회귀분석입문(1분반) 과제1 : 소프트웨어학부 컴퓨터과학전공 2016133 이유진

#2.1 : 소수점아래 5자리까지 표기, 반올림

$\Sigma x_i = 1285.21$, $\Sigma y_i = 53.7$, $\Sigma x_i^2 = 140168.7$, $\Sigma y_i^2 = 248.29$, $\Sigma x_i y_i = 5880.88$

$n = 12$

$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - (\Sigma x_i)^2/n = 140168.7 - (1285.21)^2/12 = 2521.63799$ ✓

$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)/n = 5880.88 - (1285.21)(53.7)/12 = 129.56525$ ✓

$S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - (\Sigma y_i)^2/n = 248.29 - (53.7)^2/12 = 7.98250$ ✓

#2.1.1

$\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}} = \dfrac{129.56525}{2521.63799} = 0.05138$ ✓

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \Sigma y_i/n - \hat{\beta}_1(\Sigma x_i/n) = \dfrac{53.7}{12} - 0.05138 \times \dfrac{1285.21}{12} = -1.02784$ ✓

$\Rightarrow \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = -1.02784 + 0.05138x$

#2.1.2

49/50

$k_1 = (x_1 - \bar{x})/S_{xx} = \dfrac{116.37 - 1285.21/12}{2521.63799} = 0.00368$ ✓

$k_2 = (x_2 - \bar{x})/S_{xx} = \dfrac{82.77 - 1285.21/12}{2521.63799} = -0.00965$ ✓

$\vdots$

$k_{12} = (x_{12} - \bar{x})/S_{xx} = \dfrac{122.30 - 1285.21/12}{2521.63799} = 0.00603$ ✓

$k_1 y_1 + k_2 y_2 + \cdots + k_{12} y_{12} = \dfrac{\sum^{12}(x_i - \bar{x})y_i}{S_{xx}} = 0.05138 = \hat{\beta}_1$ ✓

#2.1.3

$\hat{y}_1 = \hat{\beta}_0 + \hat{\beta}_1 x_1 = -1.02784 + 0.05138 \times 116.37 = 4.95125$

$\hat{y}_2 = \hat{\beta}_0 + \hat{\beta}_1 x_2 = -1.02784 + 0.05138 \times 82.77 = 3.22488$

$\vdots$

$\hat{y}_{12} = \hat{\beta}_0 + \hat{\beta}_1 x_{12} = -1.02784 + 0.05138 \times 122.30 = 5.25593$

$\downarrow$

$e_1 = y_1 - \hat{y}_1 = 5.6 - 4.95125 = 0.64875$

$e_2 = y_2 - \hat{y}_2 = 3.2 - 3.22488 = -0.02488$

$\vdots$

$e_{12} = y_{12} - \hat{y}_{12} = 5.4 - 5.25593 = 0.14407$

$\downarrow$

$SSE = e_1^2 + e_2^2 + \cdots + e_{12}^2 = 1.32541$

#2.1.4

$SSE = S_{yy} - \hat{\beta}_1 S_{xy} = 7.98250 - 0.05138 \times 129.56525 = 1.32543$ (반올림으로 인한 오차 있음)

#2.1.5

$\hat{\sigma}^2 = \dfrac{SSE}{n-2} = \dfrac{1.32543}{12-2} = 0.13254$ (#2.1.4 값 이용)

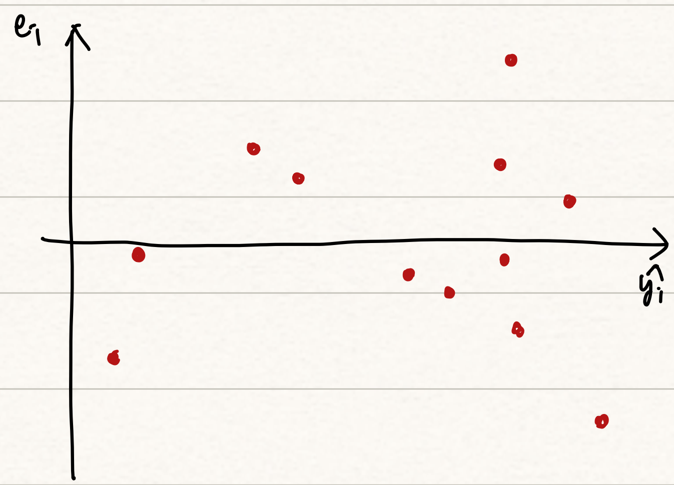
#2.1.6

$\sum e_i = e_1 + e_2 + \cdots + e_{12} = -9.8 \times 10^{-6} = -0.0000098 \approx 0$ (반올림으로 인한 오차 있음)

