

# 계량분석 Causal Diagram

김현우, PhD<sup>1</sup>

<sup>1</sup>충북대학교 사회학과 조교수

December 5, 2024



# 진행 순서

- 1 표준화 회귀계수
- 2 인과도표의 작성
- 3 인과도표 작성시 주의사항

## 표준화 회귀계수

# 표준화 회귀계수

여러 독립변수를 고려하다 보면 어떤 변수가 가장 중요한지 알고 싶어진다.

- 회귀계수 값의 크기로 변수 간 영향력의 차이를 비교할 수 없다(Why?).
- 이를 비교하기 위한 한 가지 방법은 **표준화 회귀계수**(standardized regression coefficient)를 사용하는 것이다.
- 표준화에도 몇 가지 방법이 있는데 대체로 다음을 사용한다.

$$\beta_k^* = \frac{s_{X_k}}{s_Y} b_k$$

- 여기서  $b_k$ 는 원(raw) 회귀계수,  $s$ 는 표준편차를 뜻한다.



# 표준화 회귀계수

연습 1. attend.csv에서 termgpa를 종속변수로, 기말시험(final), 예전 학점(priGPA), 과제(hwrte), 결석(skipped)을 독립변수로 하는 회귀모형을 구성하시오. 결과를 해석하고 어떤 변수가 termgpa를 가장 잘 설명하는지 논하시오.



# 표준화 회귀계수

- 표준화 회귀계수의 단위는 표준편차이므로(Why?) 그 해석은 다소 색다르다.
- “기말시험 점수의 1표준편차 증가는 학기말 GPA를 0.317 표준편차만큼 증가시킨다.”
- 여기서 중요한 것은 모든 변수들의 단위가 이제 표준편차로 통일되었다는 사실이다 (Why?).
- 특히 합성지수같은 경우에는 단위가 애매했는데 이제 모두 똑같다.
- 단위가 똑같아졌으므로 표준화 회귀계수는 값의 크기로 어떤 변수가 얼마나 중요한지 알려준다(Why?).



## 인과도표의 작성

# 인과도표의 작성

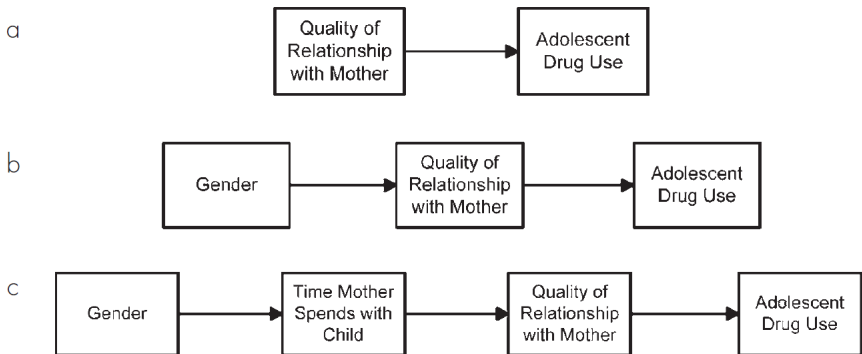
연구질문에 대한 답을 구체화시키는 과정이 필요하다.

- 가령 어머니와의 관계(quality of relationship with mother)가 청소년의 약물 남용(adolescent drug use)과 어떻게 연결되어 있는지를 사례로 살펴보자.
- 처음에는 어머니와의 관계이라는 변수가 청소년의 약물 남용이라는 변수와 어떻게 연결되어 있는가를 그린다.
- 좀 더 나아가, 혹시 청소년의 성별이 어머니와의 관계에 영향을 미치고, 궁극적으로 약물 남용의 위험을 높일 수 있음을 고민해본다.
- 만약 아들은 딸보다 어머니와 함께 보내는 시간이 적고, 이것이 어머니와의 관계를 악화시키고, 다시 청소년의 약물 남용으로 이어지지 않는지 의심해 볼 수 있다.





# 인과도표의 작성



# 인과도표의 작성

연구질문에 대한 나름의 대답을 구체화하면서 연구모형을 구상한다.

- 실증적 사회학 연구의 맥락에서 주장(argument)이란 곧 어떤 사회현상의 원인과 결과에 대한 설명이다.
- 그러므로 자신의 주장을 요약하라는 요청을 받으면 **인과도표(causal diagram)**를 그려가며 설명할 수 있어야 한다!
- 인과도표는 변수 사이의 관계를 간단히 요약하는 그림이다.
- 아주 단순한 인과도표에서 시작하여 조금씩 확장하면서 연구주제를 발전시켜 나갈 수 있다.
- 끈질긴 고민이 허술한 생각을 좋은 연구질문으로 만든다. 인과도표는 그 지팡이가 된다.



