不均一分散

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$

$\mathrm{Var}(u \mid \mathbf{x}) = \sigma_i^2$

Ch.8 不均一分散

1. OLSでの不均一分散の帰結
2. OLS推定後の不均一分散頑健的推論
3. 不均一分散の検定
4. 加重最小二乗WLS推定
5. 線形確率モデル(再)*

8-1 OLSでの不均一分散の帰結

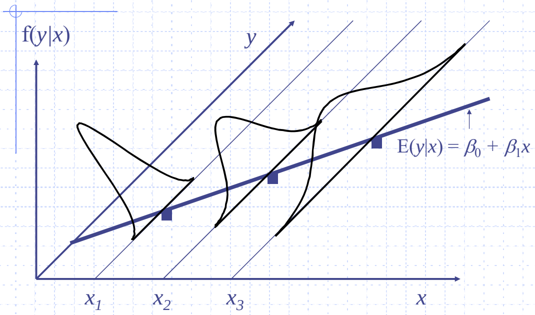
- ◆不均一分散のもとでのOLS推定量
  - OLS推定量は依然として不偏&一致推定量
  - R^2の解釈も変化しない

$$R^2 \approx 1 - \frac{\sigma_u^2}{\sigma_y^2}$$

無条件誤差分散は不均一分散(条件付き誤差分散)の影響を受けない

  - 注意点
    - OLS推定量の分散は有効ではない
    - 通常のt検定&F検定は有効でない
    - OLS推定量はBLUEではない
      - より効率的な線形推定量の存在

8-1 OLSでの不均一分散の帰結



8-2 不均一分散の頑健的推論

- ◆不均一分散頑健統計量
  - すべての計算式は大標本でのみ有効
  - 不均一分散頑健標準誤差HRSE

$$\widehat{Var}(\hat{\beta}_j) = \frac{\sum_{i=1}^{n} \hat{r}_{ij}^2 \hat{u}_i^2}{SSR_j^2}$$

    - White-Huber-Eicker標準誤差とも呼称
    - 元の回帰式とx_jを他の説明変数に回帰したときの式、それぞれの残差の2乗からなる。
  - この計算式を用いた通常のt検定は漸近的に有効
  - 通常のF統計は不均一分散では有効ではないが、この計算式は多くのソフトウェアで使用可能

8-2 不均一分散の頑健的推論

$$\hat{\beta}_1 = \beta_1 + \frac{\sum (x_i - \bar{x}) u_i}{SST_x}, \text{ where } SST_x = \sum (x_i - \bar{x})^2$$

$$Var\left(\hat{\beta}_1\right) = \frac{\sum (x_i - \bar{x})^2 \sigma_i^2}{SST_x^2} \qquad (8.2),$$

A valid estimator for this when $\sigma_i^2 \neq \sigma^2$ is

$$Est.Var\left(\hat{\beta}_1\right) = \frac{\sum (x_i - \bar{x})^2 \hat{u}_i^2}{SST_x^2} \qquad (8.3),$$

where $\hat{u}_i$ are the OLS residuals

8-2 不均一分散の頑健的推論

例：賃金方程式

$$\widehat{\log}(wage) = -.128 + .0904\,educ + .0410\,exper - .0007\,exper^2$$

	(.105)	(.0075)	(.0052)	(.0001)
	[.107]	[.0078]	[.0050]	[.0001]

- HRSEは通常の標準誤差よりも大きくも小さくもなるが、(大標本のもとで)その差は実際には小さい
- 不均一分散によってはその差が大きいこともあるため、常にHRSEを計算する方が良い

- F検定
  $H_0 : \beta_{exper} = \beta_{exper^2} = 0$
  $F = 17.95 \quad F_{robust} = 17.99$
  - F値も大きな違いはない

8-3 不均一分散の検定

不均一分散の検定

- 不均一分散の有無は、OLSが最も効率的な線形推定量であるかどうかと関連するため重要
- MLR.5
  $\boxed{H_0 : Var(u|x_1, x_2, \ldots, x_k) = Var(u|\mathbf{x}) = \sigma^2}$
  - 分散の定義とMLR.4より
  $Var(u|\mathbf{x}) = E(u^2|\mathbf{x}) - [E(u|\mathbf{x})]^2 = E(u^2|\mathbf{x})$
    - Var(a) = E[a – E(a)]² = E(a²) – E[E(a)]²
    - MLR.4: E(u|x) = 0
  $\Rightarrow\ E(u^2|x_1, \ldots, x_k) = E(u^2) = \sigma^2$
    - u^2の平均は$x_1, x_2, \ldots, x_k$の変動に影響を受けない

