

重回帰分析における予測区間

重回帰分析の予測区間を求める手順を以下に記します。せっかくですので、131 ページで登場した「伊勢橋店」における予測区間を求める手順もあわせて記します。

Step1

$$\begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{pmatrix} \text{の逆行列である} \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{pmatrix}^{-1} = \begin{pmatrix} S^{11} & S^{12} & \cdots & S^{1p} \\ S^{21} & S^{22} & \cdots & S^{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S^{p1} & S^{p2} & \cdots & S^{pp} \end{pmatrix} \text{を求める。}$$

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}^{-1} = \begin{pmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{pmatrix} = \begin{pmatrix} 20.1 & -792 \\ -792 & 128840 \end{pmatrix}^{-1} = \begin{pmatrix} 0.0657 & 0.0004 \\ 0.0004 & 0.00001 \end{pmatrix}$$

Step2

マハラノビスの汎距離の 2 乗である

$$\begin{aligned} D^2 = & \{(x_1 - \bar{x}_1)(x_1 - \bar{x}_1)S^{11} + (x_1 - \bar{x}_1)(x_2 - \bar{x}_2)S^{12} + \cdots + (x_1 - \bar{x}_1)(x_p - \bar{x}_p)S^{1p} \\ & + (x_2 - \bar{x}_2)(x_1 - \bar{x}_1)S^{21} + (x_2 - \bar{x}_2)(x_2 - \bar{x}_2)S^{22} + \cdots + (x_2 - \bar{x}_2)(x_p - \bar{x}_p)S^{2p} \\ & \vdots \\ & + (x_p - \bar{x}_p)(x_1 - \bar{x}_1)S^{p1} + (x_p - \bar{x}_p)(x_2 - \bar{x}_2)S^{p2} + \cdots + (x_p - \bar{x}_p)(x_p - \bar{x}_p)S^{pp}\} (\text{個体の個数} - 1) \end{aligned}$$

を求める。

$$\begin{aligned} D^2 = & \{(x_1 - \bar{x}_1)(x_1 - \bar{x}_1)S^{11} + (x_1 - \bar{x}_1)(x_2 - \bar{x}_2)S^{12} \\ & + (x_2 - \bar{x}_2)(x_1 - \bar{x}_1)S^{21} + (x_2 - \bar{x}_2)(x_2 - \bar{x}_2)S^{22}\} (\text{データの個数} - 1) \\ = & \{(10 - 7.7)(10 - 7.7) \times 0.0657 + (10 - 7.7)(110 - 156) \times 0.0004 \\ & + (110 - 156)(10 - 7.7) \times 0.0004 + (110 - 156)(110 - 156) \times 0.00001\} (10 - 1) \\ = & 2.6 \end{aligned}$$

Step3

予測区間を求める。

$$a_1x_1 + a_2x_2 + \dots + a_px_p + b \pm \sqrt{F(1, \text{個体の個数} - \text{説明変数の個数} - 1; 0.05) \times \left(1 + \frac{1}{\text{個体の個数}} + \frac{D^2}{\text{個体の個数} - 1}\right) \times \frac{S_e}{\text{個体の個数} - \text{説明変数の個数} - 1}}$$

$$a_1 \times 10 + a_2 \times 110 + b \pm \sqrt{F(1, 10 - 2 - 1; 0.05) \times \left(1 + \frac{1}{10} + \frac{2.6}{10 - 1}\right) \times \frac{4173.0}{10 - 2 - 1}} = 443.0 \pm 68$$

四捨五入の関係でおかしくなりましたが、お店の面積が 10 坪で最寄り駅からの距離が 110m の 1 ヶ月の売上額は、130 ページで求められた 447.3 ではなくて、本当は 443.0 です。

信頼率 99% の場合は、

$$F(1, \text{個体の個数} - \text{説明変数の個数} - 1; 0.05) = F(1, 10 - 2 - 1; 0.05) = 5.6 \text{ の部分を}$$

$$F(1, \text{個体の個数} - \text{説明変数の個数} - 1; 0.01) = F(1, 10 - 2 - 1; 0.01) = 12.2 \text{ にするだけです。}$$