STSCI 5080
Probability Models and Inference
Lecture 22: Confidence Intervals

November 15, 2018

Confidence intervals

Definition

Suppose that θ is one-dim. and let $\alpha \in (0, 1)$.

- A data dependent interval $[A_n, B_n]$, where $A_n = A_n(X_1, \ldots, X_n)$ and $B_n = B_n(X_1, \ldots, X_n)$, is a confidence interval (CI) with level $1 - \alpha$ for θ if

$$\underbrace{P_\theta(A_n \le \theta \le B_n)}_{\text{coverage probability}} \ge 1 - \alpha$$

  for any $\theta \in \Theta$.

- The interval $[A_n, B_n]$ is a confidence interval with asymptotic level $1 - \alpha$ for θ if

$$\lim_{n \to \infty} P_\theta(A_n \le \theta \le B_n) \ge 1 - \alpha$$

  for any $\theta \in \Theta$.

Rule of thumb

We should construct a CI $[A_n, B_n]$ in such a way that

$$P_\theta(A_n \leq \theta \leq B_n) = 1 - \alpha \quad (*)$$

for any θ, or

$$\lim_{n\to\infty} P_\theta(A_n \leq \theta \leq B_n) = 1 - \alpha$$

for any θ if the requirement (*) is too stringent.

Example 22.1

Let

$$X_1, \ldots, X_n \sim N(\mu, \sigma_0^2) \text{ i.i.d.}$$

where μ is unknown but σ_0^2 is known. The MLE is

$$\widehat{\mu} = \overline{X} \sim N(\mu, \sigma_0^2/n) \quad \text{i.e.} \quad \frac{\sqrt{n}(\widehat{\mu} - \mu)}{\sigma_0} = Z \sim N(0, 1).$$

We note that

$$P_\mu \left\{ \left| \frac{\sqrt{n}(\widehat{\mu} - \mu)}{\sigma_0} \right| \leq z \right\} = P(|Z| \leq z) = 2\Phi(z) - 1,$$

where $\Phi(z)$ is the cdf of $N(0, 1)$. In addition, we note that

$$\left| \frac{\sqrt{n}(\widehat{\mu} - \mu)}{\sigma_0} \right| \leq z \Leftrightarrow \widehat{\mu} - \frac{z\sigma_0}{\sqrt{n}} \leq \mu \leq \widehat{\mu} + \frac{z\sigma_0}{\sqrt{n}}.$$

We should choose z in such a way that

$$2\Phi(z) - 1 = 1 - \alpha \quad \text{i.e.} \quad z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2).$$

For example,

$$z_{\alpha/2} \approx \begin{cases} 1.96 & \text{if } \alpha = 0.05 \\ 2.58 & \text{if } \alpha = 0.01 \end{cases}.$$

CI for μ with level $1 - \alpha$

A CI for μ with level $1 - \alpha$ is given by

$$\left[\widehat{\mu} - \frac{z_{\alpha/2}\sigma_0}{\sqrt{n}}, \widehat{\mu} + \frac{z_{\alpha/2}\sigma_0}{\sqrt{n}}\right].$$

If $\alpha = 0.05$, this CI will be

$$\left[\widehat{\mu} - \frac{1.96\sigma_0}{\sqrt{n}}, \widehat{\mu} + \frac{1.96\sigma_0}{\sqrt{n}}\right].$$

General case

Suppose that

$$\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2(\theta)) \quad \text{as } n \to \infty$$

for some $\sigma^2(\theta) > 0$ (asymptotic variance). This implies

$$Z_n = \frac{\sqrt{n}(\widehat{\theta} - \theta)}{\sigma(\theta)} \xrightarrow{d} Z \sim N(0, 1).$$

So,

$$P_\theta(|Z_n| \leq z) \approx P(|Z| \leq z) = 2\Phi(z) - 1.$$

If we choose $z = z_{\alpha/2}$, then

$$P_\theta(|Z_n| \leq z_{\alpha/2}) \approx 1 - \alpha.$$

We note that

$$|Z_n| \le z_{\alpha/2} \Leftrightarrow \widehat{\theta} - \frac{z_{\alpha/2}\sigma(\theta)}{\sqrt{n}} \le \theta \le \widehat{\theta} + \frac{z_{\alpha/2}\sigma(\theta)}{\sqrt{n}},$$

which implies that

$$P_\theta \left\{ \widehat{\theta} - \frac{z_{\alpha/2}\sigma(\theta)}{\sqrt{n}} \le \theta \le \widehat{\theta} + \frac{z_{\alpha/2}\sigma(\theta)}{\sqrt{n}} \right\} \approx 1 - \alpha.$$

