経済統計学 - 宿題 No.3

蛭川雅之

提出期限：2025年12月16日（火）11:00（厳守）

注意事項

- 学生相互間で相談をしても構いませんが、答案は必ず**個人で作成**してください。
- 答案は手書きでもタイプ打ちでも構いませんが、**最初のページに必ず学籍番号と氏名を明記**してください。
- 答案を作成する際には、忘れずに**問題番号を記入**してください。具体的な答案の作成要領は次の通りです。

  1. 第1部の各問題については、計算根拠などをできるだけ具体的に説明してください。必要に応じて途中計算を書いても構いません。なお、説明がなくただ結果だけが書いてある場合、その問題の点数は零点とします。また、**数値は電卓を用いて計算**してください。
  2. 第2部の各問題については、**指示に沿った資料を添付**してください。コメントが必要な場合は、それらに直接書き込んで構いません。

- 答案は第1部・第2部をまとめて**左肩にホチキス止め**をした上で提出してください。
- **提出期限を厳守**してください。いかなる理由があろうと、期限後に提出された答案は受け取りません。
- ルールに違反して答案を書いた（例：本講義に履修登録していない学生に助力を求める、他の学生の答案もしくは過去問の解答例を丸写しする）場合、内容の如何を問わず減点の対象となりますので注意してください。

第 1 部

問題 1-1

n 個の観測値 $\{X_i\}_{i=1}^{n}$ を用いて切片のみの回帰モデル

$$X_i = \mu + \epsilon_i \ (i = 1, \ldots, n) \tag{1}$$

を最小 2 乗推定する。なお、誤差項は $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$（ただし、$\sigma^2 > 0$）を満たすものとする。このとき、以下の問いに答えよ。

(1)　μ の最小 2 乗推定量 $\hat{\mu}$ を観測値 $\{X_i\}_{i=1}^{n}$ を用いて表せ。なお、この問題については、結果のみ書いてあれば十分である。

(2)　$\hat{\mu}$ の期待値および分散を求めよ。

(3)　ノート No.11 の 18〜19 ページの手順に従い、$\hat{\mu}$ の標準誤差 $SE(\hat{\mu})$ を求めよ。【ヒント：σ^2 の一致推定量 s^2 がどのように表現されるかを考えること。】

(4)　既に学んだ通り、観測値 $\{X_i\}_{i=1}^{n}$ が正規母集団 $N(\mu, \sigma^2)$ からの n 個の無作為標本である場合、帰無仮説 $H_0 : \mu = \mu_0$ の下での検定統計量は

$$t = \frac{\sqrt{n-1}\left(\bar{X} - \mu_0\right)}{\hat{\sigma}}$$

で与えられる。ただし、

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

は X の標本分散である。この検定統計量は、回帰モデル (1) に対する帰無仮説 $H_0 : \mu = \mu_0$ の下での検定統計量

$$t = \frac{\hat{\mu} - \mu_0}{SE(\hat{\mu})}$$

に書き換えられることを示せ。

問題 1-2

ある学期の出席率（$atndrte$）を前学期までの累積 GPA（$priGPA$）および ACT（＝米国の大学への進学希望者が大学を受験する際の学力を測る共通テストの一つ）の得点（ACT）で説明する回帰式

$$atndrte_i = \beta_0 + \beta_1 priGPA_i + \beta_2 ACT_i + \epsilon_i$$

を Stata を利用して最小 2 乗推定し、次のような結果を得た。このとき、以下の問いに答えよ。なお、標準誤差は均一分散を前提とするものである。また、設問箇所以外にも意図的に空欄となっている箇所が多数あることに注意せよ。

```
. reg atndrte prigpa act

      Source |       SS           df       MS      Number of obs   =
-------------+----------------------------------   F(2, 677)       =
       Model |  57336.7612                         Prob > F        =
    Residual |  139980.564                         R-squared       =
-------------+----------------------------------   Adj R-squared   =
       Total |                   679               Root MSE        =

------------------------------------------------------------------------------
     atndrte |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      prigpa |   17.26059   1.083103
         act |  -1.716553    .169012
       _cons |    75.7004   3.884108
------------------------------------------------------------------------------
```

(1)　観測値の個数 n を求めよ。

(2)　決定係数 R^2 および自由度修正済み決定係数 $\bar{R}^2$ をそれぞれ小数第 4 位まで求めよ。

(3)　回帰式の標準誤差を小数第 3 位まで求めよ。

(4)　回帰式全体の有意性に関する検定統計量の値（F 値）を小数第 3 位まで求めよ。

(5)　最小 2 乗推定値 $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)$ に対する t 値をそれぞれ小数第 2 位まで求めよ。

(6)　係数 β_2 の 95% 信頼区間を小数第 3 位まで求め、この信頼区間を利用して帰無仮説 $H_0 : \beta_2 = -1$ を対立仮説 $H_1 : \beta_2 \neq -1$ に対し有意水準 5% で検定せよ。

(7)　新たに当該学期の GPA（$termGPA$）および期末試験の素点（$final$）を説明変数として加えた回帰式

$$atndrte_i = \beta_0 + \beta_1 priGPA_i + \beta_2 ACT_i + \beta_3 termGPA_i + \beta_4 final_i + \epsilon_i$$

を最小 2 乗推定したところ、決定係数 $R^2 = 0.4400$ が得られた。このとき、帰無仮説 $H_0 : \beta_3 = \beta_4 = 0$ を有意水準 5% で検定せよ。なお、臨界値（＝自由度 $(q, n-k-1) = (2, 675)$ の F 分布の右側 5% 点）3.0091 は説明なしに使用してよい。

