

# 계량분석

## Path Analysis and Mediation Effect

김현우, PhD<sup>1</sup>

<sup>1</sup>충북대학교 사회학과 조교수

December 5, 2024



# 진행 순서

- 1 경로분석의 이해
- 2 매개효과의 검정
- 3 누락변수편의
- 4 매개분석 연구 사례

# 경로분석의 이해

# 경로분석의 이해

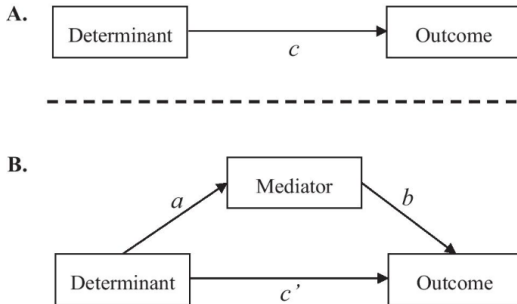
사회현상의 내부적 매커니즘을 어떻게 탐구할 수 있을까?

- 늘 새로운 주제를 연구하는 것은 아니므로 “ $X$ 는  $Y$ 에 영향을 미친다”라는 가설을 세우면 보통 재미없는 경우가 많다.
- 이것을 재미있게 만드는 전략 중 하나는 그러한 연관성을 만들어내는 핵심적인 내부 매커니즘을 밝혀내는 것이다!
- “ $X$ 는 사실  $M_e$ 를 거쳐  $Y$ 라는 결과를 이끌어낸다.”에서  $M_e$ 는  $X$ 와  $Y$  사이에서 매개효과(mediation effect)를 갖는다고 말할 수 있다.



# 경로분석의 이해

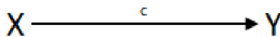
- $c$ 는 총효과(total effect)를 보여준다(그림 A).
- $a$ 와  $b$ 는 결정요인과 결과 사이 간접효과(indirect effect)를,  $c'$ 는 결정요인과 결과 사이 간접효과(indirect effect)를 보여준다(그림 B).



# 경로분석의 이해

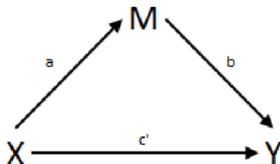
- 부분매개(partial mediation)와 전체매개(full mediation)는 구분된다.

Model 1:  
No mediation



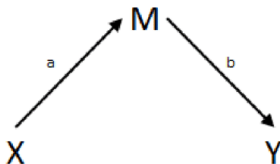
Path  $c = c'$   
 $a \&/or b = 0$

Model 2:  
General model  
(Partial mediation)



Path  $c = c' + (a \cdot b)$

Model 3:  
Full mediation



Path  $c = (a \cdot b)$   
 $c' = 0$



## 매개효과의 검정

# 매개효과의 검정

전체효과는 직접효과와 간접효과의 합이다.

- 연구자가 정의한 맥락에 따라 총효과, 직접효과, 간접효과가 달라진다. 직접효과와 간접효과를 구분하는 것이 이론적으로 의미가 있어야 한다.
- 이렇게 **인과도표(causal diagram)**를 그리면서 이론적 매커니즘을 구체화하는 과정이 바로 **경로분석(path analysis)**이다.
- 가령 “최종학력이 정치적 보수성에 영향을 미치고, 정치적 보수성이 다시 사회기관에 대한 신뢰로 이어진다”라는 연구가설을 생각해보자.
- 먼저 이 관계를 인과 도표로 그려보자.
- 다시 말해, 최종학력은 (1) 사회기관 신뢰에 대한 직접효과와 (2) 정치적 보수성을 통해 사회기관 신뢰로 이어지는 간접효과를 함께 가진다고 주장해 볼 수 있다.
- 직접효과와 간접효과의 합은 총효과와 동일하다!





