

사회통계연습

양적변수×가변수 조절효과

김현우, PhD¹

¹충북대학교 사회학과 부교수



진행 순서

- 1 조절효과의 직관적 이해
- 2 조절변수와 상호작용효과

조절효과의 직관적 이해

조절효과의 직관적 이해

X 와 Y 의 연관성은 사람에 따라 상이할지도 모른다.

- 아래와 같은 회귀식에서 회귀계수 β_1 는 X 와 Y 사이의 연관성을 보여준다.

$$E(Y|X) = \beta_0 + \beta_1 X$$

- β_1 가 보여주는 효과는 표본 안의 모든 사람(혹은 그 밖의 관찰단위)에 대해 X 와 Y 의 연관성이 똑같다는 것을 전제로 한다(Why?).
- 그런데 이 가정은 지나친 단순화일수도 있다. 가령 교육수준 X_1 이 소득 Y 에 미치는 영향은 부모의 교육수준 X_2 에 따라 달라질지도 모른다!

조절효과의 직관적 이해

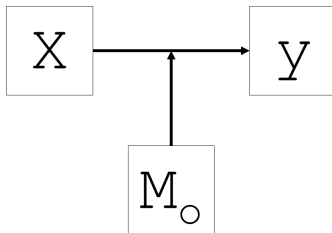
- 회귀식을 아래처럼 회귀계수 β_2 와 X_2 를 추가했다고 이런 가능성을 테스트해볼 수 있는 것은 아니다(Why?).

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

- β_1 의 해석은 그저 “ X_2 를 통제하였을 때, X_1 이 한 단위 증가하면 Y 는 β_1 만큼 증가 (또는 감소)한다”이기 때문이다.
- 우리가 원하는 해석은 “부모의 교육수준 X_2 에 따라, X_1 와 Y 의 연관성에는 차이가 있다” 같은 것이다!

조절효과의 직관적 이해

- 결론부터 말하자면, **조절변수(moderating variables)**를 활용하여 **조건부 효과(conditional effects)**를 살펴볼 수 있다.
- 그림으로 나타내보자면 결국 조절변수 M_o 의 **개입(intervention)**에 따라 독립변수 X 가 종속변수 Y 와 맺는 관계가 달라짐을 보는 것이다.



조절효과의 직관적 이해

조절효과는 연구가설을 세련화하는데 매우 중요하다.

- 단순한 선형관계로 이루어진 연구가설로 만족하지 않고, 한발 더 나아가 상호작용 효과(interaction effect)에 관해 연구가설을 세우고 테스트할 때, 연구는 훨씬 세련된 형태로 발전한다.
- 가령 노조원 지위 M_o 가 재직경력 X 과 임금 Y 사이의 연관성을 어떻게 조절하는지 살펴보자.
- “연구가설 1. 재직경력 X 가 증가하면 임금 Y 은 더 커질 것이다(이른바 연공급).”
- “연구가설 2. 노조원($M_o = 1$)의 경우 비노조원($M_o = 0$)보다 더 많은 임금을 받을 것이다(이른바 노조 프리미엄).”

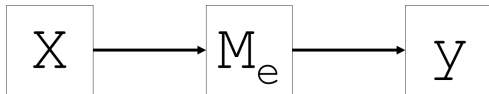
조절효과의 직관적 이해

- 더 좋은 사회학적 설명을 위해 한 걸음 더 내딛어보자!
- 노조원의 경우, 재직경력이 없거나 짧더라도 상대적으로 더 많은 임금을 보장받을 것이다. 그러므로 Y 축(임금) 절편이 비노조원보다 높으리라고 예상할 수 있다 (Why?).
- 다른 한편, 노조원은 이미 초봉을 높게 출발했으므로 재직경력이 쌓이더라도 그 상승 추세는 비노조원보다야 상대적으로 적을 것이다(Why?).
- “연구가설 3. 노조원의 경우 재직경력 X 의 증가에 따른 임금 Y 의 상승폭은 비노조원보다 오히려 더 작을 것이다.”
- 인과도표 그림을 그려가면서 연구가설을 세우는 것이 유용하다!

조절효과의 직관적 이해

조절효과와 매개효과를 혼동하지 않도록 주의해야 한다.

- 지난 주에 배운 매개변수(mediating variables) M_e 는 아래와 같이 인과도식을 그린다.



