



豊富な実証例から
計量経済学の
「生きた」知識を身につける！

第8章 標準的仮定の妥当性

藪友良

『入門 実践する計量経済学』

(東洋経済新報社)

PPT

・ ガウス＝マルコフの定理

--- 標準的仮定のもとで、最小2乗推定量は、
最良線形不偏推定量となる

・ 標準的仮定の妥当性

仮定1 X_i は確率変数ではない

仮定2 n が大きくなると $\sum_{i=1}^n (X_i - \bar{X})^2$ は ∞ に近づく

仮定3 $E[u_i] = 0$

仮定4 $V(u_i) = E[u_i^2] = \sigma^2$

仮定5 $Cov(u_i, u_j) = E[u_i u_j] = 0$

仮定6 誤差項は正規分布に従う

ガウス＝マルコフの定理

$$Y = \alpha + \beta X + u$$

- 線形推定量

$$\beta^* = \sum_{i=1}^n d_i Y_i$$

---推定量が被説明変数 $Y_1, Y_2, \dots, Y_n$ の線形関数である

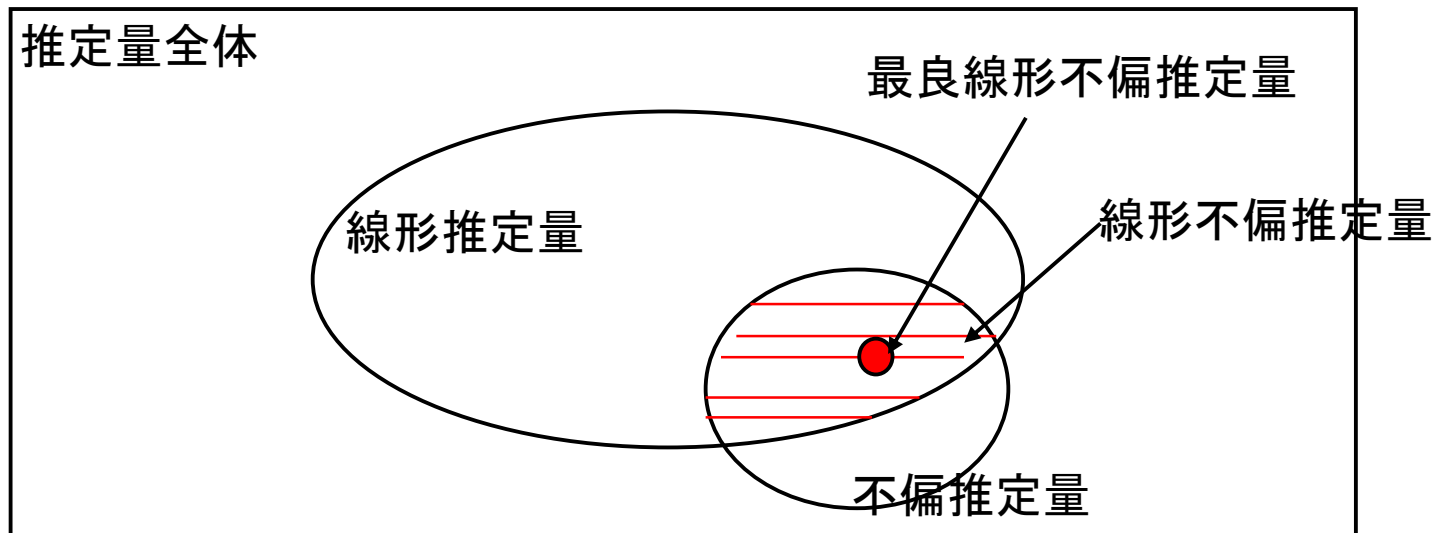
--- d_i は加重であり、何らかの固定した値をとる

- 線形不偏推定量

$$E[\beta^*] = \beta$$

- 最良線形不偏推定量 (Best Linear Unbiased Estimator: BLUE)**

---線形不偏推定量の中で最小の分散となる



例1) 2つの観測値だけを用いた推定量 ($X_2 - X_1 \neq 0$)

