Lecture 24: Asymptotic Evaluations of Interval Estimators

Mathematical Statistics II, MATH 60062/70062

Thursday April 28, 2022

Reference: Casella & Berger, 10.4

Recap: Large sample tests

Suppose $X_1, \dots, X_n$ are iid from $f_X(x \mid \theta)$, where $\theta \in \Theta \subseteq \mathbb{R}$. Suppose the regularity conditions needed for MLEs to be consistent and asymptotically normal hold. We have discussed three large sample procedures to test

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0.$$

- **Wald test:**

$$Z_n^W = \frac{\hat{\theta} - \theta_0}{\sqrt{\frac{v(\hat{\theta})}{n}}} = \frac{\hat{\theta} - \theta_0}{\sqrt{\frac{1}{nI_1(\hat{\theta})}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

- **Score test:**

$$Z_n^S = \frac{S(\theta_0 \mid \boldsymbol{X})}{\sqrt{I_n(\theta_0)}} \xrightarrow{d} \mathcal{N}(0, 1)$$

- **LRT:**

$$-2 \log \lambda(\boldsymbol{X}) = -2[\log L(\theta_0 \mid \boldsymbol{X}) - \log L(\hat{\theta} \mid \boldsymbol{X})] \xrightarrow{d} \chi_1^2.$$

Large sample confidence intervals

We will introduce three “large sample likelihood based confidence intervals.”

- Wald intervals
- Score intervals
- Likelihood ratio intervals

Large sample pivot

Suppose $X_1, \ldots, X_n$ are iid from $f_X(x \mid \theta)$, where $\theta \in \Theta \subseteq \mathbb{R}$.
The random variable

$$Q_n = Q_n(\boldsymbol{X}, \theta)$$

is called a **large sample pivot** if its asymptotic distribution is free of all unknown parameters.

If Q_n is a large sample pivot and if

Large Sample Approximation

$$P_\theta(Q_n(\boldsymbol{X}, \theta) \in \mathcal{A}) \approx 1 - \alpha,$$

then $C(\boldsymbol{X}) = \{\theta : Q_n(\boldsymbol{X}, \theta) \in \mathcal{A}\}$ is called an approximate $1 - \alpha$ confidence set for θ.

Wald intervals

Suppose $X_1, \ldots, X_n$ are iid from $f_X(x \mid \theta)$, where $\theta \in \Theta \subseteq \mathbb{R}$. Under certain regularity conditions, an MLE $\hat{\theta}$ satisfies

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

where

$$v(\theta) = \frac{1}{I_1(\theta)}.$$

If $v(\theta)$ is a continuous function of θ, then $v(\hat{\theta}) \xrightarrow{p} v(\theta)$, for all θ. By Slutsky's Theorem,

$$Q_n(\boldsymbol{X}, \theta) = \frac{\hat{\theta} - \theta}{\sqrt{\frac{v(\hat{\theta})}{n}}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Therefore, $Q_n(\boldsymbol{X}, \theta)$ is a large sample pivot.

Write

$$
\begin{aligned}
1-\alpha &\approx P_\theta(-z_{\alpha/2} \le Q_n(\boldsymbol{X}, \theta) \le z_{\alpha/2}) \\
&= P_\theta\left(-z_{\alpha/2} \le \frac{\hat{\theta}-\theta}{\sqrt{\frac{v(\hat{\theta})}{n}}} \le z_{\alpha/2}\right) \\
&= P_\theta\left(\hat{\theta} - z_{\alpha/2}\sqrt{\frac{v(\hat{\theta})}{n}} \le \theta \le \hat{\theta} + z_{\alpha/2}\sqrt{\frac{v(\hat{\theta})}{n}}\right).
\end{aligned}
$$

Therefore,

$$
\hat{\theta} \pm z_{\alpha/2}\sqrt{\frac{v(\hat{\theta})}{n}}
$$

is an approximate $1-\alpha$ confidence interval for θ. We call it a **Wald confidence interval**.

