第 8 回：回帰係数の検定
【教科書第 6 章第 4 節～第 5 節】

北村　友宏

2025 年 11 月 18 日

本日の内容

1. 単一の回帰係数の検定
2. 線形制約の検定
3. OLS 推定量の性質（大標本の場合）

実証分析例：ミンサー方程式の推定

就業可能年数をコントロールしたうえで，「修学年数が増えると，年収がどれだけ増えるのか」を分析するためのモデル（ミンサー方程式）

$$\ln income_i = \beta_0 + \beta_1 yeduc_i + \beta_2 exper_i + \beta_3 exper_i^2 + u_i$$

- $income_i$：年収（万円）
- $yeduc_i$：修学年数（年）
- $exper_i$：就業可能年数（年）
- i：個人番号

を推定する.

➥ 「年収の対数値」を「修学年数」と「就業可能年数」と「就業可能年数の２乗」に回帰する.

- 就業可能年数
    - 最後の学校を卒業してからの年数

$$\text{就業可能年数} = \text{年齢} - \text{修学年数} - 6.$$

※ 小学校に入学する年齢が 6 歳のため，6 を引いている.

    - 熟練度を表す.
      $\Rightarrow$ 賃金に影響を与える.

$$\Downarrow$$

修学年数が年収に与える純粋な効果を計測するには，熟練度（を表す就業可能年数) をコントロールする必要がある.

モデル推定結果

gretl: モデル1

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

モデル 1: 最小二乗法(OLS), 観測: 1-4299
従属変数: lincome

	係数	標準誤差	t値	p値	
const	2.48550	0.110782	22.44	1.64e-105	***
yeduc	0.117547	0.00706026	16.65	2.31e-060	***
exper	0.196174	0.00749354	26.18	2.75e-140	***
exper2	−0.00638115	0.000316188	−20.18	1.32e-086	***

Mean dependent var	5.290452	S.D. dependent var	0.895883
Sum squared resid	2736.906	回帰の標準誤差	0.798267
R-squared	0.206603	Adjusted R-squared	0.206049
F(3, 4295)	372.8097	P-value(F)	3.4e-215
Log-likelihood	−5129.400	Akaike criterion	10266.80
Schwarz criterion	10292.27	Hannan-Quinn	10275.79

出力結果の見方

- 係数: （偏）回帰係数推定値
- 標準誤差: （偏）回帰係数の標準誤差
- t 値: 「（偏）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）
- p 値: 両側 p 値
- 回帰の標準誤差：誤差項の標準誤差
- R-squared: 決定係数
- Adjusted R-squared: 自由度修正済み決定係数

標準誤差

- 標準偏差の推定値を**標準誤差（standard error）**という．
- 重回帰モデル

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{X\beta} + \boldsymbol{u}, \\ E(\boldsymbol{u} \mid \boldsymbol{X}) &= \boldsymbol{0}, \\ V(\boldsymbol{u} \mid \boldsymbol{X}) &= \sigma^2 \boldsymbol{I}_n, \end{aligned}$$

における j 番目の（偏）回帰係数の OLS 推定量 $\hat{\beta}_j$ の（デフォルトの）標準誤差は，

$$\mathrm{s.e.}\left(\hat{\beta}_j\right) = \sqrt{\left[\frac{\boldsymbol{e}'\boldsymbol{e}}{n-k-1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\right]_{j,j}}.$$

$\Rightarrow$ この標準誤差は，任意の i について $V(u_i \mid \boldsymbol{X})$ が一定（均一分散）の場合のみ正しい．

仮説検定

- 被説明変数や説明変数は様々な値をとり，観測される前はどのような値になるかが不確定（確率変数）.
- 被説明変数や説明変数を用いて計算する（偏）回帰係数の値も不確定.

⇒ 例えば回帰係数 β_1 の推定値として $\hat{\beta}_1 = 0.117547$ という値が得られても，「推定値 $\hat{\beta}_1$ は真の β_1 の値と必ずしも同じではなく，真の β_1 は 0 で，その推定値 $\hat{\beta}_1$ は様々な値をとりうる中でたまたま 0.117547 になった」可能性もある.

⇒ 仮説検定を行い，「真の（偏）回帰係数が 0 かどうか」を検証する.

gretl などの統計解析ソフトで線形回帰モデルを推定すると，各（偏）回帰係数 β_j について，

$$H_0 : \beta_j = 0 \quad \text{vs} \quad H_1 : \beta_j \neq 0$$

を検定するのに必要な情報が出力される．

- 回帰分析では，通常は両側検定を行う．

- $\boxed{H_0\text{（係数は 0）が棄却された}}$
  ➡「その（偏）回帰係数は統計的に有意に 0 と異なる」と判断.
  - 「その説明変数は被説明変数と統計的に有意に相関している」と解釈.
  - 定数項の検定の場合は「定数項は統計的に有意に 0 と異なる」と解釈.

