計量経済 II：期末試験

村澤 康友

2017 年 1 月 24 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) 棄却域 (b) χ^2 統計量 (c) 適合度検定 (d) 決定係数
2. （30 点）ある将棋の対局の中終盤 50 手の将棋ソフトとの一致率は X 棋士が 88%（22/25），Y 棋士が 60%（15/25）であった．両棋士の（無限仮説母集団における）真の一致率を p_X, p_Y，この対局の中終盤各 25 手における一致率（＝標本平均）を $\hat{p}_X, \hat{p}_Y$ とする．次の片側検定問題を考える．

$$H_0 : p_X = p_Y \quad \text{vs} \quad H_1 : p_X > p_Y.$$

   各対局者と将棋ソフトの指し手の一致／不一致は独立かつ同一なベルヌーイ試行であり，また互いの指し手についても独立と仮定する．
   (a) $\hat{p}_X, \hat{p}_Y$ の漸近分布を求めなさい．また $\hat{p}_X - \hat{p}_Y$ の漸近分布を求めなさい．
   (b) 検定統計量を与えなさい．それは H_0 の下でどのような分布に近似的に従うか？
   (c) 有意水準 1 %の検定の棄却域を定め，標本の大きさを $n_X = n_Y = 25$ として検定を実行しなさい．
   ※数値例はフィクションです．また分析の目的は両棋士の棋風の差異の検証であり，本データのみで不正行為の有無の検証は不可能です．
3. （50 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 の OLS 推定量を b，σ^2 の不偏推定量を s^2 とする．次の片側検定問題を考える．

$$H_0 : \beta = 1 \quad \text{vs} \quad H_1 : \beta > 1$$

   (a) OLS 問題を書きなさい．
   (b) b を式で与えなさい．
   (c) b の分布を求めなさい．
   (d) s^2 を式で与えなさい．
   (e) 検定統計量を与えなさい．それは H_0 の下でどのような分布に従うか？

解答例

1. 統計学の基本用語
   (a) 標本（検定統計量）の値域で H_0 を棄却する領域.
      - 「H_0 を棄却する領域」のみは 2 点.
      - 「棄却する領域」のみは 0 点.
   (b) H_0 の下で χ^2 分布にしたがう検定統計量.
      - 「χ^2 分布にしたがう検定統計量」のみは 1 点.
      - 「χ^2 分布にしたがう統計量」のみは 0 点.
   (c) 母集団分布に対する標本の適合度の検定.
      - 「適合度の検定」のみは 0 点.
   (d) $R^2 := \mathrm{ESS}/\mathrm{TSS}$.
2. 2 標本の母比率の差の検定
   (a)

$$\hat{p}_X \overset{a}{\sim} \mathrm{N}\left(p_X, \frac{p_X(1-p_X)}{n_X}\right)$$
$$\hat{p}_Y \overset{a}{\sim} \mathrm{N}\left(p_Y, \frac{p_Y(1-p_Y)}{n_Y}\right)$$

両者は独立なので

$$\hat{p}_X - \hat{p}_Y \overset{a}{\sim} \mathrm{N}\left(p_X - p_Y, \frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}\right)$$

   - $\hat{p}_X, \hat{p}_Y$ の漸近分布で 5 点，$\hat{p}_X - \hat{p}_Y$ の漸近分布で 5 点.
   - 統計量の漸近分布に統計量があったら 0 点.

   (b) 検定統計量は

$$Z := \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}_X(1-\hat{p}_X)/n_X + \hat{p}_Y(1-\hat{p}_Y)/n_Y}}$$

H_0 の下で $Z \overset{a}{\sim} \mathrm{N}(0,1)$.
   - 検定統計量で 5 点，H_0 の下での分布で 5 点.
   - 統計量の定義式に未知母数があったら 0 点.

   (c) 棄却域は $[2.33, \infty)$. 検定統計量の値は

$$\begin{aligned}
Z &:= \frac{.88 - .60}{\sqrt{.88(1-.88)/25 + .60(1-.60)/25}} \\
&= \frac{.28}{\sqrt{(.1056 + .24)/25}} \\
&= \frac{.28 \cdot 5}{\sqrt{.3456}} \\
&= \frac{1.40}{\sqrt{.3456}} \\
&\approx 2.38
\end{aligned}$$

したがって H_0 を棄却して H_1 を採択.

- 棄却域で 5 点，検定統計量の値で 5 点.

3. OLS

（a）OLS 問題は

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

（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

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

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

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

- 期待値で 2 点，分散で 4 点，分布で 4 点．

(d)

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

(e) 標準化すると

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

σ^2 を s^2 に置き換えると

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

検定統計量は

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

H_0 の下で

$$
t \sim \mathrm{t}(n-1)
$$

- 検定統計量で 5 点，H_0 の下での分布で 5 点．
- 統計量の定義式に未知母数があったら 0 点．