

# 第 10 回 : 回帰分析結果表の作成

【教科書該当箇所なし】

北村 友宏

2020 年 12 月 11 日

# 本日の内容

1. 前回の分析の再実行
2. モデル推定結果表の作成

# gretl での重回帰分析

「駅へのアクセスのよさがマンション価値に与える影響」を分析するためのモデル

$$price_i = \beta_0 + \beta_1 minutes_i + \beta_2 age_i + u_i$$

- ▶  $price_i$  : 中古マンション価格 (万円)
- ▶  $minutes_i$  : 最寄り駅までの所要時間 (分)
- ▶  $age_i$  : 築年数 (年)
- ▶  $i$  : 中古マンション番号

を推定する (今回は, 「面積」の変数は使わない).

教科書第6章4節に載っているモデルのうち、「モデル1」と「モデル2」を推定する。

- ▶ **モデル1**：価格を所要時間のみに回帰（単回帰）。
  - ▶ 「所要時間」以外の要因はコントロールしない。
  - ▶  $\beta_2 = 0$  と仮定。
- ▶ **モデル2**：価格を所要時間と築年数に回帰。
  - ▶ 「築年数」をコントロール。

# 実習 1

まず、モデル 1（価格を所要時間のみに回帰）を推定する。

1. gretl を起動.
2. 「ファイル」→「データを開く」→「ユーザー・ファイル」と操作.
3. setagayaapartment.gdt を選択し、「開く」をクリック.
4. gretl のメニューバーから「モデル」→「通常の最小二乗法」と操作.
5. 出てきたウィンドウ左側の変数リストにある price\_10th をクリックし、3つの矢印のうち上の青い右向き矢印をクリック.
  - ▶ 推定式の左辺の変数（被説明変数，従属変数）が price\_10th（万円単位の中古マンション価格）となる.

6. 「デフォルトとして設定」にチェック。
  - ▶ gretl を終了するまでの間、次回以降「通常の最小二乗法」での推定を行う際に、いま選択した変数が自動的に被説明変数（従属変数）に入力される。
7. ウィンドウ左側の変数リストにある minutes をクリックし、3つの矢印のうち真ん中の緑の右向き矢印をクリック。
  - ▶ 推定式の右辺の変数（説明変数、独立変数）が minutes（最寄り駅までの所要時間）となる。
  - ▶ 最初から説明変数リストに入っている const は推定式の切片（定数項）のこと。
8. 「頑健標準誤差を使用する」にチェック。これで、推定式の誤差項  $u_i$  のバラつき（分散）に関する仮定が誤っていても、より厳密な分析ができるようになる。
9. 「OK」をクリックすると、結果が新しいウィンドウに表示される。

gretl: モデル

ファイル 編集(E) 検定(D) 保存(S) グラフ(G) 分析(A) LaTeX

モデル 1

モデル 1: 最小二乗法 (OLS), 観測: 1-194  
 従属変数: price\_10th  
 不均一分散頑健標準誤差, バリエーション HCl

	係数	標準誤差	t値	p値	
const	3092.68	245.524	12.60	6.55e-027	***
minutes	74.5608	22.0194	3.386	0.0009	***
Mean dependent var	3762.577	S.D. dependent var	2150.961		
Sum squared resid	8.62e+08	S.E. of regression	2118.252		
R-squared	0.035207	Adjusted R-squared	0.030182		
F(1, 192)	11.46597	P-value(F)	0.000860		
Log-likelihood	-1759.988	Akaike criterion	3523.976		
Schwarz criterion	3530.512	Hannan-Quinn	3526.623		

このような画面が表示されれば成功。まだ作業があるので、「gretl: モデル」のウィンドウは**まだ閉じない!**

## 実習 2

続いて、モデル 2（価格を所要時間と築年数に回帰）を推定する。

1. gretl のメニューバーから「モデル」→「通常の最小二乗法」と操作。説明変数（独立変数）は必ず前回の選択内容が記録されており、被説明変数（従属変数）は前回「デフォルトとして設定」にチェックしていれば前回の選択内容が記録されている。
2. 従属変数の入力ボックスに price\_10th が入力されていなければ、出てきたウィンドウ左側の変数リストにある price\_10th をクリックし、3つの矢印のうち上の青い右向き矢印をクリック。
  - ▶ 推定式の左辺の変数（被説明変数，従属変数）が price\_10th（万円単位の中古マンション価格）となる。



