Asymptotic Normality

Asymptotic Efficiency

Hyun Min Kang

March 19th, 2013
Last Lecture

- What is a Bayes Risk?
Last Lecture

- What is a Bayes Risk?
- What is the Bayes rule Estimator minimizing squared error loss?
Recap

Asymptotic Normality

Asymptotic Efficiency

Summary

Last Lecture

- What is a Bayes Risk?
- What is the Bayes rule Estimator minimizing squared error loss?
- What is the Bayes rule Estimator minimizing absolute error loss?
Last Lecture

- What is a Bayes Risk?
- What is the Bayes rule Estimator minimizing squared error loss?
- What is the Bayes rule Estimator minimizing absolute error loss?
- What are the tools for proving a point estimator is consistent?
What is a Bayes Risk?
What is the Bayes rule Estimator minimizing squared error loss?
What is the Bayes rule Estimator minimizing absolute error loss?
What are the tools for proving a point estimator is consistent?
Can a biased estimator be consistent?
Bayes Estimator based on absolute error loss

Suppose that $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$. 
Bayes Estimator based on absolute error loss

Suppose that $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$. The posterior expected loss is

$$E[L(\theta, \hat{\theta}(x))] = \int_{\Omega} |\theta - \hat{\theta}(x)| \pi(\theta|x) d\theta$$
Bayes Estimator based on absolute error loss

Suppose that $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$. The posterior expected loss is

$$E[L(\theta, \hat{\theta}(x))] = \int_{\Omega} |\theta - \hat{\theta}(x)| \pi(\theta | x) d\theta$$

$$= E[|\theta - \hat{\theta}| | X = x]$$
Bayes Estimator based on absolute error loss

Suppose that $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$. The posterior expected loss is

$$
E[L(\theta, \hat{\theta}(x))] = \int_{\Omega} |\theta - \hat{\theta}(x)| \pi(\theta|x) d\theta
$$

$$
= E[|\theta - \hat{\theta}|\big| x] = \int_{-\infty}^{\hat{\theta}} -(\theta - \hat{\theta}) \pi(\theta|x) d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta}) \pi(\theta|x) d\theta
$$
Bayes Estimator based on absolute error loss

Suppose that \( L(\theta, \hat{\theta}) = |\theta - \hat{\theta}| \). The posterior expected loss is

\[
\begin{align*}
\mathbb{E}[L(\theta, \hat{\theta}(x))] &= \int_{\Omega} |\theta - \hat{\theta}(x)| \pi(\theta|x) \, d\theta \\
&= \mathbb{E}[|\theta - \hat{\theta}| | X = x] \\
&= \int_{-\infty}^{\hat{\theta}} - (\theta - \hat{\theta}) \pi(\theta|x) \, d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta}) \pi(\theta|x) \, d\theta
\end{align*}
\]

\[
\frac{\partial}{\partial \hat{\theta}} \mathbb{E}[L(\theta, \hat{\theta}(x))] = \int_{-\infty}^{\hat{\theta}} \pi(\theta|x) \, d\theta - \int_{\hat{\theta}}^{\infty} \pi(\theta|x) \, d\theta = 0
\]
Bayes Estimator based on absolute error loss

Suppose that $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$. The posterior expected loss is

$$E[L(\theta, \hat{\theta}(x))] = \int_\Omega |\theta - \hat{\theta}(x)| \pi(\theta|x) d\theta$$

$$= E[|\theta - \hat{\theta}||X = x]$$

$$= \int_{-\infty}^{\hat{\theta}} -(\theta - \hat{\theta}) \pi(\theta|x) d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta}) \pi(\theta|x) d\theta$$

$$\frac{\partial}{\partial \hat{\theta}} E[L(\theta, \hat{\theta}(x))] = \int_{-\infty}^{\hat{\theta}} \pi(\theta|x) d\theta - \int_{\hat{\theta}}^{\infty} \pi(\theta|x) d\theta = 0$$

Therefore, $\hat{\theta}$ is posterior median.
Asymptotic Evaluation of Point Estimators

When the sample size $n$ approaches infinity, the behaviors of an estimator are unknown as its *asymptotic* properties.
Asymptotic Evaluation of Point Estimators

When the sample size $n$ approaches infinity, the behaviors of an estimator are unknown as its *asymptotic* properties.

**Definition - Consistency**

Let $W_n = W_n(X_1, \cdots, X_n) = W_n(X)$ be a sequence of estimators for $\tau(\theta)$. We say $W_n$ is consistent for estimating $\tau(\theta)$ if $W_n \xrightarrow{P} \tau(\theta)$ under $P_\theta$ for every $\theta \in \Omega$. 
Recap

Asymptotic Normality

Asymptotic Efficiency

Summary

Asymptotic Evaluation of Point Estimators

When the sample size $n$ approaches infinity, the behaviors of an estimator are unknown as its *asymptotic* properties.

**Definition - Consistency**

Let $W_n = W_n(X_1, \ldots, X_n) = W_n(X)$ be a sequence of estimators for $\tau(\theta)$. We say $W_n$ is consistent for estimating $\tau(\theta)$ if $W_n \xrightarrow{P} \tau(\theta)$ under $P_\theta$ for every $\theta \in \Omega$.

$W_n \xrightarrow{P} \tau(\theta)$ (converges in probability to $\tau(\theta)$) means that, given any $\epsilon > 0$.

$$\lim_{n \to \infty} \Pr(|W_n - \tau(\theta)| \geq \epsilon) = 0$$

$$\lim_{n \to \infty} \Pr(|W_n - \tau(\theta)| < \epsilon) = 1$$
Asymptotic Evaluation of Point Estimators

When the sample size $n$ approaches infinity, the behaviors of an estimator are unknown as its *asymptotic* properties.

**Definition - Consistency**

Let $W_n = W_n(X_1, \cdots, X_n) = W_n(\mathbf{X})$ be a sequence of estimators for $\tau(\theta)$. We say $W_n$ is consistent for estimating $\tau(\theta)$ if $W_n \xrightarrow{P} \tau(\theta)$ under $P_\theta$ for every $\theta \in \Omega$.

