

사회통계연습

가변수 \times 가변수 조절효과

김현우, PhD¹

¹충북대학교 사회학과 부교수



진행 순서

- 1 가변수 \times 가변수
- 2 양적변수 \times 양적변수
- 3 상호작용효과 연습



가변수 × 가변수



가변수 × 가변수

상호작용효과도 두 변수의 측도를 고려해서 살펴보아야 한다.

- (상호작용효과를 살펴보려는) 두 변수의 측도가 어떻게 구성되어 있는가에 따라 다음 세 가지 유형이 있다.
 - (1) 양적변수 × 가변수
 - (2) 가변수 × 가변수
 - (3) 양적변수 × 양적변수
- 앞 수업에서 다룬 노조원 신분 사례는 (1)의 유형이었다(Why?).
- 이제 나머지 두 개를 살펴보자. 다행히 내용은 거의 똑같다.



가변수 × 가변수

연습 1. UNION.sav에서 대졸 여부(COLLGRAD)와 노조원 지위(UNION)가 임금(LN_WAGE)에 미치는 영향력을 회귀분석으로 추정하시오. 그 다음 대졸 여부와 노조원 지위의 (임금에 대한) 상호작용효과를 검토하시오.



가변수 × 가변수

- 주요항은 대졸 여부(COLLGRAD), 조절변수는 노조원 지위(UNION)이므로 상호작용항을 포함한 회귀모형을 다음과 같이 구성한다.

$$\begin{aligned} \text{LN_WAGE} = & \beta_0 + \beta_1 \text{COLLGRAD} + \beta_2 \text{UNION} \\ & + \beta_3 (\text{COLLGRAD} \cdot \text{UNION}) \end{aligned}$$

- 이 유형도 간단히 해석할 수 있다.

$$\text{LN_WAGE} = \hat{\beta}_0 \quad (\text{if } \text{COLLGRAD}==0 \ \& \ \text{UNION}==0)$$

$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_1 \quad (\text{if } \text{COLLGRAD}==1 \ \& \ \text{UNION}==0)$$

$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_2 \quad (\text{if } \text{COLLGRAD}==0 \ \& \ \text{UNION}==1)$$

$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 \quad (\text{if } \text{COLLGRAD}==1 \ \& \ \text{UNION}==1)$$



가변수 × 가변수

- 해석할 때는 각각의 네 가지 상황별로 예측된 임금에 대해 정리하는 편이 쉽다.
- 상호작용항을 넣었을 때는 주요항을 해석하지 않도록 주의해야 한다(Why?).
- 대졸 효과(β_1)와 노조 효과(β_2) 둘 다 임금 상승에 기여한다.
- 그런데 대졸 효과와 노조 효과가 동시에 상호작용할 때는 두 개별적 효과가 약간 상쇄되어 줄어드는 것으로 나타난다(Why?).

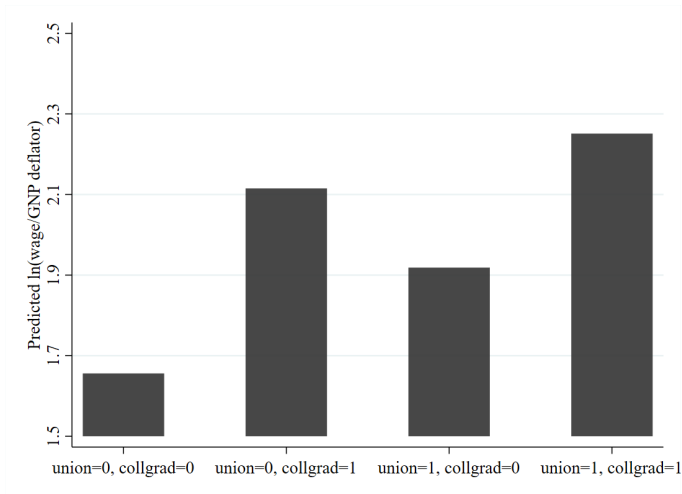


가변수 × 가변수

- 시각화할 때는 막대차트(bar chart)를 활용하는 편이 좋다(Why?).
- 회귀분석할 때 \hat{Y} 를 저장하자.
- 시각화 단계에서 SPSS의 경우, [차트 작성기]의 '수평누적 막대도표'를 선택한다. Y 는 \hat{Y} , X 는 UNION이다. 이때 'X에 군집: 색상 설정'에 COLLGRAD를 넣는다 (Why?).
- Jamovi의 경우 [기술통계]-[탐구]에서 '막대도표'를 선택하면 된다.
- 적절하게 표를 꾸민다.



가변수 × 가변수



양적변수 × 양적변수



양적변수 × 양적변수

연습 2. UNION.sav에서 재직경력(TTL_EXP)과 나이(AGE)가 임금(LN_WAGE)에 미치는 영향력을 회귀분석으로 추정하시오. 그 다음 재직경력과 나이의 상호작용효과를 검토하시오.



