

제 4 강

최소제곱 추정량의 특성 (Properties of Least Squares Estimators)

계량경제학
4.1

단순 선형회귀 모형 가정

계량경제학
4.2

1. $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$
2. $E(\varepsilon_t) = 0 \iff E(y_t) = \beta_1 + \beta_2 x_t$
3. $\text{var}(\varepsilon_t) = \sigma^2 = \text{var}(y_t)$
4. $\text{cov}(\varepsilon_i, \varepsilon_j) = \text{cov}(y_i, y_j) = 0$
5. x_t 는 확률변수가 아니며 적어도 두 개의 다른 값을 가진다.
6. $\varepsilon_t \sim N(0, \sigma^2) \iff y_t \sim N(\beta_1 + \beta_2 x_t, \sigma^2)$

(강 외생성 가정하의) 단순 선형회귀 모형의 가정

계량경제학
4.3

- | | |
|-----------------------------------------------------------------|-----------------------------------------------------------------|
| 1. $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, t=1, \dots, T$ | 1. $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, t=1, \dots, T$ |
| 2. (x_t, y_t) 는 확률표본 | 2. $E(\varepsilon_t x) = 0, x=(x_1, \dots, x_T)$ |
| 3. $E(\varepsilon_t x_t) = 0$ | 3. $\text{Var}(\varepsilon_t x) = \sigma^2$ |
| 4. $\text{var}(\varepsilon_t x_t) = \sigma^2$ | 4. $\text{cov}(\varepsilon_i, \varepsilon_j x) = 0$ |
| 5. x 는 반드시 두 개의 다른 값을 가진다. | 5. x 는 반드시 두 개의 다른 값을 가진다. |
| 6. (optional) $\varepsilon_t x_t \sim N(0, \sigma^2)$ | 6. (optional) $\varepsilon_t x_t \sim N(0, \sigma^2)$ |

단순 선형회귀 모형 추정

계량경제학
4.4

모수들 β_1 과 β_2 는 알려지지 않은 모집단 상수들임

추정치(estimate) b_1 과 b_2 를 계산해내는 공식을 β_1 과 β_2 의 추정량(estimator)이라고 함

b_1 과 b_2 가 특정한 값이 아니라

공식을 나타낼 때, 이들을 β_1 과 β_2 의 추정량이라고 하며 이들은 확률변수의 함수로서 확률변수들임

추정량은 확률변수

계량경제학
4.5

- 최소제곱추정량 b_1 and b_2 가 확률변수들이라면 그들의 분포들은 무엇이며, 특히 평균, 분산, 공분산은 무엇인가?
- 다른 대안의 추정량들의 특성들을 최소제곱추정량의 특성(properties)들과 어떻게 비교할 수 있는가?

최소제곱 추정량의 기대값

계량경제학
4.6

단순회귀모형에 있어서의 최소제곱추정량:

$$b_2 = \frac{T \sum x_t y_t - \sum x_t \sum y_t}{T \sum x_t^2 - (\sum x_t)^2} \quad (4-1)$$

$$b_1 = \bar{y} - b_2 \bar{x} \quad (4-2)$$

단 $\bar{y} = \sum y_t / T, \quad \bar{x} = \sum x_t / T$

최소제곱 추정량의 기대값 계량경제학 4.7

기울기

$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$ 를 대입하여
다음을 얻음:

$$b_2 = \beta_2 + \frac{T \sum x_t \varepsilon_t - \sum x_t \sum \varepsilon_t}{T \sum x_t^2 - (\sum x_t)^2} \quad (4-3)$$

b_2 의 평균은:

$$E b_2 = \beta_2 + \frac{T \sum x_t E \varepsilon_t - \sum x_t \sum E \varepsilon_t}{T \sum x_t^2 - (\sum x_t)^2} = 0$$

왜냐하면 $E \varepsilon_t = 0$, 따라서 $E b_2 = \beta_2$.

