알찬대학 따뜻한 동행

The Best Glocal University

고급회귀분석론

이영미 교수님

학 과	통계학과
학 번	202250926
이 름	김보람
과 제 명	HW 03
제 출 일	2022.11.21.(월)

전북대학교

JEONBUK NATIONAL UNIVERSITY

1(1) $X = \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 6 \end{pmatrix} \quad y = \begin{pmatrix} 39 \\ 24 \\ \vdots \\ 126 \end{pmatrix}$

$$X^T X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 3 & 1 & \cdots & 6 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 14 & 62 \\ 62 & 360 \end{pmatrix}$$

$$X^T y = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 3 & 1 & \cdots & 6 \end{pmatrix} \begin{pmatrix} 39 \\ 24 \\ \vdots \\ 126 \end{pmatrix} = \begin{pmatrix} \Sigma y \\ \Sigma xy \end{pmatrix} = \begin{pmatrix} 1253 \\ 6714 \end{pmatrix}$$

$$y^T y = (39 \quad 24 \quad \cdots \quad 126) \begin{pmatrix} 39 \\ 24 \\ \vdots \\ 126 \end{pmatrix} = \Sigma y^2 = 138197$$

$$(X^T X)^{-1} = \begin{pmatrix} \frac{360}{14\cdot 360-(360)^2} & \frac{-62}{14\cdot 360-(360)^2} \\ \frac{-62}{14\cdot 360-(360)^2} & \frac{14}{14\cdot 360-(360)^2} \end{pmatrix} \doteqdot \begin{pmatrix} 0.3010 & -0.0518 \\ -0.0518 & 0.0117 \end{pmatrix}$$

(2) $$\hat{\beta} = (X^T X)^{-1} X^T y = \begin{pmatrix} 0.3010 & -0.0518 \\ -0.0518 & 0.0117 \end{pmatrix} \begin{pmatrix} 1253 \\ 6714 \end{pmatrix} = \begin{pmatrix} 29.10702 \\ 13.63712 \end{pmatrix}$$

$$\hat{y} = 29.10702 + 13.63712 x_1$$

(3) $H_0 : \beta_1 = \beta_2 = 0$ vs $H_1 : \beta_1 \neq \beta_2$

$$SST = y^T y - n(\bar{y})^2 = 138197 - 14(89.5)^2 = 26053.5$$

$$SSR = \beta^T X^T y - n(\bar{y})^2$$

$$= (29.10702 \quad 13.63712) \begin{pmatrix} 1253 \\ 6714 \end{pmatrix} - 14 \cdot (89.5)^2$$

$$= 128030.7 - 112143.5 = 15887.2$$

$$SSE = SST - SSR = 10166.3$$

1 (3) 이어서..

ANOVA

요인	제곱합	자유도	평균제곱	F_0	$F(0.05)$
SSR	15887.2	1	15887.2	18.75	4.75
SSE	10166.3	12	847.2		
SST	26053.5	13			

$MSE = 847.2$

$$\widehat{var}(\hat{\beta}) = (X^T X)^{-1} MSE$$

$$= \begin{pmatrix} 255.01003 & -43.918395 \\ -43.918395 & 9.917057 \end{pmatrix}$$

2. $X = \begin{pmatrix} 1 & 195 & 57 \\ 1 & 179 & 61 \\ \vdots & \vdots & \vdots \\ 1 & 209 & 61 \end{pmatrix} \qquad y = \begin{pmatrix} 81.4 \\ 122.2 \\ \vdots \\ 113.8 \end{pmatrix}$

(1) $X^TX = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 195 & 179 & \cdots & 209 \\ 57 & 61 & \cdots & 61 \end{pmatrix} \begin{pmatrix} 1 & 195 & 57 \\ 1 & 179 & 61 \\ \vdots & \vdots & \vdots \\ 1 & 209 & 61 \end{pmatrix} = \begin{pmatrix} 8 & 1587 & 474 \\ 1587 & 315745 & 94108 \\ 474 & 94108 & 315745 \end{pmatrix}$

