계량경제학 강의 2

표본평균과 표본분산

1. 모형의 설정

- 모평균 : $E(y_i) = \mu$
- 모분산 : $Var(y_i) = E\left((y_i - \mu)^2\right) = \sigma^2$

우리의 모형(가정) 하에서,

Q1) $\boxed{\hat{\mu}}$, $\boxed{\hat{\sigma}^2}$ 어디서 나온 공식인가?

Q2) 우리가 믿을 수 있는 공식인가?

Q3) $\underline{\hat{\mu}} \neq \underline{\mu}$, 단 한 번의 기회라면

과 관련한 불확실성?

- 표본평균 : $\hat{\mu} =$ ~~~
- 표본분산 : $\hat{\sigma}^2 =$ ~~~

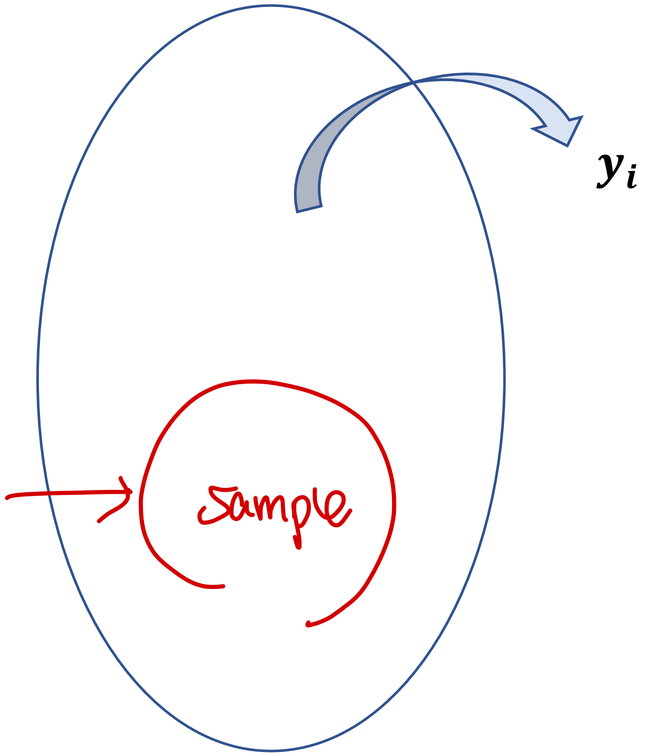
통계 + 경제 : 불확실성을 다루는 학문

→ ① 예측, ② 가설 검정

1. 모형의 설정

우리의 관심대상

'**population**'



random variable

- 확률변수 $\boxed{y_i}$가 취할 수 있는 모든 값들의 평균이 μ : $E(y_i) = \mu$

true value (N)

모수 $= \frac{1}{N}\sum_{i=1}^{N} Y_i$

- 분산이 σ^2 : $Var(y_i) = \sigma^2$

이라고 할 때 다음의 간단한 모형으로 나타낼 수 있다.

$y_i = \mu + e_i$

$y_i - \mu = e_i$

$E(y_i) - E(\mu) = 0$

Population equation : $Y_i = \mu + e_i$ $\quad i = 1, 2, \ldots N$

잔차항 (error term)

Assumption :

