Lecture 18: Delta Method

Mathematical Statistics I, MATH 60061/70061

Thursday November 18, 2021

Reference: Casella & Berger, 5.5.4

Continuity

- Suppose $X_n \xrightarrow{p} X$, as $n \to \infty$ and let $h : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then $h(X_n)$ converges in probability to $h(X)$.
- Suppose $X_n \xrightarrow{a.s.} X$, as $n \to \infty$ and let $h : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then $h(X_n)$ converges almost surely to $h(X)$.
- Suppose $X_n \xrightarrow{d} X$, as $n \to \infty$ and let $h : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then $h(X_n)$ converges in distribution to $h(X)$.

Estimating the odds

Suppose $X_1, X_2, \ldots, X_n$ are iid $\mathrm{Bern}(p)$ random variables, where $0 < p < 1$. The CLT gives

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, p(1-p)),$$

as $n \to \infty$.

A popular parameter is $\frac{p}{1-p}$, the **odds**. In biostatistics application, if the data represent the outcomes of a medical treatment with $p = 2/3$, then a person has odds $2 : 1$ of getting better.

If there is another treatment with success probability r, the **odds ratio**, $\frac{p}{1-p} / \frac{r}{1-r}$ is often estimated, giving the relative odds of one treatment over the other.

Estimating the odds

Suppose $X_1, X_2, \ldots, X_n$ are iid $\mathrm{Bern}(p)$ random variables, where $0 < p < 1$. The CLT gives

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, p(1-p)),$$

as $n \to \infty$.

A popular parameter is $\frac{p}{1-p}$, the **odds**. In biostatistics application, if the data represent the outcomes of a medical treatment with $p = 2/3$, then a person has odds $2:1$ of getting better.

If there is another treatment with success probability r, the **odds ratio**, $\frac{p}{1-p} / \frac{r}{1-r}$ is often estimated, giving the relative odds of one treatment over the other.

Typically, we would use $\frac{\hat{p}}{1-\hat{p}}$ as an estimate of $\frac{p}{1-p}$. How can we find its sampling distribution?

Slutsky's Theorem

Suppose that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} a$, where a is a constant. Then

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

Delta Method

Suppose X_n is a sequence of random variables that satisfy

$$\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

as $n \to \infty$. For a given function g, suppose that $g'(\theta)$ exists and $g'(\theta) \neq 0$. Then

$$\sqrt{n}[g(X_n) - g(\theta)] \xrightarrow{d} \mathcal{N}(0, [g'(\theta)]^2\sigma^2),$$

as $n \to \infty$.

In other words, the distribution of $g(X_n)$ can be approximated by

$$\mathcal{N}\left(g(\theta), \frac{[g'(\theta)]^2\sigma^2}{n}\right)$$

for large n.

Expanding $g(X_n)$ in a Taylor series around θ, we have

$$g(X_n) = g(\theta) + g'(\theta)(X_n - \theta) + \frac{g''(\xi)}{2}(X_n - \theta)^2,$$

where ξ is between X_n and θ. Multiplying by $\sqrt{n}$ and then rearranging, we have

$$\sqrt{n}[g(X_n) - g(\theta)] = g'(\theta)\sqrt{n}(X_n - \theta) + \frac{\sqrt{n}g''(\xi)}{2}(X_n - \theta)^2,$$

where $g'(\theta)\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, [g'(\theta)]^2\sigma^2)$.

Expanding $g(X_n)$ in a Taylor series around θ, we have

$$g(X_n) = g(\theta) + g'(\theta)(X_n - \theta) + \frac{g''(\xi)}{2}(X_n - \theta)^2,$$

where ξ is between X_n and θ. Multiplying by $\sqrt{n}$ and then rearranging, we have

$$\sqrt{n}[g(X_n) - g(\theta)] = g'(\theta)\sqrt{n}(X_n - \theta) + \frac{\sqrt{n}g''(\xi)}{2}(X_n - \theta)^2,$$

where $g'(\theta)\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, [g'(\theta)]^2\sigma^2)$.

If we can show

$$R_n = \frac{\sqrt{n}g''(\xi)}{2}(X_n - \theta)^2 \xrightarrow{p} 0,$$

then the result will follow from Slutsky's Theorem.

Rewrite the remainder term:

$$R_n = \frac{g''(\xi)}{2}(X_n - \theta)\sqrt{n}(X_n - \theta),$$

where $\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$. Provided that $g''(\xi)$ converges to something finite, if we can show $X_n - \theta \xrightarrow{p} 0$, then we have $R_n \xrightarrow{p} 0$.

Rewrite the remainder term:

$$R_n = \frac{g''(\xi)}{2}(X_n - \theta)\sqrt{n}(X_n - \theta),$$

where $\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$. Provided that $g''(\xi)$ converges to something finite, if we can show $X_n - \theta \xrightarrow{p} 0$, then we have $R_n \xrightarrow{p} 0$.

For every $\epsilon > 0$, consider

$$\lim_{n\to\infty} P(|X_n - \theta| \geq \epsilon) = \lim_{n\to\infty} P(\sqrt{n}|X_n - \theta| \geq \sqrt{n}\epsilon).$$

Since $\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$,

$$\sqrt{n}|X_n - \theta| = |\sqrt{n}(X_n - \theta)| \xrightarrow{d} |X|,$$

where $X \sim \mathcal{N}(0, \sigma^2)$, by continuity.