$$\beta^* = \frac{Y_2 - Y_1}{X_2 - X_1}$$

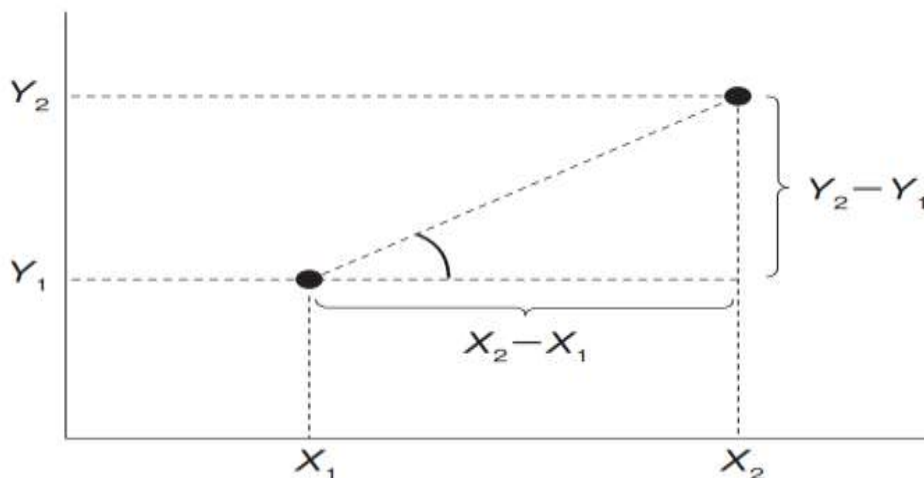
- 線形推定量

$$\beta^* = \underbrace{\frac{-1}{X_2 - X_1}}_{d_1} Y_1 + \underbrace{\frac{1}{X_2 - X_1}}_{d_2} Y_2$$

- 線形不偏推定量

$$\begin{aligned}\beta^* &= \frac{(\alpha + \beta X_2 + u_2) - (\alpha + \beta X_1 + u_1)}{X_2 - X_1} \\ &= \frac{\beta(X_2 - X_1)}{X_2 - X_1} + \frac{u_2 - u_1}{X_2 - X_1} = \beta + \frac{u_2 - u_1}{X_2 - X_1}\end{aligned}$$

図 8 - 1 線形推定量



例2) OLS推定量は線形不偏推定量

$$\begin{aligned}\hat{\beta} &= \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \\&= \frac{(X_1 - \bar{X})Y_1 + (X_2 - \bar{X})Y_2 + \cdots + (X_n - \bar{X})Y_n}{\sum_{i=1}^n (X_i - \bar{X})^2} \\&= \underbrace{\frac{(X_1 - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}}_{c_1} Y_1 + \underbrace{\frac{(X_2 - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}}_{c_2} Y_2 + \cdots + \underbrace{\frac{(X_n - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}}_{c_n} Y_n \\&= c_1 Y_1 + c_2 Y_2 + \cdots + c_n Y_n\end{aligned}$$

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

--- $E[\hat{\beta}] = \beta$ であるから、線形不偏推定量

ガウス＝マルコフの定理

標準的仮定1、3、4、5が満たされるとき、**OLS推定量は最良線形不偏推定量(BLUE)**となる

仮定1: X_i は確率変数ではない

仮定3: $E[u_i] = 0$ (誤差項の期待値が0)

仮定4: $V(u_i) = E[u_i^2] = \sigma^2$ (誤差項の分散が σ^2 で一定)

仮定5: $Cov(u_i, u_j) = E[u_i u_j] = 0$ (誤差項の共分散が0)

Hansen (2022)
最良不偏推定量

--- OLS推定量の有用性を示す定理である

--- 仮定2(n が大きくなると $\sum_{i=1}^n (X_i - \bar{X})^2$ は ∞ に近づく)と

仮定6(誤差項は正規分布に従う)は定理の証明に不要である

[証明]

・線形推定量が不偏推定量であるための条件

$$\begin{aligned}\beta^* &= \sum_{i=1}^n d_i Y_i \\ &= \sum_{i=1}^n d_i (\alpha + \beta X_i + u_i) \\ &= \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i + \sum_{i=1}^n d_i u_i\end{aligned}$$

$$Y_i = \alpha + \beta X_i + u_i$$

ここで期待値をとると

$$\begin{aligned}E[\beta^*] &= \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i + \sum_{i=1}^n d_i E[u_i] \\ &= \alpha \sum_{i=1}^n d_i + \beta \sum_{i=1}^n d_i X_i\end{aligned}$$

仮定1

仮定3

不偏性の条件は

$$\textcircled{1} \sum_{i=1}^n d_i = 1 \quad \textcircled{2} \sum_{i=1}^n d_i X_i = \beta$$

つまり、線形不偏推定量は

$$\beta^* = \beta + \sum_{i=1}^n d_i u_i$$

・線形不偏推定量 $\beta^* = \beta + \sum_{i=1}^n d_i u_i$ の分散

$$V(\beta^*) = E[(\beta^* - \beta)^2]$$

$$= E[(\sum_{i=1}^n d_i u_i)^2] = \sigma^2 \sum_{i=1}^n d_i^2$$

--- $n = 2$ なら

仮定4

仮定5

$$E[(d_1 u_1 + d_2 u_2)^2] = d_1^2 E[u_1^2] + d_2^2 E[u_2^2] + 2d_1 d_2 E[u_1 u_2]$$