양적변수 × 양적변수

- 마지막 유형은 주요항과 상호작용항 모두가 양적변수인 경우이다.
- 이번엔 재직경력(TTL_EXP)과 나이(AGE) 뿐만 아니라, 상호작용항도 포함한 회귀모형을 다음과 같이 구성한다.

$$\text{LN_WAGE} = \beta_0 + \beta_1 \text{TTL_EXP} + \beta_2 \text{AGE} + \beta_3 (\text{TTL_EXP} \cdot \text{AGE})$$

- β_3 가 통계적으로 유의하면 상호작용효과를 해석할 수 있다.
- 해석은 혼동하기 쉬우므로 약간 주의해야 한다. 요령은 조절변수를 마치 범주형처럼 읽는 것이다!



양적변수 × 양적변수

- 조절변수를 양적변수 그대로 두면 해석이 직관적이지 않으므로 대표적인 범주 몇 개만을 골라 거기에 초점을 두고 해석하는 것이 요령이다!
- 해석의 초점을 나이에 둔다면, 30세, 40세, 50세로 나누어 볼 수 있다.

$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_1 \text{TTL_EXP} + \hat{\beta}_2 \cdot \mathbf{30} + \hat{\beta}_3 (\text{TTL_EXP} \cdot \mathbf{30})$$

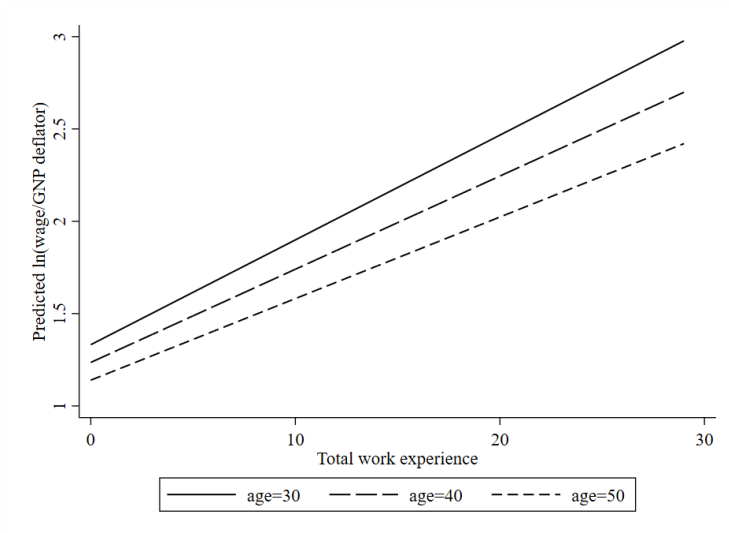
$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_1 \text{TTL_EXP} + \hat{\beta}_2 \cdot \mathbf{40} + \hat{\beta}_3 (\text{TTL_EXP} \cdot \mathbf{40})$$

$$\text{LN_WAGE} = \hat{\beta}_0 + \hat{\beta}_1 \text{TTL_EXP} + \hat{\beta}_2 \cdot \mathbf{50} + \hat{\beta}_3 (\text{TTL_EXP} \cdot \mathbf{50})$$

- 사실 양적변수 × 양적변수의 상호작용 효과를 해석하기는 까다로우므로, 조절변수를 아예 처음부터 범주화하는 경우도 종종 있다(여러분에게는 이쪽을 더 추천한다).



양적변수 × 양적변수



상호작용효과 연습



상호작용효과 연습

송리라 · 이민아(2013)를 보고 조절변수가 활용되는 방식을 살펴보자.

- 어떤 상호작용효과를 보고 있는가?
- 독립변수와 조절변수의 측도는 각각 무엇인가?
- 본문의 <표 4>에 따르면 상호작용항(성역할태도 \times 성별)이 통계적으로 유의한 것으로 나타나 있다. 이를 104페이지 그림과 함께 해석해 보자.
- 종속변수, 독립변수, 그리고 조절변수는 각각 무엇인가(97-98페이지)?
- 본문의 <표 2>에서 우울 및 성역할 태도는 각각 남녀 어느 쪽에서 높은가(165페이지)? 연령집단별 통계적으로 유의한 차이가 있는가?
- 본문의 <표 4>에 따르면 상호작용항(성역할태도 \times 성별)이 통계적으로 유의한 것으로 나타나 있다. 이를 <그림 1>과 함께 해석해 보라(103-104페이지).

송리라 · 이민아. 2012. “성역할 태도와 우울.” 한국인구학 35(3): 87-116.



