과제 Review

3.1 (예 1.1 계속) 포도수확량 자료

3.1.1 단순선형회귀모형에 대한 분산분석표와 R^2의 값을 구하라.

3.1.2 $\hat{\beta}_0$과 $\hat{\beta}_1$의 표준오차값을 구하라.

3.1.3 β_0와 β_1의 95% 신뢰구간을 구하라.

3.1.4 6월의 포도 열매개수가 100개로 나타났다. 가을철 포도의 수확량의 기대값과 예측값의 95% 신뢰구간을 구하라.

3.1.5 σ^2의 95% 신뢰구간을 구하라.

32

과제 Review

3.1 $\sum x_i = 1285.21$ $\sum y_i = 53.7$ $\sum x_i^2 = 140168.7$ $\sum y_i^2 = 248.29$ $\sum x_i y_i = 5880.88$

$\bar{x} = 107.101$ $\bar{y} = 4.475$ $S_{xx} = 2521.64$ $S_{xy} = 129.564$ $S_{yy} = 7.98250$

$\hat{\beta}_1 = 0.05138$ $\hat{\beta}_0 = -1.0279$

$SSE = 1.3254$ $\hat{\sigma}^2 = MSE = 0.13254$ $SSR = (0.05138)(129.564) = 6.6570$

Source	df	SS	MS	F^*
Regression	1	6.6571	6.6571	50.2265
Residual	10	1.3254	0.1325	
Total	11	7.9825		

$$R^2 = \frac{6.6571}{7.9825} = 0.8340$$

33

과제 Review

3.1.2 $s.e.(\hat{\beta}_1) = \sqrt{MSE/S_{xx}} = \sqrt{0.13254/2521.64} = \sqrt{.00005256} = .007250$

$$s.e.(\hat{\beta}_0) = \sqrt{MSE\left(\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right)} = \sqrt{.13254\left(\frac{1}{12}+\frac{107.1^2}{2521.64}\right)} = \sqrt{0.6139} = 0.7835$$

3.1.3 $\hat{\beta}_1 \pm t(.025,10)\ s.e.(\hat{\beta}_1) = 0.0514 \pm 2.228 \times 0.007250 = 0.0514 \pm 0.0162$

$\Rightarrow (0.352, 0.676)$

$\hat{\beta}_0 \pm t(.025,10)\ s.e.(\hat{\beta}_0) = -1.0279 \pm 2.228 \times 0.7835 = -1.0279 \pm 1.7456$

$\Rightarrow (-2.7735, 0.7177)$

34

과제 Review

3.1.4 $\hat{E}(y|X=100) = -1.0278 + 0.05138 \times 100 = 4.1102 = \hat{y}_0$

$$\hat{y}_0 \pm t(.025,10)\sqrt{MSE\left(\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right)} = 4.1102 \pm 2.228 \times \sqrt{.13254\left(\frac{1}{12}+\frac{(100-107.1)^2}{2521.64}\right)}$$

$= 4.1102 \pm 0.261 \quad \Rightarrow \quad (3.849, 4.371)$

$$\hat{y}_0 \pm t(.025,10)\sqrt{MSE\left(1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right)} = 4.1102 \pm 2.228 \times \sqrt{.13254\left(1+\frac{1}{12}+\frac{(100-107.1)^2}{2521.64}\right)}$$

$= 4.1102 \pm 0.852 \quad \Rightarrow \quad (3.258, 4.962)$

3.1.5

$$\left(\frac{SSE}{\chi^2(.025,10)}, \frac{SSE}{\chi^2(.975,10)}\right) = \left(\frac{1.3254}{20.483}, \frac{1.3254}{3.247}\right) = (0.0647, 0.4082)$$

35

과제 Review

3.2 산소섭취량 자료

다음의 자료는 24명의 중년남자를 대상으로 2마일을 뛰게 한 후 걸린 시간 X와 이 기간 중 섭취한 최고산소량 Y를 기록한 것이다.

