8　標準的仮定の意味と不均一分散

8.0　標準的仮定と仮定 1 と 2 について

・標準的仮定　1　説明変数 X は非確率変数
2　n→∞のとき X の偏差 2 乗和→∞
3　誤差項の期待値は 0
4　誤差項の分散は σ^2
5　誤差項の（異時点）共分散は 0
6　誤差項は正規分布に従う

・基本モデル　$Y_i = \alpha + \beta X_i + u_i$　(8.1)

・仮定 1　「説明変数 X は非確率変数」　→　不成立のケースは 10 章で扱う

・仮定 2　「n→∞のとき、$\sum_{i=1}^{n}(X_i - \overline{X})^2 \to \infty$」　→　通常成立する

8.1　仮定 3 について

・仮定 3　「誤差項の期待値 $E(u_i) = 0$」

$E(u_i) = \gamma \neq 0$ (8.2)のもとで $\alpha' = \alpha + \gamma$ 、$u' = u - \gamma$　(8.4)として定義し直すと(8.1)は

$Y_i = \alpha' + \beta X_i + u_i'$　(8.4)

ここで(8.4)について仮定 3 を考えると、以下のようになる。

$E(u_i') = E(u_i - \gamma) = E(u_i) - \gamma = \gamma - \gamma = 0$　(8.6)

したがって、切片のあるモデルは結果的に仮定 3 を成立させる。

・不偏性 $\hat{\beta} = \dfrac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sum(X_i - \overline{X})^2} = \beta + \dfrac{\sum(X_i - \overline{X})u_i}{\sum(X_i - \overline{X})^2} = \beta + \sum w_i u_i$

仮定 1 のとき、$E(\hat{\beta}) = \beta + \sum w_i E(u_i)$　⇔「仮定 1 不成立」 $\beta + E(\sum w_i u_i)$

仮定 3 のとき、$E(\hat{\beta}) = \beta + \sum w_i \cdot 0 = \beta$

→すなわち仮定 1 と 3 が成立すれば OLS 推定量は不偏性を持つ

（仮定 2,4,5,6 は OLS 推定量の不偏性に影響しない）

8.2　仮定 4&5 と β の分散

・仮定 4　$Var(u_i) = E(u_i - E(u_i))^2 = E(u_i - 0)^2 = E(u_i^2) = \sigma^2$

・仮定 5　$Cov(u_i, u_j) = E((u_i - 0)(u_j - 0)) = E(u_i u_j) = 0$

・ β の分散　ここで仮定 4、仮定 5 の OLS 推定量分散に対する影響を考える。

（→OLS 推定量の分散は α 、β の *t* 値などに直接影響を及ぼすため極めて重要）

$$Var(\hat{\beta}) = E\left(\hat{\beta} - E(\hat{\beta})\right)^2 = E\left(\hat{\beta} - \beta\right)^2 = E\left(\sum w_i u_i\right)^2 \qquad (8.7')$$

$$= E\left(\sum w_i^2 u_i^2 + \sum_{i \neq j} w_i w_j u_i u_j\right) = \sum w_i^2 E(u_i^2) + \sum_{i \neq j} w_i w_j E(u_i u_j) \qquad (8.8')$$

$$= \sum w_i^2 \sigma^2 + \sum_{i \neq j} w_i w_j \cdot 0 = \sigma^2 \sum w_i^2 \qquad \leftarrow \text{仮定 4\&5 より}$$

$$= \sigma^2 \sum \left(\frac{X_i - \overline{X}}{\sum (X_i - \overline{X})^2} \right)^2 = \frac{\sigma^2 \sum (X_i - \overline{X})^2}{\left\{\sum (X_i - \overline{X})^2\right\}^2} = \frac{\sigma^2}{\sum (X_i - \overline{X})^2}$$

