금융인을 위한

통계와 데이터 분석 입문

통계적 유의성 이해하기

한국금융연수원
KOREA BANKING INSTITUTE

학습 내용

1. 유의확률(p-value)
2. 유의확률(p-value)과 기각역 활용한 통계적 가설검정 비교
3. 파이썬을 활용한 모평균에 대한 가설검정
   - 쌍체표본(paired Sample)
   - 두 개의 독립인 표본

유의확률(p-value)

- 귀무가설 하에서 관찰된 통계량만큼의 극단적인 값을 관찰할 확률 : 꼬리부분의 확률
  - 확률을 구하는 방향이 대립가설에 따라 달라짐
- 한 개의 표본에서의 모평균에 대한 가설 검정에서

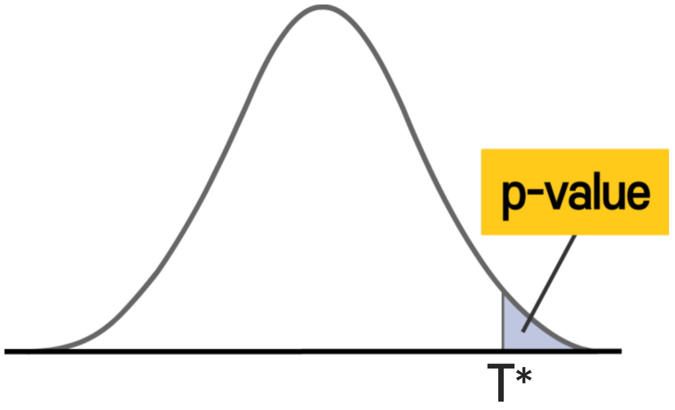
  검정통계량 : $T = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$

유의확률(p-value)

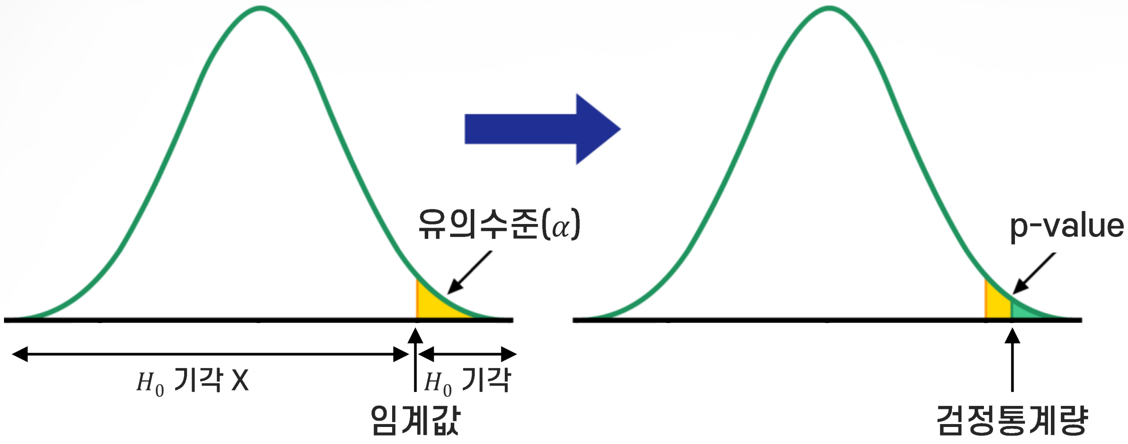
모평균에 대한 가설 1

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

- 이 단측검정에서 p-value 값은 검정통계량(T*)보다 **큰** 값을 가질 확률임

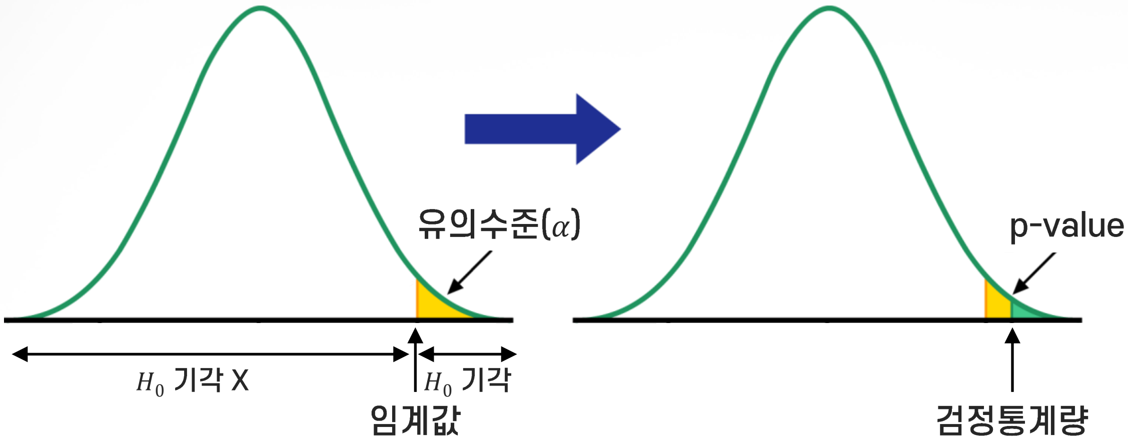


유의확률(p-value)