최고 산소 섭취량	시간 (초)	최고 산소 섭취량	시간 (초)	최고 산소 섭취량	시간 (초)
42.33	918	36.23	1045	53.29	743

자료원 : Ribisl과 Kachdadorian(1969)

3.2.1 최고 산소섭취량과 시간의 산점도를 그려라.

3.2.2 모형 $Y=\beta_0+\beta_1 X+\varepsilon$을 적합시키고 회귀계수의 추정값, $\hat{\sigma}$ 및 R^2의 값을 구하라.

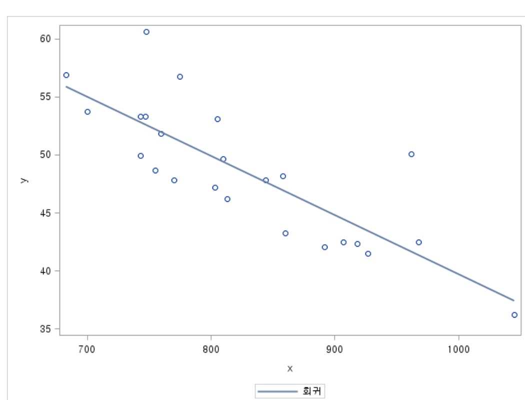
3.2.3 산점도의 형태로 보아 Y는 X의 역의 관계에 놓여 있다고 보일 것이다. 반응변수의 값에 역을 취한 새로운 반응변수와 기존의 설명변수 사이의 모형 $\frac{1}{Y}=\beta_0+\beta_1 X+\varepsilon$을 적합시키고 이 모형에서의 회귀계수의 추정값, $\hat{\sigma}$ 및 R^2의 값 등을 3.2.2의 결과와 비교하라.

36

과제 Review

3.2

3.2.1



37

과제 Review

3.2.2 $\bar{x}=826.5$ $\bar{y}=48.5467$ $S_{xx}=197658$ $S_{yy}=784.239$ $S_{xy}=-10086.2$

$\hat{\beta}_0=90.70$ $\hat{\beta}_1=-0.0510$ $SSE=784.239-(-0.0510)(-10086.2)=269.84$

$\hat{\sigma}^2=MSE=284.84/22=12.265$ $\hat{\sigma}=\sqrt{12.265}=3.50$

$R^2=1-269.84/784.239=1-0.344=0.656$

3.2.3 $\bar{x}=826.5$ $\bar{y}=0.0208947$ $S_{xx}=197658$ $S_{yy}=0.000155186$ $S_{xy}=4.62803$

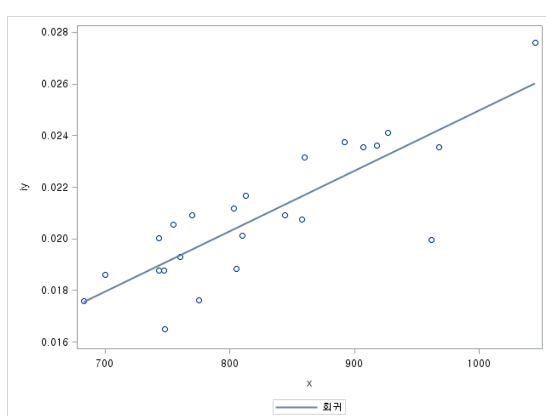
$\hat{\beta}_0=0.00154$ $\hat{\beta}_1=-0.00002341$ $SSE=0.00004684$

$\hat{\sigma}^2=MSE=0.00004684/22=0.000002129$ $\hat{\sigma}=\sqrt{0.000002129}=0.0015$

