STSCI 5080
Probability Models and Inference

Lecture 23: Testing

November 27, 2018

Two ways to construct CIs for MLEs

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

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 $g(\theta)$ such that $\sqrt{n}\{g(\widehat{\theta}) - g(\theta)\} \xrightarrow{d} N(0, 1)$ and use

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

Discussion on $z_{\alpha/2}$

- $z_{\alpha/2}$ is defined by
$$z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2),$$
where Φ is the cdf of $N(0,1)$.

- Alternatively, $z_{\alpha/2}$ is the point satisfying
$$P(|Z| \le z_{\alpha/2}) = 1 - \alpha,$$
where $Z \sim N(0,1)$.

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

List of CIs for MLEs based on the first method

- $N(\mu, \sigma_0^2)$ with known σ_0^2:

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

  where $\widehat{\mu} = \overline{X}$.

- $Po(\lambda)$:

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

  where $\widehat{\lambda} = \overline{X}$.

List of CIs for MLEs based on the first method

- $N(0, \theta)$:

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

  where $\widehat{\theta} = n^{-1}\sum_{i=1}^{n} X_i^2$.

- $Ex(\lambda)$:

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

  where $\widehat{\lambda} = 1/\overline{X}$.

List of CIs for MLEs based on the first method

- $Bin(n,p)$ ($X \sim Bin(n,p)$ is given):

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

  where $\widehat{p} = X/n$.

List of CIs for MLEs based on the second method

- $N(0,\theta)$: $g(\theta) = \frac{1}{\sqrt{2}} \log \theta$ and

$$\left[\widehat{\theta} e^{-z_{\alpha/2}\sqrt{2}/\sqrt{n}}, \widehat{\theta} e^{z_{\alpha/2}\sqrt{2}/\sqrt{n}}\right].$$

- $Po(\lambda)$: $g(\lambda) = 2\sqrt{\lambda}$ and

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

- $Ex(\lambda)$: $g(\lambda) = \log \lambda$ and

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

Chapter 9 Testing Hypotheses and Assessing Goodness of Fit

Baaack to the very first example

- There is a theory that people can postpone their death until after an important event.
- To test the theory, Phillips and Smith[1] (1990) collected data on deaths around some (important!) festival for a certain group of people.
- Of 103 deaths, 33 died the week before the festival and 70 died the week after.

[1] D.P. Phillips and D.G. Smith. (1990). "Postponement of death until symbolically meaningful occasions". *JAMA* **263** 1947-1951.

- Suppose that each person dies after the festival with probability p.
- The total number of deaths after the festival X follows $Bin(n, p)$ where $n = 103$.
- In this example, $X = 70$, and so the MLE of p is

$$\widehat{p} = \frac{X}{n} = \frac{70}{103} = 0.68...$$

- If they can postpone their deaths, $p > 0.5$; otherwise $p = 0.5$.
- We want to verify that $p = 0.5$ or $p > 0.5$. This can be formulated as a statistical testing problem:

$$H_0 : p = 0.5 \quad \text{vs.} \quad H_1 : p > 0.5.$$

- H_0 is called the null hypothesis, and H_1 is called the alternative hypothesis.
- We want to decide whether to reject H_0 based on the data.

General formulation

- Let $\{f_\theta \mid \theta \in \Theta\}$ with $\Theta \subset \mathbb{R}$ be a statistical model and suppose that we have
$$X_1, \ldots, X_n \sim f_\theta \text{ i.i.d.}$$
for some $\theta \in \Theta$.

- Let θ_0 be a single point in Θ, and $\Theta_1 \subset \Theta$ a set that does not contain θ_0: $\theta_0 \notin \Theta_1$. We know θ_0 and Θ_1.

- Consider the testing problem:
$$H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta \in \Theta_1.$$

- We want to decide whether to reject H_0 based on the data.

Test statistic

- We will construct a statistic $T_n = T_n(X_1, \ldots, X_n)$ (called a test statistic) such that
$$T_n > c \Rightarrow \text{reject } H_0, \qquad (*)$$
where c is some threshold (non-random constant).
- The procedure (*) is called a test.

Type I and II errors

- Type I error: rejecting H_0 when it is actually true. The probability of the type I error is
$$P_{\theta=\theta_0}(T_n > c).$$
- Type II error: not rejecting H_0 although H_1 is actually true.

How to choose c?

- The test (*) is said to have level α if

$$P_{\theta=\theta_0}(T_n > c) \leq \alpha.$$

  In words, we choose c in such a way that the probability of the type I error is at most α.

