第 25 回　回帰係数の推定と検定（13.2.2–13.3）

村澤 康友

2024 年 1 月 22 日

今日のポイント

1. 誤差項が無相関で分散が均一な線形回帰モデルを古典的線形回帰モデルという．誤差項が独立に $\mathrm{N}\left(0,\sigma^2\right)$ に従う線形回帰モデルを古典的正規線形回帰モデルという．
2. 被説明変数の線形関数で表される推定量を線形推定量という．不偏な線形推定量を線形不偏推定量という．分散が最小となる線形不偏推定量を最良線形不偏推定量（BLUE）という．古典的線形回帰モデルの回帰係数の OLS 推定量は BLUE（ガウス＝マルコフ定理）．
3. 誤差分散 σ^2 の不偏推定量は $s^2 := [1/(n-k)]\sum_{i=1}^n (y_i-\hat{y}_i)^2$．ただし k は推定する係数の数．古典的正規線形回帰モデルなら $(n-k)s^2/\sigma^2 \sim \chi^2(n-k)$．
4. H_0 :「個別の回帰係数＝ 0」を検定する t 統計量の値を t 値という．H_0 の下で $t \sim \mathrm{t}(n-k)$．
5. H_0 :「定数項を除く全ての回帰係数＝ 0」を検定する F 統計量の値を F 値という．H_0 の下で $F \sim \mathrm{F}(k-1,n-k)$．

目次

1　古典的線形回帰モデル

1.1　線形回帰モデル（p. 259）

2 変量データを $((y_1,x_1),\dots,(y_n,x_n))$ とする．簡単化のため x_i は非確率変数とする（実験データ）．y_i の x_i 上への線形回帰モデルは

$$\mathrm{E}(y_i)=\alpha+\beta x_i$$

または

$$\begin{aligned} y_i &= \alpha+\beta x_i+u_i \\ \mathrm{E}(u_i) &= 0 \end{aligned}$$

1.2 古典的線形回帰モデル（p. 259）

定義 1. $u_1, \dots, u_n$ が無相関で分散が均一な線形回帰モデルを**古典的線形回帰モデル**という．

注 1. すなわち

$$\begin{aligned} y_i &= \alpha + \beta x_i + u_i \\ \mathrm{E}(u_i) &= 0 \\ \mathrm{var}(u_i) &= \sigma^2 \\ \mathrm{cov}(u_i, u_j) &= 0 \quad \text{for } i \neq j \end{aligned}$$

1.3 古典的正規線形回帰モデル（p. 267）

定義 2. $u_1, \dots, u_n$ が独立に $\mathrm{N}\left(0, \sigma^2\right)$ に従う線形回帰モデルを**古典的正規線形回帰モデル**という．

注 2. すなわち

$$\begin{aligned} y_i &= \alpha + \beta x_i + u_i \\ \{u_i\} &\sim \mathrm{IN}\left(0, \sigma^2\right) \end{aligned}$$

または

$$\{y_i\} \sim \mathrm{IN}\left(\alpha + \beta x_i, \sigma^2\right)$$

ただし $\mathrm{IN}(.,.)$ は独立な $\mathrm{N}(.,.)$ の意味.

2 回帰係数の推定

2.1 OLS 推定量（p. 260）

簡単化のため定数項のない単回帰モデルで考える．y_i の x_i 上への線形回帰モデルは

$$\begin{aligned} y_i &= \beta x_i + u_i \\ \mathrm{E}(u_i) &= 0 \end{aligned}$$

β の OLS 推定量は

$$\begin{aligned} b &= \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \\ &= \frac{\sum_{i=1}^n x_i (\beta x_i + u_i)}{\sum_{i=1}^n x_i^2} \\ &= \beta + \frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2} \end{aligned}$$

2.2 OLS 推定量の分布（p. 266）

定理 1.

$$\mathrm{E}(b) = \beta$$

証明. 期待値の線形性より

$$\begin{aligned} \mathrm{E}(b) &= \mathrm{E}\left(\beta + \frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2}\right) \\ &= \beta + \frac{\sum_{i=1}^n x_i \mathrm{E}(u_i)}{\sum_{i=1}^n x_i^2} \\ &= \beta \end{aligned}$$