3. ウィンドウ左側の変数リストにある age をクリックし，3つの矢印のうち真ん中の緑の右向き矢印をクリック。
  - ▶ 推定式の右辺の変数（説明変数，独立変数）が minutes（最寄り駅までの所要時間）と age（築年数）の2つとなる。
  - ▶ 最初から説明変数リストに入っている const は推定式の切片（定数項）のこと。
4. 「頑健標準誤差を使用する」にチェック．これで，推定式の誤差項  $u_i$  のバラつき（分散）に関する仮定が誤っていても，より厳密な分析ができるようになる．
5. 「OK」をクリックすると，結果が表示される．

gretl: モデル

ファイル 編集(E) 検定(I) 保存(S) グラフ(G) 分析(A) LaTeX

モデル 1 ✕ モデル 2 ✕

モデル 2: 最小二乗法(OLS), 観測: 1-194  
 従属変数: price\_10th  
 不均一分散頑健標準誤差, バリエーション HC1

	係数	標準誤差	t値	p値	
const	4091.40	292.830	13.97	5.05e-031	***
minutes	69.0150	19.7832	3.489	0.0006	***
age	-62.4229	9.52150	-6.556	5.03e-010	***
Mean dependent var	3762.577	S.D. dependent var	2150.961		
Sum squared resid	7.64e+08	S.E. of regression	1999.784		
R-squared	0.144584	Adjusted R-squared	0.135627		
F(2, 191)	36.63869	P-value(F)	3.40e-14		
Log-likelihood	-1748.317	Akaike criterion	3502.633		
Schwarz criterion	3512.437	Hannan-Quinn	3506.603		

このような画面が表示されれば成功。まだ作業があるので、「gretl: モデル」のウィンドウは**まだ閉じない!**

# レポートや論文での推定結果表の作成

見やすく，理解しやすい表を載せるには，

- ▶ 表番号と表のタイトルをつける。
- ▶ 最低限，以下の情報を載せる（線形回帰モデルの場合）。
  - ▶ 係数推定値
  - ▶  $t$  値または標準誤差または  $p$  値のどれか  
（学術論文では標準誤差を載せる場合が多い。）
  - ▶ 決定係数
  - ▶ 観測値数
- ▶ 有意性を示すアスタリスクを付けた場合は，表の下に「注（Note）」として「表中の\*\*\*，\*\*，\* はそれぞれ有意水準 1%，5%，10%で統計的に有意であることを表す」などと注記する。

- ▶ 仮説検定に用いた標準誤差の種類や、頑健なら何に対して頑健なのかを、表中に明記するか、表の下に「注 (Note)」として「不均一分散に対して頑健な標準誤差を用いている」などと注記する.
- ▶ 観測値数は、表の下に「注 (Note)」として「観測値数は〇〇〇である」などと注記してもよい
  - ▶ e.g., 1つの表に複数のモデルを載せ、観測値数が全モデルについて同じ場合など.
- ▶ 変数名は統計解析ソフトでの変数名そのままではなく、分かりやすいように書き直す.
- ▶ 小数の数値はあまり細かく表示せず、小数第2~4位程度まで示せば十分. 特に係数推定値, 標準誤差,  $t$  値,  $p$  値は縦方向に見たとき、可能な限り小数点の位置が揃うようにする.
  - ▶  $t$  値は小数第2位まで,  $p$  値, 決定係数は小数第3位まで示せば十分.

## 実習 3

1. Word を起動し，results20201211.docx という名前で 2020microdatag フォルダに保存.
2. 「挿入」→「表」と操作して 6 行 7 列の表を作る.
3. 表全体をドラッグし，「参考資料」（Mac 版では「参照設定」）→「図表番号の挿入」と操作.
4. ラベルを「表」に，位置を「選択した項目の上」して OK をクリックすると，表のすぐ上の行に「表 1」と入力される.
  - ▶ ラベルに「表」がなければ，「新しいラベル...」をクリックして出てくるダイアログボックスの入力ボックスに表と入力して OK をクリック.
5. 「表 1」の後に全角スペースを入れてモデル推定結果と入力し，その行を中央揃えにする.

6. 表の 2 行 2 列目に回帰係数，2 行 3 列目に標準誤差，2 行 5 列目に偏回帰係数，2 行 6 列目に標準誤差と入力.
7. 表の 2 行 2 列目から 2 行 7 列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「下揃え（中央）」と操作.
8. 表の 1 列目に，以下のように入力.
  - ▶ 3 行 1 列目：最寄り駅所要時間
  - ▶ 4 行 1 列目：築年数
  - ▶ 5 行 1 列目：定数項
  - ▶ 6 行 1 列目：自由度修正済み決定係数
9. 表の 3 行 1 列目から 6 行 1 列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「中央揃え」と操作.

