Answer Key for Problem Set 5

계량경제학 남 준우 교수

1.

(1) False. 독립변수의 분산이 크다는 것은 $\sum_{i=1}^{n}(X_i - \bar{X})^2$ 이 크다는 것으로 $V(b_2) = \dfrac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$ 의 값을 감소시킨다. 따라서 회귀계수 β_2 에 대한 신뢰구간은 다른 조건이 일정할 때 좁아진다.

(2) False. 오차항 ε_i 의 정규분포 여부는 최소자승추정량의 분포와 상관이 있는 것이며 최소자승추정량이 BLUE 인 것과는 아무런 상관이 없다. 최소자승추정량이 BLUE 이기 위해서는 [가정 1]~[가정 6]만으로 충분하다.

(3) False. 예측에 있어서 회귀계수 β_1, β_2 와 미래의 독립변수 값을 안다고 하더라도 미래의 종속변수에 영향을 미치는 오차항의 값에 대한 불확실성이 존재하므로 정확한 예측은 불가능하다. 또한 이는 $V(e_f) = V(\varepsilon_f) = \sigma^2 \neq 0$ 인 것을 의미한다.

(4) False. 독립변수의 분산이 크다는 것은 (1)번에서 보듯이 $V(b_2)$ 의 값을 감소시키는 것으로 이는 최소자승추정량 b_2 의 정확도가 커짐을 의미한다.

(5) False. 독립변수의 값이 $\bar{X}$ 에 가까울수록 $\sum_{i=1}^{n}(X_i - \bar{X})^2$ 값은 작아지며 따라서 $V(b_2)$ 가 커져 기울기에 대한 정확도가 떨어진다.

(6) True. 오차항의 분산 (σ^2)이 클수록 단순회귀분석에서 최소자승추정량의 분산 $V(b_1) = \sigma^2 \dfrac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2}$, $V(b_2) = \dfrac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$ 가 커진다.

(7) False. 유의확률(p-value)의 값이 크면 클수록 추정치의 값이 0 이라는 가설을 기각하지 못하게 된다.

(8) True. 만약 오차항의 분산이 정규분포를 가진다면 최소자승추정량도 정규분포를 가지지만, 최소자승추정량의 분산값에는 아무런 상관이 없다.

2 .

타당하지 않다. t-ratio를 구해보면 그 값이

$\frac{0.000045}{0.000015} = 3$

이며, p-value의 값이 실질적으로 0이므로 계수가 0과 유의적으로 다르다. 즉 변수 X의 영향은 유의적이다.

이는 제6장에서 설명하는 바와 같이 회귀분석에서 독립변수와 종속변수의 단위를 변환함에 따라 회귀계수의 추정치의 값은 달라지나 t-통계치와 유의확률, 결정계수의 값은 변하지 않는다. 이러한 바는 어떤 특정 독립변수가 유의적인지 아닌지는 회귀계수의 추정치의 크기가 아니라 t-통계치, 유의확률 등의 통계적인 추론을 통해 설명되기 때문이다.

3.

(1) $b_2 = 0.00060$

은 기온이 1도 상승함에 따라 평균적으로 오존농도는 0.0006 ppm상승함을 의미.

(2) $R^2 = 0.362$

는 이 모형에서 기온의 변화가 설명하는 오존농도의 변화는 약 36.2%에 해당한다는 것을 의미한다.

(3)

$t-\text{통계치}=\frac{b_2}{s_{b_2}}$, $3.447=\frac{0.00060}{(A)}$
에서 0.00017을 구한다.
(4) 2-t 실용기준에 의하면 t-통계치 값이 3.447로 절대값에 있어서 2보다 크므로 기온은 오존농도 발생에 중요한 변수로 판단된다.

(5) β_2에 대한 유의확률이 0.0024로 유의수준 0.05보다 작으므로 $H_0:\beta_2=0$의 가설을 기각한다.

(6) $H_0:\beta_2\geq 0$

에 대한 대립가설은

$H_1:\beta_2<0$

으로 이는 단측검정에 해당한다. 이 경우 관찰치 23개로 부터 자유도는 21이며 위 단측검정에 대한 5% 임계치는 –1.721이다. t-통계치=3.447>임계치=-1.721이므로 귀무가설을 기각하지 못한다.

(7) 식 4.16에 따라 $s^2=\frac{1}{n-2}\sum_{i=1}^{n}e_i^2$의 관계가 성립하므로

$$\sum_{i=1}^{n}e_i^2=(n-2)\times s^2=(23-2)\times 0.0074^2=0.00115$$

를 구한다.

(8) $\widehat{O3}_{94:12}=0.0074+0.00060\,TEM_{94:12}=0.0074+0.00060\times(-2)=0.0062$

로 예측된다.

4.

(1) $\overline{Price}$=1093.655, S_{price}=535.7035, $\overline{Mileage}$=45971.545, $S_{mileage}$=32380.01

(2) $\gamma_{price,mileage}$= -0.436 : Price 와 Mileage 사이에는 음의 상관관계가 있다.

(3) b_1= 1425.145 : 주행거리가 0Km 일 때 중고차의 판매가격은 평균적으로

1425 만원이다.
(4) b_2 = -0.0072 : 주행거리가 1Km 가 증가할 때 중고차의 판매가격이 평균적으로 0.0072 만원 하락한다.

(5) 본 회귀식은 절편을 포함하며 정규방정식에 의해 $\sum_{i=1}^{n} e_i = 0$ 이 성립하므로 표본회귀선이 $(\bar{X}, \bar{Y})$ 를 지난다.

(6) β_2 에 대한 95% 신뢰구간은

$$\left(b_2 \pm t(143, \frac{0.05}{2}) \bullet SE(b_2) \right) = (-0.007211 \pm 1.98 \times 0.001245) = (-0.01, -0.0047)$$

(7) $R^2 = 0.19$: 주행거리변수가 중고차의 판매가격변동을 19% 설명한다.

(8) 일치한다. (Problem set 3 참조)

(9) $\widehat{Y_f} = 1425.145 + (-0.007211) \times 20000 = 1280.925$