

# 제 16 강

## 추가 주제들-최우추정법 - Maximum Likelihood estimation

### 최우추정법(Maximum Likelihood Estimation Method)

#### 우도함수(Likelihood Function):

(확률)표본의 결합확률 분포를 추정하고자 하는 모수의 함수로 볼 때, 이를 우도함수라 함

- 확률표본  $X_1, X_2, \dots, X_n \sim i.i.d.$  의 pdf가  $f(x_i | \theta_1, \theta_2, \dots, \theta_k)$  로 주어진다고 하면, 그 joint pdf는  $g(x_1, x_2, \dots, x_n | \theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k)$
- 이를 모수  $\theta_1, \theta_2, \dots, \theta_k$  의 함수로 볼 때, 우도 함수라고 함  $L(\theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k)$
- 로그우도 함수 :  $\ln L(\theta_1, \theta_2, \dots, \theta_k) = \sum_{i=1}^n \ln f(x_i | \theta_1, \theta_2, \dots, \theta_k)$

### 최우추정법(Maximum Likelihood Estimation Method)

#### 최우추정법:

우도함수를 극대화 시키는 모수의 값을 해당 모수의 추정치로 삼는 방법이며, 이를 통해 얻어지는 추정량을 최우추정량(Maximum likelihood estimator: MLE)이라고 함

- $(\theta_1, \theta_2, \dots, \theta_k)$  에 대한 최우추정량  $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)$  은 다음과 같이 주어짐  $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k) = \arg \max_{\theta_1, \theta_2, \dots, \theta_k} L(\theta_1, \theta_2, \dots, \theta_k) = \arg \max_{\theta_1, \theta_2, \dots, \theta_k} \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k)$
- 경우에 따라서는 우도함수의 극대화 해(solution)를 찾는 것 보다 우도함수에 자연대수를 취한 로그우도함수를 극대화 하는 해를 찾는 것이 용이함

### 최우추정법(Maximum Likelihood Estimation Method)

#### 최우추정량의 성질:

최우추정량은 점근적 불편성, 일치성, 그리고 점근적 유효성을 갖는다.

- 모수  $\theta$  에 대한 최우추정량을  $\hat{\theta}_n = \hat{\theta}(X_1, X_2, \dots, X_n)$  라고 하면,  $\lim_n E(\hat{\theta}_n) = \theta, \hat{\theta}_n \xrightarrow{p} \theta, \lim_n V(\sqrt{n}\hat{\theta}_n) = \sigma^2(\theta)$
- 단  $\sigma^2(\theta)$  는 추정량의 분산이 도달할 수 있는 이론적 하한(Cramer-Rao lower bound)
- 충분히 큰 n에 대해 다음과 같은 분포로 근사할 수 있음,  $\hat{\theta}_n \sim N\left(\theta, \frac{\sigma^2(\theta)}{n}\right)$

### 최우추정법의 응용(정규분포 확률표본)

- 확률표본이  $X_1, X_2, \dots, X_n \sim i.i.d.N(\mu, \sigma^2)$  로 주어졌을 때 로그우도함수는 다음과 같이 쓸 수 있다.

$$\ln L(\mu, \sigma^2) = \sum_{i=1}^n \ln f(x_i | \mu, \sigma^2) = \sum_{i=1}^n \ln \left[ \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{x_i - \mu}{\sigma} \right)^2 \right\} \right]$$

$$= n \ln \frac{1}{\sigma\sqrt{2\pi}} + \sum_{i=1}^n -\frac{1}{2} \left( \frac{x_i - \mu}{\sigma} \right)^2$$

- 이를  $(\mu, \sigma^2)$  에 대해 극대화

$$\frac{\partial \ln L(\mu, \sigma^2)}{\partial \mu} = \sum_{i=1}^n -\left( \frac{x_i - \mu}{\sigma^2} \right) = 0 \Rightarrow \hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$$

$$\frac{\partial \ln L(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^4} = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

### 최우추정법의 응용(단순선형회귀모형)

- 단순선형회귀모형에서 오차항의 분포를 정규분포라고 하면, 로그우도함수는 다음과 같이 쓸 수 있다.

$$\ln L(\beta_1, \beta_2, \sigma^2) = \sum_{i=1}^n \ln f((y_i, x_i) | \beta_1, \beta_2, \sigma^2) = \sum_{i=1}^n \ln \left[ \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{y_i - \beta_1 - \beta_2 x_i}{\sigma} \right)^2 \right\} \right]$$

$$= n \ln \frac{1}{\sigma\sqrt{2\pi}} + \sum_{i=1}^n -\frac{1}{2} \left( \frac{y_i - \beta_1 - \beta_2 x_i}{\sigma} \right)^2$$

