김은영

제3장 중선형회귀모형

3.6 중회귀분석에서의 추론 Ⅱ

3.6.2 일반적 가설검정

회귀계수 β_j들에 대한 가설은 다양한 형태를 가질 수 있다.

가설이 β_j들의 임의의 선형함수로 주어지는 경우에 대한 검정방법을 살펴보자.

$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \epsilon_i$

회귀계수 β_j들에 대한 가설	$H_0\ :\ C_{(q\times p)}\beta_{(p\times 1)} = m_{(q\times 1)}$
$\beta_1 + \beta_2 - 2\beta_3 = 0$	$C = [0\ \ 1\ \ 1\ \ -2\ \ 0],\ m = 0$
$\beta_1 - \beta_3 = 0\ \&\ \beta_2 - 2\beta_3 = 3$	$C = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{bmatrix},\ m = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$
$\beta_3 = \beta_4 \Leftrightarrow \beta_3 - \beta_4 = 0$	$C = [0\ \ 0\ \ 0\ \ 1\ \ -1],\ m = 0$

$H_0 : C\beta = m$을 검정하기 위한 통계량을 구해 보자.

$H_0 : C\beta = m$이 옳다는 가정 하에 β에 대한 최소제곱추정량(LSE)을 계산해 보자.

$\underset{\beta}{\arg\min}(y - X\beta)^t(y - X\beta)\quad s.t.\ C\beta = m$

$Q = (y - X\beta)^t(y - X\beta) - \lambda(C\beta - m)$ 여기서, $\lambda_{(q\times 1)}$: 라그랑지 승수

$0 = \dfrac{\partial Q}{\partial \beta} = -2X^t y + 2(X^t X)\beta - C^t\lambda$ ----- ①

$0 = \dfrac{\partial Q}{\partial \lambda} = C\beta - m$ ----- ②

$\hat{\beta}_R$: ①과 ②를 만족하는 해

①에서, $(X^t X)\hat{\beta}_R = X^t y - \dfrac{1}{2}C^t\lambda$

$\rightarrow\quad \hat{\beta}_R = (X^t X)^{-1}X^t y - \dfrac{1}{2}(X^t X)^{-1}C^t\lambda$ ----- ③

$= \hat{\beta} - \dfrac{1}{2}(X^t X)^{-1}C^t\lambda$

②에서, $C\hat{\beta}_R = m$

③에서, $\hat{\beta}_R = \hat{\beta} - \frac{1}{2}(X^tX)^{-1}C^t\lambda \quad \rightarrow \quad C\hat{\beta}_R = C\hat{\beta} - \frac{1}{2}C(X^tX)^{-1}C^t\lambda$

$\Rightarrow \frac{1}{2}\lambda = \left[C(X^tX)^{-1}C^t\right]^{-1}C(\hat{\beta}-\hat{\beta}_R) = \left[C(X^tX)^{-1}C^t\right]^{-1}(C\hat{\beta}-m)$

* $C\hat{\beta} - C\hat{\beta}_R = C(X^tX)^{-1}C^t\frac{1}{2}\lambda \rightarrow C(\hat{\beta}-\hat{\beta}_R) = C(X^tX)^{-1}C^t\frac{1}{2}\lambda \rightarrow (\hat{\beta}-\hat{\beta}_R) = (X^tX)^{-1}C^t\frac{1}{2}\lambda$

$\Rightarrow (\hat{\beta}-\hat{\beta}_R) = (X^tX)^{-1}C^t\left[C(X^tX)^{-1}C^t\right]^{-1}(C\hat{\beta}-m)$

$SSE(R) = e_R^t e_R = (y - X\hat{\beta}_R)^t(y - X\hat{\beta}_R) = (y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}_R)^t(y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}_R)$

$= (e + X\hat{\beta} - X\hat{\beta}_R)^t(e + X\hat{\beta} - X\hat{\beta}_R) = \left[e + X(\hat{\beta}-\hat{\beta}_R)\right]^t\left[e + X(\hat{\beta}-\hat{\beta}_R)\right]$

$= e^te + (\hat{\beta}-\hat{\beta}_R)^tX^te + e^tX(\hat{\beta}-\hat{\beta}_R) + (\hat{\beta}-\hat{\beta}_R)^tX^tX(\hat{\beta}-\hat{\beta}_R)$ 여기서, $X^te = 0$, $e^tX = 0$[1)]

$= SSE(F) + (C\hat{\beta}-m)^t\left[C(X^tX)^{-1}C^t\right]^{-1}C(X^tX)^{-1}(X^tX)(X^tX)^{-1}C^t\left[C(X^tX)^{-1}C^t\right]^{-1}(C\hat{\beta}-m)$

$= SSE(F) + (C\hat{\beta}-m)^t\left[C(X^tX)^{-1}C^t\right]^{-1}(C\hat{\beta}-m)$

$\Rightarrow SSE(R) - SSE(F) = (C\hat{\beta}-m)^t\left[C(X^tX)^{-1}C^t\right]^{-1}(C\hat{\beta}-m)$

$df_R - df_F = [n-(p-q)] - (n-p) = q$

$\frac{SSE(R)-SSE(F)}{\sigma^2} \sim \chi^2(q)$

$\frac{SSE(F)}{\sigma^2} \sim \chi^2(n-p)$

$SSE(R)-SSE(F)$와 $SSE(F)$ 간의 독립성을 검증하려면

$SSE(R)-SSE(F)$: $\hat{\beta}$의 함수, $SSE(F)$: e의 함수 → $Cov(\hat{\beta}, e) = 0$

$H_0 : C\beta = m$ 하에서

$$F = \frac{SSE(R)-SSR(F)/q}{s^2} = \frac{(C\hat{\beta}-m)^t\left[C(X^tX)^{-1}C^t\right]^{-1}(C\hat{\beta}-m)/q}{s^2} \sim F(q, n-p)$$

※ 부분 F−검정은 일반적 가설검정의 특수한 경우다.

- 부분 F−검정: $H_0 :\ \beta_2 = 0$
- 일반적 선형 F−검정: $H_0 : C\beta = m$

---

1) $X^te = X^t\left[I - X(X^tX)^{-1}X^t\right]y = X^ty - X^tX(X^tX)^{-1}X^ty = X^ty - X^ty = 0$

[예] $y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \epsilon_i$

$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$ ← 부분 F−검정

(Q) 이 귀무가설을 $C\beta = m$ 으로 표현해 보자. 이때 C는 어떻게 표현되는가?

(A) $\beta_2 = \beta_3 = \beta_4 = 0 \iff \beta_2 - \beta_3 = 0$ & $\beta_3 - \beta_4 = 0$

→ $C = \begin{bmatrix} 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \ m = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

(Q) $H_0 : \beta_1 = \beta_2^2$ 에 대해서도 상기 결과를 사용할 수 있는가?

(A) 사용할 수 없다. β에 대해 선형이 아니기 때문이다.