練習問題 3 解答

(1) [35.91, 826.45]

正規母集団 $N(\mu, \sigma^2)$ から無作為に観測した n 個の標本の標本分散を s^2 とすると、σ^2 の $100(1-\alpha)$ %信頼区間は

$$\left[\ \frac{(n-1)s^2}{y_{n-1,\frac{\alpha}{2}}},\ \frac{(n-1)s^2}{y_{n-1,1-\frac{\alpha}{2}}}\ \right]$$

によって求まる(スライド参照)。ここで $y_{n-1,p}$ は自由度 $n-1$ のカイ二乗分布に従う確率変数 W に対して $\Pr(W > y_{n-1,p}) = p$ を満たす定数であり、カイ二乗分布の分布表より求める事ができる。今 $n = 5, s^2 = 10, \alpha = 0.05$ であるので、上記の公式より、σ^2 の 95 %信頼区間は

$$\left[\frac{4\cdot 10^2}{y_{4,0.025}},\ \frac{4\cdot 10^2}{y_{4,0.975}}\right] = \left[\frac{4\cdot 10^2}{11.14},\ \frac{4\cdot 10^2}{0.484}\right] \approx [\ 35.91,\ 826.45\ ]$$

となる。

(2) $[-0.00655, 0.20655]$

正規母集団 $N(\mu, \sigma^2)$ から無作為に観測した n 個の標本の標本平均を $\overline{X}$ 、標本分散を s^2 とする。統計量 T を

$$T = \frac{\overline{X}-\mu}{\sqrt{s^2/n}} = \frac{\sqrt{n}(\overline{X}-\mu)}{s}$$

と定義すると、T は自由度 $n-1$ の t 分布に従う事を示す事ができる(スライド参照)。これより μ の $100(1-\alpha)$%信頼区間は

$$\left[\ \overline{X} - t_{n-1,\alpha}\sqrt{\frac{s^2}{n}}\ ,\ \overline{X} + t_{n-1,\alpha}\sqrt{\frac{s^2}{n}}\ \right]$$

によって与えられる事を示す事ができる。ここで $t_{n-1,\alpha}$ は自由度 $n-1$ の t 分布に従う確率変数 T に対して $\Pr(-t_{n-1,\alpha} < T < t_{n-1,\alpha}) = 1-\alpha$ を満たす定数であり、t 分布の分布表より求める事ができる。今、$n = 16$, $\overline{X} = 0.1, s^2 = 0.2^2 = 0.04, \alpha = 0.05$ であるので、上記の公式より、μ の 95 %信頼区間は

$$\left[0.1 - t_{15,0.05}\frac{0.2}{4},\ 0.1 + t_{15,0.05}\frac{0.2}{4}\right] = [\ 0.1 - 2.131\times 0.05,\ 0.1 + 2.131\times 0.05\ ]$$
$$= [-0.00655,\ 0.20655\ ]$$

となる。

(3)

(i) $\hat{\alpha}=-1$, $\hat{\beta}=2$

β と α の最少二乗推定値はスライドの公式より

$$\hat{\beta}=\frac{\sum_{i=1}^{3} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{3} X_i^2 - n\bar{X}^2}=\frac{2+10}{1+4+1}=2\text{, and }\ \hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=-1.$$

となる。よって $\hat{\alpha}=-1$, $\hat{\beta}=2$ となる。

(ii) $\frac{12}{13}\approx 0.92$

R^2 の定義は

$$R^2=\frac{\text{説明された変動}}{\text{全変動}}=\frac{\sum_{i=1}^{3}(\hat{Y}_i-\bar{Y})^2}{\sum_{i=1}^{3}(Y_i-\bar{Y})^2},\ \hat{Y}_i=\hat{\alpha}+\hat{\beta}X_i,$$

である。今、$\bar{Y}=-1$, $\hat{Y}_1=-1+2\times 1=1$, $\hat{Y}_2=-1+2\times(-2)=-5$, $\hat{Y}_3=-1+2\times 1=1$ であるので

$$R^2=\frac{(1+1)^2+(-5+1)^2+(1+1)^2}{(2+1)^2+(-5+1)^2+(0+1)^2}=\frac{4+16+4}{9+16+1}=\frac{24}{26}=\frac{12}{13}\approx 0.92$$

となる。

(4)

