가설검정

■ 가설검정 핵심

1. 목적: 모집단의 모수에 대한 특정 주장이 옳은지 판정

2. 대상: 모평균, 모비율, 모분산 등 모수(parameter)

3. 절차

- 귀무가설(H_0)과 대립가설(H_1) 설정
- 표본으로부터 검정통계량 계산
- 유의수준(α)과 p값 비교
- H_0 기각 여부 결정

4. 결과:

- "귀무가설 기각" → 가설에 반대할 근거 있음
- "기각하지 못함" → 가설을 부정할 근거 부족

√ 가설검정은 표본을 근거로 모수에 대한 가설을 통계적으로 판정하는 절차입니다.

■ 검정통계량과 분포별 사용 상황

1. Z분포 (표준정규분포)

- 사용 조건
  · 모집단 분산(σ^2)을 알고 있거나
  · 표본 크기 n이 충분히 커서(보통 $n \geq 30$) 중심극한정리 적용 가능

- 대표 상황
  · 모평균 검정 (σ^2 알려진 경우)
  · 모비율 검정
  · 두 집단 평균 차이 검정 (σ^2 알려진 경우, 큰 표본)

○ 사례
- 공장 불량률 검정

  · 한 공장에서 생산된 부품 불량률이 5% 이하인지 알고 싶음.
  · 대량 표본(예: 500개)을 뽑아 불량품 비율을 계산 → Z검정으로 모비율 판정.

- 대학생 키 평균 비교

  · 전국 대학생 평균 키가 172cm라고 알려져 있고, 모집단 분산(σ^2)이 알려져 있음.
  · 표본 50명을 조사했을 때 평균이 171cm → Z검정으로 평균 차이 검정

○ 문제

전국 대학생들의 평균 키는 172cm라고 알려져 있다. 모집단 분산은 $\sigma^2 = 9$ (즉, 표준편차 $\sigma = 3$)이다. 어느 대학에서 대학생 50명을 무작위 추출하여 조사한 결과, 표본 평균은 171cm였다. 이 대학생들의 평균 키가 전국 평균과 다르다고 할 수 있는지, 유의수준 0.05에서 검정하시오.

√ 풀이
1. 가설 설정

  - H_0 (귀무가설): $\mu = 172$ (전국 평균과 같다)
  - H_1 (대립가설): $\mu \neq 172$ (전국 평균과 다르다)

2. 검정통계량 (Z값)

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

- 표본평균 $\bar{x} = 171$
- 모집단 평균 μ_0 = 172
- 모집단 표준편차 σ = 3
- 표본 크기 n = 50

$$Z = \frac{171 - 172}{3/\sqrt{50}} = \frac{-1}{3/7.071} = \frac{-1}{0.424} \approx -2.36$$

---

3. 기각역 결정

- 유의수준 α = 0.05, 양측검정
- Z 임계값 = ±1.96

---

4. 판정

- 계산된 Z = -2.36
- -2.36 < -1.96 → 기각역에 속함
- 따라서 귀무가설 기각

---

5. 결론

유의수준 5%에서 **이 대학생들의 평균 키는 전국 평균 172cm와 다르다고 할 수 있다.**

2. t분포 (Student’s t)

- 사용 조건
  · 모집단 분산(σ^2) 알려지지 않음
  · 표본 크기 n이 작거나 중간 규모

- 대표 상황
  · 단일 모평균 검정 (σ^2 모름)
  · 두 집단 평균 차이 검정 (독립표본 t-test, 대응표본 t-test)
  · 회귀분석에서 회귀계수 유의성 검정

○ 사례

- 소규모 표본 평균 검정
  · 신약 복용 환자 12명의 혈압 평균이 기존 120mmHg와 차이가 있는지 확인.
  · 모집단 분산은 모름 → 단일표본 t검정.

- 두 반 시험 성적 비교
  · 반 A(15명)와 반 B(17명)의 수학 평균 점수가 다른지 알고 싶음.
  · 분산은 모름, 표본 크기도 작음 → 독립표본 t검정.

📌 두 집단 평균 차이 검정 (독립표본 t-검정)

1. 목적

- 두 개의 독립된 집단의 평균이 **통계적으로 유의하게 다른지** 판정하는 것.
  예:
  - 반 A와 반 B의 평균 시험 점수 비교
  - 남녀 평균 키 차이 검정
  - 신약 복용군 vs 위약군 평균 혈압 비교

---

2. 가설 설정

- **귀무가설(H_0)**: 두 집단의 평균이 같다 ($\mu_1 = \mu_2$)
- **대립가설(H_1)**: 두 집단 평균이 다르다 ($\mu_1 \neq \mu_2$)

> 경우에 따라 **단측 검정**으로도 가능 ($\mu_1 > \mu_2$ 또는 $\mu_1 < \mu_2$)

---

3. 검정통계량

두 집단 평균 차이를 비교할 때 사용하는 통계량은 다음과 같습니다.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

여기서,

- $\bar{x}_1, \bar{x}_2$: 두 집단의 표본평균
- n_1, n_2 : 두 집단의 표본 크기
- s_p : 풀드(pooled) 표준편차

