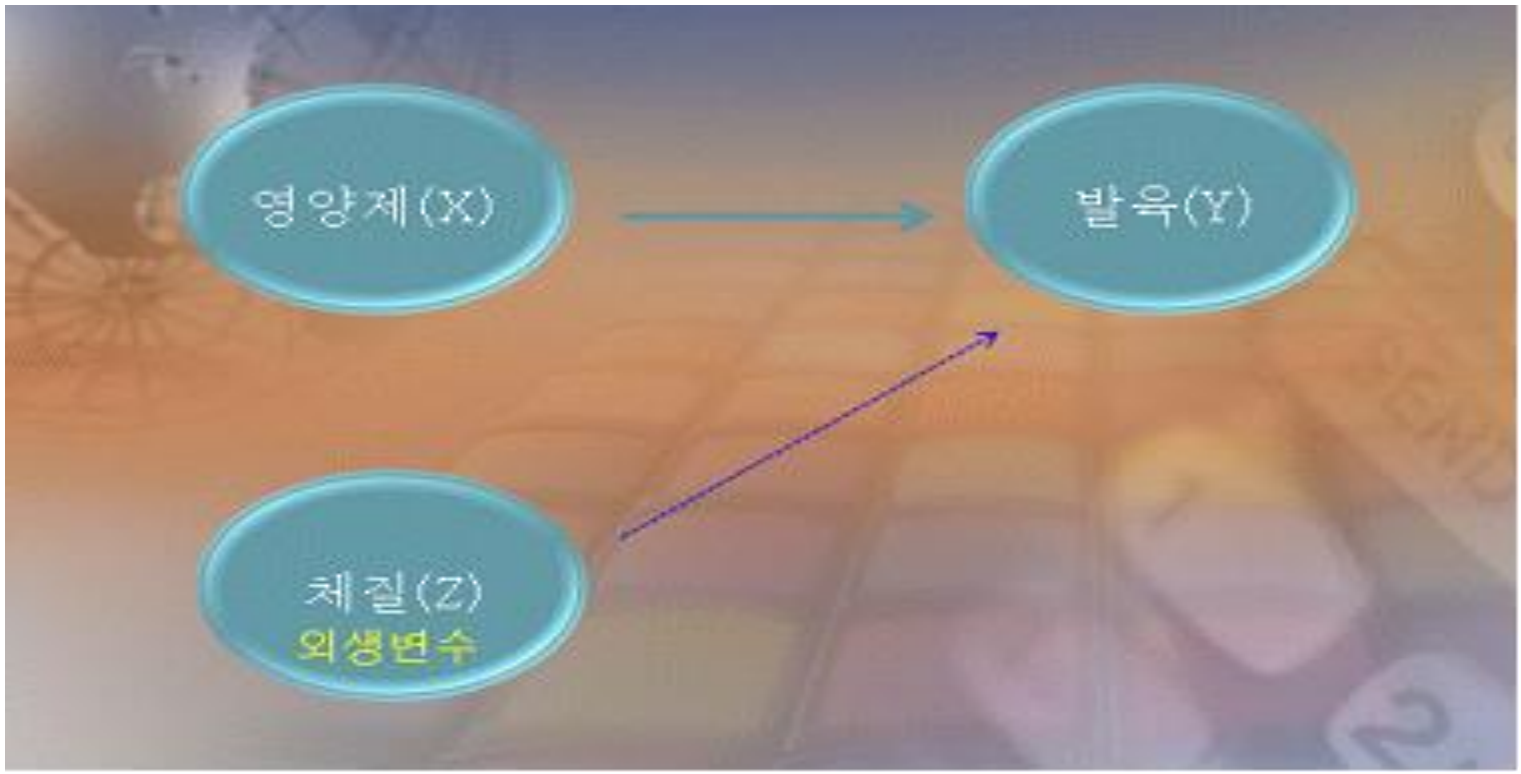
제2절 가 설

[그림 2-2] 독립변수, 종속변수

• 출처 : 신민철 저 “경영경제 통계학의 기초, 창민출판사, 2010년.