- 두 효과 모두 세련된 연구가설을 세우는데 매우 중요한 접근방법이 된다!
- 다음 학기 사회조사방법론을 수강할 때는 두 효과를 진지하게 고려하여 연구가설을 만들수 있어야 한다.

조절변수와 상호작용효과

조절변수와 상호작용효과

회귀모형에서 조절효과를 살펴볼 수 있는 단순한 방법이 있다.

- 노조원 지위 M_o 가 임금 Y 과 재직경력 X 사이의 관계를 어떻게 조절하는지 살펴보자.
- 위의 연구가설에서는 두 개의 주효과(main effects)와 하나의 상호작용효과(interaction effect)에 관해 아래와 같이 서술하고 있다.
- $TENURE \rightarrow LN_WAGE$ (주효과 #1)
- $UNION \rightarrow LN_WAGE$ (주효과 #2)
- $UNION$ 에 따라 상이한 $TENURE \rightarrow LN_WAGE$ (상호작용효과)

조절변수와 상호작용효과

- 아래의 회귀식을 꼼꼼히 살펴보자.

$$\text{LN_WAGE} = \beta_0 + \beta_1 \text{TENURE} + \beta_2 \text{UNION}$$

- 비록 TENURE와 UNION 변수가 포함되어 있지만, 그 회귀계수를 통해서는 각각의 효과만을 볼 수 있을 뿐이다.
- 그러나 만일 UNION에 따라 TENURE의 영향력이 변화한다면, TENURE의 효과($=\beta_1$)는 다음과 같이 모형화될 수 있다(Why?).

$$\beta_1 = \gamma_0 + \gamma_1 \text{UNION}$$

- 위의 두 식을 결합하면 아래와 같다.

$$\begin{aligned}\text{LN_WAGE} &= \beta_0 + (\gamma_0 + \gamma_1 \text{UNION}) \cdot \text{TENURE} + \beta_2 \text{UNION} \\ &= \beta_0 + \gamma_0 \text{TENURE} + \beta_2 \text{UNION} + \gamma_1 (\text{TENURE} \cdot \text{UNION})\end{aligned}$$

조절변수와 상호작용효과

- 독립변수인 TENURE와 조절변수인 UNION 외에도 두 변수를 곱한 새로운 변수 $TENURE \cdot UNION$ 이 포함되어 있다.
- 이 세번째 변수를 특별히 상호작용항(interaction term)이라고 한다.
- 나머지 두 변수를 주요항(main terms)이라고 한다.
- 물론 상호작용항 회귀계수에 대해서도 유의성 검정($H_0 : \beta_3 = 0$)을 한다!
- 만약 통계적으로 유의하다면 재직경력의 임금 효과가 노조원 지위에 따라 달라지는 상호작용 효과가 모집단에서도 존재한다고 말할 수 있다.

조절변수와 상호작용효과

- 상호작용항을 포함한 회귀모형이 아래와 같았다.

$$\text{LN_WAGE} = \beta_0 + \beta_1 \text{TENURE} + \beta_2 \text{UNION} + \beta_3 (\text{TENURE} \cdot \text{UNION})$$

- 상호작용항을 포함한 회귀모형의 해석은 가변수의 해석과 크게 다르지 않다.

$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_1 \text{TENURE} \quad (\text{if UNION}=0)$$

$$\begin{aligned} \text{LN_WAGE} &= \hat{\beta}_0 + \hat{\beta}_1 \text{TENURE} + \hat{\beta}_2 + \hat{\beta}_3 \text{TENURE} \quad (\text{if UNION}=1) \\ &= (\hat{\beta}_0 + \hat{\beta}_2) + (\hat{\beta}_1 + \hat{\beta}_3) \cdot \text{TENURE} \end{aligned}$$

- 노조원이 아닌 경우(UNION==0), $\hat{\beta}_0$ 와 $\hat{\beta}_1$ 로만 해석한다.
- 노조원인 경우(UNION==1), (1) $\hat{\beta}_0$ 위에 $\hat{\beta}_2$ 만큼의 상수 부스터가, (2) $\hat{\beta}_1$ 위에 $\hat{\beta}_3$ 만큼의 회귀계수 부스터가 더해진다!

조절변수와 상호작용효과

정리해보자.

