第 8 回：単回帰モデルの推定

北村　友宏

2019 年 6 月 26 日

本日の内容

1. 単回帰モデル

2. 消費関数の推定

単回帰

大きさ n の 2 変量データ
$((y_1, x_1), (y_2, x_2), \cdots, (y_n, x_n))$ を用いて，線形回帰モデル（linear regression model）

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + u_i, \\
E(u_i \mid x_i) &= 0, \\
E(u_i u_j \mid x_i) &= 0 \quad (i \neq j), \\
V(u_i \mid x_i) &= \sigma^2, \\
i &= 1, 2, \cdots, n
\end{aligned}
$$

を推定することを考える.
これを推定すれば，2 つの変数間の関係（x_i が増加すると y_i はどの程度変化する傾向があるか？）を定量的に検証できる.

- y_i：被説明変数（explained variable）
  - e.g., 消費支出
  - 従属変数（dependent variable）ともいう.
- x_i：説明変数（explanatory variable）
  - e.g., 可処分所得
  - 独立変数（independent variable）ともいう.
- β_0, β_1：回帰係数（regression coefficient）
  - 特に，β_0 は定数項（constant term）.
  - x_i が 1 単位増加すると y_i は β_1 単位増加する，という解釈.
- u_i：誤差項（error term）
  - 撹乱項（disturbance term）ともいう.

説明変数 x_i は確率的（stochastic）とする.

- 定数項以外の説明変数が 1 つである回帰モデルを単回帰モデル（simple regression model）という．

$E(u_i \mid x_i) = 0$ の仮定より，

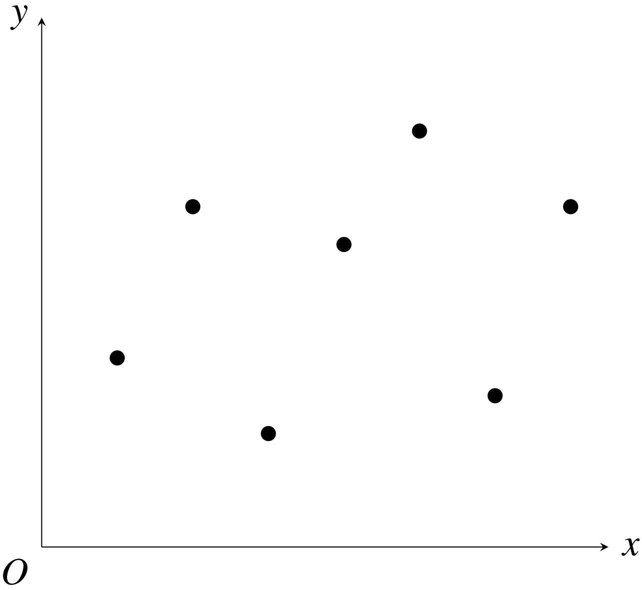
$$E(y_i \mid x_i) = \beta_0 + \beta_1 x_i.$$

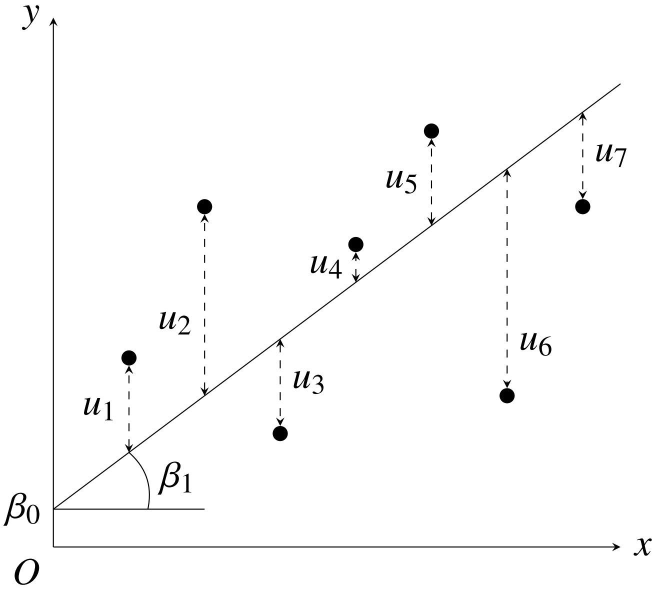
$\Longrightarrow$ これは x_i が与えられたときの y_i の条件付き期待値（conditional mean）．

- $E(y_i \mid x_i)$ を求めることを，y_i を x_i に回帰する（regress）という．

$$\Downarrow$$

β_0 と β_1 を求める（推定する）には？


y
x
O


y
x
O
β_0
β_1
u_1
u_2
u_3
u_4
u_5
u_6
u_7

モデルを

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$

と書き換え，

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2$$

が最小になるような $\hat{\beta}_0$ と $\hat{\beta}_1$ を求める．

- e_i：残差（residual）
  - 誤差項 u_i とは別物．
- $\sum_i^n e_i^2$ が最小になるように回帰係数を求める方法を通常の最小二乗法（Ordinary Least Squares, OLS）という．

- OLS によって推定される統計量を OLS 推定量（OLS estimator）といい，その実現値を OLS 推定値（OLS estimate）という．

この場合の OLS 推定量は，

$$
\begin{aligned}
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x}, \\
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}.
\end{aligned}
$$

- $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$.
- $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$.