$R^2=1-0.00004684/0.000155186=1-0.302=0.698$

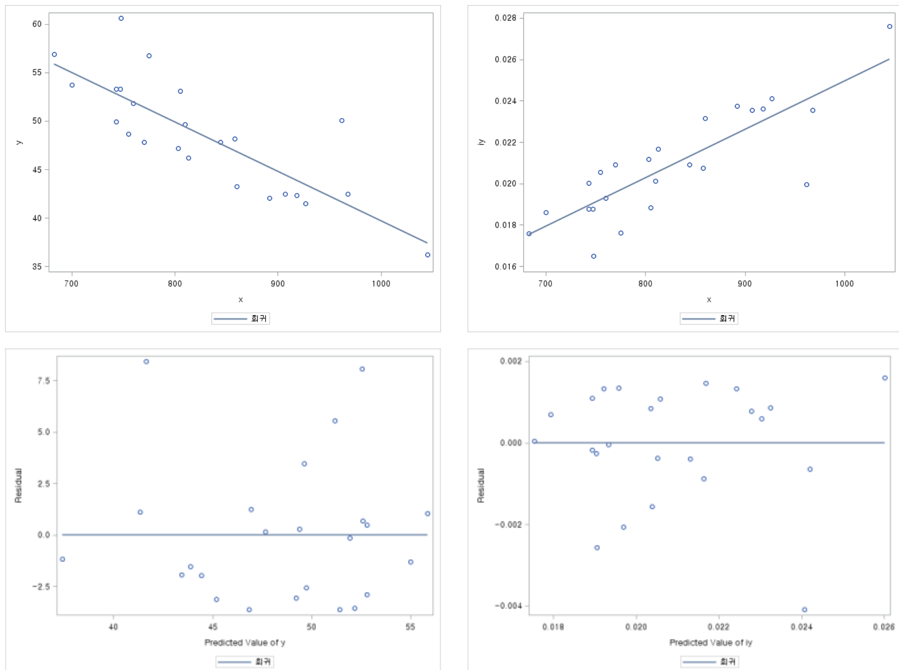
38

과제 Review



39

과제 Review



40

과제 Review

3.4 공해물질 자료

다음의 자료는 공해물질이 함유된 유화액에 물고기를 투입하고 공해물질의 농도와 물고기의 생존시간(단위 : 분)을 측정하였다. 공해물질의 농도와 생존시간에 상용로그를 취한 값을 각각 설명변수(X)와 반응변수(Y)로 한다.

$\log_{10}$(공해물질농도)	$\log_{10}$(생존시간)		
5	2.516	2.572	2.438
4.8	2.621	2.742	2.689
4.6	2.830	2.910	2.983
4.4	3.175	3.056	3.095
4.2	3.332	3.221	3.293
4.0	3.447	3.523	3.551

자료원 : Nagasawa 등(1964)

3.4.1 단순선형모형의 적합성 여부를 검정하라.

3.4.2 분산분석표, β_0와 β_1의 95% 신뢰구간을 구하라.

41

과제 Review

3.4 $\sum x_i = 81$ $\sum y_i = 53.994$ $\sum x_i^2 = 366.6$ $\sum y_i^2 = 164.106$ $\sum x_i y_i = 240.877$

$\bar{x} = 4.5$ $\bar{y} = 3.000$ $S_{xx} = 2.1$ $S_{yy} = 2.142$ $S_{xy} = -2.096$

$\hat{\beta}_1 = -2.096/2.1 = -0.9981$ $\hat{\beta}_0 = 3.000 - (-0.9981)(4.5) = 7.491$

x	n_i		y		$\bar{y}$	$\Sigma(y-\bar{y})^2$
5.0	3	2.516	2.572	2.438	2.509	0.00906
4.8	3	2.621	2.742	2.689	2.684	0.00736
4.6	3	2.830	2.910	2.983	2.908	0.01171
4.4	3	3.175	3.056	3.095	3.109	0.00736
4.2	3	3.332	3.221	3.293	3.282	0.00634
4.0	3	3.447	3.523	3.551	3.507	0.00579
합계	18					0.04762

$SSE = S_{yy} - \hat{\beta}_1 S_{xy} = 2.142 - (-0.9981)(-2.096) = 2.142 - 2.092 = 0.050$

$SS_{PE} = 0.04762$ $SS_{LOF} = SSE - SS_{PE} = 0.050 - 0.04762 = 0.00238$

$F^* = \dfrac{SS_{LOF}/(m-2)}{SS_{PE}/(n-m)} = \dfrac{0.00238/(6-2)}{0.04762/(18-6)} = 0.150$ $F(0.05; 4, 12) = 3.26$

42