# IE605: Engineering Statistics

## Lecture 13

Manjesh K. Hanawal

Previous Lecture:

- Evaluating Estimators
- Cramer's Rao inequality

This Lecture:

- Cramer's Rao inequality
- Information inequality
- Fischer Information
- Hypothesis testing
- Likelihood Ratio Test (LRT)

# Cramer-Rao's Bound

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be a sample with pdf $f(\boldsymbol{x}|\theta)$, and let $W(\boldsymbol{X})$ be any estimator such that $Var_\theta W(\boldsymbol{X}) < \infty$. Then

$$Var_\theta W(\boldsymbol{X}) \geq \frac{\left(\frac{d}{d\theta}\mathbb{E}_\theta W(\boldsymbol{X})\right)^2}{\mathbb{E}_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)^2\right)}$$

provided

$$\mathbb{E}_\theta W(\boldsymbol{X}) = \int_{\boldsymbol{x}} \frac{\partial}{\partial\theta}[W(\boldsymbol{x})f(\boldsymbol{x})]d\boldsymbol{x}.$$

# Proof of Cramer-Rao's Bound

- $[Cov(X,Y)]^2 \leq Var(X)Var(Y) \implies Var(X) \geq \frac{[Cov(X,Y)]^2}{Var(Y)}$
- $$\frac{d}{d\theta}\mathbb{E}_\theta W(\boldsymbol{X}) = \int_{\boldsymbol{x}} \frac{\partial}{\partial\theta}[W(\boldsymbol{x})f(\boldsymbol{x})]d\boldsymbol{x}$$
- $$\mathbb{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)$$
- $$Cov\left(W(\boldsymbol{X}), \frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right) = \mathbb{E}_\theta\left(W(\boldsymbol{X})\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)$$

# Cramer-Rao Bound for iid case

$$Var_\theta W(\boldsymbol{X}) \geq \frac{\left(\frac{d}{d\theta}\mathbb{E}_\theta W(\boldsymbol{X})\right)^2}{\mathbb{E}_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)^2\right)}$$
$$= \frac{\left(\frac{d}{d\theta}\mathbb{E}_\theta W(\boldsymbol{X})\right)^2}{n\mathbb{E}_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2\right)}$$

$$\mathbb{E}_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)^2\right) = \mathbb{E}_\theta\left(\left(\sum_i \frac{\partial}{\partial\theta}\log f(X_i|\theta)\right)^2\right)$$

# Fisher Information and Information Inequality

The quantity $\mathbb{E}_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)^2\right)$ is called **Information number** or **Fisher Information of sample**

- Larger the information number, smaller is the bound in the Cramer-Rao's bound
- Larger the information number, we have more information about about $\theta$.
- Cramer-Rao bound is also called as Information inequality.

# Hypotesis Testing

**Defintion:** A Hypothesis is a statement about a population parameter

**Defintion:** Two complementary hypothesis in a hypothesis testing are called *null hypothesis* and *alternate hypothesis*, denoted as $H_0$ and $H_1$, respectively.

General form of hypothesis testing

$$H_0 : \theta \in \Theta_0 \quad \text{null hypotesis}$$
$$H_1 : \theta \in \Theta_0^c \quad \text{null hypotesis}$$

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$

for some threshold $\theta_0$.

# Hypothesis Testing contd..

Hypothesis testing procedure or hypotesis test is rule that prescribes

1. For which sample values the **decision** is made to accept $H_0$ as true
2. For which sample values $H_0$ is rejected and $H_1$ is accepted as true

# Methods of tests: Likelihood Ratio Test (LRT)

Likelihood ration test statistic for testing $H_0 : \theta \in \Theta_0$ and $H_1 : \theta \in \Theta_0^c$ is

$$\lambda(\boldsymbol{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta|\boldsymbol{x})}{\sup_{\theta \in \{\Theta_0, \Theta_0^c\}} L(\theta|\boldsymbol{x})}$$

A Likelihood Ration Test (LRT) is any any test that has the reject region of the form $\{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq c\}$ for some $c \in (0, 1)$.

# Example:

# Sufficient Statistics and LRT

Let $T(\boldsymbol{X})$ be a sufficient statistic of sample $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ drawn from population $f(\boldsymbol{x}|\theta)$

$$\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{\sup_{\theta \in \Theta_0} L(\theta|\boldsymbol{x})}{\sup_{\theta \in \Theta} L(\theta|\boldsymbol{x})} \\
&= \frac{\sup_{\theta \in \Theta_0} f(\boldsymbol{x}|\theta)}{\sup_{\theta \in \Theta} f(\boldsymbol{x}|\theta)} \\
&= \frac{\sup_{\theta \in \Theta_0} h(\boldsymbol{x}) g(T(\boldsymbol{x})|\theta))}{\sup_{\theta \in \Theta} h(\boldsymbol{x}) g(T(\boldsymbol{x})|\theta))} \\
&= \frac{\sup_{\theta \in \Theta_0} g(T(\boldsymbol{x})|\theta))}{\sup_{\theta \in \Theta} g(T(\boldsymbol{x})|\theta))} \\
&= \frac{\sup_{\theta \in \Theta_0} L^*(\theta|T(\boldsymbol{x}))}{\sup_{\theta \in \Theta} L^*(\theta|T(\boldsymbol{x}))} = \lambda(T(\boldsymbol{x}))
\end{aligned}$$

# Bayes Tests:

In Baye's method, we have prior probability $P(\theta)$ and posterior probability $\pi(\theta|\boldsymbol{x})$ after observing sample $\boldsymbol{x}$.

- $\pi(\theta|\boldsymbol{x})$ can be used to find

$$\Pr(\theta \in \Theta_0|\boldsymbol{x}) \quad \text{and} \quad \Pr(\theta \in \Theta_0^c|\boldsymbol{x})$$

- We can assign

$$\Pr(\theta \in \Theta_0|\boldsymbol{x}) = \Pr(H_0 \text{ is true } |\boldsymbol{x})$$

$$\Pr(\theta \in \Theta_0^c|\boldsymbol{x}) = \Pr(H_1 \text{ is true } |\boldsymbol{x})$$

- Decision criteria

$$\begin{aligned} H_0 \quad & \text{is true if } \Pr(\theta \in \Theta_0|\boldsymbol{x}) \geq \Pr(\theta \in \Theta_0^c|\boldsymbol{x}) \\ H_1 \quad & \text{is true otherwise} \end{aligned}$$

Rejection set $\{\boldsymbol{x} : \Pr(\theta \in \Theta_0^c|\boldsymbol{x}) \geq 1/2\}$.