経済統計：期末試験

村澤 康友

2016 年 8 月 1 日

注意：3 問とも解答すること．結果より思考過程を重視するので，途中計算等も必ず書くこと（部分点は大いに与えるが，結果のみの解答は 0 点とする）．

1. （20 点）以下の用語の定義を式または言葉で書きなさい（各 20 字程度）．
   （a）t 統計量
   （b）正規方程式
   （c）線形推定量
   （d）標準誤差
2. （30 点）2 変量データを $((y_1, x_1), \ldots, (y_n, x_n))$ とする．y_i の x_i 上への定数項のない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 を OLS で推定したい．
   （a）OLS 問題を書きなさい．
   （b）β の OLS 推定量を求めなさい．
   （c）OLS 推定量の分布を求めなさい．
3. （50 点）某大学経済学部 2 回生の男女について，どちらが経済学の理解度のばらつきが大きいかを知りたい．そこで（復元）無作為抽出した男子 64 人，女子 32 人に対して試験を行い，次の結果を得た．

	平均点	（標本）標準偏差
男子	50	28
女子	60	24

   正規母集団を仮定し，母数と統計量の区別に注意して，以下の問いに答えなさい．
   （a）検定問題を定式化しなさい（問題意識を踏まえること）．
   （b）標本分散の比の分布を求めなさい．
   （c）検定統計量を与えなさい．
   （d）H_0 の下で検定統計量の分布を求め，有意水準 5 ％の検定の棄却域を定めなさい．
   （e）検定を実行し，結果を説明しなさい．

解答例

1. 統計学の基本用語
   (a) H_0 の下で t 分布にしたがう検定統計量
      - 「H_0 の下で」がなければ 1 点減
      - 「検定」がなければ 0 点
   (b) OLS 問題の 1 階の条件を整理した式
   (c) 被説明変数の線形関数で表される推定量
   (d) 推定量の標準偏差の推定値
2. OLS
   (a) OLS 問題は

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

      - min がなければ 0 点

   (b) OLS 問題の 1 階の条件は

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

すなわち

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

正規方程式は

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

OLS 推定量は

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

   (c) β の OLS 推定量は

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

したがって

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

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

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

3. 2 標本問題

(a) 母平均を μ_X, μ_Y，母分散を σ_X^2, σ_Y^2 とすると

$$
H_0 : \sigma_X^2 \le \sigma_Y^2,\ \mu_X, \mu_Y \in \mathbb{R} \quad \text{vs} \quad H_1 : \sigma_X^2 > \sigma_Y^2,\ \mu_X, \mu_Y \in \mathbb{R}
$$

- 両側検定は 0 点
- 逆の不等号の片側検定は明らかに H_0 が棄却されないので 0 点.

(b) 標本の大きさを m, n，標本分散を s_X^2, s_Y^2 とすると

$$
\begin{aligned}
\frac{(m-1)s_X^2}{\sigma_X^2} &\sim \chi^2(m-1) \\
\frac{(n-1)s_Y^2}{\sigma_Y^2} &\sim \chi^2(n-1) \\
\frac{s_X^2/s_Y^2}{\sigma_X^2/\sigma_Y^2} &\sim \mathrm{F}(m-1, n-1)
\end{aligned}
$$

$\mathrm{F}(m-1, n-1)$ の cdf を $F(.)$ とすると，任意の x について

$$
\begin{aligned}
\Pr\left[\frac{s_X^2}{s_Y^2} \le x\right] &= \Pr\left[\frac{s_X^2/s_Y^2}{\sigma_X^2/\sigma_Y^2} \le \frac{x}{\sigma_X^2/\sigma_Y^2}\right] \\
&= F\left(\frac{x}{\sigma_X^2/\sigma_Y^2}\right)
\end{aligned}
$$

- $s_X^2/s_Y^2 \sim \mathrm{F}(m-1, n-1)$ は H_0 の下でなければ間違いなので 0 点

（c）検定統計量は

$$F := \frac{s_X^2}{s_Y^2}$$

（d）H_0 の下で $F \sim \mathrm{F}(63, 31)$．F 分布表より F に関する棄却域は $[1.726, \infty)$．より正確には

$$\begin{aligned} p &:= \frac{1/63 - 1/120}{1/60 - 1/120} \\ &= \frac{57/(63 \cdot 120)}{1/120} \\ &= \frac{57}{63} \end{aligned}$$

として

$$\begin{aligned} F_{.05}(63, 31) &= pF_{.05}(60, 31) + (1-p)F_{.05}(120, 31) \\ &= \frac{57}{63} \cdot 1.726 + \frac{6}{63} \cdot 1.670 \\ &\approx 1.721 \end{aligned}$$

と補間する．

（e）検定統計量の値は

$$\begin{aligned} F &:= \frac{28^2}{24^2} \\ &= \frac{7^2}{6^2} \\ &= \frac{49}{36} \\ &\approx 1.361 \end{aligned}$$

$F < 1.726$ より有意水準 5 %で H_0 は H_1 に対して棄却されない．すなわち男子の方が理解度のばらつきが大きいとは言えない．