2. 가설검정

① "사회현상, 자연현상 또는 교육현상 등"에 대하여 연구자들은 "어떤 현상은 이럴 것이다"라는 잠정적 진술을 하거나 그 잠정적 진술에 대한 옳고 그름을 판단하는 의사결정을 내려야 한다.
② 조사자의 판단이 항상 맞을까? 때로는 틀릴 수도 있다. 즉 판단의 오류가 있을 수 있다.

[예] A 제약회사에서 항생제를 병에 담은 제품을 생산. 항생제를 최소 기준 용량 이상으로 채우는 것에 관심 & 표본을 추출
㉠ 평균 용량 a가 최소 허용량을 넘지 않는다.
㉡ 평균 용량 a가 최소 허용량을 넘는다.

③ 의사결정의 문제를 통계적 '가설검정 (statistical test of hypothesis)'이라 한다.

2-1. 귀무가설과 대립가설

	귀무가설(H_0, 귀무가설)	대립가설(H_1, 연구가설)
정의	㉠ 새로운 주장이 타당한 것으로 볼 수 없을 때 저절로 원상이나 현재 믿어지는 가설로 돌아가게 된다고 주장하는 가설 ㉡ 조사자가 지지하고 싶지 않은 가설 ㉢ 조사자가자 거짓이라는 것을 보임	㉠ 귀무가설이 부정되었을 때 진리로 남는 잠정적 진술 ㉡ 새롭게 주장하고자 하는 가설 ㉢ 조사자가 지지하고 싶어 하는 가설 ㉣ 조사자가 참이라는 것을 보임
표현	" = ", "~ 같다", " ~ 차이가 없다", "유의하지 않다", "독립이다", "연관성이 없다"	" ≠ ", " > ", " < ", " ~ 같지 않다", " ~ 차이가 있다", "모두 같은 것은 아니다", "유의하다", "독립이 아니다", "연관성이 있다"

* 통계학자들은 항상 귀무가설이 참이라고 가정하고 표본을 이용하여 얻은 정보들이 귀무가설보다 대립가설에 더 유리한 증거인지를 결정하게 된다.

[예] 성별(1:남자, 2:여자)에 따라 결혼만족도(연속형 자료)에 차이가 있는지를 검정하고자 한다.
㉠ H_0 : $\mu_1 = \mu_2$　　㉡ H_1 : $\mu_1 \neq \mu_2$

[예] 토론식 수업 전과 후의 국어점수의 평균에 차이가 있는지를 검정하고자 한다.
㉠ H_0 : 토론식 수업 전과 후의 국어점수의 평균에 차이가 없다($\mu_1 - \mu_2 = 0$)
㉡ H_1 : 토론식 수업 전과 후의 국어점수의 평균에 차이가 있다($\mu_1 - \mu_2 \neq 0$)

[예] 나트륨 섭취량이 3300mg을 넘는지 검정하고자 한다.
㉠ H_0 : $\mu = 3300$　　㉡ H_1 : $\mu > 3300$

2-2. 가설검정

(예제) 일반적으로 중학교 교사들의 주당 수업시수는 14시간이고 표준편차는 2시간이라 한다. A 지역의 중학교 교사들 중 임의로 100명을 조사한 결과, 평균 15시간 수업이 진행되는 것으로 나왔을 때, 중학교 교사들의 주당 수업시수가 14시간인지 유의수준 $\alpha = 0.05$에서 검정하여라.

[풀이] 만약 귀무가설이 참이라면 표본평균은 모평균과 많은 차이가 없을 것이다.

㉠ 1단계 : 귀무가설과 대립가설을 설정한다.

$H_0 : \mu = 14$ $\qquad$ $H_1 : \mu \neq 14$

㉡ 2단계 : 귀무가설 하에서 표집분포를 그리고 $\overline{X}$의 위치를 확인한다.