- $\boxed{H_0\text{（係数は 0）が採択された}}$
  ➡「その（偏）回帰係数は 0 と異なるとは言えない」と判断.
  - 「その説明変数は被説明変数と相関しているとは言えない」と解釈.
  - 定数項の検定の場合は「定数項は統計的に有意に 0 と異なるとは言えない」と解釈.

p 値による判断

- 検定統計量（の絶対値）が実現値（検定統計値）を超える（以上になる）確率を p 値（p-value）という.
  - p 値が 0.1 以下（未満）：有意水準 10%で H_0 を棄却.
  - p 値が 0.05 以下（未満）：有意水準 5%で H_0 を棄却（10%でも棄却できる）.
  - p 値が 0.01 以下（未満）：有意水準 1%で H_0 を棄却（5%, 10%でも棄却できる）.

$\Rightarrow$ p 値を見て，帰無仮説の採択・棄却を判断できる.

※ 検定統計量が連続型の確率分布（正規分布，t 分布，カイ二乗分布，F 分布など）に従う場合，「以上」と「超える」，「以下」と「未満」は区別しなくて良い.

gretl では，モデル推定結果の各説明変数の行の右端にアスタリスク（＊）が表示され，＊の個数を見れば，「有意水準何%で『(偏）回帰係数は 0』の H_0 を棄却できるか」が分かる．

- （アスタリスクなし）: 有意水準 10%でも「係数は 0」の H_0 採択．
- *: 有意水準 10%で，「係数は 0」の H_0 棄却．
- **: 有意水準 5%で，「係数は 0」の H_0 棄却（10%でも棄却できる）．
- ***: 有意水準 1%で，「係数は 0」の H_0 棄却（5%, 10%でも棄却できる）．

t 値による判断

$\beta_j = 0$ という H_0 を検定するための t 検定統計量は，

$$t = \frac{\hat{\beta}_j}{\mathrm{s.e.}\left(\hat{\beta}_j\right)} \sim t(n-k-1).$$

- 観測値数が十分に大きいとき，t 値の絶対値がほぼ 2 を超えていれば，H_0 を棄却と判断（大雑把な判断）．
  ➡ 「有意水準何%で H_0 を棄却できるか」を厳密に判断するには，t 値ではなく p 値を見る．

モデル推定結果

- 修学年数の係数
  - 0.117547（符号は正）
  - t 値は 16.65, p 値は 2.31×10^{-60}.
    ➡ 仮に「yeduc の係数が 0」だとすると，16.65 という t 値は 2.31×10^{-60}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない.
    ➡ 有意水準 1%で，「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）.
    ➡ 修学年数は年収と統計的に有意に相関している.
  - 年収と修学年数についてはログ＝レベル・モデルの関係.
    ➡ 就業可能年数（とその 2 乗）を一定としたうえで，修学年数が 1 年長くなると，年収が平均して 11.7547%高くなる傾向がある.

- 就業可能年数の 1 乗の係数
  - 0.196174（符号は正）
  - t 値は 26.18, p 値は 2.75×10^{-140}.
    ➡ 仮に「exper の係数が 0」だとすると，26.18 という t 値は 2.75×10^{-140}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない.
    ➡ 有意水準 1%で，「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）.
    ➡ 就業可能年数の 1 乗は年収と統計的に有意に相関している.

この係数の解釈は，次回の授業で説明する.

- 就業可能年数の 2 乗の係数
  - −0.00638115（符号は負）
  - t 値は −20.18, p 値は 1.32×10^{-86}.
    ➡ 仮に「exper2 の係数が 0」だとすると，−20.18 という t 値は 1.32×10^{-86}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない.
    ➡ 有意水準 1%で，「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）.
    ➡ 就業可能年数の 2 乗は年収と統計的に有意に相関している.

この係数の解釈は，次回の授業で説明する.

- 定数項
  - 2.4855（符号は正）
  - t 値は 22.44, p 値は 1.64×10^{-105}
    ➡ 仮に「定数項が 0」だとすると，22.44 という t 値は 1.64×10^{-105}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない．
    ➡ 有意水準 1%で，「係数は 0」の H_0 が棄却される（5%や 10%でも棄却される）．
    ➡ 定数項は統計的に有意に 0 と異なる．

線形制約の検定

ミンサー方程式

$$\ln income_i = \beta_0 + \beta_1 yeduc_i + \beta_2 exper_i + \beta_3 exper_i^2 + u_i$$

について,

$$H_0 : \beta_2 = 0 \text{ かつ } \beta_3 = 0 \quad \text{vs} \quad H_1 : \beta_2 \neq 0 \text{ または } \beta_3 \neq 0$$

を検定することを考える.

