

# 計量経済 I：期末試験

村澤 康友

提出期限：2020 年 8 月 2 日 (日)

提出方法：My KONAN

**注意：**指定のワードファイルの解答用紙に解答を入力し、pdf ファイルに変換して提出すること。計算には計算機を使用すること。何を参照してもよいが、決して他人と相談しないこと。

- (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).
  - 操作変数
  - 差分の差分 (DID) 法
  - 個別効果
  - 回帰不連続デザイン (RDD)
- (30 点) ストレスが喫煙を増やす効果をデータで検証したい。そこで 2007 年と 2009 年に同じ調査対象者を調査して集めたデータをプールして、「1 日の喫煙本数」を「生活の満足度 (0~4 の 5 段階評価)」と「所得 (万円)」で説明する次の重回帰モデルを推定した。

$$\text{喫煙本数} = \alpha + \beta \cdot \text{生活の満足度} + \gamma \cdot \text{所得} + \text{誤差}$$

分析結果のコンピューター出力は以下の通りであった (一部割愛)。

モデル 1: Pooled OLS, 観測数: 6044

従属変数: ncig

	係数	標準誤差		
const	4.36875	0.201810		
life	-0.669773	0.0681275		
income	0.00450884	0.000271076		
Mean dependent var	3.829004	S.D. dependent var	5.160509	
Sum squared resid	151988.9	S.E. of regression	5.015931	
R-squared	0.055560	Adjusted R-squared	0.055248	
F(2, 6041)	177.6926	P-value(F)	1.03e-75	

- ストレスと喫煙の関係の推定結果を、単位も含めて正確に説明しなさい。
- 問題意識を踏まえて検定問題を定式化しなさい。
- 検定統計量を計算し、検定結果を簡潔に説明しなさい。

3. (50点) 前問の重回帰分析には問題があると考え、同じデータを用いて分析をやり直した。分析の修正点は以下の通り。

- 重回帰モデルを固定効果モデルに変更
- 生活の満足度 (0~4 の 5 段階評価) をダミー変数に変換
- 時点ダミーを追加
- 頑健標準誤差を使用

新たな分析結果のコンピューター出力は以下の通りであった。<sup>\*1</sup>

モデル 1: 固定効果モデル, 観測数: 6044

クロスセクションユニット数: 3022

時系列の長さ= 2

従属変数: ncig

頑健 (HAC) 標準誤差

	係数	標準誤差	t 値	p 値
const	3.82546	0.578651	6.611	4.50e-011 ***
Dlife_1	-0.455154	0.607350	-0.7494	0.4537
Dlife_2	-0.545349	0.566432	-0.9628	0.3357
Dlife_3	-0.970031	0.576918	-1.681	0.0928 *
Dlife_4	-0.978094	0.615740	-1.588	0.1123
income	0.00354489	0.000758326	4.675	3.07e-06 ***
dt_2	-0.346303	0.100095	-3.460	0.0005 ***
Mean dependent var	3.829004	S.D. dependent var	5.160509	
Sum squared resid	42278.82	S.E. of regression	3.744085	
LSDV R-squared	0.737285	Within R-squared	0.014496	

- (a) 重回帰モデルを固定効果モデルに変更すべき理由は何か？前問の分析の問題点を踏まえ、キーワードを適切に用いて簡潔に説明しなさい。
- (b) 生活の満足度をダミー変数に変換すべき理由は何か？前問の分析の問題点を踏まえ、具体的な数値を参照して簡潔に説明しなさい。
- (c) 時点ダミーの係数の推定結果を、単位も含めて正確に説明しなさい。
- (d) ストレスと喫煙の関係の推定結果を、単位も含めて正確に説明しなさい。
- (e) ストレスが喫煙を増やす効果の検定結果を、具体的な数値を参照して簡潔に説明しなさい。

<sup>\*1</sup> ※この分析にも問題がある。このデータには喫煙者と非喫煙者が混在しており、喫煙本数が0の観測値が多い。この問題の解決には非線形モデル (トービット・モデル) や連立方程式モデル (標本選択モデル) が必要となる。

## 解答例

### 1. 計量経済学の基本用語

- (a) 線形モデルの説明変数と相関があり，誤差項と相関がない変数.
- 「 $\text{cov}(Z, X) \neq 0$  かつ  $\text{cov}(Z, U) = 0$  を満たす  $Z$ 」「 $E(ZX) \neq 0$  かつ  $E(ZU) = 0$  を満たす  $Z$ 」でも OK.
- (b) 処置前後の処置群と対照群の変化の差で ATE を推定する手法.
- (c) 個体に固有で観測を通じて一定の効果.
- (d) 処置の割当ルールがもたらす回帰式の不連続性を利用して ATE を推定する手法.

### 2. 重回帰

- (a) 生活の満足度が 1 段階上がると 1 日の喫煙本数は 0.669773 本減る = 生活の満足度が 1 段階下がると 1 日の喫煙本数は 0.669773 本増える.
- 単位の誤りは 5 点減. 左辺の喫煙本数に 0 があり，対数変換できないので，回帰係数  $\neq$  弾力性.
- (b) 片側検定問題は

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

- 仮説を  $\beta$  で表さなければ 0 点.
- (c)  $t$  値は  $-9.831$ .  $H_0$  の下での検定統計量の分布は  $t(6041) \approx N(0, 1)$ . 有意水準を 5% とすると，棄却域は  $(-\infty, -1.645]$ . したがって  $H_0$  を棄却し， $H_1$  (ストレスは喫煙を増やす) を採択する.
- $t$  値で 5 点.

### 3. 固定効果モデル

- (a) 前問の分析には共変量が 1 つしかなく，欠落変数バイアスの恐れがある. 固定効果モデルとして個別ダミーを説明変数に加えれば，すべての時間を通じて変化しない共変量を個別効果として調整できる.
- 問題点で 5 点，解決策で 5 点.
  - 問題点のキーワードは「欠落変数バイアス」，解決策のキーワードは「個別効果」.
- (b) 生活の満足度は量的な意味がない順序変数なので，1 段階上がる限界効果を  $-0.669773$  本で一定とするのは無理がある. ダミー変数に変換すれば限界効果は必ずしも一定でなく，実際に  $-0.455154$  本， $-0.090195$  本， $-0.424682$  本， $-0.008063$  本と変化している.
- 数値なしは 0 点.
- (c) 時点ダミーの係数の推定値は  $-0.346303$ . すなわち 2007 年から 2009 年にかけて 1 日の喫煙本数が平均 0.346303 本減ったことを示している.
- 数値なしは 0 点.
  - 単位の誤りは 5 点減.
- (d) 生活の満足度が 0 の人を基準にすると，満足度が 1~4 の人の 1 日の喫煙本数は，それぞれ 0.455154 本，0.545349 本，0.970031 本，0.978094 本少ない.
- 数値なしは 0 点.
  - 単位の誤りは 5 点減.
  - 基準との比較なしは 5 点減.
- (e) 生活の満足度が 0 の人を基準にすると，満足度が 3 の人の  $t$  値は  $-1.681$ .  $H_0$  の下での検定統計量の分布は  $t(3016) \approx N(0, 1)$  ( $3016 = 6044 - 3021 - 7$ ). 有意水準を 5% とすると，棄却域は

$(-\infty, -1.645]$ . したがって  $H_0$  を棄却し,  $H_1$  (ストレスは喫煙を増やす) を採択する. (※ただし基準の取り方で検定結果が変わる可能性がある. また厳密には多次元母数の片側検定が必要.)

- 数値なしは 0 点.
- 検定なしは 0 点.
- 両側検定は 5 点減.