

### \*\*\* 추론통계 \*\*\*

통계학이란 논리적 사고와 객관적 사실을 바탕으로 일반적인 확률적 결정론에 의해서 인과관계를 규명하는 학문이다. 특히 연구목적에 의해 설정된 가설들에 대해서 분석결과가 어떤 결과를 뒷받침하고 있는지를 통계적 방법으로 검정할 수 있다.

기술통계가 수집된 자료의 특성을 쉽게 파악하기 위해 자료정리 및 요약하는 통계학 분야라고 한다면, 집단간 차이분석은 모집단에서 추출한 표본정보를 통해 모집단의 다양한 특성을 추론하는 추론통계 분야다.

- **추정** : 모집단에서 추출한 표본에서 얻은 정보를 이용하여 모집단의 특성을 나타내는 값을 확률적으로 추정한다. 점추정, 구간추정
- **검정** : 유의수준과 표본의 검정 통계량을 비교하여 통계가설의 진위를 입증한다.

\* **통계적 추론** : 어느 집단(모집단)의 특징을 알고자 자료조사를 진행하려 할 때, 집단의 규모나 성격에 따라 차이는 있으나 대개의 경우에 모집단 전체를 대상으로 자료조사를 하기란 쉽지 않다. 그래서 집단의 특징을 손쉽게 파악하기 위해 모집단의 일부자료만 표본으로 뽑아서 추정조사를 하게 된다.

추정은 점추정과 구간추정으로 나뉜다. 예를 들어 키 175는 유동적이다. 구간 추정 시 신뢰구간 설정이 중요한데 아무리 신뢰구간이라도 모수가 신뢰구간 안에 포함되지 않을 확률은 항상 존재하게 되는데 보통 이러한 확률을  $\alpha$ (알파)라고 한다. 이 때 신뢰구간은  $xy$  축에서  $y$ 축을 기준으로 양쪽을 다루게 되므로  $\alpha / 2$ 가 된다. 모수가 구간에 포함되지 않을 확률이  $\alpha$ 이므로, 모수가 구간에 포함될 확률은  $1 - \alpha$ 가 된다. 이를 신뢰구간 이라고 하며, 일반적으로 90%, 95%, 99%의 확률을 많이 사용한다.

\* **통계적 가설** : 세상에는 관념(정설)이 있으나 그 자체가 늘 정답은 아니며 대개의 사람들이 인정하는 생각이다. 하지만 이러한 생각은 영원하지 않으며 불완전을 포함하고 있는데, 시간이 지남에 따라 이러한 불완전을 해결하기 위한 새로운 생각, 즉 가설이 생겨나게 된다. 발생한 가설은 세상에 인정받지 못하고 사라질 수도 있고, 기존의 정설을 대신해서 새로운 정설로 인정될 수도 있다. 이는 다시 새로운 가설에 의해 사라질 가능성도 있다.

예를 들어 지구를 보더라도 "지구는 네모 모양이다 --> 천동설 --> 지동설 --> ?(새로운 가설)로 변화해 왔으며 새로운 가설이 만들어 질수도 있다.

\* **가설 검정** : 일반화된 사실은 쉽게 바뀌지 않는다. 그래서 어떠한 사실을 주장하려면 일반화된 사실이 잘못되었다는 것을 증명해야 한다. 이것이 가설검정을 하는 이유다. 모수에 대한 주장을 가설로 정립하고 가설이 맞는지를 자료를 통해 판단하게 된다. 이 때는 물론 확률을 이용한다.

\* **귀무 / 대립가설** : 추론통계에서는 현재의 정설을 "귀무가설( $H_0$ )"이라 하고, 새로운 가설을 "대립가설( $H_1$ )"이라고 하며, 이 두 개의 가설 중 어느 것이 더 타당한지를 판단하게 된다.

추론통계에서 "검정 통계량은 표본 통계량을 통해 모수는 '이러 이러할 것이다'라는 가설을 추정하는 것으로, 언제나 틀릴확률이 있다."라는 전제 조건을 갖는다. 어떠한 자료에 대해 일반적으로 인정되는 지금까지의 정설에 대한 새로운 의견, 즉 아직은 일반적이지 못한 가설을 수용 또는 기각하는 과정을 거치게 된다.

\* **가설의 종류**

귀무가설 (영가설, $H_0$ )	현재의 정설(가설)로 관습적, 보수적 주장. 두 통계치 간에 차이가 없다는 가설 새 가설에 의해 대체될 것을 예상하는 가설
대립가설 (연구가설, $H_1$ )	귀무가설에 대립되는 설로, 새롭게 검정하고자 하는 주장이다. 적극적으로 입증하고자 하는 주장이다.