① $E(e_i) = 0, \ \forall i$

② $E(e_i^2) = \sigma^2 \ \forall i$ ← var. definition 대입 ✓

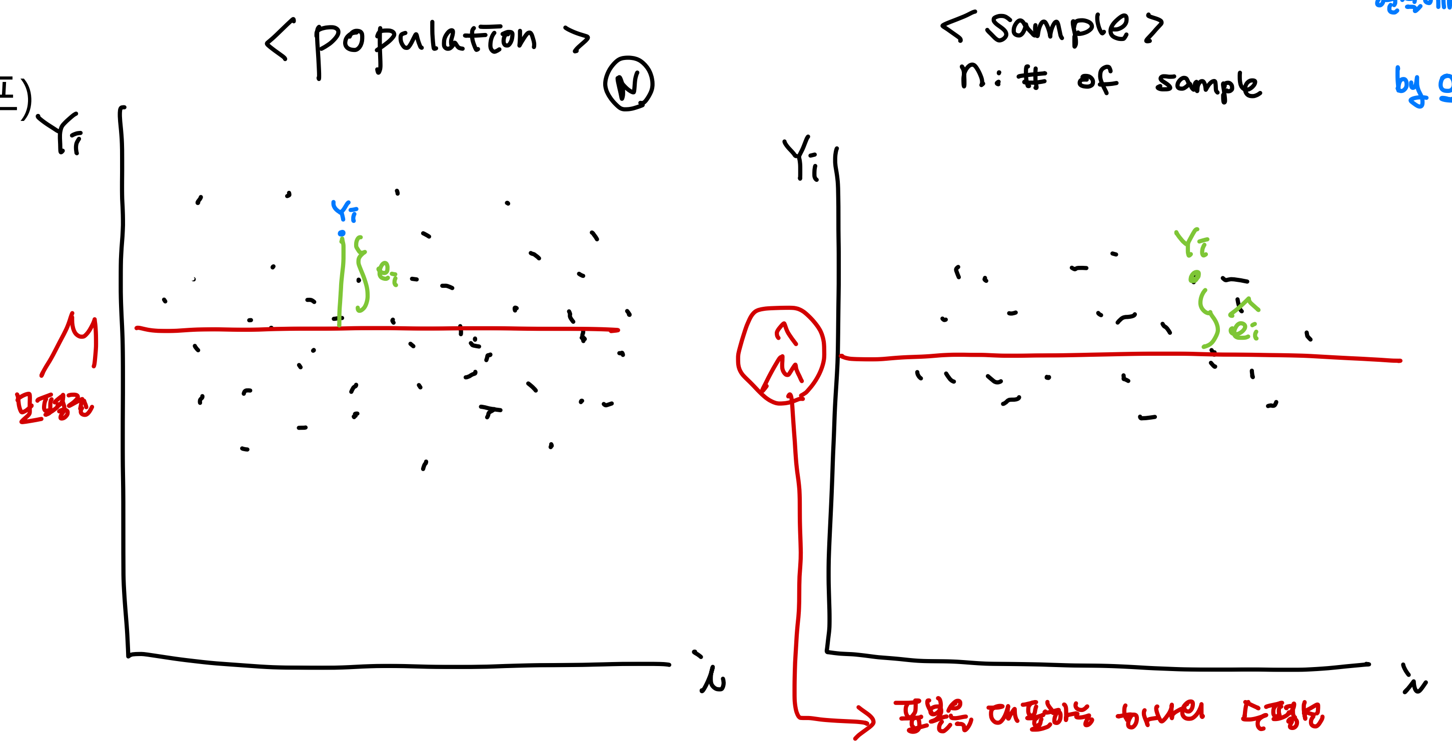
③ $E(e_i \cdot e_j) = 0, \ \forall i \neq j$ ← 공분산 ✓

2. ~~최소제곱법(OLS)~~

- **최소제곱법(OLS: Ordinary Least Squares)**

: 잔차항의 제곱의 합을 극소화 하는 방법

(그래프)



현실에서는 $\hat{\mu} \neq \mu$

by OLS $\quad y = \hat{\mu} + \hat{e}_i$ ⤳ 최소화

$\hat{e}_1 + \hat{e}_2 + \dots + \hat{e}_n$

$\text{Min} \sum_{i=1}^{n} \hat{e}_i^2$ with respect to $\hat{\mu}$

2. 최소제곱법(OLS)

- 표본회귀식(sample regression equation)

가정

: $y_i = \hat{\mu} + \hat{e}_i$

① $E(e_i) = 0 \quad \forall i$
② $E(e_i^2) = \sigma^2 \quad \forall i$
③ $E(e_i \cdot e_j) = 0 \quad \forall i \neq j$