This type of large sample interval is called a **Wald confidence interval** as it can also be obtained by inverting the large sample test of

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0,$$

using the Wald test statistic

$$Z_n^W = \frac{\hat{\theta} - \theta_0}{\sqrt{\frac{v(\hat{\theta})}{n}}}$$

and rejection region

$$R = \{\boldsymbol{x} \in \mathcal{X} : |z_n^W| \geq z_{\alpha/2}\}.$$

Wald confidence intervals for functions of θ

With an MLE $\hat{\theta}$, if $g : \mathbb{R} \to \mathbb{R}$ is differentiable at θ and $g'(\theta) \neq 0$, then

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} \mathcal{N}\left(0, [g'(\theta)]^2 v(\theta)\right).$$

If $[g'(\theta)]^2 v(\theta)$ is a continuous function of θ, then $[g'(\hat{\theta})]^2 v(\hat{\theta})$ is a consistent estimator for it, because MLEs are consistent and consistency is preserved under continuous mappings. Therefore,

$$Q_n(\boldsymbol{X}, \theta) = \frac{g(\hat{\theta}) - g(\theta)}{\sqrt{\frac{[g'(\hat{\theta})]^2 v(\hat{\theta})}{n}}} \xrightarrow{d} \mathcal{N}(0, 1),$$

by Slutsky's Theorem and

$$g(\hat{\theta}) \pm z_{\alpha/2} \sqrt{\frac{[g'(\hat{\theta})]^2 v(\hat{\theta})}{n}}$$

is an approximate $1 - \alpha$ confidence interval for $g(\theta)$.

Bernoulli Wald confidence interval

Suppose $X_1, \ldots, X_n$ are iid $\mathrm{Bern}(p)$, where $0 < p < 1$.
The MLE of p is given by

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

In Lecture 22, we showed that

$$v(p) = \frac{1}{I_1(p)} = p(1-p).$$

Therefore,

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

is an approximate $1 - \alpha$ Wald confidence interval for p.

Wald confidence interval for log odd

Suppose $X_1, \ldots, X_n$ are iid $\mathrm{Bern}(p)$, where $0 < p < 1$. Derive a $1 - \alpha$ Wald confidence interval for

$$g(p) = \log\left(\frac{p}{1-p}\right),$$

the log odds of p.

Wald confidence interval for log odd

Suppose $X_1, \ldots, X_n$ are iid $\mathrm{Bern}(p)$, where $0 < p < 1$. Derive a $1 - \alpha$ Wald confidence interval for

$$g(p) = \log\left(\frac{p}{1-p}\right),$$

the log odds of p.

First note that $g(p) = \log[p/(1-p)]$ is a differentiable function and

$$g'(p) = \frac{1}{p(1-p)} \neq 0.$$

The Delta Method gives

$$\sqrt{n}\left[\log\left(\frac{\hat{p}}{1-\hat{p}}\right) - \log\left(\frac{p}{1-p}\right)\right] \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{p(1-p)}\right).$$

Because the asymptotic variance $1/p(1-p)$ can be consistently estimated by $1/\hat{p}(1-\hat{p})$, we have

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

by Slutsky's Theorem.

Therefore,

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) \pm z_{\alpha/2}\sqrt{\frac{1}{n\hat{p}(1-\hat{p})}}$$

is an approximate $1-\alpha$ Wald confidence interval for the log odd $g(p) = \log[p/(1-p)]$.

Score confidence intervals

Suppose $X_1, \ldots, X_n$ are iid from $f_X(x \mid \theta)$, where $\theta \in \Theta \subseteq \mathbb{R}$. Consider testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0.$$

Under H_0, the score statistic

$$Q_n(\boldsymbol{X}, \theta_0) = \frac{S(\theta_0 \mid \boldsymbol{X})}{\sqrt{I_n(\theta_0)}} \xrightarrow{d} \mathcal{N}(0,1),$$

where $I_n(\theta) = nI_1(\theta)$ is the Fisher information based on the sample.

Therefore,

$$R = \{\boldsymbol{x} \in \mathcal{X} : |Q_n(\boldsymbol{X}, \theta_0)| \geq z_{\alpha/2}\}$$

is an approximate size α rejection region, and the acceptance region is

$$A = R^c = \{\boldsymbol{x} \in \mathcal{X} : |Q_n(\boldsymbol{X}, \theta_0)| < z_{\alpha/2}\}.$$

Inverting this acceptance region,

$$C(\boldsymbol{x}) = \{\theta : |Q_n(\boldsymbol{X}, \theta)| < z_{\alpha/2}\}$$

is an approximate $1 - \alpha$ confidence set for θ. If $C(\boldsymbol{x})$ is an interval, then we call it a **score confidence interval**.

