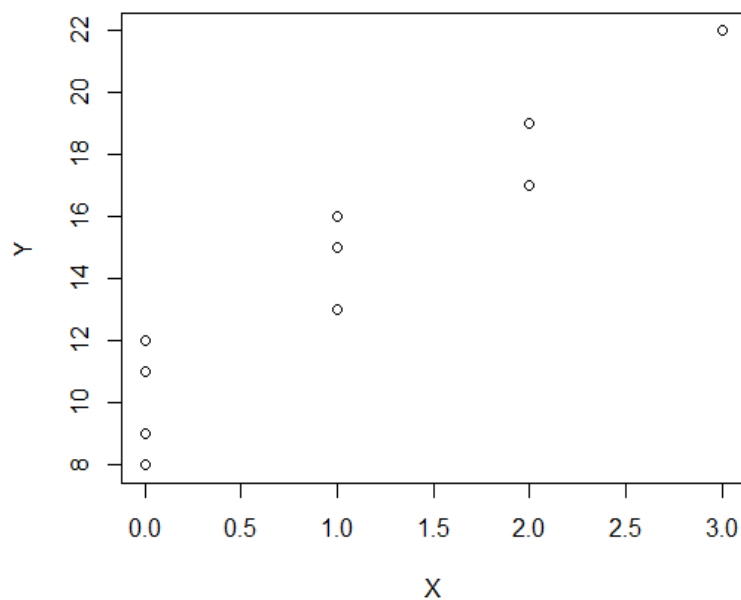


확률과통계 11주차 모범답안

1. 바이오생명 연구에 사용되는 물질이 담긴 앰플을 비행기를 이용하여 운송한다. 1 박스에 1000개의 앰플이 들어있다. 독립변수 X 는 비행기의 환승 수를 나타내고, 종속변수 Y 는 도착했을 때 부서진 앰플 수를 나타낸다. 10회의 운송 데이터는 다음과 같다. 다음의 질문에 답하라. (Kutner, 2005)

i	1	2	3	4	5	6	7	8	9	10
x_i	1	0	2	0	3	1	0	1	2	0
Y_i	16	9	17	12	22	13	8	15	19	11

- (1) 산점도를 작성하라. (2점)



- (2) 회귀분석을 위한 통계적모형을 기술하고 필요한 가정을 제시하라. (3점)

$$Y = \beta_0 + \beta_1 X + \epsilon_i$$

$$\epsilon_i \sim iid N(0, \sigma^2)$$

- (3) 회귀계수를 추정하고 회귀식을 제시하라. (3점)

$$\hat{\beta}_0 = 10.2, \hat{\beta}_1 = 4$$

$$\hat{Y}_i = 10.2 + 4x_i$$

- (4) 기울기에 대한 가설검정을 유의수준 0.05에서 실시하라. (3점)

$$H_0 : \beta_1 = 0, H_1 : \beta_1 \neq 0$$

$$P\text{-value} = 2.75e - 5$$

$2.75e - 5 < 0.025$ 이므로 귀무가설 기각

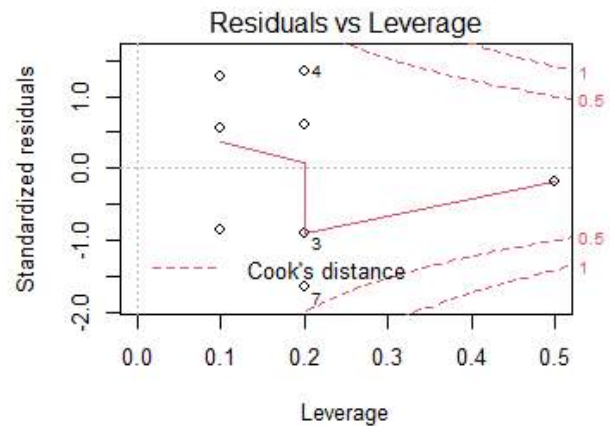
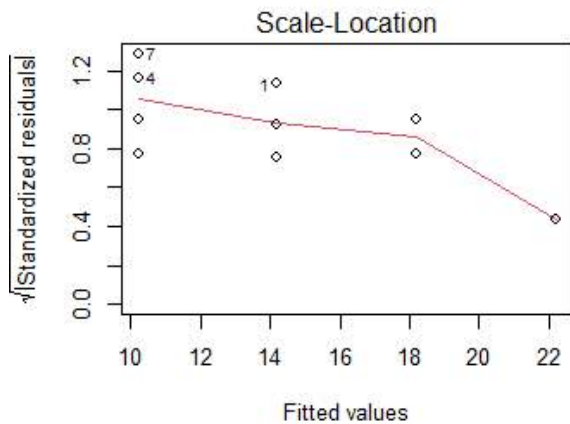
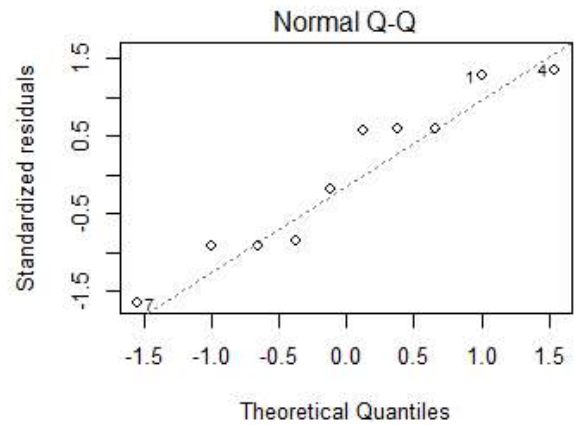
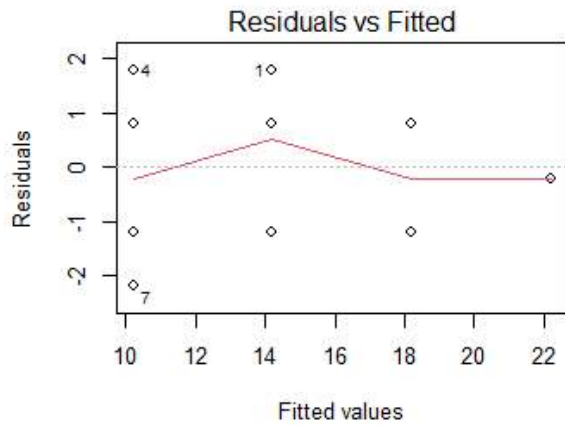
$\therefore$ 따라서 x 는 유의미한 독립변수이다.

(5) 결정계수를 구하라. (3점)

Multiple R-squared: 0.9009

Adjusted R-squared: 0.8885

(6) 회귀진단을 실시하라. (3점)



회귀진단 결과를 해석하면 선형성, 정규성, 독립성을 만족하는 것으로 보이나 일부 아웃라이어의 영향을 받는 것으로 보인다. 그러나 등분산성에는 문제가 있다. Cook's distance는 1보다 작으므로 influential case는 없다.

(7) 만약 환승을 1회 한다면 손상된 애플은 몇 개 인지 예측하라. (3점)

$$\hat{Y}_i = 10.2 + 4x_i$$

$$x_i = 1$$

$$\hat{Y}_i = 10.2 + 4 \cdot 1 = 10.2 + 4 = 14.2$$

∴ 손상된 애플은 약 15개가 된다.