Biostatistics 602 - Statistical Inference
Lecture 12
Cramer-Rao Theorem

Hyun Min Kang

February 19th, 2013

Last Lecture

1. If you know MLE of θ, can you also know MLE of $\tau(\theta)$ for any function τ?

Last Lecture

1. If you know MLE of θ, can you also know MLE of $\tau(\theta)$ for any function τ?
2. What are plausible ways to compare between different point estimators?

Last Lecture

1. If you know MLE of θ, can you also know MLE of $\tau(\theta)$ for any function τ?
2. What are plausible ways to compare between different point estimators?
3. What is the *best unbiased estimator* or *uniformly unbiased minimium variance estimator (UMVUE)*?

Last Lecture

1. If you know MLE of θ, can you also know MLE of $\tau(\theta)$ for any function τ?
2. What are plausible ways to compare between different point estimators?
3. What is the *best unbiased estimator* or *uniformly unbiased minimium variance estimator (UMVUE)*?
4. What is the Cramer-Rao bound, and how can it be useful to find UMVUE?

Recap : Cramer-Rao inequality

Theorem 7.3.9 : Cramer-Rao Theorem

Let $X_1, \cdots, X_n$ be a sample with joint pdf/pmf of $f_{\mathbf{X}}(\mathbf{x}|\theta)$. Suppose $W(\mathbf{X})$ is an estimator satisfying

1. $E[W(\mathbf{X})|\theta] = \tau(\theta), \ \forall \theta \in \Omega.$
2. $\mathrm{Var}[W(\mathbf{X})|\theta] < \infty.$

For $h(\mathbf{x}) = 1$ and $h(\mathbf{x}) = W(\mathbf{x})$, if the differentiation and integrations are interchangeable, i.e.

$$\frac{d}{d\theta} E[h(\mathbf{x})|\theta] \quad = \quad \frac{d}{d\theta} \int_{x \in \mathcal{X}} h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x} = \int_{x \in \mathcal{X}} h(\mathbf{x}) \frac{\partial}{\partial \theta} f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x}$$

Then, a lower bound of $\mathrm{Var}[W(\mathbf{X})|\theta]$ is

$$\mathrm{Var}[W(\mathbf{X})|\theta] \geq \frac{[\tau'(\theta)]^2}{E\left[\{\frac{\partial}{\partial \theta} \log f_{\mathbf{X}}(\mathbf{X}|\theta)\}^2 |\theta\right]}$$

Recap : Cramer-Rao bound in iid case

Corollary 7.3.10

If $X_1, \cdots, X_n$ are iid samples from pdf/pmf $f_X(x|\theta)$, and the assumptions in the above Cramer-Rao theorem hold, then the lower-bound of $\text{Var}[W(\mathbf{X})|\theta]$ becomes

$$\text{Var}[W(\mathbf{X})|\theta] \geq \frac{[\tau'(\theta)]^2}{nE\left[\{\frac{\partial}{\partial\theta}\log f_X(X|\theta)\}^2|\theta\right]}$$

Recap : Score Function

Definition: Score or Score Function for X

$$
\begin{aligned}
X_1, \cdots, X_n &\overset{\text{i.i.d.}}{\sim} f_X(x|\theta) \\
S(X|\theta) &= \frac{\partial}{\partial\theta} \log f_X(X|\theta) \\
E\left[S(X|\theta)\right] &= 0 \\
S_n(X|\theta) &= \frac{\partial}{\partial\theta} \log f_{\mathbf{X}}(\mathbf{X}|\theta)
\end{aligned}
$$

Recap : Fisher Information Number

Definition: Fisher Information Number

$$
\begin{aligned}
I(\theta) &= E\left[\left\{\frac{\partial}{\partial\theta}\log f_X(X|\theta)\right\}^2\right] = E\left[S^2(X|\theta)\right] \\
I_n(\theta) &= E\left[\left\{\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right\}^2\right] \\
&= nE\left[\left\{\frac{\partial}{\partial\theta}\log f_X(X|\theta)\right\}^2\right] = nI(\theta)
\end{aligned}
$$

The bigger the information number, the more information we have about θ, the smaller bound on the variance of unbiased estimates.

Recap : Simplified Fisher Information

Lemma 7.3.11

If $f_X(x|\theta)$ satisfies the two interchangeability conditions

$$\frac{d}{d\theta}\int_{x\in\mathcal{X}} f_X(x|\theta)\,dx = \int_{x\in\mathcal{X}} \frac{\partial}{\partial\theta} f_X(x|\theta)\,dx$$

$$\frac{d}{d\theta}\int_{x\in\mathcal{X}} \frac{\partial}{\partial\theta} f_X(x|\theta)\,dx = \int_{x\in\mathcal{X}} \frac{\partial^2}{\partial\theta^2} f_X(x|\theta)\,dx$$

which are true for exponential family, then

$$I(\theta) = E\left[\left\{\frac{\partial}{\partial\theta}\log f_X(X|\theta)\right\}^2\right] = -E\left[\frac{\partial^2}{\partial\theta^2}\log f_X(X|\theta)\right]$$