# 인과도표의 작성

인과도표에서 나타낸 각각의 항목은 변수를 지시한다.

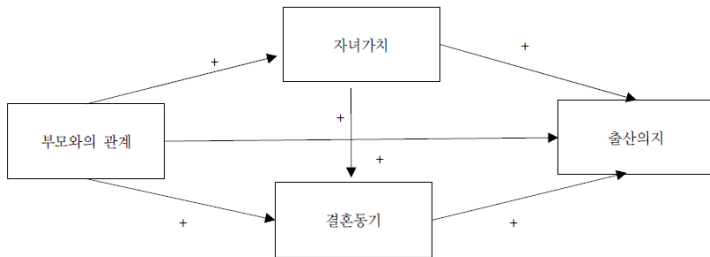
- 만일 다차원적 개념이라면 여러 개의 변수로 측정되어야 할 수도 있다.
- 이런 경우 (하나의 다차원적 개념을 측정하는) 여러 개의 변수들을 인과도표 상에 나타내도록 하자.
- 이렇게 변수로 측정하는 방식은 기존 문헌을 훑어내도록 하자.
- 김은진 외(2002)의 연구처럼 부모와의 관계, 자녀가치, 결혼동기, 출산의지 등 여러 개념들의 복합적인 관계를 인과도표로 나타낼 수 있다.

김은진, 정혜정, 정유진. 2022. “대학생이 지각한 자녀가치와 결혼동기 및 부모와의 관계가 출산의지에 미치는 영향: 성별에 따른 다집단 경로분석.” 가정과삶의질연구 40(1): 27-39.



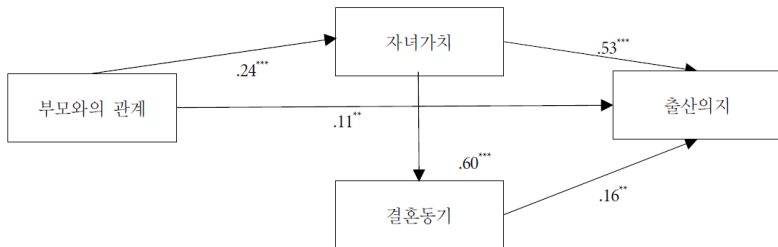
# 인과도표의 작성

- 인과도표는 연구 보고서의 첫머리에 연구모형으로 제시할 수도 있다.



# 인과도표의 작성

- 하지만 분석이 끝나고 화살표 위에 회귀계수를 써넣는 방식으로 연구 결과를 요약하는데 쓰이기도 한다!



\*\*  $p < .01$ , \*\*\*  $p < .001$

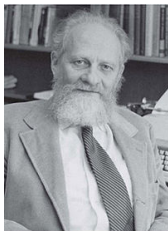
모든 경로계수는 표준화계수임



# 인과도표의 작성

## 세대간 직업이동의 매커니즘은 무엇일까?

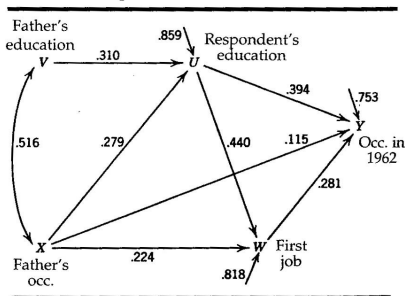
- 세대간 직업이동(inter-generational occupational mobility)의 매커니즘에 관한 고전적 설명은 **지위획득모형(status attainment model)**에 의해 이루어졌다.
- 1960년대에 이미 Peter Blau와 Otis Dudley Duncan는 미국 센서스 원자료를 분석하여 이른바 **블라우-덩컨 모형(Blau-Duncan Model)**을 발표하였다.