- The test is said to have asymptotic level α if

$$\lim_{n\to\infty} P_{\theta=\theta_0}(T_n > c) \leq \alpha.$$

- A common choice of α: $\alpha = 0.05$ or 0.01.

Trivial test

- If you choose c to be very large then you always do not reject H_0. In such a case, the probability of the type I error is zero, but you always make the type II error.
- In practice, we should choose c in such a way that

$$P_{\theta=\theta_0}(T_n > c) = \alpha$$

or

$$\lim_{n\to\infty} P_{\theta=\theta_0}(T_n > c) = \alpha.$$

Two sided alternative hypothesis

- Consider the testing problem:

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

Suppose that the MLE $\widehat{\theta}$ is such that

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

for any value of θ.

- Consider the statistic

$$T_n = \frac{\sqrt{n}(\widehat{\theta} - \theta_0)}{\sigma(\theta_0)}.$$

Note: θ_0 and not θ!

If $\theta = \theta_0$, then

$$T_n \xrightarrow{d} N(0, 1).$$

On the other hand, if $\theta \neq \theta_0$, then

$$|T_n| \xrightarrow{P} \infty.$$

So, a reasonable test will be

$$|T_n| > c \Rightarrow \text{ reject } H_0.$$

The threshold c is chosen in such a way that

$$\lim_{n\to\infty} P_{\theta=\theta_0}(|T_n| > c) = \alpha.$$

If $\theta = \theta_0$, then $T_n \xrightarrow{d} Z \sim N(0,1)$, and so

$$P_{\theta=\theta_0}(|T_n| > c) \approx P(|Z| > c) = 1 - P(|Z| \le c).$$

Solving

$$1 - P(|Z| \le c) = \alpha, \text{ i.e.}, P(|Z| \le c) = 1 - \alpha,$$

we have

$$c = z_{\alpha/2}.$$

If $\alpha = 0.05$, we can choose $z_{\alpha/2} = 1.96$.

Recap

For the testing problem,

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

a test with asymptotic level α is given by

$$\left| \frac{\sqrt{n}(\widehat{\theta} - \theta_0)}{\sigma(\theta_0)} \right| > z_{\alpha/2} \Rightarrow \text{reject } H_0.$$

If $\alpha = 0.05$, we can choose $z_{\alpha/2} = 1.96$.

Example 23.1

Let

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

and consider the testing problem

$$H_0 : \lambda = 1 \quad \text{vs.} \quad H_1 : \lambda \neq 1.$$

The MLE is $\widehat{\lambda} = \overline{X}$ and $\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda)$. Hence, a test with asymptotic level α is given by

$$|\sqrt{n}(\widehat{\lambda} - 1)| > z_{\alpha/2} \Rightarrow \text{reject } H_0.$$

Why does this test have asymptotic level α?
If $\lambda = 1$,

$$\sqrt{n}(\widehat{\lambda} - 1) \xrightarrow{d} Z \sim N(0, 1),$$

and so

$$P_{\lambda=1}\left\{\left|\sqrt{n}(\widehat{\lambda} - 1)\right| > z_{\alpha/2}\right\} \approx P(|Z| > z_{\alpha/2}) = \alpha.$$

Example 23.2

Let

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

where $0 < p < 1$ is unknown. Consider the testing problem

$$H_0 : p = p_0 \quad \text{vs.} \quad H_1 : p \neq p_0$$

for some known p_0 (e.g. $p_0 = 0.5$). The MLE is $\widehat{p} = X/n$ and $\sqrt{n}(\widehat{p} - p) \xrightarrow{d} N(0, p(1-p))$. Hence, a test with asymptotic level α is given by

$$\left| \frac{\sqrt{n}(\widehat{p} - p_0)}{\sqrt{p_0(1-p_0)}} \right| > z_{\alpha/2} \Rightarrow \text{reject } H_0.$$

Why does this test have asymptotic level α?

Final exam

- The coverage of the final exam is the material covered in Lectures 18-24, Discussion sessions during the same period, and the homeworks.
- On the Thursday lecture, I will talk more about testing (for about 30 minutes). In addition, I will discuss practice problems.
- Questions in the final exam are straightforward, and Problem 5 in the practice final exam will be asked in the actual exam. Be prepared:)
- You can bring notes and textbooks (but no electronic devices are allowed to use). Don't forget to bring the solutions of Discussion problems. If you are well prepared, then you can get full mark without much difficulty.