

## 計量分析 2：宿題 9

村澤 康友

提出期限：2024 年 7 月 9 日

注意：すべての質問に解答しなければ提出とは認めない。授業の HP の解答例を正確に再現すること（乱数は除く）。グループで取り組んでよいが、個別に提出すること。解答例をコピーしたり、他人の名前で提出した場合は、提出点を 0 点とし、再提出も認めない。すべての結果をワードに貼り付けて印刷し（A4 縦・両面印刷可・手書き不可）、2 枚以上の場合は向きを揃えて問題番号順に重ね、左上隅をホッチキスで留めること。

1. (教科書 p. 236, 実証分析問題 9-A) データセット「9\_1\_cig\_xt.dta」を gretl に読み込み、以下の分析を行いなさい。
    - (a) 教科書 p. 225 の「喫煙本数」を「生活の満足度」で説明する単回帰モデルの 2007 年と 2009 年の推定結果を再現しなさい。
    - (b) 教科書 p. 226 の「喫煙本数」の差分を「生活の満足度」の差分で説明する単回帰モデルの推定結果を再現しなさい。※変数の変換方法は資料「gretl 入門」を参照。
    - (c) 前問の分析に所得を共変量として加えた重回帰モデルを推定しなさい。
  2. (教科書 p. 236, 実証分析問題 9-B) データセット「9\_2\_life\_xt.dta」を gretl に読み込み、以下の分析を行いなさい。
    - (a) 教科書 p. 228 の「生活の満足度」を「怪我・病気ダミー」と「年収」で説明する重回帰モデルの 2009 年の推定結果を再現しなさい。
    - (b) 教科書 p. 229 の「生活の満足度」の差分を「怪我・病気ダミー」と「年収」の差分で説明する重回帰モデルの推定結果を再現しなさい。
    - (c) パネル・データによる線形回帰モデルの推定は以下の手順で実行する。
      - i. メニューから「モデル」→「パネル」→「固定効果あるいは変量効果」を選択。
      - ii. 「従属変数」を 1 つ選択。
      - iii. 「説明変数（回帰変数）」を選択。
      - iv. 「固定効果」か「変量効果（ランダム効果）」をチェック。
      - v. 「OK」をクリック。
- 教科書 p. 231 の固定効果モデルと教科書 p. 233 の変量効果モデルの推定結果を再現しなさい。

解答例

1. (a) 2007 年

モデル 1: 最小二乗法 (OLS), 観測: 1–3022

従属変数: ncig

	係数	Std. Error	t-ratio	p 値
const	5.73678	0.281568	20.37	0.0000
life	−0.685414	0.102435	−6.691	0.0000
Mean dependent var	3.966992	S.D. dependent var	5.346335	
Sum squared resid	85088.68	S.E. of regression	5.308018	
$R^2$	0.014609	Adjusted $R^2$	0.014282	
$F(1, 3020)$	44.77234	P-value( $F$ )	2.63e−11	
Log-likelihood	−9331.411	Akaike criterion	18666.82	
Schwarz criterion	18678.85	Hannan–Quinn	18671.15	

2009 年

モデル 2: 最小二乗法 (OLS), 観測: 1–3022

従属変数: ncig

	係数	Std. Error	t-ratio	p 値
const	5.07249	0.272702	18.60	0.0000
life	−0.507946	0.0946631	−5.366	0.0000
Mean dependent var	3.691016	S.D. dependent var	4.964790	
Sum squared resid	73761.82	S.E. of regression	4.942109	
$R^2$	0.009444	Adjusted $R^2$	0.009116	
$F(1, 3020)$	28.79208	P-value( $F$ )	8.67e−08	
Log-likelihood	−9115.560	Akaike criterion	18235.12	
Schwarz criterion	18247.15	Hannan–Quinn	18239.44	

(b) 差分の単回帰

モデル 3: Pooled OLS, 観測数: 3022

クロスセクションユニット数: 3022

時系列の長さ = 1

従属変数: d\_ncig

	係数	Std. Error	t-ratio	p 値
d_life	-0.287266	0.109858	-2.615	0.0090
Mean dependent var	-0.275976	S.D. dependent var		5.322167
Sum squared resid	85607.61	S.E. of regression		5.323299
Uncentered $R^2$	0.002258	Centered $R^2$		-0.000425
$F(1, 3021)$	6.837607	P-value( $F$ )		0.008970
Log-likelihood	-9340.598	Akaike criterion		18683.20
Schwarz criterion	18689.21	Hannan-Quinn		18685.36

(c) 差分の重回帰

モデル 4: Pooled OLS, 観測数: 3022

クロスセクションユニット数: 3022

時系列の長さ = 1

従属変数: d\_ncig

	係数	Std. Error	t-ratio	p 値
d_life	-0.300654	0.109478	-2.746	0.0061
d_income	0.00314364	0.000643058	4.889	0.0000
Mean dependent var	-0.275976	S.D. dependent var		5.322167
Sum squared resid	84935.49	S.E. of regression		5.303238
Uncentered $R^2$	0.010092	Centered $R^2$		0.007429
$F(2, 3020)$	15.39386	P-value( $F$ )		2.23e-07
Log-likelihood	-9328.688	Akaike criterion		18661.38
Schwarz criterion	18673.40	Hannan-Quinn		18665.70