산점도

(x축: $\hat{y}_i$, y축: e_i)

선형관계를 찾을 수 없다



#2.3

#2.3.1

$\Sigma x_i = 3450.2$, $\Sigma y_i = 426$, $\Sigma x_i^2 = 700759$, $\Sigma y_i^2 = 10821$, $\Sigma x_i y_i = 86735.5$

$n = 17$, $\bar{x} = (\Sigma x_i)/n = 202.95294$, $\bar{y} = (\Sigma y_i)/n = 25.05882$

$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - (\Sigma x_i)^2/n = 700759 - (3450.2)^2/17 = 530.76235$ ✓

$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)/n = 86735.5 - \frac{3450.2 \times 426}{17} = 277.54706$ ✓

$S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - (\Sigma y_i)^2/n = 10821 - 426^2/17 = 145.94118$ ✓

$\Downarrow$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{277.54706}{530.76235} = 0.52292$ ✓

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 25.05882 - 0.52292 \times 202.95294 = -81.06933$ ✓

$SSE = S_{yy} - \hat{\beta}_1 S_{xy} = 145.94118 - 0.52292 \times 277.54706 = 0.80627$ ✓

$\downarrow$

$\hat{\sigma}^2 = \frac{SSE}{n-2} = \frac{0.80627}{17-2} = 0.05375$ ✓

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = -81.06933 + 0.52292 \cdot x_i$ ✓

$e_i = y_i - \hat{y}_i = y_i + 81.06933 - 0.52292 \cdot x_i$ ⇒ 모든 i에 대해 e_i 제시필요 −0.5

#2.3.2

$\Sigma x_i = 3450.2$, $\Sigma y_i = 2373.29$, $\Sigma x_i^2 = 700759$, $\Sigma y_i^2 = 331751.6$, $\Sigma x_i y_i = 482141.5$

$n = 17$, $\bar{x} = (\Sigma x_i)/n = 202.95294$, $\bar{y} = (\Sigma y_i)/n = 139.60529$

$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - (\Sigma x_i)^2/n = 700759 - (3450.2)^2/17 = 530.76235$ ✓

$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)/n = 482141.5 - \frac{3450.2 \times 2373.29}{17} = 475.31424$ ✓

$S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - (\Sigma y_i)^2/n = 331751.6 - (2373.29)^2/17 = 427.75152$ ✓

$\Downarrow$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{475.31424}{530.76235} = 0.89553$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 139.60529 - 0.89553 \times 202.95294 = -42.14516$$

$$SSE = S_{yy} - \hat{\beta}_1 S_{xy} = 427.75152 - 0.89553 \times 475.31424 = 2.09336$$

↓

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = \frac{2.09336}{17-2} = 0.13956$$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = -42.14516 + 0.89553 \cdot x_i$$

$e_i = y_i - \hat{y}_i = y_i + 42.14516 - 0.89553 \cdot x_i$ ⇒ 모든 i에 대해 e_i 제시 필요 -0.5

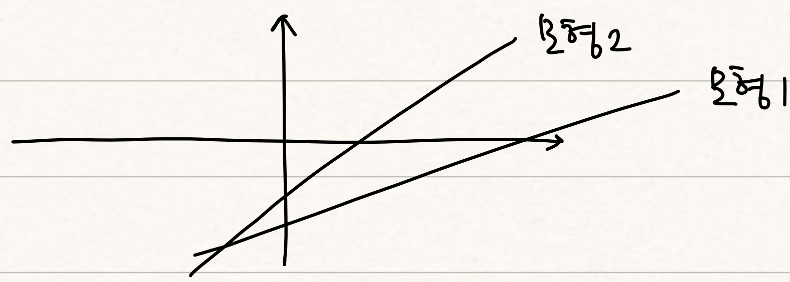
#2.3.3

두 모형의 비교 ; 설명변수가 '온도' 일때

모형1) 반응변수가 '압력' 인경우 → 추정회귀식 $\hat{y}_i = -81.06933 + 0.52292 \cdot x_i$

모형2) 반응변수가 ' $100 \times \log_{10}$(압력) ' 인 경우 → 추정회귀식 $\hat{y}_i = -42.14516 + 0.89553 \cdot x_i$

⇓



모형1보다 log를 취해 보정된 반응변수를 갖는 모형2의 추정회귀식의 기울기 (β_1)가 더 크므로 모형2가 설명변수에 따른 반응변수 값의 차이를 더 드러낸다.