

今回も、前回のガソリン価格の需要関数の推定のためのデータをもとに説明を行う。

1. 誤差項の諸問題

1.1 系列相関（時系列データの分析にのみ適用される）

以下では、誤差項が過去の誤差項と相関（系列相関）を持っているかどうかを調べる。これは、系列相関の存在によって、推定法を改善する余地が発生するからである。これは、あくまでも改善であって、たとえ系列相関の存在を無視しても、推定量は一致性をもつ（つまり、観測数を増やしたら本当の値に収束する）。ただ、この場合、推定値の標準誤差が OLS のコマンドから得られるものとは異なるので、 t 値、 F 値などによる検定は誤ってしまう。

以上は理論的な話であり、実務家が系列相関の存在に関心を持つのは、それによって、回帰式の説明変数に入れなかった、omitted variable の存在があきらかになると信じているからである。理論的には、これは誤りで、系列相関が検出される場合は、1) モデルの関数形がただしいが系列相関がある場合、2) 関数形が間違っていて、omitted variable が存在し、それが誤差項に入ってしまったからかのいずれかである。1) の場合は、コクラン - オーカット法によってより推定誤差の小さい推定量を得ることができる、2) の場合は、新たな変数を加えて回帰を行うことになる。変数選択については、AIC, BIC, 自由度修正済み決定係数による選択を行うことになる。

a) ダービン・ワトソン検定

gretl を起動する。

gretl メイン・ウィンドウのメニューバーの「ファイル」「データを開く」「sample file」を選択する。出てきたウィンドウの「Greene」のタブをクリックし、greene8_3 をダブルクリックする。

メニューバーの「Add」「Define new variable」とたどり、左クリックして選択する。出てきたウィンドウの入力ボックスに、「 $\ln G_{percapita} = \log(G/Pop)$ 」と入力する。

Pg, Y, Pnc, Puc, Ppt, Pd, Pn, Ps を選択し、メニューバーの「Add」「logs of selected variables」とたどり、左クリックする。

gretl メイン・ウィンドウのメニューバーの「モデル」「Ordinary least squares」をたどり、出てきたウィンドウの「dependent variable」に $\ln G_{percapita}$ 、「independent variables」に $\ln Pg$, $\ln Y$, $\ln Puc$, $\ln Pd$, $\ln Pn$, $\ln Ps$ を指定する（BIC 最小のモデル）。そのあと、「OK」ボタンを左クリックする。

出力結果の以下の行に注目する。

Durbin-Watson statistic = 1.18968

First-order autocorrelation coeff. = 0.362461

教科書の巻末付録 6 (p.297) を見る。観測個数は 36, $k = 6$ なので, $n = 35$. $k = 6$ の交差しているところを見る。L=1.10, U=1.88 となっている。これと、Durbin-Watson statistic (DW 統計量) = 1.18968 を比較する。この場合、 $L < DW$ 統計量 $< U$ なので、系列相関がないともあるとも判定できない境界である。詳しくは、教科書の p.162~p.164 参照。

なお，教科書で理論では， U を境界値と考え， DW 統計量 $< U$ なら系列相関が存在する
 と考え，実務家は DW 統計量 $< L$ で系列相関の存在と判定するというのは，実務家がこ
 れを omitted variable の存在の検出に使っているという事情も関連していると思われる．

b) その他の検定

モデルの推定結果のウィンドウのメニューバーの「Tests」 「Autocorrelation」を選択
 する．

出てきたウィンドウの「Lag order for test」(lag 次数)に「2」をいれる．この場合，入
 力ボックス右端の をクリックすることで値を変更する．その後，「OK」ボタンを押
 す．

結果を見る．

Breusch-Godfrey test for autocorrelation up to order 2

OLS estimates using the 34 observations 1962-1995

Dependent variable: uhat

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|----------|-------------|-----------|--------|-----------|
| const | -1.06161 | 0.905330 | -1.173 | 0.25200 |
| L_Pg | 0.0416246 | 0.0575445 | 0.723 | 0.47618 |
| L_Y | 0.105269 | 0.0911091 | 1.155 | 0.25884 |
| L_Puc | 0.0338085 | 0.0602526 | 0.561 | 0.57972 |
| L_Pd | -0.0916754 | 0.210202 | -0.436 | 0.66649 |
| L_Pn | -0.0983807 | 0.205150 | -0.480 | 0.63571 |
| L_Ps | 0.0364323 | 0.136421 | 0.267 | 0.79161 |
| uhat_1 | 0.413003 | 0.202464 | 2.040 | 0.05206 * |
| uhat_2 | -0.0671333 | 0.196223 | -0.342 | 0.73511 |

Unadjusted R-squared = 0.200679

Test statistic: LMF = 3.138268,

with p-value = $P(F(2,25) > 3.13827) = 0.0608$

Alternative statistic: $TR^2 = 6.823077$,

with p-value = $P(\text{Chi-square}(2) > 6.82308) = 0.033$

Ljung-Box $Q' = 4.9457$ with p-value = $P(\text{Chi-square}(2) > 4.9457) = 0.0843$

