中級統計学：復習テスト 25

学籍番号＿＿＿＿＿＿＿＿＿＿＿氏名＿＿＿＿＿＿＿＿＿＿＿

2026 年 1 月 9 日

注意：すべての質問に解答しなければ提出とは認めない．正答に修正した上で，復習テスト 21〜26 を順に重ねて左上でホチキス止めし，定期試験実施日（1 月 27 日の予定）に提出すること．

1. 2 変量データを $((y_1, x_1), \dots, (y_n, x_n))$ とする．y_i の x_i 上への定数項なしの古典的線形回帰モデルは

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

β の OLS 推定量を b とする．

(a) b を式で与えなさい．

(b) b の期待値を求めなさい．

(c) b の分散を求めなさい．

2. 2 変量データを $((y_1, x_1), \ldots, (y_n, x_n))$ とする．y_i の x_i 上への定数項なしの古典的正規線形回帰モデルは

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

β の OLS 推定量を b とする．σ^2 を既知として次の片側検定問題を考える．

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

(a) b の分布を求めなさい.

(b) 検定統計量を与えなさい.

(c) 検定統計量の H_0 の下での分布を与えなさい.

(d) 分布表を利用して，有意水準 5 %の検定の棄却域を定めなさい.

(e) 分布表を利用して，検定統計量の値が 2.0 のときの p 値を求めなさい.

解答例

1. (a) OLS 問題は

$$\begin{aligned}&\min_{b} \quad \sum_{i=1}^{n}(y_i - bx_i)^2 \\ &\text{and} \quad b \in \mathbb{R}\end{aligned}$$

1 階の条件より

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

すなわち

$$\sum_{i=1}^{n} x_i(y_i - bx_i) = 0$$

正規方程式は

$$\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} x_i^2 = 0$$

$\sum_{i=1}^{n} x_i^2 \neq 0$ なら

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

(b) 前問の式に回帰式 $y_i = \beta x_i + u_i$ を代入すると

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

期待値の線形性より

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

(c) $u_1, \dots, u_n$ は無相関で分散が均一なので

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

2. (a) b は $(y_1, \dots, y_n)$ の線形変換だから正規分布．Q1(b)(c) より

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

(b) 前問の結果を標準化すると

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

$H_0 : \beta = c$ を代入すると，検定統計量は

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

(c) H_0 の下で

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

(d) 標準正規分布表より H_0 の下で

$$\Pr[Z \geq 1.65] = .05$$

したがって棄却域は $[1.65, \infty)$.

(e) 標準正規分布表より H_0 の下で

$$\Pr[Z \geq 2.00] = .02275$$

したがって p 値=.02275.