

제3절 추리통계분석은 어떻게 실행하는가?

◆ 양측/양방향 검정(two-tailed test)

- ◆ 연구가설(귀무가설)의 내용이 예측방향을 포함하지 않는 경우 수행하는 검정방법

⇒ 일반적으로 단측검정이 양측검정보다 더 강력한(엄격한) 결론을 보여 줌

예) 한편 [그림 8-13]은 표본분포가 정규분포를 따를 때, 단측검정과 양측검정의 차이 점을 정규분포곡선을 이용하여 설명

⇒ [그림 8-13]의 단측검정의 가상의 예에서는 <표 8-3>의 (a)의 경우를 도시

- ◆ 영가설이 기각되려면, 표본에서의 '구독자소득평균(\bar{X}_1)-비구독자평균(\bar{X}_2)의 값이 $\mu_1 - \mu_2 = 0$ (모집단에서 소득평균의 차이)에서 최소 +1.65표준오차($\alpha=0.05$ 수준에 부합하는 Z값 * 표준오차)만큼 떨어져 있어야 함을 보여주고 있음

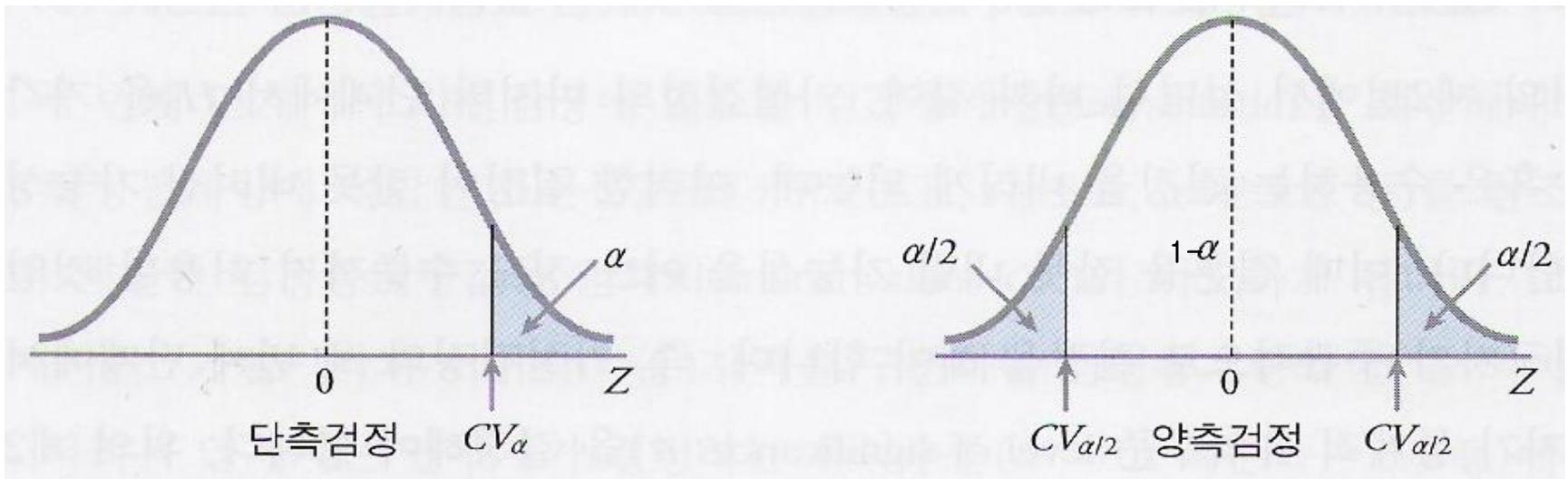
⇒ 양측검정의 경우(<표 8-3>의 (c))에는 +1.96표준오차보다($\alpha = .025$ 수준에 부합하는 Z값 * 표준오차, 즉 임계치[$=CV_{\alpha/2}$]의 크기) 크거나 혹은 -1.96표준오차보다 작아야 함을 보여주고 있음

제3절 추리통계분석은 어떻게 실행하는가?