- 외생변수(extraneous variable)
  - 독립변수와 종속변수를 제외한 다른 모든 변수, 특히 실험연구에서 중시됨

[그림 2-3] 독립, 종속, 외생변수

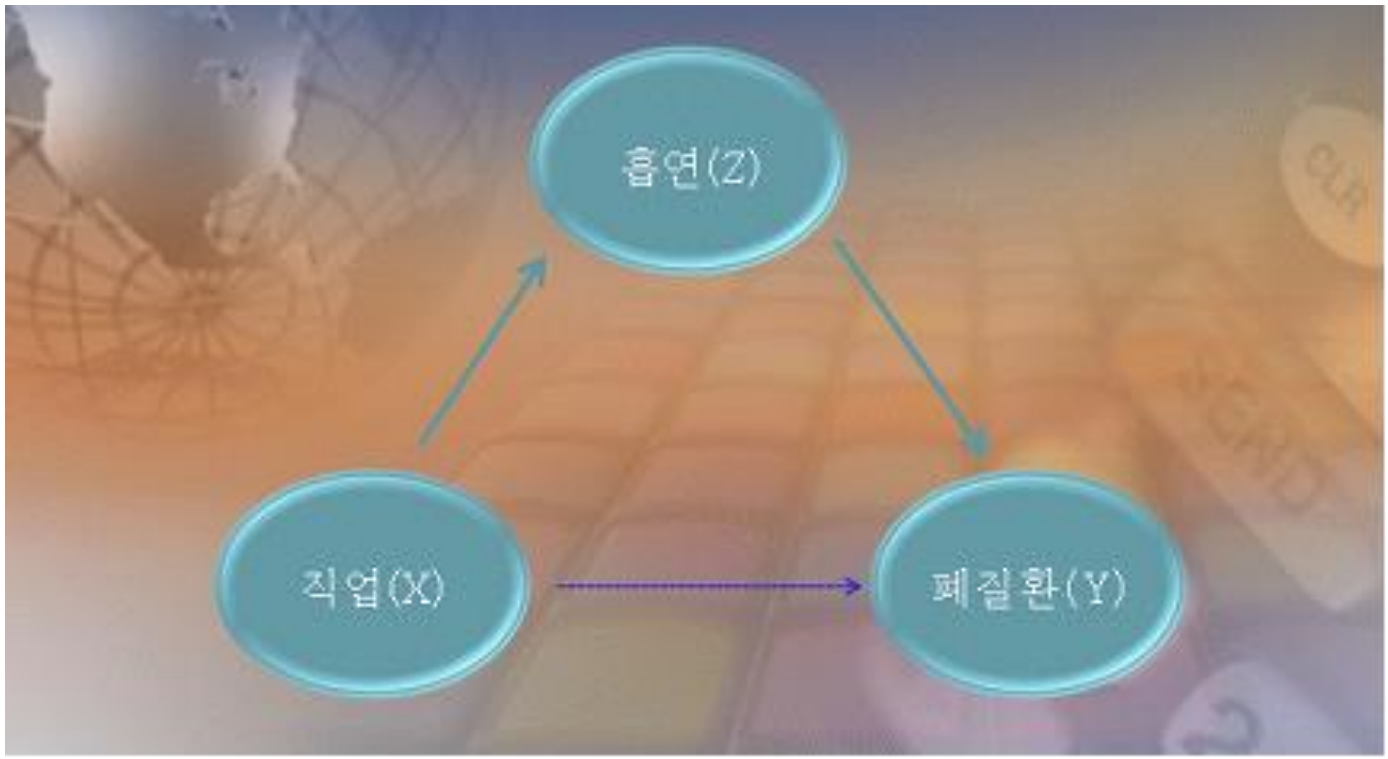


✦**매개변수**(intervening variable 혹은 mediating variable)

- 독립변수와 종속변수 사이의 관계를 중간에서 설명해 주는 변수

✦ **의사관계**(擬似關係: spurious relationship)

[그림 2-4] 독립, 종속, 매개변수



- **조절변수**(moderating variable)
  - 독립변수가 종속변수에 미치는 영향을 강화해 주거나 약화해 주는 변수
  - **조건변수**(conditional variable), **상호작용변수**(interaction variable)

- 조절변수 vs. 매개변수

[그림 3-5] 독립변수, 종속변수, 조절변수

(3) 이산변수와 연속변수

✦ 이산변수(discrete variable)

