第 5 回　単回帰モデル③

[5] 最小二乗推定量の分散

最小二乗推定量の分散を求めるため，(S1)～(S3)に加えて次の二つの仮定をおく。

(S4) 誤差項の均一分散： $\mathrm{Var}(u_i) = \sigma^2$ （一定）

(S5) 誤差項の独立性　： $i \neq j$ について u_i と u_j は確率的に独立である。

（例 1）貯蓄関数と不均一分散（sav：貯蓄，inc：所得）

$$\mathrm{sav}_i = \beta_0 + \beta_1 \mathrm{inc}_i + u_i \quad (i = 1, \ldots, n)$$

（例 2）時系列データの回帰分析と誤差項の「系列相関」

誤差項の系列相関：異なる時点 t と s について $\mathrm{Corr}(u_t, u_s) \neq 0$ となること。

正の系列相関： $\mathrm{Corr}(u_t, u_{t+1}) > 0$

※ 時系列データの回帰分析で誤差項に系列相関が生じる理由

1）慣性（持続的，循環的な動き） ⇒ 正の系列相関

2）反動（前期と反対方向への動き） ⇒ 負の系列相関

3）除外変数（モデルに含まれない変数の自己相関） ⇒ 正または負の系列相関

4）回帰モデルの関数形の誤り ⇒ 正または負の系列相関

（復習事項）

・ $\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} = \sum_{i=1}^{n}(x_{i1}+x_{i2}+\cdots+x_{in}) = (x_{11}+x_{12}+\cdots+x_{1n})+\cdots+(x_{n1}+x_{n2}+\cdots+x_{nn})$

・ $\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j = x_1(x_1+x_2+\cdots+x_n)+x_2(x_1+x_2+\cdots+x_n)+\cdots+x_n(x_1+x_2+\cdots+x_n)$

$$= x_1^2+x_2^2+\cdots+x_n^2+(x_1x_2+x_1x_3+\cdots+x_1x_n)+\cdots+(x_nx_1+x_nx_2+\cdots+x_nx_{n-1})$$

$$= \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \quad (j \neq i)$$

・仮定 S4： $\mathrm{Var}(u_i) = \sigma^2$ （一定） ⇒ $E(u_i^2) = \sigma^2$

・仮定 S5： u_i と u_j （$j \neq i$）は確率的に独立 ⇒ $E(u_i u_j) = 0$

定理 2.2（最小二乗推定量の分散）

仮定(S1)～(S5)が成り立てば， $\mathrm{Var}(\hat{\alpha}) = \sigma^2\left(\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right)$， $\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{S_{xx}}$

（証明）

(3) 最小二乗推定量の分散の推定（誤差分散 σ^2 の推定）

・σ^2 の不偏推定量： $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2} = \dfrac{RSS}{n-2}$

・$\hat{\alpha}$ と $\hat{\beta}$ の分散の推定量と「標準誤差」： $\hat{\sigma}^2$ を定理 2.2 に代入すれば,

$$\hat{\alpha} \text{ の標準誤差 } \quad se(\hat{\alpha}) = \sqrt{\hat{Var}(\hat{\alpha})} = \sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}$$

$$\hat{\beta} \text{ の標準誤差 } \quad se(\hat{\beta}) = \sqrt{\hat{Var}(\hat{\beta})} = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

（例）前回の数値例について， $\hat{\sigma}^2$， $se(\hat{\alpha})$， $se(\hat{\beta})$ を求める。

[5] ガウス＝マルコフの定理

> **定理 2.3（ガウス＝マルコフの定理）**
> 仮定(S1)～(S5)の下で， α と β の最小二乗推定量 $\hat{\alpha}$ と $\hat{\beta}$ は最良線形不偏推定量（BLUE）である。

<u>B</u>est（最良）　：（線形不偏推定量の中で）分散が最小である。

<u>L</u>inear（線形）　： $\tilde{\beta} = \sum_{i=1}^{n} w_i y_i$ とかけるとき， $\tilde{\beta}$ は線形推定量という。

<u>U</u>nbiased（不偏）： $E(\hat{\beta}_j) = \beta_j$ が成り立つ。

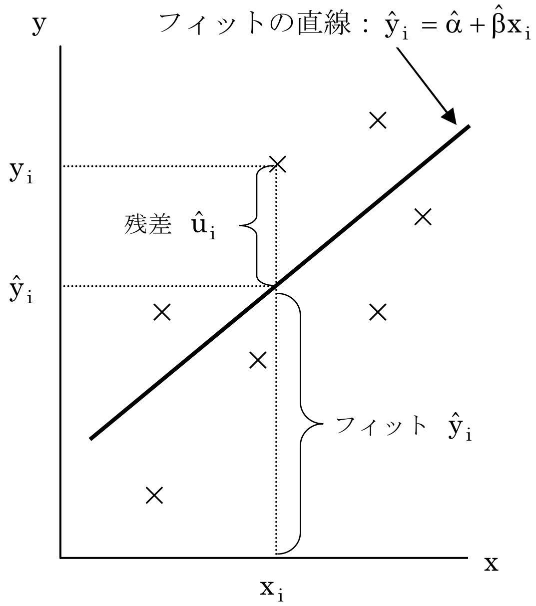
<u>E</u>stimator（推定量）

[6] 単回帰モデルの復習

・単回帰モデル：説明変数 x が従属変数 y に与える影響は関係： $y = \alpha + \beta x + u$ で表せる。

⇒ パラメータ α と β を知れば，x が y に与える影響を測ることができる。

・観察する y_i，フィット $\hat{y}_i$，残差 $\hat{u}_i$ の関係



パラメータ α と β の推定量を $\hat{\alpha}$，$\hat{\beta}$ とするとき，

フィット：$\hat{y}_i = \hat{\alpha} + \hat{\beta}x_i$

残差：$\hat{u}_i = y_i - \hat{y}_i$

・最小二乗法

$\mathrm{RSS} = \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta}x_i)^2$ を $\hat{\alpha}$ と $\hat{\beta}$ について最小化して求められる α と β の推定量

⇒ 最小化の条件 1：

最小化の条件 2：

$$\Rightarrow \quad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \quad \hat{\beta} = \frac{S_{xy}}{S_{xx}}$$

※ $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$，$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$

・最小二乗推定量が望ましい性質をもつための仮定

(S1) $E(u_i) = 0$

(S2) x_i は確率変数ではない。

(S3) $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 > 0$

（S1）〜（S3）：不偏性に必要

(S4) 誤差項の均一分散： $\mathrm{Var}(u_i) = \sigma^2$ （一定）

(S5) 誤差項の独立性 ： $i \neq j$ について u_i と u_j は確率的に独立である。

（S1）〜（S5）：ガウス＝マルコフの定理に必要