$X^Ty = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 195 & 179 & \cdots & 209 \\ 57 & 61 & \cdots & 61 \end{pmatrix} \begin{pmatrix} 81.4 \\ 122.2 \\ \vdots \\ 113.8 \end{pmatrix} = \begin{pmatrix} 901.9 \\ 179676.4 \\ 54034.3 \end{pmatrix}$

$(X^TX)^{-1} = \begin{pmatrix} 82.87 & -0.135 & -0.946 \\ -0.135 & 0.001 & -0.002 \\ -0.946 & -0.002 & 0.022 \end{pmatrix}$

$\hat{\beta} = (X^TX)^{-1}X^Ty = \begin{pmatrix} -554.311 \\ -0.1797 \\ 11.85999 \end{pmatrix}$

$\hat{y} = -554.311 - 0.1797\,x_1 + 11.85999\,x_2$

(2) $\mathrm{Var}(\hat{\beta}) = (X^TX)^{-1}\sigma^2 = (X^TX)^{-1} \times 3$

$= \begin{pmatrix} 248.625 & -0.404 & -2.838 \\ -0.404 & 0.004 & -0.006 \\ -2.838 & -0.006 & 0.067 \end{pmatrix}$

$\mathrm{Var}(\hat{\beta}_0) = 248.625$

$\mathrm{Var}(\hat{\beta}_1) = 0.004$

$\mathrm{Var}(\hat{\beta}_2) = 0.067$

$\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2) = -0.006$

(3)

$\mathrm{Var}(\hat{y}_s) = [x^T(X^TX)^{-1}x]\,\sigma^2$

$= [(1\ \ 200\ \ 59)\,(X^TX)^{-1} \begin{pmatrix} 1 \\ 200 \\ 59 \end{pmatrix}] \cdot 3$

$= 0.3936204$

2-14)

$\hat{\beta}_1$은 x_2가 변하지않을때 x_1이 1만큼증가하면 y는 $\hat{\beta}_1$만큼증가한다

$\hat{\beta}_2$는 x_1이 변하지않을때 x_2가 1만큼증가하면 y는 $\hat{\beta}_2$만큼 증가한다.

(5) $H_0 : \beta_1 = \beta_2 = 0$ vs $H_1 : \beta_1 \neq \beta_2 \neq 0$

$$SST = y^T y - n(\bar{y})^2 = 91.4^2 + 122.2^2 + \cdots + 113.8^2 - 8 \cdot (112.7375)^2$$
$$= 110959 - 101678 = 9281.079$$

$$SSR = \hat{\beta}^T X^T y - n(\bar{y})^2$$
$$= \begin{pmatrix} -554.311 & -0.1797 & 11.85999 \end{pmatrix} \begin{pmatrix} 901.9 \\ 179676.4 \\ 54034.3 \end{pmatrix} - 101678$$
$$= 108618.1 - 101678 = 6940.146$$

$$SSE = SST - SSR = 2340.933$$

ANOVA	요인	df	평균제곱	F_0	$F_{0.05}(2,5)$
SSR	6940.146	2	3470.0975	7.412	5.179
SSE	2340.933	5	468.1866		
SST	9281.079	7			

$F_0 > F$ 이므로 귀무가설을 기각한다. 즉 β_1과 β_2가 동시에 0이되지 않는다.

(6) $R^2 = \dfrac{SSR}{SST} = \dfrac{6940.146}{9281.079} \approx 0.7477736$

(7) $R^2_{adj} = 1 - \dfrac{SSE/5}{SST/7} = 1 - \dfrac{468.1866}{1325.868} \approx 0.6468829$

(8) $\hat{\sigma}^2 = MSE = 468.1866$

?? (9) $E(SSR) = 2\sigma^2 + \beta^T X^T \left(I - \frac{J}{n}\right) X \beta$ ...??