We are tempted to use

$$\left[\widehat{\theta} - \frac{z_{\alpha/2}\sigma(\theta)}{\sqrt{n}}, \widehat{\theta} + \frac{z_{\alpha/2}\sigma(\theta)}{\sqrt{n}} \right]$$

as a CI, but

$\sigma(\theta)$ **is in general unknown!**

Implication of Slutsky theorem

Theorem

If $\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$ *and* $\sigma^2(\theta)$ *is continuous in* θ*, then*

$$\frac{\sqrt{n}(\widehat{\theta} - \theta)}{\sigma(\widehat{\theta})} \xrightarrow{d} N(0, 1).$$

Implication of Slutsky theorem

Theorem

If $\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$ *and* $\sigma^2(\theta)$ *is continuous in* θ*, then*

$$\frac{\sqrt{n}(\widehat{\theta} - \theta)}{\sigma(\widehat{\theta})} \xrightarrow{d} N(0, 1).$$

Instead of $\sigma(\theta)$, use $\sigma(\widehat{\theta})$!

Recap

CI for θ with level $1-\alpha$

A CI for θ with asymptotic level $1-\alpha$ is given by

$$\left[\widehat{\theta} - \frac{z_{\alpha/2}\sigma(\widehat{\theta})}{\sqrt{n}}, \widehat{\theta} + \frac{z_{\alpha/2}\sigma(\widehat{\theta})}{\sqrt{n}}\right].$$

If $\alpha = 0.05$, this CI will be

$$\left[\widehat{\theta} - \frac{1.96\sigma(\widehat{\theta})}{\sqrt{n}}, \widehat{\theta} + \frac{1.96\sigma(\widehat{\theta})}{\sqrt{n}}\right].$$

Example 22.2

Example

Let

$$X_1, \ldots, X_n \sim Po(\lambda) \text{ i.i.d.}$$

for some $\lambda > 0$. The MLE is $\widehat{\lambda} = \overline{X}$, and by CLT

$$\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda).$$

So a CI for λ with asymptotic level 95% is given by

$$\left[\widehat{\lambda} - \frac{1.96\sqrt{\widehat{\lambda}}}{\sqrt{n}}, \widehat{\lambda} + \frac{1.96\sqrt{\widehat{\lambda}}}{\sqrt{n}}\right].$$

Example 22.3

Example

Let

$$X \sim Bin(n,p)$$

for some $0 < p < 1$. The MLE is $\widehat{p} = X/n$, and by CLT

$$\sqrt{n}(\widehat{p} - p) \xrightarrow{d} N(0, p(1-p)) \quad \text{as } n \to \infty.$$

So a CI for p with asymptotic level 95% is given by

$$\left[\widehat{p} - \frac{1.96\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}, \widehat{p} + \frac{1.96\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}\right].$$

See also Brown, Cai, and DasGupta (2001, Statistical Science) for a discussion on this CI.

How to determine the sample size?

- CIs can be used to determine the sample size.
- The basic idea is to determine the sample size in such a way that the radius of the CI is less than a desired accuracy.

Example 22.3 (cont.)

- Suppose that you are interested in the proportion p of the population who supports Party A.
- You draw a sample of n potential voters and X of them support Party A.
- A CI for p with asymptotic level 95% is given by

$$\left[\widehat{p} - \frac{1.96\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}, \widehat{p} + \frac{1.96\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}\right].$$

- You can determine the sample size n in such a way that

the radius of the CI $\leq$ a desired accuracy.

- For example, suppose that we choose n in such a way that

$$\text{the radius of the CI} \leq 0.01.$$

- The radius is

$$\frac{1.96\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}} \leq \frac{1.96}{2\sqrt{n}}$$

where we have used the fact that $\widehat{p}(1-\widehat{p}) \leq 1/4$.