$W_n \xrightarrow{P} \tau(\theta)$ (converges in probability to $\tau(\theta)$) means that, given any $\epsilon > 0$.

$$\lim_{n \to \infty} \Pr( |W_n - \tau(\theta)| \geq \epsilon ) = 0$$
$$\lim_{n \to \infty} \Pr( |W_n - \tau(\theta)| < \epsilon ) = 1$$

When $|W_n - \tau(\theta)| < \epsilon$ can also be represented that $W_n$ is close to $\tau(\theta)$. 

Hyun Min Kang
Biostatistics 602 - Lecture 16
March 19th, 2013 4 / 33
Asymptotic Evaluation of Point Estimators

When the sample size $n$ approaches infinity, the behaviors of an estimator are unknown as its *asymptotic* properties.

**Definition - Consistency**

Let $W_n = W_n(X_1, \cdots, X_n) = W_n(\mathbf{X})$ be a sequence of estimators for $\tau(\theta)$. We say $W_n$ is consistent for estimating $\tau(\theta)$ if $W_n \xrightarrow{P} \tau(\theta)$ under $P_\theta$ for every $\theta \in \Omega$.

$W_n \xrightarrow{P} \tau(\theta)$ (converges in probability to $\tau(\theta)$) means that, given any $\epsilon > 0$,

\[
\lim_{n \to \infty} \Pr(|W_n - \tau(\theta)| \geq \epsilon) = 0
\]

\[
\lim_{n \to \infty} \Pr(|W_n - \tau(\theta)| < \epsilon) = 1
\]

When $|W_n - \tau(\theta)| < \epsilon$ can also be represented that $W_n$ is close to $\tau(\theta)$. Consistency implies that the probability of $W_n$ close to $\tau(\theta)$ approaches to 1 as $n$ goes to $\infty$. 
Tools for proving consistency

- Use definition (complicated)
Tools for proving consistency

- Use definition (complicated)
- Chebychev’s Inequality

\[
\Pr(|W_n - \tau(\theta)| \geq \epsilon) = \Pr((W_n - \tau(\theta))^2 \geq \epsilon^2) \\
\leq \frac{\text{E}[(W_n - \tau(\theta))^2]}{\epsilon^2} \\
= \frac{\text{MSE}(W_n)}{\epsilon^2} = \frac{\text{Bias}^2(W_n) + \text{Var}(W_n)}{\epsilon^2}
\]
Tools for proving consistency

- Use definition (complicated)
- Chebychev’s Inequality

\[
\Pr(|W_n - \tau(\theta)| \geq \epsilon) = \Pr((W_n - \tau(\theta))^2 \geq \epsilon^2) \leq \frac{\text{E}[W_n - \tau(\theta)]^2}{\epsilon^2} = \frac{\text{MSE}(W_n)}{\epsilon^2} = \frac{\text{Bias}^2(W_n) + \text{Var}(W_n)}{\epsilon^2}
\]

Need to show that both \text{Bias}(W_n) and \text{Var}(W_n) converges to zero
Theorem for consistency

**Theorem 10.1.3**

If $W_n$ is a sequence of estimators of $\tau(\theta)$ satisfying

- $\lim_{n \to \infty} \text{Bias}(W_n) = 0$.
- $\lim_{n \to \infty} \text{Var}(W_n) = 0$.

for all $\theta$, then $W_n$ is consistent for $\tau(\theta)$.
Weak Law of Large Numbers

**Theorem 5.5.2**

Let $X_1, \cdots, X_n$ be iid random variables with $E(X) = \mu$ and $\text{Var}(X) = \sigma^2 < \infty$. Then $\overline{X}_n$ converges in probability to $\mu$.

I.e. $\overline{X}_n \xrightarrow{P} \mu$. 
Consistent sequence of estimators

Theorem 10.1.5

Let $W_n$ is a consistent sequence of estimators of $\tau(\theta)$. Let $a_n, b_n$ be sequences of constants satisfying

1. $\lim_{n \to \infty} a_n = 1$
2. $\lim_{n \to \infty} b_n = 0$. 

Continuous Map Theorem

If $W_n$ is consistent for $\tau$ and $g$ is a continuous function, then $g(W_n)$ is consistent for $g(\tau)$. 
Consistent sequence of estimators

**Theorem 10.1.5**

Let $W_n$ is a consistent sequence of estimators of $\tau(\theta)$. Let $a_n, b_n$ be sequences of constants satisfying

1. $\lim_{n \to \infty} a_n = 1$
2. $\lim_{n \to \infty} b_n = 0$.

Then $U_n = a_n W_n + b_n$ is also a consistent sequence of estimators of $\tau(\theta)$. 
Consistent sequence of estimators

Theorem 10.1.5

Let $W_n$ is a consistent sequence of estimators of $\tau(\theta)$. Let $a_n, b_n$ be sequences of constants satisfying

1. $\lim_{n \to \infty} a_n = 1$
2. $\lim_{n \to \infty} b_n = 0$.

Then $U_n = a_n W_n + b_n$ is also a consistent sequence of estimators of $\tau(\theta)$.

Continuous Map Theorem

If $W_n$ is consistent for $\theta$ and $g$ is a continuous function, then $g(W_n)$ is consistent for $g(\theta)$. 
Example - Exponential Family

Problem

Suppose \( X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exponential}(\beta) \).
Example - Exponential Family

Problem

Suppose $X_1, \cdots, X_n \overset{i.i.d.}{\sim} \text{Exponential}(\beta)$.

1. Propose a consistent estimator of the median.
Example - Exponential Family

Problem

Suppose \( X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exponential}(\beta) \).

