Ⅱ　回帰分析（Regression Analysis）

1 回帰分析とは

理論モデル(因果関係の明示)→数学モデル(特定化)→確率モデル(撹乱項の導入)

理論 $Y_i = F(X_i)$ → 線形モデル $Y_i = \alpha + \beta X_i$

→ 線形回帰モデル(Linear Regression Model) $Y_i = \alpha + \beta X_i + \varepsilon_i$

(1) 原因 X_i:説明変数(独立変数) → 結果 Y_i:被説明変数(従属変数)

回帰：Galton(1822-1911) 親の身長→子の身長

(2)確率モデル(確率的かく乱項 ε)の必要性

①理論の不完全性(Xのリストアップが不完全) ②線形近似の不完全性 ③測定誤差

$\varepsilon_i = \sum_{j=1}^{\infty} \varepsilon_i^j$ →Central Limit Theorem→$N(0, \sigma^2)$ かく乱項 ε は正規分布に従う

(3)線形モデルの一般性（p-40）

1)両対数(double-logarithmic)[$\ln Y = \alpha + \beta \ln X$]:弾力性(elasticity)一定

2)片対数(semi-logarithmic)[$Y = \alpha + \beta \ln X$]:X の最大値(最小値)が存在する

3)片対数[$\ln Y = \alpha + \beta X$]: Y の最大値(最小値)が存在する

4)逆数(reciprocal)[$Y = \alpha + \beta / X$]:Y の最大値(最小値)$= \alpha$ が存在する,philips curve

5)逆数[$Y = \beta / (X - \alpha)$]:X の最大値(最小値)$= \alpha$ が存在する,$1/Y = -\alpha/\beta + 1/\beta X$

6)高次(kth-order power)[$Y = \alpha + \beta_1 X + \ldots + \beta_k X^k$]:変曲点の存在,wage curve, cost curve

7)ロジット(logit)[$Y = \dfrac{e^{\alpha+\beta X}}{1+e^{\alpha+\beta X}}$] → logit 変換[$y = \log(Y/(1-Y)) = \alpha + \beta X$]

2. 点推定(point estimation)の方法

標本(X,Y)から母集団（$Y = \alpha + \beta X + \varepsilon$ ）を推測する方法： ①最小２乗法　②最尤法

2.1 最小２乗法(Least Squares Method) テキスト p-14

残差２乗和を最小化 $\min_{\alpha,\beta} S(\alpha,\beta) = \sum \left(Y_i - (\alpha + \beta X_i)\right)^2 \to \partial S/\partial\alpha = \partial S/\partial\beta = 0$

$$S_\alpha = \sum 2(Y_i - (\alpha + \beta X_i))(-1) = 0, S_\beta = \sum 2(Y_i - (\alpha + \beta X_i))(-X_i) = 0$$

これより**正規方程式** $n\alpha + \beta \sum X_t = \sum Y_t,\ \alpha \sum X_t + \beta \sum X_t^2 = \sum X_t Y_t$ が得られる。

X, Y の平均､分散､共分散が $\overline{X} = \Sigma X_t/n$, $s_X^2 = \Sigma X_t^2/n - \overline{X}^2$, $\overline{Y} = \Sigma Y_t/n$, $s_Y^2 = \Sigma Y_t^2/n - \overline{Y}^2$,

$s_{XY} = \Sigma X_t Y_t/n - \overline{X} \cdot \overline{Y}$ のように書けることから、正規方程式は

$\overline{Y} - \alpha - \beta\overline{X} = 0,\ s_{xy} + \overline{X} \cdot \overline{Y} - \overline{X}\alpha - (s_x^2 + \overline{X}^2)\beta = 0$ となり、これを解くと

$b = \dfrac{s_{XY}}{s_X^2} = \dfrac{\mathrm{cov}(X,Y)}{\mathrm{var}(X)},\ a = \overline{Y} - b\overline{X}$ が得られる。

Y_t とその推定値 $\hat{Y}_t = a + bX_t$ との**回帰残差** $e_t = Y_t - (a + bX_t)$ は $\sum e_t = 0,\ \sum e_t x_t = 0$ となる。

2.2* 最尤法(Maximum Likelihood Method)　テキスト p-201

中心極限定理(p-48)から $\varepsilon_i \sim N(0, \sigma^2)$ とすると $f(\varepsilon_i) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)}$ となる。

n 個の標本観測誤差が生じる同時確率(尤度)は

$$f(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = f(\varepsilon_1) \cdot f(\varepsilon_2) \ldots f(\varepsilon_n) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{\left(-\frac{1}{2\sigma^2}\sum \varepsilon_i^2\right)} = L(\alpha, \beta, \sigma^2)$$ となる。