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

$$I(\mu) = -E\left[\frac{\partial^2}{\partial \mu^2} \log f_X(X|\mu)\right]$$

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

$$\begin{aligned} I(\mu) &= -E\left[\frac{\partial^2}{\partial \mu^2} \log f_X(X|\mu)\right] \\ &= -E\left[\frac{\partial^2}{\partial \mu^2} \log\left\{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right)\right\}\right] \end{aligned}$$

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

$$
\begin{aligned}
I(\mu) &= -E\left[\frac{\partial^2}{\partial\mu^2}\log f_X(X|\mu)\right] \\
&= -E\left[\frac{\partial^2}{\partial\mu^2}\log\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right)\right\}\right] \\
&= -E\left[\frac{\partial^2}{\partial\mu^2}\left\{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X-\mu)^2}{2\sigma^2}\right\}\right]
\end{aligned}
$$

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

$$\begin{aligned} I(\mu) &= -E\left[\frac{\partial^2}{\partial\mu^2}\log f_X(X|\mu)\right] \\ &= -E\left[\frac{\partial^2}{\partial\mu^2}\log\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right)\right\}\right] \\ &= -E\left[\frac{\partial^2}{\partial\mu^2}\left\{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X-\mu)^2}{2\sigma^2}\right\}\right] \\ &= -E\left[\frac{\partial}{\partial\mu}\left\{\frac{2(X-\mu)}{2\sigma^2}\right\}\right] = \frac{1}{\sigma^2} \end{aligned}$$

The Cramer-Rao bound for μ is $[nI(\mu)]^{-1} = \frac{\sigma^2}{n}$

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

$$\begin{aligned}
I(\mu) &= -E\left[\frac{\partial^2}{\partial\mu^2}\log f_X(X|\mu)\right] \\
&= -E\left[\frac{\partial^2}{\partial\mu^2}\log\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right)\right\}\right] \\
&= -E\left[\frac{\partial^2}{\partial\mu^2}\left\{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X-\mu)^2}{2\sigma^2}\right\}\right] \\
&= -E\left[\frac{\partial}{\partial\mu}\left\{\frac{2(X-\mu)}{2\sigma^2}\right\}\right] = \frac{1}{\sigma^2}
\end{aligned}$$

The Cramer-Rao bound for μ is $[nI(\mu)]^{-1} = \frac{\sigma^2}{n} = \text{Var}(\overline{X})$.

Recap - Normal Distribution

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where σ^2 is known.

$$\begin{aligned} I(\mu) &= -E\left[\frac{\partial^2}{\partial \mu^2} \log f_X(X|\mu)\right] \\ &= -E\left[\frac{\partial^2}{\partial \mu^2} \log\left\{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right)\right\}\right] \\ &= -E\left[\frac{\partial^2}{\partial \mu^2}\left\{-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X-\mu)^2}{2\sigma^2}\right\}\right] \\ &= -E\left[\frac{\partial}{\partial \mu}\left\{\frac{2(X-\mu)}{2\sigma^2}\right\}\right] = \frac{1}{\sigma^2} \end{aligned}$$

The Cramer-Rao bound for μ is $[nI(\mu)]^{-1} = \frac{\sigma^2}{n} = \text{Var}(\overline{X})$. Therefore $\overline{X}$ attains the Cramer-Rao bound and thus the best unbiased estimator for μ.

Example of Cramer-Rao lower bound attainment

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Is $\overline{X}$ the best unbiased estimator of p? Does it attain the Cramer-Rao lower bound?

Example of Cramer-Rao lower bound attainment

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Is $\overline{X}$ the best unbiased estimator of p? Does it attain the Cramer-Rao lower bound?

Solution

$$E(\overline{X}) = p$$

Example of Cramer-Rao lower bound attainment

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Is $\overline{X}$ the best unbiased estimator of p? Does it attain the Cramer-Rao lower bound?

Solution

$$\begin{aligned} E(\overline{X}) &= p \\ \text{Var}(\overline{X}) &= \frac{1}{n}\text{Var}(X) = \frac{p(1-p)}{n} \end{aligned}$$

Example of Cramer-Rao lower bound attainment

Problem

$X_1, \cdots, X_n \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Is $\overline{X}$ the best unbiased estimator of p? Does it attain the Cramer-Rao lower bound?

Solution

$$\begin{aligned} E(\overline{X}) &= p \\ \text{Var}(\overline{X}) &= \frac{1}{n}\text{Var}(X) = \frac{p(1-p)}{n} \\ I(p) &= E\left[\left\{\frac{\partial}{\partial\theta}\log f_X(X|\theta)\right\}^2 \middle| p\right] \end{aligned}$$

Example of Cramer-Rao lower bound attainment

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Is $\overline{X}$ the best unbiased estimator of p? Does it attain the Cramer-Rao lower bound?