まず，Breusch-Godfrey (ブロイシュ - ゴッドfrey) 検定の結果を見る．F 検定の結果
 は，p 値が 0.0608 で，0.05 以上なので，系列相関があるとまではいえない結果になって

いる。次に、LM 検定の p 値は、0.033 で 0.05 より小さいので、系列相関の存在を示している。また、もう一つの検定は、Ljung-Box の Q 検定で、この p 値は、0.0843 で、0.05 以上なので、系列相関があるとまではいえない。このように、conflicting な結果がでたとはいえるが、系列相関がある程度存在するとはいえるだろう。存在を否定した検定の p 値も十分小さいからである。

また、何次の自己相関があるかであるが、

| | | | | |
|--------|------------|----------|--------|-----------|
| uhat_1 | 0.413003 | 0.202464 | 2.040 | 0.05206 * |
| uhat_2 | -0.0671333 | 0.196223 | -0.342 | 0.73511 |

の部分を見ればよい。5%有意に近いのは、uhat_1、すなわち、 \hat{u}_{t-1} である。従って、一次の自己相関のみを考慮すればいいことになる。

c) コ克蘭・オーカット法 (実際は反復コ克蘭・オーカット法)

b)で検出された誤差項の系列相関に対処して、より分散の小さい、また、標準誤差も正しく計算できる推定量を得よう。

gretl のメイン・ウィンドウにもどり、メニューバーの「モデル」 「Time Series」 「Cochrane - Orcutt」を選択する。

先ほどの説明変数が dependent variables の入力ボックスに入っていることを確認して、「OK」ボタンを押す。

出力の評価

Performing iterative calculation of rho...

| ITER | RHO | ESS |
|-------|---------|------------|
| 1 | 0.36246 | 0.00682884 |
| 2 | 0.45611 | 0.00669689 |
| 3 | 0.48507 | 0.00668342 |
| 4 | 0.49462 | 0.00668192 |
| 5 | 0.49785 | 0.00668175 |
| 6 | 0.49895 | 0.00668173 |
| final | 0.49933 | |

Model 8: Cochrane-Orcutt estimates using the 35 observations 1961-1995

Dependent variable: l_G_percapita

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|----------|-------------|-----------|---------|--------------|
| const | -11.3476 | 1.11162 | -10.208 | <0.00001 *** |
| l_Pg | -0.403825 | 0.0620294 | -6.510 | <0.00001 *** |

| | | | | |
|-------|------------|-----------|--------|--------------|
| l_Y | 1.30418 | 0.115144 | 11.326 | <0.00001 *** |
| l_Puc | -0.0901031 | 0.0647719 | -1.391 | 0.17516 |
| l_Pd | 0.956633 | 0.216200 | 4.425 | 0.00013 *** |
| l_Pn | 0.697448 | 0.249851 | 2.791 | 0.00935 *** |
| l_Ps | -0.928190 | 0.166752 | -5.566 | <0.00001 *** |

Statistics based on the rho-differenced data:

Sum of squared residuals = 0.00668173
Standard error of residuals = 0.0154478
Unadjusted R-squared = 0.990417
Adjusted R-squared = 0.988364
F-statistic (6, 28) = 116.763 (p-value < 0.00001)
Durbin-Watson statistic = 1.91497
First-order autocorrelation coeff. = 0.0410312
Akaike information criterion (AIC) = -186.405
Schwarz Bayesian criterion (BIC) = -175.517
Hannan-Quinn criterion (HQC) = -182.646

Excluding the constant, p-value was highest for variable 15 (l_Puc)

まず、OLSとの係数推定値の違いを見てみよう。

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|----------|-------------|-----------|---------|--------------|
| const | -9.48676 | 0.766315 | -12.380 | <0.00001 *** |
| l_Pg | -0.501872 | 0.0580896 | -8.640 | <0.00001 *** |
| l_Y | 1.12201 | 0.0760662 | 14.750 | <0.00001 *** |
| l_Puc | -0.138068 | 0.0636618 | -2.169 | 0.03844 ** |
| l_Pd | 1.11874 | 0.224532 | 4.983 | 0.00003 *** |
| l_Pn | 0.978631 | 0.207686 | 4.712 | 0.00006 *** |
| l_Ps | -1.07458 | 0.139958 | -7.678 | <0.00001 *** |

係数値の変化が微妙にあることが分かるだろう。また、OLSでは、l_Pucが有意だったが、コクラン・オーカット法では、有意ではないことがわかる。これが、系列相関の存在による検定のゆがみである。