- 범주형 변수/유목변수(categorical variable)
- 취할 수 있는 값(values)이 단절적인 혹은 일정한 수로 한정된 변수

예) 성별(남, 여), 직업(전문직, 비전문직, ..., 기타) 등

✦ 연속변수(continuous variable)

- 연속적인 값(values)이 의미를 가지는 변수

예) 체중, 키, 시간 등

✦ 모수통계기법(parametric statistical analysis) vs.
비모수통계분석기법(non parametric statistical analysis)

* 가능하다면 연속변수로 측정하는 것이 자료수집 후 다양한 통계분석에 유리

1. 검증대상 가설 설정

연구문제 = 언어로 표시한 의문문
가설 = 연구문제에 대한 잠정적인 답

⇒ 가설이 참인지 여부를 분석할 필요
⇒ 잠정적인 답이 연구문제를 해결하는 진실된 답이 되는지 여부를 확인

⇒ 검증대상 가설은 어떻게 설정해야 할까요?

- ✦ 기본적으로 '만약 ~라면 ~할 것이다(if A, then B)'라는 형태로 설정
- ✦ 검정대상이 되는 가설 유형은?

제3절 가설검정의 논리적 기초

<표 2-5> 연구문제와 가설의 유형

구 분	예
연구문제	학력에 따른 소득의 차이가 있을까? (왜 자녀교육을 중요시 하는 것일까?)
가설 (a)	교육연수는 연소득에 영향을 줄 것이다.
가설 (b)	교육연수는 연소득에 영향을 주지 않을 것이다.
가설 (c)	대졸 이상자 연평균소득 > 대졸 이하자 연평균소득
가설 (d)	대졸 이상자 연평균소득 = 대졸 이하자 연평균소득

가설 (a)= 연구문제에 대한 잠정적인 해답 중 가장 직접적인 형태

가설 (b)= 가설 (a)를 부정하는 형태의 문장

가설 (c)= 가설 (a)를 통계적인 표현(즉, 수식)으로 나타낸 형태

가설 (d)= 가설 (c)를 부정한 형태 => *검정 대상이 되는 가설은?*

2. 가설의 유형

1) 연구가설과 영가설

(1) 연구가설(research hypothesis)/대립가설(alternative hypothesis) -> H_1

- 검증하려고 하는 경영/경제 현상간의 관계를 그대로 나타내는 가설형태

예1) '교육연수는 연소득에 영향을 줄 것이다'
예2) '광고비 증가는 매출을 증가시킬 것이다'

(2) 영(零)가설/귀무(歸無)가설(null hypothesis) -> H_0

- 검증하려고 하는 현상간의 관계를 부정하는 형태의 가설

예1) '교육연수는 연소득에 영향을 주지 않을 것이다'
예2) '광고비 증가가 매출을 증가시키는 것은 아니다'

<표 2-6> 연구가설과 영가설

연구/대립가설(H_1)	영/귀무가설(H_0)
교육연수는 연소득에 영향을 줄 것이다.	교육연수는 연소득에 영향을 주지 않을 것이다.
광고비 증가는 매출을 증가시킬 것이다.	광고비 증가가 매출을 증가시키는 것은 아니다.

제3절 가설검정의 논리적 기초

(3) 왜 영가설을 가설검정에 사용?

- ✦ Sir Fisher
- ✦ Popper(1963)의 반증주의(Falsificationism)

⇒ '후건(後件) 긍정의 오류(fallacy of affirming the consequent)'를 회피하기 위해 (경험적으로 관찰된 Y는 X가 아닌 다른 요인의 영향에 의한 것일 수도 있기 때문)

<표 2-7> 후건 긍정의 오류

* 연구가설의 반대관계(즉, 현상간의 관계가 없다는)를 표현한 영가설을 '기각'하는 것이 논리적인 오류를 최소화시킬 수 있는 방법임-> 영가설을 기각하는 것이 가설검정의 목표임, 영가설이 기각되면 연구가설이 수용된 것으로 추론하게 됨

(가) If X, then Y	(나) Y, therefore X
만일 공장이 강을 오염시킨다면, 물고기의 죽음이 증가할 것이다.	물고기의 죽음이 증가했다. 공장이 강을 오염시켰기 때문이다.

