第5章　統計学の回帰分析への応用

5.1　確率的モデル：単回帰モデル

再び，話を簡単にするために単回帰モデルを考えることにしよう。すなわち，(X_1, Y_1)，(X_2, Y_2)，$\cdots$，(X_n, Y_n) のように n 組のデータがあり，X_i と Y_i との間に線型関係を想定する。

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

最小二乗法を用いて，データに直線のあてはめを行った。その結果，$\hat{\alpha}$，$\hat{\beta}$，$\hat{Y}_i$ を求めるための公式は，

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} = \frac{\sum_{i=1}^{n} X_i Y_i - n\overline{XY}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}$$

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

であった。

$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i$ とするとき，Y_i，$\hat{Y}_i$，$\hat{u}_i$，$\hat{\alpha}$，$\hat{\beta}$ の関係は以下の通りである。

$$Y_i = \hat{Y}_i + \hat{u}_i = \hat{\alpha} + \hat{\beta}X_i + \hat{u}_i$$

残差 $\hat{u}_i$ が必ず含まれることから，回帰モデルを

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

として誤差項 (または，攪乱項) u_i を含め，それを確率変数として考える。

u_i は平均 0，分散 σ^2 の正規分布が仮定されることが多い。

ある確率密度分布（ここでは正規分布）があって，その分布に従い，データ（ここでは Y_i）が生成されるモデルのことを確率的モデルと呼ぶ。

u_i は確率変数なので，Y_i も確率変数となる。

Y_i： 被説明変数，従属変数

X_i： 説明変数，独立変数

α，β： 未知母数 (未知パラメータ)

$\hat{\alpha}$，$\hat{\beta}$： 推定量 (ここでは，最小二乗推定量)，時には，推定値（最小二乗推定値）

(* 復習) 推定量と推定値：

統計学では，

母平均 μ の推定量は $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$，推定値は $\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$

母分散 σ^2 の推定量は $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$，推定値は $s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2$

$x_1, x_2, \cdots, x_n$ は観測値（または，実現値）

$X_1, X_2, \cdots, X_n$ はそれぞれの観測値に対応する確率変数

---

1. 残差 $\hat{u}_i$ は u_i の実現値としてみなすことができる。

2. 推定量 $\hat{\alpha}$，$\hat{\beta}$ の性質を統計学的に考察可能となる。

5.2　回帰モデルの仮定

回帰モデル

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

の仮定：

1. X_i は確率変数でないと仮定する（固定された値）。

2. **すべての i について，$\mathbf{E}(u_i) = 0$ とする。**

3. **すべての i について，$\mathbf{V}(u_i) = \sigma^2$ とする（$\mathbf{V}(u_i) = \mathbf{E}(u_i^2) = \sigma^2$ に注意）。**

(* 復習) 分散：

確率変数 X の平均 $\mu = \mathbf{E}(X)$，分散 $\sigma^2 = \mathbf{V}(X)$

分散の定義： $\sigma^2 = \mathbf{V}(X) = \mathbf{E}((X-\mu)^2) = \mathbf{E}(X^2) - \mu^2$

もし $\mu = \mathbf{E}(X) = 0$ であれば，$\sigma^2 = \mathbf{V}(X) = \mathbf{E}(X^2)$

---

4. **すべての $i \neq j$ について，$\mathbf{Cov}(u_i, u_j) = 0$ とする（$\mathbf{Cov}(u_i, u_j) = \mathbf{E}(u_i u_j) = 0$ に注意）。**

(* 復習) 共分散：

確率変数 X**，**Y **の平均** $\mu_X = \mathbf{E}(X)$**，**$\mu_Y = \mathbf{E}(Y)$**，共分散** $\sigma_{XY} = \mathbf{Cov}(X, Y)$

共分散の定義： $\sigma_{XY} = \mathbf{Cov}(X, Y) = \mathbf{E}((X - \mu_X)(Y - \mu_Y)) = \mathbf{E}(XY) - \mu_X\mu_Y$

