김은영

# 제3장 중선형회귀모형

## 3.5 중회귀분석에서의 추론 Ⅰ

- 중회귀분석에서의 관심 대상 모수: 회귀계수 $\beta_0, \beta_1, \cdots, \beta_{p-1}$과 $y$의 평균값

- 오차항 $\epsilon_i$들이 서로 독립이고 모두 $N(0, \sigma^2)$의 분포를 가진다는 가정 하에

$$\epsilon_i \sim i.i.d. N(0, \sigma^2) \qquad (i = 1, \cdots, n)$$

관측치 벡터 $y$는 평균벡터가 $X\beta$이고, 분산-공분산행렬이 $I_n\sigma^2$인 $n$차원 정규분포를 갖는다.

$$y \sim N(X\beta, I_n\sigma^2)$$

### 3.5.1 회귀계수에 대한 추론

$$\hat{\beta} \sim N_p\left(\beta,\ (X^t X)^{-1}\sigma^2\right)$$

다변량정규분포에서 각 주변확률분포도 정규분포가 된다는 성질에 따라 $\hat{\beta}_{LSE}$의 각 원소는 다음과 같은 정규분포를 따른다.

$$\hat{\beta}_j \sim N\left(\beta_j,\ c_{j+1,\, j+1}\sigma^2\right) \quad j = 0,\ 1,\ 2,\ \cdots,\ p-1$$

여기서, $c_{j+1,\, j+1}$: $(X^t X)^{-1}$의 $j+1$번째 대각원소

$$\frac{\hat{\beta}_j - \beta_j}{\sigma\sqrt{c_{j+1,\, j+1}}} \sim N(0,\ 1) \quad (j = 0,\ 1,\ \cdots,\ p-1)$$

$s^2 = \dfrac{SSE}{n-p}$: $\sigma^2$의 불편추정량

$$\frac{\dfrac{\hat{\beta}_j - \beta_j}{\sigma\sqrt{c_{j+1,\, j+1}}}}{\sqrt{\dfrac{SSE/\sigma^2}{n-p}}} = \frac{\hat{\beta}_j - \beta_j}{s\sqrt{c_{j+1,\, j+1}}} \sim t(n-p) \quad (j = 0,\ 1,\ \cdots,\ p-1)$$

$\hat{\beta}$과 $SSE$ 간의 독립성 성립 여부를 확인하면, 여기서, $\hat{\beta} = (X^t X)^{-1} X^t y$, $SSE = y^t(I-H)y$

$$BVA = (I-H)\sigma^2 I\left[(X^t X)^{-1} X^t\right]^t = \sigma^2\left[I - X(X^t X)^{-1} X^t\right]\left[(X^t X)^{-1} X^t\right]^t$$

$$= \sigma^2\left[I - X(X^t X)^{-1} X^t\right]\left[(X^t X)^{-1} X^t\right]^t$$

$$= \sigma^2\left[X(X^t X)^{-1} - X(X^t X)^{-1} X^t X (X^t X)^{-1}\right] = 0$$

< $\beta_j$에 대한 가설검정 >

$H_0$ : $\beta_j = \beta_{j0}$

$H_1$ : $\beta_j > \beta_{j0}$ , $\beta_j < \beta_{j0}$ , $\beta_j \neq \beta_{j0}$

검정통계량: $t = \dfrac{\hat{\beta}_j - \beta_{j0}}{s\sqrt{c_{j+1,\, j+1}}} \sim t(n-p)$

< $\beta_j$에 대한 $100(1-\alpha)\%$ 신뢰구간 >

$$t_{\alpha/2} < \frac{\hat{\beta}_j - \beta_{j0}}{s\sqrt{c_{j+1,\, j+1}}} < t_{\alpha/2}$$

$$\left[\ \hat{\beta}_j - t_{\alpha/2} \cdot s\sqrt{c_{j+1,\, j+1}},\ \hat{\beta}_j + t_{\alpha/2} \cdot s\sqrt{c_{j+1,\, j+1}}\ \right]$$