$\Sigma \hat{e}_i^2 = \Sigma (y_i - \mu)^2$

- $\boxed{\text{OLS}}$ 추정량 유도

$\hat{\mu}$을 추정하는 하나의 criterion

$$\boxed{\text{Min} \sum_{i=1}^{n} \hat{e}_i^2 \quad w.r.t \quad \hat{\mu}}$$

$$\rightarrow \boxed{\frac{d \Sigma \hat{e}_i^2}{d \hat{\mu}} = 0, \quad (F.O.C)}$$

$\Rightarrow$ F.O.C)

$\sum_{i=1}^{n} 2(y_i - \hat{\mu})(-1) = 0$

$\Rightarrow \Sigma (Y_i - \hat{\mu}) = 0$

$\Sigma Y_i - n \cdot \hat{\mu} = 0$

$\therefore \hat{\mu} = \frac{1}{n} \Sigma Y_i$

사전적 ex-ante	실현	사후적 ex-post
$\underbrace{\hat{\mu}}_{} = \frac{1}{n} \Sigma Y_i$ estimator, 공식 확률 변수		$\underbrace{\hat{\mu}}_{} = \frac{1}{n} \Sigma Y_i$ estimate. 추정량 / 추정치 값.

3. OLS 추정량의 불편성

Repeated Sampling 에서, 평균적으로 $\hat{\mu}$은 μ라는 값을 갖는다.

Q $\boxed{E(\hat{\mu}) = \mu}$?

성립한다면 $\hat{\mu}$을 μ의 Unbiased Estimator 라 부른다. (불편 추정량)

- **불편성(unbiasedness)** : $E(\hat{\theta}) = \theta$
- OLS 추정량의 불편성:

$E(\hat{\mu}) = \mu$?

① $E\left(\frac{1}{n}\sum Y_i\right)$

$= E\left(\frac{1}{n}(Y_1 + Y_2 + \cdots + Y_n)\right)$

$= \frac{1}{n}\left(\underbrace{E(Y_1)}_{\mu} + \cdots + \underbrace{E(Y_n)}_{\mu}\right)$

$= \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \mu \quad \therefore \; E(\hat{\mu}) = \mu$

② $E(\hat{\mu})$

$= E\left(\frac{1}{n}\sum Y_i\right)$ ← population Equation

$= \frac{1}{n} E\left(\sum(\mu + e_i)\right)$

$= \frac{1}{n}\left(n \cdot \mu + E\left(\sum e_i\right)\right)$

$= \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \mu + \frac{1}{n}\underbrace{\cancel{\left(E(e_1) + E(e_2) + \cdots + E(e_n)\right)}}_{0 \text{ by assumption ①}}$

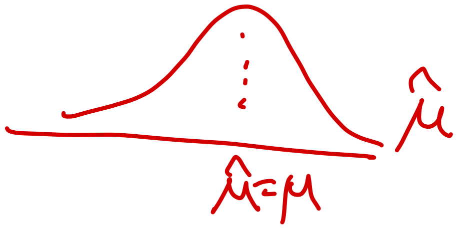
$= \mu$

$\boxed{\therefore E(\hat{\mu}) = \mu}$

4. OLS 추정량의 분산

$\mu \neq \hat{\mu}$, 한 번의 실험

→ $\hat{\mu}$과 관련한 불확실성



+가정 : $E(e_i e_j) = 0\ ,\ i \neq j$

- $Var(\hat{\mu}) = E[(\hat{\mu} - E(\hat{\mu}))^2]$

by def. $E(\hat{\mu}) = \mu$

$$= E[(\hat{\mu} - \mu)^2]$$

by OLS

$$= E[(\frac{1}{n}\sum Y_i - \mu)^2]$$

$$= E[(\frac{\sum Y_i - n\cdot\mu}{n})^2]$$

$\sum(Y_i - \mu)$, e_i

$$= E[(\frac{\sum e_i}{n})^2]$$

$$= \frac{1}{n^2} E[e_1^2 + e_2^2 + \dots + e_n^2]$$

$$= \frac{1}{n^2}\cdot n\sigma^2$$

$$= \frac{\sigma^2}{n}$$

$\therefore\ Var(\hat{\mu}) = \frac{\sigma^2}{n}$

$E(e_i^2) = \sigma^2\ \forall i$

σ^2의 추정?