もし $\mu_X = \mathbf{E}(X) = 0$**，または，**$\mu_Y = \mathbf{E}(Y) = 0$ **であれば，**$\sigma_{XY} = \mathbf{Cov}(X, Y) = \mathbf{E}(XY)$

---

5. すべての i **について，**$u_i \sim N(0, \sigma^2)$ **とする（正規分布）。**

6. $n \longrightarrow \infty$ **のとき（データ数が無限大になると），**$\sum_{i=1}^{n}(X_i - \overline{X})^2 \longrightarrow \infty$ **とする。**

攪乱項 $u_1, u_2, \cdots, u_n$ はそれぞれ互いに独立で，それぞれは平均ゼロ，分散 σ^2 の正規分布を仮定する。

再度，まとめて，回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i \qquad i = 1, 2, \cdots, n$$

ただし，$u_1, u_2, \cdots, u_n$ はそれぞれ互いに独立で，

すべての $i = 1, 2, \cdots, n$ について，$u_i \sim N(0, \sigma^2)$ を仮定する。

ただし，

Y_i**： 被説明変数，従属変数**

X_i**： 説明変数，独立変数**

α，β，σ^2： 未知母数 (未知パラメータ)

$\hat{\alpha}$，$\hat{\beta}$，s^2： 推定量 (最小二乗推定量)，s^2（σ^2 の推定量）については後述。

(* 復習) 期待値：

定数 a，b，確率変数 X について，$\mathbf{E}(aX \pm b) = a\mathbf{E}(X) \pm b$

特に，回帰直線 $Y_i = \alpha + \beta X_i + u_i$ について，

$$\mathbf{E}(Y_i) = \mathbf{E}(\alpha + \beta X_i + u_i) = \alpha + \beta X_i + \mathbf{E}(u_i) = \alpha + \beta X_i$$

として解釈される（α，β，X_i は非確率変数，u_i は確率変数）。

5.2.1　誤差項（攪乱項）の経済学的意味

1. 経済理論自身が不完全：X 以外にも他の説明変数が必要であるにもかかわらず，それを誤って除いている可能性がある。

2. モデルの定式化が不完全：Y と X との間の線形関係が誤りかもしれない。

3. 理論モデルとデータとの対応：理論モデルで考えられる変数と実際に用いたデータが適当でないかもしれない。例：所得のデータについては国民総生産，国民所得，可処分所得，労働所得 $\cdots$，金利では公定歩合，国債利回り，定期預金金利，全国銀行平均約定金利 $\cdots$

4. 測定上の誤差：経済データは一般的に推計されているため完全ではない。誤差を含む。

5.3　$\hat{\alpha}$，$\hat{\beta}$ の統計的性質

もう一度，2 つの式を並べて比べる。

$Y_i = \alpha + \beta X_i + u_i$　　$u_1, u_2, \cdots, u_n$ はそれぞれ互いに独立で，$u_i \sim N(0, \sigma^2)$ を仮定

$Y_i = \hat{\alpha} + \hat{\beta} X_i + \hat{u}_i$

$(\hat{\alpha}, \hat{\beta})$ は (α, β) の最小二乗法による推定量である。

すなわち，

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \qquad \hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

となる。

ただし，

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

とする。

$\hat{u}_i$ は残差で，$\hat{u}_i = Y_i - \hat{\alpha} - \hat{\beta}X_i$ と計算される。

5.3.1 $\hat{\beta}$ について

β の最小二乗推定量 $\hat{\beta}$ は次のように変形される。分母の添え字を j に変更する。

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

$$
\begin{aligned}
&= \frac{\beta \sum_{i=1}^{n}(X_i - \overline{X})X_i - \beta \sum_{i=1}^{n}(X_i - \overline{X})\overline{X}}{\sum_{j=1}^{n}(X_j - \overline{X})^2} + \frac{\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\sum_{j=1}^{n}(X_j - \overline{X})^2} \\
&= \frac{\beta \sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{j=1}^{n}(X_j - \overline{X})^2} + \frac{\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\sum_{j=1}^{n}(X_j - \overline{X})^2} \\
&= \beta + \frac{\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\sum_{j=1}^{n}(X_j - \overline{X})^2} \\
&= \beta + \sum_{i=1}^{n} \omega_i u_i
\end{aligned}
$$