- 이를  $(\beta_1, \beta_2, \sigma^2)$  에 대해 극대화

$$\frac{\partial \ln L(\beta_1, \beta_2, \sigma^2)}{\partial \beta_1} = 0, \frac{\partial \ln L(\beta_1, \beta_2, \sigma^2)}{\partial \beta_2} = 0 \Rightarrow \text{MLE} = \text{LSE}$$

$$\frac{\partial \ln L(\beta_1, \beta_2, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} - \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{2\sigma^4} = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{n}$$

추가적 주제들 - 최우추정법 계량경제학  
16.7

**최우추정법의 응용(이진반응모형, Binary Response Model)**

$$y = \begin{cases} 1 & \text{individual drives to work} \\ 0 & \text{individual takes bus to work} \end{cases}$$

$$f(y) = p^y(1-p)^{1-y}, \quad y = 0, 1 \Rightarrow E(y) = p \Rightarrow$$

$$y = E(y) + \varepsilon = p + \varepsilon$$

We are assuming that the probability of driving is related to the difference in driving times,  $x$ , in the transportation example.

- Linear Probability Model :
 
$$E(y) = p = \beta_1 + \beta_2 x$$

추가적 주제들 - 최우추정법 계량경제학  
16.8

**최우추정법의 응용(이진반응모형)**

Standard Normal cumulative distribution function:

$$F(z) = P[Z \leq z] = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-0.5u^2} du$$

$$p = P[Z \leq \beta_1 + \beta_2 X_{i2}] = F(\beta_1 + \beta_2 X)$$

For  $\beta_2 > 0$ ,  $p$  will approach 1 as  $X \longrightarrow +\infty$

For  $\beta_2 > 0$ ,  $p$  will approach 0 as  $X \longrightarrow -\infty$

추가적 주제들 - 최우추정법 계량경제학  
16.9

**최우추정법의 응용(이진반응모형)**

Using the cumulative distribution function of logistic distribution:

$$p = P[Z \leq \beta_1 + \beta_2 X_{i2}] = F(\beta_1 + \beta_2 X)$$

$$= \frac{1}{1 + e^{-(\beta_1 + \beta_2 X)}}$$

For  $\beta_2 > 0$ ,  $p$  will approach 1 as  $X_{i2} \longrightarrow +\infty$

For  $\beta_2 > 0$ ,  $p$  will approach 0 as  $X_{i2} \longrightarrow -\infty$

추가적 주제들 - 최우추정법 계량경제학  
16.10

**최우추정법의 응용(이진반응모형)**

- 프로빗 모형의 추정 : 최우추정법

예컨대, 세 명에 대해  $y_1 = 1, y_2 = 1$  and  $y_3 = 0$  이 관측되고 이들의 설명변수 값이  $x_1 = 15, x_2 = 20$  and  $x_3 = 5$  관측됨

$\rightarrow y$  의 확률함수는 베르누이 분포로부터 주어지고 이를 프로빗 모형과 결합하면, 확률함수는 다음과 같이 쓸 수 있음

$$f(y_i) = [F(\beta_1 + \beta_2 x_i)]^{y_i} [1 - F(\beta_1 + \beta_2 x_i)]^{1-y_i}, \quad y_i = 0, 1$$

추가적 주제들 - 최우추정법 계량경제학  
16.11

**최우추정법의 응용(이진반응모형)**

$y_1 = 1, y_2 = 1$  and  $y_3 = 0$  를 관측하게 될 확률은?: 세 명이 무작위로 추출될 경우  $y_1, y_2$  및  $y_3$  에 대한 결합확률분포함수는  $f(y_1, y_2, y_3) = f(y_1)f(y_2)f(y_3)$  로 주어짐

$y_1 = 1, y_2 = 1$  및  $y_3 = 0$  이 관찰될 확률은

$$P[y_1 = 1, y_2 = 1, y_3 = 0] = f(1, 1, 0) = f(1)f(1)f(0)$$

$$= F[\beta_1 + \beta_2(15)] \cdot F[\beta_1 + \beta_2(20)] \cdot \{1 - F[\beta_1 + \beta_2(5)]\}$$

: 우도함수(likelihood function)

주어진 표본의 값이 관측될 확률, 즉 우도함수를 극대화하는 값  $b_1$  과  $b_2$  를  $\beta_1$  와  $\beta_2$  의 추정치로 구함