Math 408 - Mathematical Statistics

Lecture 26-27. Confidence Intervals from MLEs

April 5-8, 2013

Agenda

- Exact Method
  - Normal distribution $N(\mu, \sigma^2)$
- Approximate Method
  - $\mathrm{Bernoulli}(p)$
- Bootstrap Method
- Summary

Three Methods

Recall the definition of a confidence interval (see also Lectures 8,17,23):

Definition

A $100(1-\alpha)\%$ **confidence interval** for a parameter θ is a random interval calculated from the sample,

$$X_1, \ldots, X_n \sim \pi(x|\theta)$$

which contains θ with probability $1-\alpha$.

There are three methods for constructing confidence intervals using MLEs $\hat{\theta}_{\mathrm{MLE}}$:

- Exact Method
- Approximate Method
- Bootstrap Method

Exact Method. Example: Normal distribution $\mathcal{N}(\mu, \sigma^2)$

Let $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$, then the MLEs for μ and σ^2 are (Lecture 24):

$$\hat{\mu}_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n \qquad \hat{\sigma}^2_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

- A confidence interval for μ is based on the following fact (Lecture 13-14):

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \sim t_{n-1}$$

where S_n^2 is the sample variance $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \frac{n}{n-1}\hat{\sigma}^2_{\mathrm{MLE}}$

Result

A $100(1-\alpha)\%$ confidence interval for μ is

$$\hat{\mu}_{\mathrm{MLE}} \pm \frac{1}{\sqrt{n-1}}\hat{\sigma}_{\mathrm{MLE}} t_{n-1}(\alpha/2)$$

where $t_{n-1}(\alpha)$ is the point beyond which the t-distribution with $(n-1)$ degrees of freedom has probability α.

Exact Method. Example: Normal distribution $N(\mu, \sigma^2)$

- A confidence interval for σ^2 is based on the following fact (Lecture 13-14):

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$$

Result

A $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\left(\frac{n\hat{\sigma}^2_{\mathrm{MLE}}}{\chi^2_{n-1}(\frac{\alpha}{2})}, \frac{n\hat{\sigma}^2_{\mathrm{MLE}}}{\chi^2_{n-1}(1-\frac{\alpha}{2})}\right)$$

where $\chi^2_{n-1}(\alpha)$ is the point beyond which the χ^2-distribution with $(n-1)$ degrees of freedom has probability α.

Remark:

The main drawback of the exact method is that in practice the sampling distributions — like t_{n-1} and χ^2_{n-1} in our example — are not known.

Approximate Method

One of the most important properties of MLE is that it is asymptotically normal:

$$\hat{\theta}_{\mathrm{MLE}} \to \mathcal{N}\left(\theta_0, \frac{1}{nI(\theta_0)}\right), \qquad \text{as } n \to \infty$$

where $I(\theta_0)$ is Fisher information

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

Since the true value θ_0 is unknown, we will use $I(\hat{\theta}_{\mathrm{MLE}})$ instead of $I(\theta_0)$:

Result

An **approximate** $100(1-\alpha)\%$ *confidence interval for* θ_0 *is*

$$\hat{\theta}_{\mathrm{MLE}} \pm \frac{z_{\alpha/2}}{\sqrt{nI(\hat{\theta}_{\mathrm{MLE}})}}$$

where z_α *is the point beyond which the standard normal distribution has probability* α*.*

Approximate Method. Example: $\mathrm{Bernoulli}(p)$

- Let $X_1, \ldots, X_n \sim \mathrm{Bernoulli}(p)$.
  Find an approximate confidence interval for p
- Answer:

$$\overline{X}_n \pm z_{\alpha/2}\sqrt{\frac{\overline{X}_n(1-\overline{X}_n)}{n}}$$

Bootstrap Method

Suppose $\hat{\theta}$ is an estimate of a parameter θ, the true unknown value of which is θ_0. $\hat{\theta}$ can be any estimate, not necessarily MLE,

$$X_1, \ldots, X_n \sim \pi(x|\theta) \qquad \hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$$

Define a new random variable

$$\Delta = \hat{\theta} - \theta_0$$

- Step 1: Assume (for the moment) that the distribution of Δ is known. Let (as before) q_α be the number such that $\mathbb{P}(\Delta > q_\alpha) = \alpha$.Then

  $$\mathbb{P}(q_{1-\frac{\alpha}{2}} \leq \hat{\theta} - \theta_0 \leq q_{\frac{\alpha}{2}}) = 1 - \alpha$$

  And therefore a $100(1-\alpha)\%$ confidence interval for θ_0 is

  $$\left(\hat{\theta} - q_{\frac{\alpha}{2}}, \hat{\theta} - q_{1-\frac{\alpha}{2}}\right)$$

  The problem is that the distribution of Δ is not known and, therefore, q_α are not known.