# 매개효과의 검정

- 학력의 총효과  $\omega_1$  은 학력만을 독립변수로 하고, 사회기관 신뢰를 종속변수로 하여 계산할 수 있다(Why?).

$$CONFIDENCE = \omega_0 + \omega_1 EDUC$$

- 학력의 직접효과  $\beta_1$  은 학력과 정치적 이데올로기를 독립변수로 하고, 사회기관 신뢰를 종속변수로 하여 계산할 수 있다(Why?).

$$CONFIDENCE = \beta_0 + \beta_1 EDUC + \beta_2 IDEOLOGY$$



# 매개효과의 검정

- 학력의 간접효과 첫번째 경로  $\gamma_1$ 은 학력을 독립변수로 하고, 이데올로기를 종속변수로 하여 계산할 수 있다.

$$IDEOLOGY = \gamma_0 + \gamma_1 EDUC$$

- 학력의 간접효과는 두번째 경로  $\beta_2$ 은 학력과 정치적 이데올로기를 독립변수로 하고, 사회기관 신뢰를 종속변수로 하여 계산할 수 있다.

$$CONFIDENCE = \beta_0 + \beta_1 EDUC + \beta_2 IDEOLOGY$$



# 매개효과의 검정

Stata에서 매개효과를 계산해보자.

- STARTUPS.csv은 143개 미국 대학에 걸쳐 조사된 대학내 창업기업에 관한 자료이다.
- 먼저 창업기업의 수 startups를 종속변수로 하여, 먼저 기술이전부서의 연령 duration을 독립변수로 한 회귀분석을 수행한다.
- 다음으로 대학 연구비 research를 독립변수로 추가한다.
- 각 단계에서 얻어진 표준화 회귀계수  $\beta$ 를 통해 총효과, 직접효과, 간접효과의 크기를 추정하여 인과도표를 완성한다.



# 매개효과의 검정

매개효과를 강박적으로 볼 필요는 없다.

- ‘무엇의’ “무엇을 통한” ‘무엇에 대한’ 매개효과인지를 먼저 분명히 할 것!
- 인과도표 속에서 ‘이론적으로’ 가치있는 매개효과만을 보면 충분하다.
- 기존문헌을 꼼꼼히 읽으면서 가치있는 매개효과가 무엇인지는 스스로 탐구해야 한다.  
매개효과를 시사하는 기존문헌 위에서 새로운 설명을 만들어내야 한다.
- 매개효과를 본다고 모든 변수 간 매개효과를 볼 필요도 없다.



# 매개효과의 검정

- 회귀분석의 맥락에서 매개효과의 탐색은 결국 변수를 넣고 빼는 문제로 귀결된다.
- 즉 **위계적 회귀분석(hierarchical regression analysis)**이 매개효과를 탐색하는 방식으로 사용된다.
- 중요한 변수를 단계적으로 추가하면서 변화하는 회귀계수를 해석할 때 매개효과를 부분적으로 언급할 수 있다.
- 그러나 이를 넘어 매개효과의 크기를 본격적으로 계산한다는 것은 무언가 총효과를 직접효과와 간접효과를 분해할만한 ‘이론적인’ 가치가 있었다는 것을 뜻한다(가령 “간접효과가 총효과의  $x\%$ 이다”라는 발견이 특별한 의미를 가져야 한다).
- 다시 말해 구태여 그럴 가치가 없다면 매개효과를 언급할 필요는 없다.



# 매개효과의 검정

간접효과를 만들어내는 경로가 두 개 이상일 수도 있다.

- 가령 “SMSA 이외 지역 거주 `not_smsa`가 대졸 여부 `collgrad` 및 노조원 여부 `union`를 매개로 임금 `ln_wage`에 영향을 미친다”라는 연구가설을 생각해보자.
- 이때 남부 거주 여부 `south`도 통제변수로 고려하자. 이 변수는 모든 모형에서 반드시 공통적으로 포함되어야 한다.
- 이 관계를 인과 도표로 그려보자(이 단계에서 잘못 그리면 안된다).
- SMSA 이외 지역 거주 `not_smsa`의 총효과는 직접효과와 두 개의 간접효과로 분해될 수 있다.
- 그러므로 간접효과도 두 파트로 각자 계산해야 한다. 상대적인 구성비율도 계산할 수 있다.



## 누락변수편의

# 누락변수편의

매개효과는 모형설정과 관련하여 다른 시사점이 있다.