Solution

$$\begin{aligned}
E(\overline{X}) &= p \\
\text{Var}(\overline{X}) &= \frac{1}{n}\text{Var}(X) = \frac{p(1-p)}{n} \\
I(p) &= E\left[\left\{ \frac{\partial}{\partial\theta} \log f_X(X|\theta) \right\}^2 \middle| p \right] \\
&= -E\left[\frac{\partial^2}{\partial\theta^2} \log f_X(X|\theta) | p \right]
\end{aligned}$$

Solution (cont'd)

$$f_X(x|\theta) = p^x(1-p)^{1-x}$$

Solution (cont'd)

$$\begin{aligned} f_X(x|\theta) &= p^x(1-p)^{1-x} \\ \log f_X(x|\theta) &= x\log p + (1-x)\log(1-p) \end{aligned}$$

Solution (cont'd)

$$\begin{aligned} f_X(x|\theta) &= p^x(1-p)^{1-x} \\ \log f_X(x|\theta) &= x\log p + (1-x)\log(1-p) \\ \frac{\partial}{\partial p}\log f_X(x|p) &= \frac{x}{p} - \frac{1-x}{1-p} \end{aligned}$$

Solution (cont'd)

$$
\begin{aligned}
f_X(x|\theta) &= p^x(1-p)^{1-x} \\
\log f_X(x|\theta) &= x \log p + (1-x)\log(1-p) \\
\frac{\partial}{\partial p} \log f_X(x|p) &= \frac{x}{p} - \frac{1-x}{1-p} \\
\frac{\partial^2}{\partial p^2} \log f_X(x|p) &= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}
\end{aligned}
$$

Solution (cont'd)

$$
\begin{aligned}
f_X(x|\theta) &= p^x(1-p)^{1-x} \\
\log f_X(x|\theta) &= x\log p + (1-x)\log(1-p) \\
\frac{\partial}{\partial p}\log f_X(x|p) &= \frac{x}{p} - \frac{1-x}{1-p} \\
\frac{\partial^2}{\partial p^2}\log f_X(x|p) &= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2} \\
I(p) &= -E\left[-\frac{X}{p^2} - \frac{1-X}{(1-p)^2}|p\right]
\end{aligned}
$$

Solution (cont'd)

$$
\begin{aligned}
f_X(x|\theta) &= p^x(1-p)^{1-x} \\
\log f_X(x|\theta) &= x\log p + (1-x)\log(1-p) \\
\frac{\partial}{\partial p}\log f_X(x|p) &= \frac{x}{p} - \frac{1-x}{1-p} \\
\frac{\partial^2}{\partial p^2}\log f_X(x|p) &= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2} \\
I(p) &= -E\left[-\frac{X}{p^2} - \frac{1-X}{(1-p)^2}|p\right] \\
&= \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}
\end{aligned}
$$

Solution (cont'd)

$$\begin{aligned}
f_X(x|\theta) &= p^x(1-p)^{1-x} \\
\log f_X(x|\theta) &= x\log p + (1-x)\log(1-p) \\
\frac{\partial}{\partial p}\log f_X(x|p) &= \frac{x}{p} - \frac{1-x}{1-p} \\
\frac{\partial^2}{\partial p^2}\log f_X(x|p) &= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2} \\
I(p) &= -E\left[-\frac{X}{p^2} - \frac{1-X}{(1-p)^2}|p\right] \\
&= \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}
\end{aligned}$$

Therefore, the Cramer-Rao bound is $\frac{1}{nI(p)} = \frac{p(1-p)}{n} = \mathrm{Var}\overline{X}$, and $\overline{X}$ attains the Cramer-Rao lower bound, and it is the UMVUE.

Regularity condition for Cramer-Rao Theorem

$$\frac{d}{d\theta}\int_{x\in\mathcal{X}} h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x} = \int_{x\in\mathcal{X}} h(\mathbf{x}) \frac{\partial}{\partial\theta} f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x}$$

Regularity condition for Cramer-Rao Theorem

$$\frac{d}{d\theta}\int_{x\in\mathcal{X}} h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x} = \int_{x\in\mathcal{X}} h(\mathbf{x}) \frac{\partial}{\partial\theta} f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x}$$

- This regularity condition holds for exponential family.

Regularity condition for Cramer-Rao Theorem

$$\frac{d}{d\theta}\int_{x\in\mathcal{X}} h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x} = \int_{x\in\mathcal{X}} h(\mathbf{x}) \frac{\partial}{\partial\theta} f_{\mathbf{X}}(\mathbf{x}|\theta)\, d\mathbf{x}$$

- This regularity condition holds for exponential family.
- How about non-exponential family, such as $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}(0,\theta)$?