1. Propose a consistent estimator of the median.
2. Propose a consistent estimator of \( \Pr(X \leq c) \) where \( c \) is constant.
Consistent estimator of \( \Pr(X \leq c) \)

\[
\Pr(X \leq c) = \int_0^c \frac{1}{\beta} e^{-x/\beta} \, dx
\]
Consistent estimator of $\Pr(X \leq c)$

\[
\Pr(X \leq c) = \int_0^c \frac{1}{\beta} e^{-x/\beta} \, dx
\]

\[
= 1 - e^{-c/\beta}
\]
Consistent estimator of $\Pr(X \leq c)$

$$
\Pr(X \leq c) = \int_0^c \frac{1}{\beta} e^{-x/\beta} \, dx \\
= 1 - e^{-c/\beta}
$$

As $X$ is consistent for $\beta$, $1 - e^{-c/\beta}$ is continuous function of $\beta$. 
Consistent estimator of $\Pr(X \leq c)$

$$\Pr(X \leq c) = \int_0^c \frac{1}{\beta} e^{-x/\beta} \, dx$$

$$= 1 - e^{-c/\beta}$$

As $\bar{X}$ is consistent for $\beta$, $1 - e^{-c/\beta}$ is a continuous function of $\beta$. By continuous mapping Theorem, $g(\bar{X}) = 1 - e^{-c/\bar{X}}$ is consistent for $\Pr(X \leq c) = 1 - e^{-c/\beta} = g(\beta)$
Consistent estimator of $\Pr(X \leq c)$ - Alternative Method

Define $Y_i = I(X_i \leq c)$. Then $Y_i \sim \text{i.i.d. Bernoulli}(p)$ where $p = \Pr(X \leq c)$. 
Define $Y_i = I(X_i \leq c)$. Then $Y_i \overset{	ext{i.i.d.}}{\sim} \text{Bernoulli}(p)$ where $p = \Pr(X \leq c)$.

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq c)$$

is consistent for $p$ by Law of Large Numbers.
Theorem 10.1.6 - Consistency of MLEs

Suppose $X_i \sim \text{i.i.d.} f(x|\theta)$. Let $\hat{\theta}$ be the MLE of $\theta$, and $\tau(\theta)$ be a continuous function of $\theta$. Then under "regularity conditions" on $f(x|\theta)$, the MLE of $\tau(\theta)$ (i.e. $\tau(\hat{\theta})$) is consistent for $\tau(\theta)$. 
Asymptotic Normality

Definition: Asymptotic Normality

A statistic (or an estimator) $W_n(X)$ is asymptotically normal if

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta))$$

for all $\theta$

where $\xrightarrow{d}$ stands for "converge in distribution"
Asymptotic Normality

Definition: Asymptotic Normality

A statistic (or an estimator) $W_n(X)$ is asymptotically normal if

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta))$$

for all $\theta$

where $\xrightarrow{d}$ stands for "converge in distribution"

- $\tau(\theta)$ : "asymptotic mean"
- $\nu(\theta)$ : "asymptotic variance"
Defintion: Asymptotic Normality

A statistic (or an estimator) $W_n(\mathbf{X})$ is *asymptotically normal* if

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta))$$

for all $\theta$, where $\xrightarrow{d}$ stands for "converge in distribution"

- $\tau(\theta)$: "asymptotic mean"
- $\nu(\theta)$: "asymptotic variance"

We denote $W_n \sim \mathcal{A}\mathcal{N}\left(\tau(\theta), \frac{\nu(\theta)}{n}\right)$. 


Central Limit Theorem

Assume \( X_i \overset{\text{i.i.d.}}{\sim} f(x|\theta) \) with finite mean \( \mu(\theta) \) and variance \( \sigma^2(\theta) \).

\[
\bar{X} \sim \mathcal{N}\left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right)
\]
Central Limit Theorem

Assume $X_i \overset{	ext{i.i.d.}}{\sim} f(x|\theta)$ with finite mean $\mu(\theta)$ and variance $\sigma^2(\theta)$.

\[
\bar{X} \sim \mathcal{N}\left(\mu(\theta), \frac{\sigma^2(\theta)}{n}\right)
\]

\[\iff \quad \sqrt{n} (\bar{X} - \mu(\theta)) \overset{d}{\to} \mathcal{N}(0, \sigma^2(\theta))\]
Central Limit Theorem

Assume $X_i \sim f(x|\theta)$ with finite mean $\mu(\theta)$ and variance $\sigma^2(\theta)$.

$\overline{X} \sim \mathcal{N} \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right)$

$\Leftrightarrow \sqrt{n} \left( \overline{X} - \mu(\theta) \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\theta))$

Theorem 5.5.17 - Slutsky’s Theorem

If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} a$, where $a$ is a constant,
Central Limit Theorem

Assume $X_i \overset{i.i.d.}{\sim} f(x|\theta)$ with finite mean $\mu(\theta)$ and variance $\sigma^2(\theta)$.

$$\bar{X} \sim \mathcal{N} \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right)$$

$$\Leftrightarrow \sqrt{n} \left( \bar{X} - \mu(\theta) \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\theta))$$

Theorem 5.5.17 - Slutsky’s Theorem

If $X_n \xrightarrow{d} X, Y_n \xrightarrow{P} a$, where $a$ is a constant,

1. $Y_n \cdot X_n \xrightarrow{d} aX$
Central Limit Theorem

Assume \( X_i \overset{i.i.d.}{\sim} f(x|\theta) \) with finite mean \( \mu(\theta) \) and variance \( \sigma^2(\theta) \).

\[
\frac{1}{\sqrt{n}} \left( \bar{X} - \mu(\theta) \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\theta) / n)
\]

Theorem 5.5.17 - Slutsky’s Theorem

If \( X_n \xrightarrow{d} X \), \( Y_n \xrightarrow{P} a \), where \( a \) is a constant,

1. \( Y_n \cdot X_n \xrightarrow{d} aX \)
2. \( X_n + Y_n \xrightarrow{d} X + a \)
Example - Estimator of $\Pr(X \leq c)$

Define $Y_i = I(X_i \leq c)$. Then $Y_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p)$ where $p = \Pr(X \leq c)$. 

Example - Estimator of $\Pr(X \leq c)$

Define $Y_i = I(X_i \leq c)$. Then $Y_i \sim_{i.i.d.} \text{Bernoulli}(p)$ where $p = \Pr(X \leq c)$.

