

13장 가설검정 (2부)

1. 가설검정 예제 (두번째)

또 하나의 가설검정을 생각해 본다.

예제: 코코아 박은 대상으로 이번에는 표본의 크기를 100으로
정확. 결과는 아래와 같음.

치료여부	예	아니오
도수	90	10

이제 가설검정 6단계 과정을 통하여 치료를 90% 라는
제약리사의 주장을 검정하고자 한다.

1) 단계 I : 영가설과 대립가설 설정

영가설 (H_0) : $p = 0.9$ ($p =$ 치료율)

대립가설 (H_1) : $p < 0.9$

2) 단계 II : 검정통계 설정

영가설이 참이라고 가정할 때, 코코아 환자에게
코코아 박이 효과가 있는지 여부는 이항분포를
따른다. 이는 마치 동전을 던졌을 때 앞면이 나올
확률을 따지는 것과 동일하다.

100명을 대상으로 할 때, 치료되는 사람의 수를
 X 라 하면, X 는 아래 모양의 이항분포를 따른다:

$$X \sim B(100, 0.9)$$

2단계 이 경우 X 는 점무분포를 근사적으로 따른다.

이는 $100 * 0.9 = 90 > 5$ 이고 $100 * 0.1 = 10 > 5$ 이기 때문이다.

즉,

$$X \sim N(90, 9) \text{ 이 성립한다.}$$

$$\begin{matrix} \nearrow & \nwarrow \\ 100 * 0.9 & 100 * 0.9 * 0.1 \end{matrix}$$

3) 단계 II: 기각역 설정.

① 유의수준을 5%로 정함.

② 기각역은 아래 공식을 만족하는 C 값들의 집합임:

$$P(X < C) < 0.05 \quad \leftarrow \text{라만크스 점}$$

4) 단계 IV: p-값 확인

진정통계는 아래 모양의 정규분포이다.

$$X \sim N(90, 9).$$

22건 표본 검사의 경우 100명이 80명이 치료되었다.

따라서 80명 이하로 치료된 확률이 p-값이 된다.

즉,

$$\begin{aligned} P(X \leq 80) &= P(Z \leq \frac{80-90}{3}) \\ &= P(Z \leq -3.333) \\ &= 0.0004 \end{aligned}$$

주의: 연속성 보정을
가설검정에서만
같이 하지 않음

5) 단계 V: 기각역 프락 여부

앞서 구한 p-값은 0.0004 이며 0.05 보다 작다.

따라서 80명 이하 숫자는 기각역에 포함된다.

6) 단계 VI: 최종 결정

100명 중에 치료된 사람의 수인 80이 기각역에 포함된다. 따라서 치료율이 90%라는
평가설을 기각한다.

2. 가설검정 오류와 검정력

- 통계적 주장(가설)은 전혀 절대적이지 않으며 확률적으로만 의미를 갖는다.
- 그러한 가설을 검증한 결과 또한 절대적이지 않으며 확률적으로만 의미를 갖는다.
- 예제 : 포도야 사과인지의 여부.

① 귀무가설이 90% 라는 영가설 (H_0) 이 참인 경우

i) 검정결과 H_0 를 채택하는 경우 : 옳은 결정

ii) " " " 기각하는 경우 : 제 1종 오류 발생

② 귀무가설이 90% 라는 영가설 (H_0) 이 거짓인 경우

i) 검정결과 H_0 를 채택하는 경우 : 제 2종 오류 발생

ii) " " " 기각하는 경우 : 옳은 결정 ("검정력" 입증)

* 주의 : 독스바제 스캔들의 경우는 ② - ii) 에 해당함.

즉, 가설검정의 검정력이 입증된 경우인.

- 위 네 경우를 정리하면 아래 표와 같다.

	H_0 를 채택	H_0 를 기각
H_0 가 참	옳은 결정	제 1종 오류
H_0 가 거짓	제 2종 오류	검정력 입증

• 제 1종 오류가 발생할 확률

- 위약수준과 같은.

- 영가설이 (예를 들어) 95%의 신뢰수준으로 참이기 때문에 그렇지 않은 경우, 즉, 영가설을 기각할 확률이 나머지 5%인.

• 제 2종 오류가 발생할 확률

- 보다 복잡한 계산 과정이 요구됨.

- 여기서는 제 2종 오류에 대한 감수만을 익히면 됨.

• 검정력

- 영가설이 거짓인데 그 사실을 맞출 확률

- 1에서 (제 2종 오류 발생 확률)을 뺀 값.