Stat 4620: Day 17

(3/26)

Sec. 6.2: Efficiency of an estimator

Definition 6.2.1 (Efficient Estimator). Let Y be an unbiased estimator of a parameter θ in the case of point estimation. The statistic Y is called an **efficient estimator** of θ if and only if $V(Y)$ attains the Rao-Cramer lower bound.
Definition 6.2.2 (Efficiency). Let Y be an unbiased estimator of a parameter θ in the case of point estimation. The **efficiency** of Y is

$$e(Y) = \frac{\text{RCLB}}{V(Y)}$$

Example (Beta(θ, 1)) Let $x_1, \ldots, x_n$ be a random sample from a distribution with continuous pdf

$$f(x;\theta) = \theta x^{\theta-1},\ 0 < x < 1$$

where $\theta \in (0, \infty)$. Then

$$\ln f = \ln\theta + (\theta - 1)\ln x$$

$$\frac{\partial}{\partial\theta}\ln f = \frac{1}{\theta} + \ln x$$

and the information is

$$\mathbf{I}(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln f\right] = -E\left[-\frac{1}{\theta^2}\right] = \frac{1}{\theta^2}$$

Next, we find the mle of θ.

$$l(\theta) = \sum_{i=1}^{n}\ln f(x_i;\theta) = n\ln\theta + \theta\sum\ln x_i - \sum\ln x_i$$

$$\frac{\partial}{\partial\theta}l(\theta) = \frac{n}{\theta} + \sum\ln x_i$$

Equating to 0 and solving for θ, the mle is

$$\hat{\theta} = \frac{-n}{\sum\ln x_i}$$

It can be shown that $-\ln X_i$ has $\Gamma(1, 1/\theta)$ distribution and consequently, $W = -\sum_{i=1}^{n}\ln X_i$ has $\Gamma(n, 1/\theta)$ distribution. It has been shown in Chapter 3 that

$$E\left(W^k\right) = \frac{(n+k-1)!}{\theta^k(n-1)!}$$

for $k > -n$. Then

$$E\left(\hat{\theta}\right) = E\left(\frac{n}{W}\right) = nE\left(W^{-1}\right) = n\left(\frac{(n-2)!}{\theta^{-1}(n-1)!}\right) = \theta\frac{n}{n-1}$$

So the mle $\hat{\theta}$ is biased, but the bias disappears as $n \to \infty$. Note, however, that the estimator

$$Y = \frac{n-1}{n}\hat{\theta} = \frac{-(n-1)}{\sum \ln x_i}$$

is unbiased for θ. The variance of this unbiased estimator can be shown to equal

$$V(Y) = \frac{\theta^2}{n-2}$$

The RCLB is

$$\text{RCLB} = \frac{1}{n\mathbf{I}(\theta)} = \frac{\theta^2}{n}$$

The efficiency of the unbiased estimator is

$$e(Y) = \frac{\text{RCLB}}{V(Y)} = \frac{n-2}{n}$$

so the estimator Y is not efficient but is *asymptotically efficient.*

Sec. 6.2: Asymptotic normality of mle

Theorem 8. *Let $x_1, \ldots, x_n$ be a random sample from $f(x;\theta), \theta \in \Omega$. Let θ_0 denote the true value. Assume that (R0)-(R5) hold, and suppose that $0 < \mathbf{I}(\theta) < \infty$. Then the mle $\hat{\theta}_n$ satisfies*

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{D} N\left(0, \frac{1}{\mathbf{I}(\theta_0)}\right)$$

Comments:

1. The theorem says: for large n,

$$\hat{\theta}_n \dot{\sim} N\left(\theta_0, \frac{1}{n\mathbf{I}(\theta_0)}\right)$$

2. The theorem's conclusion is read as "*...converges in distribution to...*" and is formally defined as:

$$\lim_{n\to\infty} P\left[\sqrt{n\mathbf{I}(\theta_0)}\left(\hat{\theta}_n - \theta_0\right) \le t\right] = \Phi(t)$$

   where Φ is cdf of $N(0,1)$.

3. The asymptotic variance of $\hat{\theta}_n$ is the RCLB, so the mle is asymptotically optimal.

Proof.

$$l(\theta) = \ln L(\theta) = \ln \prod_{i=1}^{n} f(x_i,\theta) = \sum_{i=1}^{n} \ln f(x_i,\theta)$$

$$\frac{1}{\sqrt{n}} l'(\theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \ln f(x_i,\theta)$$

The score functions $\left\{\frac{\partial}{\partial\theta} \ln f(x_i,\theta)\right\}$ are independent and identically distributed random variables with mean 0 and variance $\mathbf{I}(\theta)$. By the Central Limit Theorem,

$$\frac{1}{\sqrt{n}} l'(\theta) \xrightarrow{D} N(0, \mathbf{I}(\theta)) \tag{1}$$

By the Law of Large Numbers

$$-\frac{1}{n}l''(\theta) = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\ln f(x_i,\theta) \xrightarrow{P} \mathbf{I}(\theta) \tag{2}$$