第 2 部

問題 2-1

　データファイル"hw3_q2-1.csv"には 1985 年における米国各州（コロンビア特別区を含む）の教師の平均給与 (S) および児童一人当たり支出額 (P) が記録されている（単位：米ドル)。ここで、(P, S) の間に線形の関係があると仮定し、回帰式

$$S_i = \beta_0 + \beta_1 P_i + \epsilon_i,\ i = 1, \ldots, 51\,(= n)$$

を考える。このとき以下の問いに答えよ。なお、本問題については Stata を利用せず、Excel を用いてノート No.11 問題 1-2 と同様の計算を実行せよ。なお、Excel ワークシートの該当箇所のプリントアウトを必ず添付せよ。

(1)　(β_0, β_1) の最小 2 乗推定値 $\left(\hat{\beta}_0, \hat{\beta}_1\right)$ を計算せよ。

(2)　(1) の結果に基づく予測値を $\hat{S}$、残差を e と表記する。さらに、$\overline{\hat{S}}$ を予測値 $\hat{S}$ の平均、$\overline{S}$ を被説明変数 S の平均とする。このとき、

$$\overline{\hat{S}} = \bar{S}$$

$$\sum_{i=1}^{n} e_i = 0$$

$$\sum_{i=1}^{n} P_i e_i = 0$$

が成り立つことを示せ。

(3)　分散分解

$$TSS = ESS + RSS$$

が成り立つことを示せ。

(4)　重相関係数 $R = r_{S\hat{S}}$、決定係数 R^2、自由度修正済み決定係数 $\bar{R}^2$、および (P, S) に関するピアソン積率相関係数 r_{PS} を計算し、

$$R^2 = r_{S\hat{S}}^2 = r_{PS}^2$$
$$\bar{R}^2 \le R^2$$

が成り立つことを示せ。

問題 2-2

データファイル“`hw3_q2-2.csv`”を用いて賃金関数

$$\ln(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + \epsilon$$

を推定する。このとき、以下の問いに答えよ。計算には Stata を用いることとし、ログファイルのプリントアウトを必ず添付せよ。なお、各変数の定義は次の通りである。

変数	定義
$wage$	1 月当たり所得（米ドル）
$educ$	教育年数（年）
$exper$	就労経験（年）
$tenure$	現在の雇用主の下での年数（年）

(1)　まず賃金関数を最小 2 乗推定し、結果をノート No.11 の 20 ページの様式に沿って報告せよ。なお、均一分散を前提とする標準誤差を計算すること。

(2)　(1) の回帰式において、「就労経験が 1 年増えることによる $\ln(wage)$ への影響」と「現在の雇用主の下でもう 1 年余計に働くことによる $\ln(wage)$ への影響」とが変わらないかどうかを仮説検定したい。$(\beta_0, \beta_1, \beta_2, \beta_3)$ を用いて帰無仮説を書き表せ。

(3)　あなたは Stata を使って (2) の帰無仮説を有意水準 5% で両側検定したい。ただし、あなたは Stata に関しては初心者のため、Stata で最小 2 乗推定値の分散共分散行列を呼び出す方法、線形制約検定用のコマンドを使用する方法のどちらも知らないとする。この場合、どのようにしてこの仮説検定を実行すればよいか。検定結果はどのようなものか。なお、検定統計量を計算する際、(1) と同様、均一分散を前提とする標準誤差を用いること。

問題 2-3

データファイル“`hw3_q2-3.csv`”は問題 1-2 で使用したものと同一のデータセットからの抜粋である。このとき、以下の問いに答えよ。計算には Stata を用いることとし、ログファイルのプリントアウトを必ず添付せよ。なお、観測値の個数 n は問題 1-2 と異なるが、これは宿題提出率が記録されていない学生のデータを除いているためである。

(1)　まず、標準化した（＝素点から平均点を引き、標準偏差で割った）期末試験の得点（$stndfnl$）をその学期の出席率（$atndrte$）および宿題提出率（$hwrte$）、前学期までの累積 GPA（$priGPA$）、ACT の得点（ACT）で説明する回帰式

$$stndfnl_i = \beta_0 + \beta_1 atndrte_i + \beta_2 hwrte_i + \beta_3 priGPA_i + \beta_4 ACT_i + \epsilon_i \tag{2}$$

を最小 2 乗推定する。次に、$stndfnl$ と $priGPA$ ないし ACT との間に非線形の関係が存在するとの前提から、回帰式 (2) に $priGPA$ および ACT の 2 次の項を加えた回帰式

$$\begin{aligned} stndfnl_i = \beta_0 &+ \beta_1 atndrte_i + \beta_2 hwrte_i + \beta_3 priGPA_i + \beta_4 ACT_i \\ &+ \beta_5 priGPA_i^2 + \beta_6 ACT_i^2 + \epsilon_i \end{aligned} \tag{3}$$

を最小 2 乗推定する。回帰式 (2)(3) の推定結果をノート No.12 の 27 ページの様式に沿って報告せよ。なお、Eicker-Huber-White の公式によるロバスト標準誤差を計算すること。

(2)　(3) で追加した 2 つの説明変数 $priGPA^2$ および ACT^2 の係数の有意性（即ち、帰無仮説 $H_0 : \beta_5 = \beta_6 = 0$）について、有意水準 1% で同時検定を行え。なお、検定統計量は Eicker-Huber-White の公式を用いて計算せよ。2 つの係数のうち少なくとも一方は有意にゼロと異なるといってよいか。