

제6장 선형회귀모형의 변환

6.1 가변수를 이용한 회귀분석

6.1.3 회귀모형을 이용한 분산분석

[예제 6.3]

어느 4년제 대학교에서 학년 별로 영어 실력을 비교하기 위하여 각 학년에서 6명을 무작위로 추출하여 같은 영어 시험을 보게 한 결과 <표 6.1>의 자료를 얻었다(단, 3학년에서 1명, 4학년에서 2명의 결시자가 있었다고 가정한다.) 이 자료에 대하여 회귀모형 (6.3)을 적합하고 F -검정을 하라.

<표 6.1>

학년	영어 점수	평균
1	81 75 69 90 72 83	78.3
2	65 80 73 79 81 69	74.5
3	72 67 62 76 80	71.4
4	89 94 79 88	87.5

【풀이】

□ 범주의 수: $c = 4 \rightarrow$ 필요한 가변수의 수: $c - 1 = 4 - 1 = 3$

범주 당 관측치의 수: $n_1 = 6, n_2 = 6, n_3 = 5, n_4 = 4$

□ 가변수에 대한 정의

$$Z_1 = \begin{cases} 1, & \text{1학년인 경우} \\ 0, & \text{o.w.} \end{cases}, \quad Z_2 = \begin{cases} 1, & \text{2학년인 경우} \\ 0, & \text{o.w.} \end{cases}, \quad Z_3 = \begin{cases} 1, & \text{3학년인 경우} \\ 0, & \text{o.w.} \end{cases}$$

4학년의 경우: $Z_1 = Z_2 = Z_3 = 0$

□ 회귀모형: $Y_i = \beta_0 + \beta_1 Z_{i1} + \beta_2 Z_{i2} + \beta_3 Z_{i3} + \epsilon_i, \quad i = 1, 2, \dots, 21$

$$\square Y = \begin{bmatrix} 81 \\ 75 \\ 69 \\ 90 \\ 72 \\ 83 \\ 65 \\ 80 \\ 73 \\ 79 \\ 81 \\ 69 \\ 72 \\ 67 \\ 62 \\ 76 \\ 80 \\ 89 \\ 94 \\ 79 \\ 88 \end{bmatrix}, \text{ 디자인행렬 } X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\square \text{ 적합된 회귀식: } \hat{Y} = 87.5 - 9.17Z_1 - 13Z_2 - 16.1Z_3$$

(3.513) (4.535) (4.535) (4.713)

$$- 1\text{학년 평균영어성적 추정치: } \hat{\beta}_0 + \hat{\beta}_1 = 87.5 - 9.17 = 78.33$$

$$- 2\text{학년 평균영어성적 추정치: } \hat{\beta}_0 + \hat{\beta}_2 = 87.5 - 13 = 74.5$$

$$- 3\text{학년 평균영어성적 추정치: } \hat{\beta}_0 + \hat{\beta}_3 = 87.5 - 16.1 = 71.4$$

$$- 4\text{학년 평균영어성적 추정치: } \hat{\beta}_0 = 87.5$$

⇒ 각 학년의 평균영어성적 추정치는 각 학년 표본들의 표본평균과 같다.

<표 6.1>

요인	제곱합	자유도	평균제곱	F
회귀	643.6	3	214.5	4.35
오차	839.0	17	49.4	
전체	1482.7	20		

$$\square \text{ 가설검정 } H_0: \beta_1 = \beta_2 = \beta_3 = 0, H_1: \text{not } H_0$$

$$F = 4.35, \text{ 기각치: } F_{0.05}(3, 17) = 3.20 \rightarrow H_0 \text{ 기각}$$

⇒ 이 대학교의 학년 간에는 영어 실력의 차이가 있다고 말할 수 있다.