- H_0 の意味：
  - β_2 と β_3 はどちらも 0.
    ➥ モデルの説明変数に $exper_i$ と $exper_i^2$ は両方とも不要.
- H_1 の意味：
  - 少なくとも β_2 か β_3 のどちらか一方は 0 でない.
    ➥ モデルの説明変数に少なくとも $exper_i$ と $exper_i^2$ のどちらかが必要.

- H_0 が真という仮定の下で，ミンサー方程式には $\beta_2 = 0$ と $\beta_3 = 0$ の 2 つの線形制約を課していることになる.
  ➥ 制約の数は 2.

別の例として，コブ・ダグラス型生産関数

$$\ln y_i = \beta_0 + \beta_k \ln k_i + \beta_l \ln l_i + u_i$$

に対して，規模に関して収穫一定の線形制約

$$\beta_k + \beta_l = 1$$

を H_0 とし，

$\beta_k + \beta_l \neq 1$（規模に関して収穫一定でない）
$\beta_k + \beta_l > 1$（規模に関して収穫逓増）
$\beta_k + \beta_l < 1$（規模に関して収穫逓減）

のどれか 1 つを H_1 とした検定をすることもできる．
➥ 制約の数は 1（「$\beta_k + \beta_l = 1$」の 1 つ）．

- y_i: 生産量または生産額
- k_i: 資本量または資本額
- l_i: 労働量

定数項以外に説明変数が k 個ある重回帰モデルにおいて，r 個の線形制約を検定するための検定統計量とその分布は，

$$F = \frac{(Q_R - Q_{UR})/r}{Q_{UR}/(n-k-1)} \sim F(r, n-k-1).$$

- Q_R: 制約付きモデルの残差二乗和
- Q_{UR}: 制約なしモデルの残差二乗和

（F 分布に従うことの証明は省略）

$$\Downarrow$$

F 検定をすればよい．

この検定統計量は，決定係数を用いて以下のように書き換えることもできる.

$$F = \frac{(R_{UR}^2 - R_R^2)/r}{(1 - R_{UR}^2)/(n-k-1)} \sim F(r, n-k-1).$$

- R_{UR}^2: 制約なしモデルの決定係数
- R_R^2: 制約付きモデルの決定係数

※ R_{UR}^2 と R_R^2 はどちらも，自由度を「修正していない」決定係数.

（証明）

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

とすると，決定係数の定義より，

$$R_R^2 = 1 - \frac{Q_R}{TSS},$$
$$R_{UR}^2 = 1 - \frac{Q_{UR}}{TSS},$$

なので，

$$\frac{Q_R}{TSS} = 1 - R_R^2,$$
$$\frac{Q_{UR}}{TSS} = 1 - R_{UR}^2.$$

よって，

$$\frac{Q_R - Q_{UR}}{TSS} = \frac{Q_R}{TSS} - \frac{Q_{UR}}{TSS}$$
$$= (1 - R_R^2) - (1 - R_{UR}^2)$$
$$= R_{UR}^2 - R_R^2.$$

したがって，

$$F = \frac{(Q_R - Q_{UR})/r}{Q_{UR}/(n-k-1)}$$
$$= \frac{\frac{Q_R - Q_{UR}}{TSS}/r}{\frac{Q_{UR}}{TSS}/(n-k-1)}$$
$$= \frac{(R_{UR}^2 - R_R^2)/r}{(1 - R_{UR}^2)/(n-k-1)}.$$（証明終）