OLS 推定における仮定（単回帰の場合）

- 説明変数を所与として，誤差項の期待値はゼロ．
  - $E(u_i \mid x_i) = 0.$

  $\Rightarrow$ 説明変数と誤差項は無相関．

- 説明変数を所与として，誤差項の分散は一定で，異なる個体の誤差項同士は無相関．
  - $V(u_i \mid x_i) = \sigma^2.$
  - $E(u_i u_j \mid x_i) = 0 \quad (i \neq j).$

- 説明変数を所与として，誤差項は正規分布に従う．
  - $u_i \mid x_i \sim N\left(0, \sigma^2\right).$

消費関数の推定

いま整理・加工・分析している都道府県別・男女別データセットを用いて，ケインズ型消費関数

$$c_i = \beta_0 + \beta_1 y_i + u_i$$

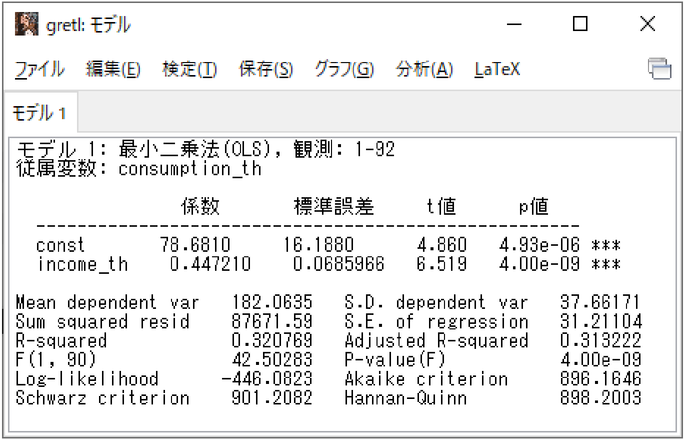
- c_i：消費支出
- y_i：可処分所得
- β_0：基礎消費
  - たとえ可処分所得がゼロであっても生活のためにどの程度消費支出をする必要があるかを表す.
- β_1：限界消費性向
  - 可処分所得が 1 単位増加すると消費支出が何単位増加する傾向があるかを表す.

を推定する.

実習 1

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. 消費 2009.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 consumption_th 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 consumption_th（千円単位の消費支出）となる.
6. ウィンドウ左側の変数リストにある income_th 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 income_th（千円単位の可処分所得）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.
7. 「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.



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！**

出力結果の見方

- 係数: 回帰係数推定値
- 標準誤差: 回帰係数の標準誤差
  - 次回の授業で説明
- t 値: 「回帰係数が 0」という帰無仮説の両側 t 検定における検定統計量の実現値（t 値）
  - 次回の授業で説明
- p 値: 両側 p 値
  - 次回の授業で説明
- R-squared: 決定係数

決定係数

決定係数（R-squared）は,

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}.$$

- 定数項ありの単回帰の場合，$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$.
- 意味 モデルの当てはまりの良さ（説明変数で，被説明変数の変動のうち，どの程度の割合を説明できているか）
- $0 \leq R^2 \leq 1$.
  - $R^2 = 0$: 全く説明できていない.
  - $R^2 = 1$: 完全に説明できている.

  $\Rightarrow R^2 = 0$ や $R^2 = 1$ になることは，実際の実証分析ではまず起こり得ない.

ケインズ型消費関数推定結果

- 所得の係数
  - 0.44721（符号は正）
  - 限界消費性向の推定値
  - 可処分所得が千円高くなると，消費支出が平均して 447.21 円（0.44721 千円）高くなる．
- 定数項
  - 78.681（符号は正）
  - 基礎消費の推定値
  - 仮に可処分所得が 0 円であっても，生活のために毎月平均 78,681 円（78.681 千円）の消費支出をする必要がある．
- 決定係数
  - $R^2 = 0.320769.$
    ➥ 所得は消費の変動の約 32%を説明できている．

$\Longrightarrow$ 経済理論と整合的．

実習 2

1. 表示された「gretl: モデル 1」のウィンドウのメニューバーから「ファイル」→「名前を付けて保存」と操作。
2. 「標準テキスト」を選び，「OK」をクリック。
3. 消費関数推定結果 1.txt という名前で「2020 ミクロデータ分析 1」フォルダに保存．すると，表示された推定結果をそのままテキストファイルで保存できる．本日の作業はここまで．