(4-3)에 대한 도출 과정 계량경제학 4.8

$$b_2 = \frac{T \sum x_t y_t - \sum x_t \sum y_t}{T \sum x_t^2 - (\sum x_t)^2}$$

$$= \frac{T \sum x_t (\beta_1 + \beta_2 x_t + \varepsilon_t) - \sum x_t \sum (\beta_1 + \beta_2 x_t + \varepsilon_t)}{T \sum x_t^2 - (\sum x_t)^2}$$

$$= \frac{T \beta_1 \sum x_t + T \beta_2 \sum (x_t)^2 + T \sum x_t \varepsilon_t - T \beta_1 \sum x_t - \beta_2 (\sum x_t)^2 - \sum x_t \sum \varepsilon_t}{T \sum x_t^2 - (\sum x_t)^2}$$

$$= \frac{T \beta_2 \sum (x_t)^2 - \beta_2 (\sum x_t)^2 + T \sum x_t \varepsilon_t - \sum x_t \sum \varepsilon_t}{T \sum x_t^2 - (\sum x_t)^2}$$

불편추정량(Unbiased Estimator) 계량경제학 4.9

기울기

$E b_2 = \beta_2$ 라는 결과는 b_2 의 분포가 β_2 를 중심으로 이루어져 있음을 의미함

b_2 의 분포가 β_2 를 중심으로 이루어 있으므로 b_2 는 β_2 의 불편추정량이라고 함

불편추정량(Unbiased Estimator) 계량경제학 4.10

앞서의 불편성에 대한 결과는 **옳은 모형(correct model)**을 사용하는 것을 전제로 함

모형이 잘못되었거나, 특히 중요한 변수가 빠져 있을 경우, $E \varepsilon_t \neq 0$, 이고 이 경우 $E b_2 \neq \beta_2$ 이다.

예:
 $Y = \beta_1 + \beta_2 X_2 + (\beta_3 X_3 + v)$
 $\varepsilon = \beta_3 X_3 + v$
 $E(\varepsilon_t) \neq 0$

불편추정량(Unbiased Estimator) 계량경제학 4.11

절편

비슷한 방법으로 상수항 또는 절편에 대한 추정량 b_1 역시 모형이 옳게 설정되어 있을 때 β_1 에 대한 불편추정량임

$$E b_1 = \beta_1$$

b_1 의 불편성에 대한 증명 계량경제학 4.12

$$b_1 = \bar{y} - b_2 \bar{x}$$

$$E(b_1) = E(\bar{y}) - \bar{x} E(b_2)$$

$$= E\left(\frac{1}{T} \sum (\beta_1 + \beta_2 x_t + \varepsilon_t)\right) - \beta_2 \bar{x}$$

$$= E\left(\beta_1 + \beta_2 \bar{x} + \frac{1}{T} \sum \varepsilon_t\right) - \beta_2 \bar{x} = \beta_1$$

b₂에 대한 다른 표현

계량경제학
4.13

β₂에 대한 최소제곱추정량은 y_t들의 선형결합으로 표현됨:

$$b_2 = \sum w_t y_t \quad (4-4)$$

where $w_t = \frac{(x_t - \bar{x})}{\sum (x_t - \bar{x})^2}$

$$b_2 = \beta_2 + \sum w_t \varepsilon_t \quad (4-5)$$

$$b_1 = \bar{y} - b_2 \bar{x}$$



(4-4)와 (4-5)에 대한 도출 과정

계량경제학
4.14

$$\begin{aligned} \sum (x_t - \bar{x})^2 &= \sum x_t^2 - 2\bar{x} \sum x_t + T\bar{x}^2 = \sum x_t^2 - 2\bar{x} \left(T \frac{1}{T} \sum x_t \right) + T\bar{x}^2 \\ &= \sum x_t^2 - 2T\bar{x} + T\bar{x}^2 = \sum x_t^2 - T\bar{x}^2 \\ &= \sum x_t^2 - \bar{x} \sum x_t = \sum x_t^2 - \frac{(\sum x_t)^2}{T} \\ \sum (x_t - \bar{x})(y_t - \bar{y}) &= \sum x_t y_t - T\bar{x}\bar{y} = \sum x_t y_t - \frac{\sum x_t \sum y_t}{T} \\ b_2 &= \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2} = \frac{T \sum x_t y_t - \sum x_t \sum y_t}{T \sum x_t^2 - (\sum x_t)^2} \end{aligned}$$