また、Durbin-Watson statistic = 1.91497 をみると、残差の自己相関がほとんど無くなっていることが分かる。

この場合，1 次の自己相関のみを扱っているが，2 次以上の自己相関の処理も知っておこう．

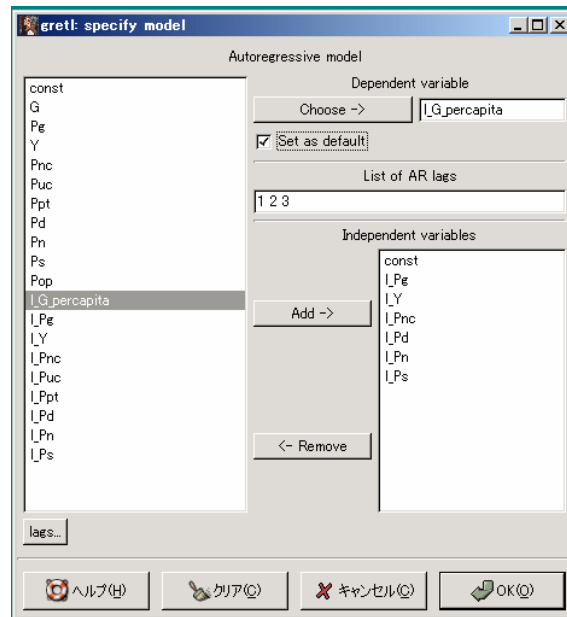


図 1

gretl のメイン・ウィンドウにもどり，メニューバーの「モデル」 「Time Series」 「Autoregressive Estimation」を選択する．

図 1 のようなウインドウがでるが，その「List of AR lags」に，考慮する誤差項の系列相関の次数が 3 なら，「1 2 3」と指定する．

結果の評価

Generalized Cochrane-Orcutt estimation

| ITER | ESS | % CHANGE |
|------|----------|-----------|
| 1 | 0.005931 | undefined |
| 2 | 0.005845 | 1.453 |
| 3 | 0.005816 | 0.494 |
| 4 | 0.005803 | 0.220 |
| 5 | 0.005797 | 0.100 |
| 6 | 0.005795 | 0.045 |
| 7 | 0.005793 | 0.019 |
| 8 | 0.005793 | 0.008 |
| 9 | 0.005793 | 0.004 |

Model 20: AR estimates using the 33 observations 1963-1995

Dependent variable: l_G_percapita

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|----------|-------------|-----------|---------|--------------|
| const | -11.7259 | 1.09451 | -10.713 | <0.00001 *** |
| l_Pg | -0.367569 | 0.0622337 | -5.906 | <0.00001 *** |
| l_Y | 1.34189 | 0.112504 | 11.927 | <0.00001 *** |
| l_Puc | -0.105601 | 0.0645595 | -1.636 | 0.11395 |
| l_Pd | 0.767696 | 0.217677 | 3.527 | 0.00158 *** |
| l_Pn | 0.769734 | 0.242981 | 3.168 | 0.00390 *** |
| l_Ps | -0.904661 | 0.153756 | -5.884 | <0.00001 *** |

Estimates of the AR coefficients:

| | | | | |
|-----|------------|----------|--------|---------|
| u_1 | 0.609794 | 0.168648 | 3.616 | 0.00108 |
| u_2 | -0.313628 | 0.195912 | -1.601 | 0.11989 |
| u_3 | -0.0663568 | 0.149358 | -0.444 | 0.66003 |

Sum of AR coefficients = 0.229810

Statistics based on the rho-differenced data:

Sum of squared residuals = 0.00579275
Standard error of residuals = 0.0149264
Unadjusted R-squared = 0.987606
Adjusted R-squared = 0.984746
F-statistic (6, 26) = 207.85 (p-value < 0.00001)
Durbin-Watson statistic = 1.97065
First-order autocorrelation coeff. = -0.114641
Akaike information criterion (AIC) = -177.723
Schwarz Bayesian criterion (BIC) = -167.247
Hannan-Quinn criterion (HQC) = -174.198

Excluding the constant, p-value was highest for variable 15 (l_Puc)

1.2 分散不均一性

分散不均一性とは、説明変数の値によって誤差項の分散が変化する場合である。クロスセクションデータにおいては、一般的に発生することである。

a) 分散不均一性の検定 (White 検定)

