Healthcare Statistics

Lecture 6 : 평균치 비교

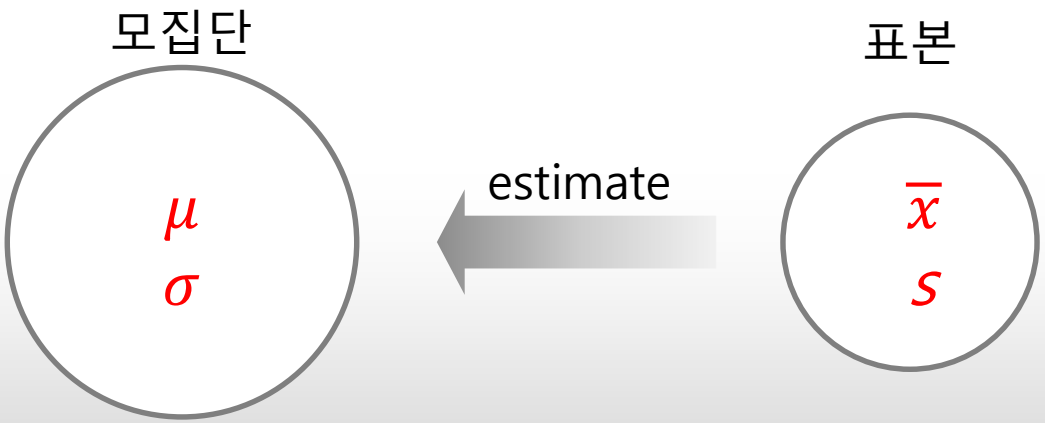
의료경영학과 김 성 수

청주대학교
CHEONGJU UNIVERSITY

T-test



- 목적
  - 1908년 흑맥주의 질을 보존하기 위해 t-검정을 개발한 화학자 윌리엄 설리 고셋(William Sealy Gosset, 1876~1937) : *동영상 클릭*
  - 모집단의 분산이나 표준편차를 알지 못할 때 모집단을 대표하는 표본으로부터 추정된 분산이나 표준편차를 가지고 모집단의 평균을 추정하는 통계기법

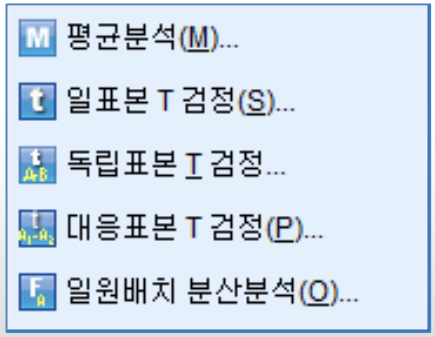


T-test

- 종류
  - 기본적으로 두 집단의 평균을 비교하기 위한 분석 기법
  - 표본의 집단수와 측정횟수에 따라서 몇 가지로 나뉜다.

종류	평균비교를 위한 분석방법		표본집단수	측정횟수
t-검정	일표본 T 검정	One-sample T test	1개	1회
	대응표본 T 검정	Paired T test	1개	2회
	독립표본 T 검정	Independent sample T test	2개	1회
분산분석	일원배치분산분석	ANOVA(Analysis of Variance)	3개 이상	1회

※ SPSS P/G : 분석 → 평균비교(M) →



T-test

- 일표본 t-검정
  - 두 집단 간 평균 차이를 비교하기 위해 집단 속성이 존재
  - But 두 집단중 한 집단의 속성을 나타내는 데이터만 구하는 경우
  - Ex) 전국 대학생 평균 키와 OO대학 OO학과 학생의 키의 비교
  - 가설
    - H_0 : 전국 대학생과 우리학과 학생들의 평균 키는 차이가 없다.
    - H_1 : 전국 대학생과 우리학과 학생들의 평균 키는 차이가 있다.
  - 통계량

$$t = \frac{\bar{X} - \mu_0}{s_{\bar{X}}} \qquad s_{\bar{X}} = \frac{s}{\sqrt{n}}$$

