3주차 1차시 : 단순회귀분석(OLS추정량의 특성)

1. 바람직한 추정량이란?

2. OLS 추정량의 특성

   (1) 선형추정량

   (2) 불편추정량

   (3) 최소분산추정량

1.바람직한 추정량이란?

-바람직한 추정량이 갖추어야 할 조건이 많이 있지만 불편성, 효율성, 일치성을 들고 있음

바람직한 추정량의 기준

- 불편성

  $E(\hat{\theta}) = \theta$

- 효율성

  $Var(\hat{\theta})$이작을수록효율적

- 일치성

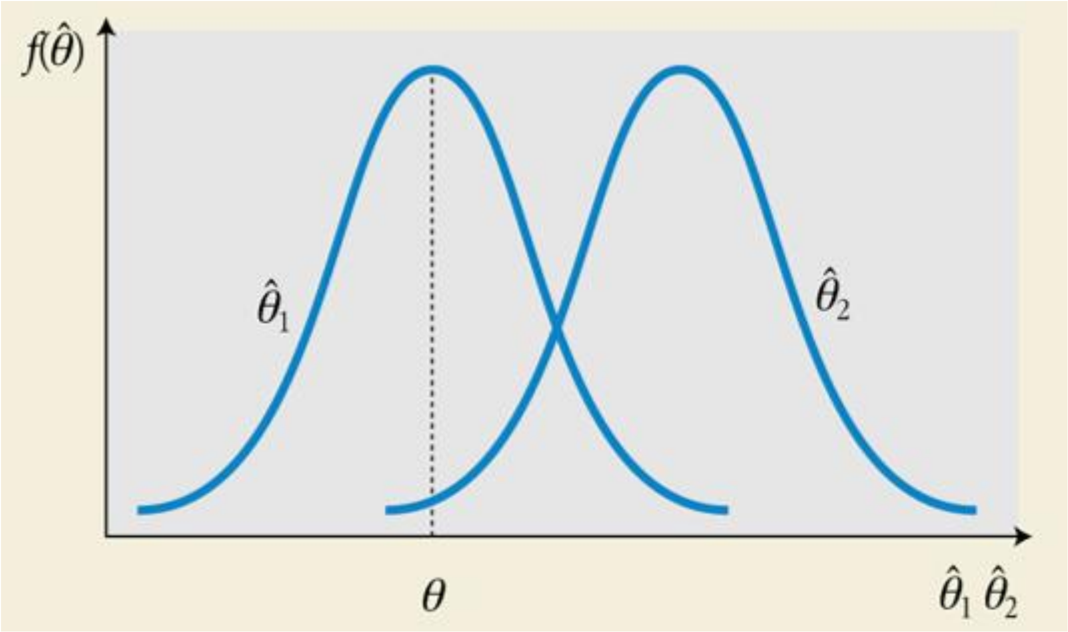
  $p\lim(\hat{\theta}) = \theta$

(1)불편성

-추정량 $\hat{\theta}$의 기댓값이 모수 θ와 일치하면, 즉 $E(\hat{\theta})=\theta$이면 $\hat{\theta}$은 모수 θ의 불편추정량이라고 함

-θ와 $E(\hat{\theta})$의 차이를 편의(bias)라고 하므로 $E(\hat{\theta})=\theta$이면 불편추정량이됨

-그림에서$\hat{\theta}_1$이 θ의 불편추정량임

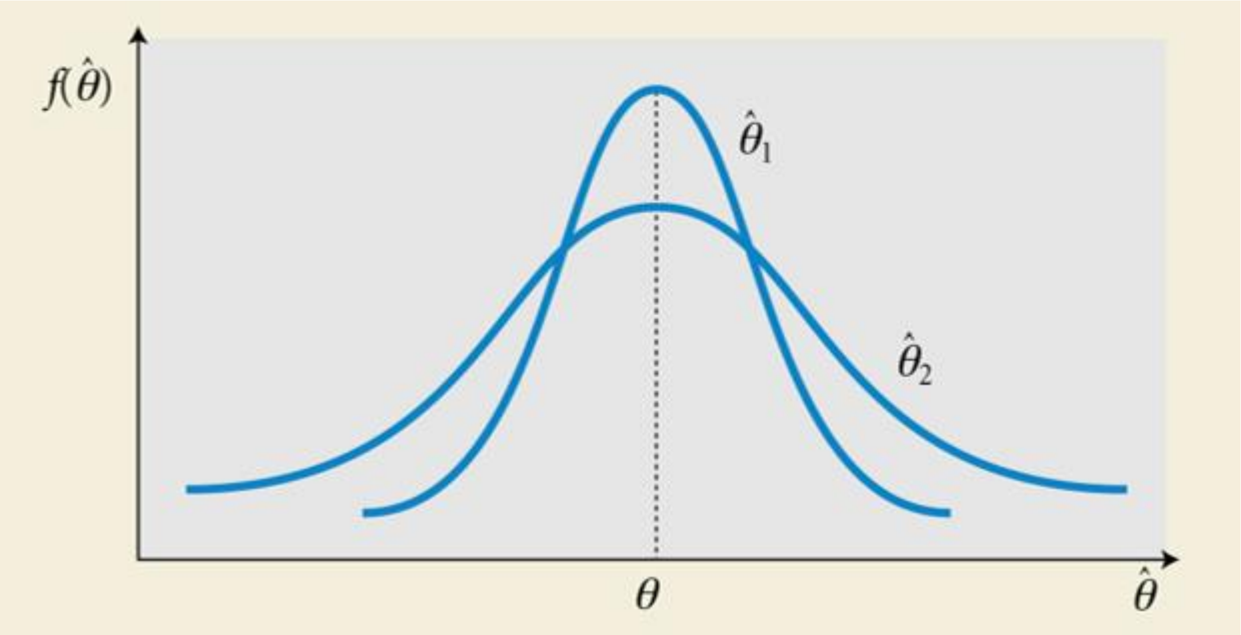


(2)효율성

-$\hat{\theta}_1$과 $\hat{\theta}_2$가 모두 불편추정량일 경우, Var($\hat{\theta}_1$)< Var($\hat{\theta}_2$)이면 $\hat{\theta}_1$이 $\hat{\theta}_2$보다 더 효율적(efficient)이라고 함