8-3 不均一分散の検定

Breusch-Pagan test for heteroskedasticity

- u^2への$x_1, x_2, \ldots, x_k$の影響を検定(BP)
  $\hat{u}^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k + error$
  $H_0 : \delta_1 = \delta_2 = \cdots = \delta_k = 0$
- 検定統計量

$$F = \frac{R^2_{\hat{u}^2}/k}{1 - R^2_{\hat{u}^2}/(n-k-1)}$$

大きな検定統計量(高いR^2)は帰無仮説に反する論拠

$$LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_k$$

LM統計量(Lagrange multiplier)においても高い統計値（高いR^2)は「u^2の期待値が説明変数と無相関」という帰無仮説を棄却

8-3 不均一分散の検定

例8.4：住宅価格の不均一分散検定

- レベル(非対数)

$$\widehat{price} = -21.77 + .0021\,lotsize + .123\,sqrft + 13.85\,bdrms$$

	(29.48)	(.0006)	(.013)	(9.01)

$\Rightarrow R^2_{\hat{u}^2} = .1601,\ p\text{-}value_F = .002,\ p\text{-}value_{LM} = .0028$

- 対数

$$\widehat{\log}(price) = -1.30 + .168\,\log(lotsize) + .700\,\log(sqrft) + .037\,bdrms$$

	(.65)	(.038)	(.093)	(.028)

$\Rightarrow R^2_{\hat{u}^2} = .0480,\ p\text{-}value_F = .245,\ p\text{-}value_{LM} = .2390$

- 対数化は不均一分散を緩和

8-3 不均一分散の検定

The White test for heteroskedasticity

- u^2への$x_1, x_2, \ldots, x_k$の影響を検定(White)
  $\hat{u}^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 + \delta_4 x_1^2 + \delta_5 x_2^2 + \delta_6 x_3^2$
  $\quad + \delta_7 x_1 x_2 + \delta_8 x_1 x_3 + \delta_9 x_2 x_3 + error$
  $H_0 : \delta_1 = \delta_2 = \cdots = \delta_9 = 0$
  - 検定統計量 $LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_9$
- White testのメリット&デメリット
  - 2乗項、交差項との関係も考慮し、BP検定より一般的
  - すべての2乗項・交差項を含むと、係数の数が膨大に（例：k=6のとき27係数）

8-3 不均一分散の検定

Alternative form of the White test

- u^2への$x_1, x_2, \ldots, x_k$の影響を検定(White 2)
  $\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$
  - u^2への影響をyとy^2の予測値を用いて間接的に検定
    - yとy^2の予測値が説明変数、2乗項、交差項を含むと仮定
  $H_0 : \delta_1 = \delta_2 = 0,\ LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_2$
- 例8.5：住宅価格の不均一分散検定(対数版)
  $R^2_{\hat{u}^2} = .0392, LM = 88(.0392) \approx 3.45, p\text{-}value_{LM} = .178$
  - 均一分散の帰無仮説を棄却できない

8-4 加重最小二乗法WLS

- Weighted least squares estimation
  - 8-4a 不均一分散が既知のケース

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + u_i$$

$$\boxed{Var(u_i|\mathbf{x}_i) = \sigma^2 h(\mathbf{x}_i),\ h(\mathbf{x}_i) = h_i > 0}$$

- 不均一分散の関数形は既知
- 変換

$$\Rightarrow \left[\frac{y_i}{\sqrt{h_i}}\right] = \beta_0\left[\frac{1}{\sqrt{h_i}}\right] + \beta_1\left[\frac{x_{i1}}{\sqrt{h_i}}\right] + \cdots + \beta_k\left[\frac{x_{ik}}{\sqrt{h_i}}\right] + \left[\frac{u_i}{\sqrt{h_i}}\right]$$

$$\Leftrightarrow\ y_i^* = \beta_0 x_{i0}^* + \beta_1 x_{i1}^* + \cdots + \beta_k x_{ik}^* + u_i^*$$

8-4 加重最小二乗法WLS

- 例：貯蓄と所得

$$sav_i = \beta_0 + \beta_1 inc_i + u_i,\ Var(u_i|inc_i) = \sigma^2 inc_i$$

$$\left[\frac{sav_i}{\sqrt{inc_i}}\right] = \beta_0\left[\frac{1}{\sqrt{inc_i}}\right] + \beta_1\left[\frac{inc_i}{\sqrt{inc_i}}\right] + u_i^*$$

- 変換後の回帰式には切片がないことに注意
- 変換後のモデルは均一分散

$$E(u_i^{*2}|\mathbf{x}_i) = E\left[\left(\frac{u_i}{\sqrt{h_i}}\right)^2 |\mathbf{x}_i\right] = \frac{E(u_i^2|\mathbf{x})}{h_i} = \frac{\sigma^2 h_i}{h_i} = \sigma^2$$