(4-4)와 (4-5)에 대한 도출 과정-계속

계량경제학
4.15

$$\begin{aligned} \sum (x_t - \bar{x}) &= 0 \\ b_2 &= \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2} = \frac{\sum (x_t - \bar{x})y_t - \bar{y} \sum (x_t - \bar{x})}{\sum (x_t - \bar{x})^2} \\ &= \frac{\sum (x_t - \bar{x})y_t}{\sum (x_t - \bar{x})^2} = \sum \left[\frac{(x_t - \bar{x})}{\sum (x_t - \bar{x})^2} \right] y_t = \sum w_t y_t \\ b_2 &= \sum w_t y_t = \sum w_t (\beta_1 + \beta_2 x_t + \varepsilon_t) \\ &= \beta_1 \sum w_t + \beta_2 \sum w_t x_t + \sum w_t \varepsilon_t \end{aligned}$$

(4-4)와 (4-5)에 대한 도출 과정-계속

계량경제학
4.16

$$\begin{aligned} \sum w_t &= \sum \left[\frac{(x_t - \bar{x})}{\sum (x_t - \bar{x})^2} \right] = \frac{1}{\sum (x_t - \bar{x})^2} \sum (x_t - \bar{x}) = 0 \\ \sum (x_t - \bar{x})^2 &= \sum (x_t - \bar{x})(x_t - \bar{x}) \\ &= \sum (x_t - \bar{x})x_t - \bar{x} \sum (x_t - \bar{x}) = \sum (x_t - \bar{x})x_t \\ \sum w_t x_t &= \frac{\sum (x_t - \bar{x})x_t}{\sum (x_t - \bar{x})^2} = \frac{\sum (x_t - \bar{x})x_t}{\sum (x_t - \bar{x})x_t} = 1 \\ b_2 &= \beta_1 \sum w_t + \beta_2 \sum w_t x_t + \sum w_t \varepsilon_t \\ &= \beta_2 + \sum w_t \varepsilon_t \end{aligned}$$

최소제곱 추정량의 분산

계량경제학
4.17

기울기

y_t와 ε_t가 모두 분산이 σ²임이 주어졌을 때, 추정량 b₂의 분산은?:

$$\text{var}(b_2) = \frac{\sigma^2}{\sum (x_t - \bar{x})^2}$$



b₂는 y_t들의 함수이지만 var(b₂)는 직접적으로 y_t를 포함하지 않음.

b₂의 분산

계량경제학
4.18

$$\begin{aligned} \text{var}(b_2) &= \text{var}(\beta_2 + \sum w_t \varepsilon_t) = \text{var}(\sum w_t \varepsilon_t) && \text{[since } \beta_2 \text{ is a constant]} \\ &= \sum w_t^2 \text{var}(\varepsilon_t) && \text{[using cov}(\varepsilon_i, \varepsilon_j) = 0\text{]} \\ &= \sigma^2 \sum w_t^2 && \text{[using var}(\varepsilon_t) = \sigma^2\text{]} \\ &= \frac{\sigma^2}{\sum (x_t - \bar{x})^2} \end{aligned}$$

$$\sum w_t^2 = \sum \left[\frac{(x_t - \bar{x})^2}{\{\sum (x_t - \bar{x})^2\}^2} \right] = \frac{1}{\sum (x_t - \bar{x})^2}$$