・仮定 4&5 が不成立のケース

✓　仮定 4 が不成立のケース（$Var(u_i) = \sigma_i^2$ 、ただし仮定 5 は成立）

$$Var(\hat{\beta}^*) = \sum w_i^2 E(u_i^2) = \sum w_i^2 \sigma_i^2 = \frac{\sum (X_i - \overline{X})^2 \sigma_i^2}{\left\{\sum (X_i - \overline{X})^2\right\}^2}$$

$$\neq \frac{\sigma^2}{\sum (X_i - \overline{X})^2}$$

✓　仮定 5 が不成立のケース（$Cov(u_i, u_j) = \sigma_{ij}^2 \neq 0$ 、ただし仮定 4 は成立）

$$Var(\hat{\beta}^*) = \sum w_i^2 E(u_i^2) + \sum_{i \neq j} w_i w_j E(u_i u_j) = \sigma^2 \sum w_i^2 + \sum_{i \neq j} w_i w_j \sigma_{ij}$$

$$\neq \sigma^2 \sum w_i^2$$

・問題点 ①t 値などの統計的推論が成立しなくなる

$$t = \frac{\hat{\beta} - \beta}{s_{\hat{\beta}}} \quad \leftarrow \quad \left(s_{\hat{\beta}} = \sqrt{s_{\hat{\beta}}^2} \quad \leftarrow \quad E(s_{\hat{\beta}}^2) = \sigma_{\hat{\beta}}^2 = Var(\hat{\beta}) \right)$$

$*Var(\hat{\beta}^*) \neq Var(\hat{\beta})$ より、「誤った分散」をもとに「誤った t 値」を算出する

②OLSE は不偏推定量の中で最小分散ではない（→最良線型不偏推定量ではない）

8.3　撹乱項の不均一分散

・不均一分散　　誤差（残差）分散が（説明変数 X とともに）変化　→　$Var(u_i) = E(u_i^2) = \sigma_i^2$

（8.16）→クロスセクション分析で多く、時系列ではあまり見られない
例）所得（X）が高い世帯ほど消費（Y）のばらつきが大きい

・対処法　　誤差分散 σ_i^2 が既知であれば、事前に $1/\sigma_i$ でウェイトを付ける（→WLS）

$$\frac{Y_i}{\sigma_i}=\alpha\frac{1}{\sigma_i}+\beta\frac{X_i}{\sigma_i}+\frac{u_i}{\sigma_i} \quad\rightarrow\quad Var\left(\frac{u_i}{\sigma_i}\right)=E\left(\frac{u_i^2}{\sigma_i^2}\right)=\frac{\sigma_i^2}{\sigma_i^2}=1$$

実際的方法・・$\sigma_i=cZ_i$ （8.19）と仮定（ここでZは多くの場合、Xの関数）

$$\frac{Y_i}{Z_i}=\alpha\frac{1}{Z_i}+\beta\frac{X_i}{Z_i}+\frac{u_i}{Z_i} \quad\rightarrow\quad Var\left(\frac{u_i}{Z_i}\right)=E\left(\frac{u_i^2}{Z_i^2}\right)=\frac{\sigma_i^2}{Z_i^2}=c^2$$

*決定係数は本来のモデルと異なる数値を算出することに注意

・WLS 有効性　$\sigma_i=cZ_i$ （8.19）と仮定

$$Var(\hat\beta_{WLS})=\frac{\sum(X_i^*-\bar X^*)^2c^2}{\left\{\sum(X_i^*-\bar X^*)^2\right\}^2}=\frac{c^2}{\sum(X_i^*-\bar X^*)^2} \qquad \leftarrow X_i^*=\frac{X_i}{Z_i}$$

$$Var(\hat\beta_{OLS})=\frac{\sum(X_i-\bar X)^2\sigma_i^2}{\left\{\sum(X_i-\bar X)^2\right\}^2}=\frac{c^2\sum(X_i-\bar X)^2Z_i^2}{\left\{\sum(X_i-\bar X)^2\right\}^2}$$