- p-value vs. 임계값(기각역)
  - 검정통계량 > 임계값(기각역에 속함) → p-value < α
  - 검정통계량 < 임계값(기각역에 속하지 않음) → p-value > α

유의확률(p-value)



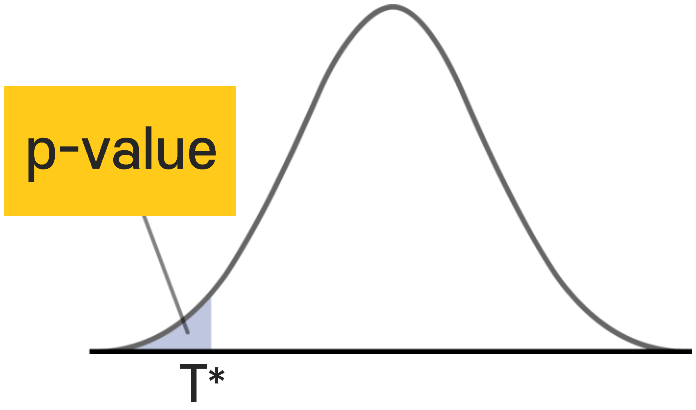
- 귀무가설 기각 여부 결정
  - p-value < α 이면, H_0를 기각하고 H_1을 채택
    : 모평균은 μ_0보다 통계적으로 유의하게 큼
  - p-value > α 이면, H_0를 기각하지 못함
    : 모평균은 μ_0보다 통계적으로 유의하게 크다고 할 수 없음

유의확률(p-value)

모평균에 대한 가설 2

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

- 이 단측검정에서 p-value 값은 검정통계량(T*)보다 작은 값을 가질 확률임



유의확률(p-value)

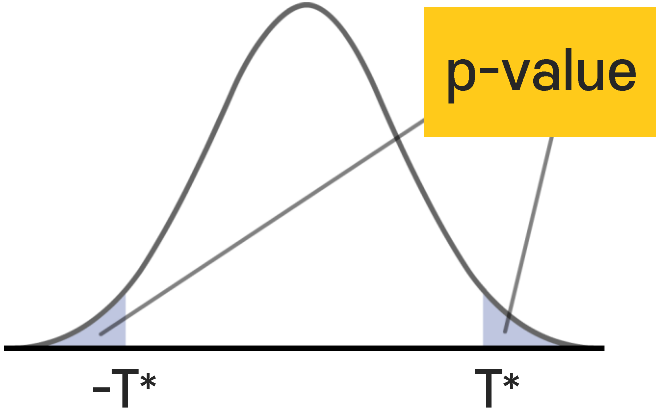
- 귀무가설 기각 여부 결정
  - p-value < α 이면, H_0를 기각하고 H_1을 채택
    : 모평균은 μ_0보다 통계적으로 유의하게 작음
  - p-value > α 이면, H_0를 기각하지 못함
    : 모평균은 μ_0보다 통계적으로 유의하게 작다고 할 수 없음

유의확률(p-value)

모평균에 대한 가설 3

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

- 양측 검정에서 p-value 값은 검정통계량(T*)이 양수이든 음수이든 한쪽 끝에서 계산된 확률의 2배임



유의확률(p-value)

- 귀무가설 기각 여부 결정
  - p-value < α 이면, H_0를 기각하고 H_1을 채택
    : 모평균은 μ_0와 통계적으로 유의하게 차이가 있음
  - p-value > α 이면, H_0를 기각하지 못함
    : 모평균은 μ_0와 통계적으로 유의하게 차이가 난다고 할 수 없음

유의확률(p-value)의 오용과 남용

- p-value의 오용과 남용 피해야 함
  - p-value가 나타내는 것은 **데이터와 귀무가설이 어느정도 맞는지 아닌지를 나타낼 뿐임**
  - 실제 효과 크기를 나타내는 것도 아니고 결과의 중요성을 나타내는 것도 아님
  - 높은 p-value를 귀무가설이 옳다는 증거로 사용하면 안됨 : **guilty vs. not guilty** 이지 innocent를 보이는 것이 아님
  - 미국통계학회에서 p-value에 대한 성명서 발표함 (The ASA's Statement on p-values: Context, Process, and Purpose)

학습정리

- 유의확률(p-value)은 귀무가설 하에서 관찰된 통계량만큼의 **극단적인 값**을 가질 확률임

- **검정통계량과 임계값**을 비교하여 기각 여부를 결정하는 방법과 **유의확률(p-value)과 유의수준**을 비교하는 방법이 결과적으로 동일한 방법임

- 파이썬을 활용하여 모평균에 대한 가설검정인 **T-test**를 쌍체표본의 경우와 두개의 독립인 표본인 경우 수행할 수 있음