

가설검정

▶ 가설

- 귀무가설(null hypothesis= H_0) : 기존에 알려진 사실. 특별한 입증 자료가 없으면 고수하려고하는 성질을 가짐
- 대립가설(alternative hypothesis= H_1) : 현재까지 알려진 사실에 대한 반박. 표본에서 확실한 근거가 제시되어야 채택이 됨

▶ 오류

- 모집단 전체를 조사하는 것이 아닌 표본을 추출하여 가설을 검정함으로써 생기는 오류
- 1종오류(type 1 error) = α : 귀무가설이 옳는데 대립가설을 채택하는 오류
- 2종오류(type 2 error) = β : 대립가설이 옳는데 귀무가설을 채택하는 오류

▶ 기각

- 기각역(rejection area) : 귀무가설을 기각하는 영역
- 임계점(critical point) : 기각역의 경계치
- 기각역은 귀무가설을 기각하고 대립가설을 채택하는 영역이기 때문에, 기각역의 부등호 방향과 대립가설의 부등호 방향은 같음
- > 표본 크기가 고정되어 있다면 제1종 오류를 범할 확률과 제2종 오류를 범할 확률은 시소.(한 확률이 커지면 다른 확률이 작아짐) => 두 오류 모두를 줄이고 싶다면 표본의 크기를 늘려야함
- > 기존의 입장을 중시하는 보수적인 관점에서 1종오류만을 기준

▶ 유의수준(α)

- 제1종 오류를 범할 확률의 최대 허용값
- 제1종 오류가 일어날 확률이 낮으면 낮을수록 좋겠지만, 너무 낮추려고 하다가는 귀무가설을 기각하지 못하는 보수적인 판단을 내리게 됨. -> 적절한 유의수준 필요

▶ 검정

- 양측검정 : 대립가설의 부등호가 양쪽 방향으로 표시 $H_0 : \theta = \theta_0$
 $H_1 : \theta \neq \theta_0 \Rightarrow \theta < \theta_0 \text{ or } \theta > \theta_0$
- 단측검정 : 대립가설의 부등호가 한쪽 방향으로 표시 $H_0 : \theta = \theta_0$
 $H_1 : \theta < \theta_0$

▶ 검정력함수, 검사특성곡선

- 검정력함수 : 기각역이 정해져 있을 때, 모평균(μ)의 변화에 따라 기각하는 확률의 함수($\gamma(\mu)$)
- 검사특성곡선 : 기각역이 정해져 있을 때, 모평균(μ)의 변화에 따라 기각하지 않는 확률의 함수(검정력함수의 여사상)

▶ 임계값

- 좌측 단측검정 : $p(X \leq x | H_0) = \alpha$ 을 만족하는 x (*등호가 있음에 주의)

- 양측 단측검정 : $p(X \leq x_1 | H_0) = \alpha/2$, $p(X \geq x_2 | H_0) = \alpha/2$ 을 만족하는 x
- 우측 단측검정 : $p(X \geq x | H_0) = \alpha$ 을 만족하는 x

▶ 검정통계량과 p-value

- 검정통계량 : 가설검정에서 사용되는 표본자료의 통계량
- p-value : 검정통계량까지의 확률값(임계값 자리에 해당 검정통계량을 대입하여 구하면 됨)

▶ 가설검정을 위한 두 가지 방법

○ 검정통계량과 기각역 간의 비교

- 검정통계량과 기각역을 비교하여 검정통계량이 기각역에 속하지 않으면 귀무가설을 채택, 기각역에 속하면 귀무가설을 기각한다.

○ p-value와 유의수준 간의 비교

- 검정통계량을 이용하여 p-value를 구한 후, 이를 유의수준과 비교하여 유의수준과 같거나 크면 귀무가설을 채택, 작으면 귀무가설을 기각한다.
- 이 때, 단측검정인지 양측검정인지에 유의하여 유의수준과 비교하여야 한다.

▶ 모평균의 검정

- 표본평균이 갖는 확률분포에 따라 이를 z혹은 t에 관한 확률분포로 바꾸어서 검정

▶ 모비율의 검정

- 확률변수가 이항분포를 띄기 때문에 해당 분포를 이용하여 검정. 만약 sample의 크기가 어느 정도 크면 중심극한정리를 이용하여 검정가능

▶ 모비율 차이 검정

- 모비율 p_1, p_2 가 있고, 표본비율 \hat{p}_1, \hat{p}_2 이 있을 때, 두 모비율을 알지는 못하지만 같다.
- > 합동추정량으로 \hat{p} 를 추정(모분산이 같을 때 합동추정량을 사용한 것처럼)
- 합동추정량 : $\frac{x_1, x_2}{n_1 + n_2}$

▶ 모분산의 검정

- 모집단이 정규분포를 따를 때, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ 을 이용

▶ 모분산 차이 검정

- 두 모집단이 정규분포를 따를 때, $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$ 을 이용