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- 자유도(df): $n_1 + n_2 - 2$

4. 사용 조건 (전제 가정)

1. 두 집단은 서로 **독립적** (겹치지 않는 집단)
2. 각 집단은 **정규분포**를 따른다고 가정
3. 두 집단의 분산이 같다고 가정 (**등분산성**)
   - 만약 등분산성을 만족하지 않으면 **Welch's t-test**를 사용

---

5. 절차

1. 귀무가설(H_0)과 대립가설(H_1) 설정
2. 두 집단의 평균, 분산, 표본크기 계산
3. 풀드 표준편차(s_p) 또는 Welch 분산 사용
4. 검정통계량 t값 계산
5. 자유도 df로 임계값 찾거나 p-값 계산
6. 유의수준 α와 비교 → H_0 기각 여부 판정

---

6. 해석 방법

- |t| > 임계값 → H_0 기각 → 평균 차이 있음
- |t| ≤ 임계값 → H_0 채택 → 평균 차이 있다고 볼 수 없음
- p-값 < α → H_0 기각

7. 구체적 예시

> **두 반 성적 비교**

- 반 A (n_1=15): 평균 71.1, 표준편차 2.4
- 반 B (n_2=17): 평균 80.5, 표준편차 2.0

검정통계량:

$$t = \frac{71.1 - 80.5}{2.24\sqrt{1/15 + 1/17}} \approx -11.46$$

- 자유도 df = 30
- 유의수준 0.05, 양측검정 임계값 ≈ ±2.042
- |t| = 11.46 > 2.042 → H_0 기각

👉 결론: 두 반의 평균은 **통계적으로 유의하게 다르다.**

---

8. 확장

- **대응표본 t-검정 (Paired t-test)**: 같은 집단을 전후 비교 (예: 운동 전후 체중 변화)
- **Welch's t-검정**: 두 집단 분산이 같지 않을 때 사용

② 독립표본 t검정 문제

문제

반 A(15명)과 반 B(17명)의 수학 시험 점수가 다음과 같다.

- 반 A: 72, 75, 68, 70, 74, 69, 71, 73, 76, 68, 72, 74, 70, 69, 71
- 반 B: 78, 80, 82, 79, 81, 77, 83, 84, 79, 80, 82, 78, 81, 83, 80, 79, 82

두 반의 평균 점수에 차이가 있는지 검정하라. (유의수준 α=0.05)

풀이

1. **가설 설정**
- H_0: $\mu_1 = \mu_2$ (두 반 평균 차이 없음)
- H_1: $\mu_1 \neq \mu_2$ (두 반 평균 차이 있음)

2. **표본 통계량**
- 반 A 평균 = 71.1, 표준편차 ≈ 2.40, n_1 = 15
- 반 B 평균 = 80.5, 표준편차 ≈ 2.02, n_2 = 17

3. **검정통계량**
- 풀드 분산

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \approx \frac{(14)(2.40^2) + (16)(2.02^2)}{30} \approx 5.02$$

$$s_p = \sqrt{5.02} \approx 2.24$$

- t값

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{71.1 - 80.5}{2.24 \cdot \sqrt{1/15 + 1/17}} \approx \frac{-9.4}{0.82} \approx -11.46$$

4. **기각역 확인**
- 자유도 df = n_1+n_2-2 = 30
- α=0.05, 양측검정 → 임계값 $t_{0.025,30} \approx \pm 2.042$

5. **판정**
- 계산된 t = -11.46 은 기각역에 포함됨.
- **결론: 귀무가설 기각 → 두 반 평균 점수에 유의한 차이가 있음**

3. 카이제곱 분포 (χ^2)

- 사용 조건
  · 비율/분할표, 분산 관련 검정에 주로 사용

- 대표 상황
  · 분산 검정 (모분산 = 특정 값인지?)
  · 적합도 검정 (데이터가 특정 분포를 따르는지?)
  · 독립성 검정 (두 범주형 변수 간 독립 여부, 교차표 분석)

○ 사례

- 주사위 적합도 검정

  · 주사위를 60번 던져 나온 눈의 분포가 "공정한 주사위(1/6씩)"인지 확인.
  · 기대빈도와 관찰빈도 비교 → χ^2 적합도 검정.

- 흡연과 질병의 독립성 검정

  · 흡연 여부(흡연/비흡연)와 질병 발생 여부(있음/없음)를 교차표로 분석.
  · 두 변수가 독립인지 여부 확인 → χ^2 독립성 검정.

○ 문제

어느 연구자가 200명을 대상으로 흡연 여부와 질병 발생 여부를 조사하여 다음과 같은 교차표를 얻었다.

	질병 있음	질병 없음	합계
흡연자	30	70	100
비흡연자	10	90	100
합계	40	160	200

이 자료를 바탕으로, 흡연 여부와 질병 발생 여부가 독립인지 유의수준 0.05에서 검정하시오.

√ 풀이
 1. 가설 설정

  - H_0 (귀무가설): 흡연 여부와 질병 발생 여부는 독립이다.
  - H_1 (대립가설): 흡연 여부와 질병 발생 여부는 독립이 아니다.

