中級統計学：定期試験

村澤 康友

2025 年 1 月 21 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）．図や数式を用いる場合は，グラフや記号の意味を明記すること．何でも持ち込み可だが，通信機器の所持と物品の貸借は厳禁.

1. （20 点）以下で定義される統計学の専門用語をそれぞれ書きなさい.
   （a）統計的仮説の真偽を標本から判定すること
   （b）X の 1 %の増加に対する Y の変化率
   （c）誤差項が無相関で分散が均一な線形回帰モデル
   （d）回帰変動÷総変動
2. （30 点）高校での成績が大学での成績の予測に役立つかどうかを検証したい．そこで大学での GPA（colgpa）を高校での GPA（hsgpa）と英語（vsat）・数学（msat）の入試得点の対数値で説明する重回帰モデルを推定し，次の結果を得た．

$$\widehat{\text{colgpa}} = \underset{(-4.258)}{-4.12477} + \underset{(6.651)}{0.403586}\,\text{hsgpa} + \underset{(2.518)}{0.323233}\,\text{l\_vsat} + \underset{(3.471)}{0.548688}\,\text{l\_msat}$$

$$T = 427 \quad \bar{R}^2 = 0.2122 \quad F(3,423) = 39.253 \quad \hat{\sigma} = 0.48002$$

(括弧内は、t 統計量)

   （a）高校での GPA から大学での GPA への限界効果の OLS 推定値とその標準誤差は幾らか？
   （b）問題意識を踏まえ，検定問題を定式化しなさい.
   （c）有意水準 5 %の検定の棄却域を定め，検定統計量の値を示し，検定結果を説明しなさい.
3. （50 点）利き手の割合の男女差の有無を検証したい．そこで 100 人を無作為抽出して調査し，次表の結果を得た．この結果を母比率の差の検定と独立性の検定の 2 つの手法で分析する．

	右利き	左利き	計
男性	43	9	52
女性	44	4	48
計	87	13	100

   （a）問題意識を踏まえ，母比率の差の検定と独立性の検定の検定問題を，それぞれ定式化しなさい.
   （b）母比率の差の検定統計量を定義し，その H_0 の下での漸近分布を示しなさい.
   （c）有意水準 5 %の検定の棄却域を定め，検定統計量の値を計算し，検定結果を説明しなさい.
   （d）独立性の χ^2 検定統計量を定義し，その H_0 の下での漸近分布を示しなさい.
   （e）有意水準 5 %の検定の棄却域を定め，χ^2 検定統計量の値を計算し，検定結果を説明しなさい.

解答例

1. 統計学の基本用語
   (a)（仮説）検定
   (b) 弾力性
       - 「弾性値」でも OK.
   (c) 古典的線形回帰モデル
   (d) 決定係数
       - 「自由度修正済み決定係数」は 1 点.
2. 回帰分析
   (a) OLS 推定値は 0.403586．標準誤差は $0.403586/6.651 \approx 0.061$.
       - 各 5 点.
   (b) hsgpa の回帰係数を β とすると，検定問題は

$$H_0 : \beta = 0 \quad \text{vs} \quad H_1 : \beta \neq 0$$

       - 記号の説明なしは 1 点減.
       - 片側検定は 1 点減.
   (c) H_0 の下での t 統計量の分布 t(423) は $\mathrm{t}(\infty) = \mathrm{N}(0, 1)$ で近似できるので，有意水準 5 %の検定の棄却域は $(-\infty, -1.96] \cup [1.96, \infty)$．t 値=6.651 は棄却域に入るので，$H_0$ を棄却して H_1 を採択する．すなわち高校での成績は大学での成績の予測に役立つ.
       - 棄却域 4 点，検定統計量 4 点，正しい棄却域と検定統計量に基づく検定結果 2 点.
3. 母比率の差の検定と独立性の検定
   (a) 男性の母集団における右利きの母比率を p_X，女性の母集団における右利きの母比率を p_Y とすると，母比率の差の検定問題は

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

   母集団における右利きの男性の母比率を $p_{X,R}$，左利きの男性の母比率を $p_{X,L}$，右利きの女性の母比率を $p_{Y,R}$，左利きの女性の母比率を $p_{Y,L}$，男性の母比率を $p_{X,.}$，女性の母比率を $p_{Y,.}$，右利きの母比率を $p_{.,R}$，左利きの母比率を $p_{.,L}$ とすると，独立性の検定問題は

$$H_0 : \begin{bmatrix} p_{X,R} & p_{X,L} \\ p_{Y,R} & p_{Y,L} \end{bmatrix} = \begin{bmatrix} p_{X,.}p_{.,R} & p_{X,.}p_{.,L} \\ p_{Y,.}p_{.,R} & p_{Y,.}p_{.,L} \end{bmatrix} \quad \text{vs} \quad H_1 : \begin{bmatrix} p_{X,R} & p_{X,L} \\ p_{Y,R} & p_{Y,L} \end{bmatrix} \neq \begin{bmatrix} p_{X,.}p_{.,R} & p_{X,.}p_{.,L} \\ p_{Y,.}p_{.,R} & p_{Y,.}p_{.,L} \end{bmatrix}$$