Bernoulli score confidence interval

Suppose $X_1, \ldots, X_n$ are iid $\mathrm{Bern}(p)$, where $0 < p < 1$.

In Lecture 22, we showed that

$$Q_n(\boldsymbol{X}, p) = \frac{S(p \mid \boldsymbol{X})}{I_n(p)} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}.$$

Therefore,

$$C(\boldsymbol{X}) = \{p : |Q_n(\boldsymbol{X}, p)| < z_{\alpha/2}\} = \left\{ p : \left| \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \right| < z_{\alpha/2} \right\}$$

forms the score interval for p. With $\boldsymbol{X} = \boldsymbol{x}$, note the boundary

$$Q_n(\boldsymbol{x}, p) = z_{\alpha/2} \iff (\hat{p} - p)^2 = z_{\alpha/2}^2 \frac{p(1-p)}{n}.$$

Upon rearrangement, we get

$$\left(1+\frac{z_{\alpha/2}^2}{n}\right)p^2-\left(2\hat{p}+\frac{z_{\alpha/2}^2}{n}\right)p+\hat{p}^2=0.$$

Using the quadratic formula, the lower and upper limits are

$$p_L=\frac{(2\hat{p}+z_{\alpha/2}^2/n)-\sqrt{(2\hat{p}+z_{\alpha/2}^2/n)^2-4(1+z_{\alpha/2}^2/n)\hat{p}^2}}{2(1+z_{\alpha/2}^2/n)}$$

$$p_U=\frac{(2\hat{p}+z_{\alpha/2}^2/n)+\sqrt{(2\hat{p}+z_{\alpha/2}^2/n)^2-4(1+z_{\alpha/2}^2/n)\hat{p}^2}}{2(1+z_{\alpha/2}^2/n)},$$

respectively.

Likelihood ratio confidence intervals

Suppose $X_1, \ldots, X_n$ are iid from $f_X(x \mid \theta)$, where $\theta \in \Theta \subseteq \mathbb{R}$. Consider testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0.$$

The LRT statistic is

$$\lambda(\boldsymbol{x}) = \frac{L(\theta_0 \mid \boldsymbol{x})}{L(\hat{\theta} \mid \boldsymbol{x})}$$

and

$$R = \{\boldsymbol{x} \in \mathcal{X} : -2 \log \lambda(\boldsymbol{x}) \geq \chi^2_{1,\alpha}\}$$

is an approximate size α rejection region for testing H_0 versus H_1. Inverting the acceptance region,

$$C(\boldsymbol{x}) = \left\{ \theta : -2 \log \left[\frac{L(\theta \mid \boldsymbol{x})}{L(\hat{\theta} \mid \boldsymbol{x})} \right] < \chi^2_{1,\alpha} \right\}$$

is an approximate $1 - \alpha$ confidence set for θ. If $C(\boldsymbol{x})$ is an interval, then we call it a **likelihood ratio confidence interval**.

Bernoulli likelihood ratio confidence interval

Suppose $X_1, \ldots, X_n$ are iid $\mathrm{Bern}(p)$, where $0 < p < 1$.

In Lecture 23, we showed that

$$-2\log\left[\frac{L(p \mid \boldsymbol{x})}{L(\hat{p} \mid \boldsymbol{x})}\right] = -2\left[n\hat{p}\log\left(\frac{p}{\hat{p}}\right) + n(1-\hat{p})\log\left(\frac{1-p}{1-\hat{p}}\right)\right].$$

Therefore, the confidence interval is

$$C(\boldsymbol{x}) = \left\{p : -2\left[n\hat{p}\log\left(\frac{p}{\hat{p}}\right) + n(1-\hat{p})\log\left(\frac{1-p}{1-\hat{p}}\right)\right] < \chi^2_{1,\alpha}\right\}.$$

This interval can be calculated using numerical search methods.