Using Leibnitz's Rule

Leibnitz's Rule

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x|\theta)\,dx = f(b(\theta)|\theta)\,b'(\theta) - f(a(\theta)|\theta)\,a'(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x|\theta)\,dx$$

Using Leibnitz's Rule

Leibnitz's Rule

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x|\theta)\,dx = f(b(\theta)|\theta)\,b'(\theta) - f(a(\theta)|\theta)\,a'(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x|\theta)\,dx$$

Applying to Uniform Distribution

$$f_X(x|\theta) = 1/\theta$$

Using Leibnitz's Rule

Leibnitz's Rule

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x|\theta)\,dx = f(b(\theta)|\theta)\,b'(\theta) - f(a(\theta)|\theta)\,a'(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x|\theta)\,dx$$

Applying to Uniform Distribution

$$f_X(x|\theta) = 1/\theta$$

$$\frac{d}{d\theta}\int_0^{\theta} h(x)\left(\frac{1}{\theta}\right)dx = \frac{h(\theta)}{\theta}\frac{d\theta}{d\theta} - h(0)f_X(0|\theta)\frac{d0}{d\theta} + \int_0^{\theta}\frac{\partial}{\partial\theta}h(x)\left(\frac{1}{\theta}\right)dx$$

Using Leibnitz's Rule

Leibnitz's Rule

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x|\theta)\,dx = f(b(\theta)|\theta)\,b'(\theta) - f(a(\theta)|\theta)\,a'(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x|\theta)\,dx$$

Applying to Uniform Distribution

$$\begin{aligned} f_X(x|\theta) &= 1/\theta \\ \frac{d}{d\theta}\int_0^{\theta} h(x)\left(\frac{1}{\theta}\right)dx &= \frac{h(\theta)}{\theta}\frac{d\theta}{d\theta} - h(0)f_X(0|\theta)\frac{d0}{d\theta} + \int_0^{\theta}\frac{\partial}{\partial\theta}h(x)\left(\frac{1}{\theta}\right)dx \\ &\neq \int_0^{\theta}\frac{\partial}{\partial\theta}h(x)\left(\frac{1}{\theta}\right)dx \end{aligned}$$

Using Leibnitz's Rule

Leibnitz's Rule

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x|\theta)\,dx = f(b(\theta)|\theta)\,b'(\theta) - f(a(\theta)|\theta)\,a'(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x|\theta)\,dx$$

Applying to Uniform Distribution

$$\begin{aligned} f_X(x|\theta) &= 1/\theta \\ \frac{d}{d\theta}\int_0^{\theta} h(x)\left(\frac{1}{\theta}\right) dx &= \frac{h(\theta)}{\theta}\frac{d\theta}{d\theta} - h(0) f_X(0|\theta)\frac{d0}{d\theta} + \int_0^{\theta} \frac{\partial}{\partial\theta} h(x)\left(\frac{1}{\theta}\right) dx \\ &\neq \int_0^{\theta} \frac{\partial}{\partial\theta} h(x)\left(\frac{1}{\theta}\right) dx \end{aligned}$$

The interchangeability condition is not satisfied.

Solving the Uniform Distribution Example

If $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}(0, \theta)$, the unbiased estimator of θ is

$$T(\mathbf{X}) = \frac{n+1}{n} X_{(n)}$$

Solving the Uniform Distribution Example

If $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}(0, \theta)$, the unbiased estimator of θ is

$$
\begin{aligned}
T(\mathbf{X}) &= \frac{n+1}{n} X_{(n)} \\
E\left[\frac{n+1}{n} X_{(n)}\right] &= \theta
\end{aligned}
$$

Solving the Uniform Distribution Example

If $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}(0, \theta)$, the unbiased estimator of θ is

$$
\begin{aligned}
T(\mathbf{X}) &= \frac{n+1}{n} X_{(n)} \\
E\left[\frac{n+1}{n} X_{(n)}\right] &= \theta \\
\text{Var}\left[\frac{n+1}{n} X_{(n)}\right] &= \frac{1}{n(n+2)}\theta^2 < \frac{\theta^2}{n}
\end{aligned}
$$

Solving the Uniform Distribution Example

If $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}(0, \theta)$, the unbiased estimator of θ is

$$\begin{aligned} T(\mathbf{X}) &= \frac{n+1}{n} X_{(n)} \\ E\left[\frac{n+1}{n} X_{(n)}\right] &= \theta \\ \text{Var}\left[\frac{n+1}{n} X_{(n)}\right] &= \frac{1}{n(n+2)}\theta^2 < \frac{\theta^2}{n} \end{aligned}$$

The Cramer-Rao lower bound (if interchangeability condition was met) is $\frac{1}{nI(\theta)} = \frac{\theta^2}{n}$.

When is the Cramer-Rao Lower Bound Attainable?

It is possible that the value of Cramer-Rao bound may be strictly smaller than the variance of any unbiased estimator

When is the Cramer-Rao Lower Bound Attainable?