<표 8-3> 단측검정과 양측검정

	연구가설(H_1)	귀무가설(H_0)
단측검정 (a)	구독자소득 > 非구독자소득	구독자소득 \leq 非구독자소득
단측검정 (b)	구독자소득 < 非구독자소득	구독자소득 \geq 非구독자소득
양측검정 (c)	구독자소득 \neq 非구독자소득	구독자소득 = 非구독자소득

[그림 8-13] $\alpha = .05$ 수준에서의 단측검정과 양측검정



제3절 추리통계분석은 어떻게 실행하는가?

(2) 통계분석방법 및 검정통계량 선택

- ◆ 설정된 가설을 검정하기 위해 적절한 통계분석방법을 선택
- ⇒ 특정 가설은 다양한 통계방법으로 검정 가능한 것이 일반적
- ⇒ 연구자는, 근원척도의 유형, 표본통계량(sample statistic)의 표본분포(sampling distribution), 검정통계량(test statistic) 등을 고려하여 가설 검정에 적절한 통계분석기법을 선택

※ 검정통계량(test statistic)

- ◆ 표본통계량(sample statistic)이 H_0 에서 모수(parameter)에 대해 예측하는 수준에 얼마나 근접했는지를 판단하게 해주는 도구
- ◆ 추리통계분석기법에 따라 다양한 유형이 개발되어 있음
- ⇒ SPSS Statistics 등을 이용하여 통계분석기법을 실행하는 경우에는 연구자가 선택한 통계기법에 맞는 검정통계량을 SPSS에서 계산해 주기 때문에 연구자는 검정결과를 적절히 해석하기만 하면 됨

가설검정의 목적은 귀무가설의 타당성 여부 판단에 있는데, 이 결정의 기준이 되는 표본통계량→ 검정통계량(test statistic)
예) 모평균에 대한 가설검정→ 검정통계량은 모평균에 대한 통계량인 표본평균으로부터 구할 수 있음

제3절 추리통계분석은 어떻게 실행하는가?

(3) 통계적 유의 수준 결정

- ◆ 가설검정의 마지막 단계에서 H_0 을 기각하거나 혹은 수용하는 결정을 내리게 되는데, 이러한 결정이 잘못 내려질 가능성 [=통계적 유의수준 (level of significance: α)]을 어느 정도 수준까지 허용할 것인가를 연구자가 주관적으로 결정
 - ⇒ 관례적으로 1종 오류(= α)가 0.01 이하 수준인 통계적 추정결과는 상대적으로 정확한 결론 : 통계적 유의도(α)를 0.01 혹은 0.05로 정하는 것이 관행
 - ⇒ 한편, SPSS Statistics를 사용하는 가설검정에서는 통계적 유의 수준을 SPSS명령에 반영하는 옵션은 없으나, “영가설(H_0)이 참이라는 전제하에 표본에서 계산된 검정통계량값이 관찰될(표본분포에서 발생할) 확률”이라고 할 수 있는 유의수준(p 값)이 바로 제시됨
 - ⇒ 연구자는 p 값이 α 보다 작거나 같으면 H_0 를 기각하는 결론을 내림

제3절 추리통계분석은 어떻게 실행하는가?

(4) 통계분석(검정통계량 계산)

- ◆ 둘째 단계에서 정한 통계분석기법을 실행하고 가설검정에 필요한 검정 통계량(test statistic)을 계산
- ⇒ 특정 통계기법의 분석명령을 SPSS가 실행하게 되면 관련된 검정통계량이 동시에 계산되고 특정한 검정통계량에 상응하는 **관찰된 유의수준** (단측/양측 검정), 즉 **p 값**도 동시에 제시됨

제3절 추리통계분석은 어떻게 실행하는가?

