Biostatistics 602 - Statistical Inference
Lecture 17
Asymptotic Evaluation of Point Estimators

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?

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?

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?
- 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

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

Bayes Estimator based on absolute error loss

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

$$\begin{aligned} \mathrm{E}[L(\theta, \hat{\theta}(\mathbf{x}))] &= \int_{\Omega} |\theta - \hat{\theta}(\mathbf{x})| \pi(\theta|\mathbf{x}) d\theta \\ &= \mathrm{E}[|\theta - \hat{\theta}||\mathbf{X} = \mathbf{x}] \end{aligned}$$

Bayes Estimator based on absolute error loss

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

$$\begin{aligned} \mathrm{E}[L(\theta, \hat{\theta}(\mathbf{x}))] &= \int_{\Omega} |\theta - \hat{\theta}(\mathbf{x})| \pi(\theta|\mathbf{x})\, d\theta \\ &= \mathrm{E}[|\theta - \hat{\theta}||\mathbf{X} = \mathbf{x}] \\ &= \int_{-\infty}^{\hat{\theta}} -(\theta - \hat{\theta})\pi(\theta|\mathbf{x})\, d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta})\pi(\theta|\mathbf{x})\, d\theta \end{aligned}$$

Bayes Estimator based on absolute error loss

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

$$\begin{aligned} \mathrm{E}[L(\theta, \hat{\theta}(\mathbf{x}))] &= \int_{\Omega} |\theta - \hat{\theta}(\mathbf{x})| \pi(\theta|\mathbf{x}) \, d\theta \\ &= \mathrm{E}[|\theta - \hat{\theta}||\mathbf{X} = \mathbf{x}] \\ &= \int_{-\infty}^{\hat{\theta}} -(\theta - \hat{\theta}) \pi(\theta|\mathbf{x}) \, d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta}) \pi(\theta|\mathbf{x}) \, d\theta \end{aligned}$$

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

Bayes Estimator based on absolute error loss

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

$$\begin{aligned} \mathrm{E}[L(\theta, \hat{\theta}(\mathbf{x}))] &= \int_\Omega |\theta - \hat{\theta}(\mathbf{x})| \pi(\theta|\mathbf{x}) \, d\theta \\ &= \mathrm{E}[|\theta - \hat{\theta}| | \mathbf{X} = \mathbf{x}] \\ &= \int_{-\infty}^{\hat{\theta}} -(\theta - \hat{\theta}) \pi(\theta|\mathbf{x}) \, d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta}) \pi(\theta|\mathbf{x}) \, d\theta \end{aligned}$$

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

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

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_θ for every $\theta \in \Omega$.

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

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

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 $W_n \xrightarrow{\text{P}} \tau(\theta)$ under P_θ for every $\theta \in \Omega$.

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

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

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

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 $W_n \xrightarrow{\text{P}} \tau(\theta)$ under P_θ for every $\theta \in \Omega$.

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

$$\begin{aligned} \lim_{n\rightarrow\infty} \Pr(|W_n - \tau(\theta)| \geq \epsilon) &= 0 \\ \lim_{n\rightarrow\infty} \Pr(|W_n - \tau(\theta)| < \epsilon) &= 1 \end{aligned}$$

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 ∞.

Tools for proving consistency

- Use definition (complicated)

Tools for proving consistency

- Use definition (complicated)
- Chebychev's Inequality

$$\begin{aligned}\Pr(|W_n - \tau(\theta)| \geq \epsilon) &= \Pr((W_n - \tau(\theta))^2 \geq \epsilon^2) \\ &\leq \frac{\mathrm{E}[W_n - \tau(\theta)]^2}{\epsilon^2} \\ &= \frac{\mathrm{MSE}(W_n)}{\epsilon^2} = \frac{\mathrm{Bias}^2(W_n) + \mathrm{Var}(W_n)}{\epsilon^2}\end{aligned}$$

Tools for proving consistency

- Use definition (complicated)
- Chebychev's Inequality

$$\begin{aligned} \Pr(|W_n - \tau(\theta)| \geq \epsilon) &= \Pr((W_n - \tau(\theta))^2 \geq \epsilon^2) \\ &\leq \frac{\mathrm{E}[W_n - \tau(\theta)]^2}{\epsilon^2} \\ &= \frac{\mathrm{MSE}(W_n)}{\epsilon^2} = \frac{\mathrm{Bias}^2(W_n) + \mathrm{Var}(W_n)}{\epsilon^2} \end{aligned}$$

