

計量経済学練習問題

劉 慶豊*

平成 23 年 6 月 27 日

*E-mail:qliu@res.otaru-uc.ac.jp。

1 Excel による回帰分析

1.1 出力結果の読み方

概要								
回帰統計								
重相関 R	0.998238							
重決定 R2	0.99648							
補正 R2	0.995718							
標準誤差	0.054565							
観測数	46							
分散分析表								
	自由度	変動	分散	割された分散	有意 F			
回帰	8	31.18198	3.897748	1309.156	5.8E-43			
残差	37	0.11016	0.002977					
合計	45	31.29214						
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.376255	0.937939	0.401151	0.690617	-1.52419	2.2767	-1.52419	2.2767
X 値 1	0.610378	0.10487	5.820312	1.1E-06	0.397891	0.822865	0.397891	0.822865
X 値 2	0.33296	0.066106	5.03677	2.26E-05	0.199017	0.466903	0.199017	0.466903
X 値 3	0.12174	0.217402	0.559977	0.57887	-0.31876	0.562239	-0.31876	0.562239
X 値 4	0.066664	0.060199	1.107401	0.275268	-0.05531	0.188639	-0.05531	0.188639
X 値 5	0.014494	0.166246	0.087187	0.930993	-0.32235	0.351341	-0.32235	0.351341
X 値 6	0.166285	0.068327	2.432655	0.019896	0.027841	0.304729	0.027841	0.304729
X 値 7	0.17635	0.103609	1.702077	0.09713	-0.03358	0.386282	-0.03358	0.386282
X 値 8	0.124641	0.0862	1.445959	0.156607	-0.05002	0.299298	-0.05002	0.299298

上から順に
決定係数の平方根
決定係数
修正済み決定係数
残差分散の平方根(s)
標本数 n

帰無仮説が「定数項以外の係数が全て 0」の時の F 検定の P 値

F 統計量

上から順に
ESS
RSS
TSS

帰無仮説は「係数=0」の時の t 値

両側 t 検定の P 値

1.2 練習問題

問題一 表 1. は Excel によるは

$$\begin{aligned}
 \log Y = & \beta_1 + \beta_2 \log L + \beta_3 \log K \\
 & + \beta_4 \log H1 + \beta_5 \log H2 + \beta_6 \log H3 \\
 & + \beta_7 \log H4 + \beta_8 \log H5 + \beta_9 \log H6
 \end{aligned} \tag{1}$$

の推定結果である。

表 1 .

概要								
回帰統計								
重相関 R	0.998238							
重決定 R2	0.99648							
補正 R2	0.995718							
標準誤差	0.054565							
観測数	46							
分散分析表								
	自由度	変動	分散	測された分散	有意 F			
回帰	8	31.18198	3.897748	1309.156	5.8E-43			
残差	37	0.11016	0.002977					
合計	45	31.29214						
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.376255	0.937939	0.401151	0.690617	-1.52419	2.2767	-1.52419	2.2767
X 値 1	0.610378	0.10487	5.820312	1.1E-06	0.397891	0.822865	0.397891	0.822865
X 値 2	0.33296	0.066106	5.03677	1.26E-05	0.199017	0.466903	0.199017	0.466903
X 値 3	0.12174	0.217402	0.559977	0.57887	-0.31876	0.562239	-0.31876	0.562239
X 値 4	0.066664	0.060199	1.107401	0.275268	-0.05531	0.188639	-0.05531	0.188639
X 値 5	0.014494	0.166246	0.087187	0.930993	-0.32235	0.351341	-0.32235	0.351341
X 値 6	0.166285	0.068327	2.433655	0.019896	0.027841	0.304729	0.027841	0.304729
X 値 7	0.17635	0.103609	1.702077	0.09713	-0.03358	0.386282	-0.03358	0.386282
X 値 8	0.124641	0.0862	1.445959	0.156607	-0.05002	0.299298	-0.05002	0.299298

1. 推定された回帰式を書いてください。
2. $\beta_8 = 0$ かどうかの両側検定を行ってください。有意水準は 5% とする。
3. 定数項以外の係数がすべて 0 であるという帰無仮説に関して検定を行ってください。

解答

1. 係数の推定結果を代入して得られた回帰式は

$$\begin{aligned} \log Y = & 0.38 + 0.61 \log L + 0.33 \log K \\ & + 0.12 \log H1 + 0.07 \log H2 + 0.01 \log H3 \\ & + 0.17 \log H4 + 0.18 \log H5 + 0.12 \log H6 \end{aligned} \quad (2)$$

