

Regression HW 1

CH03, CH04

1. 원점을 지나는 회귀모형은 다음과 같이 정의할 수 있다.

$$y_i = \beta_1 x_i + \epsilon_i, \quad \epsilon_i \sim_{i.i.d.} N(0, \sigma^2), \quad i = 1, \dots, n$$

오차제곱합을 정의하고 β_1 의 최소제곱추정량 ($\hat{\beta}_1$)을 구하여라.

2. 자동차의 무게가 무거우면 이를 움직이는 데 더 많은 연료가 소모된다는 것은 알려진 사실이다. 자동차의 무게와 자동차를 1km 움직이는 데 필요한 에너지량과의 함수관계를 정확히 판단하기 위하여 A 자동차회사는 다음의 자료를 실험을 통하여 얻었다. 실험 비용이 많이 드는 관계로 9번만 실험하였다.

무게 X (단위: 1,000kg)	에너지 소모량 Y (단위: 1,000Btu)	무게 X (단위:1,000kg)	에너지 소모량 Y (단위 : 1,000Btu)
0.9	2.0	1.7	3.2
1.3	2.6	0.7	1.8
2.1	4.3	1.2	2.3
2.5	5.8	1.6	3.0
2.4	5.1		

- (1) 이 데이터의 산점도를 그리시오.
- (2) 최소제곱법의 의한 회귀직선을 적합시키시오.
- (3) 데이터의 산점도를 그리고 추정한 회귀직선을 (1)에서 그린 산점도 위에 그리시오.
- (4) 결정계수와 상관계수를 구하시오.
- (5) 분산분석표를 작성하고 회귀직선의 유의 여부를 검정하시오 (유의수준 $\alpha = 0.05$ 사용).
- (6) β_0, β_1 에 대한 90% 신뢰구간을 구하시오.
- (7) $H_0 : \beta_1 = 1$ vs. $H_1 : \beta_1 \neq 1$ 의 가설검정을 유의수준 $\alpha = 0.1$ 에서 수행하시오.

- (8) 무게가 3,000kg 이 되는 차량의 평균 에너지 소모량을 예측하시오. 이것은 무게가 1,000kg이 되는 차량의 에너지 소모량의 몇 배인가?
 - (9) 무게가 3,000kg 이 되는 차량의 평균 에너지 소모량과 하나의 개별 y 값의 90% 신뢰구간을 각각 구하시오.
 - (10) 잔차 $e_i = y_i - \hat{y}_i$ 를 구하고 잔차의 합이 0 임을 확인하시오.
 - (11) 잔차들의 x_i 에 대한 가중합, $\sum x_i e_i$ 를 구하시오.
 - (12) 잔차들의 \hat{y} 에 대한 가중합 $\sum \hat{y}_i e_i$, 를 구하시오.
 - (13) 원점을 지나는 회귀직선을 구하시오.
 - (14) 위 회귀직선에서 회귀계수(기울기)의 90% 신뢰구간을 구하시오.
 - (15) 원점을 지나는 회귀직선의 결정계수를 구하시오.
 - (16) 원점을 포함한 회귀직선과 포함하지 않은 회귀직선의 결과를 비교하여라.
3. (강의노트 CH04, p9) 다음이 성립함을 증명하시오.

$$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{(xx)}}\right)\right)$$