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq c)$$

is consistent for $p$. Therefore,
Example - Estimator of $\Pr(X \leq c)$

Define $Y_i = I(X_i \leq c)$. Then $Y_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p)$ where $p = \Pr(X \leq c)$.

$$
\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq c)
$$

is consistent for $p$. Therefore,

$$
\frac{1}{n} \sum_{i=1}^{n} I(X_i \leq c) \sim \mathcal{AN} \left( \text{E}(Y), \frac{\text{Var}(Y)}{n} \right)
$$

$$
= \mathcal{AN} \left( p, \frac{p(1-p)}{n} \right)
$$
Example

Let \( X_1, \ldots, X_n \) be iid samples with finite mean \( \mu \) and variance \( \sigma^2 \). Define

\[
S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]

By Central Limit Theorem, \( \bar{X}_n \overset{\text{d}}{\rightarrow} N(\mu, \sigma^2/n) \) as \( n \to \infty \).
Example

Let $X_1, \cdots, X_n$ be iid samples with finite mean $\mu$ and variance $\sigma^2$. Define

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

By Central Limit Theorem,

$$\bar{X}_n \sim \mathcal{AN}\left(\mu, \frac{\sigma^2}{n}\right)$$
Example

Let $X_1, \ldots, X_n$ be iid samples with finite mean $\mu$ and variance $\sigma^2$. Define

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

By Central Limit Theorem,

$$\bar{X}_n \sim \mathcal{AN} \left( \mu, \frac{\sigma^2}{n} \right)$$

$$\iff \sqrt{n}(\bar{X} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$
**Example**

Let \( X_1, \ldots, X_n \) be iid samples with finite mean \( \mu \) and variance \( \sigma^2 \). Define

\[
S^2_n = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2
\]

By Central Limit Theorem,

\[
\overline{X}_n \sim \mathcal{N} \left( \mu, \frac{\sigma^2}{n} \right)
\]

\[
\iff \sqrt{n}(\overline{X} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)
\]

\[
\iff \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1)
\]
Example (cont’d)

\[
\frac{\sqrt{n}(\bar{X} - \mu)}{S_n} = \frac{\sigma}{S_n} \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}
\]
Asymptotic Normality

$$\sqrt{n(\overline{X} - \mu)} \frac{S_n}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}(\overline{X} - \mu)}$$

We showed previously $S_n^2 \xrightarrow{p} \sigma^2 \Rightarrow S_n \xrightarrow{p} \sigma \Rightarrow \sigma / S_n \xrightarrow{p} 1$. 

Asymptotic Efficiency

Recap

Summary
Example (cont’d)

\[
\frac{\sqrt{n} (\overline{X} - \mu)}{S_n} = \frac{\sigma \sqrt{n} (\overline{X} - \mu)}{\sigma}
\]

We showed previously \( S_n^2 \xrightarrow{P} \sigma^2 \Rightarrow S_n \xrightarrow{P} \sigma \Rightarrow \sigma / S_n \xrightarrow{P} 1 \).

Therefore, By Slutsky’s Theorem \( \frac{\sqrt{n} (\overline{X} - \mu)}{S_n} \xrightarrow{P} \mathcal{N}(0, 1) \).
Delta Method

Theorem 5.5.24 - Delta Method

Assume $W_n \sim \mathcal{AN} \left( \theta, \frac{\nu(\theta)}{n} \right)$. If a function $g$ satisfies $g'(\theta) \neq 0$, then

$$g(W_n) \sim \mathcal{AN} \left( g(\theta), \left[ g'(\theta) \right]^2 \frac{\nu(\theta)}{n} \right)$$
Delta Method - Example

\[ X_1, \cdots, X_n \overset{i.i.d.}{\sim} \text{Bernoulli}(p) \text{ where } p \neq \frac{1}{2}, \text{ we want to know the asymptotic distribution of } \overline{X}(1 - \overline{X}). \]
Delta Method - Example

\(X_1, \cdots, X_n \sim \text{i.i.d.} \text{ Bernoulli}(p)\) where \(p \neq \frac{1}{2}\), we want to know the asymptotic distribution of \(\bar{X}(1 - \bar{X})\). By central limit Theorem,

\[
\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1 - p)}} \overset{d}{\rightarrow} N(0, 1)
\]
Delta Method - Example

\( X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Bernoulli}(p) \) where \( p \neq \frac{1}{2} \), we want to know the asymptotic distribution of \( \bar{X}(1 - \bar{X}) \). By central limit Theorem,

\[
\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1 - p)}} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\( \Leftrightarrow \bar{X}_n \sim \mathcal{AN} \left( p, \frac{p(1 - p)}{n} \right) \)
Delta Method - Example

\(X_1, \cdots, X_n \overset{i.i.d.}{\sim} \text{Bernoulli}(p)\) where \(p \neq \frac{1}{2}\), we want to know the asymptotic distribution of \(\bar{X}(1 - \bar{X})\). By central limit Theorem,

\[
\frac{\sqrt{n} (\bar{X}_n - p)}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)
\]

\[
\iff \bar{X}_n \sim \mathcal{AN} \left( p, \frac{p(1-p)}{n} \right)
\]

Define \(g(y) = y(1 - y)\), then \(\bar{X}(1 - \bar{X}) = g(\bar{X})\).
Delta Method - Example

\(X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)\) where \(p \neq \frac{1}{2}\), we want to know the asymptotic distribution of \(\bar{X}(1 - \bar{X})\). By central limit Theorem,

\[
\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1 - p)}} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\[
\Leftrightarrow \bar{X}_n \sim \mathcal{AN} \left( p, \frac{p(1 - p)}{n} \right)
\]

Define \(g(y) = y(1 - y)\), then \(\bar{X}(1 - \bar{X}) = g(\bar{X})\).

\[
g'(y) = (y - y^2)' = 1 - 2y
\]
Delta Method - Example

\( X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Bernoulli}(p) \) where \( p \neq \frac{1}{2} \), we want to know the asymptotic distribution of \( \bar{X}(1 - \bar{X}) \). By central limit Theorem,

\[
\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1 - p)}} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\[ \iff \bar{X}_n \sim \mathcal{AN} \left( p, \frac{p(1 - p)}{n} \right) \]

