

데이터 분석 방법론

1

□ 연속형 변수와 이산형 변수

- 1) 연속형 변수(continuous variable): 키, 몸무게 또는 5 점 척도의 문항, 평균이 의미가 있는 데이터
예) 키, 몸무게, 수학성적, 자기효능감 점수, 5 점 척도 문항 등
- 2) 이산형 데이터(discrete variable): 성별, 학년, 직업, 종교와 같이 평균이 의미 없는 데이터
예) 일련의 인구사회학 변수, 성별, 학년, 직급, 직업, 가족형태 등

□ 데이터 속성에 따른 분석 방법론

독립변수와 종속변수가 어떤 데이터 속성을 가지냐에 따라 분석기법이 다름

□ 독립적인 데이터

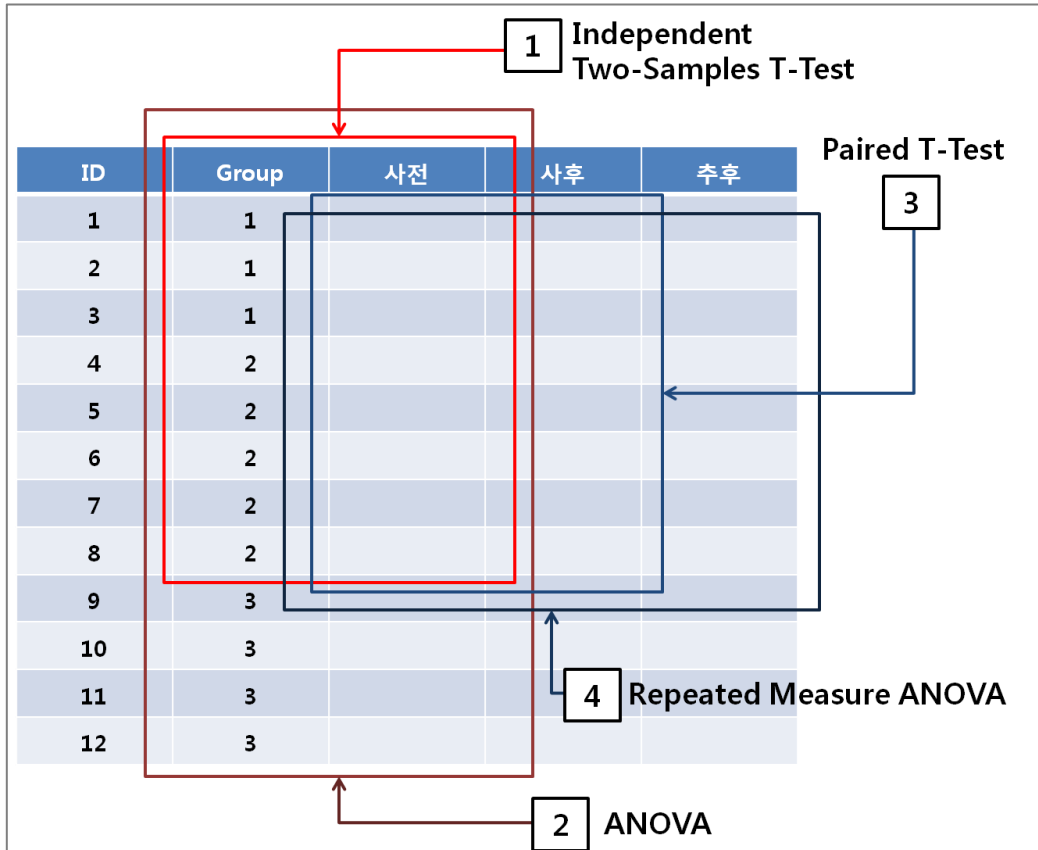
변수 구분		종속변수	
		연속형	이산형
독립변수	연속형	상관분석, 회귀분석	로지스틱 회귀분석 순서형 프로빗 모형
	이산형	T-Test(Level 2) ANOVA(Level 3+)	Chi-square Test

□ 반복 실험 데이터

반복 실험회수	분석기법
2 회(사전-사후)	대응표본 T-Test(Paired T-Test)
2 회(사전-사후) - 사전동질성 위배	공분산분석(ANCOVA)
3 회 이상(1 회차~n 회차)	반복측정(Repeated Measure)

□ 데이터 구조로 본 차이검정 방법론

2



1) Independent Two-Samples T-Test(독립 이표본 t-검정)