- 他のGauss-Markov仮定も同様に成立する場合、変換されたモデルに適用されるOLSは、最良線形不偏推定量を得る

8-4 加重最小二乗法WLS

- WLSについて
  - 加重残差平方和の最小化

$$\min \sum_{i=1}^{n}\left(\left[\frac{y_i}{\sqrt{h_i}}\right] - b_0\left[\frac{1}{\sqrt{h_i}}\right] - b_1\left[\frac{x_{i1}}{\sqrt{h_i}}\right] - \cdots - b_k\left[\frac{x_{ik}}{\sqrt{h_i}}\right]\right)^2$$

$$\Leftrightarrow\ \min \sum_{i=1}^{n}(\underbrace{y_i - b_0 - b_1 x_{i1} - \cdots - b_k x_{ik}}_{=u_i})^2/h_i$$

- 最適化問題において、PRFからの乖離が大きい観測値ほどその重要性を低くするように計算（=より大きいh_i）
- なぜWLSはOLSよりも効率的なのか？
  - 大きなばらつきを伴う観測値は、小さなばらつきを伴う観察値ほど有益でないため、より少なく加重
- WLSは、一般化最小二乗（GLS）の特殊ケース

8-4 加重最小二乗法WLS

- 例8.6：金融資産

純金融資産

不均一分散の関数形を想定

TABLE 8.1 Dependent Variable: *nettfa*

Independent Variables	(1) OLS	(2) WLS	(3) OLS	(4) WLS
inc	.821	.787	.771	.740
	(.104)	(.063)	(.100)	(.064)
$(age-25)^2$	—	—	.0251	.0175
			(.0043)	(.0019)
male	—	—	2.48	1.84
			(2.06)	(1.56)
e401k	—	—	6.89	5.19
			(2.29)	(1.70)
intercept	−10.57	−9.58	−20.98	−16.70
	(2.53)	(1.65)	(3.50)	(1.96)
Observations	2,017	2,017	2,017	2,017
R-squared	.0827	.0709	.1279	.1115

401K年金制度参加

WLS推定値の標準誤差はかなり小さい（より効率的になるという理論と整合的）

8-4 加重最小二乗法WLS

- 重要な特殊ケース
  - 観測値が都市/郡/州/国/会社ごとの平均値である場合は、その構成要素の数で加重が必要

$$\overline{contrib}_i = \beta_0 + \beta_1 \overline{earns}_i + \beta_2 \overline{age}_i + \beta_3 mrate_i + \bar{u}_i$$

企業iの年金制度への平均寄与　企業iの平均収入&年齢　企業の制度への寄与率　各企業での誤差の平均

$$\Rightarrow\ Var(\bar{u}_i) = Var\left(\frac{1}{m_i}\sum_{i=1}^{m_i} u_{i,e}\right) = \boxed{\sigma^2/m_i}$$

- 個人レベルで均一分散→グループ平均では不均一分散
- 企業規模m_iと等しいウェイトを使用したWLSで推定
- 個人レベルでの均質分散の仮定が正確でない場合、WLS推定後にHRSEを計算

8-4 加重最小二乗法WLS

- 8-4b 未知の不均一分散関数（feasible GLS）
  - 不均一分散の一般形を想定

$$\boxed{Var(u|\mathbf{x}) = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) = \sigma^2 h(\mathbf{x})}$$

- 誤差分散>0であるため、指数関数を使用

$$u^2 = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) \cdot v$$

- v：説明変数とは独立している部分の誤差

$$\Rightarrow\ \log(u^2) = \alpha_0 + \delta_1 x_1 + \cdots + \delta_k x_k + e$$

$$\log(\hat{u}^2) = \hat{\alpha}_0 + \hat{\delta}_1 x_1 + \cdots + \hat{\delta}_k x_k + error$$

$$\Rightarrow\ \hat{h}_i = \exp(\hat{\alpha}_0 + \hat{\delta}_1 x_1 + \cdots + \hat{\delta}_k x_k)$$

- 推定した不均一分散関数の逆数をWLSのウェイトとして使用
- Feasible GLSは一致推定量&OLSよりも漸近的効率性

8-4 加重最小二乗法WLS

◆例8.7 タバコの需要

- OLSによる推定

$$\widehat{cigs} = -\underset{(24.08)}{3.64} + \underset{(.728)}{.880}\log(income) - \underset{(5.773)}{.751}\log(cigpric)$$

