第12章　補足：最尤法の前に追加

12.1　誤差項：非正規分布のケース（P.222 の 9.5 節の前に追加）

(* 復習) 中心極限定理：

確率変数 X_1，X_2，$\cdots$，X_n は互いに独立とする。$\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ とする。

$n \longrightarrow \infty$ のとき，$\dfrac{\overline{X} - \mathrm{E}(\overline{X})}{\sqrt{\mathrm{V}(\overline{X})}} \longrightarrow N(0,1)$ となる。ただし，$n\mathrm{V}(\overline{X}) < \infty$ とする。

(a) すべての i について $\mathrm{E}(X_i) = \mu$，$\mathrm{V}(X_i) = \sigma^2$ とする。

$n \longrightarrow \infty$ のとき，$\dfrac{\overline{X} - \mathrm{E}(\overline{X})}{\sqrt{\mathrm{V}(\overline{X})}} = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \longrightarrow N(0,1)$ となる。

$\mathrm{E}(\overline{X}) = \mu$，$\mathrm{V}(\overline{X}) = \dfrac{\sigma^2}{n}$ に注意。

(b) $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \longrightarrow N(0,1)$ は $\sqrt{n}(\overline{X}-\mu) \longrightarrow N(0,\sigma^2)$ と同じ意味である

なぜなら，$\mathrm{V}\Big(\sqrt{n}(\overline{X}-\mu)\Big) = (\sqrt{n})^2\mathrm{V}(\overline{X}-\mu) = n\mathrm{V}(\overline{X}) = \sigma^2$ となる。

(c) $\mathrm{E}(X_i)=\mu$，$\mathrm{V}(X_i)=\sigma_i^2$（不均一分散），$\sigma^2 = \lim_{n\to\infty}\dfrac{1}{n}\sum_{i=1}^{n}\sigma_i^2 < \infty$ のとき，

$\sqrt{n}(\overline{X}-\mu) \longrightarrow N(0,\sigma^2)$ となる。

● 中心極限定理を回帰分析に当てはめる。

単回帰：

$$Y_i = \alpha + \beta X_i + u_i$$

を考える。

- 誤差項 u_1，u_2，$\cdots$，u_n は互いに独立で，それぞれ平均ゼロ・分散 σ^2 とする（正規分布を仮定しない）。
- $n \longrightarrow \infty$ のとき，$\dfrac{1}{n}\sum_i (X_i-\overline{X})^2 \longrightarrow M < \infty$（一定の値）とする（すな

わち，$\lim_{n\to\infty}\dfrac{1}{n}\sum_i (X_i - \overline{X})^2 = M < \infty$)。

この場合，最小二乗推定量はどうなるか？
β の最小二乗推定量 $\hat\beta$ は次のように書き換えられる。

$$\begin{aligned}\hat\beta &= \frac{\sum_i (X_i-\overline{X})(Y_i-\overline{Y})}{\sum_i (X_i-\overline{X})^2} = \beta + \frac{\sum_i (X_i-\overline{X})u_i}{\sum_j (X_j-\overline{X})^2} \\ &= \beta + \frac{\frac{1}{n}\sum_i (X_i-\overline{X})u_i}{\frac{1}{n}\sum_j (X_j-\overline{X})^2}\end{aligned}$$

$\hat\beta$ の分布を求めるために，まず，分子の $\dfrac{1}{n}\sum_i (X_i-\overline{X})u_i$ の分布を求める。
$\dfrac{1}{n}\sum_i (X_i-\overline{X})u_i$ が「(* 復習) 中心極限定理」の $\overline{X} = \dfrac{1}{n}\sum_i X_i$ に対応する。
すなわち，$(X_i-\overline{X})u_i$ が「(* 復習) 中心極限定理」の X_i に対応する。
「(* 復習) 中心極限定理」の (c) に当てはめる。

$(X_i-\overline{X})u_i$ の平均・分散は下記の通りとなる。

- $\mathrm{E}((X_i - \overline{X})u_i) = (X_i - \overline{X})\mathrm{E}(u_i) = 0$
- $\mathrm{V}((X_i - \overline{X})u_i) = (X_i - \overline{X})^2\mathrm{V}(u_i) = \sigma^2(X_i - \overline{X})^2$