절편

$$b_1 = \bar{y} - b_2 \bar{x} \Rightarrow$$

추정량 b_1 의 분산은?:

$$\text{var}(b_1) = \sigma^2 \left[\frac{\sum x_t^2}{T \sum (x_t - \bar{x})^2} \right]$$



$$\begin{aligned} \text{Var}(b_1) &= E[(b_1 - \beta_1)^2] = E[(\bar{y} - b_2 \bar{x} - (\bar{y} - \beta_2 \bar{x} - \bar{\epsilon}))^2] \\ &= E[(-\bar{x}(b_2 - \beta_2) + \bar{\epsilon})^2] \\ &= \bar{x}^2 E[(b_2 - \beta_2)^2] + E[\bar{\epsilon}^2] \quad (\because E[\bar{\epsilon}(b_2 - \beta_2)] = 0) \\ &= \bar{x}^2 \frac{\sigma^2}{\sum (x_t - \bar{x})^2} + \frac{1}{T^2} E[\sum \epsilon_t \sum \epsilon_s] = \frac{\sigma^2 \bar{x}^2}{\sum (x_t - \bar{x})^2} + \frac{\sigma^2}{T} \\ &= \frac{\sigma^2 [T\bar{x}^2 + \sum (x_t - \bar{x})^2]}{T \sum (x_t - \bar{x})^2} \\ &= \frac{\sigma^2 [T\bar{x}^2 + \sum x_t^2 - 2\bar{x} \sum x_t + T\bar{x}^2]}{T \sum (x_t - \bar{x})^2} = \frac{\sigma^2 [T\bar{x}^2 + \sum x_t^2 - 2T\bar{x}^2 + T\bar{x}^2]}{T \sum (x_t - \bar{x})^2} \\ &= \frac{\sigma^2 [\sum x_t^2]}{T \sum (x_t - \bar{x})^2} \end{aligned}$$



$$\text{cov}(b_1, b_2) = \sigma^2 \left[\frac{-\bar{x}}{\sum (x_t - \bar{x})^2} \right]$$



$\bar{x} = 0$ 이면, 기울기는 공분산에 영향을 주지 않고 변화할 수 있음

$$\begin{aligned} \text{Cov}(b_1, b_2) &= E\{[b_1 - E(b_1)][b_2 - E(b_2)]\} \\ &= E\{[b_1 - \beta_1][b_2 - \beta_2]\} \\ &= E\{[(\bar{y} - b_2 \bar{x}) - (\bar{y} - \beta_2 \bar{x} - \bar{\epsilon})][b_2 - \beta_2]\} \\ &= E\{[(\bar{y} - b_2 \bar{x}) - (\bar{y} - \beta_2 \bar{x})][b_2 - \beta_2]\} \\ &= E\{[-\bar{x}(b_2 - \beta_2)][b_2 - \beta_2]\} \\ &= -\bar{x} E(b_2 - \beta_2)^2 \\ &= -\bar{x} \left[\frac{\sigma^2}{\sum (x_t - \bar{x})^2} \right] \\ E(\bar{\epsilon}(b_2 - \beta_2)) &= \frac{1}{T} E(\sum \epsilon_i \sum w_i \epsilon_i) \\ &= \frac{1}{T} E\left(\sum w_i \epsilon_i^2 + \sum_{i \neq s} w_i \epsilon_i \epsilon_s\right) \\ &= \frac{1}{T} E(\sum w_i \epsilon_i^2) = \frac{1}{T} \sigma^2 (\sum w_i) = 0 \end{aligned}$$



- σ^2 : y_t 값들에 내재해 있는 불확실성 (오차항의 원천에서 비롯됨)이 b_1 과 b_2 의 불확실성 및 그 관계에 영향을 줌
- x_t 값들이 퍼져 있을수록 b_1, b_2 에 대한 신뢰를 높일 수 있음
- 표본의 크기 T 가 클수록 분산과 공분산이 작아짐
- b_1 의 분산은 x_t 의 값들이 0으로부터 멀어질 수록 커짐 (즉 제곱값들의 합이 클수록 커짐).
- x_t 의 표본평균이 0일 경우 기울기 b_2 의 변화는 절편 b_1 에 영향을 미치지 않음. 그러나 표본 평균이 +인 경우 b_1 와 b_2 의 공분산은 음이며, -인 경우엔 양임

단순선형회귀모형의 가정 1 - 5하에서, (표준, ordinary)최소제곱추정량 b_1 과 b_2 는 β_1 와 β_2 에 대한 모든 선형불편추정량 가운데 가장 작은 분산을 갖는다.