It is possible that the value of Cramer-Rao bound may be strictly smaller than the variance of any unbiased estimator

Corollary 7.3.15 : Attainment of Cramer-Rao Bound

Let $X_1, \cdots, X_n$ be iid with pdf/pmf $f_X(x|\theta)$, where $f_X(x|\theta)$ satisfies the assumptions of the Cramer-Rao Theorem.

When is the Cramer-Rao Lower Bound Attainable?

It is possible that the value of Cramer-Rao bound may be strictly smaller than the variance of any unbiased estimator

Corollary 7.3.15 : Attainment of Cramer-Rao Bound

Let $X_1, \cdots, X_n$ be iid with pdf/pmf $f_X(x|\theta)$, where $f_X(x|\theta)$ satisfies the assumptions of the Cramer-Rao Theorem. Let $L(\theta|\mathbf{x}) = \prod_{i=1}^n f_X(x_i|\theta)$ denote the likelihood function. If $W(\mathbf{X})$ is unbiased for $\tau(\theta)$, then $W(\mathbf{X})$ attains the Cramer-Rao lower bound if and only if

When is the Cramer-Rao Lower Bound Attainable?

It is possible that the value of Cramer-Rao bound may be strictly smaller than the variance of any unbiased estimator

Corollary 7.3.15 : Attainment of Cramer-Rao Bound

Let $X_1, \cdots, X_n$ be iid with pdf/pmf $f_X(x|\theta)$, where $f_X(x|\theta)$ satisfies the assumptions of the Cramer-Rao Theorem. Let $L(\theta|\mathbf{x}) = \prod_{i=1}^{n} f_X(x_i|\theta)$ denote the likelihood function. If $W(\mathbf{X})$ is unbiased for $\tau(\theta)$, then $W(\mathbf{X})$ attains the Cramer-Rao lower bound if and only if

$$\frac{\partial}{\partial\theta} \log L(\theta|\mathbf{x}) = S_n(\mathbf{x}|\theta) = a(\theta)[W(\mathbf{X}) - \tau(\theta)]$$

for some function $a(\theta)$.

Proof of Corollary 7.3.15

We used Cauchy-Schwarz inequality to prove that

$$\left[\mathrm{Cov}\{W(\mathbf{X}), \frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\}\right]^2 \leq \mathrm{Var}[W(\mathbf{X})]\mathrm{Var}\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right]$$

Proof of Corollary 7.3.15

We used Cauchy-Schwarz inequality to prove that

$$\left[\mathrm{Cov}\{W(\mathbf{X}), \frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\}\right]^2 \leq \mathrm{Var}[W(\mathbf{X})]\mathrm{Var}\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right]$$

In Cauchy-Schwarz inequality, the equality satisfies if and only if there is a linear relationship between the two variables, that is

$$\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{x}|\theta) = \frac{\partial}{\partial\theta}\log L(\theta|\mathbf{x}) = a(\theta)W(\mathbf{x}) + b(\theta)$$

Proof of Corollary 7.3.15 (cont'd)

$$E\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right] = E[S_n(\mathbf{X}|\theta)] = 0$$

Proof of Corollary 7.3.15 (cont'd)

$$\begin{aligned} E\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right] &= E[S_n(\mathbf{X}|\theta)] = 0 \\ E[a(\theta)W(\mathbf{X}) + b(\theta)] &= 0 \end{aligned}$$

Proof of Corollary 7.3.15 (cont'd)

$$
\begin{aligned}
E\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right] &= E[S_n(\mathbf{X}|\theta)] = 0 \\
E[a(\theta)W(\mathbf{X}) + b(\theta)] &= 0 \\
a(\theta)E[W(\mathbf{X})] + b(\theta) &= 0
\end{aligned}
$$

Proof of Corollary 7.3.15 (cont'd)

$$\begin{aligned}
E\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right] &= E[S_n(\mathbf{X}|\theta)] = 0 \\
E[a(\theta)W(\mathbf{X}) + b(\theta)] &= 0 \\
a(\theta)E[W(\mathbf{X})] + b(\theta) &= 0 \\
a(\theta)\tau(\theta) + b(\theta) &= 0
\end{aligned}$$

Proof of Corollary 7.3.15 (cont'd)

$$\begin{aligned}
E\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right] &= E[S_n(\mathbf{X}|\theta)] = 0 \\
E[a(\theta)W(\mathbf{X}) + b(\theta)] &= 0 \\
a(\theta)E[W(\mathbf{X})] + b(\theta) &= 0 \\
a(\theta)\tau(\theta) + b(\theta) &= 0 \\
b(\theta) &= -a(\theta)\tau(\theta)
\end{aligned}$$

Proof of Corollary 7.3.15 (cont'd)

$$
\begin{aligned}
E\left[\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\right] &= E[S_n(\mathbf{X}|\theta)] = 0 \\
E[a(\theta)W(\mathbf{X}) + b(\theta)] &= 0 \\
a(\theta)E[W(\mathbf{X})] + b(\theta) &= 0 \\
a(\theta)\tau(\theta) + b(\theta) &= 0 \\
b(\theta) &= -a(\theta)\tau(\theta) \\
\frac{\partial}{\partial\theta}\log L(\theta|\mathbf{x}) &= a(\theta)W(\mathbf{x}) - a(\theta)\tau(\theta) = a(\theta)[W(\mathbf{x}) - \tau(\theta)]
\end{aligned}
$$

Revisiting the Bernoulli Example

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Is $\overline{X}$ the best unbiased estimator of p? Does it attain the Cramer-Rao lower bound?