Need to show that both $\mathrm{Bias}(W_n)$ and $\mathrm{Var}(W_n)$ converges to zero

Theorem for consistency

Theorem 10.1.3

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

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

for all θ, 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 $\mathrm{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2 < \infty$. Then $\overline{X}_n$ converges in probability to μ.
i.e. $\overline{X}_n \xrightarrow{\mathrm{P}} \mu$.

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

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 θ and g is a continuous function, then $g(W_n)$ is consistent for $g(\theta)$.

Example - Exponential Family

Problem

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

Example - Exponential Family

Problem

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

1. Propose a consistent estimator of the median.

Example - Exponential Family

Problem

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

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 $\Pr(X \leq c)$

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

Consistent estimator of $\Pr(X \leq c)$

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

As $\overline{X}$ is consistent for β, $1 - e^{-c/\beta}$ is continuous function of β.

Consistent estimator of $\Pr(X \leq c)$

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

As $\overline{X}$ is consistent for β, $1 - e^{-c/\beta}$ is continuous function of β.
By continuous mapping Theorem, $g(\overline{X}) = 1 - e^{-c/\overline{X}}$ 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 $Y_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$ where $p = \Pr(X \leq c)$.

Consistent estimator of $\Pr(X \leq c)$ - Alternative Method

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

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

is consistent for p by Law of Large Numbers.

Consistency of MLEs

Theorem 10.1.6 - Consistency of MLEs

Suppose $X_i \overset{\text{i.i.d.}}{\sim} f(x|\theta)$. Let $\hat{\theta}$ be the MLE of θ, and $\tau(\theta)$ be a continuous function of θ. 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) $W_n(\mathbf{X})$ is *asymptotically normal* if

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

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

Asymptotic Normality

Definition: Asymptotic Normality

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

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

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

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

Asymptotic Normality

Definition: Asymptotic Normality

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

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

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

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

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

Central Limit Theorem

Central Limit Theorem

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

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

Central Limit Theorem

Central Limit Theorem

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

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

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

Central Limit Theorem

Central Limit Theorem

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

$$\begin{aligned} \overline{X} &\sim \mathcal{AN}\left(\mu(\theta), \frac{\sigma^2(\theta)}{n}\right) \\ \Leftrightarrow \sqrt{n}\left(\overline{X} - \mu(\theta)\right) &\xrightarrow{\text{d}} \mathcal{N}(0, \sigma^2(\theta)) \end{aligned}$$

Theorem 5.5.17 - Slutsky's Theorem

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

Central Limit Theorem

Central Limit Theorem

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

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

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

Theorem 5.5.17 - Slutsky's Theorem

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

1. $Y_n \cdot X_n \xrightarrow{\text{d}} aX$

Central Limit Theorem

Central Limit Theorem

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

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

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

Theorem 5.5.17 - Slutsky's Theorem

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

1. $Y_n \cdot X_n \xrightarrow{\text{d}} aX$
2. $X_n + Y_n \xrightarrow{\text{d}} X + a$

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

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

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

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

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

is consistent for p. Therefore,

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

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

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

is consistent for p. Therefore,

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

Example

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

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

Example

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

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

By Central Limit Theorem,

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

Example

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

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

By Central Limit Theorem,

$$\begin{aligned} \overline{X}_n &\sim \mathcal{AN}\left(\mu, \frac{\sigma^2}{n}\right) \\ \Leftrightarrow \sqrt{n}(\overline{X} - \mu) &\xrightarrow{\text{d}} \mathcal{N}(0, \sigma^2) \end{aligned}$$

Example

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

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

By Central Limit Theorem,

$$\begin{aligned} \overline{X}_n &\sim \mathcal{AN}\left(\mu, \frac{\sigma^2}{n}\right) \\ \Leftrightarrow \sqrt{n}(\overline{X} - \mu) &\xrightarrow{\text{d}} \mathcal{N}(0, \sigma^2) \\ \Leftrightarrow \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} &\xrightarrow{\text{d}} \mathcal{N}(0, 1) \end{aligned}$$