- 총효과 추정모형과 직접·간접효과 추정모형의 차이점은 매개변수(mediating variable)를 넣는가 빼는가의 차이 뿐이다.
- 만일  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$  가 완전모형 설정이고,  $Y = \beta_0 + \beta_1 X_1$  로는 불완전하다고 하자.
- 이때 두 모형의 비교는 마치  $X_1$  이  $X_2$  를 매개로  $Y$  에 어떤 영향을 미치는가 매개효과를 추정하는 것과 같다.
- 다시 말해, 간접효과의 크기는 곧 누락변수편의(omitted variable bias)의 크기와 같다(Why?).





# 누락변수편의

Stata에서 실습해보자.

- 남부 거주 south가 노조 가입 union에 영향을 미치고, 이것이 다시 임금 ln\_wage에 영향을 미친다는 연구가설을 검정하기로 하자.
- 노조 가입 union를 독립변수로 포함한 회귀모형을 완전하게 설정된 모형(fully specified model), 즉 **완전모형(full model)**이라고 가정한다.
- 노조 가입 union를 누락한 회귀모형은 **축약모형(reduced model)**이 된다.
- 이때 두 모형에서 통제변수는 모두 똑같이 유지하고 표본도 똑같이 유지해야 한다.



# 누락변수편의

- 축약모형에서 얻은  $b_{south}^{[R]}$  은 남부 거주의 총효과가 된다(Why?).
- 완전모형에서 얻은  $b_{south}^{[F]}$  은 남부 거주의 직접효과가 된다(Why?).
- 따라서  $b_{south}^{[R]} - b_{south}^{[F]}$  는 남부 거주의 간접효과가 된다(Why?).
- 당연히  $b_{south}^{[R]}$  은 왜곡된 추정값이고(Why?), 간접효과 크기는 정확히 누락변수편의 크기를 반영한다.
- 매개효과를 검정할 때, 두 모형 간에는 기본적으로 **내포성(nestedness)**이 확보되어야 한다. 다른 통제변수는 변화가 없어야 하고 표본도 동일하게 유지해야 한다.



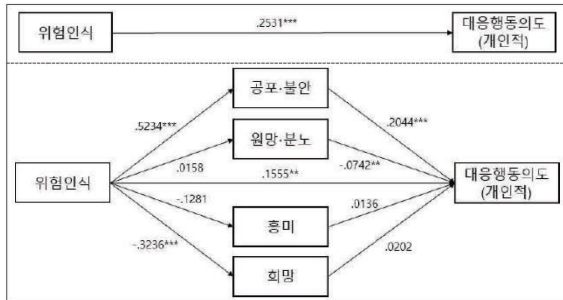
## 매개분석 연구 사례

# 매개분석 연구 사례

매개효과를 지지하는 이론적 모형이 필요하다.

- 가령 오수빈·윤순진 (2022)은 (1) 기후변화 위험 인식과 대응행동에 관한 선행연구를 먼저 검토하였고, 그 다음 (2) 위험에 대한 감정적 반응(affective response)과 인지적 평가이론(cognitive appraisal theory)을 검토했다.

〈그림 2〉 경로별 효과 크기('개인적 대응 행동 의도' 모형)



주: 위의 그림은 경로 총 효과, 아래 그림은 직접 효과를 의미함.  
( $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ )

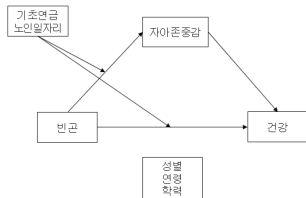
오수빈·윤순진. 2022. “기후위기 위험 인식과 대응 행동 의도 간 감정적 반응의 매개효과.” 『환경정책』 30(4): 53-97.

# 매개분석 연구 사례

- 김태환(2024)의 연구는 빈곤이 사회심리적 요인을 거쳐 노인의 건강수준을 결정하는 경로를 살펴보고, 그 경로에 있어서 재분배정책의 영향과 역할을 검토하였다.