Define \( g(y) = y(1 - y) \), then \( \bar{X}(1 - \bar{X}) = g(\bar{X}) \).

\[ g'(y) = (y - y^2)' = 1 - 2y \]

By Delta Method,

\[ g(\bar{X}) = \bar{X}(1 - \bar{X}) \sim \mathcal{AN} \left( g(p), \left[ g'(p) \right]^2 \frac{p(1 - p)}{n} \right) \]
Delta Method - Example

\(X_1, \ldots, X_n \sim \text{i.i.d.} \) Bernoulli\((p)\) where \(p \neq \frac{1}{2}\), we want to know the asymptotic distribution of \(\overline{X}(1 - \overline{X})\). By central limit Theorem,

\[
\frac{\sqrt{n}(\overline{X}_n - p)}{\sqrt{p(1 - p)}} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\(\Leftrightarrow \overline{X}_n \sim \mathcal{AN} \left( p, \frac{p(1 - p)}{n} \right) \)

Define \(g(y) = y(1 - y)\), then \(\overline{X}(1 - \overline{X}) = g(\overline{X})\).

\(g'(y) = (y - y^2)' = 1 - 2y\)

By Delta Method,

\[
g(\overline{X}) = \overline{X}(1 - \overline{X}) \sim \mathcal{AN} \left( g(p), \left[g'(p)\right]^2 \frac{p(1 - p)}{n} \right)
\]

\[
= \mathcal{AN} \left( p(1 - p), (1 - 2p)^2 \frac{p(1 - p)}{n} \right)
\]
Asymptotic Normality

Given a statistic $W_n(X)$, for example $\bar{X}$, $s_X^2$, $e^{-\bar{X}}$
Asymptotic Normality

Given a statistic $W_n(X)$, for example $\bar{X}$, $s^2_X$, $e^{-\bar{X}}$

$\sqrt{n}( W_n - \tau(\theta)) \overset{d}{\rightarrow} \mathcal{N}(0, \nu(\theta))$ for all $\theta$

$\iff W_n \sim \mathcal{AN} \left( \tau(\theta), \frac{\nu(\theta)}{n} \right)$
Asymptotic Normality

Given a statistic $W_n(X)$, for example $\bar{X}, s^2_X, e^{-X}$

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta)) \quad \text{for all } \theta$$

$$\iff W_n \sim \mathcal{AN}\left(\tau(\theta), \frac{\nu(\theta)}{n}\right)$$

Tools to show asymptotic normality

1. Central Limit Theorem
Asymptotic Normality

Given a statistic $W_n(\mathbf{X})$, for example $\overline{X}$, $s^2_{\mathbf{X}}$, $e^{-\overline{X}}$

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta)) \quad \text{for all } \theta$$

$$\iff W_n \sim \mathcal{AN}\left(\tau(\theta), \frac{\nu(\theta)}{n}\right)$$

Tools to show asymptotic normality

1. Central Limit Theorem
2. Slutsky Theorem
Given a statistic $W_n(X)$, for example $\overline{X}$, $s^2_X$, $e^{-X}$

$$\sqrt{n}( W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta)) \quad \text{for all } \theta$$

$$\iff W_n \sim \mathcal{AN}\left( \tau(\theta), \frac{\nu(\theta)}{n} \right)$$

Tools to show asymptotic normality

1. Central Limit Theorem
2. Slutsky Theorem
3. Delta Method (Theorem 5.5.24)
Using Central Limit Theorem

\[ \bar{X} \sim \mathcal{N} \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right) \]

where \( \mu(\theta) = E(X) \), and \( \sigma^2(\theta) = \text{Var}(X) \).
Using Central Limit Theorem

\[ \bar{X} \sim \mathcal{N} \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right) \]

where \( \mu(\theta) = \mathbb{E}(X) \), and \( \sigma^2(\theta) = \text{Var}(X) \).

For example, in order to get the asymptotic distribution of \( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \),
Using Central Limit Theorem

\[ \bar{X} \sim AN \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right) \]

where \( \mu(\theta) = E(X) \), and \( \sigma^2(\theta) = Var(X) \).

For example, in order to get the asymptotic distribution of \( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \), define \( Y_i = X_i^2 \), then

\[ \frac{1}{n} \sum_{i=1}^{n} X_i^2 = \frac{1}{n} \sum_{i=1}^{n} Y_i = \bar{Y} \]
Using Central Limit Theorem

\[ \bar{X} \sim \mathcal{AN} \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right) \]

where \( \mu(\theta) = \mathbb{E}(X) \), and \( \sigma^2(\theta) = \text{Var}(X) \).

For example, in order to get the asymptotic distribution of \( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \), define \( Y_i = X_i^2 \), then

\[ \frac{1}{n} \sum_{i=1}^{n} X_i^2 = \frac{1}{n} \sum_{i=1}^{n} Y_i = \bar{Y} \]

\[ \sim \mathcal{AN} \left( \mathbb{E}Y, \frac{\text{Var}(Y)}{n} \right) \]
Using Central Limit Theorem

\[ \overline{X} \sim \mathcal{N} \left( \mu(\theta), \frac{\sigma^2(\theta)}{n} \right) \]

where \( \mu(\theta) = \mathbb{E}(X) \), and \( \sigma^2(\theta) = \text{Var}(X) \).

For example, in order to get the asymptotic distribution of \( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \), define \( Y_i = X_i^2 \), then

\[ \frac{1}{n} \sum_{i=1}^{n} X_i^2 = \frac{1}{n} \sum_{i=1}^{n} Y_i = \overline{Y} \]

\[ \sim \mathcal{N} \left( \mathbb{E}Y, \frac{\text{Var}(Y)}{n} \right) \]

\[ \sim \mathcal{N} \left( \mathbb{E}X^2, \frac{\text{Var}(X^2)}{n} \right) \]
Using Slutsky Theorem

When \( X_n \overset{d}{\rightarrow} X, \ Y_n \overset{P}{\rightarrow} a, \) then

1. \( Y_nX_n \overset{d}{\rightarrow} aX \)
2. \( X_n + Y_n \overset{d}{\rightarrow} X + a. \)
Using Delta Method (Theorem 5.5.24)

Assume $W_n \sim \mathcal{AN}(\theta, \frac{\nu(\theta)}{n})$. If a function $g$ satisfies $g'(\theta) \neq 0$, then

$$g( W_n ) \sim \mathcal{AN} \left( g(\theta), [g'(\theta)]^2 \frac{\nu(\theta)}{n} \right)$$
Example

Problem

\[ X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2) \quad \mu \neq 0 \]

Find the asymptotic distribution of MLE of \( \mu^2 \).
Example

Problem

\[ X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2) \quad \mu \neq 0 \]

Find the asymptotic distribution of MLE of \( \mu^2 \).