Method Using Corollary 7.3.15

$$L(p|\mathbf{x}) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

Method Using Corollary 7.3.15

$$
\begin{aligned}
L(p|\mathbf{x}) &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
\log L(p|\mathbf{x}) &= \log \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}
\end{aligned}
$$

Method Using Corollary 7.3.15

$$
\begin{aligned}
L(p|\mathbf{x}) &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
\log L(p|\mathbf{x}) &= \log \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= \sum_{i=1}^{n} \log[p^{x_i}(1-p)^{1-x_i}]
\end{aligned}
$$

Method Using Corollary 7.3.15

$$
\begin{aligned}
L(p|\mathbf{x}) &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
\log L(p|\mathbf{x}) &= \log \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= \sum_{i=1}^{n} \log[p^{x_i}(1-p)^{1-x_i}] \\
&= \sum_{i=1}^{n} [x_i \log p + (1-x_i)\log(1-p)]
\end{aligned}
$$

Method Using Corollary 7.3.15

$$
\begin{aligned}
L(p|\mathbf{x}) &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
\log L(p|\mathbf{x}) &= \log \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= \sum_{i=1}^{n} \log[p^{x_i}(1-p)^{1-x_i}] \\
&= \sum_{i=1}^{n} [x_i \log p + (1-x_i)\log(1-p)] \\
&= \log p \sum_{i=1}^{n} x_i + \log(1-p)(n - \sum_{i=1}^{n} x_i)
\end{aligned}
$$

Method Using Corollary 7.3.15 (cont'd)

$$\frac{\partial}{\partial p}\log L(p|\mathbf{x}) = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p}$$

Method Using Corollary 7.3.15 (cont'd)

$$\begin{aligned}\frac{\partial}{\partial p}\log L(p|\mathbf{x}) &= \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} \\ &= \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p}\end{aligned}$$

Method Using Corollary 7.3.15 (cont'd)

$$\begin{aligned}
\frac{\partial}{\partial p}\log L(p|\mathbf{x}) &= \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} \\
&= \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p} \\
&= \frac{(1-p)n\overline{x} - np(1-\overline{x})}{p(1-p)}
\end{aligned}$$

Method Using Corollary 7.3.15 (cont'd)

$$\begin{aligned}
\frac{\partial}{\partial p} \log L(p|\mathbf{x}) &= \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} \\
&= \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p} \\
&= \frac{(1-p)n\overline{x} - np(1-\overline{x})}{p(1-p)} \\
&= \frac{n(\overline{x}-p)}{p(1-p)}
\end{aligned}$$

Method Using Corollary 7.3.15 (cont'd)

$$\begin{aligned}
\frac{\partial}{\partial p}\log L(p|\mathbf{x}) &= \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n-\sum_{i=1}^{n} x_i}{1-p} \\
&= \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p} \\
&= \frac{(1-p)n\overline{x} - np(1-\overline{x})}{p(1-p)} \\
&= \frac{n(\overline{x}-p)}{p(1-p)} \\
&= a(p)[W(\mathbf{x}) - \tau(p)]
\end{aligned}$$

Method Using Corollary 7.3.15 (cont'd)

$$\begin{aligned}
\frac{\partial}{\partial p}\log L(p|\mathbf{x}) &= \frac{\sum_{i=1}^n x_i}{p} - \frac{n-\sum_{i=1}^n x_i}{1-p} \\
&= \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p} \\
&= \frac{(1-p)n\overline{x} - np(1-\overline{x})}{p(1-p)} \\
&= \frac{n(\overline{x}-p)}{p(1-p)} \\
&= a(p)[W(\mathbf{x}) - \tau(p)]
\end{aligned}$$

where $a(p) = \frac{n}{p(1-p)}$, $W(\mathbf{x}) = \overline{x}$, $\tau(p) = p$.

Method Using Corollary 7.3.15 (cont'd)

$$\begin{aligned}
\frac{\partial}{\partial p} \log L(p|\mathbf{x}) &= \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} \\
&= \frac{n\overline{x}}{p} - \frac{n(1-\overline{x})}{1-p} \\
&= \frac{(1-p)n\overline{x} - np(1-\overline{x})}{p(1-p)} \\
&= \frac{n(\overline{x}-p)}{p(1-p)} \\
&= a(p)[W(\mathbf{x}) - \tau(p)]
\end{aligned}$$

where $a(p) = \frac{n}{p(1-p)}$, $W(\mathbf{x}) = \overline{x}$, $\tau(p) = p$. Therefore, $\overline{X}$ is the best unbiased estimator for p and attains the Cramer-Rao lower bound.