Bootstrap Method

- Step 2: Assume that the distribution of Δ is not known, but θ_0 is known. Then we can approximate the distribution of Δ as follows:

$$
\begin{aligned}
X_1^{(1)}, \dots, X_n^{(1)} \sim \pi(x|\theta_0) \quad &\rightsquigarrow \quad \hat{\theta}^{(1)} - \theta_0 = \Delta^{(1)} \\
X_1^{(2)}, \dots, X_n^{(2)} \sim \pi(x|\theta_0) \quad &\rightsquigarrow \quad \hat{\theta}^{(2)} - \theta_0 = \Delta^{(2)} \\
&\cdots\cdots \\
X_1^{(B)}, \dots, X_n^{(B)} \sim \pi(x|\theta_0) \quad &\rightsquigarrow \quad \hat{\theta}^{(B)} - \theta_0 = \Delta^{(B)}
\end{aligned}
$$

From these realizations $\Delta^{(1)}, \dots, \Delta^{(B)}$ of Δ we can approximate the distribution of Δ by its empirical distribution, and, therefore, we can approximate q_α. The problem is that θ_0 is not known!

Bootstrap Method

- Step 3: **Bootstrap strategy**: Use $\hat{\theta}$ instead of θ_0.

$$X_1^{(1)}, \dots, X_n^{(1)} \sim \pi(x|\theta_0) \quad \rightsquigarrow \quad \hat{\theta}^{(1)} - \hat{\theta} \approx \Delta^{(1)}$$
$$X_1^{(2)}, \dots, X_n^{(2)} \sim \pi(x|\theta_0) \quad \rightsquigarrow \quad \hat{\theta}^{(2)} - \hat{\theta} \approx \Delta^{(2)}$$
$$\dots\dots$$
$$X_1^{(B)}, \dots, X_n^{(B)} \sim \pi(x|\theta_0) \quad \rightsquigarrow \quad \hat{\theta}^{(B)} - \hat{\theta} \approx \Delta^{(B)}$$

Distribution of Δ is approximated from realizations $\Delta^{(1)}, \dots, \Delta^{(B)}$.

Remark:

$\hat{\theta}^{(i)}$ is the estimate of θ that is obtained from $X_1^{(i)}, \dots, X_n^{(i)}$ by the same method (for example, MLE) as $\hat{\theta}$ was obtained from $X_1, \dots, X_n$.

Summary

- We considered three methods for constructing confidence intervals using MLEs: Exact Method, Approximate Method, Bootstrap Method
- Exact Method provides exact confidence intervals, but it is difficult to use in practice
  - Example: $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$

$$\mu : \quad \hat{\mu}_{\mathrm{MLE}} \pm \frac{1}{\sqrt{n-1}} \hat{\sigma}^2_{\mathrm{MLE}} t_{n-1}(\alpha/2)$$

$$\sigma^2 : \quad \left(\frac{n\hat{\sigma}^2_{\mathrm{MLE}}}{\chi^2_{n-1}(\frac{\alpha}{2})}, \frac{n\hat{\sigma}^2_{\mathrm{MLE}}}{\chi^2_{n-1}(1-\frac{\alpha}{2})} \right)$$

- Approximate method provides an approximate confidence interval for θ_0, which is constructed using asymptotical properties of MLE:

$$\hat{\theta}_{\mathrm{MLE}} \pm \frac{z_{\alpha/2}}{\sqrt{nI(\hat{\theta}_{\mathrm{MLE}})}}$$

- Bootstrap Method provides an approximate confidence interval. Bootstrap is the most popular method in practice since it is easy to implement.