（有効性証明）　ここで$a_i=(X_i-\bar X)Z_i$、$b_i=(X_i-\bar X)/Z_i$ とすると、次の関係が成立する。

$$\frac{Var(\hat\beta_{WLS})}{Var(\hat\beta_{OLS})}=\frac{c^2/\sum b_i^2}{c^2\sum a_i^2/(\sum a_ib_i)^2}=\frac{\left(\sum a_ib_i\right)^2}{\sum a_i^2\sum b_i^2}\le 1$$

* Cauchy-Schwartz の不等式（[相乗平均≦相加平均]の関係の応用型）より

→したがって<u>適切な加重変数を</u>用いたWLSEはOLSEより有効な推定量である。

8.4　不均一分散モデルの推定と検定

・一般的推定（Feasible Generalized Least Squares; FGLS）

基本モデル　$$Y_i=\beta_1+\beta_2X_{2i}+\cdots+\beta_kX_{ki}+u_i \tag{8.30}$$

誤差分散　$$Var(u_i)=\sigma_i^2=\exp(\alpha_1+\alpha_2Z_{2i}+\cdots+\alpha_pZ_{pi}) \tag{8.31}$$

$$\ln(\hat u_i^2)=\alpha_1+\alpha_2Z_{2i}+\cdots+\alpha_pZ_{pi}+w_i \tag{8.32}$$

①　ここで（8.30）を回帰して残差$\hat u_i$を求める。

②　$\hat u_i^2$をσ_i^2の代理変数として（8.32）を回帰し、α_iを求める。

((8.32)の右辺説明変数を$\hat Y_i$、$\hat Y_i^2$とするやり方もある。)

③　$\hat\alpha_i$を用いて（8.31）からσ_i^2の理論値$\hat\sigma_i^2$を求める。

④ $\hat{\sigma}_i$ をウェイトとして（8.30）を推定

・一般的検定

不均一分散の検定とは、下記の帰無仮説を設定し、検定するのが一般的である。

H_0: $Var(u \mid X_1, X_2, \ldots, X_k) = E(u^2 \mid X_1, X_2, \ldots, X_k) = E(u^2) = \sigma^2$ （仮定 4）

✓ Breusch-Pagan's LM test

ここで u^2 と X_j が線形の関係にあると仮定すると、次の帰無仮説を設定できる。

$u^2 = a_1 + a_2X_2 + \ldots + a_k X_k + v$ (8.32a)'

H_0: $a_2 = \ldots = a_k = 0$ H_1：H_0 でない

これを F-test あるいは LM-test にて検定。

*LM-test・・(8.32a)'の回帰結果 R^2 を用いて nR^2 を計算（LM〜$\chi^2(k-1)$）

✓ White test ($k = 3$ のケース)

$u^2 = a_1 + a_2X_2 + a_3 X_3 + a_4X_2^2 + a_5 X_3^2 + a_6X_2X_3 + v$ (8.32b)'

$u^2 = a_1 + a_2(est.Y) + a_3 (est.Y^2) + v$ (8.32c)'

(8.32b)'、もしくは(8.32c)'を用いて、BP-test と同じ手順で検定。

（参考）不均一分散の対処法

・対数化 $Y_i = \alpha + \beta X_i + u_i \quad \rightarrow \quad \ln Y_i = \alpha' + \beta' \ln X_i + u'_i$

・頑健標準偏差 (White) Heteroskedasticity -Robust Standard Errors

（単回帰）
$$Est.Var(\hat{\beta}) = \frac{\sum (X_i - \bar{X})^2 \sigma_i^2}{\left\{\sum (X_i - \bar{X})^2\right\}^2} \approx \frac{\sum (X_i - \bar{X})^2 \hat{u}_i^2}{\left\{\sum (X_i - \bar{X})^2\right\}^2}$$