6. 표본분산의 평균

$Var(\hat{\mu}) = \frac{\sigma^2}{n}$

σ^2의 추정?

(1) μ가 알려져 있는 경우

→ 모집단 평균 known

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} e_i^2$$

Recall ii) $E(e_i^2) = \sigma^2 .\ \forall i$

WTS) $E(\hat{\sigma}^2) = \sigma^2$을 만족하는 $\hat{\sigma}^2$은?

$E(e_1^2 + e_2^2 + \dots + e_n^2) = n\sigma^2$

$E\left(\frac{\Sigma e_i^2}{n}\right) = \sigma^2$

(Pop. eqn) $Y_i = \mu + e_i$ ✓

Sample Data known

→ e_i를 안다

$\hat{\sigma}^2 = \frac{1}{n} e_i^2$

$\hat{\sigma}^2 = \frac{1}{n}\Sigma(Y_i - \mu)^2$

이 case $\hat{\sigma}^2$은 σ^2의 unbiased Estimator

6. 표본분산의 평균

(2) μ가 알려져 있지 않은 경우

$$\boxed{\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} e_i^2}$$

증명) $\sum_{i=1}^{n} e_i^2 =$

Q) $E(\hat{\sigma}^2) = \sigma^2$이 되는 $\hat{\sigma}^2$을 구하자.

→ 우리는 μ가 unknown 이므로 대신 $\hat{\mu}$을 이용한다.

sample) $\boxed{Y_i = \hat{\mu} + \hat{e}_i}$, $\hat{e}_i = Y_i - \hat{\mu}$

$\hat{e}_i^2 = (Y_i - \hat{\mu})^2$을 사용해야 한다

$$e_i = \underbrace{(Y_i - \hat{\mu})}_{\hat{e}_i} + (\hat{\mu} - \mu)$$

$e_i = \hat{e}_i + (\hat{\mu} - \mu)$ — 식 Ⓐ

6. 표본분산의 평균

(2) μ가 알려져 있지 않은 경우

$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} e_i^2$

증명) $\sum_{i=1}^{n} e_i^2 =$

$E\left(\sum_{i=1}^{n} e_i^2\right) = E\left(\sum_{i=1}^{n}(\hat{e}_i + (\hat{\mu}-\mu))^2\right)$

$= E\left[\sum\{\hat{e}_i^2 + (\hat{\mu}-\mu)^2 + 2\cdot\hat{e}_i(\hat{\mu}-\mu)\}\right]$

$2\cdot\sum_{i=1}^{n}\hat{e}_i(\hat{\mu}-\mu)$

(OLS) F.O.C $\frac{\partial\sum e_i^2}{\partial\hat{\mu}} = 0$;

$\sum\hat{e}_i = 0$

$= E\left[\sum\hat{e}_i^2 + \sum(\hat{\mu}-\mu)^2\right]$

$= E(\sum\hat{e}_i^2) + \sum\cdot var(\hat{\mu})$

$= E(\sum\hat{e}_i^2) + \cancel{n}\cdot\frac{\sigma^2}{\cancel{n}}$

$E(e_i^2) = \sigma^2 \ \forall i$

$E(\sum e_i^2) = E(\sum\hat{e}_i^2) + \sigma^2$

$n\cdot\sigma^2 = E(\sum\hat{e}_i^2) + \sigma^2$

6. 표본분산의 평균

(n−1) $\sigma^2 = E(\hat{e}_i^2)$

$\therefore E(\hat{\sigma}^2) = \sigma^2$

$\hat{\sigma}^2 = \frac{1}{n-1}\Sigma \hat{e}_i^2$

(2) μ가 알려져 있지 않은 경우

$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} e_i^2$

증명) $\sum_{i=1}^{n} e_i^2 =$

μ : unknown

$\hat{\sigma}^2 = \frac{1}{n-1}\Sigma \hat{e}_i^2$ ← $\hat{\mu}$의 추정

μ : known ⟺ e_i도 known

$\hat{\sigma}^2 = \frac{1}{n}\Sigma e_i^2$