ここで $V(\beta^*)$ を書き換えると

$$\sigma^2 \sum_{i=1}^n d_i^2 = \sigma^2 \sum_{i=1}^n \{c_i + (d_i - c_i)\}^2$$

$$= \sigma^2 \sum_{i=1}^n \{c_i^2 + (d_i - c_i)^2 + 2(d_i - c_i)c_i\}$$

$$= \sigma^2 \left\{ \sum_{i=1}^n c_i^2 + \sum_{i=1}^n (d_i - c_i)^2 + 2 \sum_{i=1}^n (d_i - c_i)c_i \right\}$$

分散は $d_i = c_i$ の
とき最小化される

--- 第3項は0となる

$$\textcircled{2} \sum_{i=1}^n d_i X_i = 1$$

$$\textcircled{1} \sum_{i=1}^n d_i = 0$$

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

$$\sum_{i=1}^n c_i^2 = \left\{ \frac{(X_1 - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}^2 + \left\{ \frac{(X_2 - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}^2 + \cdots + \left\{ \frac{(X_n - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \right\}^2$$

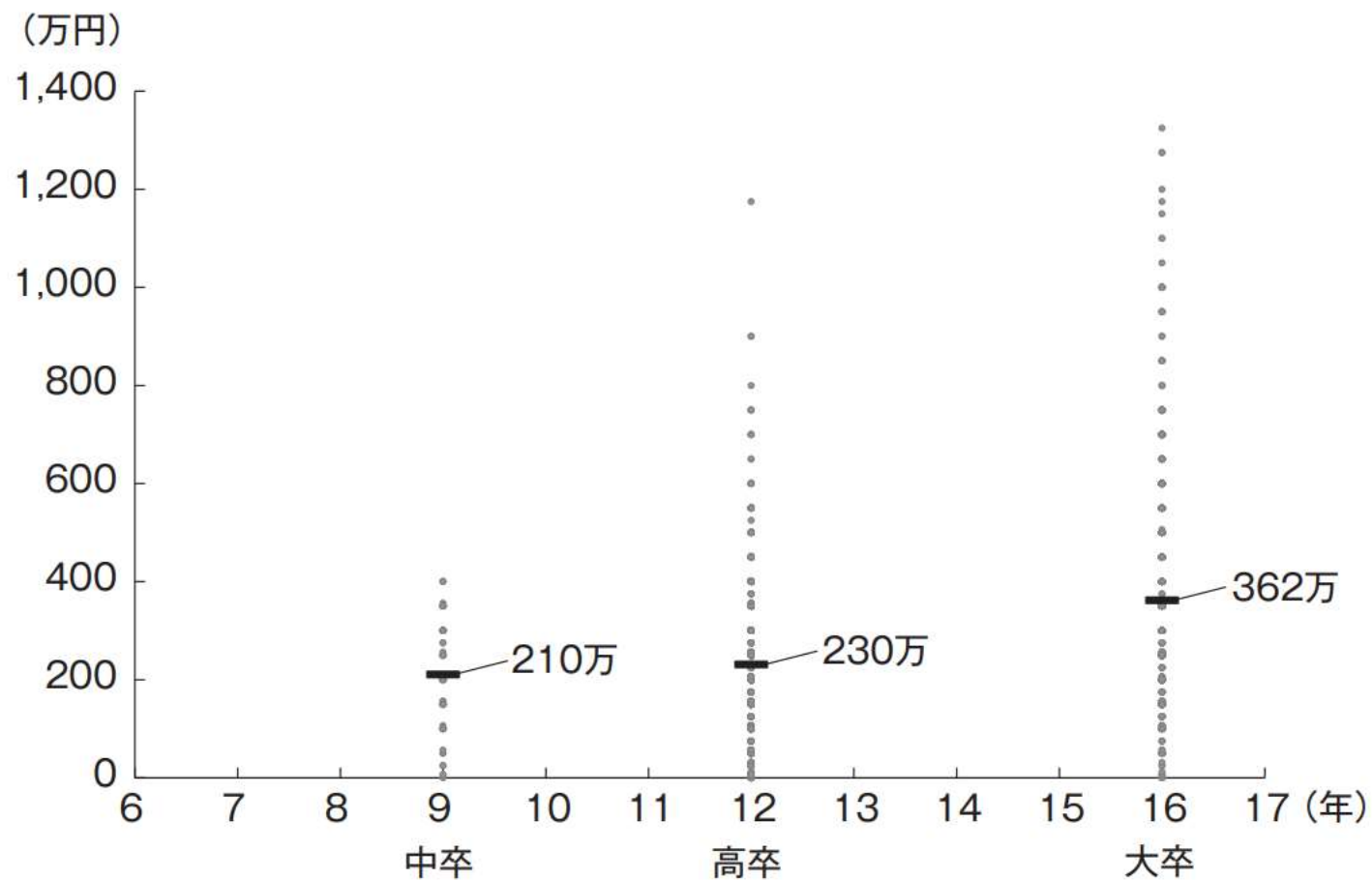
$$= \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{\left\{ \sum_{i=1}^n (X_i - \bar{X})^2 \right\}^2} = \frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