\* 가설검정 시에 이 두개의 가설은 정반대로 설정되어야 한다.

예) 서로 같다 - 서로 같지 않다

집단1의 분산이 더 크다 - 집단1의 분산이 더 작거나 같다.

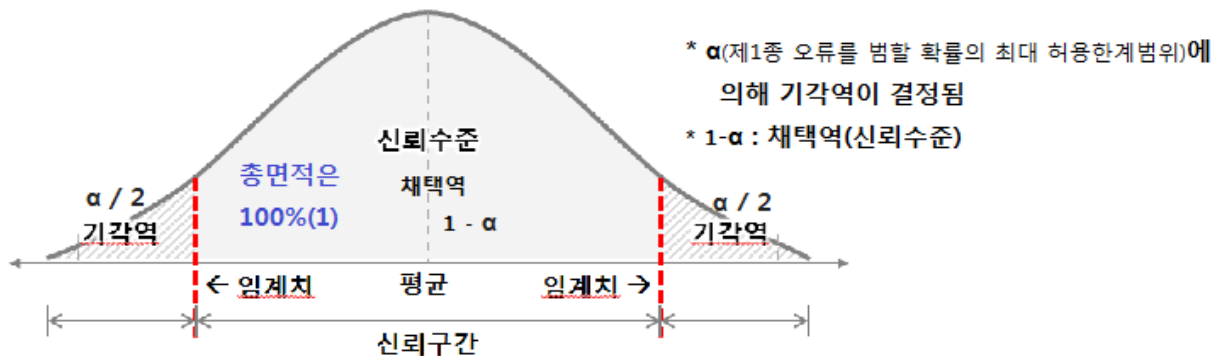
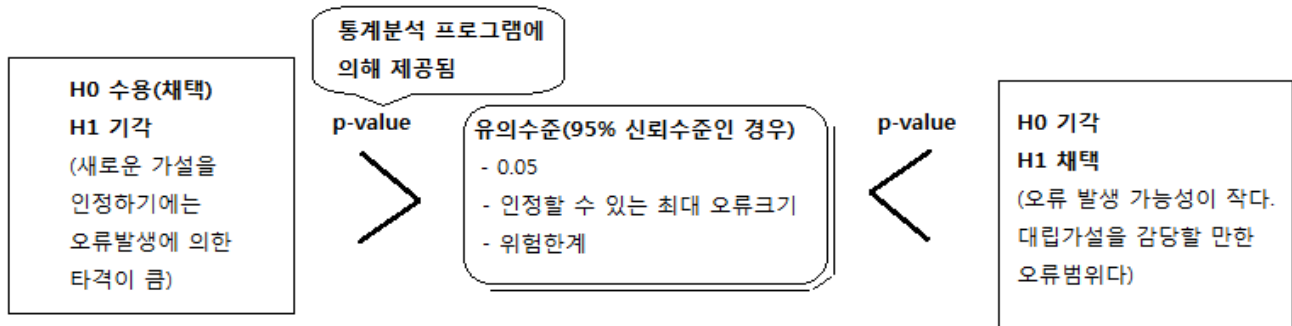
집단1의 비율이 더 크다 - 집단1의 비율이 더 작거나 같다.

- 논문이나 보고서에서는 귀무가설을 기각하고, 대립가설을 채택하는 것이 목적이다.

예) 이상 - 이하,      같지 않다 - 초과, 미만

\* **유의수준** : 표본을 가지고 오류를 판단할 때 오류를 범해도 감당 가능한 수준을 유의수준(0.05, 0.1, 0.01)이라 한다.

**제1종 오류발생확률(p-value)** : 귀무가설이 맞는데 틀렸다고 하는 경우의 오류(이런 확률결과가 나올 확률값)



**귀무 대립 가설 예)** -----

- 귀무가설 : 한우 1인분은 150g이다.
- 대립가설 : 한우 1인분은 150g이 아니다.

이러한 의견에 대해 새로운 가설을 택할 것인가? 유의수준 0.05( $\alpha$ 값)라고 했을 때 분석방법을 통해 얻은 p-value(검정통계량, 유의확률, 정설이 틀릴 확률)가 0.07 이 나온 경우를 생각해 보자. 가설 검정을 통해 p값이 도출되었을 때

- 1) **p-value > 0.05** : 귀무가설 채택
  - 오류를 인정하기에는 위험확률이 너무 커서 기존의 생각을 그대로 갖기로 한다.
- 2) **p-value < 0.05** : 대립가설 채택
  - 감당할 만한 오류이니 '한우 1인분은 150g이 아니다'라는 의견을 인정하기로 한다. 기존 생각을 버리고 새로운 의견을 받아 들인다.