□

定理 2. 古典的線形回帰モデルなら

$$\mathrm{var}(b) = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}$$

証明. $u_1, \dots, u_n$ は無相関で分散が均一なので

$$\begin{aligned} \mathrm{var}(b) &= \mathrm{var}\left(\beta + \frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2}\right) \\ &= \mathrm{var}\left(\frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2}\right) \\ &= \frac{\mathrm{var}(x_1 u_1) + \cdots + \mathrm{var}(x_n u_n)}{(\sum_{i=1}^n x_i^2)^2} \\ &= \frac{x_1^2 \mathrm{var}(u_1) + \cdots + x_n^2 \mathrm{var}(u_n)}{(\sum_{i=1}^n x_i^2)^2} \\ &= \frac{\sigma^2 \sum_{i=1}^n x_i^2}{(\sum_{i=1}^n x_i^2)^2} \\ &= \frac{\sigma^2}{\sum_{i=1}^n x_i^2} \end{aligned}$$

□

定理 3. 古典的正規線形回帰モデルなら

$$b \sim \mathrm{N}\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

証明. b は $(y_1, \dots, y_n)$ の線形変換だから正規分布.

□

例 1 (母平均の OLS 推定)**.** 平均 μ, 分散 σ^2 の母集団分布から抽出した無作為標本を $(y_1, \dots, y_n)$ とする．$\mathrm{E}(y_i) = \mu$, $\mathrm{var}(y_i) = \sigma^2$ より

$$\begin{aligned} y_i &= \mu \cdot 1 + u_i \\ \mathrm{E}(u_i) &= 0 \\ \mathrm{var}(u_i) &= \sigma^2 \\ \mathrm{cov}(u_i, u_j) &= 0 \quad \text{for } i \neq j \end{aligned}$$

μ の OLS 推定量は

$$
\begin{aligned}
\hat{\mu} &= \frac{\sum_{i=1}^{n} 1 \cdot y_i}{\sum_{i=1}^{n} 1^2} \\
&= \frac{\sum_{i=1}^{n} y_i}{n} \\
&= \bar{y}
\end{aligned}
$$

すなわち母平均の OLS 推定量は標本平均．$\hat{\mu}$ の期待値と分散は

$$
\begin{aligned}
\mathrm{E}(\hat{\mu}) &= \mathrm{E}(\bar{y}) \\
&= \mu \\
\mathrm{var}(\hat{\mu}) &= \mathrm{var}(\bar{y}) \\
&= \frac{\sigma^2}{n}
\end{aligned}
$$

2.3　OLS 推定量の性質（p. 266）

定義 3. 被説明変数の線形関数で表される推定量を**線形推定量**という．

注 3. OLS 推定量は線形推定量.

定義 4. 不偏な線形推定量を**線形不偏推定量**という．

注 4. OLS 推定量は線形不偏推定量.

定義 5. 分散が最小となる線形不偏推定量を**最良線形不偏推定量**（*Best Linear Unbiased Estimator, BLUE*）という．

定理 4 (ガウス＝マルコフ定理)**.** *古典的線形回帰モデルの回帰係数の OLS 推定量は BLUE.*

証明.「計量経済学」で扱うので省略. □

3　誤差分散の推定

3.1　誤差分散の不偏推定量（p. 267）

$\hat{y}_i := bx_i$ とすると，σ^2 の推定量は

$$
s^2 := \frac{1}{n-k} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
$$

ただし k は推定する係数の数.

定理 5.

$$
\mathrm{E}\left(s^2\right) = \sigma^2
$$

証明.「計量経済学」で扱うので省略. □

定理 6. 古典的正規線形回帰モデルなら

$$
\frac{(n-k)s^2}{\sigma^2} \sim \chi^2(n-k)
$$

証明.「計量経済学」で扱うので省略. □

定理 7. 古典的正規線形回帰モデルなら b と s^2 は独立.

証明.「計量経済学」で扱うので省略. □

3.2　標準誤差（p. 267）

定義 6. 推定量の標準偏差の推定値を**標準誤差**という.

注 5. b の標準誤差は

$$
\mathrm{s.e.}(b) = \sqrt{\frac{s^2}{\sum_{i=1}^{n} x_i^2}}
$$

4　回帰係数の t 検定

4.1　分散が既知

簡単化のため定数項のない単回帰モデルで考える．y_i の x_i 上への古典的正規線形回帰モデルは

$$
\begin{aligned}
y_i &= \beta x_i + u_i \\
\{u_i\} &\sim \mathrm{IN}\left(0, \sigma^2\right)
\end{aligned}
$$

次の片側検定問題を考える．

$$
H_0 : \beta = c \quad \text{vs} \quad H_1 : \beta > c
$$

β の OLS 推定量は

$$
b = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}
$$

したがって

$$
b \sim \mathrm{N}\left(\beta, \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}\right)
$$

標準化すると

$$
\frac{b-\beta}{\sqrt{\sigma^2 / \sum_{i=1}^{n} x_i^2}} \sim \mathrm{N}(0,1)
$$

