HW_1

모평균 가설검정

우리나라 중학교 **1**학년 학생의 평균키는 **159cm**로 알려져 있다**.** 서울지역의 중학교 **1**학년 학생 **50**명을 대상으로 조사한 결과 표본평균은 **160cm**、표본분산은 **6cm**으로 나타났다**.** 서울지역 중학교 **1**학년 학생의 평균키와 우리나라 중학교 **1**학년 학생의 평균키 차이가 있는지 가설검정을 진행하여라**.**

문제 설정

- 전국 중학교 1학년 평균 키: 159 cm
- 서울 지역 표본(50명) 평균 키: 160 cm
- 표본 분산: 6 cm^2 → 표본 표준편차 $s \approx 2.449$ cm

가설

- H_0(귀무가설): 서울 지역 평균 = 159 cm
- H_1(대립가설): 서울 지역 평균 ≠ 159 cm (양측)

검정 통계량

- 표준 오차 = $s / \sqrt{n} = 2.449 / \sqrt{50} \approx 0.3464$ cm
- t값 = (표본 평균 − 모집단 평균) / 표준 오차
  = $(160 - 159) / 0.3464 \approx 2.886$
- 자유도 = $n - 1 = 49$

기각 기준**(**유의수준 **5 %)**

- 두 꼬리 기준 임계값: ±2.009 (자유도 49)
- 계산된 t값 2.886은 임계값 2.009를 넘음 → 귀무가설 기각
- p-값 ≈ 0.006(두 꼬리) < 0.05 → 동일 결론

결론
 서울 지역 중학교 1학년 평균 키는 전국 평균 159 cm와 통계적으로 유의한 차이가 있다 (유의수준 5 %).

참고**: 95 %** 신뢰구간

- 160 cm ± 0.70 cm → 159.30 cm ~ 160.70 cm
- 전국 평균 159 cm는 이 구간 밖에 있으므로 위 결론과 일치한다.

모비율 가설검정

민주당 대통령 후보 득표율에 대한 문제에 대하여 정규분포 근사가 가능한지에 대해서 검토하고**,** 민주당의 대통령 후보자 득표율이 **45%**라는 주장에 대하여 유의수준 **1%**에서 검정하여라**. (**데이터는 유권자 **200**명 표본 중 **81**명이 민주당 대통령 후보자를 지지한다고 응답한 자료이다**.)**

구분	값
표본 크기 n	200명
민주당 후보 지지 응답자	81명
표본비 $\hat{p}$	81 ÷ 200 = **0.405**
검정하려는 비율 p_0	**0.45 (45 %)**
유의수준	1 %(양측)

1. 정규근사 가능 여부

- $np_0 = 200 \times 0.45 =$ **90**
- $n(1-p_0) = 200 \times 0.55 =$ **110**
 두 값 모두 5(혹은 10)보다 충분히 크므로 정규분포 근사가 적합하다.

2. 가설 설정

- **H_0 (**귀무가설**)**: 실제 득표율 p = 45 %

- **H_1 (**대립가설**)**: 실제 득표율 $p \neq 45$ % (양측)

3. 검정 통계량 계산

1. 귀무가설 하 표준오차
   $\sqrt{[p_0(1-p_0)/n]} = \sqrt{[0.45 \times 0.55/200]} \approx \mathbf{0.03518}$

2. Z 값
   $(\hat{p} - p_0)$ / 표준오차
   $= (0.405 - 0.45) \div 0.03518 \approx \mathbf{-1.28}$

4. 기각 기준과 **p-**값

- 1 % 양측 검정 임계 Z 값: ± 2.576
- $|-1.28| < 2.576$ → 귀무가설을 기각하지 못함
- 두 꼬리 p-값 ≈ 0.20 (20 %) > 1 % → 동일 결론

5. 결론

유권자 200명 표본에서 관측된 지지율 40.5 %는 “실제 득표율이 45 %”라는 주장과 통계적으로 유의한 차이가 있다고 보기 어렵다(유의수준 1 %). 즉, 현재 자료만으로는 45 %라는 주장을 반박할 근거가 없다.

모분산 가설검정

정밀기계로부터 생산되는 볼트길이의 표준편차는 **2mm**를 넘지 못하도록 규정되어 있다**.** 만일 표준편차가 **2mm**를 넘는다고 판단되면 납품받은 볼트 전체를 반납하도록 계약되어 있다**.** 볼트에 대한 정밀품질검사를 진행하기 전에 사전적으로 **10**개 볼트를 무작위로 선택하여 볼트품질에 문제가 있다고 판단되면 정말품질검사 단계로 들어가지 않고 바로 반납하려고 한다**. 10**개 볼트에 대한 측정데이터는 다음과 같다**.**

(10개 표본의 볼트길이**,** 단위 **mm)**

80.30, 74.66,79.72, 82.49, 76.61, 78.58, 80.38, 82.19, 81.49

(표본평균**: 79.525mm,** 표본분산**: 6.074mm^2)**
볼트의 표준편차 **2mm** 요건을 만족하지 못한다는 증거를 이 데이터를 통해서 확인할 수 있는지에 대한 검정을 유의수준 **5%**로 진행하여라**.** 단 볼트 길이는 정규분포를 따른다고 가정한다**.**

항목	값
표본 수 n	10개
표본 분산 s^2	6.074 mm²
표본 표준편차 s	약 2.465 mm
관리 기준(귀무가설) σ_0	2 mm → $\sigma_0^2 = 4$ mm²
유의수준	5 %(단측, “초과 여부” 확인)

1. 가설 설정

- **H_0**: $\sigma \le 2$ mm (요건을 만족)
- **H_1**: $\sigma > 2$ mm (요건 미달)

2. 검정 통계량

정규분포 가정에서 모분산 검정은 카이제곱 분포를 이용한다.

$\chi^2 = (n - 1) \times s^2 / \sigma_0^2$
$= 9 \times 6.074 / 4$
$\approx \mathbf{13.667}$

자유도(df) = $n - 1 = 9$.

3. 기각 기준

- 5 % 단측 검정의 임계 χ^2 값(df = 9): **16.92**
- 계산된 χ^2(13.667) < 16.92 → 기각역에 들지 않음

4. p-값

- $P(\chi^2 > 13.667 \mid df = 9) \approx$ **0.135** (13.5 %)
- $p > 0.05$ → 귀무가설 기각 못 함

5. 결론

유의수준 5 %에서 볼트 길이의 표준편차가 **2 mm**를 초과한다고 단정할 충분한 증거가 없다.
따라서 이 작은 사전 표본만으로 바로 반납 결정을 내리기는 어렵다. (표본 자체의 표준편차는 2 mm보다 크지만, 통계적으로 확실하지 않다.)