이는 b_1 과 b_2 가 β_1 and β_2 에 대한 최우수 선형 불편 추정량 (BLUE, the Best Linear Unbiased Estimator)임을 의미함

Gauss Markov Theorem의 의미

계량경제학 4.25

1. b_1 과 b_2 는 선형이고 불편인 추정량의 범주내에서 “최우수(best)”임
2. “최우수(Best)”라 함은 선형/불편 추정량 가운데 가장 작은 분산을 가짐을 의미함
3. 가우스 마코프 정리가 성립하기 위해서는 처음 다섯개의 가정들 모두가 만족되어야 함.
4. 가우스 마코프 정리는 여섯번째 가정, 즉 정규분포에 대한 가정은 필요로하지 않음
5. 가우스 마코프 정리는 추정량 b_1 및 b_2 에 적용되는 것이며 특정한 표본값(즉 추정치) b_1 및 b_2 에 적용되는 것이 아님

Gauss Markov Theorem의 의미

계량경제학 4.26

7. 선형불편추정량의 범주에 추정을 국한시키지 않는다면 가우스 마코프 정리는 더 이상 의미를 갖지 않으며 비선형 혹은(그리고) 편향(biased) 추정량을 대신 사용할 수도 있음 (Note: 편향 혹은 비선형 추정량이 가우스 마코프 정리를 만족시키는 추정량보다 더 작은 분산을 가질 수 있음). 하지만 선형불편추정량에 우리의 관심을 국한시키는 한 더 이상 다른 추정량을 탐구할 필요는 없음



Proof of the Gauss–Markov Theorem

계량경제학 4.27

$$b_2^* = \sum k_i y_i$$

(where the k_i are constants) be any other linear estimator of β_2 .

- Suppose that

$$k_i = w_i + c_i$$

, where c_i is another constant and w_i is given in equation (4.4).

- Into this new estimator substitute y_i and simplify, using the properties of w_i .

$$\begin{aligned} b_2^* &= \sum k_i y_i = \sum (w_i + c_i) y_i = \sum (w_i + c_i)(\beta_1 + \beta_2 x_i + \varepsilon_i) \\ &= \sum (w_i + c_i) \beta_1 + \sum (w_i + c_i) \beta_2 x_i + \sum (w_i + c_i) \varepsilon_i \\ &= \beta_1 \sum w_i + \beta_1 \sum c_i + \beta_2 \sum w_i x_i + \beta_2 \sum c_i x_i + \sum (w_i + c_i) \varepsilon_i \\ &= \beta_1 \sum c_i + \beta_2 \sum c_i x_i + \sum (w_i + c_i) \varepsilon_i \end{aligned} \tag{4.6}$$

since $\sum w_i = 0$ and $\sum w_i x_i = 1$.

Proof of the Gauss–Markov Theorem

계량경제학 4.28

$$\begin{aligned} E(b_2^*) &= \beta_1 \sum c_i + \beta_2 \sum c_i x_i + \sum (w_i + c_i) E(\varepsilon_i) \\ &= \beta_1 \sum c_i + \beta_2 \sum c_i x_i \end{aligned} \tag{4.7}$$

In order for the linear estimator $b_2^* = \sum k_i y_i$ to be unbiased it must be true that

$$\sum c_i = 0 \text{ and } \sum c_i x_i = 0 \tag{4.8}$$

These conditions must hold in order for

$$b_2^* = \sum k_i y_i$$

to be in the class of linear and unbiased estimators.