- 大標本の場合，以下の検定統計量を用いて線形制約の検定を行うこともできる．

$$\chi^2 = (n-k-1)\cdot\frac{Q_R - Q_{UR}}{Q_{UR}} \sim \chi^2(r),$$

$$\chi^2 = (n-k-1)\cdot\frac{R_{UR}^2 - R_R^2}{1-R_{UR}^2} \sim \chi^2(r).$$

（証明は省略）

$$\Downarrow$$

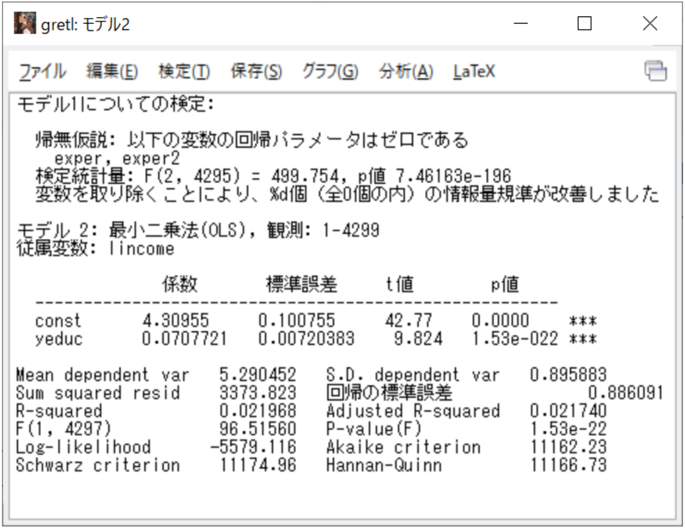
カイ二乗検定をすればよい．

gretl で複数の係数が 0 の H_0 を検定する方法

1. モデルを推定した後，結果ウィンドウの「検定」→「変数を取り除く」と操作．
2. 出てきたウィンドウ左側の変数リストの中から，H_0 において係数を 0 にしたい説明変数を選び，2 つの矢印のうち上の緑の右向きの矢印をクリック．
   - 右側のリストに，H_0 が真であるという仮定の下で係数を 0 とする説明変数が追加される．
3. 「変数を減らしたモデルを推定する」をクリック．
4. 「OK」をクリック．

制約なしモデルを通常の最小二乗法で推定した場合，F 検定が行われる．

線形制約の検定結果



gretl: モデル2

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

モデル1についての検定:

帰無仮説: 以下の変数の回帰パラメータはゼロである
exper, exper2
検定統計量: F(2, 4295) = 499.754, p値 7.46163e-196
変数を取り除くことにより、%d個 (全0個の内) の情報量規準が改善しました

モデル 2: 最小二乗法(OLS), 観測: 1-4299
従属変数: lincome

	係数	標準誤差	t値	p値	
const	4.30955	0.100755	42.77	0.0000	***
yeduc	0.0707721	0.00720383	9.824	1.53e-022	***

Mean dependent var	5.290452	S.D. dependent var	0.895883
Sum squared resid	3373.823	回帰の標準誤差	0.886091
R-squared	0.021968	Adjusted R-squared	0.021740
F(1, 4297)	96.51560	P-value(F)	1.53e-22
Log-likelihood	-5579.116	Akaike criterion	11162.23
Schwarz criterion	11174.96	Hannan-Quinn	11166.73

- 検定統計量の実現値（検定統計値）は 499.754, p 値は 7.46163×10^{-196}
  ➡ 仮に「exper と exper2 の係数がどちらも 0」だとすると，499.754 という検定統計値は 7.46163×10^{-196}, つまりほぼ 0%の確率（1%を下回る確率）でしか出てこない.
  ➡ 有意水準 1%で，「exper と exper2 の係数がどちらも 0」の H_0 が棄却される（5%や 10%でも棄却される）.
  ➡ exper と exper2 の係数のうち少なくとも一方は 0 でないと判断される.

$$\Downarrow$$

就業可能年数およびその 2 乗項をミンサー方程式に含めることには統計的にも意味がある.

OLS 推定量の性質（大標本の場合）

大標本の場合，

- 誤差項の（条件なし）期待値は 0

$$E(\boldsymbol{u}) = \mathbf{0}.$$

- すべての説明変数と誤差項は無相関

$$\mathrm{Cov}(x_{ji}, u_i) = 0, \quad j = 1, 2, \cdots, k.$$

の仮定の下で，重回帰モデルの OLS 推定量 $\hat{\boldsymbol{\beta}}$ は一致性をもつ．

$$\operatorname*{plim}_{n\to\infty} \hat{\boldsymbol{\beta}} = \boldsymbol{\beta}.$$

⇒ 標本サイズが十分に大きい（観測値数が十分に多い）とき，重回帰モデルの偏回帰係数の OLS 推定量は真の偏回帰係数に確率収束する．（証明は省略）

前スライドの 2 つの仮定に加え，

- 任意の i について，すべての説明変数を所与とした誤差項の条件付き分散は一定

$$V(u_i \mid \boldsymbol{X}) = \sigma^2 \quad \text{for all } i.$$

の仮定の下で，重回帰モデルの OLS 推定量 $\hat{\boldsymbol{\beta}}$ は漸近正規性をもつ．

$$\hat{\boldsymbol{\beta}} \overset{a}{\sim} N\left(\boldsymbol{\beta}, \frac{\sigma^2}{n} \boldsymbol{M}_{\boldsymbol{XX}}^{-1}\right).$$

- $\boldsymbol{M}_{\boldsymbol{XX}}^{-1} = \operatorname*{plim}_{n \to \infty} \left(\frac{1}{n} \boldsymbol{X}' \boldsymbol{X}\right)^{-1}.$

⇒ 標本サイズが十分に大きい（観測値数が十分に多い）とき，重回帰モデルの偏回帰係数の OLS 推定量は近似的に正規分布に従う．（証明は省略）

今日のキーワード

標準誤差，p 値

次回までの準備

- 今回の講義スライドを読み直す.
- 「提出課題 5」に取り組む.
- 教科書第 7 章第 1 節〜第 3 節の「例 7.3：教育の収益率の男女差をチョウ検定で調べる」までを読む.