Solution

1. It can be easily shown that MLE of \( \mu \) is \( \bar{X} \).
Example

Problem

\[ X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2) \quad \mu \neq 0 \]

Find the asymptotic distribution of MLE of \( \mu^2 \).

Solution

1. It can be easily shown that MLE of \( \mu \) is \( \overline{X} \).
2. By the invariance property of MLE, MLE of \( \mu^2 \) is \( \overline{X}^2 \).
Example

Problem

\[ X_1, \cdots, X_n \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma^2) \quad \mu \neq 0 \]

Find the asymptotic distribution of MLE of \( \mu^2 \).

Solution

1. It can be easily shown that MLE of \( \mu \) is \( \bar{X} \).
2. By the invariance property of MLE, MLE of \( \mu^2 \) is \( \bar{X}^2 \).
3. By central limit theorem, we know that
   \[
   \bar{X} \sim \mathcal{AN} \left( \mu, \frac{\sigma^2}{n} \right)
   \]
Define \( g(y) = y^2 \), and apply Delta Method.
Solution (cont’d)

4 Define \( g(y) = y^2 \), and apply Delta Method.

\[
g'(y) = 2y
\]
Solution (cont’d)

4 Define $g(y) = y^2$, and apply Delta Method.

$$g'(y) = 2y$$

$$\bar{X}^2 \sim \mathcal{N} \left( g(\mu), \frac{[g'(\mu)]^2 \sigma^2}{n} \right)$$
Solution (cont’d)

Define $g(y) = y^2$, and apply Delta Method.

$$g'(y) = 2y$$

$$\bar{X}^2 \sim \mathcal{AN} \left( g(\mu), [g'(\mu)]^2 \frac{\sigma^2}{n} \right)$$

$$\sim \mathcal{AN} \left( \mu^2, (2\mu)^2 \frac{\sigma^2}{n} \right)$$
Asymptotic Relative Efficiency (ARE)

If both estimators are consistent and asymptotic normal, we can compare their asymptotic variance.
Asymptotic Relative Efficiency (ARE)

If both estimators are consistent and asymptotic normal, we can compare their asymptotic variance.

**Definition 10.1.16 : Asymptotic Relative Efficiency**

If two estimators $W_n$ and $V_n$ satisfy

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} N(0, \sigma^2_W)$$

$$\sqrt{n}(V_n - \tau(\theta)) \xrightarrow{d} N(0, \sigma^2_V)$$

The asymptotic relative efficiency (ARE) of $V_n$ with respect to $W_n$ is

$$\text{ARE}(V_n; W_n) = \frac{\sigma^2_W}{\sigma^2_V}$$

If $\text{ARE}(V_n; W_n) > 1$ for every $\theta \in \Omega$, then $V_n$ is asymptotically more efficient than $W_n$. 
Asymptotic Relative Efficiency (ARE)

If both estimators are consistent and asymptotic normal, we can compare their asymptotic variance.

Definition 10.1.16: Asymptotic Relative Efficiency

If two estimators $W_n$ and $V_n$ satisfy

$$\sqrt{n}[W_n - \tau(\theta)] \xrightarrow{d} \mathcal{N}(0, \sigma^2_W)$$

$$\sqrt{n}[V_n - \tau(\theta)] \xrightarrow{d} \mathcal{N}(0, \sigma^2_V)$$

The asymptotic relative efficiency (ARE) of $V_n$ with respect to $W_n$ is

$$\text{ARE}(V_n, W_n) = \frac{\sigma^2_W}{\sigma^2_V}$$
Asymptotic Relative Efficiency (ARE)

If both estimators are consistent and asymptotic normal, we can compare their asymptotic variance.

**Definition 10.1.16 : Asymptotic Relative Efficiency**

If two estimators $W_n$ and $V_n$ satisfy

\[
\sqrt{n}[W_n - \tau(\theta)] \xrightarrow{d} \mathcal{N}(0, \sigma^2_W)
\]

\[
\sqrt{n}[V_n - \tau(\theta)] \xrightarrow{d} \mathcal{N}(0, \sigma^2_V)
\]

The asymptotic relative efficiency (ARE) of $V_n$ with respect to $W_n$ is

\[
\text{ARE}(V_n, W_n) = \frac{\sigma^2_W}{\sigma^2_V}
\]

If $\text{ARE}(V_n, W_n) \geq 1$ for every $\theta \in \Omega$, then $V_n$ is asymptotically more efficient than $W_n$. 

Hyun Min Kang
Biostatistics 602 - Lecture 16
March 19th, 2013 26 / 33
Example

**Problem**

Let $X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. consider estimating

$$\Pr(X = 0) = e^{-\lambda}$$
Example

Problem

Let $X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. Consider estimating
\[ \Pr(X = 0) = e^{-\lambda} \]

Our estimators are
\[ W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) \]
Example

Problem

Let $X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. consider estimating

$$\Pr(X = 0) = e^{-\lambda}$$

Our estimators are

$$W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0)$$

$$V_n = e^{-\bar{X}}$$
Example

Problem

Let \( X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda) \). Consider estimating

\[
\Pr(X = 0) = e^{-\lambda}
\]

Our estimators are

\[
W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0)
\]

\[
V_n = e^{-\bar{X}}
\]