Proof of the Gauss–Markov Theorem

계량경제학 4.29

So we will assume the conditions (4.8) hold and use them to simplify expression (4.6):

$$b_2^* = \sum k_i y_i = \beta_2 + \sum (w_i + c_i) \varepsilon_i \tag{4.9}$$

We can now find the variance of the linear unbiased estimator

Using the additional fact that

$$\sum c_i w_i = \sum \left[\frac{c_i (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right] = \frac{1}{\sum (x_i - \bar{x})^2} \sum c_i x_i - \frac{\bar{x}}{\sum (x_i - \bar{x})^2} \sum c_i = 0$$

- Use the properties of variance to obtain:

$$\begin{aligned} \text{var}(b_2^*) &= \text{var}(\beta_2 + \sum (w_i + c_i) \varepsilon_i) = \sum (w_i + c_i)^2 \text{var}(\varepsilon_i) \\ &= \sigma^2 \sum (w_i + c_i)^2 = \sigma^2 \sum w_i^2 + \sigma^2 \sum c_i^2 \\ &= \text{var}(b_2) + \sigma^2 \sum c_i^2 \\ &\geq \text{var}(b_2) \text{ since } \sum c_i^2 \geq 0 \end{aligned} \tag{4.10}$$



최소제곱추정량의 확률분포

계량경제학 4.30

y_i 와 ε_i 가 정규분포를 하는 경우

β_2 의 최소제곱추정량은 y_i 들의 선형결합으로 표현될 수 있음

$$b_2 = \sum w_t y_t \text{ where } w_t = \frac{(x_t - \bar{x})}{\sum (x_t - \bar{x})^2}$$

$$b_1 = \bar{y} - b_2 \bar{x}$$

정규분포를 하는 확률변수의 선형결합은 정규분포를 하게 되므로 b_1 와 b_2 가 정규분포를 함을 의미함

중심극한정리(CLT)하에서 정규분포

가우스 마코프 정리의 다섯개의 가정들이 성립하고 표본의 크기 T가 충분히 크다면, 최소제곱추정량 b_1 과 b_2 의 분포는 정규분포로 근사(approximation)할 수 있으며 표본의 크기가 클 수록 근사의 정확도는 커짐

$$b_1 \sim N \left[\beta_1, \frac{\sigma^2 \sum x_t^2}{T \sum (x_t - \bar{x})^2} \right]$$

$$b_2 \sim N \left[\beta_2, \frac{\sigma^2}{\sum (x_t - \bar{x})^2} \right]$$

표본의 크기 T가 무한히 커질 경우 추정량 b_1 과 b_2 가 그 모수값(true population value), β_1 과 β_2 로 (확률적으로)수렴하는 것이 바람직함 - 일치성

이러한 일치성(consistency)를 달성하는 하나의 방법은 b_1 과 b_2 의 분산이 T가 무한대로 감에 따라 0이 되는 것임

최소제곱추정량 b_1 과 b_2 의 분산의 공식은 그들의 분산이 실제로 T가 무한히 커짐에 따라 0으로 가게 되는 것을 보여주고 있음. 따라서 b_1 과 b_2 는 β_1 과 β_2 의 일치추정량임

$$E(\varepsilon_t^2) = \sigma^2 \Rightarrow \sum \varepsilon_t^2 / T \quad ?$$

$$\hat{\varepsilon}_t = y_t - b_1 - b_2 x_t \Rightarrow \sum \hat{\varepsilon}_t^2 / T \quad ?$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^T \hat{\varepsilon}_t^2}{T-2}$$