$s_{\bar{X}}$ = 표준오차

T-test

- 독립표본 t-검정
  - 목적
    - 독립된 2집단간의 평균차이를 검증하는 분석방법 즉, 두 집단의 평균이 통계적으로 유의한 차이를 보이고 있는지 여부를 검증
      ex) 노인환자와 비노인 환자의 평균재원일수 차이?
  - 원리
    - 각 표본의 분산과 두 표본을 합한 전체 집단의 분산을 이용하여 평균의 차이가 어느 정도 유의 한가를 검정하는 것이다.
    - 두 집단간의 평균의 차이를 검정하기 위해서 t값을 계산한다.
    - 즉, 평균의 차이 값이 단지 우연에 의한 것인지 아니면 통계적인 유의한 차이에 의한 것인지를 밝히는 과정

T-test

- 통계량
  - t-검증 통계량 각 표본의 분산과 두 표본을 합산한 전체집단의 분산을 이용
  - 두 집단의 평균차이 검증

$$\text{각집단의표준오차} = \frac{\text{각집단의표준편차}}{\sqrt{\text{사례의수}}}$$

$$\text{전체의표준오차} = \sqrt{\sum(\text{각집단의표준편차})^2}$$

$$t = \frac{\text{제1평균} - \text{제2평균}}{\text{전체의표준오차}}$$

T-test

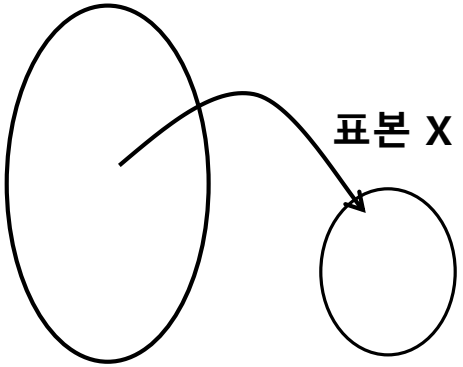
- 가설설정
  - H_0 : 두 집단간에 평균의 차이가 없다.
  - H_1 : 두 집단간에 평균의 차이가 있다.
- 주의사항
  - T-test는 반드시 분산의 동질성을 먼저 검정하여야 함
  - 두 집단의 모분산이 같을 때 : Pooled Variances Estimate 이용
  - 두 집단의 모분산이 다를 때 : Separate Variances Estimate 이용
- 표작성의 예