(i) 1.0

Y_i と X_i が完全な負の線形関係を持っているので、R^2 は 1.0 である。

(ii) 0

切片(定数項) α の最少二乗推定量は $\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}$ (講義スライド参照)。 よって $\hat{\beta}=0$ の時、$\hat{\alpha}=\bar{Y}$ となる。これは $\hat{Y}_i=\hat{\alpha}+\hat{\beta}X_i=\bar{Y}$ を意味している。これを R^2 の定義に代入すると

$$R^2=\frac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}=\frac{\sum_{i=1}^{n}(\bar{Y}-\bar{Y})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}=0.$$

となる。

(iii) $R^2=0.2$

R^2 は

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

と表せる。よって

$$R^2=1-\frac{120}{150}=1-0.8=0.2$$

となる。

(5)

(i) 2

残差平方和は $SSR = \sum_{i=1}^{3}(Y_i - bX_i)^2$ であり、正規方程式は

$$\frac{\partial SSR}{\partial b} = -2\sum_{i=1}^{3} X_i(Y_i - bX_i) = 0$$

これを解くと

$$\sum_{i=1}^{3} X_i(Y_i - bX_i) = 0 \iff \sum_{i=1}^{3} X_i Y_i = b\sum_{i=1}^{3} X_i^2 \iff b = \frac{\sum_{i=1}^{3} X_i Y_i}{\sum_{i=1}^{3} X_i^2} = \frac{15+5}{9+1} = \frac{20}{10} = 2$$

よって $\hat{\beta} = 2.$

(ii) 14

SSR は

$$SSR = \sum_{i=1}^{3}(Y_i - \hat{\beta}_{ols} X_i)^2 = (5-6)^2 + (-5+2)^2 + 2^2 = 1+9+4 = 14$$

となる。

(6)

(i) 3

この時 β の最少二乗推定値は

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{20}{100} = \frac{1}{5}$$

である。よってα の最少二乗推定値は

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} = 5 - \frac{1}{5}10 = 3$$

となる。

(ii) 0.25

定数項を含んだ単回帰モデルの場合 R^2 は相関係数の 2 乗であるので(スライド参照)

$$R^2 = r_{xy}^2 = \left(\frac{20}{4\times 10}\right)^2 = 0.25$$

となる。

(iii) 4

まず、

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i = \bar{Y} - \hat{\beta}\bar{X} + \hat{\beta} X_i \quad (\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \text{ を代入})$$

より

$$\hat{Y}_i - \bar{Y} = \hat{\beta}(X_i - \bar{X})$$

となるが、これより

$$(\hat{Y}_i - \bar{Y})^2 = \hat{\beta}^2 (X_i - \bar{X})^2$$

を得る。この両辺を $i=1,\ldots n$ まで足し合わせる事により

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

$$\Leftrightarrow \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = \hat{\beta}^2 \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \quad \text{(両辺を } \textstyle\sum_{i=1}^{n} (Y_i - \bar{Y})^2 \text{ で割る)}$$

$$\Leftrightarrow \hat{\beta}^2 = R^2 \frac{\frac{1}{n-1}\sum_{i=1}^{n} (Y_i - \bar{Y})^2}{\frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2} = 0.8 \frac{10}{2} = 4$$

となる。

(7)

(i) $t_\beta = 7.5$

7.5

帰無仮説 H_0: $\beta = \beta_0$ を検定するための t 検定統計量は $t_\beta = \dfrac{\hat{\beta} - \beta_0}{\sqrt{\hat{\sigma}_\beta^2}}$ であるここで $\hat{\beta}$ はβの最小二乗推定量、$\sqrt{\hat{\sigma}_\beta^2}$ は標準誤差(の推定値)である。今、$\beta_0 = 0$, $\hat{\beta} = 1.5$, $\sqrt{\hat{\sigma}_\beta^2} = 0.2$ であるので t値(t検定統計量の実現値)は$t_\beta = 1.5/0.2 = 7.5$ である。

(ii) 棄却する

この回帰モデルには定数項が存在し、定数項以外の説明変数の数が 1 つなので、t 検定統計量は自由度 $n - 2$ の t 分布に従う。今 $n = 25$ なので $t_\beta \sim t(23)$ である。有意水準 1%で検定するという事は、分布表より、$|t_\beta| > 2.807$ であれば棄却するという事になるので、この場合帰無仮説は棄却となる。

(iii) 棄却できない

この場合の t 値は $t_\beta = \dfrac{1.5 - 1.7}{0.2} = -1$ となるので、帰無仮説は棄却できない(採択する)となる。