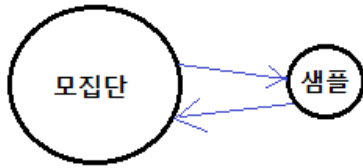
가설예2) 새로운 포장방법으로 생산된 과자들의 평균 보관시간이 기존의 포장방법 보다 효율적인가?

가설 예3) 새로운 치료방법이 제공된 환자들의 치유율이 높았는가?

가설 예4) A상표 비료에 비해 새로운 B상품의 비료가 식물의 평균 수확량을 증가시키는가?

## P-value 란?

“샘플 데이터는 극히 우연한 기회에 얻어진 것이다. 예를 들어, 2025년 7월 7일 12시 12분 30초에 강남역 근처에서 추출된 데이터는 매우 드문 확률( $1/\infty$ )로 수집된 샘플 데이터이다. 이 데이터를 바탕으로 검정 통계량을 산출하고, 그 결과를 모집단에 대한 가설 검정에 활용한다.”



일반적인 산술 규칙에서 해석할 수 없는 식이긴 하나

$$\frac{\infty}{\infty} - \frac{1}{\infty} = \frac{\infty}{\infty} \text{ 으로 엄청난 경우의 수가 나올 수 있다.}$$

이 검정 통계량으로 모집단에 적용하려고 하는 것이다. 이때  $\frac{\infty}{\infty}$ 에서 겨우  $\frac{1}{\infty}$ 의 데이터로 얻은 값으로 모집단을 추론한다는 것은 대단히 황당한 상황이라고 할 수 있다. 그러니까 이 검정 통계량을 바탕으로 모집단에 대한 결론을 내리려는 것이다. 그런데 겨우 하나의 시점에서 얻은 극히 일부의 데이터를 가지고 전체 모집단을 추론한다는 것은, 어찌 보면 매우 무모한 일일 수 있다.

그래서 통계학에서는 ‘유의확률(p-value)’이라는 개념을 도입한다. 우리가 흔히 말하는 “p-value가 0.05보다 작다”는 말은, 얻어진 결과가 95% 이상의 신뢰 수준을 가진다고 해석된다. 여기서 ‘신뢰할 수 있는 값’이라는 것은, 동일한 실험이나 샘플링을 무수히 반복했을 때, 해당 검정 통계량이 일정한 구간(신뢰구간, confidence interval) 안에 포함될 확률이 95% 이상이라는 의미다.

예를 들어 카이제곱 검정을 했더니 검정 통계량(카이제곱 값)이 0.3이 나왔다고 하자.

이 값에 해당하는 p값이 0.05보다 작게, 예를 들어 0.01이라고 가정한다면, 이는 귀무가설이 참일 때 이렇게 극단적인 결과가 우연히 나올 확률이 1%밖에 되지 않는다는 뜻이다. 즉, p값이 0.05보다 작다는 것은 우리가 설정한 유의수준(5%)보다 낮은 확률로 이 결과가 나타났다는 의미이며, 이를 바탕으로 우리는 귀무가설을 기각하고, 대립가설을 지지할 만한 통계적 근거가 있다고 해석할 수 있다.

여기서 중요한 것은 카이제곱 값 자체가 믿을 만하다는 것이 아니라, 해당 통계량이 유의수준보다 작은 p값을 만들어냈기 때문에 통계적으로 유의하다는 것이다.

또한, p값은 “앞으로 다른 샘플을 뽑아도 같은 결과가 나올 가능성”을 의미하는 것이 아니라, 현재 샘플에서 관측된 결과가 귀무가설 하에서 우연히 발생할 수 있는 수준인가를 판단하는 기준이다.

결론적으로, p값이 충분히 작다면, “이번 분석 결과는 귀무가설과는 통계적으로 의미 있는 차이가 있으며, 현재 데이터는 모집단의 특성을 어느 정도 잘 반영하고 있을 가능성이 있다”고 해석할 수 있다.

-----이를 다른 관점으로 말하면,

\* p-value > 0.05 : 관측된 결과가 우연히 발생했을 가능성이 5%보다 크다 → 즉, 통계적으로 유의하지 않다 (우연일 수 있다)

\* p-value < 0.05 : 관측된 결과가 우연히 발생했을 가능성이 5%보다 작다 → 즉, 통계적으로 유의하다 (우연이라 보기 어렵다) 라고 말할 수 있다. --