Lec 8. Interval Estimation

Zhang W.P.

Contents

8.1 Confidence Interval 2
8.1.1 The pivotal method 4
8.1.2 The asymptotic method 5
8.1.3 Other methods 6

Learning objectives

After this section, you should be able to ...

- understand the confidence interval
- understand the pivotal method
- understand the asymptotic method

Statistical Estimation

- **Point Estimation:** using the data to calculate a single estimate of the parameter of interest. For example, we often use the sample mean $\bar{x}$ to estimate the population mean μ.
- **Interval Estimation:** provides a range of values (an interval) that may contain the unknown parameter (such as the population mean μ).

8.1 Confidence Interval

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a sample from population $\{f(x \mid \theta), \theta \in \Theta\}$ with support $\mathcal{X}$. Let $L_n(\boldsymbol{X})$ and $U_n(\boldsymbol{X})$ be statistics such that $L_n(x) \leq U_n(x), \forall x \in \mathcal{X}$. Then $[L_n(\boldsymbol{X}), U_n(\boldsymbol{X})]$ is a $100(1-\alpha)\%$ **confidence interval (CI)** for θ if

$$P_\theta \left(L_n(\boldsymbol{X}) \leq \theta \leq U_n(\boldsymbol{X}) \right) \geq 1 - \alpha$$

Def.

for all $\theta \in \Theta$. The interval $(L_n(\boldsymbol{x}), U_n(\boldsymbol{x}))$ is a **large sample** (asymptotic) $100(1-\alpha)\%$ CI for θ if

$$P_\theta \left(L_n(\boldsymbol{X}) < \theta < U_n(\boldsymbol{X}) \right) \to 1 - \alpha$$

for all $\theta \in \Theta$ as $n \to \infty$. α is called confidence level.

Remark:

- It will sometimes be more natural to use an open interval $(L_n(x), U_n(x))$ or even a half-open and half-closed interval
- The interval can also be one-sided interval: $(-\infty, U_n(X)]$ or $[L_n(X), \infty)$
- This definition can be extended to multi-dimensional cases: 2D, confidence band; $\geq$ 3D, confidence region.
- Naturally, we want our set to have **small size** and **large coverage probability**, but such sets are usually difficult to construct.
- A typical CI form of θ is $\hat{\theta} \pm sd \times F_q$, where F_q is the quantile and sd is the standard error of $\hat{\theta}$. The value $sd \times qf$ is called the margin of error

Find the CI

8.1.1 The pivotal method

The random variable $g(\boldsymbol{X}, \theta)$ is a pivot or pivotal quantity if its distribution is independent θ.

Def.

- For pivotal quantity $g(\boldsymbol{X}, \theta)$, find $a < b$ such that

$$P_\theta(a \leq g(\boldsymbol{X}, \theta) \leq b) \geq 1 - \alpha$$

- Re-write the interval as

$$C(\boldsymbol{x}) = \{\theta : a \leq g(\boldsymbol{x}, \theta) \leq b\} = \{\theta : L_n(\boldsymbol{x}, a) \leq \theta \leq U_n(\boldsymbol{x}, b)\}$$

and $C(\boldsymbol{X})$ is a $1 - \alpha$ confidence set for θ.

Find the CI

8.1.2 The asymptotic method

The quantity $g(\boldsymbol{X}, \theta)$ is an asymptotic pivot if its limiting distribution is independent of θ.

Def.

A large sample $100(1-\alpha)\%$ CI for θ can be easily obtained.

8.1.3 Other methods

We will discuss the following methods later.

- Inverting a test Statistic
- Bayesian interval