10. 表の1行1列目から2行1列目までをドラッグし、「レイアウト」タブ（右端の、色が濃いほう）から「結合」→「セルの結合」と操作.
11. 表の1行2列目から1行4列目までをドラッグし、「レイアウト」タブ（右端の、色が濃いほう）から「結合」→「セルの結合」と操作して、「配置」→「中央揃え」と操作. 結合したセルにモデル1と入力.
12. 表の1行5列目から1行7列目までをドラッグし、「レイアウト」タブ（右端の、色が濃いほう）から「結合」→「セルの結合」と操作して、「配置」→「中央揃え」と操作. 結合したセルにモデル2と入力.

13. gretl に出力されていた，minutes のみを説明変数とするモデル（モデル 1）の推定結果の数値を，Word で作成した表の**モデル 1**の対応するセルにコピー・貼り付けする．数値をドラッグして選択し，右クリック→「コピー」と操作すればコピーできる．
- ▶ const は定数項，minutes は最寄り駅所要時間，係数は回帰係数，Adjusted R-squared は自由度修正済み決定係数．
  - ▶ 築年数に対応するセルは空欄にする．
  - ▶ 自由度修正済み決定係数は，回帰係数の列（6 行 2 列目）に入力する．
14. 回帰係数，標準誤差，自由度修正済み決定係数は**小数第 3 位を四捨五入**．



15. gretl に出力されていた，minutes のみを説明変数とするモデル（モデル 1）の推定結果を見て，標準誤差の右隣のセルには，その変数の係数の  $p$  値が 0.01 未満なら\*\*\*，0.05 未満なら\*\*，0.10 未満なら\*と入力する.
16. 表の 3 行 2 列目から 6 行 3 列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「**下揃え（右）**」と操作.
17. 表の 3 行 4 列目から 5 行 4 列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「**下揃え（左）**」と操作.

18. gretl に出力されていた，minutes と age を説明変数とするモデル（モデル 2）の推定結果の数値を，Word で作成した表の**モデル 2**の対応するセルにコピー・貼り付けする．数値をドラッグして選択し，右クリック→「コピー」と操作すればコピーできる．
- ▶ const は定数項，minutes は最寄り駅所要時間，age は築年数，係数は偏回帰係数，Adjusted R-squared は自由度修正済み決定係数．
  - ▶ 自由度修正済み決定係数は，偏回帰係数の列（6 行 5 列目）に入力する．
19. 偏回帰係数，標準誤差，自由度修正済み決定係数は**小数第 3 位を四捨五入**．

20. gretl に出力されていた，minutes と age を説明変数とするモデル（モデル 2）の推定結果を見て，標準誤差の右隣のセルには，その変数の係数の  $p$  値が 0.01 未満なら\*\*\*，0.05 未満なら\*\*，0.10 未満なら\*と入力する.
21. 表の 3 行 5 列目から 6 行 6 列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「**下揃え（右）**」と操作.
22. 表の 3 行 7 列目から 5 行 7 列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「**下揃え（左）**」と操作.

## 23. Word で作成した表のすぐ下の行に,

(注1) 表中の\*\*\*は有意水準 1%で統計的に有意であることを表す.

(注2) 不均一分散に対して頑健な標準誤差を用いている.

(注3) 観測値数は 194 である.

と入力して上書き保存.

- ▶ 「アスタリスク1つ (有意水準 10%) と2つ (有意水準 5%)」はこの表では出てこなかったので省略.
- ▶ 観測値数は, 出力結果の「観測数」と記載されている箇所を見れば分かる.

表 1 モデル推定結果

	モデル 1			モデル 2		
	回帰係数	標準誤差		偏回帰係数	標準誤差	
最寄り駅 所要時間	74.56	22.02	***	69.02	19.78	***
築年数				-62.42	9.52	***
定数項	3092.68	245.52	***	4091.40	292.83	***
自由度修 正済み決 定係数	0.03			0.14		

(注1) 表中の\*\*\*は有意水準 1%で統計的に有意であることを表す。  
(注2) 不均一分散に対して頑健な標準誤差を用いている。  
(注3) 観測値数は 194 である。

このような表を作成できればよい。

本日の作業はここまで.

今回は gretl のデータセットに変更を加えていないので、gretl のデータセット (setagayaapartment.gdt) を上書き保存する必要はない.