- So it is enough to take n in such a way that

$$\frac{1.96}{2\sqrt{n}} \leq 0.01 \Leftrightarrow n \geq 98^2 = 9604.$$

Variance stabilizing transformation

Definition

Suppose that $\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$ and let $g(\theta)$ be a smooth function in θ. The function $g(\theta)$ is called a variance stabilizing transformation for $\widehat{\theta}$ if

$$\sqrt{n}\{g(\widehat{\theta}) - g(\theta)\} \xrightarrow{d} N(0, 1).$$

Example 22.3

Let

$$X_1, \ldots, X_n \sim Po(\lambda) \text{ i.i.d.}$$

for some $\lambda > 0$. The MLE is $\widehat{\lambda} = \overline{X}$, and by CLT

$$\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda).$$

Consider the function $g(\lambda) = 2\sqrt{\lambda}$. Since $g'(\lambda) = 1/\sqrt{\lambda}$, we have

$$\sqrt{n}(2\sqrt{\widehat{\lambda}} - 2\sqrt{\lambda}) \xrightarrow{d} N(0, (1/\lambda) \times \lambda) = N(0, 1).$$

So a CI for $2\sqrt{\lambda}$ with asymptotic level $1 - \alpha$ is given by

$$\left[2\sqrt{\widehat{\lambda}} - \frac{z_{\alpha/2}}{\sqrt{n}}, 2\sqrt{\widehat{\lambda}} + \frac{z_{\alpha/2}}{\sqrt{n}}\right].$$

Now, we note that

$$
\begin{aligned}
2\sqrt{\lambda} &\in \left[2\sqrt{\widehat{\lambda}} - \frac{z_{\alpha/2}}{\sqrt{n}}, 2\sqrt{\widehat{\lambda}} + \frac{z_{\alpha/2}}{\sqrt{n}}\right] \\
\Leftrightarrow \lambda &\in \left[\left(\sqrt{\widehat{\lambda}} - \frac{z_{\alpha/2}}{2\sqrt{n}}\right)^2, \left(\sqrt{\widehat{\lambda}} + \frac{z_{\alpha/2}}{2\sqrt{n}}\right)^2\right].
\end{aligned}
$$

So, the following is also a CI for λ with asymptotic level $1-\alpha$:

$$
\left[\left(\sqrt{\widehat{\lambda}} - \frac{z_{\alpha/2}}{2\sqrt{n}}\right)^2, \left(\sqrt{\widehat{\lambda}} + \frac{z_{\alpha/2}}{2\sqrt{n}}\right)^2\right].
$$

General case

In general, suppose that $\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$. Then

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

by the delta method. So it is enough to choose $g(\theta)$ to satisfy

$$\{g'(\theta)\}^2\sigma^2(\theta) = 1 \quad \text{i.e.} \quad g(\theta) = \underbrace{\int \frac{1}{\sigma(\theta)} d\theta}_{\text{indefinite integral}} .$$

Since this $g(\theta)$ is increasing in θ,

$$\left[g^{-1}\left(g(\widehat{\theta}) - z_{\alpha/2}/\sqrt{n}\right), g^{-1}\left(g(\widehat{\theta}) + z_{\alpha/2}/\sqrt{n}\right)\right]$$

will be a CI for θ with asymptotic level $1 - \alpha$.

Example 22.4

Let

$$X_1, \ldots, X_n \sim Ex(\lambda) \text{ i.i.d.}$$

for some $\lambda > 0$. The MLE is $\widehat{\lambda} = 1/\overline{X}$, and

$$\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda^2).$$

1. Find a variance stabilizing transformation for $\widehat{\lambda}$.
2. Use the variance stabilizing transformation derived in Part (a) to find a CI for λ with asymptotic level $1 - \alpha$.

1. The variance stabilizing transformation is

$$g(\lambda) = \int \frac{1}{\lambda} d\lambda = \log \lambda.$$

2. Since $g^{-1}(x) = e^x$,

$$\left[\exp\left(\log \widehat{\lambda} - z_{\alpha/2}/\sqrt{n}\right), \exp\left(\log \widehat{\lambda} + z_{\alpha/2}/\sqrt{n}\right)\right] \qquad (*)$$

   is a CI for λ with asymptotic level $1-\alpha$. We note that

$$(*) = \left[\widehat{\lambda} e^{-z_{\alpha/2}/\sqrt{n}}, \widehat{\lambda} e^{z_{\alpha/2}/\sqrt{n}}\right].$$

Recap

We have discussed two ways to find CIs for MLEs:

1. The first method is to plug in $\widehat{\theta}$ for $\sigma(\theta)$ and use

$$\left[\widehat{\theta} - \frac{z_{\alpha/2}\sigma(\widehat{\theta})}{\sqrt{n}}, \widehat{\theta} + \frac{z_{\alpha/2}\sigma(\widehat{\theta})}{\sqrt{n}}\right].$$

2. The second method is to find a variance stabilizing transformation $g(\theta)$ and use

$$\left[g^{-1}\left(g(\widehat{\theta}) - z_{\alpha/2}/\sqrt{n}\right), g^{-1}\left(g(\widehat{\theta}) + z_{\alpha/2}/\sqrt{n}\right)\right].$$