OLSの結果を示したウィンドウにもどり、メニューバーの「Tests」「Heteroskedasticity」を選択する。
出力の診断

White's test for heteroskedasticity

OLS estimates using the 36 observations 1960-1995

Dependent variable: uhat^2

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|------------|-------------|-----------|--------|---------|
| const | 0.226931 | 9.66366 | 0.023 | 0.98184 |
| l_Pg | 0.0322570 | 1.65074 | 0.020 | 0.98489 |
| l_Y | 0.141301 | 1.94281 | 0.073 | 0.94381 |
| l_Puc | -1.60433 | 1.33690 | -1.200 | 0.26445 |
| l_Pd | 3.54493 | 6.19921 | 0.572 | 0.58314 |
| l_Pn | -1.35508 | 5.61779 | -0.241 | 0.81546 |
| l_Ps | 0.460758 | 4.07943 | 0.113 | 0.91286 |
| sq_l_Pg | -0.0128664 | 0.0553414 | -0.232 | 0.82199 |
| l_Pg_l_Y | -0.00433983 | 0.164593 | -0.026 | 0.97961 |
| l_Pg_l_Puc | 0.0377390 | 0.0752058 | 0.502 | 0.62932 |
| l_Pg_l_Pd | 0.0574694 | 0.256485 | 0.224 | 0.82832 |
| l_Pg_l_Pn | 0.0648200 | 0.354209 | 0.183 | 0.85935 |
| l_Pg_l_Ps | -0.107187 | 0.242044 | -0.443 | 0.66961 |
| sq_l_Y | -0.0167879 | 0.0991531 | -0.169 | 0.86975 |
| l_Y_l_Puc | 0.155231 | 0.128408 | 1.209 | 0.26122 |
| l_Y_l_Pd | -0.347940 | 0.618864 | -0.562 | 0.58937 |
| l_Y_l_Pn | 0.126036 | 0.559063 | 0.225 | 0.82729 |
| l_Y_l_Ps | -0.0285685 | 0.419666 | -0.068 | 0.94740 |
| sq_l_Puc | 0.0541886 | 0.0688960 | 0.787 | 0.45422 |
| l_Puc_l_Pd | -0.370221 | 0.633811 | -0.584 | 0.57524 |
| l_Puc_l_Pn | 0.0907689 | 0.338818 | 0.268 | 0.79556 |
| l_Puc_l_Ps | -0.0453176 | 0.143119 | -0.317 | 0.75962 |
| sq_l_Pd | 0.467874 | 0.920659 | 0.508 | 0.62503 |
| l_Pd_l_Pn | -0.752348 | 1.32192 | -0.569 | 0.58489 |
| l_Pd_l_Ps | 0.475583 | 0.800691 | 0.594 | 0.56895 |
| sq_l_Pn | 0.103048 | 0.490634 | 0.210 | 0.83889 |
| l_Pn_l_Ps | 0.118503 | 0.655096 | 0.181 | 0.86095 |
| sq_l_Ps | -0.122945 | 0.325353 | -0.378 | 0.71535 |

Unadjusted R-squared = 0.588896

Test statistic: $TR^2 = 21.200246$,

with p-value = $P(\text{Chi-square}(27) > 21.200246) = 0.776821$

with p-value = $P(\text{Chi-square}(27) > 21.200246) = 0.776821$ をみて p 値で判定する。0.05 より小さければ、5%有意で分散不均一性があると判定する。

b) Newey-West 推定量

では、分散不均一性が会った場合の対処である。まず、分散不均一性を考慮して、正しい標準誤差を計算し、推定量の分散も OLS より小さい推定量である Newey-West の推定量を求めてみよう。もちろん、分散不均一性を考慮しなくても推定量は一致性をもつが、検定のゆがみを避けるためには、b) や c) の方法が必要である。

gretl のメイン・ウィンドウにもどり、メニューバーの「モデル」 「Other linear Models」 「Heteroskedasticity Corrected」を選択する。

でてきたウィンドウで先ほどの 6 つの説明変数と被説明変数が指定されていることを確認し、「OK」ボタンを押す。

出力結果の診断

Model XX: Heteroskedasticity-corrected estimates using the 36 observations 1960-1995

Dependent variable: l_G_percapita

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|----------|-------------|-----------|---------|--------------|
| const | -10.2459 | 0.697717 | -14.685 | <0.00001 *** |
| l_Pg | -0.431229 | 0.0422166 | -10.215 | <0.00001 *** |
| l_Y | 1.18693 | 0.0684081 | 17.351 | <0.00001 *** |
| l_Puc | -0.0763973 | 0.0612522 | -1.247 | 0.22228 |
| l_Pd | 0.960768 | 0.178479 | 5.383 | <0.00001 *** |
| l_Pn | 0.783334 | 0.133430 | 5.871 | <0.00001 *** |
| l_Ps | -0.969149 | 0.0913706 | -10.607 | <0.00001 *** |

Statistics based on the weighted data:

Sum of squared residuals = 82.9662

Standard error of residuals = 1.69142

Unadjusted R-squared = 0.991056

Adjusted R-squared = 0.989206

F-statistic (6, 29) = 535.589 (p-value < 0.00001)

Durbin-Watson statistic = 0.79457
 First-order autocorrelation coeff. = 0.586884
 Akaike information criterion (AIC) = 146.221
 Schwarz Bayesian criterion (BIC) = 157.305
 Hannan-Quinn criterion (HQC) = 150.089

Statistics based on the original data:

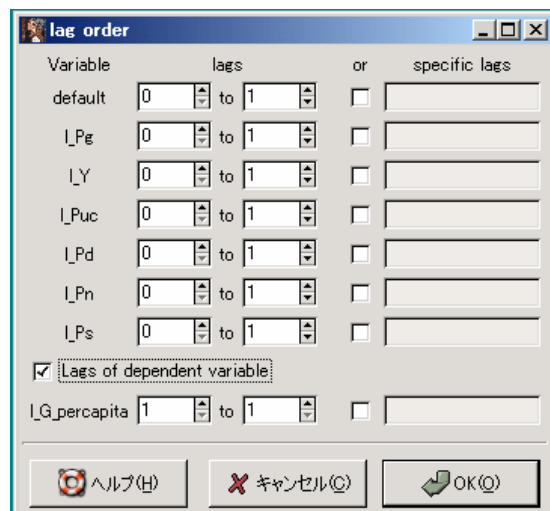
Mean of dependent variable = -0.00370861
 Standard deviation of dep. var. = 0.151691
 Sum of squared residuals = 0.0120008
 Standard error of residuals = 0.0203426

Excluding the constant, p-value was highest for variable 15 (l_Puc)

以上の結果では，残差の系列相関が考慮されていない．それを考慮するためにはどうするか？
 gretl のメイン・ウィンドウにもどり，メニューバーの「モデル」 「Other linear Models」
 「Heteroskedasticity Corrected」を選択する．そして，でてきたウィンドウに先ほどの 6 つの説明変
 数と被説明変数が指定されていることを確認し，「lags」のボタンを左クリックする．

図 2 の様に lags の to の後の入力ボックスの右端の を左クリックして，ラグを 0 から 1 までと
 指定する．さらに，Lags of dependent variables の左のチェックボックスを左クリックして，チェッ
 クを入れ，「OK」ボタンを押す．これは，自己相関の次数を 1 次と考えたために，説明，被説明変数の両
 方に関して，1 次のラグをとったのである．もし，自己相関の次数が 3 次ならば，すべての説明変数及
 び被説明変数に関して，右側のボックスの値が 3 になるように指定する．被説明変数のラグも右側のボ
 ックスの右端にある を左クリックして指定する．

説明変数，被説明変数を指定するウィンドウに戻り，「OK」ボタンを押す．



出力結果の判定

Model 24: Heteroskedasticity-corrected estimates using the 35 observations 1961-1995

Dependent variable: l_G_percapita

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|---------------|-------------|-----------|---------|--------------|
| const | -5.06599 | 0.655995 | -7.723 | <0.00001 *** |
| l_Pg | -0.430864 | 0.0366466 | -11.757 | <0.00001 *** |
| l_Pg_1 | 0.0994227 | 0.0383827 | 2.590 | 0.01707 ** |
| l_Y | 0.581837 | 0.112787 | 5.159 | 0.00004 *** |
| l_Y_1 | 0.00634677 | 0.154445 | 0.041 | 0.96761 |
| l_Puc | 0.135267 | 0.0479985 | 2.818 | 0.01030 ** |
| l_Puc_1 | -0.107568 | 0.0378107 | -2.845 | 0.00970 *** |
| l_Pd | 0.302396 | 0.154522 | 1.957 | 0.06378 * |
| l_Pd_1 | 0.367778 | 0.195536 | 1.881 | 0.07393 * |
| l_Pn | 0.339871 | 0.133363 | 2.548 | 0.01871 ** |
| l_Pn_1 | -0.134343 | 0.133483 | -1.006 | 0.32566 |
| l_Ps | 0.545914 | 0.360381 | 1.515 | 0.14472 |
| l_Ps_1 | -0.958193 | 0.321272 | -2.982 | 0.00710 *** |
| l_G_percapi_1 | 0.503089 | 0.0601964 | 8.357 | <0.00001 *** |

Statistics based on the weighted data:

Sum of squared residuals = 61.8547
 Standard error of residuals = 1.71624
 Unadjusted R-squared = 0.999473
 Adjusted R-squared = 0.999146
 F-statistic (13, 21) = 3060.73 (p-value < 0.00001)
 Durbin-Watson statistic = 2.21626
 First-order autocorrelation coeff. = -0.108573
 Akaike information criterion (AIC) = 147.256
 Schwarz Bayesian criterion (BIC) = 169.031
 Hannan-Quinn criterion (HQC) = 154.773

Statistics based on the original data:

Mean of dependent variable = 0.00566012
 Standard deviation of dep. var. = 0.142948
 Sum of squared residuals = 0.00473906
 Standard error of residuals = 0.0150223

Excluding the constant, p-value was highest for variable 21 (l_Y_1)