조절변수의 연습

<표 3> 우울에 미치는 성별과 성역할 태도의 영향

		모형1		모형2		모형3	
		계수	표준오차	계수	표준오차	계수	표준오차
성역할 태도		0.01**	0.010	0.01**	0.01	0.01*	0.01
성별(1 = 여성)		0.12**	0.050	0.15***	0.05	0.10**	0.05
연령		-0.01**	0.004	-0.01**	0.01	-0.01***	0.01
결혼유무(1 = 결혼)		-0.15**	0.060	-0.06	0.09	0.02	0.09
교육수준	고졸(준거집단)						
	2~4년제 대졸	0.11*	0.060	-0.01	0.07	-0.06	0.06
	석사이상	-0.20**	0.100	-0.20*	0.11	-0.15	0.10
연평균가구소득		-0.02***	0.010	-0.01**	0.01	-0.01	0.01
직업형태	정규직(준거집단)						
	비정규직	0.02	0.080	0.03	0.08	0.01	0.08
	자영업	0.03	0.090	0.012	0.10	0.003	0.09
	무직	-0.09	0.070	-0.08	0.07	-0.12*	0.07
	학생	-0.22**	0.090	-0.23**	0.09	-0.23	0.09
	없음(준거집단)						
자녀유무	미취학아동			-0.13	0.10	-0.2**	0.09
	초등학생이상			-0.14	0.10	-0.17*	0.10
집안일 정도	저녁식사준비			0.03	0.02	0.03	0.02
	빨래			-0.05*	0.02	-0.04	0.02
	청소			0.002	0.03	0.05	0.03
주관적 건강평가						-0.28***	0.03
상수항		2.45***	0.190	2.38***	0.24	1.74***	0.24
	표본수	969		912		912	
	F	5.03***		3.52***		9.61***	
	결정계수	0.0547		0.0593		0.1545	

주: * p<0.10, **p<0.05, ***p<0.01



조절변수의 연습

<표 4> 성별과 성역할 태도 간의 상호작용 효과분석

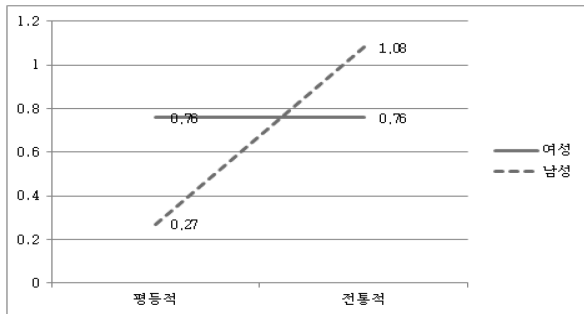
		모형	
		계수	표준오차
성역할 태도*성별		-0.030**	0.01
성역할 태도		0.030***	0.01
성별(1=여성)		0.760***	0.26
연령		-0.010***	0.01
결혼여부(1=기혼)		0.010	0.09
교육수준	고졸(준거집단)		
	2-4년제 대졸	-0.070	0.06
	석사이상	-0.170*	0.10
연평균가구소득		-0.010	0.01
직업형태	정규직(준거집단)		
	비정규직	0.020	0.08
	자영업	0.002	0.09
	무직	-0.100	0.07
	학생	-0.230	0.09***
자녀유무	없음(준거집단)		
	미취학아동	-0.190**	0.09
	초등학생이상	-0.180*	0.10
집안일 정도	저녁식사준비	0.030	0.02
	빨래	-0.030	0.02
	청소	0.010	0.03
주관적 건강평가		-0.280***	0.03
상수항		1.330***	0.29
표본수		912	
F		9.50***	
결정계수		0.1607	

주: * p<0.10, **p<0.05, ***p<0.01



조절변수의 연습

<그림 1> 성별과 성역할 태도에 따른 우울

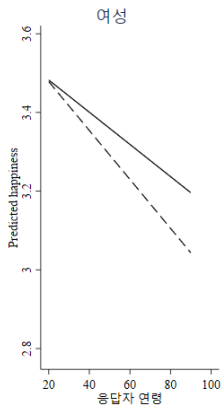
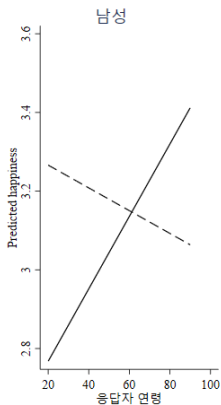
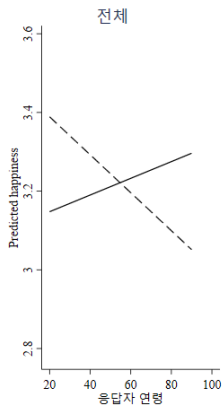


상호작용효과 연습

연습 3. HAPPY.SAV에서 행복감(HAPPY)을 종속변수로, 성별(SEX), 나이(AGE), 도시 거주여부(URBAN)를 독립변수로 한 회귀분석을 수행하시오. 그 다음 나이와 도시 거주여부 간 상호작용효과에 관한 연구가설을 제시하고 이를 검증하시오. 같은 내용의 분석을 성별에 따라 따로따로 수행한 뒤, 그 시사점에 대해서도 논하시오.



상호작용효과 연습



— 농촌 - - - 도시