- A. 두 집단간의 평균에 대한 차이검정으로 독립변수는 두 개의 범주로 구성된 그룹변수이며, 종속변수는 연속형 변수로 설정된 경우
- B. 예시
 - i. 성별에 따른 학업성취도의 차이
 - ii. 재발여부에 따른 비만도의 차이

2) ANOVA(분산분석)

- A. 독립적인 세 집단 이상간의 평균차이 검정으로, 독립변수는 세 개 이상의 범주로 구성된 명목변수이고 종속변수는 연속형 변수로 설정된 경우
- B. 예시
 - i. 1~3 학년의 자아존중감 점수에 대한 차이
 - ii. 강수량 수준(3 개 수준 이상)에 따른 쌀 수확량에 대한 차이

3) Paired Samples T-Test(대응표본 t-검정)

- A. 단일 또는 그 이상의 집단 내에서 사전-사후 점수의 변화가 유의한 지를 검증하는 방법
- B. 어떤 효과(effect)를 주기 전과 주고 난 이후의 변화를 검증하는데 많이 적용됨
- C. 예시
 - i. 100m 달리기 전과 그 이후의 심장 박동수의 차이
 - ii. 음악치료 프로그램 이전과 이후의 음향 이해 정도의 차이

4) Repeated Measure ANOVA

- A. 동일 개체가 3 회 이상의 반복적인 측정을 통해 생성된 데이터가 있을 경우, 실험수준의 변화 또는 시간 추이에 따른 측정치의 변화에 대한 유의성 검증
- B. 사전-사후 또는 이전과 그 이후에 대한 차이는 Paired T-Test 를 적용하며, 실험처치(treatment)가 3 회 이상일 경우는 반복측정 ANOVA 로 구현

□ 교육학 또는 실험연구에서 자주 적용되는 분석 예

4

ID	Group	사전	사후
1	실험군		
2	실험군		
3	실험군		
4	실험군		
5	실험군		
6	실험군		
7	실험군		
8	통제군		
9	통제군		
10	통제군		
11	통제군		
12	통제군		
13	통제군		
14	통제군		

Diagram illustrating the experimental design with annotations:

- 2**: Points to the first 7 rows (Experimental Group).
- 2'**: Points to the last 7 rows (Control Group).
- 1**: Points to the '사전' (Before) column for the Experimental Group.
- 3**: Points to the '사후' (After) column for the Control Group.

1) 사전 동질성 검증

- 실험군과 통제군의 사전점수가 동질한 지를 검증하며 독립 이표본 t-검정을 통해 검증함
- 사전 동질성이 가정되어야만, 사후점수의 변화를 공정하게 검증 가능함
- 사전 동질성이 위배될 경우, 공분산 분석(ANCOVA) 또는 PSM(경향점수 매칭: propensity score matching)을 통해 검증하게 됨

2) 사전-사후 점수에 대한 차이검증

- A. 실험군과 통제군 각각의 사전-사후 차이검증을 대응표본 t-검정(paired samples t-test)을 통해 검증함
- B. 일반적으로 실험군에서는 사전-사후 차이검정에서 통계적으로 유의한 결과를 기대하게 되며, 통제군의 경우는 사전-사후 차이검정에서 통계적으로 유의한 결과가 실험군 대비 상대적으로 적을 것을 기대함

3) 사후점수에 대한 프로그램 효과검증

- A. 실험군과 통제군의 사후점수에 대한 차이검증을 독립 이표본 t-검정을 통해 검증하게 됨.
- B. 프로그램 효과검증은 사전동질성을 가정한 실험군과 통제군의 사후점수에 대한 차이검증으로 진행하게 됨.

Main Tip> 분석기법의 적용

데이터가 존재한다면, 먼저 종속변수와 독립변수에 대한 정의를 해야 하며, 독립변수가 연속형 데이터인지, 명목형 데이터인지, 그리고 종속변수가 연속형 데이터인지, 명목형 데이터인지에 따라 분석기법의 적용이 달라진다.

독립변수가 명목형이고 종속변수가 연속형 변수인 경우, 일반적으로 차이검정을 하게 된다. 남녀에 따른 비만도에 대한 차이, 학년에 따른 학업성취도에 대한 차이가 이런 예시에 해당한다. **결과적으로 독립변수와 종속변수의 데이터의 속성이 어떠냐에 따라 분석기법을 결정짓게 된다.**