中級計量経済学２０２４練習問題(1)

以下の問いに答えなさい。（提出の必要はありません。）

１． X, Y を確率変数とし、$(Y_i, X_i),\ i = 1, 2, \cdots, n$ をデータとして、以下の問いに答えなさい。

(i) 確率変数 X, Y の密度関数を $f(x, y)$ とする。X の周辺密度関数 $f_X(x)$ と Y の周辺密度関数 $f_Y(x)$ を $f(x, y)$ で表しなさい。また、X を条件とする Y の条件付き密度関数はどのように定義されるか？

(ii) それを用いて、繰り返し期待値の法則 $E(Y) = E\{E(Y|X)\}$ が成り立つことを示しなさい。

(iii)　分散について $E[\{X - E(X)\}^2] = E[X\{X - E(X)\}]$ が成り立つことを示しなさい。

(iv)　共分散について

$$E[\{X - E(X)\}\{Y - E(Y)\}] = E[X\{Y - E(Y)\}] = E[\{X - E(X)\}Y]$$

が成り立つことを示しなさい。

(v)　標本平均 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ について、$\sum_{i=1}^{n}(X_i - \bar{X}) = 0$ が成り立つことを示しなさい。

(vi) 標本分散について $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i(X_i - \bar{X})$ が成り立つことを示しなさい。

(vii) 標本共分散について、

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) = \frac{1}{n-1}\sum_{i=1}^{n} X_i(Y_i - \bar{Y}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})Y_i$$

が成り立つことを示しなさい。

２．説明変数がひとつの回帰モデル $Y_i = \beta_0 + \beta_1 X_i + u_i$ から無作為標本 $(Y_i, X_i),\ i = 1, 2, \cdots, n$ が得られたとする。β_0, β_1 の最小二乗推定量を $(\hat{\beta}_0, \hat{\beta}_1)$、$Y_i$ の予測値を $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ とし、残差を $\hat{u}_i = Y_i - \hat{Y}_i$ とする。

(i) 最小二乗法の一階の条件を導出し、

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

と表されることを示しなさい。ただし、 $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ とする。

(ii) (i) と問 1 の結果を用いて、

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

と表されることを示しなさい。

(iii) $E(X_i) = \mu_X$ とする。 $X_i - \bar{X} = (X_i - \mu_X) - (\bar{X} - \mu_X)$ と分解し、中心極限定理と大数の法則、スラツキーの補題を用いて、$\sqrt{n}(\hat{\beta}_1 - \beta_1)$ の漸近正規性を示しなさい。

(iv) (i) で得た一階の条件から $\sum_{i=1}^{n}\hat{u}_i = 0$、$\sum_{i=1}^{n} X_i\hat{u}_i = 0$ となることを示しなさい。

(v) $Y_i - \bar{Y} = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y})$ と分解し、以上の結果を用いて

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{n}\hat{u}_i^2$$

が成り立つことを示し、その結果 $0 \leq R^2 \leq 1$ であることを示しなさい。

3．真の定数項が0の回帰モデル $Y_i = \beta_1 X_i + \epsilon_i,\ E(\epsilon_i|X_i) = 0$ から無作為標本 $(Y_1, X_1), \cdots, (Y_n, X_n)$ が得られたとする。A さんは定数項が 0 であることを知らずに定数項を含めずに

$$\min_{b_0, b_1} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$$

という最小二乗推定を行い、$\hat{\beta}_1^A$ を得たとする。他方、B さんは定数項が 0 であることを知っており

$$\min_{b_1} \sum_{i=1}^{n} (Y_i - b_1 X_i)^2$$

から $\hat{\beta}_1^B$ を得たとする。

(i) $\hat{\beta}_1^A$ と $\hat{\beta}_1^B$ が共に一致性をもつことを示しなさい。

(ii) $\hat{\beta}_1^A$ と $\hat{\beta}_1^B$ の漸近分布を求めなさい。

(iii) (ii) で求めた漸近分散の結果から、 $\hat{\beta}_1^A$ と $\hat{\beta}_1^B$ のどちらの推定量が望ましいか？

4．本当の関係が $Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_i,\ E(\epsilon_i|X_i, Z_i) = 0$ である時に、無作為標本 $(Y_i, X_i),\ i = 1, 2, \cdots, n$ が得られたとする。Z_i のデータが得られなかったため、回帰モデル $Y_i = b_0 + b_1 X_i + u_i$ を用いて、最小二乗推定を行うことを考える。$E(X) = \mu_X$, $E(Z) = \mu_Z$, $Var(X) = \sigma_X^2$, $Var(Z) = \sigma_Z^2$, $Cov(X, Z) = \sigma_{XZ}$ とする。最小二乗推定量 $\hat{\beta}_0, \hat{\beta}_1$ のそれぞれについて、$n \to \infty$ としたときのバイアス（確率収束の収束先と真の値の差）を求めなさい。その結果、どのような場合にバイアスが大きくなるか述べなさい。

5．説明変数が、0または1の値を取るダミー変数 D_i である回帰モデル

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

から 無作為標本 $(Y_i, D_i),\ i = 1, 2, \cdots, n$ が得られたとする。また、データは $i = 1, 2, \cdots, n_1$ について $D_i = 1$、$i = n_1 + 1, \cdots, n_1$ について $D_i = 0$ となるように並べられているとする。そのとき、β_1 の OLS 推定量 $\hat{\beta}_1$ は、$D_i = 1$ の Y の平均と $D_i = 0$ の Y の平均の差となることを示しなさい。つまり

$$\hat{\beta}_1 = \bar{Y}_1 - \bar{Y}_0$$

ただし $\bar{Y}_1 = \sum_{i=1}^{n_1} Y_i / n_1$、$\bar{Y}_0 = \sum_{i=n_1+1}^{n} Y_i / (n - n_1)$ である。