2024/04/11

藪友良

第 2 章の答え

練習問題 1

OLS 推定量$\hat{\alpha}$は、$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$であり、これを書き換えると次式となる。

$$\bar{Y} = \hat{\alpha} + \hat{\beta}\bar{X}$$

よって、$X = \bar{X}$のとき回帰直線$\hat{Y} = \hat{\alpha} + \hat{\beta}X$の値は$\bar{Y} = \hat{\alpha} + \hat{\beta}\bar{X}$であり、回帰直線は点$(\bar{X}, \bar{Y})$を通ることが確認できる。

練習問題 2

標本平均は、それぞれ次のようになる。

$$\bar{Y} = \frac{1}{5}(3 + 6 + 5 + 7 + 9) = 6$$

$$\bar{X} = \frac{1}{5}(5 + 6 + 5 + 8 + 11) = 7$$

また、標本分散は、それぞれ次のようになる。

$$s_Y^2 = \frac{1}{5-1}[(3-6)^2 + (6-6)^2 + (5-6)^2 + (7-6)^2 + (9-6)^2] = 5$$

$$s_X^2 = \frac{1}{5-1}[(5-7)^2 + (6-7)^2 + (5-7)^2 + (8-7)^2 + (1-7)^2] = 6.5$$

最後に、標本共分散s_{XY}は、

$$\frac{1}{5-1}[(3-6)(5-7) + (6-6)(6-7) + (5-6)(5-7) + (7-6)(8-7) + (9-6)(1-7)]$$

であり、これを計算すると、5.25 となる。標本相関係数は、これまでの情報を用いて、次のようになる。

$$r_{XY} = \frac{5.25}{\sqrt{5 \times 6.5}} = 0.9209$$

練習問題 3

予測値$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i$と練習問題 1 で求めた式$\bar{Y} = \hat{\alpha} + \hat{\beta}\bar{X}$を用いると、次の関係が得られる。

$$\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2=\sum_{i=1}^{n}(\hat{\alpha}+\hat{\beta}X_i-(\hat{\alpha}+\hat{\beta}\bar{X}))^2$$

$$=\sum_{i=1}^{n}(\hat{\beta}X_i-\hat{\beta}\bar{X})^2$$

$$=\hat{\beta}^2\sum_{i=1}^{n}(X_i-\bar{X})^2$$

この偏差 2 乗和を決定係数の式に代入すると、

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

となる。この式から、$\hat{\beta}=0$の場合、決定係数も 0 となることがわかる。$\hat{\beta}=0$は、説明変数Xで被説明変数Yの動きを全く説明できない状況であり、このとき、決定係数は 0 になるといえる。

練習問題 4

(a) OLS 推定量は次のようになる。

$$\hat{\beta}=-\frac{430}{400}=-1.075$$

$$\hat{\alpha}=65-(-1.075)\times 25=91.875$$

クラスの人数が 0 人は概念的にありえないため、定数項は数学的切片と解釈される。また、$\hat{\beta}=-1.075$から、クラス人数が 1 人増えると、クラスの平均点が 1.075 点下がる。

(b) クラスの人数が 20 のとき、予測値は次のようになる。

$$\hat{Y}=91.875-1.075\times 20=70.375$$

(c)決定係数は、練習問題 3 の関係式を用いると、次のようになる。

$$R^2=(1.075)^2\frac{400}{500}=0.9245$$

平均点の全変動のうち 92.45%はクラスサイズの変動で説明できる。

練習問題 5

(a)OLS 推定量は次のようになる。

$$\hat{\beta} = \frac{1000}{250} = 4$$

$$\hat{\alpha} = 90 - 4 \times 20 = 10$$

定数項は、気温が 0 度のときの冷麺の売り上げ個数である。係数$\hat{\beta} = 4$から、気温が 1 度上がると、売り上げ個数は 4 個増える。

(b)気温が 30 度のとき、予測値は次のようになる。

$$\hat{Y} = 10 + 4 \times 30 = 130$$

(c) 決定係数は、練習問題 3 の関係式を用いると、次のようになる。

$$R^2 = \hat{\beta}^2 \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 4^2 \frac{250}{5000} = 0.8$$

売り上げ個数の全変動のうち 80%は気温の変動により説明できる。

練習問題 6

練習問題 3 の結果から、

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

となる。右辺$\hat{\beta}$に、OLS 推定量の公式である

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

を代入して展開すると次のようになる。

$$R^2 = \left(\frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)^2 \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

$$= \frac{\left(\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})\right)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

$$= \left\{\frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}\right\}^2$$

ここで、右辺の分母と分子を$n-1$で割ると、次式となり、これはまさにX_iとY_iの標本相関係数の 2 乗である。

$$\left\{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2}}\right\}^2$$

練習問題 7

係数 β を 0 としたモデルでは、残差 2 乗和は次のようになる。

$$\sum_{i=1}^{n}\tilde{u}_i^2=\sum_{i=1}^{n}(Y_i-\tilde{\alpha})^2$$

定数項$\tilde{\alpha}$に関して残差 2 乗和を微分して 0 と置くと、

$$\frac{\partial\sum_{i=1}^{n}\tilde{u}_i^2}{\partial\tilde{\alpha}}=\sum_{i=1}^{n}\frac{\partial\tilde{u}_i^2}{\partial\tilde{\alpha}}=\sum_{i=1}^{n}\frac{\partial\tilde{u}_i^2}{\partial\tilde{u}_i}\frac{\partial\tilde{u}_i}{\partial\tilde{\alpha}}=-2\sum_{i=1}^{n}(Y_i-\hat{\alpha})=0$$