Normal distribution example

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Consider estimating σ^2, assuming μ is known. Is Cramer-Rao bound attainable?

Normal distribution example

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Consider estimating σ^2, assuming μ is known. Is Cramer-Rao bound attainable?

Solution

$$I(\sigma^2) = -E\left[\frac{\partial^2}{\partial(\sigma^2)^2}\log f_X(X|\mu,\sigma)|p\right]$$

Normal distribution example

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Consider estimating σ^2, assuming μ is known. Is Cramer-Rao bound attainable?

Solution

$$\begin{aligned} I(\sigma^2) &= -E\left[\frac{\partial^2}{\partial(\sigma^2)^2}\log f_X(X|\mu,\sigma)|p\right] \\ f(x|\mu,\sigma^2) &= \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \end{aligned}$$

Normal distribution example

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Consider estimating σ^2, assuming μ is known. Is Cramer-Rao bound attainable?

Solution

$$
\begin{aligned}
I(\sigma^2) &= -E\left[\frac{\partial^2}{\partial(\sigma^2)^2}\log f_X(X|\mu,\sigma)|p\right] \\
f(x|\mu,\sigma^2) &= \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \\
\log f(x|\mu,\sigma^2) &= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}
\end{aligned}
$$

Normal distribution example

Problem

$X_1, \cdots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Consider estimating σ^2, assuming μ is known. Is Cramer-Rao bound attainable?

Solution

$$
\begin{aligned}
I(\sigma^2) &= -E\left[\frac{\partial^2}{\partial(\sigma^2)^2} \log f_X(X|\mu,\sigma)|p\right] \\
f(x|\mu,\sigma^2) &= \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \\
\log f(x|\mu,\sigma^2) &= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2} \\
\frac{\partial}{\partial(\sigma^2)} \log f(x|\mu,\sigma^2) &= -\frac{1}{2}\frac{1}{\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2}
\end{aligned}
$$

Solution (cont'd)

$$\frac{\partial^2}{\partial(\sigma^2)^2} \log f(x|\mu, \sigma^2) = \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3}$$

Solution (cont'd)

$$\begin{aligned}
\frac{\partial^2}{\partial(\sigma^2)^2} \log f(x|\mu,\sigma^2) &= \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3} \\
I(\sigma^2) &= -E\left[\frac{1}{2\sigma^4} - \frac{2(x-\mu)^2}{2\sigma^6}\right]
\end{aligned}$$

Solution (cont'd)

$$\begin{aligned}
\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x|\mu,\sigma^2) &= \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3} \\
I(\sigma^2) &= -E\left[\frac{1}{2\sigma^4} - \frac{2(x-\mu)^2}{2\sigma^6}\right] \\
&= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}E[(x-\mu)^2] = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}\sigma^2 = \frac{1}{2\sigma^4}
\end{aligned}$$

Solution (cont'd)

$$\begin{aligned}
\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x|\mu,\sigma^2) &= \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3} \\
I(\sigma^2) &= -E\left[\frac{1}{2\sigma^4} - \frac{2(x-\mu)^2}{2\sigma^6}\right] \\
&= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}E[(x-\mu)^2] = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}\sigma^2 = \frac{1}{2\sigma^4}
\end{aligned}$$

Cramer-Rao lower bound is $\frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n}$.

Solution (cont'd)

$$\frac{\partial^2}{\partial(\sigma^2)^2} \log f(x|\mu, \sigma^2) = \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3}$$

$$I(\sigma^2) = -E\left[\frac{1}{2\sigma^4} - \frac{2(x-\mu)^2}{2\sigma^6}\right]$$

$$= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}E[(x-\mu)^2] = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}\sigma^2 = \frac{1}{2\sigma^4}$$

Cramer-Rao lower bound is $\frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n}$. The unbiased estimator of $\hat{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$, gives

Solution (cont'd)

$$\begin{aligned}\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x|\mu,\sigma^2) &= \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3} \\ I(\sigma^2) &= -E\left[\frac{1}{2\sigma^4} - \frac{2(x-\mu)^2}{2\sigma^6}\right] \\ &= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}E[(x-\mu)^2] = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}\sigma^2 = \frac{1}{2\sigma^4}\end{aligned}$$

Cramer-Rao lower bound is $\frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n}$. The unbiased estimator of $\hat{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$, gives

$$\mathrm{Var}(\hat{\sigma^2}) = \frac{2\sigma^4}{n-1} > \frac{2\sigma^4}{n}$$

Solution (cont'd)

$$\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x|\mu,\sigma^2) = \frac{1}{2}\frac{1}{(\sigma^2)^2} - \frac{2(x-\mu)^2}{2(\sigma^2)^3}$$

$$I(\sigma^2) = -E\left[\frac{1}{2\sigma^4} - \frac{2(x-\mu)^2}{2\sigma^6}\right]$$

$$= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}E[(x-\mu)^2] = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}\sigma^2 = \frac{1}{2\sigma^4}$$

Cramer-Rao lower bound is $\frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n}$. The unbiased estimator of $\hat{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$, gives

$$\mathrm{Var}(\hat{\sigma^2}) = \frac{2\sigma^4}{n-1} > \frac{2\sigma^4}{n}$$

So, $\hat{\sigma^2}$ does not attain the Cramer-Rao lower-bound.