まず、定数項の見方だが、これは、 $const/(1-l_G_percapi_1)$ で計算しなおす必要がある（詳しくは教科書 p.169 を見よ）。一般に系列相関として考慮した次数が長くなれば、 $const/(1-l_G_percapi_1-l_G_percapi_2-\dots)$ として計算される。これは、この結果のウィンドウのメニューバーの「Save」「Define new variable」とたどり、その結果でてきた、ウィンドウに「 $c=\$coeff[1]/(1-\$coeff[\$ncoeff])$ 」と入力することで計算することができる。もし次数を2ととったなら、「 $c=\$coeff[1]/(1-\$coeff[\$ncoeff-1]-\$coeff[\$ncoeff])$ 」となるだろう。有意性を判断するのはより高度な統計学の知識を必要とする。

他の係数は、ラグをとらない変数の係数を見る。有意性も、その係数の p 値を見ればよい。それは、出力結果で網掛けが入っている部分である。自己相関と分散不均一性の両方を考慮した場合、l_Pd が有意でなくなっている。

最後に誤差項の自己相関の有意性の判定であるが、

l_G_percapi_1 0.503089 0.0601964 8.357 <0.00001 ***

をみる。つまり、被説明変数のラグをとった変数の係数の有意性で見ればよい。この場合、かなり有意である。

さて、誤差項の系列相関と分散不均一性をとったモデルと、そうでないモデルの残差を比較してみよう。図3がこのモデルの残差の時系列プロット、図4がこのモデルの残差の l_Y を横軸にとった散布図である。図5は OLS モデルの残差の時系列プロット、図6は OLS の残差の l_Y を横軸にとった散布図である。系列相関もまた、l_Y との分散の相関もいくつかの特異値を除けばかなり改善されているのがみてとれる。これらの図を出力する手順は今まで学んだ手順でできるので、自分でもやってみてください。

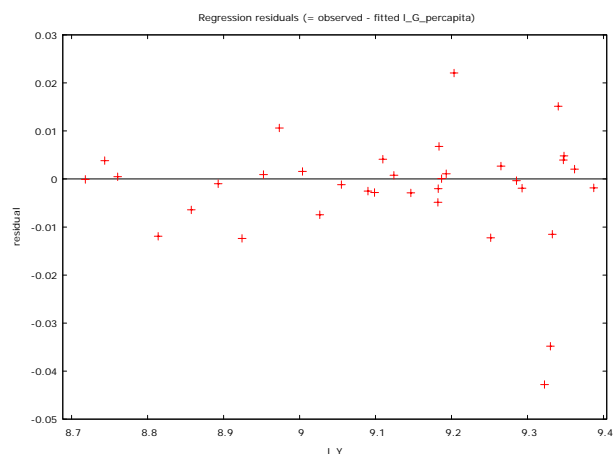
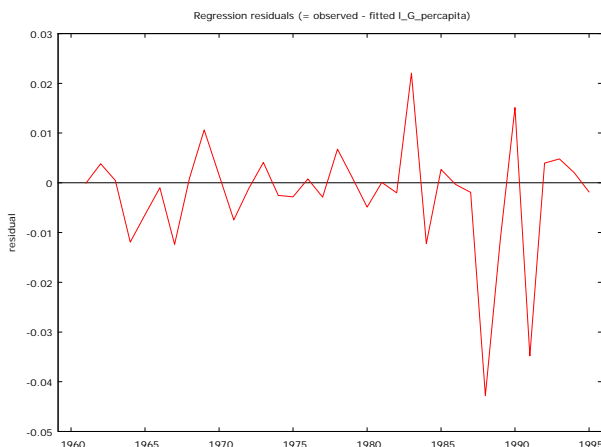