즉, 중심극한정리에 의해 표집분포의 평균은 14, 표준오차는

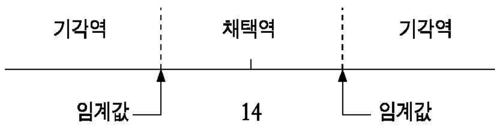
$$SE = \frac{s}{\sqrt{n}} = \frac{2}{10} = 0.2$$

이며, 100명의 평균 15에 대한 검정통계량의 값이

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{15 - 14}{0.2} = 5$$

이므로 모평균으로부터 5배의 표준편차 거리에 떨어져 있다.

㉢ 3단계 : 귀무가설을 기각할지, 채택할지를 정하기 위하여 유의수준 α에 맞는 임계값(critical value)을 정하여 기각역(rejection region)과 채택역(acceptance region)을 구분한다.



이 때, 양측검정이냐 단측검정이냐에 따라 임계값을 정한다.

즉, 유의수준 $\alpha = 0.05$인 양측검정이므로, 표준정규분포의 오른쪽 꼬리와 왼쪽 꼬리의 영역이 $\alpha/2 = 0.025$가 되는 점을 임계값으로 잡아 기각역과 채택역을 나눈다. 이 값들은 $Z = \pm 1.96$이고, 만약 $Z > 1.96$ 또는 $Z < -1.96$이면 귀무가설을 기각할 것이다.

㉣ 4단계 : 검정통계량 값이 기각역에 속하는지 아닌지 확인한다. 만약 검정통계량 값이 기각역에 속한다면 귀무가설을 기각하여 대립가설을 주장하며, 채택역에 속한다면 귀무가설을 채택하고 대립가설을 포기하거나 다른 검정법을 선택하여 다시 검정한다.

즉, 검정통계량 값이 5로 기각역에 속하므로, $H_0 : \mu = 14$를 기각하고 주당 교사의 수업시수가 14시간이 아니라는 결론을 내린다.

① 위의 과정에서 검정통계값, 유의수준, 기각역, 채택역, 임계값을 알아야 한다.

㉠ 검정통계값 : 표본으로부터 계산된 값

㉡ 유의수준 : 귀무가설이 옳은데도 불구하고 틀린 것으로 치고 기각하는 확률의 허용 수준

㉢ 채택역 : 귀무가설을 채택하는 영역

㉣ 기각역 : 귀무가설을 기각하고 대립가설을 채택하는 영역

㉤ 임계값 : 채택역과 기각역을 나누는 기준이 되는 값

② 양측 검정과 단측 검정

㉠ 양측 가설검정(two-tailed test of hypothesis)

- 귀무가설을 '...이 같다', 대립가설을 '...같지 않다'고 설정하여 검정하는 방법이다.
- 기각역은 양쪽으로 존재한다.
- 유의수준이 양쪽으로 분할되어 $\frac{\alpha}{2}$에 해당하는 Z값이 기각값이 된다.

검정방법 \ α	.10	.05	.01
양측 검정	± 1.645	± 1.96	± 2.58

[예] 어느 학교 교사들의 주 당 수업시수의 평균이 전국의 주당 수업시수인 14시간과 다른지를 보여주려 한다.

- $H_0 : \mu = 14$
- $H_1 : \mu \neq 14$

[설명] 귀무가설을 기각한다는 것(대립가설이 참이라는 것)은 이 학교 교사들의 평균 주당 수업시수는 14시간이 아니라는 것이다. 이는 평균의 값이 14시간보다 크거나 작다라는 것을 의미한다.

㉡ 단측 가설검정(one-tailed test of hypothesis)

- 단일표본검정시 귀무가설이 모집단의 평균 모수치가 특정수보다 작거나 같다 혹은 크거나 같다를 또는 두 독립표본 검정시 어느 모집단의 평균 모수치가 다른 모집단의 평균 모수치보다 크거나 같다 혹은 작거나 같다를 검정하는 방법이다.
- 기각역이 한쪽에 존재한다.
- 유의수준이 한쪽으로만 고려되므로 α에 해당되는 Z값이 기각값이 된다.