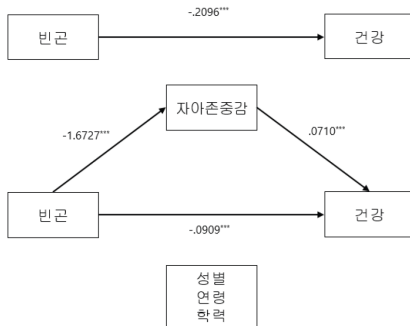
〈표 3〉 빈곤과 건강의 관계에서 자아존중감의 매개효과

변수	B	SE	t	p	LLCI	ULCI
종속변수: 자아존중감						
빈곤	-1.6727	.1075	-15.5525	.0000	-1.8835	-1.4618
성별	.2649	.1056	2.5083	.0122	.0579	.4720
연령	-.0705	.0073	-9.5892	.0000	-.0849	-.0561
학력	.4846	.0839	5.7767	.0000	.3201	.6490
$R^2=.1041$ , $F=173.4503$ , $p=.0000$						
종속변수: 건강						
빈곤	-.0909	.0231	-3.9274	.0001	-.1361	-.0455
성별	.0972	.0223	4.3638	.0000	.0535	.1408
연령	-.0189	.0016	-12.1080	.0000	-.0220	-.0158
학력	.0601	.0177	3.3874	.0007	.0253	.0948
자아존중감	.0710	.0027	26.0402	.0000	.0657	.0763
$R^2=.1873$ , $F=275.3677$ , $p=.0000$						
총효과, 직접효과						
구분	Effect	SE	t	p	LLCI	ULCI
총효과	-.2096	.0239	-8.7626	.0000	-.2564	-.1627
직접효과	-.0908	.0231	-3.9274	.0000	-.1361	-.0455
간접효과						
경로	Effect	BootSE	BootLLCI	BootULCI		
빈곤→자아존중감→건강	-.1189	.0091	-.1373	-.1012		



〔그림 2〕 연구모형

# 매개분석 연구 사례



[그림 3] 빈곤과 건강의 관계에서 자아존중감의 매개효과

김태환. 2024. “노인 빈곤과 건강의 경로 메커니즘 분석: 기초연금과 노인일자리의 조절된 매개효과.” 『노인복지연구』 79(3): 325-357.

- 김재엽·최지현(2009)은 여성 청소년을 조사하여 성폭력 피해 경험이 성폭력 피해에 대한 자기 귀인(attribution)을 거쳐 우울함에 이르는 인과 관계를 살펴보았다.

## 1) 성폭력

성폭력은 상대가 원하지 않는 성관련 언어나 행동을 일방적으로 표출하여 상대에게 정신적·신체적 피해를 주는 행위로 정의하는 바, 이에 성폭력 피해경험으로 성적놀림, 성추행, 강간미수와 강간을 포함하여 측정하였다. 성폭력 피해 경험 문항은 지난 1년간 '성적으로 놀림(농담 등)을 당한 적이 있다', '추근거리거나 의도적인 신체접촉(의도적으로 부딪히기, 몸 만지기 등)을 당한 적이 있다', '강제로 키스나 포옹, 애무를 당한 적이 있다', '강제로 성교를 당하려다가 실패한 적이 있다', '강제로 성교를 당한 적이 있다'의 5개 문항으로 구성되었다. 각 문항은 '없음(1점)', '1~2번(2점)', '3~5번(3점)', '6~9번(4점)', '10번 이상(5점)'의 5점 척도로 측정되었다. 분석에서는 각 문항의 점수를 합산하여 사용하였고, 점수가 높을수록 성폭력피해 경험이 많은 것으로 해석한다.

## 2) 성통념

본 연구에서는 왜곡된 성통념을 측정하기 위해 총 10개의 문항을 사용하였다. 왜곡된 성통념을 측정하기 위해 기존 연구에서 강간통념 척도를 활용하였으며, 여기서 강간통념이란 강간에 대해 가지고 있는 그릇된 통념(rape myth)을 말한다. 본 연구에 사용된 왜곡된 성통념 척도는 이석재(1999)의 강간통념 수용도 척도와 남미정(2004)의 강간통념 수용도 척도를 합한 후 중복되는 문항을 제외하고 본 연구의 목적에 따라 재구성하여 사용하였다. 측정문항은 '여자들이 조심하면 성폭력은 줄일 수 있다', '강제로 성관계 당한 여자는 그 사건에 대해 어느 정도 죄책감을 느껴야 한다', '성폭력은 노출이 심한 여름에 주로 젊은 여자들에게만 일어난다', '여자가 키스나 애무를 허용하는 것은 성관계를 허용하는 뜻이다' 등을 포함한 10문항이며, 각 문항은 '매우 아니다(1점)'에서 '매우 그렇다(5점)'의 5점 척도를 사용하였다. 분석에서는 각 문항의 점수를 합산하여 사용하였으며, 점수가 높을수록 왜곡된 성통념을 가진 것으로 해석한다. 본 연구에서 척도의 신뢰도는 Cronbach's  $\alpha=.782$ 로 나타났다.