Example (cont'd)

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

Example (cont'd)

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

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

Example (cont'd)

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

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

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

Delta Method - Example

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

Delta Method - Example

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

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

Delta Method - Example

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

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

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

Delta Method - Example

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

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

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

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

Delta Method - Example

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

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

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

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

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

Delta Method - Example

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

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

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

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,

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

Delta Method - Example

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

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

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

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,

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

$$= \mathcal{AN}\left(p(1-p), (1-2p)^2 \frac{p(1-p)}{n}\right)$$

Asymptotic Normality

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

Asymptotic Normality

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

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{d} \mathcal{N}(0, \nu(\theta)) \quad \text{for all } \theta$$
$$\iff W_n \sim \mathcal{AN}\left(\tau(\theta), \frac{\nu(\theta)}{n}\right)$$

Asymptotic Normality

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

$$\sqrt{n}(W_n - \tau(\theta)) \xrightarrow{\text{d}} \mathcal{N}(0, \nu(\theta)) \quad \text{for all } \theta$$
$$\iff W_n \sim \mathcal{AN}\left(\tau(\theta), \frac{\nu(\theta)}{n}\right)$$

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

Asymptotic Normality

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

$$\begin{aligned}\sqrt{n}(W_n - \tau(\theta)) &\xrightarrow{d} \mathcal{N}(0, \nu(\theta)) \quad \text{for all } \theta \\ \iff W_n &\sim \mathcal{AN}\left(\tau(\theta), \frac{\nu(\theta)}{n}\right)\end{aligned}$$

Tools to show asymptotic normality

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

Using Central Limit Theorem

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

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

Using Central Limit Theorem

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

where $\mu(\theta) = \mathrm{E}(X)$, and $\sigma^2(\theta) = \mathrm{Var}(X)$.
For example, in order to get the asymptotic distribution of $\frac{1}{n}\sum_{i=1}^n X_i^2$,

Using Central Limit Theorem

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

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

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

Using Central Limit Theorem

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

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

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n X_i^2 &= \frac{1}{n}\sum_{i=1}^n Y_i = \overline{Y} \\ &\sim \mathcal{AN}\left(\mathrm{E}\,Y, \frac{\mathrm{Var}(Y)}{n}\right) \end{aligned}$$

Using Central Limit Theorem

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

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

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n X_i^2 &= \frac{1}{n}\sum_{i=1}^n Y_i = \overline{Y} \\ &\sim \mathcal{AN}\left(\mathrm{E}\,Y, \frac{\mathrm{Var}(Y)}{n}\right) \\ &\sim \mathcal{AN}\left(\mathrm{E}X^2, \frac{\mathrm{Var}(X^2)}{n}\right)\end{aligned}$$

Using Slutsky Theorem

When $X_n \xrightarrow{\text{d}} X$, $Y_n \xrightarrow{\text{P}} a$, then

1. $Y_n X_n \xrightarrow{\text{d}} aX$
2. $X_n + Y_n \xrightarrow{\text{d}} X + a$.

Using Delta Method (Theorem 5.5.24)

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

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

Example

Problem

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

Find the asymptotic distribution of MLE of μ^2.

Example

Problem

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

Find the asymptotic distribution of MLE of μ^2.

Solution

1. It can be easily shown that MLE of μ is $\overline{X}$.

Example

Problem

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

Find the asymptotic distribution of MLE of μ^2.

Solution

1. It can be easily shown that MLE of μ is $\overline{X}$.
2. By the invariance property of MLE, MLE of μ^2 is $\overline{X}^2$.

Example

Problem

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

Find the asymptotic distribution of MLE of μ^2.

Solution

1. It can be easily shown that MLE of μ is $\overline{X}$.
2. By the invariance property of MLE, MLE of μ^2 is $\overline{X}^2$.
3. By central limit theorem, we know that

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

Solution (cont'd)

4. 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.