喫煙本数/日　所得　煙草価格

$$-\underset{(.167)}{.501}\,educ - \underset{(.160)}{.771}\,age - \underset{(.0017)}{.0090}\,age^2 - \underset{(1.11)}{2.83}\,restaurn$$

禁煙レストランの有無

$$n = 807, R^2 = .0526, p\text{-}value_{Breusch-Pagan} = .000$$

  - 不均一分散の検定:「均一分散」の仮説を棄却

8-4 加重最小二乗法WLS

- FGLSによる推定

$$\widehat{cigs} = -\underset{(17.80)}{5.64} + \boxed{\underset{(.44)}{1.30}}\log(income) - \underset{(4.46)}{2.94}\log(cigpric)$$

$$-\underset{(.120)}{.463}\,educ + \underset{(.097)}{.482}\,age - \underset{(.0009)}{.0056}\,age^2 - \underset{(.80)}{3.46}\,restaurn$$

$$n = 807, R^2 = .1134$$

- 考察
  - 所得弾力性がFGLSでは統計的有意
  - FGLSは全係数の標準誤差がOLSより小さく、より正確な推定を示唆

8-4 加重最小二乗法WLS

◆8-4c 不均一分散関数の誤った特定化

- 不均一分散関数が誤って特定化されると…
  - MLR.1-4のもとでWLSは依然として一致推定量
    - しかし頑健な標準誤差HRSEを計算すべき
    - また必ずしもMLR.4が成立しているとは限らない

$$E(u_i|\mathbf{x}_i) = 0 \quad \Rightarrow \quad E\left(\left(u_i/\sqrt{h(\mathbf{x}_i)}\right)|\mathbf{x}_i\right) = 0$$

  - OLSとWLSの推定値が非常に異なる場合、通常、仮定(例えばMLR.4)が間違っていることを示唆
  - 強い不均一分散がある場合、効率性のために誤った特定化関数を使う方が良いケースが多い

8-5 線形確率モデル(再)

◆線形確率モデルでのWLS

$$P(y = 1|\mathbf{x}) = p(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

$$\Rightarrow \; Var(y|\mathbf{x}) = p(\mathbf{x})\,[1 - p(\mathbf{x})]$$

LPMでは不均一分散の関数形は既知

$$\Rightarrow \; \hat{h}_i = \hat{y}_i(1 - \hat{y}_i)$$

この逆数をWLSのウェイトとして使用

- 考察
  - LPM予測がゼロ以下または1より大きい場合、実行不可能
  - そのようなケースがまれである場合、数値を0.01 / 0.99に調整
  - それ以外の場合は、OLSで頑健な標準誤差HRSEを計算する方が良い

付論 不均一分散の対数

◆Under heteroskedasticity, we assume:

- $\mathrm{Var}[y_i|x_i] = \mathrm{E}[y_i - \mathrm{E}[y_i|x_i]\ |x_i]^2 = \sigma_i^2$ (a.1)
- $\mathrm{E}[y_i|x_i] = \beta_0 + \beta_1 x_i \equiv \mu_i$. (a.2)

◆By ***Taylor series expansion***, we can get

- $f(y_i) \approx f(\mu_i) + f'(\mu_i)(y_i - \mu_i)$. (a.3)

◆Here, since $f(\mu_i)$ can be treated as nonrandom variable under conditional $\boldsymbol{x}$,

- $\mathrm{E}[f(y_i)|x_i] = f(\mu_i)$. (a.4)

Cont. 不均一分散の対数

$$\begin{aligned}
\mathrm{Var}[f(y_i)|x_i] &= \mathrm{E}[f(y_i) - \mathrm{E}[f(y_i)|x_i]\ |x_i]^2 && \leftarrow(a.1)\\
&= \mathrm{E}[f(y_i) - f(\mu_i)\ |x_i]^2 && \leftarrow(a.4)\\
&= \mathrm{E}[f'(\mu_i)(y_i - \mu_i)\ |x_i]^2 && \leftarrow(a.3)\\
&= \{f'(\mu_i)^2\}\mathrm{E}[(y_i - \mu_i)\ |x_i]^2 && \\
&= \{f'(\mu_i)\}^2\sigma_i^2 && \leftarrow(a.1)
\end{aligned}$$

If we assume $f(y_i) = \ln(y_i)$ & $\sigma_i^2 = \sigma^2\mathrm{E}[y_i|x_i]^2$,

$$\begin{aligned}
\{f'(\mu_i)\}^2\sigma_i^2 &= \{1/\mu_i\}^2\sigma^2\mu_i^2 && \leftarrow(a.2)\\
&= \sigma^2 && \text{(Homoskedasticity)}
\end{aligned}$$