標準的仮定の妥当性

仮定4($V(u_i) = E[u_i^2] = \sigma^2$)

- **均一分散**(誤差項の分散は全ての i で同じ)
 - 無作為抽出により選ばれた横断面データなら均一分散となる
- **不均一分散**(誤差項の分散は i に依存してかわる)
 - 通常のデータでは $V(u_i) = E[u_i^2] = \sigma_i^2$
 - 例1) 所得の分散は性別に依存して変わる。男性の方が分散は大きい傾向がある
 - 例2) 時系列データでは誤差項の分散は時間に依存してかわる
 - 金融危機、コロナ危機のとき変数の分散は大きい

図 8 - 6 教育年数と所得との関係



仮定5($Cov(u_i, u_j) = E[u_i u_j] = 0$)

- 無作為抽出した横断面データでは問題とならない
 - 無作為抽出なら個々のデータは相互に独立となる
- 時系列データ(パネルデータ)なら満たされない
 - ショックの影響は時間を通じて波及していく(系列相関がある)
例) 金融危機は2008年の景気を減速させただけでなく、その後、数年の間、影響を与え続けていた
- 仮定4、5が満たされないとOLS推定量はBLUEではない
 - OLS推定量の分散は $\sigma_{\hat{\beta}}^2 = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$ ではない。よって、不均一分散や系列相関に頑健な分散を用いる
 - 一般化最小2乗(GLS)推定量はBLUEとなる

仮定1(X_i は確率変数ではない)

- 仮定1のもとで、説明変数 X_i は誤差項 u_i と無相関になる

$$E[\sum_{i=1}^n (X_i - \bar{X}) u_i] = \sum_{i=1}^n (X_i - \bar{X}) E[u_i] = 0$$

- 説明変数が確率変数でも、説明変数と誤差項が無相関なら
(**外生性**) 一致性は満たされる。説明変数と誤差項に相関があると(**内生性**)、不偏性も一致性も満たされない。

[証明]

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K + u$$

X_1 を1単位増やすと

$$Y' = \alpha + \beta_1 (X_1 + 1) + \beta_2 X_2 + \cdots + \beta_K X_K + u'$$

このとき、 Y の変化分は

$$\begin{aligned} Y' - Y &= [\alpha + \beta_1 (X_1 + 1) + \beta_2 X_2 + \cdots + \beta_K X_K + u'] \\ &\quad - [\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_K X_K + u] = \beta_1 + \underbrace{(u' - u)} \end{aligned}$$

バイアス

仮定6(誤差項は正規分布に従う)

- ・ ガウス=マルコフの定理に仮定6は不要である
- ・ 小標本で、信頼区間や仮説検定に必要な仮定である
 - 仮定6のもとで、小標本でもOLS推定量は正規分布に従う
- ・ サンプルサイズが大きいのら仮定6はいらない

$$\hat{\beta} = \beta + \frac{\sum_{i=1}^n (X_i - \bar{X}) u_i}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

ここで $v_i = (X_i - \bar{X}) u_i$ と定義し、 $\sqrt{n}$ を両辺にかけると

$$\sqrt{n}(\hat{\beta} - \beta) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^n v_i}{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

--- (ゆるい条件のもとで) サンプルサイズが大きいと、

- ① X_i の標本分散は分散に収束する
- ② 中心極限定理より、 $\sum_{i=1}^n v_i$ は正規分布に従う

- サンプルサイズが大きいと、OLS推定量は正規分布する
- t 統計量は標準正規分布に従う

まとめ

- ・ **ガウス＝マルコフの定理**

- OLS推定量はBLUEとなる

- ・ **仮定4(均一分散)と仮定5(系列相関なし)**

- これらが満たされないと、OLS推定量はBLUEではない

- ・ **仮定1(説明変数は確率変数ではない)**

- 説明変数と誤差項が無相関なら一貫性はある

- 相関しているとき、OLS推定量は一貫性を満たさない

- ・ **仮定6(誤差項は正規分布に従う)**

- サンプルサイズが大きいもとで問題とならない