

9주차 1차시 : 가변수모형(개념 및 종류)

1. 가변수 모형

- (1) 가변수의 본질 및 필요성
- (2) 가변수의 종류

1.가변수 모형

(1)가변수의 본질 및 필요성

- 가변수(dummy variable)는 독립변수 중의 일부가 성질을 달리하는 질적인 자료로 되어 있을 경우 사용됨
- 가변수란 질적인 면을 나타내 주는 변수이며 가변수를 사용하면 더욱 정확한 통계적 추론을 할 수 있음
- 질적인 면을 고려해야 함에도 불구하고 이를 고려하지 않을 경우 이는 모형 내 있어야 할 변수를 뺀 경우이고 이 경우 추정량은 불편성을 가지지 못함

(2)가변수의 종류

- 가변수는 질적 범주를 구별하기 위해 사용되는 변수이므로 모형에서 어느 한 질적 변수의 2개의 질적 범주를 구분하기 위해서 하나의 가변수가 통상적으로 쓰임
- 둘 이상의 질적 변수나 어느 한 질적 변수의 두 개 이상의 질적 범주를 모형이 포함하고 있을 경우 두 개 이상의 가변수가 필요함

(예1)전시와 소비시의 소비행태

$$C_t = \beta_0 + \beta_1 Y_t + u_t \quad (\text{보통회귀식})$$

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 D_t + u_t \quad (\text{가변수가 포함된 회귀식})$$

$$\text{단, } D_t = \begin{cases} 1, t \text{가 평화시} \\ 0, t \text{가 전쟁시} \end{cases}$$

(예2)성별 및 인종별 임금격차(두 개의 질적 변수를 가진 경우)

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (\text{보통회귀식})$$

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_{1i} + \beta_3 D_{2i} + u_i \quad (\text{가변수가 포함된 회귀식})$$

$$\text{단, } D_{1i} = \begin{cases} 1, i \text{가 남자} \\ 0, i \text{가 여자} \end{cases}$$

$$D_{2i} = \begin{cases} 1, i \text{가 백인} \\ 0, i \text{가 유색인} \end{cases}$$

(예3)학력별 임금격차(두 개 이상의 질적 범주를 가진 경우)

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (\text{보통회귀식})$$

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_{Hi} + \beta_3 D_{Ci} + u_i \quad (\text{가변수가 포함된 회귀식})$$

단, $D_{Hi} = 0, D_{Ci} = 0$, i 가 중졸 이하

$$D_{Hi} = 1, D_{Ci} = 0, i \text{가 고졸}$$

$$D_{Hi} = 0, D_{Ci} = 1, i \text{가 대졸 이상}$$