- 상호작용항은 결국 독립변수와 조절변수를 곱해서 만든다. 그 뒤, 상호작용항을 기존의 두 변수인 주요항과 함께 모형에 넣으면 된다.
- 상호작용항을 해석할 때 무엇이 조절변수인지는 여러분 스스로가 결정할 몫이다. 이것은 완전히 주관적인 해석의 문제이다(Why?).
- 상호작용 효과는 연구가설을 보다 발전시킨 것이다. 그러므로 상호작용항이 없는 회귀모형을 먼저 추정한 뒤 이를 보여주고, 표 안에서 나란히 비교하는 것이 중요하다. 결정계수의 변화에 관해서도 언급해야 한다!

조절변수와 상호작용효과

연습 1. UNION.SAV에서 재직경력(TENURE)과 노조원 지위(UNION)가 임금(LN_WAGE)에 미치는 영향력을 회귀분석으로 추정하시오. 그리고 노조원 지위와 재직경력 간 상호작용효과를 검토하시오.

조절변수와 상호작용효과

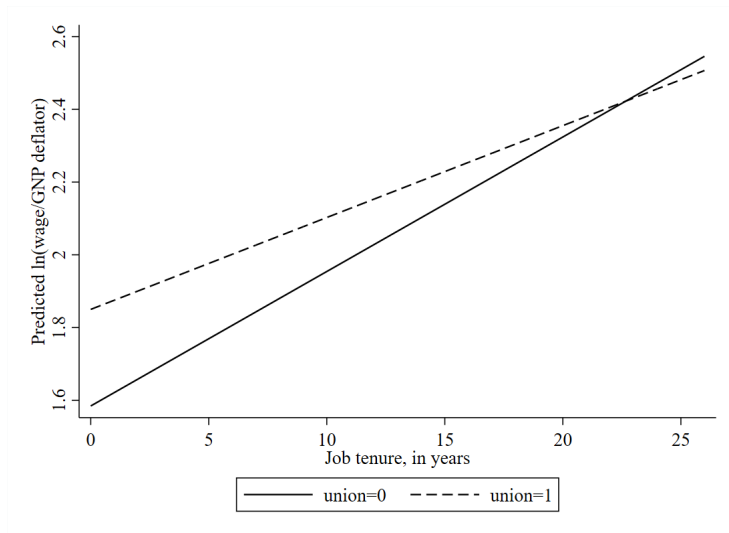
- 먼저 상호작용항이 통계적으로 유의한지 확인한다!
- 통계적으로 유의하지 않으면 상호작용효과 가설은 지지되지 않은 셈이므로 이것을 버리고 상호작용항이 없는 회귀식을 해석해야 한다.
- “노조원의 경우, 비노조원보다 재직경력이 1년 증가에 따른 임금 상승분이 0.012만큼 적다.”
- 상호작용항을 넣을 때는 그 상호작용항을 이루는 주요항들도 반드시 함께 회귀식 안에 넣어야 한다(Why?).
- 그러나 주요항은 해석하거나 유의성 검정을 따로 수행하지 않는다(Why?).

조절변수와 상호작용효과

상호작용효과는 꼭 시각화하는 편이 바람직하다.

- 회귀분석 결과표만 가지고는 의미 해석이 어렵다(Why?).
- Jamovi나 SPSS에서 (상호작용항이 포함된) 선형회귀분석을 수행할 때 비표준화 예측값을 “저장”한다.
- 이것을 Y 축으로 TENURE를 X 축으로 한 집단별 산점도를 그린다.
- SPSS에서라면 집단화 산점도 메뉴 안에서 우측 상단의 “색상 설정”에 UNION을 넣는다.
- 만들어진 집단화 산점도에서 “부집단 회귀선 적합 추가” 아이콘을 눌러 적합선 두 개를 추가한다(깨끗하게 만드려면 적합선만 남기고 산점도는 오히려 삭제하는 것도 요령이다).
- Jamovi에서라면 “집단” 안에 UNION을 넣는다.

조절변수와 상호작용효과



조절변수와 상호작용효과

연습 2. diabetes.sav에서 연령(age), 흑인(black), 여성(female)와 당뇨병 유무(diabetes)의 연관성을 회귀분석으로 살펴보시오. 흑인과 연령의 상호작용효과를 추가적으로 살펴보고 이를 시각화시오.

조절변수와 상호작용효과