(5) 계산된 검정통계량과 임계치(검정통계량의 p 값과 유의수준 α) 비교

- ◆ 산출된 검정통계량에 적합한 표본분포표(정규분포표, t 분포표, χ^2 분포표 등)를 참조하여 이미 계산된 검정통계량값을 임계치와 비교

※ 임계치(critical value)

- ◆ 가설기각과 비기각(수용)을 구분하는 검정통계량의 값
 - ⇒ 연구자가 결정하는 유의수준($\alpha = .01$ 혹은 $\alpha = 0.5$)에 의해 결정
 - ⇒ SPSS를 실행하면 검정통계량값, 임계치, 검정통계량값에 부응하는 p 값이 바로 제시됨
 - ⇒ 계산된 검정통계량의 값이 그 검정통계량의 임계치보다 크면 H_0 를 기각
- ◆ 검정통계량의 수준에 부응하는 p 값이 연구자가 미리 생각해둔 유의수준 (예: $\alpha = .05$)보다 더 작으면(예: $p=0.01$) 영가설(H_0)을 기각

제3절 추리통계분석은 어떻게 실행하는가?

(6) 가설검정 결론

- ◆ 연구자가 미리 결정한 통계적 유의수준(예: $\alpha = .05$ 혹은 $\alpha < .05$) 혹은 이 수준에 부응하는 임계치를 기준으로, 통계분석을 통해 계산된 검정통계량이나 검정통계량의 p 값(예: $p = .015$)을 확인하여 H_0 을 기각("검정통계량의 p 값 \leq 유의수준" 혹은 "검정통계량 값 \geq 임계치")하거나 혹은 수용("검정통계량의 p 값 $>$ 유의수준" 혹은 "검정통계량 값 $<$ 임계치") 결정을 하게 됨
- ⇒ 통계가설을 기각/수용하게 된 후에는, 실질가설의 입장(즉, 언어적인 표현)으로도 가설검정결과를 설명하는 것이 바람직함

제3절 추리통계분석은 어떻게 실행하는가?

2) 추리통계분석(가설검정)을 위한 통계기법의 유형

(1) 상호관계분석

- ◆ 현상(변수)들간의 관계를 검정

(2) 인과관계분석

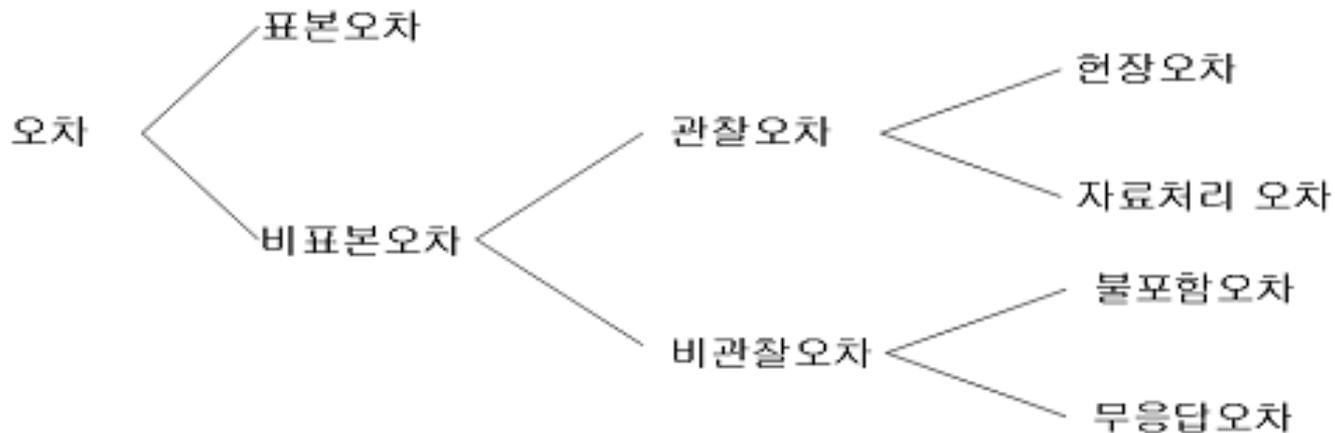
- ◆ 현상들간의 '원인-결과' 관계를 검정
- ⇒ 통계적 증거, 이론적 지지(theoretical support), 적절한 연구설계 (research design) 등 수반 필요
- ⇒ 본서에서 사용하는 인과관계분석이라는 용어의 의미는, "인과관계를 지지할 수 있는 통계적 증거를 제공하는 것"으로만 한정

제4절 추리통계분석의 기본 가정은 무엇인가?