である。ただし，$\omega_i = \dfrac{(X_i - \overline{X})}{\sum_{j=1}^{n}(X_j - \overline{X})^2}$ とする。

途中の計算（2 行目の分子第 2 項目，5 行目の分子第 1 項，6 行目の分子第 2 項）で，$\sum_{i=1}^{n}(X_i - \overline{X}) = 0$ に注意せよ（$\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ から得られる）。

3 行目から 4 行目では，$Y_i = \alpha + \beta X_i + u_i$ が代入されている。

3 行目では，ω_i を使って，

$$
\begin{aligned}
\hat{\beta} &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{j=1}^n (X_j - \overline{X})^2} \\
&= \frac{\sum_{i=1}^n (X_i - \overline{X})Y_i}{\sum_{j=1}^n (X_j - \overline{X})^2} \\
&= \sum_{i=1}^n \omega_i Y_i
\end{aligned}
$$

と書き直すこともできる。

$\longrightarrow$ β の最小二乗推定量 $\hat{\beta}$ は Y_i の線形推定量となっている。

よって，まとめると，

$$
\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{j=1}^n (X_j - \overline{X})^2}
$$

$$= \sum_{i=1}^{n} \omega_i Y_i$$

$$= \beta + \sum_{i=1}^{n} \omega_i u_i$$

となる。$\omega_i = \dfrac{(X_i - \overline{X})}{\sum_{j=1}^{n}(X_j - \overline{X})^2}$ である。

5.3.2　$\hat{\alpha}$ について

α の最小二乗推定量 $\hat{\alpha}$ については，

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

$$= \frac{1}{n}\sum_{i=1}^{n} Y_i - \overline{X}\sum_{i=1}^{n} \omega_i Y_i$$

$$= \sum_{i=1}^{n}(\frac{1}{n} - \overline{X}\omega_i)Y_i$$

$$= \sum_{i=1}^{n} \lambda_i Y_i$$

となる。ただし， $\lambda_i = \dfrac{1}{n} - \overline{X}\omega_i$ **である。**

$\longrightarrow$ α **の最小二乗推定量** $\hat{\alpha}$ **は** Y_i **の線形推定量となっている。**

さらに，書き換える。

$$\hat{\alpha} = \sum_{i=1}^{n}(\frac{1}{n} - \overline{X}\omega_i)Y_i$$

$$= \sum_{i=1}^{n}(\frac{1}{n} - \overline{X}\omega_i)(\alpha + \beta X_i + u_i)$$

$$= \sum_{i=1}^{n}(\frac{1}{n} - \overline{X}\omega_i)\alpha + \sum_{i=1}^{n}\beta(\frac{1}{n} - \overline{X}\omega_i)X_i + \sum_{i=1}^{n}(\frac{1}{n} - \overline{X}\omega_i)u_i$$

$$= \alpha + \sum_{i=1}^{n}(\frac{1}{n} - \overline{X}\omega_i)u_i$$

$$= \alpha + \sum_{i=1}^{n}\lambda_i u_i$$

下記が途中で，

$$\sum_{i=1}^{n}\omega_i = \frac{\sum_{i=1}^{n}(X_i - \overline{X})}{\sum_{j=1}^{n}(X_j - \overline{X})^2} = 0$$

$$\sum_{i=1}^{n}\omega_i X_i = \sum_{i=1}^{n}\omega_i X_i - \sum_{i=1}^{n}\omega_i\overline{X} = \sum_{i=1}^{n}\omega_i(X_i - \overline{X}) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{j=1}^{n}(X_j - \overline{X})^2} = 1$$