Determine which one is more asymptotically efficient estimator.
Solution - Asymptotic Distribution of $V_n$

$$V_n(X) = e^{-\bar{X}}$$, by CLT,
Solution - Asymptotic Distribution of $V_n$

\[ V_n(X) = e^{-\bar{X}}, \text{ by CLT,} \]

\[ \bar{X} \sim \mathcal{N}(EX, \text{Var}X/n) \sim \mathcal{N}(\lambda, \lambda/n) \]
Solution - Asymptotic Distribution of $V_n$

\[ V_n(X) = e^{-\bar{X}}, \text{ by CLT}, \]

\[ \bar{X} \sim \mathcal{N}(EX, \text{Var}X/n) \sim \mathcal{N}(\lambda, \lambda/n) \]

Define $g(y) = e^{-y}$, then $V_n = g(\bar{X})$ and $g'(y) = -e^{-y}$. By Delta Method
Solution - Asymptotic Distribution of $V_n$

$V_n(\mathbf{X}) = e^{-\bar{X}}$, by CLT,

$$\bar{X} \sim \mathcal{A}\mathcal{N}(EX, \text{Var}X/n) \sim \mathcal{A}\mathcal{N}(\lambda, \lambda/n)$$

Define $g(y) = e^{-y}$, then $V_n = g(\bar{X})$ and $g'(y) = -e^{-y}$. By Delta Method

$$V_n = e^{-\bar{X}} \sim \mathcal{A}\mathcal{N} \left( g(\lambda), [g'(\lambda)]^2 \frac{\lambda}{n} \right)$$
Solution - Asymptotic Distribution of $V_n$

$V_n(\mathbf{X}) = e^{-\bar{X}}$, by CLT,

$\bar{X} \sim \mathcal{N}(EX, VarX/n) \sim \mathcal{N}(\lambda, \lambda/n)$

Define $g(y) = e^{-y}$, then $V_n = g(\bar{X})$ and $g'(y) = -e^{-y}$. By Delta Method

$V_n = e^{-\bar{X}} \sim \mathcal{N} \left( g(\lambda), [g'(\lambda)]^2 \frac{\lambda}{n} \right)$

$\sim \mathcal{N} \left( e^{-\lambda}, e^{-2\lambda} \frac{\lambda}{n} \right)$
Solution - Asymptotic Distribution of $W_n$

Define $Z_i = I(X_i = 0)$
Solution - Asymptotic Distribution of $W_n$

Define $Z_i = I(X_i = 0)$

$$W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n$$
Solution - Asymptotic Distribution of $W_n$

Define $Z_i = I(X_i = 0)$

\[
W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n
\]

$Z_i \sim \text{Bernoulli}(E(Z))$
Solution - Asymptotic Distribution of $W_n$

Define $Z_i = I(X_i = 0)$

$$W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) = \bar{Z}_n$$

$$Z_i \sim \text{Bernoulli}(E(Z))$$

$$E(Z) = \Pr(X = 0) = e^{-\lambda}$$
Define $Z_i = I(X_i = 0)$

$$W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n$$

$Z_i \sim$ Bernoulli($E(Z)$)

$$E(Z) = \Pr(X = 0) = e^{-\lambda}$$

$$\text{Var}(Z) = e^{-\lambda}(1 - e^{-\lambda})$$
Solution - Asymptotic Distribution of $W_n$

Define $Z_i = I(X_i = 0)$

$$W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) = \bar{Z}_n$$

$Z_i \sim$ Bernoulli$(E(Z))$

$E(Z) = Pr(X = 0) = e^{-\lambda}$

$Var(Z) = e^{-\lambda}(1 - e^{-\lambda})$

By CLT,

$$W_n = \bar{Z}_n \sim AN (E(Z), Var(Z)/n)$$
Solution - Asymptotic Distribution of $W_n$

Define $Z_i = I(X_i = 0)$

$$W_n = \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) = \bar{Z}_n$$

$Z_i \sim \text{Bernoulli}(E(Z))$

$$E(Z) = Pr(X = 0) = e^{-\lambda}$$

$$\text{Var}(Z) = e^{-\lambda}(1 - e^{-\lambda})$$

By CLT,

$$W_n = \bar{Z}_n \sim \mathcal{A}\mathcal{N} \left( E(Z), \frac{\text{Var}(Z)}{n} \right)$$

$$\sim \mathcal{A}\mathcal{N} \left( e^{-\lambda}, \frac{e^{-\lambda}(1 - e^{-\lambda})}{n} \right)$$
Solution - Calculating ARE

\[
\text{ARE}(W_n, V_n) = \frac{e^{-2\lambda} \lambda/n}{e^{-\lambda} (1 - e^{-\lambda})/n}
\]
Solution - Calculating ARE

\[ \text{ARE}(W_n, V_n) = \frac{e^{-2\lambda} \lambda/n}{e^{-\lambda}(1 - e^{-\lambda})/n} \]

\[ = \frac{\lambda}{e^{\lambda}(1 - e^{-\lambda})} \]

Therefore, \( W_n \) is less efficient than \( V_n \) (MLE), and ARE attains maximum at \( \lambda = 0 \).
Solution - Calculating ARE

\[
\text{ARE}(W_n, V_n) = \frac{e^{-2\lambda} \lambda / n}{e^{-\lambda} (1 - e^{-\lambda}) / n} \\
= \frac{\lambda}{e^\lambda (1 - e^{-\lambda})} \\
= \frac{\lambda}{e^\lambda - 1}
\]
Solution - Calculating ARE

\[ \text{ARE}(W_n, V_n) = \frac{e^{-2\lambda} \lambda/n}{e^{-\lambda}(1 - e^{-\lambda})/n} \]

\[ = \frac{\lambda}{e^\lambda(1 - e^{-\lambda})} \]

\[ = \frac{\lambda}{e^\lambda - 1} \]

\[ = \frac{\lambda}{\left(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \cdots\right) - 1} \]
Solution - Calculating ARE

\[
\text{ARE}(W_n, V_n) = \frac{e^{-2\lambda} \frac{\lambda}{n}}{e^{-\lambda} \left(1 - e^{-\lambda}\right) / n} = \frac{\lambda}{e^\lambda \left(1 - e^{-\lambda}\right)} = \frac{\lambda}{e^\lambda - 1} \\
\leq 1 \quad (\forall \lambda \geq 0)
\]
Solution - Calculating ARE

\[
\text{ARE}(W_n, V_n) = \frac{e^{-2\lambda} \frac{\lambda}{n}}{e^{-\lambda} \left(1 - e^{-\lambda}\right) / n} \leq 1 \quad (\forall \lambda \geq 0)
\]