$E(MSR) = E(SSR/2) = \sigma^2 + \beta^T X^T \left(I - \frac{J}{n}\right) X \beta \,/ 2$

3. $X^TX = \begin{pmatrix} 10 & 80 & 90 & 638 \\ 80 & 686 & 784 & 5085 \\ 90 & 784 & 932 & 5544 \\ 638 & 5085 & 5544 & 43324 \end{pmatrix}$ $\quad y^Ty = 3828$

$X^Ty = \begin{pmatrix} 186 \\ 1608 \\ 1866 \\ 11696 \end{pmatrix}$ $\quad (X^TX)^{-1} = \begin{pmatrix} 3.4546 & -0.0413 & -0.1050 & -0.0326 \\ -0.0413 & 0.1076 & -0.0629 & -0.0040 \\ -0.1050 & -0.0629 & 0.0461 & 0.0030 \\ -0.0326 & -0.0040 & 0.0030 & 0.0006 \end{pmatrix}$

(1) $\hat{\beta} = (X^TX)^{-1}X^Ty$

$= \begin{pmatrix} -0.9484 \\ 1.5151 \\ 0.7876 \\ 0.0053 \end{pmatrix}$ $\quad \hat{y} = -0.9484 + 1.5151x_1 + 0.7876x_2 + 0.0053x_3$

(2) $\hat{\beta}_1$: $x_2 \cdot x_3$가 변하지 않을때 x_1이 1만큼변하면 y는 β_1 (-0.09484)만큼 변화

$\hat{\beta}_2$: $x_1 \cdot x_3$가 변하지 않을때 x_2가 1만큼변하면 y는 β_2 (1.5151) 만큼 변화

$\hat{\beta}_3$: $x_1 \cdot x_2$가 변하지 않을 때 x_3가 1만큼변하면 y는 β_3(0.0053) 만큼 변화

(3) $Var(\hat{\beta}) = (X^TX)^{-1}\sigma^2$이므로, $\quad \hat{\sigma}^2 = MSE$이다.

먼저 SSE를 구하면

$SSE = \Sigma(y_i - \hat{y})^2 = y^Ty - \hat{\beta}^TX^Ty$

$= 3828 - (-0.9484 \ 1.5151 \ 0.7876 \ 0.0053)\begin{pmatrix} 186 \\ 1608 \\ 1866 \\ 11696 \end{pmatrix}$

$= 36.2776$

$MSE = SSE/6 = 6.0463$

$Var(\hat{\beta}) = (X^TX)^{-1} MSE = \begin{pmatrix} 20.8876 & -0.2494 & -0.6350 & -0.1971 \\ -0.2494 & 0.6507 & -0.3803 & -0.0240 \\ -0.6350 & -0.3803 & 0.2788 & 0.0183 \\ -1.1971 & -0.0240 & 0.0183 & 0.0035 \end{pmatrix}$

$Var(\hat{\beta}_4) = 0.0035$

3(4) $SST = y^T y - n(\bar{y})^2 = 2.8^2 + 3.9^2 + \cdots + 3.3^2 - 10\cdot(3.46)^2$

$= 368.4$

$SSE = 36.2776$

$SSR = SST - SSE = 332.1224$

ANOVA

요인	제곱합	df	평균제곱	F_0
SSR	332.1224	3	110.7075	18.31
SSE	36.2776	6	6.0463	
SST	368.4	9		

$R^2 = \dfrac{SSR}{SST} \approx \dfrac{332.1224}{368.4} \approx 0.9015$

(5) $\hat{y} = \widehat{E(y|X=x)} = x^T \hat{\beta}$

$Var(\hat{y}) = Var(X^T \hat{\beta}) = X^T Var(\hat{\beta}) x$

$= (1 \quad 20 \quad 27 \quad 60)(X^T X)^{-1} \begin{pmatrix} 1 \\ 20 \\ 27 \\ 60 \end{pmatrix} \sigma^2$

$= 3.3188 \cdot \sigma^2$

$\because$ σ^2을 모르므로 $\hat{\sigma}^2 = MSE$ 이용하면

$= 3.3188 \times MSE = 3.3188 \times 6.0463 \approx 20.0665$

(6) $Var(\hat{y}_s) = [1 + X^T (X^T X^{-1}) x]\sigma^2$이므로 $\hat{\sigma}^2 = MSE$ 이용

