計量分析数学資料

1　基本事項

- 標本平均 $\bar{X}$
$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$
- 平均周りの偏差
$$\Sigma(X_i - \bar{X})$$
- 偏差二乗和 S_{xy}
$$S_{xx} = \Sigma(X_i - \bar{X})^2 = (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2$$
- 標本分散 ${s_x}^2$
$${s_x}^2 = \frac{S_{xy}}{n-1} = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$$
- 標本標準偏差 s_x
$$s_x = \sqrt{{s_x}^2} = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{n-1}}$$
- 偏差の積和 S_{xy}
$$S_{xy} = \Sigma(X_i - \bar{X})(Y_i - \bar{Y})$$
- 標本共分散 s_{xy}
$$s_{xy} = \frac{S_{xy}}{n-1} = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

1.1　さまざまな式変形

標本平均、標本分散の変形

$$\Sigma X_i = n\bar{X}$$

$$\Sigma(X_i - \bar{X}) = n\bar{X} - n\bar{X} = 0$$

分散の分母を変形

$$\begin{aligned} \Sigma(X_i - \bar{X})^2 &= \Sigma({X_i}^2 - 2X_i\bar{X} + \bar{X}^2) \\ &= \Sigma{X_i}^2 - \Sigma 2X_i\bar{X} + \Sigma\bar{X}^2 \end{aligned} \tag{1}$$

1. ここで右辺の第二項は $\Sigma 2X_i\bar{X} = 2\bar{X}\Sigma X_i$ であり、
2. $\Sigma X_i = n\bar{X}$ より、$\Sigma 2X_i\bar{X} = 2\bar{X}\Sigma X_i = 2n\bar{X}^2$
3. また第 3 項は $\Sigma\bar{X}^2 = n\bar{X}^2$ であるから

$$\Sigma(X_i - \bar{X})^2 = \Sigma{X_i}^2 - n\bar{X}^2$$

さらに $\Sigma(X_i - \bar{X})\bar{X} = \bar{X}\Sigma(X_i - \bar{X}) = 0$ より、

$$\begin{aligned} \Sigma(X_i - \bar{X})^2 &= \Sigma(X_i - \bar{X})X_i - \Sigma(X_i - \bar{X})\bar{X} \\ &= \Sigma(X_i - \bar{X})X_i \end{aligned} \tag{2}$$

と変形できる。共分散も同様に変形でき、

$$
\begin{aligned}
\Sigma(X_i-\bar{X})(Y_i-\bar{Y}) &= \Sigma X_iY_i - n\bar{X}\bar{Y} \\
&= \Sigma(X_i-\bar{X})Y_i \\
&= \Sigma(Y_i-\bar{Y})X_i
\end{aligned} \tag{3}
$$

2 最小二乗法 (単回帰モデル)

以下に定義する J を最小にする α,β を考える。

$$
J = \Sigma {u_i}^2 = \Sigma(Y_i-\alpha-\beta X_i)^2 \tag{4}
$$

J を α,β で偏微分すると、

$$
\begin{aligned}
\frac{\partial J}{\partial \alpha} &= \Sigma\frac{\partial}{\partial \alpha}(Y_i-\alpha-\beta X_i)^2 \\
&= \Sigma\{-2(Y_i-\alpha-\beta X_i)\}
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial \beta} &= \Sigma\frac{\partial}{\partial \beta}(Y_i-\alpha-\beta X_i)^2 \\
&= \Sigma\{-2X_i(Y_i-\alpha-\beta X_i)\}
\end{aligned} \tag{6}
$$

ここで、J を最小にする α,β を、それぞれ $\hat{\alpha},\hat{\beta}$ とおく。$\hat{\alpha},\hat{\beta}$ は最小二乗推定量と呼ばれる。α,β が $\hat{\alpha},\hat{\beta}$ を取るとき、(4) 式を偏微分した (5) 式、(6) 式はそれぞれ 0 になる。従って、

$$
\begin{cases}
\Sigma(Y_i-\alpha-\beta X_i) = \Sigma Y_i - n\hat{\alpha} - \hat{\beta}\Sigma X_i = 0 & (7) \\
\Sigma X_i(Y_i-\alpha-\beta X_i) = \Sigma X_iY_i - \hat{\alpha}\Sigma X_i - \hat{\beta}\Sigma {X_i}^2 = 0 & (8)
\end{cases}
$$

これを変形すると正規方程式が得られる。

$$
\begin{cases}
n\hat{\alpha} + (\Sigma X_i)\hat{\beta} = \Sigma Y_i & (9) \\
(\Sigma X_i)\hat{\alpha} + (\Sigma {X_i}^2)\hat{\beta} = \Sigma X_iY_i & (10)
\end{cases}
$$

これを解いて $\hat{\beta}$ の推定量

$$
\begin{aligned}
\hat{\beta} &= \frac{\Sigma X_iY_i - n\bar{X}\bar{Y}}{(\Sigma {X_i}^2) - n\bar{X}^2} \\
&= \frac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})}{\Sigma(X_i-\bar{X})^2} \\
&= \frac{\Sigma(X_i-\bar{X})Y_i}{\Sigma(X_i-\bar{X})^2}
\end{aligned} \tag{11}
$$

を得る。これはさらに以下のようにも変形できる。

$$
\begin{aligned}
\hat{\beta} &= \frac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})}{\Sigma(X_i-\bar{X})^2} = \frac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})/(n-1)}{\Sigma(X_i-\bar{X})^2/(n-1)} \\
&= \frac{S_{xy}}{S_x^2} = \frac{s_{xy}}{s_{xy}^2}
\end{aligned} \tag{12}
$$

正規方程式の第一式を変形すると、α を β で表すことができ、

$$
\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}_i
$$

となる。ここでまとめると、

$$
\begin{cases}
\hat{\beta} = \dfrac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{\Sigma(X_i - \bar{X})^2} = \dfrac{\Sigma X_i Y_i - n\bar{X}\bar{Y}}{\Sigma {X_i}^2 - n\bar{X}^2} & (13) \\
\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}_i & (14)
\end{cases}
$$

■回帰直線

$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i$ を回帰直線とよび、$\hat{Y}_i$ は Y_i の理論値、もしくは予測値と呼ばれる。
Y の実績値 Y_i はこの回帰直線を用いて表現すると以下のように書ける。

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

■決定係数

$$R^2 = \frac{\Sigma(\hat{Y}_i - \bar{Y})}{\Sigma(Y_i - \bar{Y})} = \frac{\Sigma(Y_i - \bar{Y}) - \Sigma{\hat{u}_i}^2}{\Sigma(Y_i - \bar{Y})} = 1 - \frac{\Sigma{\hat{u}_i}^2}{\Sigma(Y_i - \bar{Y})}$$

■誤差項の仮定

1. 残差の総和が 0 ($\Sigma\hat{u}_i = 0$)
2. 説明変数 X_i と残差の積和が 0 ($\Sigma X_i\hat{u}_i = 0$)
3. 理論値 $\hat{Y}_i$ と残差の積和が 0 ($\Sigma\hat{Y}_i\hat{u}_i = 0$)

■OLS の仮定

1. 線形モデルとして書ける
2. 説明変数 X が散らばっている
3. 標本 (x_i, y_i) は無作為抽出
4. 誤差項 u_i は説明変数 x_i と相関しない
5. 誤差項は互いに独立 $E(u_i, u_j) = \forall i \neq j$
6. 誤差項の分散は均一 $V(u_i|x_i) = E(u_iu_i|x_i) = \sigma^2$
7. (小標本の場合) 誤差分布の正規性 u_iN