が使われている（$\sum_{i=1}^{n} \beta(\frac{1}{n} - \overline{X}\omega_i)X_i = 0$ に注意）。

下記のように書き換えても同じ結果が得られる。

$$
\begin{aligned}
\hat{\alpha} &= \overline{Y} - \hat{\beta}\overline{X} \\
&= \alpha - (\hat{\beta} - \beta)\overline{X} + \overline{u} \\
&= \alpha - \overline{X}\sum_{i=1}^{n} \omega_i u_i + \frac{1}{n}\sum_{i=1}^{n} u_i \\
&= \alpha + \sum_{i=1}^{n} (\frac{1}{n} - \overline{X}\omega_i)u_i \\
&= \alpha + \sum_{i=1}^{n} \lambda_i u_i
\end{aligned}
$$

となる。$\lambda_i = \frac{1}{n} - \overline{X}\omega_i$ としている。

1 行目の $\overline{Y}$ に $\overline{Y} = \alpha + \beta\overline{X} + \overline{u}$ が代入されている（$Y_i = \alpha + \beta X_i + u_i$ を i について足し合わせて，n で割ると，この式が得られる)。ただし，$\overline{Y} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} Y_i$，$\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$，$\overline{u} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} u_i$ である。

2 行目の $\hat{\beta} - \beta$ に $\hat{\beta} = \beta + \sum_{i=1}^{n} \omega_i u_i$ が使われている。

まとめると，

$$
\begin{aligned}
\hat{\alpha} &= \alpha + \sum_{i=1}^{n} (\frac{1}{n} - \overline{X}\omega_i) u_i \\
&= \alpha + \sum_{i=1}^{n} \lambda_i u_i
\end{aligned}
$$

となる。

5.3.3　$\hat{\alpha}$，$\hat{\beta}$ の期待値（平均）

$\hat{\beta}$ は次のように書き換えられた。

$$\hat{\beta} = \beta + \sum_{i=1}^{n} \omega_i u_i$$

の両辺に期待値をとる。

(* 復習) 期待値：

定数 a**，**b**，確率変数** X **について，**$\mathbf{E}(aX \pm b) = a\mathbf{E}(X) \pm b \longrightarrow$ **(再掲)**

(* 復習) 確率変数の和の期待値：

2 つの確率変数 X**，**Y **について，**$\mathbf{E}(X \pm Y) = \mathbf{E}(X) \pm \mathbf{E}(Y)$

$$\mathbf{E}(\hat{\beta}) = \mathbf{E}(\beta + \sum_{i=1}^{n} \omega_i u_i) = \beta + \sum_{i=1}^{n} \mathbf{E}(\omega_i u_i) = \beta + \sum_{i=1}^{n} \omega_i \mathbf{E}(u_i) = \beta$$

となる。

$\hat{\beta}$ は β の不偏推定量であると言える。

(* 復習) 不偏推定量について：

n 個の確率変数 X_1，X_2，$\cdots$，X_n は互いに独立で，

それぞれは母数 θ に依存するものとする（例えば，$\theta = (\mu,\ \sigma^2)$ である）。

θ の推定量を $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ とする。

$\mathbf{E}(\hat{\theta}) = \theta$ となるとき，$\hat{\theta}$ は θ の不偏推定量であるという。

$\hat{\alpha}$ については，

$$\hat{\alpha} = \alpha + \sum_{i=1}^{n} \lambda_i u_i$$

を利用して，辺々に期待値をとると，

$$\mathbf{E}(\hat{\alpha}) = \mathbf{E}(\alpha + \sum_{i=1}^{n} \lambda_i u_i) = \alpha + \sum_{i=1}^{n} \lambda_i \mathbf{E}(u_i) = \alpha$$

$\lambda_i = \dfrac{1}{n} - \overline{X}\omega_i$ は非確率変数であることに注意。

$\hat{\alpha}$ は α の不偏推定量であると言える。