$= [1 + 3.3188]\hat{\sigma}^2$

$= 4.3188 \times MSE$

$= 4.3188 \times 6.0463$

$\doteqdot 26.1128$

3(7)* $E(SSR) = 3\sigma^2 + \beta^T X^T (I - \frac{J}{n}) X \beta$

(8) β_3의 95% 신뢰구간

$\beta_3 \pm (6;0.025)\sqrt{C_{33} \cdot MSE}$

$\Rightarrow 0.0053 \pm t(6;0.025)\sqrt{0.0006 \times 6.0463}$

$\Rightarrow 0.0053 \pm 2.447 \times 0.6023$

$\Rightarrow 0.0053 \pm 0.1474$

$\Rightarrow (-0.1421,\ 0.1527)$

(9) β_1의 99% 신뢰구간

$\beta_1 \pm t(6;0.005)\sqrt{C_{11} \times MSE}$

$\Rightarrow -0.9484 \pm t(6;0.005)\sqrt{0.1076 \times 6.0463}$

$\Rightarrow -0.9484 \pm 3.707 \times 0.8066$

$\Rightarrow (-3.9384,\ 2.0416)$

(10) $x_1=20,\ x_2=27,\ x_3=60$에서 $E(y)$의 95% 신뢰구간

$x_1=20,\ x_2=27,\ x_3=60$에서 $\hat{y}$는

$\hat{y} = -0.9484 + (1.5151)(20) + (0.7876)(27) + (0.0053)(60)$

$= 50.9368$

$Var(\hat{y}) = x^T (X^T X)^{-1} x \sigma^2$

$= (1\ \ 20\ \ 27\ \ 60)(X^T X)^{-1} \begin{pmatrix} 1 \\ 20 \\ 27 \\ 60 \end{pmatrix} \sigma^2$

$= 3.3188 \cdot \sigma^2 \doteq 3.3188\, \hat{\sigma}^2 = 3.3188 \cdot MSE \approx 20.0665$

$\because \hat{\sigma}^2 = MSE$ 이므로 95% 신뢰구간은

$\hat{y} \pm t(6;0.025)\sqrt{x^T (X^T X)^{-1} x \cdot MSE}$

$\Rightarrow 50.9368 \pm 2.447\sqrt{(3.3188)(6.0463)}$

$\Rightarrow 50.9368 \pm 2.447 \times 4.4796$

$\Rightarrow (39.9753,\ 61.8983)$

3(11) $H_0: \beta_1 = 0$ vs $H_1: \beta_1 > 0$ $\alpha = 0.05$

$q^T = (0, 1, 0, 0)$, $C = 0$이므로 $q^T \beta = \beta_1 = 0$

$t_0 = \beta_1 / \sqrt{q^T (X^T X)^{-1} q \cdot MSE}$

$= 1.5151 / \sqrt{0.1076 \times 6.0463}$

$= 1.5151 / 0.8066 \approx 1.8783$

$t(6; 0.025) = 2.447$이고 $t_0 < t$ 이므로 H_0를 기각할수 없다.

(12) $H_0: \beta_1 = \beta_2 = \beta_3$ $\alpha = 0.05$ 이고, H_1: not H_0.

FM: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$

$H_0: \beta_1 = \beta_2 = \beta_3$이면 $H = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$ 이다.

RM: $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_1 x_3 + \varepsilon$

$= \beta_0 + \beta_1 (x_1 + x_2 + x_3) + \varepsilon$

$= \gamma_0 + \gamma_1 z_1 + \varepsilon$

축소모형 $SST(R) = 2.8^2 + 3.9^2 + \cdots + 3.3^2 - 10 \cdot (\bar{y}) = 368.4$

$SSR(R) = \beta^T X^T y = 8.03$

$SSE(R) = SST - SSE = 360.37$

완전모형 $SSE(F) = SSE = 36.2776$

$$F_0 = \frac{SSE(R) - SSE(F)}{p} \div \frac{SSE(F)}{n-k}$$

$$= \frac{360.37 - 36.2776}{1} \div \frac{36.2776}{6} = 31.93562$$

$F(1, 6; 0.05) = 5.9874$ 이며

$F_0 > F$ 이므로 귀무가설은 기각한다.