Is Cramer-Rao lower-bound for σ^2 attainable?

$$L(\sigma^2|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

Is Cramer-Rao lower-bound for σ^2 attainable?

$$L(\sigma^2|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\log L(\sigma^2|\mathbf{x}) = -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}$$

Is Cramer-Rao lower-bound for σ^2 attainable?

$$L(\sigma^2|\mathbf{x}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\log L(\sigma^2|\mathbf{x}) = -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial \log L(\sigma^2|\mathbf{x})}{\partial \sigma^2} = -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} + \sum_{i=1}^n \frac{(x_i-\mu)^2}{2(\sigma^2)^2}$$

Is Cramer-Rao lower-bound for σ^2 attainable?

$$
\begin{aligned}
L(\sigma^2|\mathbf{x}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] \\
\log L(\sigma^2|\mathbf{x}) &= -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \\
\frac{\partial \log L(\sigma^2|\mathbf{x})}{\partial \sigma^2} &= -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2(\sigma^2)^2} \\
&= -\frac{n}{2\sigma^2} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^4}
\end{aligned}
$$

Is Cramer-Rao lower-bound for σ^2 attainable?

$$
\begin{aligned}
L(\sigma^2|\mathbf{x}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] \\
\log L(\sigma^2|\mathbf{x}) &= -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2} \\
\frac{\partial \log L(\sigma^2|\mathbf{x})}{\partial\sigma^2} &= -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2(\sigma^2)^2} \\
&= -\frac{n}{2\sigma^2} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^4} \\
&= \frac{n}{2\sigma^4}\left(\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} - \sigma^2\right)
\end{aligned}
$$

Is Cramer-Rao lower-bound for σ^2 attainable?

$$\begin{aligned}
L(\sigma^2|\mathbf{x}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] \\
\log L(\sigma^2|\mathbf{x}) &= -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2} \\
\frac{\partial \log L(\sigma^2|\mathbf{x})}{\partial\sigma^2} &= -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2(\sigma^2)^2} \\
&= -\frac{n}{2\sigma^2} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^4} \\
&= \frac{n}{2\sigma^4}\left(\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} - \sigma^2\right) \\
&= a(\sigma^2)(W(\mathbf{x}) - \sigma^2)
\end{aligned}$$

Is Cramer-Rao lower-bound for σ^2 attainable? (cont'd)

Therefore,

1. If μ is known, the best unbiased estimator for σ^2 is $\sum_{i=1}^n (x_i - \mu)^2/n$, and it attains the Cramer-Rao lower bound, i.e.

Is Cramer-Rao lower-bound for σ^2 attainable? (cont'd)

Therefore,

1. If μ is known, the best unbiased estimator for σ^2 is $\sum_{i=1}^{n}(x_i - \mu)^2/n$, and it attains the Cramer-Rao lower bound, i.e.

$$\mathrm{Var}\left[\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n}\right] = \frac{2\sigma^4}{n}$$

Is Cramer-Rao lower-bound for σ^2 attainable? (cont'd)

Therefore,

1. If μ is known, the best unbiased estimator for σ^2 is $\sum_{i=1}^n (x_i - \mu)^2/n$, and it attains the Cramer-Rao lower bound, i.e.

$$\text{Var}\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}\right] = \frac{2\sigma^4}{n}$$

2. If μ is not known, the Cramer-Rao lower-bound cannot be attained.

Is Cramer-Rao lower-bound for σ^2 attainable? (cont'd)

Therefore,

1. If μ is known, the best unbiased estimator for σ^2 is $\sum_{i=1}^n (x_i - \mu)^2/n$, and it attains the Cramer-Rao lower bound, i.e.

$$\text{Var}\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}\right] = \frac{2\sigma^4}{n}$$

2. If μ is not known, the Cramer-Rao lower-bound cannot be attained.

At this point, we do not know if $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$ is the best unbiased estimator for σ^2 or not.

Summary

Today : Cramero-Rao Theorem

- Recap of Cramer-Rao Theorem and Corollary
- Examples with Simple Distributions
- Regularity Condition
- Attainability

Summary

Today : Cramero-Rao Theorem

- Recap of Cramer-Rao Theorem and Corollary
- Examples with Simple Distributions
- Regularity Condition
- Attainability

Next Lecture

- Rao-Blackwell Theorem