# 인과도표의 작성

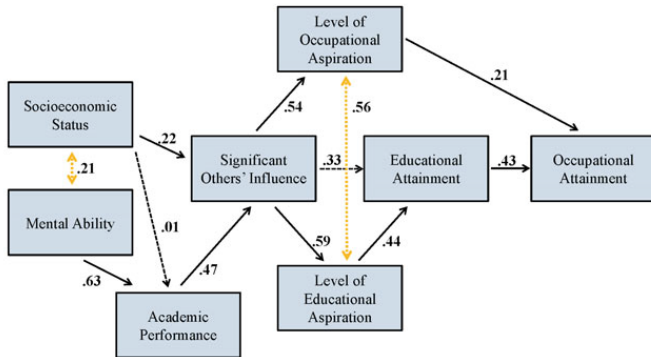
- 본인의 교육수준, 부친의 교육수준, 부친의 직업위세가 청년구직자의 노동시장 성과에 가장 중요한 변인이다.
- 그런데 본인의 교육수준은 다시 부친의 교육수준과 직업위세에 크게 좌우된다.
- 지위를 귀속 지위(ascribed status)와 성취 지위(achieved status)로 구분해왔지만, 지위획득모형에 따르면 성취 지위는 귀속 지위에 크게 영향받는다.

FIGURE 1  
Path coefficients in basic model of the  
process of stratification



# 인과도표의 작성

- 이후 블라우-딩컨 모형이 사회심리학적 요인을 무시하고 있다는 비판이 제기되어 William Sewell을 중심으로 하여 이른바 위스컨신 모형(Wisconsin Model)이 발표된다.
- 두뇌, 자신의 교육과 직업에 대한 열망(aspiration), 파트너(SO)의 영향력 등이 추가적으로 고려된다.



Source: Sewell et al. 1969, p. 85 (simplified depiction)



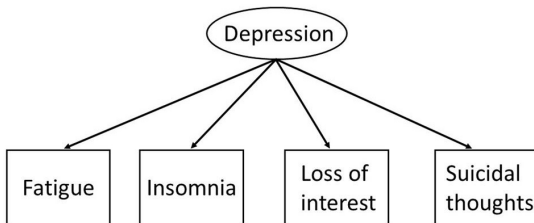


## 인과도표 작성시 주의사항

# 인과도표 작성시 주의사항

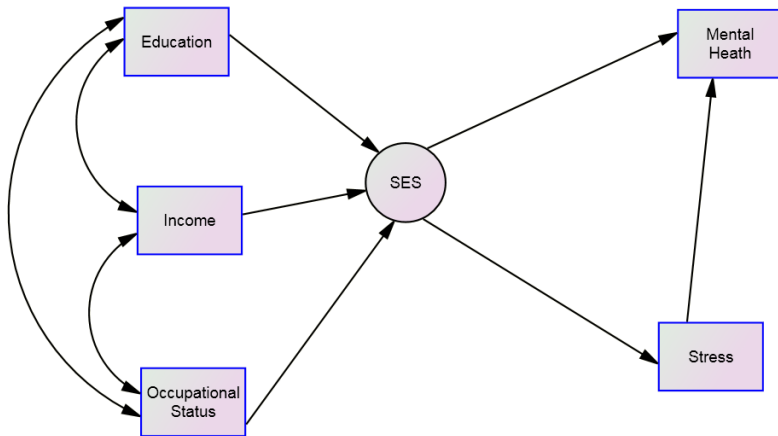
인과도표를 그릴 때 따라야 하는 몇 가지 규약이 있다.

- 변수는 사각형으로 표시하고, 인과관계는 화살표로 표시한다.
- 여러 변수로 구성된 합성지수인 경우에는 동그라미로 표시할 수 있다.
- 구조방정식 모형(structural equations model)의 관례에 따라 합성지수와 개별 변수와의 관계를 그릴수도 있다(우리에게 필수는 아니다).



# 인과도표 작성시 주의사항

- 독립변수는 가능한 왼쪽에, 종속변수는 가능한 오른쪽에 그리자.



# 인과도표 작성시 주의사항

- 개념을 대충 몽뚱그려 그림을 그리면 안된다. 설명 처음엔 그렇게 그렸더라도 나중엔 훨씬 구체화시켜야 한다.
- 단순히 “ $X$ 가  $Y$ 에 영향을 미친다”라는 가설을 설정하면 보통 재미가 없다(Why?).
- 그러한 연관성의 내부(구체적으로 어떤 매커니즘)와 외부(외적으로 주어지는 특정한 조건)에 주목해야 한다!
- 혹은 “의외의 선”을 그려보아 전혀 이어져 있지 않은 것들 사이에서 숨겨진 연관성을 파악할 수도 있다.



# 인과도표 작성시 주의사항

- 특히 **인과순환(causal loop)**을 포함하지 않아야 한다(Why?).