The mle $\hat{\theta}_n$ satisfies $l'(\hat{\theta}_n) = 0$. Using Taylor's expansion of $l'(\theta)$,

$$0 = l'(\hat{\theta}_n) = l'(\theta_0) + \left(\hat{\theta}_n - \theta_0\right) l''(\theta_0) + \frac{1}{2}\left(\hat{\theta}_n - \theta_0\right)^2 l'''(\theta_n^*)$$

where θ_n^* is between $\hat{\theta}_n$ and θ_0. Rearranging terms,

$$\left(\hat{\theta}_n - \theta_0\right) = \frac{-l'(\theta_0)}{l''(\theta_0) + \frac{1}{2}\left(\hat{\theta}_n - \theta_0\right) l'''(\theta_n^*)}$$

Furthermore,

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) = \frac{\frac{1}{\sqrt{n}}l'(\theta_0)}{-\frac{1}{n}l''(\theta_0) - \frac{1}{n}\frac{1}{2}\left(\hat{\theta}_n - \theta_0\right) l'''(\theta_n^*)} \equiv \frac{A_n}{B_n + C_n}$$

By equations (1) and (2), $A_n \xrightarrow{D} N(0, \mathbf{I}(\theta))$, $B_n \xrightarrow{P} \mathbf{I}(\theta)$, and $C_n \xrightarrow{P} 0$.

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{D} \frac{N(0,\mathbf{I}(\theta))}{\mathbf{I}(\theta) + 0} \sim N\left(0, \frac{1}{\mathbf{I}(\theta)}\right)$$

□

Definition 6.2.3 (Asymptotic Efficiency) Let $x_1, \ldots, x_n$ be independent and identically distributed random variables from a distribution with density $f(x;\theta)$. Suppose that $\hat{\theta}_{1n}(x_1, \ldots, x_n)$ is an estimator of θ_0 such that $\sqrt{n}\left(\hat{\theta}_{1n} - \theta_0\right) \xrightarrow{D} N\left(0, \sigma^2_{\hat{\theta}_{1n}}\right)$. Then

1. The **asymptotic efficiency** of $\hat{\theta}_{1n}$ is defined to be

$$e(\hat{\theta}_{1n}) = \frac{1/\mathbf{I}(\theta_0)}{\sigma^2_{\hat{\theta}_{1n}}}$$

2. The estimator $\hat{\theta}_{1n}$ is said to be **asymptotically efficient** if its asymptotic efficiency is 1.

3. Let $\hat{\theta}_{2n}(x_1, \ldots, x_n)$ be another estimator of θ_0 such that $\sqrt{n}\left(\hat{\theta}_{2n} - \theta_0\right) \xrightarrow{D} N\left(0, \sigma^2_{\hat{\theta}_{2n}}\right)$. Then the **asymptotic relative efficiency** (ARE) of $\hat{\theta}_{1n}$ to $\hat{\theta}_{2n}$ is

$$e\left(\hat{\theta}_{1n}, \hat{\theta}_{2n}\right) = \frac{\sigma^2_{\hat{\theta}_{2n}}}{\sigma^2_{\hat{\theta}_{1n}}}$$

Example 6.2.5 (ARE of the sample median to the mean)
Q: What is the ARE of the sample median to the sample mean?
A: *It depends on the underlying distribution.*
Case 1: Laplace

$$f(x_i) = \frac{1}{2}e^{-|x_i - \theta|}, \ -\infty < x_i < \infty, \ \ i = 1, \ldots, n$$

We already know that the sample median is mle of θ. By theorem, the asymptotic variance is

$$\sigma^2_{\hat{\theta}_{1n}} = \frac{1}{\mathbf{I}_\theta} = 1$$

since $\mathbf{I}_\theta = 1$ (show this). The asymptotic variance of the sample mean is

$$\sigma^2_{\hat{\theta}_{2n}} = 2$$

(show this). Then

$$\text{ARE (median, mean)} = \frac{\sigma^2_{\hat{\theta}_{2n}}}{\sigma^2_{\hat{\theta}_{1n}}} = \frac{2}{1} = 2.$$

Case 2: Normal

$$f(x_i) = \frac{1}{\sqrt{2\pi}} e^{-(x_i-\theta)^2/2}, \ -\infty < x_i < \infty, \ \ i = 1, \ldots, n$$

By Theorem 10.2.3, the asymptotic variance of the sample median is

$$\sigma^2_{\hat{\theta}_{1n}} = \frac{1}{4f^2(\theta_0)} = \frac{\pi}{2}$$

For the sample mean,

$$\sigma^2_{\hat{\theta}_{2n}} = 1$$

so

$$\text{ARE (median, mean)} = \frac{\sigma^2_{\hat{\theta}_{2n}}}{\sigma^2_{\hat{\theta}_{1n}}} = \frac{1}{\pi/2} = 2/\pi = .636$$