$H_0 : \beta = c$ を代入すると，検定統計量は

$$
Z := \frac{b-c}{\sqrt{\sigma^2 / \sum_{i=1}^{n} x_i^2}}
$$

H_0 の下で

$$
Z \sim \mathrm{N}(0,1)
$$

標準正規分布表より棄却域を定める．

4.2　分散が未知（p. 268）

σ^2 の推定量 s^2 について

$$\frac{(n-k)s^2}{\sigma^2} \sim \chi^2(n-k)$$

また b と s^2 は独立．したがって σ^2 を s^2 に置き換えると

$$\frac{b-\beta}{\sqrt{s^2/\sum_{i=1}^n x_i^2}} \sim \mathrm{t}(n-k)$$

$H_0 : \beta = c$ を代入すると，検定統計量（t 統計量）は

$$t := \frac{b-c}{\sqrt{s^2/\sum_{i=1}^n x_i^2}}$$

H_0 の下で

$$t \sim \mathrm{t}(n-k)$$

t 分布表より棄却域を定める．

4.3　t 値（p. 270）

定義 7. $H_0 : \beta = 0$ を検定する t 統計量の値を **t 値**という．

注 6. すなわち

$$t = \frac{b}{\mathrm{s.e.}(b)}$$

5　回帰係数の F 検定

5.1　分散が既知

簡単化のため定数項のない単回帰モデルで考える．y_i の x_i 上への古典的正規線形回帰モデルは

$$\begin{aligned} y_i &= \beta x_i + u_i \\ \{u_i\} &\sim \mathrm{IN}\left(0, \sigma^2\right) \end{aligned}$$

次の両側検定問題を考える．

$$H_0 : \beta = c \quad \text{vs} \quad H_1 : \beta \neq c$$

β の OLS 推定量は

$$b = \beta + \frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2}$$

したがって

$$b \sim \mathrm{N}\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

標準化すると

$$\frac{b-\beta}{\sqrt{\sigma^2/\sum_{i=1}^n x_i^2}} \sim \mathrm{N}(0,1)$$

2 乗すると

$$\frac{(b-\beta)^2\sum_{i=1}^n x_i^2}{\sigma^2} \sim \chi^2(1)$$

$H_0 : \beta = c$ を代入すると，検定統計量は

$$\chi^2 := \frac{(b-c)^2\sum_{i=1}^n x_i^2}{\sigma^2}$$

H_0 の下で

$$\chi^2 \sim \chi^2(1)$$

χ^2 分布表より棄却域を定める．

5.2　分散が未知（p. 273）

σ^2 の推定量 s^2 について

$$\frac{(n-k)s^2}{\sigma^2} \sim \chi^2(n-k)$$

また b と s^2 は独立．したがって σ^2 を s^2 に置き換えると

$$\begin{aligned} \frac{(b-\beta)^2\sum_{i=1}^n x_i^2}{s^2} &= \frac{(b-\beta)^2\sum_{i=1}^n x_i^2/\sigma^2}{[(n-k)s^2/\sigma^2]/(n-k)} \\ &\sim \mathrm{F}(1, n-k) \end{aligned}$$

$H_0 : \beta = c$ を代入すると，検定統計量（F 統計量）は

$$F := \frac{(b-c)^2\sum_{i=1}^n x_i^2}{s^2}$$

H_0 の下で

$$F \sim \mathrm{F}(1, n-k)$$

F 分布表より棄却域を定める．

5.3　F 値（p. 274）

定義 8. $H_0 : \beta = 0$ を検定する F 統計量の値を **F 値**という．

注 7. 次の重回帰モデルを考える．

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{i,2} + \dots + \beta_k x_{i,k} + u_i \\ \{u_i\} &\sim \mathrm{IN}\left(0, \sigma^2\right) \end{aligned}$$

ただし β_1 は定数項．次の両側検定問題を考える．

$$H_0 : \begin{pmatrix} \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = \mathbf{0} \quad \text{vs} \quad H_1 : \begin{pmatrix} \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \neq \mathbf{0}$$

このとき H_0 の下で

$$F \sim \mathrm{F}(k-1, n-k)$$

6　今日のキーワード

古典的線形回帰モデル，古典的正規線形回帰モデル，線形推定量，線形不偏推定量，最良線形不偏推定量（BLUE），ガウス＝マルコフ定理，標準誤差，t 値，F 値

7　次回までの準備

提出　宿題 7
復習　教科書第 13 章 2.2–3 節，復習テスト 25
予習　教科書第 13 章 4 節