$\hat{\sigma}^2$ 는 σ^2 의 불편추정량임



$$\begin{aligned} \varepsilon_t &= y_t - \beta_1 - \beta_2 x_t \Rightarrow \bar{\varepsilon} = \bar{y} - \beta_1 - \beta_2 \bar{x} \\ &\Rightarrow \varepsilon_t - \bar{\varepsilon} = y_t - \bar{y} - \beta_2 (x_t - \bar{x}) \\ \hat{\varepsilon}_t &= y_t - b_1 - b_2 x_t, \quad 0 = \bar{y} - b_1 - b_2 \bar{x} \Rightarrow \hat{\varepsilon}_t = y_t - \bar{y} - b_2 (x_t - \bar{x}) \\ \hat{\varepsilon}_t &= \varepsilon_t - \bar{\varepsilon} + \beta_2 (x_t - \bar{x}) - b_2 (x_t - \bar{x}) \\ &= -(b_2 - \beta_2)(x_t - \bar{x}) + \varepsilon_t - \bar{\varepsilon} \\ \sum (\hat{\varepsilon}_t)^2 &= (b_2 - \beta_2)^2 \sum (x_t - \bar{x})^2 + \sum (\varepsilon_t - \bar{\varepsilon})^2 \\ &\quad - 2(b_2 - \beta_2) \sum (x_t - \bar{x})(\varepsilon_t - \bar{\varepsilon}) \end{aligned}$$

$$\begin{aligned} E\left(\sum (\hat{\varepsilon}_t)^2\right) &= E\left((b_2 - \beta_2)^2 \sum (x_t - \bar{x})^2 + E\left(\sum (\varepsilon_t - \bar{\varepsilon})^2\right) - 2E\left((b_2 - \beta_2) \sum (x_t - \bar{x})(\varepsilon_t - \bar{\varepsilon})\right)\right) \\ E\left(\sum (\varepsilon_t - \bar{\varepsilon})^2\right) &= \sum E(\varepsilon_t^2 - 2\varepsilon_t \bar{\varepsilon} + \bar{\varepsilon}^2) = \sum E(\varepsilon_t^2) - 2 \sum E(\varepsilon_t \bar{\varepsilon}) + T E(\bar{\varepsilon}^2) \\ &= T\sigma^2 - \frac{2}{T} \sum E\left(\varepsilon_t^2 + \varepsilon_t \sum_{j \neq t} \varepsilon_j\right) + \frac{1}{T} E\left(\sum \varepsilon_t^2 + \sum_{i \neq j} \varepsilon_i \varepsilon_j\right) \\ &= T\sigma^2 - 2\sigma^2 + \sigma^2 = (T-1)\sigma^2 \\ -2E\left((b_2 - \beta_2) \sum (x_t - \bar{x})(\varepsilon_t - \bar{\varepsilon})\right) &= -2 \sum (x_t - \bar{x}) E\left((\varepsilon_t - \bar{\varepsilon}) \sum w_j \varepsilon_j\right) \\ &= -2 \sum (x_t - \bar{x}) E\left(w_t \varepsilon_t^2 - \frac{1}{T} \sum w_j \varepsilon_j^2\right) \\ &= -2\sigma^2 \sum (x_t - \bar{x})(w_t - \bar{w}) = -2\sigma^2 \end{aligned}$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^T \hat{\varepsilon}_t^2}{T-2} \Rightarrow \widehat{\text{Var}}(b_1), \widehat{\text{Var}}(b_2)$$

$$\text{S.E.}(b_i) = \sqrt{\widehat{\text{Var}}(b_i)}$$

표준오차(Standard Error) ▶

gretl: model 1

Model 1: OLS, using observations 1-40
Dependent variable: y

	coefficient	std. error	t-ratio	p-value
const	40.7676	22.1387	1.841	0.0734 *
x	0.128289	0.0305333	4.201	0.0002 ***

	Mean dependent var	S.D. dependent var	Sum squared resid	S.E. of regression	R-squared	Adjusted R-squared	F(1, 38)	P-value(F)	Log-likelihood	Akaike criterion	Schwarz criterion	Hannan-Quinn
	130.3130	45.15857	54311.33	37.80536	0.317118	0.299148	17.64653	0.000155	-201.0297	406.0594	409.4372	407.2807

설명변수의 값 x_0 가 주어졌을 때,
종속변수의 값 y_0 를 예측하고자 할 경우

최소제곱예측치는:

$$\hat{y}_0 = b_1 + b_2 x_0$$