（重回帰）
$$Est.Var(\hat{\beta}_j) = \frac{\sum \hat{v}_{ij}^2 \sigma_i^2}{\left\{\sum \hat{v}_{ij}{}^2\right\}^2} \approx \frac{\sum \hat{v}_{ij}^2 \hat{u}_i^2}{\left\{\sum \hat{v}_{ij}{}^2\right\}^2}$$

*ここで $\hat{v}_{ij}$ は X_jを他の説明変数に回帰したときの残差

$\rightarrow Est.s.e.\left(\hat{\beta}_j\right) = \sqrt{Est.Var(\hat{\beta}_j)}$。ただしこの統計量は大標本のときのみ有効

階層化されたデータ

所得階層別消費支出額　　（円、人）

標本番号	所得階層（万円）			消費支出費 C	可処分所得 Y	世帯人員 P	集計世帯数 n
1	0	～	100	116098	124280	2.69	47
2	100	～	150	128390	124804	2.87	200
3	150	～	200	147125	148657	3.04	445
4	200	～	250	163339	178820	3.26	925
5	250	～	300	182814	199268	3.46	1633
6	300	～	350	198436	220161	3.53	2488
7	350	～	400	217572	242309	3.71	2947
8	400	～	450	233169	263882	3.82	3250
9	450	～	500	248818	284857	3.90	3164
10	500	～	550	263260	307069	3.97	2893
11	550	～	600	279973	323893	3.95	2401
12	600	～	650	294536	345861	3.96	2066
13	650	～	700	316420	368988	4.04	1642
14	700	～	750	331743	390789	4.03	1473
15	750	～	800	348293	403944	4.13	1115
16	800	～	900	367518	436126	4.16	1714
17	900	～	1000	393330	486895	4.18	999
18	1000	～		462660	569612	4.28	1617

（出所）総務庁統計局『全国消費実態調査報告』、勤労者世帯（昭和59年）
*消費支出額および可処分所得は各階層における世帯あたりの平均値

いま次のようなモデルを考える。

$$\text{消費支出}（C_i）=\beta_0+\beta_1\text{可処分所得}（Y_i）+\beta_2\text{世帯人員}（P_i）+u_i \qquad \text{(a8.1)}$$

ここでは各世帯についてこのモデルが成り立っているとする。つまり i は個別世帯を示す。またこのとき OLS の標準的仮定はすべて成り立っているとする。

ところが実際に入手可能なデータは上表のように集計（平均）データであることが多く、変数は各階層における世帯あたりの平均値となっている。つまり k 階層に属する個別世帯のデータを x_{ki}（$i=1,2,\cdots,n_k$）、その階層に属する世帯数を n_k とすれば、k 階層の変数（世帯あたり平均値）は、

$$\bar{x}_k=\frac{1}{n_k}\sum_{i=1}^{n_k}x_{ki}$$

となる。当然、(a8.1）の誤差項も

$$\bar{u}_k=\frac{1}{n_k}\sum_{i=1}^{n_k}u_{ki}$$

となる。(a8.1）の誤差項の標準的仮定から $E(u_i^2)=\sigma^2$ を考慮すると、

$$E(\bar{u}_k^2)=E\left[\left(\frac{1}{n_k}\sum u_{ki}\right)^2\right]=E\left[\left(\frac{u_{k1}}{n_k}\right)^2+\left(\frac{u_{k2}}{n_k}\right)^2+\cdots+\left(\frac{u_{knk}}{n_k}\right)^2\right] \qquad \leftarrow E\left(u_iu_j\right)=0$$

$$=\frac{E(u_{k1}^2)}{n_k^2}+\frac{E(u_{k2}^2)}{n_k^2}+\cdots+\frac{E(u_{knk}^2)}{n_k^2}=\frac{n_k\sigma^2}{n_k^2}=\frac{1}{n_k}\sigma^2 \qquad \leftarrow E(u_i^2)=\sigma^2$$

このように誤差分散は階層に属する世帯数に依存する。よって（a8.1）を集計データによって推定する場合には、不均一分散に注意する必要がある。