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

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

Solution (cont'd)

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

$$\begin{aligned} g'(y) &= 2y \\ \overline{X}^2 &\sim \mathcal{AN}\left(g(\mu), [g'(\mu)]^2 \frac{\sigma^2}{n}\right) \\ &\sim \mathcal{AN}\left(\mu^2, (2\mu)^2 \frac{\sigma^2}{n}\right) \end{aligned}$$

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

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

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

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

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

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

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

$$\mathrm{ARE}(V_n, W_n) = \frac{\sigma_W^2}{\sigma_V^2}$$

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

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

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

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

$$\mathrm{ARE}(V_n, W_n) = \frac{\sigma_W^2}{\sigma_V^2}$$

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

Example

Problem

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

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

Example

Problem

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

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

Our estimators are

$$\begin{aligned} W_n &= \frac{1}{n}\sum_{i=1}^{n} I(X_i = 0) \\ V_n &= e^{-\overline{X}} \end{aligned}$$

Example

Problem

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

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

Our estimators are

$$\begin{aligned} W_n &= \frac{1}{n} \sum_{i=1}^{n} I(X_i = 0) \\ V_n &= e^{-\overline{X}} \end{aligned}$$

Determine which one is more asymptotically efficient estimator.

Solution - Asymptotic Distribution of V_n

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

Solution - Asymptotic Distribution of V_n

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

$$\overline{X} \sim \mathcal{AN}(\mathrm{E}X, \mathrm{Var}X/n) \sim \mathcal{AN}(\lambda, \lambda/n)$$

Solution - Asymptotic Distribution of V_n

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

$$\overline{X} \sim \mathcal{AN}(\mathrm{E}X, \mathrm{Var}X/n) \sim \mathcal{AN}(\lambda, \lambda/n)$$

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

Solution - Asymptotic Distribution of V_n

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

$$\overline{X} \sim \mathcal{AN}(\mathrm{E}X, \mathrm{Var}X/n) \sim \mathcal{AN}(\lambda, \lambda/n)$$

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

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

Solution - Asymptotic Distribution of V_n

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

$$\overline{X} \sim \mathcal{AN}(\mathrm{E}X, \mathrm{Var}X/n) \sim \mathcal{AN}(\lambda, \lambda/n)$$

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

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

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 $Z_i = I(X_i = 0)$

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

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

Solution - Asymptotic Distribution of W_n

Define $Z_i = I(X_i = 0)$

$$
\begin{aligned}
W_n &= \frac{1}{n}\sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n \\
Z_i &\sim \text{Bernoulli}(\text{E}(Z)) \\
\text{E}(Z) &= \Pr(X = 0) = e^{-\lambda}
\end{aligned}
$$

Solution - Asymptotic Distribution of W_n

Define $Z_i = I(X_i = 0)$

$$\begin{aligned} W_n &= \frac{1}{n}\sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n \\ Z_i &\sim \mathrm{Bernoulli}(\mathrm{E}(Z)) \\ \mathrm{E}(Z) &= \Pr(X = 0) = e^{-\lambda} \\ \mathrm{Var}(Z) &= e^{-\lambda}(1 - e^{-\lambda}) \end{aligned}$$

Solution - Asymptotic Distribution of W_n

Define $Z_i = I(X_i = 0)$

$$\begin{aligned} W_n &= \frac{1}{n}\sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n \\ Z_i &\sim \text{Bernoulli}(\text{E}(Z)) \\ \text{E}(Z) &= \Pr(X = 0) = e^{-\lambda} \\ \text{Var}(Z) &= e^{-\lambda}(1 - e^{-\lambda}) \end{aligned}$$

By CLT,

$$W_n = \overline{Z}_n \sim \mathcal{AN}\left(\text{E}(Z), \text{Var}(Z)/n\right)$$

Solution - Asymptotic Distribution of W_n

Define $Z_i = I(X_i = 0)$

$$\begin{aligned} W_n &= \frac{1}{n}\sum_{i=1}^{n} I(X_i = 0) = \overline{Z}_n \\ Z_i &\sim \text{Bernoulli}(\text{E}(Z)) \\ \text{E}(Z) &= \Pr(X = 0) = e^{-\lambda} \\ \text{Var}(Z) &= e^{-\lambda}(1 - e^{-\lambda}) \end{aligned}$$

By CLT,

$$\begin{aligned} W_n = \overline{Z}_n &\sim \mathcal{AN}\left(\text{E}(Z), \text{Var}(Z)/n\right) \\ &\sim \mathcal{AN}\left(e^{-\lambda}, \frac{e^{-\lambda}(1 - e^{-\lambda})}{n}\right) \end{aligned}$$