$\sigma^2(X_i - \overline{X})^2$ は「(* 復習) 中心極限定理」(c) の σ_i^2 に対応する。

まずは,「(* 復習) 中心極限定理」(c) の $\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 < \infty$ が成り立つかどうかをチェックする。

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathrm{V}((X_i - \overline{X})u_i) = \lim_{n\to\infty}\sigma^2\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sigma^2 M < \infty$$

よって，分子に $\sqrt{n}$ を掛けて,

$$\sqrt{n}\frac{1}{n}\sum_{i}(X_i - \overline{X})u_i \longrightarrow N(0,\ \sigma^2 M)$$

となる。

分母に関しては，仮定から $\displaystyle\frac{1}{n}\sum_{j}(X_j - \overline{X})^2 \longrightarrow M$ となる。

$\hat{\beta}-\beta$ に $\sqrt{n}$ を掛けて，次のように書き直すことができる。

$$\sqrt{n}(\hat{\beta}-\beta)=\frac{\sqrt{n}\frac{1}{n}\sum_i(X_i-\overline{X})u_i}{\frac{1}{n}\sum_j(X_j-\overline{X})^2}$$

分子の分散は $\sigma^2 M$ に収束し，分母は M に収束する。

よって，分子分母合わせて，分散は $\left(\frac{1}{M}\right)^2\times\sigma^2 M=\frac{\sigma^2}{M}$ に収束する。

したがって，

$$\sqrt{n}(\hat{\beta}-\beta)=\frac{\sqrt{n}\frac{1}{n}\sum_i(X_i-\overline{X})u_i}{\frac{1}{n}\sum_j(X_j-\overline{X})^2}\longrightarrow N\left(0,\frac{\sigma^2}{M}\right)$$

となる。

すなわち，n が大きいとき，近似的に，

$$\hat{\beta}\sim N\left(\beta,\frac{\sigma^2}{\sum(X_i-\overline{X})^2}\right)$$

となる。

また，σ^2 をその推定量 s^2 に置き換えても，n が大きいとき，近似的に，

$$\hat{\beta} \sim N\Big(\beta,\ \frac{s^2}{\sum(X_i - \overline{X})^2}\Big)$$

となる（実践では，これが用いられる）。

結論：　誤差項 u_i の分布を仮定しなくても，データ数 n が大きければ，$\hat{\beta}$ の分布は正規分布で近似できる。

重回帰 $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i$ でも同様で，u_1，u_2，$\cdots$，u_n の分布を仮定しなくても（ただし，互いに独立で，平均ゼロ・分散 σ^2 の仮定は必要），n が大きいとき，$\hat{\beta}_j$ は近似的に下記のような正規分布を用いることができる。

$$\hat{\beta}_j \sim N(\beta_j,\ s^2 a_{jj})$$

$s\sqrt{a_{jj}}$ は $\hat{\beta}_j$ の標準誤差を表す。

a_{jj} については，講義ノート P.165 参照のこと。

12.2　説明変数：確率変数のケース（P.222 の 9.5 節に追加）

本節では，簡単化のために，単回帰 $Y_i = \alpha + \beta X_i + u_i$ を考える。

(1) X_i は非確率変数 $\longrightarrow$ 今までの最小二乗推定量

(2) X_i は確率変数

(2a) X_i と u_i は相関なし $\longrightarrow$ $\mathrm{Cov}(X_i, u_i) = 0$（12.2.1 節）

(2b) X_i は u_i は相関あり $\longrightarrow$ $\mathrm{Cov}(X_i, u_i) \neq 0$（12.2.2 節）

12.2.1　説明変数と誤差項に相関がない場合

最小二乗法による推定量：

$$\hat{\beta} = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sum(X_i - \overline{X})^2} = \beta + \sum \omega_i u_i$$