2) 실질가설과 통계가설

(1) 실질가설(substantive hypothesis)

- 연구하고자 하는 현상간의 관계에 대한 추측성 문장(conjectural statement)

  예) "광고비를 증가시키면 매출이 증가할 것이다"

(2) 통계가설(statistical hypothesis)

- 연구 현상간의 관계를 계량적으로 표현하는 가설

⇒ 검증 가능성 고려

<표 2-8> 실질가설과 통계가설

실질가설	통계가설
교육연수는 연소득에 영향을 줄 것이다.	대졸이상자 연소득평균 > 대졸미만자 연소득평균
가톨릭신자의 자녀수가 개신교신자의 자녀수보다 많을 것이다.	가톨릭신자 자녀수평균 > 개신교신자 자녀수평균

3) 검증대상 가설

- 연구문제에 대한 잠정적인 해답인 가설에는 4개 유형 존재

⇒ 4가지 유형 중 영가설이자 통계가설 형태로 검정

<표 2-9> 가설의 유형

	연구/대립가설(H_1)	영/귀무가설(H_0)
실질가설	교육연수는 연소득에 영향을 줄 것이다.(a)	교육연수는 연소득에 영향을 주지 않을 것이다.(b)
통계가설	대졸이상자 연소득평균 > 대졸미만자 연소득평균(c)	대졸이상자 연소득평균 = 대졸 미만자 연소득평균(d)

3. 가설의 특징

(1) 가설의 명확성

- 가설을 구성하는 모든 변수를 명확히 정의
  ⇒ 개념적 정의(conceptual definition)와 조작적 정의(operational definition)를 포함

(2) 가설의 구체성(specific)

- 현상간 관계의 방향(direction) 및 현상간의 관계 성립 조건 등을 특정(specify) 해야 함 -> 현상간의 관계가 이럴 수도 있고 저럴 수도 있다는 문장은 가설이 될 수 없음

(3) 가설의 검증 가능성

- 현재 사용 가능한 방법론(available methods)을 이용한 진위의 검정이 가능

(4) 당연한 관계는 가설이 될 수 없음

예1) "사람은 모두 죽는다"
예2) "사람은 먹지 않고 살 수 있다" -> 참일 수도 있고 거짓일 수도 있는 문장만이 가설 후보가 될 수 있음

(5) 가설은 동어반복(tautological)을 피해야 함

예) "그는 젊지 않으면 늙을 것이다" -> 가설은 두 개의 서로 다르지만 관련이 있다고 기대되는 개념간의 관계를 표시해야 함

[도입사례]

(A) 국내 기업의 경우 삼성전자, SK Telecom을 포함한 몇몇 대기업은 연간 국내 광고비만 1,000억원 이상 집행하고 있습니다. 여러분이 광고관리자라고 가정하지요. 광고인지도가 높아지면 광고의 설득효과가 높아지는지 알아보라는 최고경영층의 지시가 있다고 상정하겠습니다. 광고인지도와 광고의 설득효과간의 관계를 알아내기 위해 '추상적인' 개념인 광고인지도나 광고의 설득효과를 어떻게 '경험적으로' 나타낼 수 있을까요?

(B) 국내총생산(Gross Domestic Product: GDP)은 일정 기간 중 한 국가가 생산한 재화와 용역의 (부가)가치를 측정하는 도구이고, 1인당 GDP는 경제적 생활수준(economic living standards)의 측정도구로 사용됩니다. GDP는 기본적으로 자국의 통화단위로 측정되는데, GDP의 국가 간 직접 비교가 의미가 있을까요?

(C) 우리는 보통 몇 가지 기준을 가지고 배우자 후보를 평가하는 경향이 있습니다. 대표적인 기준은 아마 외모, 능력, 건강, 사람됨 등이 되겠지요. 그렇다면 우리는 어떻게 배우자감을 구체적으로 판단하게 될까요?

생각해 볼 문제

① 경영/경제 현상을 측정할 수 있는 도구는 어떻게 만들 수 있을까요?
② 측정은 통계학과 어떠한 관계가 있을까요?

1. 측정이란 무엇인가?

