.

Example 1

Let $X_1, \ldots, X_n$ be a random sample from a population with mean μ and variance σ^2.

$$
\begin{aligned}
P(|\overline{X} - \mu| > \epsilon) &\leq \frac{Var(\overline{X})}{\epsilon^2} \qquad \text{(by Chebyshev inequality)} \\
&= \frac{\sigma^2}{n\epsilon^2} \\
&\longrightarrow 0 \text{ as } n \longrightarrow \infty
\end{aligned}
$$

Thus, sample mean $\overline{X}$ is a consistent estimator of population mean μ.

Chebyshev inequality

Suppose that a random variable X has finite mean μ and variance σ^2. Then for any $k > 0$,

$$P(\{|X - \mu| \geq k\}) \leq \frac{\sigma^2}{k^2}.$$

In other words, *if the variance is small, then X is unlikely to be too far from the mean.*
For $k = c\sigma$, we get

$$P(\{|X - \mu| \geq c\sigma\}) \leq \frac{\sigma^2}{c^2\sigma^2} = \frac{1}{c^2},$$

and for $c = 3$, we get

$$P(\{|X - \mu| \geq 3\sigma\}) \leq \frac{1}{9}.$$

Example 2

Let $X_1, \ldots, X_n$ be a sequence of i.i.d. r.v.s from a population with mean μ, then by Weak Law of Large Numbers,

$$\overline{X} \xrightarrow{p} \mu$$

In general, whenever the population mean exists, the sample mean is consistent for it.

In case of Cauchy distribution, $E(X)$ does not exist. In fact, distribution of $\overline{X}$ is same as that of X_1.

Invariance Property

- If T_n is consistent for θ and h is a continuous function, then $h(T_n)$ is consistent estimator for $h(\theta)$.
- Example: If T is consistent for θ, then T^2 is consistent for θ^2.
- Such a property does not hold in Unbiasedness.

  If T is UE for p, then $\frac{1}{T}$ is not an UE for $\frac{1}{p}$.

Important results

- If $E(T_n) = \theta_n \longrightarrow \theta$ and $V(T_n) = \sigma_n^2 \longrightarrow 0$, then T_n is consistent for θ.

- If T_n is consistent for θ and $a_n \to 1$, $b_n \to 0$, then $a_n T_n + b_n$ is also consistent for θ.

Examples:

$$\frac{n}{n+1} T_n + \frac{1}{n+1}, \qquad \frac{n+2}{n+4} T_n + \frac{1}{n}$$

Example 3: Shifted Exponential Distribution

Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with a location parameter μ. The pdf is

$$f_{X_i}(x) = \begin{cases} e^{-(x-\mu)}, & x > \mu \\ 0, & \text{otherwise.} \end{cases}$$

In general, the pdf of Shifted Exponential Distribution is:

$$f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{(x-\mu)}{\lambda}}, & x > \mu, \lambda > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Here, $E(X) = \lambda + \mu$, and $Var(X) = \lambda^2$.

Example 3 cont'd

In this example, $E(X_i) = \mu + 1$, and so, $E(\overline{X}) = \mu + 1$.

By WLLN, we get $\overline{X}$ as a consistent estimator of $\mu + 1$.

Thus, $T_1 = \overline{X} - 1$ is unbiased and consistent for estimating μ.

Example 3 cont'd

- Consider $X_{(1)} = \min\{X_1, \ldots, X_n\}$.
  The distribution function is

$$\begin{aligned} F_{X_{(1)}}(x) &= P(X_{(1)} \leq x) \\ &= 1 - P(X_{(1)} > x) \\ &= 1 - P(X_1 > x, \ldots, X_n > x) \\ &= 1 - P(X_1 > x) \ldots P(X_n > x) \\ &= 1 - [P(X_1 > x)]^n \\ &= 1 - [1 - F_{X_1}(x)]^n \\ &= 1 - e^{n(\mu - x)},\ x > \mu \end{aligned}$$

$$\because F_{X_i}(x) = \int_{\mu}^{x} e^{\mu - t} dt = 1 - e^{\mu - x}.$$

Example 3 cont'd

So the pdf of $X_{(1)}$ is

$$f_{X_{(1)}}(x) = ne^{n(\mu - x)}, \; x > \mu.$$

Here, $E(X_{(1)}) = \mu + \frac{1}{n}$ and $Var(X_{(1)}) = \frac{1}{n^2}$.

Consider $T_2 = X_{(1)} - \frac{1}{n}$. Then T_2 is unbiased for μ.

$Var(T_2) = \frac{1}{n^2} \longrightarrow 0$ as $n \longrightarrow \infty$.

Thus, T_2 is also unbiased and consistent for estimating μ.

Thanks for your patience!