である。

2. 帰無仮説を $H_0 : \beta_8 = 0$, 対立仮説を $H_A : \beta_8 \neq 0$ として t 検定を行う。結論は同じだが検定を行う方法は以下の 2 種類がある。

(a) 表の結果により β_8 の両側検定の P 値が $0.097 = 9.7\%$ であり、有意水準 5% より大きいであるため、帰無仮説を棄却できない。 $\beta_8 = 0$ と考えられる。

(b)

$$t = \frac{\hat{\beta}_8 - 0}{s_{\beta_8}} = \frac{0.18 - 0}{0.10} = 1.8.$$

この t 値が自由度が $n - K = 46 - 9 = 37$ の t 分布の 2.5% (両側検定であるため、元の有意水準 5% を 2 で割る) 有意水準点 1.96 より小さいため、帰無仮説は棄却できない。 $\beta_8 = 0$ と考えられる。

3. F 検定になる。有意水準を 5% とする。帰無仮説は $H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$. 対立仮説は $H_1 : \text{帰無仮説ではない}$ 。表から検定統計量を読み取る。 $f = 1309.16$ 、対応する P 値は $5.8E - 43 = 5.8 \times 10^{-43}$ 殆どゼロとなっており、有意水準 5% より小さいため、帰無仮説は棄却される。

問題二 表 2. は Excel による

$$\log Y = \beta_1 + \beta_2 \log L + \beta_3 \log K + \beta_7 \log H4 + \beta_8 \log H5 \quad (3)$$

の推定結果である。

表 2.

概要								
回帰統計								
重相関 R	0.998006							
重決定 R2	0.996015							
補正 R2	0.995626							
標準誤差	0.055148							
観測数	46							
分散分析表								
	自由度	変動	分散	割された分散	有意 F			
回帰	4	31.16745	7.791862	2562.028	1.38E-48			
残差	41	0.124693	0.003041					
合計	45	31.29214						
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.948916	0.619966	1.530593	0.13355	-0.30313	2.200964	-0.30313	2.200964
X 値 1	0.707843	0.060668	11.66756	1.32E-14	0.585322	0.830363	0.585322	0.830363
X 値 2	0.336529	0.057582	5.844375	7.2E-07	0.22024	0.452817	0.22024	0.452817
X 値 3	0.235699	0.041541	5.673919	1.26E-06	0.151805	0.319592	0.151805	0.319592
X 値 4	0.190143	0.097784	1.94453	0.058714	-0.00733	0.387621	-0.00733	0.387621

1. 推定された回帰式を書いてください。

- 修正済みの決定係数で判断する場合、問題一と問題二の回帰式のどちらを選択するかに関して答えなさい。
- $\beta_8 = 0$ かどうかの両側検定を行ってください。有意水準は 5% とする。
- 問題一にある表 1. の結果とあわせて、

$$\begin{aligned} \log Y = & \beta_1 + \beta_2 \log L + \beta_3 \log K \\ & + \beta_4 \log H1 + \beta_5 \log H2 + \beta_6 \log H3 \\ & + \beta_7 \log H4 + \beta_8 \log H5 + \beta_9 \log H6 \end{aligned} \quad (4)$$

に関して $H_0 : \beta_4 = \beta_5 = \beta_6 = \beta_9 = 0$ に関して検定しなさい。(ヒント : F 検定になる。)

解答

- 推定された回帰式は

$$\log Y = 0.95 + 0.71 \log L + 0.34 \log K + 0.24 \log H4 + 0.19 \log H5$$

となる。

- 表 2 の修正済みの決定係数は 0.9956 で表 1 の修正済みの決定係数 0.9957 より小さいため、問題一の回帰式を選ぶ。
- 帰無仮説を $H_0 : \beta_8 = 0$, 対立仮説を $H_A : \beta_8 \neq 0$ として t 検定を行う。表 2 の結果により β_8 の両側検定の P 値が $0.059 = 5.9\%$ であり、有意水準 5% より大きいであるため、帰無仮説を棄却できない。 $\beta_8 = 0$ と考えられる。
- 有意水準を 5% として、 $H_0 : \beta_4 = \beta_5 = \beta_6 = \beta_9 = 0$ に関して F 検定を行う。帰無仮説の元ではモデルが

$$\log Y = \beta_1 + \beta_2 \log L + \beta_3 \log K + \beta_7 \log H4 + \beta_8 \log H5 \quad (5)$$