✦ 측정(measurement)

- "현상이 지니고 있는 추상적(abstract)인 특징(예: 배우자 후보의 능력의 정도)을 일정한 규칙에 따라 우리가 경험할 수 있는 구체적(concrete)인 사물(예: 연봉 1억원)과 연결하는 과정"
- "특정한 규칙에 따라 현상에 숫자(數字, numeral)를 부여하는 것"
- 추상적인 개념으로 파악한 사회현상의 속성을 구체적인, 즉 우리가 경험할 수 있는 숫자(numeral)로 나타내는 것

⇨ 사회현상의 연구에 다른 기호보다 양적인 의미(quantitative meaning)를 보유한 기호인 수(number)를 자주 사용하는 이유는 수들(numbers)간의 관계는 통계학을 이용하여 엄밀하게 분석할 수 있기 때문

- 숫자(numeral)
  - 아라비아 숫자(Arabic numerals), 로마숫자(Roman numerals) 등의 예에서 알 수 있듯이 '1, 2, 3...' 혹은 'I, II, III...' 등의 형태를 가지는 부호(sign)

- 수(number)
  - 양적인 의미(quantitative meaning)가 부여된 숫자

  cf. 직접측정 vs. 간접측정

2. 측정의 수준

✦ 측정의 수준(levels of measurement)<- 각 기호가 보유하는 정보의 내용에 따라 구분 가능

- 다양한 사회현상의 속성을 재는 기본적인 자
- 근원척도의 유형(types of primary scale) or 종류

1) 명목척도

✦ 명목척도(nominal scale)

- 현상의 속성에 부여된 수(number)가 그 현상의 속성이 어떠한 유형으로 구분되는지만을 알려주는 경우, 이러한 정보만을 가진 자

⇨ 명목척도의 특징을 가진 자로 잰 수가 제공하는 정보는, 유형(category)만을 구분, 이렇게 수집된 자료에 적용 가능한 통계분석기법이 제한됨

⇨ 사칙연산 무의미

예) 성별

<표 3-1> 명목척도의 예

변 수 / 응답자	성 별	소유자동차
응답자 1	남 (=1)	Sonata (=1)
응답자 2	여 (=2)	SM5 (=2)
:	:	:
:	:	:
응답자 99	남 (=1)	SM5 (=2)
응답자 100	여 (=2)	Sonata (=1)

2) 서열척도

- 서열척도(ordinal scale)
  - 사용되는 수가 유형(category)의 구분은 물론 조사현상의 순위정보까지 알려주고 있는 자

⇨ 측정된 수들(numbers)간의 크기 비교는 무의미

<표 3-2> 서열척도의 예

변 수 / 응답자	화장품 A	화장품 B	화장품 C
응답자 1	1(위)	2(위)	3(위)
응답자 2	1	3	2
:	:	:	:
:	:	:	:
응답자 99	2	1	3
응답자 100	1	2	3

3) 등간척도

- 등간척도(interval scale)
  - 사용되는 수가 집단, 순위 및 동등 간격이라고 하는 3종류의 정보를 가지고 있는 자, 즉 척도

    예) 상품의 선호도, 온도계, IQ 점수, 심리검사점수 등

<표 3-3> 등간척도의 예

변 수 / 응답자	선호도	만족도	재구매의도
응답자 1	5	5	5
응답자 2	5	4	5
:	:	:	:
:	:	:	:
응답자 99	4	4	3
응답자 100	2	2	1

4) 비율척도

✦ 비율척도(ratio scale)

- 사용되는 수(number)에 집단, 순서, 동일간격의 3종의 정보에다가 절대적인 0(absolute zero)이라고 하는 정보까지 포함한 가장 정보량이 많은 유형의 자, 즉 척도
- 0이 의미를 가지는 유형의 자 cf. 온도계(0도가 절대적인 기준점이 아님)

  예) 무게, 길이, 소득, 매출액, 이자율, 시험점수 등

<표 3-4> 비율척도의 예

변 수 / 응답자	국 어	영 어	수 학
응답자 1	100	100	100
응답자 2	95	94	85
:	:	:	:
:	:	:	:
응답자 99	34	64	53
응답자 100	12	0	1

• 출처 : 신민철 저 “경영경제 통계학의 기초, 창민출판사, 2010년.