3(13) $H_0: \beta_1 = \beta_2 + 3$ $\alpha = 0.05$로 검정.

$H_1: \beta_1 \neq \beta_2 + 3$

$$H_0: (0, 1, -1, 0)\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{pmatrix} = 3$$

$$Q = (C\beta - m)^T [C(X^TX)^{-1}C^T]^{-1} (C\beta - m)$$

$$= (\beta_1 - \beta_2 - 3)^T \left[(0, 1, -1, 0)(X^TX)^{-1}\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}\right]^{-1} (\beta_1 - \beta_2 - 3)$$

$$= \frac{(1.5151 - 0.7876 - 3)^2}{0.1076 + 0.4641 - 2(-0.0629)} \fallingdotseq 7.4040$$

검정통계량 F값은

$$F_0 = \frac{Q/P}{SSE/n-k} = \frac{7.4040/1}{36.2776/6} \fallingdotseq 1.2246$$

$\alpha = 0.05$에서 $F(1, 6; 0.05) = 4.76$

$F_0 < F$ 이므로 귀무가설을 기각할수없다

(14) $H_0: E(y) = 3.5$ vs $H_1: E(y) \neq 3.5$

$$t_0 = \frac{\hat{y} - 3.5}{\sqrt{x^T(X^TX)^{-1}x\, MSE}} \fallingdotseq \frac{50.9368 - 3.5}{\sqrt{20.0665}} \fallingdotseq 10.5896$$

$\because$ ($x_1 = 20$, $x_2 = 27$, $x_3 = 60$에서 $\hat{y} = 50.9368$)

$t(6; 0.025) = 2.447$이고 $t_0 > t$ 이므로

귀무가설을 기각한다.

4. $\sum_{j=1}^{n} Var(\hat{y}_j) = (p+1)\sigma^2$ 증명

$$\sum_{j=1}^{n} var(\hat{y}_j) = tr[Var(\hat{y}_j)]$$
$$= tr(X(X^TX)^{-1}X^T\sigma^2]$$
$$= \sigma^2 \underbrace{tr(X(X^TX)^{-1}X^T)}_{= (p+1)}$$ ← X는 $n\times(p+1)$행렬.
$$= (p+1)\sigma^2$$

5. (1) $e = [I - X(X^TX)^{-1}X^T]y$ 이고

$I - X(X^TX)^{-1}X^T = A$라고 하자. 즉 $e = Ay$

$$cov(e, y) = cov(Ay, y) = A\,cov(y, y) = A\,Var(y)$$

$Var(y) = I_n\sigma^2$이므로

$$cov(e, y) = [I_n - X(X^TX)^{-1}X^T]\sigma^2$$

(3) $cov(e, \hat{\beta})$

$e = (I_n - X(X^TX)^{-1}X^T)y$ 이고, $\hat{\beta} = X(X^TX)^{-1}X^Ty$이다.

$$cov(e, \hat{\beta}) = Cov((I_n - X(X^TX)^{-1}X^T)y,\ X(X^TX)^{-1}X^Ty)$$
$$= (I_n - X(X^TX)^{-1}X^T)(X(X^TX)^{-1}X^T)\ Cov(y, y)$$
$$= [X(X^TX)^{-1}X^T - X(X^TX)^{-1}X^TX(X^TX)^{-1}X^T]\ cov(y, y)$$
$$= O_{n\times(k+1)}$$

(6) $\sum_{j=1}^{n} e_j\hat{y}_j$

$$= \sum_{j=1}^{n} (\beta_0 + \sum \beta_i x_{ij})e_j$$
$$= \beta_0 \underbrace{\sum e_j}_{=0} + \sum \beta_j \underbrace{\sum x_{ij} e_j}_{=0} = 0$$

(2). (4). (5)는 모르겠습니다..ㅠㅠ