この尤度 L を最大にする問題は単調増加変換した対数尤度 lnL 最大化問題と同値である

lnL(α, β, σ²)=$n\ln\left(\frac{1}{\sqrt{2\pi\sigma}}\right)-\frac{1}{2\sigma^2}\sum\left(Y_i-\alpha-\beta X_i\right)^2$ 最大化の１階の条件は最小２乗法の条件と同一であるから、最小２乗推定量(LSE)と最尤推定量(MLE)は同じになる。
（σ²の MLE は$\sum e_i^2/n$となる）

3. 当てはまり(goodness of fit)の尺度: 決定係数(Coefficient of Determination) p-17

$e_i=Y_i-\hat{Y}_i$，$\hat{Y}_i$=a+bX_iとするとΣe_i=0, ΣX_ie_i=0, Σ$\hat{Y}_i e_t$=aΣe_i+bΣX_ie_i=0 が成り立つ。

TSS=$\sum\left(Y_t-\bar{Y}\right)^2=\sum\left(\hat{Y}_t-\bar{Y}\right)^2+\sum\left(Y_t-\hat{Y}_t\right)^2=\sum\left(\hat{Y}_t-\bar{Y}\right)^2+\sum e_t^2$=ESS+RSS

両辺を TSS で割ると 1 = ESS/TSS + RSS/TSS が得られるが、

$0\leqq r^2=\frac{ESS}{TSS}=1-\frac{RSS}{TSS}\leqq 1$ を決定係数とよぶ。

また相関係数 r は r=$\pm\sqrt{r^2}$ からも得ることができる

4. 重回帰分析(multiple regression analysis) p-53

4.1 最小２乗法(K 変数モデル)

$Y_t=\beta_1+\beta_2X_{2t}+\beta_3X_{3t}+\ldots+\beta_KX_{Kt}+\varepsilon_t$, t=1,…n

$$\min_{\beta 1,\ldots,\beta K}\sum_{t=1}^{n}(Y_t-\beta_1-\beta_2X_{2t}\ldots-\beta_KX_{Kt})^2$$

$$\sum 2(Y_t-\beta_1-\beta_2X_{2t}\ldots-\beta_KX_{Kt})\cdot-1=0 \quad\rightarrow\sum Y_t=n\beta_1+\beta_2\sum X_{2t}+\ldots+\beta_K\sum X_{Kt}$$

$$\sum 2(Y_t-\beta_1-\beta_2X_{2t}\ldots-\beta_KX_{Kt})\cdot(-X_{2t})=0\rightarrow\sum X_{2t}Y_t=\beta_1\sum X_{2t}+\beta_2\sum X_{2t}^2+\ldots+\beta_K\sum X_{2t}X_{Kt}$$

$$\vdots \qquad\qquad \vdots$$

$$\sum 2(Y_t-\beta_1-\beta_2X_{1t}\ldots-\beta_KX_{Kt})\cdot(-X_{Kt})=0\rightarrow\sum X_{Kt}Y_t=\beta_1\sum X_{1t}+\beta_2\sum X_{2t}X_{Kt}+\ldots+\beta_K\sum X_{Kt1}^2$$

この K 本の連立方程式を解けばよい。
※行列で考えるとすっきりするかも…

$Y_t=\beta_1+\beta_2X_{2t}+\beta_3X_{3t}+\ldots+\beta_KX_{Kt}+\varepsilon_t$, t=1,…n → **y=Xβ+ε → b_{OLSE}=(X'X)$^{-1}$X'y**

min S=Σε_t^2=**ε'ε=(y-Xβ)'(y-Xβ)=y'y-2β'X'y+βX'Xb→S_β=-2X'y+2X'Xβ=0**

4.2 当てはまりの尺度 p-68

(1)決定係数、重相関係数

$$R^2=1-\frac{\sum e_t^2}{\sum(Y_t-\bar{Y})^2}$$ =(y と $\hat{y}$ との重相関係数)²

(2)自由度修正済決定係数

$$\bar{R}^2=1-\frac{RSS/(n-k)}{TSS/(n-1)}=1-(1-R^2)\frac{n-1}{n-k}$$

(3)その他の指標

赤池情報量基準 AIC=ln(e'e/n)+2k/n, Schwarz-Bayes 基準 SC=ln(e'e/n)+k/n・ln(n)