1) 추리통계분석의 기본가정

- (1) 모수추정(parameter estimation)의 기반이 되는 표본(sample)은 당연히 모집단으로부터 추출되어야 함
- (2) 확률표본추출(probability sampling)에 따른 표본추출 필요
- (3) 추리통계에서는 표본오차 혹은 표본오류(sampling error)만을 다룸
⇒ 표본오류의 가능성이 상대적으로 아주 작을 경우에는 오히려 비표본 오차가 상대적으로 더 연구의 타당성을 낮출 가능성

[그림 8-14] 표본조사의 오차



제4절 추리통계분석의 기본 가정은 무엇인가?

2) 비표본오차

(1) 관찰오차

◆ 현장오차

- 사회현상의 관찰과정에서의 면접원과 응답자의 상호작용에서 발생하는 오차
예) 응답자가 자신이 느끼는 대로가 아니고 면접원이 원한다고 생각하는 대로 응답을 하게 되는 경우에는 조사대상 현상을 제대로 반영하지 못하는 현장오차가 발생

◆ 자료처리오차

- 조사결과를 잘못 기록하거나 기록된 자료를 잘못 처리할 때 발생하는 오차

(2) 비관찰오차

◆ 불포함오차

- 표본프레임이 불완전해서 모집단을 제대로 반영하지 못하는 경우 발생
예) 전화번호부를 표본프레임으로 사용했을 경우 전화번호에 누락된 응답자들은 표본에 추출될 수 없게 됨

◆ 무응답오차

- 응답자가 응답을 거부했을 때 발생하는 오차

표 9-1 귀무가설에 대한 판정과 오류

검정결과	실제	
	H_0 가 참	H_0 가 거짓
채택	옳은 결정 확률 = $1 - \alpha$	제2종 오류 확률 = β (β 위험)
기각	제1종 오류 확률 = α (유의수준)	옳은 결정 확률 = $1 - \beta$ (검정력)

$$\text{검정통계량} = \frac{\text{표본 통계량} - \text{귀무가설에서 설정된 모수값}}{\text{표본 통계량의 표준오차}}$$

가설검정의 목적은 귀무가설의 타당성 여부 판단에 있는데, 이 결정의 기준이 되는 표본통계량 → 검정통계량(test statistic)

예) 모평균에 대한 가설검정 → 검정통계량은 모평균에 대한 통계량인 표본평균으로부터 구할 수 있음

결정규칙

표본정보에 근거하여 귀무가설을 채택하는 경우와 대립가설이 보다 타당하다고 판단되어 귀무가설을 기각하는 경우 중 하나를 결정하는 기준

제1종 오류

귀무가설이 참일 때 귀무가설을 기각하는 오류

제2종 오류

귀무가설이 거짓일 때 귀무가설을 채택하는 오류

α 위험 혹은 유의수준

제1종 오류의 발생확률

$\alpha = P(\text{제1종 오류}) = P(H_0 \text{를 기각} \mid H_0 \text{가 참})$

β 위험

제2종 오류의 발생확률

$$\beta = P(\text{제2종 오류}) = P(H_0 \text{를 채택} \mid H_0 \text{가 거짓})$$

검정력

거짓 귀무가설을 기각할 확률($1 - \beta$)

검정통계량

귀무가설의 타당성 여부를 결정하는 기준이 되는 표본 통계량

임계값 CV(critical value)

주어진 유의수준에서 귀무가설을 채택하거나 기각하는 의사결정을 할 때 검정 통계량과의 비교기준이 되는 값