となる。式展開では、次の関係式を用いた。

$$\frac{\partial\tilde{u}_i^2}{\partial\tilde{u}_i}=2\tilde{u}_i=2(Y_i-\tilde{\alpha})$$

$$\frac{\partial\tilde{u}_i}{\partial\tilde{\alpha}}=\frac{\partial(Y_i-\tilde{\alpha})}{\partial\tilde{\alpha}}=-1$$

ここで、両辺を－2 で割ると、次の正規方程式が得られる。

$$\sum_{i=1}^{n}(Y_i-\hat{\alpha})=0$$

上式を満たす$\tilde{\alpha}$は最小 2 乗推定量であるため、$\hat{\alpha}$と表記している。この場合、残差は$\hat{u}_i=Y_i-\hat{\alpha}$であるため、正規方程式から残差の和は 0 となる。そして、正規方程式を展開すると、

$$\sum_{i=1}^{n}Y_i-n\,\hat{\alpha}=0$$

であり、さらに両辺をnで割ると、$\hat{\alpha}$が標本平均$\bar{Y}$になることを確認できる。

$$\hat{\alpha}=\frac{1}{n}\sum_{i=1}^{n}Y_i=\bar{Y}$$

練習問題 8

(a) 残差 2 乗和は、次のように表現できる。

$$\sum_{i=1}^{n}(Y_i-\tilde{\alpha}-\tilde{\beta}X_i)^2=\sum_{i=1}^{n}(Y_i-\hat{\alpha}-\hat{\beta}X_i+(\hat{\alpha}-\tilde{\alpha})+(\hat{\beta}-\tilde{\beta})X_i)^2$$

$Y_i=\hat{\alpha}+\hat{\beta}X_i+\hat{u}_i$を書き換えると、$\hat{u}_i=Y_i-\hat{\alpha}-\hat{\beta}X_i$となる。上式の右辺に、$\hat{u}_i=Y_i-\hat{\alpha}-\hat{\beta}X_i$を代入すると、右辺は次のようになる。

$$\sum_{i=1}^{n}(\hat{u}_i+(\hat{\alpha}-\tilde{\alpha})+(\hat{\beta}-\tilde{\beta})X_i)^2$$

(b)　上式右辺は、次のように展開できる。

$$\sum_{i=1}^{n}(\hat{u}_i+\{(\hat{\alpha}-\tilde{\alpha})+\left(\hat{\beta}-\tilde{\beta}\right)X_i\})^2=\sum_{i=1}^{n}\hat{u}_i^2+\sum_{i=1}^{n}\left\{(\hat{\alpha}-\tilde{\alpha})+(\hat{\beta}-\tilde{\beta})X_i\right\}^2$$

$$+2(\hat{\alpha}-\tilde{\alpha})\underbrace{\sum_{i=1}^{n}\hat{u}_i}_{=0}+2\left(\hat{\beta}-\tilde{\beta}\right)\underbrace{\sum_{i=1}^{n}\hat{u}_iX_i}_{=0}$$

$$=\sum_{i=1}^{n}\hat{u}_i^2+\sum_{i=1}^{n}((\hat{\alpha}-\tilde{\alpha})+(\hat{\beta}-\tilde{\beta})X_i)^2$$

式展開では、残差の性質($\sum_{i=1}^{n}\hat{u}_i=0$、$\sum_{i=1}^{n}\hat{u}_iX_i=0$)を用いた。

(c)これまでの結果をまとめると、次のようになる。

$$\sum_{i=1}^{n}(Y_i-\tilde{\alpha}-\tilde{\beta}X_i)^2=\sum_{i=1}^{n}\hat{u}_i^2+\sum_{i=1}^{n}((\hat{\alpha}-\tilde{\alpha})+(\hat{\beta}-\tilde{\beta})X_i)^2$$

ここで、右辺第 2 項は 2 乗和なので 0 以上となる。このため、次式が成立する。

$$\sum_{i=1}^{n}\hat{u}_i^2\leq\sum_{i=1}^{n}(Y_i-\tilde{\alpha}-\tilde{\beta}X_i)^2$$

左辺は OLS 推定量の残差 2 乗和であり、右辺は任意の$\tilde{\alpha}$と$\tilde{\beta}$に対する残差 2 乗和である。このため、OLS 推定量($\hat{\alpha}$、$\hat{\beta}$)は残差 2 乗和$\sum_{i=1}^{n}(Y_i-\tilde{\alpha}-\tilde{\beta}X_i)^2$を最小にしていることが確認できる。

練習問題 9、 10

ウェブサイトから、再現に必要なデータ、STATA と R のコードがダウンロードできる。