<표00> 성별에 따른 건강수준의 차이

sex	mean	SD	t	p
male	3.80	.94	-1.11	.269
female	4.00	.64		

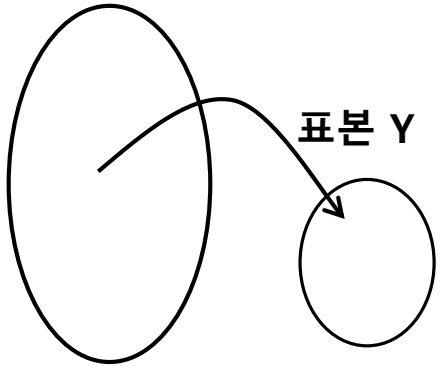
T-test

모집단 A



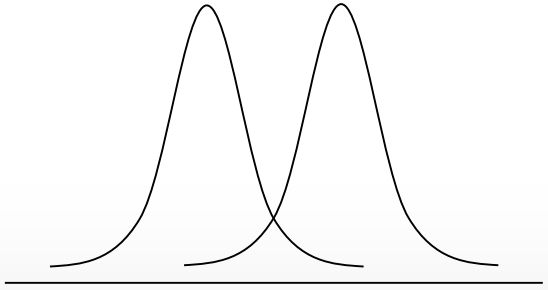
정규성 : $A \sim N(\mu_A, \sigma_A^2)$

모집단 B

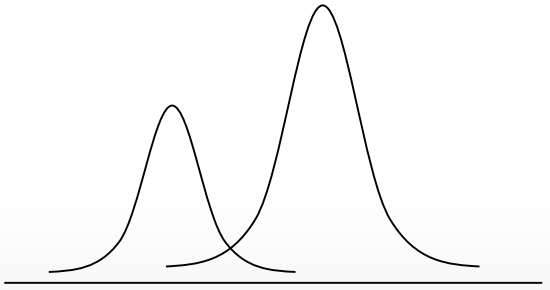


$B \sim N(\mu_B, \sigma_B^2)$

등분산



이분산



기본 가정 : 두 모집단의 분포는 정규분포이고 분산은 $\sigma_A^2 = \sigma_B^2$ 으로 동일해야 한다.

즉, 위의 등분산 Graph의 형태이다.

T-test

- 대응표본 t-검정
  - 동일한 표본을 대상으로 다른 2개 변수의 평균치를 비교
  - 같은 대상에 대하여 전후 측정치의 비교에 사용
  - Ex) 새로운 고혈압 약제를 개발하여 이 약의 투약 전후 혈압의 평균을 비교
  - 가설
    - H_0 : A와 B간에는 평균의 차이가 없다.
    - H_1 : A와 B간에는 평균의 차이가 있다.

T-test

- 주요 t-test의 비교

독립표본 T 검정(Independent T)	대응표본 T 검정(Paired T)

• 독립된 2 표본의 평균의 차이 분석
• 기본 가정 : 정규성, 등분산성
• 가설
  • H_0 : $\mu_A = \mu_B$
  • H_1 : $\mu_A \neq \mu_B$
• 종속변수 : 등간, 비율
• 독립변수 : 명목 or 서열(수준 2)
	• 짝을 이룬 자료의 평균 변화의 분석
	• 기본 가정 : 정규성
	• 가설
	  • H_0 : $\mu_{A-B} = 0$
	  • H_1 : $\mu_{A-B} \neq 0$
	• 종속변수 : 등간, 비율
	• 독립변수 : 등간, 비율

비모수 : Mann-Whitney U Test	비모수 : Wilcoxon 순위합 검정

T-test

- 실습
  - P245 : 단일 평균치 분석
    (일표본 T검정 : One-sample T test)
  - P252 : 독립된 두 평균치 비교
    (독립표본 T 검정 : Independent-samples T test)
  - P259 : 짝지은 두 평균치 비교
    (대응표본 T 검정 : Paired T test)

실습문제

- 자료실 파일 다운로드 : 환자만족1.xlsx 파일 이용

1) 20세 이상 입원환자를 대상으로 아래와 같은 분석을 실시하시오. (15점)

성별	전반적 만족도		t값	p값
	평균	표준편차		
남자				
여자				

- 귀무가설 : ( )
- 대립가설 : ( )
- 5%유의수준을 기준으로 가설 검정한 결과를 기술하시오.

_____

_____

_____

실습문제

2) 전체 입원환자를 대상으로 분석하여 아래의 물음에 답하시오. (10점)

변수	평균	표준편차	t값	p값
친절만족				
치료만족				

- 귀무가설 : ( )
- 대립가설 : ( )
- 5%유의수준을 기준으로 가설 검정한 결과를 기술하시오.

_____

_____

_____

실습문제

3) 전체 입원환자를 대상으로 아래와 같은 분석을 실시하시오. (15점)

- 평균만족도 : 시설만족, 이용만족, 친절만족, 치료만족 변수를 모두 더하여 평균한 것.

거주지	평균만족도		t값	p값
	평균	표준편차		
서울, 경기				
서울/경기 이외 지역				

- 귀무가설 : ( )
- 대립가설 : ( )
- 5%유의수준을 기준으로 가설 검정한 결과를 기술하시오.

_____

_____

_____