となるため、表 2 が帰無仮説に対応している、表 2 から $RSS(H_0) = 0.1247$ を読み取る。そして、表 1 は対立仮説に対応しているので、表 2 から $RSS(H_A) = 0.1102$ を読み取る。 f 値の公式に代入して

$$\begin{aligned} f &= \frac{n - K}{m} \frac{RSS(H_0) - RSS(H_A)}{RSS(H_A)} \\ &= \frac{46 - 9}{4} \frac{0.1247 - 0.1102}{0.1102} \\ &= 1.2171 \end{aligned}$$

となる。自由度が $(m, n - K) = (4, 37)$ の F 分布の 5 % 有意水準点が F 分布表から読み取り、2.63 である (自由度が $(4, 37)$ の f 値がテキストの付表にないため、自由度が $(4, 37)$ と $(4, 50)$ の f 値を読み取って、その平均値を利用した)。 $f = 1.2171 < 2.63$ であるので、帰無仮説は棄却できない。

問題三 講義用ホームページにある「性別ダミー」のデータを利用して身長 (X) と体重 (Y) の関係を表す線形モデル ($Y_i = \alpha + \beta X_i + u_i$) を推定し推定式を書きなさい。データの前半 25 人を女性、後半 25 人を男性として、性別ダミーのデータを作成して、ダミー変数のデータを取り入れたモデルを再推定し推定式を書きなさい。ダミー変数の係数に関して t 検定を行い、さらに \bar{R}^2 を利用してダミーを入れたと入れていないモデルのどちらを選択するかを判断しなさい。さらに推定式を男性式と女性式に分けて書きなさい。その上、ダミー変数の係数が表す男女差について説明しなさい。

解答

表 3 .

回帰統計	
重相関 R	0.715644
重決定 R2	0.512146
補正 R2	0.501983
標準誤差	8.77458
観測数	50

分散分析表

	自由度	変動	分散	調整された分散	有意 F
回帰	1	3879.704	3879.704	50.39019	5.19E-09
残差	48	3695.676	76.99325		
合計	49	7575.38			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	-56.2212	16.30417	-3.44827	0.001184	-89.0029	-23.4394	-89.0029	-23.4394
身長(cm)	0.698723	0.098431	7.098605	5.19E-09	0.500815	0.896632	0.500815	0.896632

推定結果：

$$\hat{Y}_i = -56.22 + 0.6987X_i$$

となった。

表 4 .

回帰統計	
重相関 R	0.759558
重決定 R ²	0.576928
補正 R ²	0.558925
標準誤差	8.257725
観測数	50

分散分析表

	自由度	変動	分散	割られた分	有意 F
回帰	2	4370.449	2185.225	32.04611	1.66E-09
残差	47	3204.931	68.19002		
合計	49	7575.38			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	-58.5731	15.36882	-3.81117	0.000401	-89.4912	-27.655	-89.4912	-27.655
男女ダミー	6.266884	2.336061	2.682672	0.010049	1.567336	10.96643	1.567336	10.96643
身長(cm)	0.693992	0.09265	7.490472	1.48E-09	0.507604	0.880379	0.507604	0.880379

ダミー変数 M_i を女性なら 1、男性なら 0 として定義する。ダミー変数を取り入れた推定結果：

$$\hat{Y}_i = -58.57 + 6.27M_i + 0.694X_i$$

となった。

M_i の t 値が 2.68 となり、対応している P 値は $0.01 < 5\%$ であるので、両側検定 5% で M_i の係数が 0 である帰無仮説が棄却される。すなわち、男女の性別による身長と体重の間の関係に差が存在する。

ダミー変数を取り入れた式の $\bar{R}^2 = 0.558925$ となり、ダミー変数を取り入れていない式の $\bar{R}^2 = 0.501983$ より大きなので、ダミー変数を取り入れたモデルを選択する。

推定結果を男性式と女性式に分けて書く。

$$\text{男性式} : \hat{Y}_i = -58.57 + 0.694X_i$$

$$\begin{aligned} \text{女性式} : \hat{Y}_i &= (-58.57 + 6.27) + 0.694X_i \\ &= -52.3 + 0.694X_i \end{aligned}$$

二つの式の定数項が異なる。女性式の定数項が男性式より 6.27 大きくなっている。すなわち、同じ身長であっても女性の体重が男性より重くなりがちである。