図 3

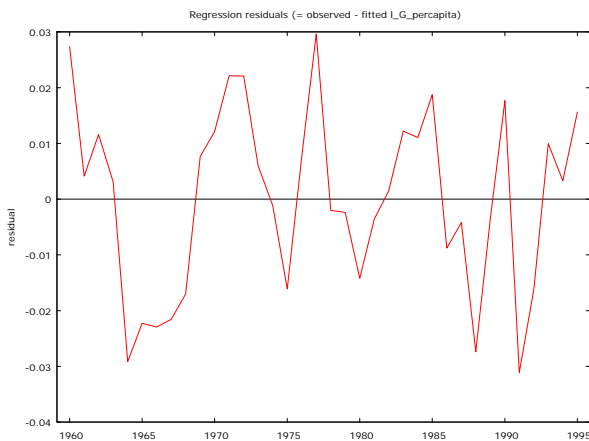


図 5

図 4

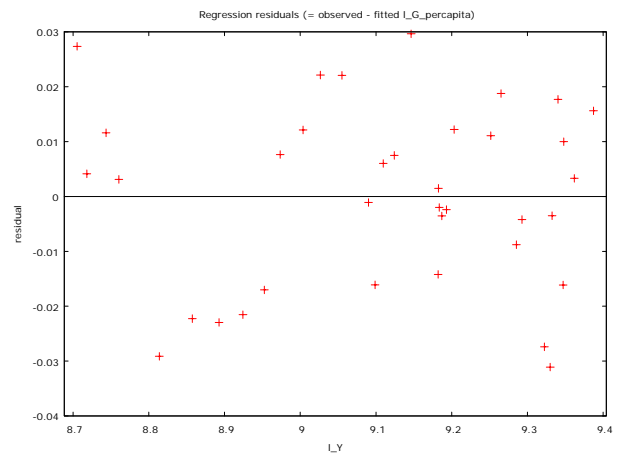


図 6

c) 分散不均一性，系列相関に対して頑健な(Robust)標準誤差の推定

上記の b) の方法は，必要な仮定さえ満たせば有効な方法である．が，その方法を成立させている仮定は，しばしば成立していないことが多い．一方，OLS 推定量というのは，弱い仮定が成立すれば，(分散不均一性や系列相関が存在しても) 一致性が成立するので，その推定誤差評価が誤っていることを補正してやれば，最も実用性が高い方法だといえる．誤差評価の補正を学ぼう．なお，時系列データの場合，分散不均一性だけではなく系列相関も考慮した分散評価の補正ができる．

gretl メイン・ウィンドウにもどり，メニューバーの「モデル」 「Other linear Models」 「Ordinary Least Squares」 を選択する．

出てきたウィンドウの左下の「Robust Standard errors」のチェックボックスを左クリックし，説明変数と被説明変数をこのプリントの最初の OLS のモデルで指定した変数に戻し，「OK」 ボタンを左クリックする．

結果の診断

Model 25: OLS estimates using the 36 observations 1960-1995

Dependent variable: l_G_percapita

HAC standard errors, bandwidth 2 (Bartlett kernel)

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE |
|----------|-------------|-----------|--------|--------------|
| const | -9.48676 | 0.955218 | -9.932 | <0.00001 *** |
| l_Pg | -0.501872 | 0.0613763 | -8.177 | <0.00001 *** |
| l_Y | 1.12201 | 0.0945532 | 11.866 | <0.00001 *** |
| l_Puc | -0.138068 | 0.0755570 | -1.827 | 0.07796 * |
| l_Pd | 1.11874 | 0.224628 | 4.980 | 0.00003 *** |
| l_Pn | 0.978631 | 0.203489 | 4.809 | 0.00004 *** |
| l_Ps | -1.07458 | 0.133593 | -8.044 | <0.00001 *** |

Mean of dependent variable = -0.00370861
 Standard deviation of dep. var. = 0.151691
 Sum of squared residuals = 0.00955231
 Standard error of residuals = 0.0181491
 Unadjusted R-squared = 0.988139
 Adjusted R-squared = 0.985685
 F-statistic (6, 29) = 322.968 (p-value < 0.00001)
 Durbin-Watson statistic = 1.18968
 First-order autocorrelation coeff. = 0.362461
 Log-likelihood = 97.1391
 Akaike information criterion (AIC) = -180.278
 Schwarz Bayesian criterion (BIC) = -169.193
 Hannan-Quinn criterion (HQC) = -176.409

この分散不均一性を考慮した標準誤差の計算により, l_Puc の有意性が10%水準まで下がったことが分かる.

2. 関数形の検定

2.1 Reset 検定

関数形がこれで正しいのかを考える必要がある.

1. 2のc)のモデルの出力ウィンドウ(OLSの結果を表示したウィンドウ)にもどる.

メニューバーの「Tests」 「Ramsey's RESET」を選択する.

出力の判定

Auxiliary regression for RESET specification test
 OLS estimates using the 36 observations 1960-1995
 Dependent variable: l_G_per capita

| VARIABLE | COEFFICIENT | STDERROR | T STAT | P-VALUE | |
|----------|-------------|-----------|---------|----------|-----|
| const | -13.1966 | 1.08403 | -12.174 | <0.00001 | *** |
| l_Pg | -0.552352 | 0.0498721 | -11.075 | <0.00001 | *** |
| l_Y | 1.52802 | 0.112460 | 13.587 | <0.00001 | *** |
| l_Puc | -0.0967313 | 0.0400401 | -2.416 | 0.02274 | ** |
| l_Pd | 1.36443 | 0.146229 | 9.331 | <0.00001 | *** |
| l_Pn | 0.896572 | 0.157505 | 5.692 | <0.00001 | *** |
| l_Ps | -1.24439 | 0.0987021 | -12.608 | <0.00001 | *** |
| yhat^2 | -1.09401 | 0.272562 | -4.014 | 0.00043 | *** |
| yhat^3 | -5.55461 | 0.831354 | -6.681 | <0.00001 | *** |