       - 各 5 点.
       - 母比率の差の片側検定は 1 点.
       - 独立性の検定問題は 1 つのみでも可（例えば $H_0 : p_{X,R} = p_{X,.}p_{.,R}$ vs $H_1 : p_{X,R} \neq p_{X,.}p_{.,R}$).
   (b) 男性の標本の大きさを n_X，女性の標本の大きさを n_X，男性の標本における右利きの標本比率を $\hat{p}_X$，女性の標本における右利きの標本比率を $\hat{p}_Y$ とすると，

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

したがって

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

標準化すると

$$\frac{\hat{p}_X - \hat{p}_Y - (p_X - p_Y)}{\sqrt{p_X(1-p_X)/n_X + p_Y(1-p_Y)/n_Y}} \overset{a}{\sim} \mathrm{N}(0,1)$$

$H_0 : p_X = p_Y$ を代入し，母比率を標本比率に置き換えると，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 5 点，漸近分布 5 点.
- 分母の標本比率 $\hat{p}_X, \hat{p}_Y$ はプールした標本比率 $\hat{p}$ でも OK.

(c) 有意水準 5 %の検定の棄却域は $(-\infty, -1.96] \cup [1.96, \infty)$．検定統計量の値は

$$\begin{aligned} Z &:= \frac{43/52 - 44/48}{\sqrt{(43/52)(9/52)/52 + (44/48)(4/48)/48}} \\ &\approx -1.36 \end{aligned}$$

Z は棄却域に入らず H_0 は棄却されない．すなわち利き手の割合に男女差があるとは言えない.

- 棄却域 4 点，検定統計量 4 点，正しい棄却域と検定統計量に基づく検定結果 2 点.
- 検定結果のみの解答は 0 点.

(d) 標本の大きさを n とする．標本における右利きの男性の標本比率を $\hat{p}_{X,R}$，左利きの男性の標本比率を $\hat{p}_{X,L}$，右利きの女性の標本比率を $\hat{p}_{Y,R}$，左利きの女性の標本比率を $\hat{p}_{Y,L}$，男性の標本比率を $\hat{p}_{X,.}$，女性の標本比率を $\hat{p}_{Y,.}$，右利きの標本比率を $\hat{p}_{.,R}$，左利きの標本比率を $\hat{p}_{.,L}$ とすると，χ^2 検定統計量は

$$\chi^2 := \frac{n(\hat{p}_{X,R} - \hat{p}_X\hat{p}_R)^2}{\hat{p}_{X,.}\hat{p}_{.,R}} + \frac{n(\hat{p}_{X,L} - \hat{p}_X\hat{p}_L)^2}{\hat{p}_{X,.}\hat{p}_{.,R}} + \frac{n(\hat{p}_{Y,R} - \hat{p}_Y\hat{p}_R)^2}{\hat{p}_{Y,.}\hat{p}_{.,L}} + \frac{n(\hat{p}_{Y,L} - \hat{p}_Y\hat{p}_L)^2}{\hat{p}_{Y,.}\hat{p}_{.,L}}$$

H_0 の下で $\chi^2 \overset{a}{\sim} \chi^2(1)$.

- 検定統計量 5 点，漸近分布 5 点.

(e) 有意水準 5 %の検定の棄却域は $[3.84146, \infty)$．χ^2 検定統計量の値は

$$\begin{aligned} \chi^2 &:= \frac{100(0.43 - 0.52 \cdot 0.87)^2}{0.52 \cdot 0.87} + \frac{100(0.09 - 0.52 \cdot 0.13)^2}{0.52 \cdot 0.13} \\ &\quad + \frac{100(0.44 - 0.48 \cdot 0.87)^2}{0.48 \cdot 0.87} + \frac{100(0.04 - 0.48 \cdot 0.13)^2}{0.48 \cdot 0.13} \\ &\approx 1.78 \end{aligned}$$

χ^2 は棄却域に入らず H_0 は棄却されない．すなわち性別と利き手の割合に関係がある（独立でない）とは言えない.

- 棄却域 4 点，検定統計量 4 点，正しい棄却域と検定統計量に基づく検定結果 2 点.
- 検定結果のみの解答は 0 点.