2. (a) 2009 年

モデル 1: 最小二乗法 (OLS), 観測: 1-3020

従属変数: life

	係数	Std. Error	t-ratio	p 値
const	2.67366	0.0305300	87.57	0.0000
shock	-0.124873	0.0346006	-3.609	0.0003
income	0.000282184	7.10062e-005	3.974	0.0001
Mean dependent var	2.703974	S.D. dependent var		0.938548
Sum squared resid	2633.179	S.E. of regression		0.934227
$R^2$	0.009842	Adjusted $R^2$		0.009186
$F(2, 3017)$	14.99441	P-value( $F$ )		3.31e-07
Log-likelihood	-4078.226	Akaike criterion		8162.452
Schwarz criterion	8180.491	Hannan-Quinn		8168.939

(b) 差分

モデル 2: Pooled OLS, 観測数: 3020

クロスセクションユニット数: 3020

時系列の長さ = 1

従属変数: d.life

	係数	Std. Error	t ratio	p 値
const	0.215365	0.0313550	6.869	0.0000
shock	-0.140117	0.0484445	-2.892	0.0039
d_income	0.000223286	0.000161409	1.383	0.1667
Mean dependent var	0.164238	S.D. dependent var		1.310712
Sum squared resid	5168.730	S.E. of regression		1.308893
$R^2$	0.003434	Adjusted $R^2$		0.002773
$F(2, 3017)$	5.197346	P-value( $F$ )		0.005581
Log-likelihood	-5096.623	Akaike criterion		10199.25
Schwarz criterion	10217.29	Hannan-Quinn		10205.73

(c) 固定効果

モデル 1: 固定効果モデル, 観測数: 6040

クロスセクションユニット数: 3020

時系列の長さ = 2

従属変数: life

	係数	標準誤差	t-ratio	p 値
const	2.48153	0.0453239	54.75	0.0000
y2	0.215365	0.0313550	6.869	0.0000
shock_y2	-0.140117	0.0484445	-2.892	0.0039
income	0.000223286	0.000161409	1.383	0.1667
Mean dependent var	2.621854	S.D. dependent var		0.952447
Sum squared resid	2584.365	回帰の標準誤差		0.925527
LSDV $R^2$	0.528256	Within $R^2$		0.018844
$F(3022, 3017)$	1.117939	P-value( $F$ )		0.001099
Log-likelihood	-6006.637	Akaike criterion		18059.27
Schwarz criterion	38331.99	Hannan-Quinn		25097.00
$\hat{\rho}$	-1.000000	Durbin-Watson		1.999449

名前の付けられた説明変数についての結合検定 (Joint test on named regressors) -

検定統計量:  $F(3, 3017) = 19.3148$

なお、p 値 (p-value) =  $P(F(3, 3017) > 19.3148) = 2.10651e-12$

定数項がクロスセクションユニット (グループ) ごとに異なるかどうかの検定 -

帰無仮説: 各クロスセクションユニットは共通の定数項を持つ

検定統計量:  $F(3018, 3017) = 1.08665$

なお、p 値 (p-value) =  $P(F(3018, 3017) > 1.08665) = 0.0112508$



(d) 変量効果

モデル 2: 変量効果モデル (GLS), 観測数: 6040

クロスセクションユニット数: 3020

時系列の長さ = 2

従属変数: life

	係数	標準誤差	$z$	p 値
const	2.45439	0.0263360	93.20	0.0000
shock	0.0153078	0.0349948	0.4374	0.6618
y2	0.212996	0.0310253	6.865	0.0000
shock_y2	-0.139777	0.0484373	-2.886	0.0039
income	0.000303324	5.20242e-005	5.830	0.0000
Mean dependent var	2.621854	S.D. dependent var	0.952447	
Sum squared resid	5393.587	回帰の標準誤差	0.945288	
Log-likelihood	-8228.545	Akaike criterion	16467.09	
Schwarz criterion	16500.62	Hannan-Quinn	16478.73	
$\hat{\rho}$	-1.000000	Durbin-Watson	1.999449	
$\hat{\sigma}_v^2 = 0.0372258$				
$\hat{\sigma}_\varepsilon^2 = 0.856601$				
$\theta = 0.0408154$				

名前の付けられた説明変数についての結合検定 (Joint test on named regressors) -

漸近的検定統計量:  $\chi^2(4) = 95.628$

なお、p 値 (p-value) = 8.37905e-20

ブロイシュ＝ペーガン (Breusch-Pagan) 検定 -

帰無仮説: 個別誤差 (unit-specific error) の分散は 0 である

漸近的検定統計量:  $\chi^2(1) = 5.22889$

なお、p 値 (p-value) = 0.0222147

ハウスマン (Hausman) 検定 -

帰無仮説: GLS 推定値は一致性を持つ

漸近的検定統計量:  $\chi^2(1) = 0.27467$

なお、p 値 (p-value) = 0.600216