Therefore \( W_n = \frac{1}{n} \sum I(X_i = 0) \) is less efficient than \( V_n \) (MLE), and ARE attains maximum at \( \lambda = 0 \).
Asymptotic Efficiency

Definition: Asymptotic Efficiency for iid samples

A sequence of estimators $W_n$ is asymptotically efficient for $\tau(\theta)$ if for all $\theta \in \Omega$,
Asymptotic Efficiency

**Definition**: Asymptotic Efficiency for iid samples

A sequence of estimators $W_n$ is asymptotically efficient for $\tau(\theta)$ if for all $\theta \in \Omega$,

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}\left(0, \frac{[\tau'(\theta)]^2}{I(\theta)}\right)$$

Note: $\frac{[\tau'(\theta)]^2}{I(\theta)}$ is the C-R bound for unbiased estimators of $\tau(\theta)$.
Asymptotic Efficiency

Definition: Asymptotic Efficiency for iid samples

A sequence of estimators \( W_n \) is asymptotically efficient for \( \tau(\theta) \) if for all \( \theta \in \Omega \),

\[
\sqrt{n}( W_n - \tau(\theta) ) \xrightarrow{d} \mathcal{N} \left( 0, \frac{[\tau'(\theta)]^2}{I(\theta)} \right)
\]

\[\iff\]
\[ W_n \sim \mathcal{AN} \left( \tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)} \right) \]
Asymptotic Efficiency

**Definition: Asymptotic Efficiency for iid samples**

A sequence of estimators \( W_n \) is asymptotically efficient for \( \tau(\theta) \) if for all \( \theta \in \Omega \),

\[
\sqrt{n}( W_n - \tau(\theta) ) \xrightarrow{d} \mathcal{N} \left( 0, \frac{[\tau'(\theta)]^2}{I(\theta)} \right)
\]

\[ \iff \quad W_n \sim \mathcal{AN} \left( \tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)} \right) \]

\[ I(\theta) = E \left[ \left\{ \frac{\partial}{\partial \theta} \log f(X|\theta) \right\}^2 | \theta \right] \]
Asymptotic Efficiency

Definition: Asymptotic Efficiency for iid samples

A sequence of estimators $W_n$ is asymptotically efficient for $\tau(\theta)$ if for all $\theta \in \Omega$,

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} N\left(0, \frac{[\tau'(\theta)]^2}{I(\theta)}\right)$$

$$\iff W_n \sim \mathcal{AN}\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)}\right)$$

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

$$= -E\left[\frac{\partial^2}{\partial \theta^2} \log f(X|\theta) | \theta\right] \text{ (if interchangeability holds)}$$
Asymptotic Efficiency

Definition: Asymptotic Efficiency for iid samples

A sequence of estimators $W_n$ is asymptotically efficient for $\tau(\theta)$ if for all $\theta \in \Omega$,

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} N\left(0, \frac{[\tau'(\theta)]^2}{I(\theta)}\right)$$

$$\iff W_n \sim \mathcal{A}\mathcal{N}\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)}\right)$$

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

$$= -E \left[\frac{\partial^2}{\partial \theta^2} \log f(X|\theta)| \theta\right] \quad (\text{if interchangeability holds})$$

Note: $\frac{[\tau'(\theta)]^2}{nI(\theta)}$ is the C-R bound for unbiased estimators of $\tau(\theta)$. 
Asymptotic Efficiency of MLEs

Theorem 10.1.12

Let \( X_1, \cdots, X_n \) be iid samples from \( f(x|\theta) \). Let \( \hat{\theta} \) denote the MLE of \( \theta \). Under same regularity conditions, \( \hat{\theta} \) is consistent and asymptotically normal for \( \theta \), i.e.

\[
\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}
\left(0, \frac{1}{I(\theta)}\right)
\text{for every } \theta \in \Omega
\]
Theorem 10.1.12

Let $X_1, \cdots, X_n$ be iid samples from $f(x|\theta)$. Let $\hat{\theta}$ denote the MLE of $\theta$. Under same regularity conditions, $\hat{\theta}$ is consistent and asymptotically normal for $\theta$, i.e.

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N} \left(0, \frac{1}{I(\theta)} \right) \text{ for every } \theta \in \Omega$$

And if $\tau(\theta)$ is continuous and differentiable in $\theta$, then

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N} \left(0, \frac{[\tau'(\theta)]}{I(\theta)} \right)$$

$$\implies \tau(\hat{\theta}) \sim \mathcal{AN} \left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)} \right)$$
Theorem 10.1.12

Let \( X_1, \cdots, X_n \) be iid samples from \( f(x|\theta) \). Let \( \hat{\theta} \) denote the MLE of \( \theta \). Under same regularity conditions, \( \hat{\theta} \) is consistent and asymptotically normal for \( \theta \), i.e.

\[
\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \frac{1}{I(\theta)}) \text{ for every } \theta \in \Omega
\]

And if \( \tau(\theta) \) is continuous and differentiable in \( \theta \), then

\[
\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N} \left( 0, \frac{[\tau'(\theta)]}{I(\theta)} \right)
\]

\[
\implies \tau(\hat{\theta}) \sim \mathcal{AN} \left( \tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)} \right)
\]

Again, note that the asymptotic variance of \( \tau(\hat{\theta}) \) is Cramer-Rao lower bound for unbiased estimators of \( \tau(\theta) \).
Summary

Today

- Central Limit Theorem
- Slutsky Theorem
- Delta Method
- Asymptotic Relative Efficiency
Summary

Today

- Central Limit Theorem
- Slutsky Theorem
- Delta Method
- Asymptotic Relative Efficiency

Next Lecture

- Hypothesis Testing