The distribution of $|X|$ does not have probability/density going to ∞, so

$$\lim_{n\to\infty} P(|X_n - \theta| \geq \epsilon) = \lim_{n\to\infty} P(\sqrt{n}|X_n - \theta| \geq \sqrt{n}\epsilon) = 0.$$

Therefore, $X_n \xrightarrow{p} \theta$, and by continuity, $X_n - \theta \xrightarrow{p} 0$.

Applying Slutsky's Theorem, we have the Delta Method:

$$\sqrt{n}[g(X_n) - g(\theta)] \xrightarrow{d} \mathcal{N}(0, [g'(\theta)]^2\sigma^2),$$

as $n \to \infty$

Variance of odds estimator

Suppose $X_1, X_2, \ldots, X_n$ are iid $\mathrm{Bern}(p)$ random variables, where $0 < p < 1$. Using $\frac{\hat{p}}{1-\hat{p}}$ as an estimate of the odds $\frac{p}{1-p}$, what is the variance of the estimate?

Variance of odds estimator

Suppose $X_1, X_2, \ldots, X_n$ are iid $\mathrm{Bern}(p)$ random variables, where $0 < p < 1$. Using $\frac{\hat{p}}{1-\hat{p}}$ as an estimate of the odds $\frac{p}{1-p}$, what is the variance of the estimate?

The CLT gives

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, p(1-p)), \quad \text{when } n \to \infty$$

Take $g(p) = \frac{p}{1-p}$, so $g'(p) = \frac{1}{(1-p)^2}$. The Delta Method says that

$$\begin{aligned} \mathrm{Var}\left(\frac{\hat{p}}{1-\hat{p}}\right) &\approx [g'(p)]^2 \mathrm{Var}(\hat{p}) \\ &= \left[\frac{1}{(1-p)^2}\right]^2 \frac{p(1-p)}{n} = \frac{p}{n(1-p)^3}. \end{aligned}$$

Second-order Delta Method

Suppose X_n is a sequence of random variables that satisfy

$$\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

as $n \to \infty$. For a given function g, suppose that g is twice differentiable at θ, $g'(\theta) = 0$ and $g''(\theta) \neq 0$. Then

$$n[g(X_n) - g(\theta)] \xrightarrow{d} \sigma^2 \frac{g''(\theta)}{2} \chi_1^2,$$

as $n \to \infty$.

The proof is similar to that for the first-order Delta Method.

Convergence of the sample mean squared

Suppose $X_1, \ldots, X_n$ are iid random variables with $E(X_1) = \mu$ and $\mathrm{Var}(X_1) = \sigma^2 < \infty$. The CLT gives

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

as $n \to \infty$. Consider $g(\bar{X}_n) = \bar{X}_n^2$. With $g(\mu) = \mu^2$, we have $g'(\mu) = 2\mu$, which is nonzero except when $\mu = 0$.

Convergence of the sample mean squared

Suppose $X_1, \ldots, X_n$ are iid random variables with $E(X_1) = \mu$ and $\mathrm{Var}(X_1) = \sigma^2 < \infty$. The CLT gives

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

as $n \to \infty$. Consider $g(\bar{X}_n) = \bar{X}_n^2$. With $g(\mu) = \mu^2$, we have $g'(\mu) = 2\mu$, which is nonzero except when $\mu = 0$.

Convergence of $\bar{X}_n^2$:

- If $\mu \neq 0$, by the first-order Delta Method

$$\sqrt{n}(\bar{X}_n^2 - \mu^2) \xrightarrow{d} \mathcal{N}(0, 4\mu^2\sigma^2).$$

- If $\mu = 0$, the result with the first-order Delta Method collapses. We can then apply the second-order Delta Method

$$n(\bar{X}_n^2 - \mu^2) = n(\bar{X}_n^2) \xrightarrow{d} \sigma^2\chi_1^2.$$

Multivariate extensions

All convergence concepts can be extended to handle sequences of random variables.

Central Limit Theorem: Suppose $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$, is a sequence of iid random vectors (of dimension k) with $E(\boldsymbol{X}_1) = \boldsymbol{\mu}_{k\times 1}$ and $\mathrm{Cov}(\boldsymbol{X}_1) = \boldsymbol{\Sigma}_{k\times k}$. Let $\bar{\boldsymbol{X}}_n = (\bar{X}_{1+}, \bar{X}_{2+}, \ldots, \bar{X}_{k+})'$ denote the vector of sample means. Then $\sqrt{n}(\bar{\boldsymbol{X}}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathrm{MVN}_k(\mathbf{0}, \boldsymbol{\Sigma})$.

Multivariate Delta Method: Suppose $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$, is a sequence of iid random vectors (of dimension k) that satisfy $\sqrt{n}(\boldsymbol{X}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathrm{MVN}_k(\mathbf{0}, \boldsymbol{\Sigma})$. For a given function $g : \mathbb{R}^k \to \mathbb{R}$, suppose that g is differentiable at $\boldsymbol{\mu}$ and is not zero. Then

$$\sqrt{n}[g(\boldsymbol{X}_n) - g(\boldsymbol{\mu})] \xrightarrow{d} \mathcal{N}\left(0, \frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{x}} \boldsymbol{\Sigma} \frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{x}'}\right)$$

where

$$\frac{\partial g(\boldsymbol{\mu})}{\partial \boldsymbol{x}} = \left(\frac{\partial g(\boldsymbol{x})}{\partial x_1}, \ldots, \frac{\partial g(\boldsymbol{x})}{\partial x_k}\right)\Big|_{\boldsymbol{x}=\boldsymbol{\mu}}.$$