## 3) 우울

본 연구에서는 우울을 측정하기 위해 Beck의 우울 척도(BDI : Beck Depression Inventory)를 사용하였다. BDI는 우울의 심각도를 측정하는 목적으로 사용되는 측정도구로 가장 널리 사용되고 있는 우울 척도 중 하나이다. 본 연구에서는 BDI의 문항 중 11개 문항을 추출한 후, 각 문항을 '전혀 그렇지 않다(1점)'에서 '항상 그렇다(5점)'의 5점 척도로 측정하였다. 우울 척도의 문항에는 '슬픔과 우울한 느낌으로 힘들었다', '나의 삶은 실패의 연속이라고 생각했다', '평소보다 많이 울었다', '앞날에 대해 기대할 것이 없다고 생각했다' 등이 포함되어 있으며, 각 문항의 값을 합산하여 그 값이 높을수록 우울한 것으로 해석한다. 본 연구에서 우울 척도의 신뢰도는 Cronbach's  $\alpha=.893$ 로 나타났다.

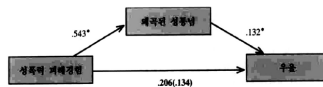
## 매개분석 연구 사례

(N=489)

	구분	B(SE)	$\beta$	sig.	VIF
통제변수	학년	.454(.512)	.041	.376	1.063
	학업성적	-.370(.248)	-.068	.136	1.036
	생활수준	.660(.446)	.068	.140	1.052
	종교 <sup>1)</sup>	.214(.651)	.015	.743	1.052
	부와의 대화시간	.007(.008)	.056	.392	2.158
	모와의 대화시간	-.008(.006)	-.086	.197	2.182
독립변수	성폭력 피해경험	.543(.171)	.141	.002**	1.002
	F		2.392*		
	R <sup>2</sup>		.034		
	Adj. R <sup>2</sup>		.020		

(*N*=489)

구분	독립변수→종속변수			독립, 매개변수→종속변수			
	B(SB)	$\beta$	VIF	B(SB)	$\beta$	VIF	
통제변수	학년	-.078(.614)	-.006	1.063	-.138(.611)	-.010	1.064
	학업성적	-.788(.297)	-.118**	1.036	-.739(.297)	-.111*	1.040
	생활수준	-1.114(.535)	-.094*	1.052	-1.200(.534)	-.101*	1.057
	종교 <sup>1)</sup>	1.706(.780)	.099*	1.052	1.678(.776)	.097*	1.053
	부와의 대화시간	-.010(.010)	-.069	2.158	-.011(.010)	-.075	2.162
	모와의 대화시간	-.015(.007)	-.142*	2.182	-.014(.007)	-.133*	2.190
독립변수	성폭력 피해경험	.206(.207)	.044	1.002	.134(.208)	.029	1.022
매개변수	애곡된 성품님				.132(.052)	.108*	1.035
F		5.317***			5.431***		
R <sup>2</sup>		.072			.083		
Adj. R <sup>2</sup>		.058			.068		
R <sup>2</sup> change					.011***		

<sup>\*\*</sup> $p < .01$  <sup>\*\*\*</sup> $p < .001$ <sup>\*\*</sup> $p < .01$  <sup>\*\*\*</sup> $p < .001$ 

〈그림 1〉 경로 분석 결과

김재업·최지현. 2009. "성폭력 경험이 청소년기 여학생의 우울에 미치는 영향과 왜곡된 성통념의 매개효과." 『아동학회지』 30(2): 231-247.