Solution - Calculating ARE

$$\mathrm{ARE}(W_n, V_n) = \frac{e^{-2\lambda}\lambda/n}{e^{-\lambda}(1 - e^{-\lambda})/n}$$

Solution - Calculating ARE

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

Solution - Calculating ARE

$$
\begin{aligned}
\text{ARE}(W_n, V_n) &= \frac{e^{-2\lambda}\lambda/n}{e^{-\lambda}(1-e^{-\lambda})/n} \\
&= \frac{\lambda}{e^{\lambda}(1-e^{-\lambda})} \\
&= \frac{\lambda}{e^{\lambda}-1}
\end{aligned}
$$

Solution - Calculating ARE

$$
\begin{aligned}
\text{ARE}(W_n, V_n) &= \frac{e^{-2\lambda}\lambda/n}{e^{-\lambda}(1-e^{-\lambda})/n} \\
&= \frac{\lambda}{e^{\lambda}(1-e^{-\lambda})} \\
&= \frac{\lambda}{e^{\lambda}-1} \\
&= \frac{\lambda}{\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{3!}+\cdots\right)-1}
\end{aligned}
$$

Solution - Calculating ARE

$$\begin{aligned}
\text{ARE}(W_n, V_n) &= \frac{e^{-2\lambda}\lambda/n}{e^{-\lambda}(1-e^{-\lambda})/n} \\
&= \frac{\lambda}{e^{\lambda}(1-e^{-\lambda})} \\
&= \frac{\lambda}{e^{\lambda}-1} \\
&= \frac{\lambda}{\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{3!}+\cdots\right)-1} \\
&\leq 1 \qquad (\forall \lambda \geq 0)
\end{aligned}$$

Solution - Calculating ARE

$$\begin{aligned}
\text{ARE}(W_n, V_n) &= \frac{e^{-2\lambda}\lambda/n}{e^{-\lambda}(1-e^{-\lambda})/n} \\
&= \frac{\lambda}{e^{\lambda}(1-e^{-\lambda})} \\
&= \frac{\lambda}{e^{\lambda}-1} \\
&= \frac{\lambda}{\left(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{3!}+\cdots\right)-1} \\
&\leq 1 \qquad (\forall \lambda \geq 0)
\end{aligned}$$

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 $\theta \in \Omega$,

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

Asymptotic Efficiency

Definition : Asymptotic Efficiency for iid samples

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

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

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

Asymptotic Efficiency

Definition : Asymptotic Efficiency for iid samples

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

$$\begin{aligned}
\sqrt{n}(W_n - \tau(\theta)) &\xrightarrow{d} \mathcal{N}\left(0, \frac{[\tau'(\theta)]^2}{I(\theta)}\right) \\
\iff W_n &\sim \mathcal{AN}\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)}\right) \\
I(\theta) &= E\left[\left\{\frac{\partial}{\partial\theta}\log f(X|\theta)\right\}^2 |\theta\right]
\end{aligned}$$

Asymptotic Efficiency

Definition : Asymptotic Efficiency for iid samples

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

$$\begin{aligned} \sqrt{n}(W_n - \tau(\theta)) &\xrightarrow{d} \mathcal{N}\left(0, \frac{[\tau'(\theta)]^2}{I(\theta)}\right) \\ \iff W_n &\sim \mathcal{AN}\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nI(\theta)}\right) \\ I(\theta) &= E\left[\left\{\frac{\partial}{\partial\theta}\log f(X|\theta)\right\}^2 |\theta\right] \\ &= -E\left[\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)|\theta\right] \text{ (if interchangeability holds)} \end{aligned}$$

Asymptotic Efficiency

Definition : Asymptotic Efficiency for iid samples

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

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

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

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

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

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 θ. Under same regularity conditions, $\hat{\theta}$ is consistent and asymptotically normal for θ, i.e.

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

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 θ. Under same regularity conditions, $\hat{\theta}$ is consistent and asymptotically normal for θ, i.e.

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

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

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

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

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 θ. Under same regularity conditions, $\hat{\theta}$ is consistent and asymptotically normal for θ, i.e.

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

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

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

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

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