ただし，$\omega_i = \dfrac{X_i - \overline{X}}{\sum(X_j - \overline{X})^2}$ とする。

　$X = (X_1,\ X_2,\ \cdots,\ X_n)$ とする。

(* 復習) 2 つの確率変数（$X,\ Y$）の独立について：

X と Y の同時密度関数 $f_{xy}(x, y)$

X の周辺密度関数 $f_x(x)$

Y の周辺密度関数 $f_y(y)$

Y を与えたもとで X の条件付き密度関数 $f_{x|y}(x|y)$

- $f_{xy}(x, y) = f_{x|y}(x|y)f_y(y)$ は必ず成り立つ。
- $f_{xy}(x, y) = f_x(x)f_y(y) \quad \Longleftrightarrow \quad X$ と Y は独立
- $f_{x|y}(x|y) = f_x(x) \quad \Longleftrightarrow \quad X$ と Y は独立

(* 復習) 2 つの確率変数（X, Y）の独立について（その 2）：

X と Y が独立のとき,

- $\mathrm{Cov}(X, Y) = 0$
- $\mathrm{E}(X|Y) = \mathrm{E}(X)$, $\mathrm{V}(X|Y) = \mathrm{V}(X)$

となる。

条件付き期待値を取ると,

$$\mathrm{E}(\hat{\beta}|X) = \mathrm{E}(\beta + \sum_i \omega_i u_i|X) = \beta + \sum_i \omega_i \mathrm{E}(u_i|X) = \beta$$

となり，$\hat{\beta}$ は β の不偏推定量となる。

ω_i は X_1, X_2, $\cdots$, X_n の関数となっている。

$X = (X_1,\ X_2,\ \cdots,\ X_n)$ と u_i に相関がない場合，すなわち，$j = 1, 2, \cdots, n$ について $\mathrm{Cov}(X_j, u_i) = 0$ の場合，$\mathrm{E}(u_i|X) = \mathrm{E}(u_i) = 0$ となる。

条件付き分散については,

$$\mathrm{V}(\hat{\beta}|X) = \mathrm{V}(\beta + \sum_i \omega_i u_i|X) = \mathrm{V}(\sum_i \omega_i u_i|X) = \sum_i \omega_i^2 \mathrm{V}(u_i|X) = \sigma^2 \sum_i \omega_i^2 = \frac{\sigma^2}{\sum_i (X_i - \overline{X})^2}$$

となる。

すなわち，説明変数が確率変数であっても，誤差項と相関がなければ，何も変更せずに，最小二乗法を適用することができる。

12.2.2　説明変数と誤差項に相関がある場合

X と u_i に相関がある場合，$\mathrm{E}(u_i|X) \neq \mathrm{E}(u_i) = 0$ となるので,

$$\mathrm{E}(\hat{\beta}|X) = \mathrm{E}(\beta + \sum_i \omega_i u_i|X) = \beta + \sum_i \omega_i \mathrm{E}(u_i|X) \neq \beta$$

となる。

したがって，$\hat{\beta}$ は β の不偏推定量とはならない。

$\hat{\beta}$ は β の一致推定量かどうか？

$$\hat{\beta} = \beta + \sum_i \omega_i u_i = \beta + \frac{\sum_i (X_i - \overline{X}) u_i}{\sum_j (X_j - \overline{X})^2}$$

$$= \beta + \frac{\frac{1}{n}\sum_i (X_i - \overline{X}) u_i}{\frac{1}{n}\sum_j (X_j - \overline{X})^2} \longrightarrow \beta + \frac{M_{xu}}{M_{xx}} \neq \beta$$

ただし，$n \longrightarrow \infty$ のとき，$\frac{1}{n}\sum_i (X_i - \overline{X}) u_i \longrightarrow M_{xu} \neq 0$，$\frac{1}{n}\sum_i (X_i - \overline{X})^2 \longrightarrow M_{xx}$ とする。$n \longrightarrow \infty$ のときは，分子・分母は別々に計算することができる（証明略）。

M_{xu} は，$n \longrightarrow \infty$ のとき，X_i と u_i の共分散に相当する。M_{xx} は，$n \longrightarrow \infty$ のとき，X_i の分散に相当する。

以上から，$\hat{\beta}$ は β の不偏推定量でも一致推定量でもない。

$\hat{\alpha}$ も同様に不偏推定量でも一致推定量でもない。

なぜなら，$\hat{\alpha}$ は，

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X} = \alpha + \sum_i \lambda_i u_i$$