Test statistic: $F = 24.102276$,
with p-value = $P(F(2,27) > 24.1023) = 9.87e-007$

p 値が 0.05 を下回ると、現在の関数形は正しくないことがいえる。この例では、それぞれの変数の 2 乗などをいれることが必要かもしれない。

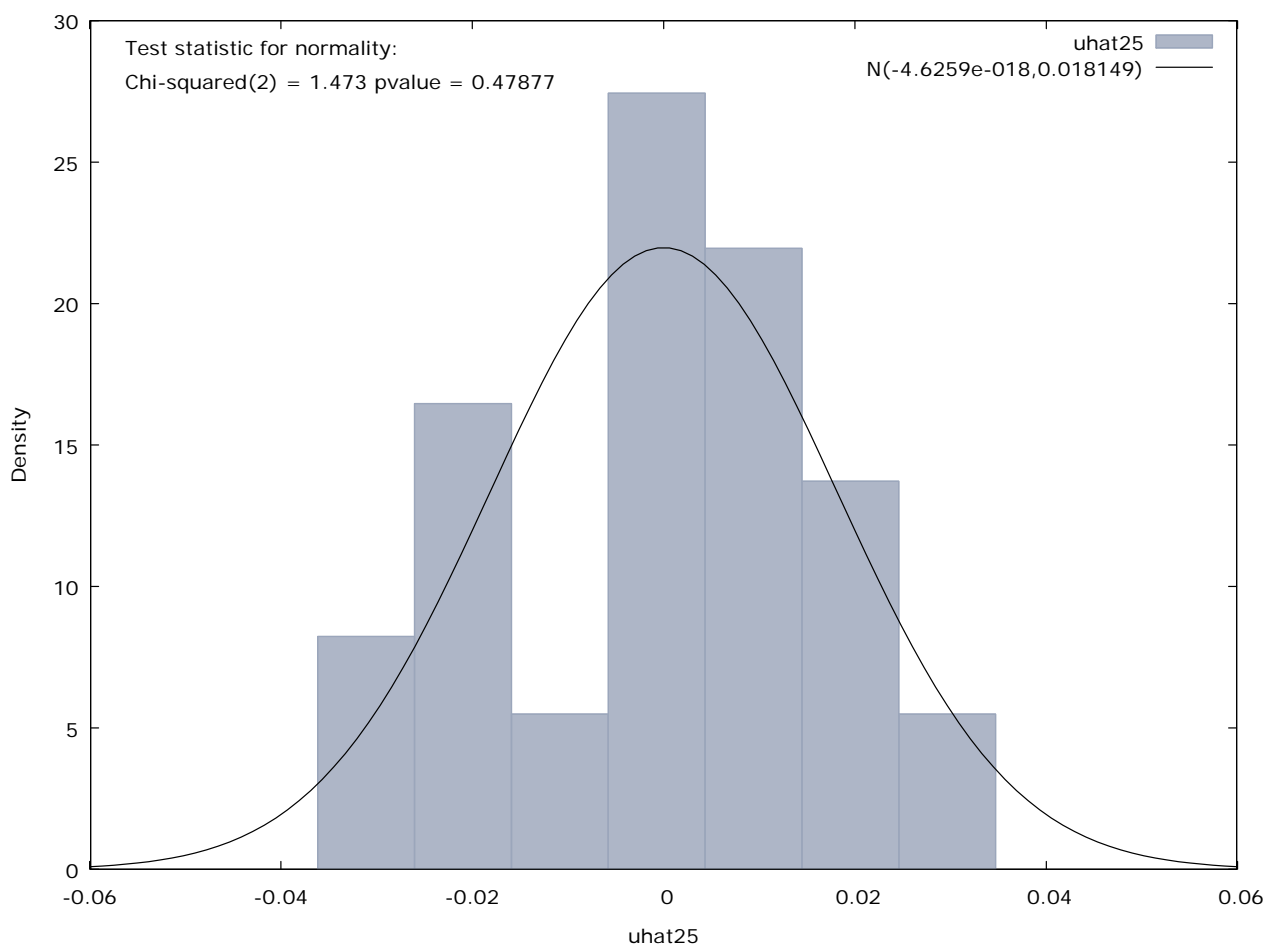
2.2 誤差項の正規性検定

1. 2 の c) のモデルの出力ウィンドウにもどる。

メニューバーの「Tests」 「Normality of residual」を選択する。

結果として図 7 のグラフが得られる。この上の p 値をみて、これが 0.05 を下回った場合、誤差項の正規性も疑われる。もしそうだった場合、標本サイズ (= 観測個数) が十分大きい場合 (50 程度以上)、F 検定を使うよりもカイ二乗検定を利用すべきだし、t 検定を利用するよりも $N(0, 1)$ 分布に収束する検定を利用すべきである。また、標本数が少ない場合は、言いたい仮説に保守的に使用する分布を選ぶ。係数が 0 ではないということを言いたい場合は、それが言いにくい t 分布による検定を利用する。係数が 0 であるということがいいければ、それがいいにくい正規分布による検定を行う。

この例では、誤差項は正規分布とって良い。



☒ 7