검정방법 \ α	.10	.05	.01
단측 검정	$+1.28$ 혹은 -1.28	$+1.645$ 혹은 -1.645	$+2.33$ 혹은 -2.33

- 단측 검정에서 귀무가설에 등호가 꼭 포함된다.

[예] 현재 A장치로 생산된 제품 중 3%가 불량품이다. 이 장치의 일부를 간단히 교정함으로써 불량률 p를 낮출 수 있는지를 알고 싶다.

- $H_0 : p = 0.03$
- $H_1 : p < 0.03$

[설명] 만약 귀무가설을 기각할 수 있다면(대립가설이 참이라면), 간단한 교정을 통해 불량률을 낮출 수 있다는 결론을 내릴 수 있다. 이는 비율의 값이 0.03보다 작다라는 것을 의미한다.

㉢ 유의수준에 따른 양측 검정과 단측 검정의 기각값을 비교할 때 단측 검정에 의한 기각값의 절댓값이 양측 검정에 의한 기각값의 절댓값보다 작다. 이는 양측 검정보다 단측 검정을 취할 때 귀무가설을 쉽게 기각할 수 있음을 의미한다.

- Z통계값이 $+1.80$일 때, 유의수준을 .05로 하면 양측 검정에서 기각값이 ± 1.96으로 귀무가설이 기각되지 않으나 단측 검정에서는 1.645로 귀무가설이 기각된다.
- 단측 검정이 귀무가설을 쉽게 기각하므로 '검정력이 강하다(powerful)'라고 할 수 있다.

2-3. 제1종 오류, 제2종 오류, 검정력, 유의수준

		진리(True)	
		H_0	H_1
의사결정 (decision making)	H_0	$1-\alpha$	제2종 오류 β
	H_1	제1종 오류 α (유의수준)	$1-\beta$ (검정력)

【참고】 귀무가설(H_0)이 진인데 그 귀무가설을 기각하는 오류와 대립가설(H_1)이 진인데 귀무가설을 채택하는 오류 중에 전자(귀무가설이 진인데 그 귀무가설을 기각하는 오류)가 보다 심각한 오판이다.

2-3-1. 제1종 오류, 제2종 오류

① 제1종 오류(type 1 error)

㉠ 귀무가설이 진인데 그 귀무가설을 기각하는, 즉 대립가설을 채택하는 판단의 오류이다.

㉡ α로 표기, 유의수준(significant level)이라 한다.

㉢ 유의수준이란 심각한 오판을 내릴 확률 즉, 제1종 오류의 수준을 의미한다.

㉣ 제1종 오류를 극소화하는 연구가 바람직한 연구이다.

② 제2종 오류(type 2 error)

㉠ 대립가설이 진인데 귀무가설을 기각하지 않고 채택하는 (귀무가설이 진이 아닌데 귀무가설을 채택하는 오판) 판단의 오류이다.

㉡ β로 표기한다.

2-3-2. 유의수준(significant level)

① 유의수준이란?

㉠ 심각한 오판을 내릴 확률, 심각한 오판을 허용하는 수준 즉, 제1종 오류의 수준을 의미한다.

㉡ α로 표기한다.

㉢ H_0이 진일 때 H_0가 진이 아니라고 오판할 최대 허용 한계인 실수의 확률이다.

㉣ 일반적으로 경험사회과학에서 유의수준은 0.05 또는 0.01로 설정하고 있다.

㉤ 연구자가 이론적 배경이 강하면 유의수준을 낮출 수 있다.

㉥ 유의수준의 설정은 연구가 시작되기 전, 연구자에 의하여 결정되어야 하는데 의사결정의 기준이 되기 때문이다.

㉦ 유의수준을 %로 나타내는 경우는 잘못된 것이다.