-$\hat{\theta}_1$이 다른 모든 불편추정량보다 작은 분산을 가지면 $\hat{\theta}_1$은 가장 효율적인(most efficient) 또는 최소분산(minimum variance) 불편추정량이라고 함

-그림에서$\hat{\theta}_1$이 $\hat{\theta}_2$보다 효율적인 불편추정량임



2.OLS 추정량의 특징

(참고)기대연산자(Expectation Operator) 및 분산연산자(Variance Operator)

A.기댓값(평균)을 계산하는 기대연산자는 다음과 같은 특징을 가지고 있다.
① E(c) = c
② E(X±c) = E(X)±c
③ E(cX) = cE(X)
④ E(aX±b) = aE(X)±b
⑤ E(aX±bY) = aE(X)±bE(Y)

B.분산을 계산하는 분산연산자는 다음과 같은 특징을 가지고 있다.
① var(c) = 0
② var(X) = E(X^2) − $[E(X)]^2$
③ var(cX) = c^2var(X)
④ var(X±c) = var(X)
⑤ var(aX±bY) = a^2var(X)+b^2var(Y)±2ab cov(X,Y)
⑥ var(aX±bY) = a^2var(X)+b^2var(Y) (만약 X,Y가 독립이면)

(1)선형추정량

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i (Y_i - \bar{Y})}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2} - \frac{\bar{Y}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} = \sum_{i=1}^{n} w_i Y_i$$

단, $w_i = \dfrac{x_i}{\sum_{i=1}^{n} x_i^2}$

(2)불편추정량

$$\begin{aligned}\hat{\beta}_1 &= \sum_{i=1}^{n} w_i Y_i \\ &= \sum_{i=1}^{n} w_i \left(\beta_0 + \beta_1 X_i + u_i\right) \\ &= \beta_0 \sum_{i=1}^{n} w_i + \beta_1 \sum_{i=1}^{n} w_i X_i + \sum_{i=1}^{n} w_i u_i \\ &= \beta_1 + \sum_{i=1}^{n} w_i u_i \quad \left(\because \sum_{i=1}^{n} w_i = 0, \sum_{i=1}^{n} w_i X_i = 1\right)\end{aligned}$$

$$\begin{aligned}E(\hat{\beta}_1) &= \beta_1 + E\left(\sum_{i=1}^{n} w_i u_i\right) \\ &= \beta_1 + \sum_{i=1}^{n} w_i E(u_i) \\ &= \beta_1\end{aligned}$$

$$\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \frac{x_i}{\sum_{i=1}^{n} x_i^2}$$

$$= \frac{\sum_{i=1}^{n} (X_i - \bar{X})}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} X_i - n\bar{X}}{\sum_{i=1}^{n} x_i^2}$$

$$= 0$$

$$\sum_{i=1}^{n} w_i X_i = \sum_{i=1}^{n} \frac{x_i (x_i - \bar{X})}{\sum_{i=1}^{n} x_i^2}$$

$$= \sum_{i=1}^{n} \frac{x_i^2}{\sum_{i=1}^{n} x_i^2} - \sum_{i=1}^{n} \frac{x_i \bar{X}}{\sum_{i=1}^{n} x_i^2}$$

$$= \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2}$$

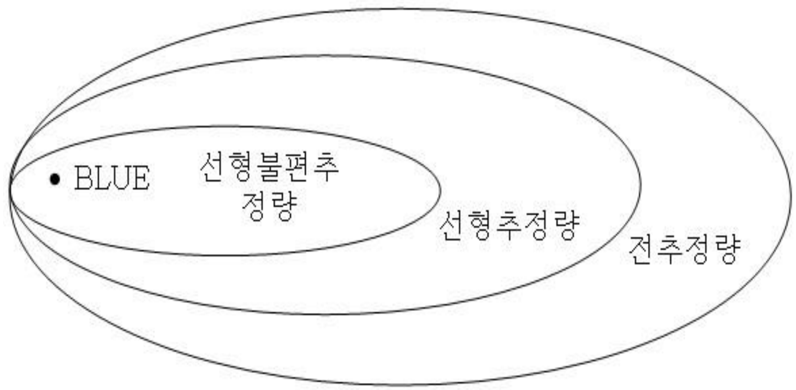
$$= 1$$

(3)최소분산추정량

$\hat{\beta}_1 = \sum_{i=1}^{n} w_i Y_i$

$\tilde{\beta}_1 = \sum_{i=1}^{n} c_i Y_i$ 단, $\sum_{i=1}^{n} c_i = 0, \sum_{i=1}^{n} c_i X_i = 1$

$\Rightarrow \mathrm{var}(\hat{\beta}_1) \le \mathrm{var}(\tilde{\beta}_1)$



Gauss-Markov 정리 : 고전적 회귀모형에서 최소자승법으로 구한 추정량 $\hat{\beta}_0$ 및 $\hat{\beta}_1$은 선형불편추정량 중에서 최소의 분산을 갖는다. 즉, 최소자승법으로 구한 추정량은 BLUE(Best Linear Unbiased Estimator)이다