2. 기대도수 계산

독립이라면 각 셀의 기대도수는

$$E_{ij} = \frac{(\text{행합계}) \times (\text{열합계})}{\text{전체합계}}$$

- 흡연 & 질병 있음: $\frac{100 \times 40}{200} = 20$
- 흡연 & 질병 없음: $\frac{100 \times 160}{200} = 80$
- 비흡연 & 질병 있음: $\frac{100 \times 40}{200} = 20$
- 비흡연 & 질병 없음: $\frac{100 \times 160}{200} = 80$

따라서 기대도수표는:

	질병 있음	질병 없음
흡연자	20	80
비흡연자	20	80

---

3. χ^2 검정통계량 계산

$$\chi^2 = \sum \frac{(\text{관측도수} - \text{기대도수})^2}{\text{기대도수}}$$

- $(30-20)^2/20 = 100/20 = 5$
- $(70-80)^2/80 = 100/80 = 1.25$
- $(10-20)^2/20 = 100/20 = 5$
- $(90-80)^2/80 = 100/80 = 1.25$

$$\chi^2 = 5 + 1.25 + 5 + 1.25 = 12.5$$

4. 기각역 결정

- 자유도 df = (행−1)×(열−1) = (2−1)×(2−1) = 1
- 유의수준 $\alpha = 0.05$
- χ^2 임계값 (df=1, α=0.05) = 3.841

---

5. 판정

- 계산된 $\chi^2 = 12.5 > 3.841$
- ⇒ 귀무가설 기각

---

6. 결론

유의수준 5%에서 **흡연 여부와 질병 발생 여부는 독립이 아니다.**

즉, **흡연과 질병은 관련성이 있다**고 볼 수 있다.

✅ 요약

- H_0: 독립이다 → 기각됨
- $\chi^2 = 12.5$, df=1, $p<0.05$
- 결론: 흡연과 질병은 **독립이 아님** (연관 있음)

4. F분포

- 사용 조건
  · 두 개 이상의 분산 비교, 분산비를 검정

- 대표 상황
  · 두 집단의 분산이 같은지 검정
  · 분산분석(ANOVA) → 집단 평균 차이 검정 시 사용
  · 회귀분석 전체 유의성 검정

○ 사례

- 세 가지 학습법 효과 비교 (ANOVA)

  · 학습법 A, B, C 세 그룹 학생들의 평균 점수 차이가 있는지 비교.
  · 분산분석(ANOVA)에서 집단 간 분산과 집단 내 분산 비율을 계산 → F검정.

○ 문제

- 어느 연구자가 세 가지 학습법(A, B, C)의 효과를 비교하기 위해 학생들을 세 그룹으로 나누어 시험 점수를 측정하였다.
  · 학습법 A: 70, 72, 68, 74, 71
  · 학습법 B: 75, 78, 74, 77, 76
  · 학습법 C: 80, 82, 78, 81, 79

- 세 그룹 학생들의 평균 점수가 동일하다고 할 수 있는지, 유의수준 0.05에서 검정하시오.

√ 풀이
  1. 가설 설정

   - H_0 (귀무가설): 세 집단의 평균은 모두 같다 ($\mu A = \mu B = \mu C$)
   - H_1 (대립가설): 적어도 한 집단의 평균은 다르다

  2. 기초 통계량 계산

   - 그룹별 표본 크기: $nA = nB = nC = 5$, 총 $n = 15$
   - 그룹 평균:
    · A = (70+72+68+74+71)/5 = 71
    · B = (75+78+74+77+76)/5 = 76
    · C = (80+82+78+81+79)/5 = 80
   - 전체 평균 (Grand mean): (355+380+400)/15 = 1135/15 ≈ 75.67

3. 제곱합 계산 (Sum of Squares)

- **집단 간 제곱합 (SSB)**

$$SSB = \sum n_i(\bar{x}_i - \bar{x})^2 = 5(71 - 75.67)^2 + 5(76 - 75.67)^2 + 5(80 - 75.67)^2$$

= 5(21.78) + 5(0.11) + 5(18.78)
≈ 108.9 + 0.55 + 93.9 = **203.35**

- **집단 내 제곱합 (SSW)**

$$SSW = \sum\sum(x_{ij} - \bar{x}_i)^2$$

(각 그룹 내 분산 계산 후 합산)

- 그룹 A 분산: $(70-71)^2+(72-71)^2+(68-71)^2+(74-71)^2+(71-71)^2$ = 1+1+9+9+0=20
- 그룹 B 분산: $(75-76)^2+(78-76)^2+(74-76)^2+(77-76)^2+(76-76)^2$ = 1+4+4+1+0=10
- 그룹 C 분산: $(80-80)^2+(82-80)^2+(78-80)^2+(81-80)^2+(79-80)^2$ = 0+4+4+1+1=10
  합계 = 40

따라서 **SSW = 40**

- 전체 제곱합 (SST) = SSB + SSW = 203.35 + 40 = 243.35

---

4. 자유도와 평균제곱

- 집단 간 자유도: dfB = k−1 = 3−1 = 2
- 집단 내 자유도: dfW = n−k = 15−3 = 12
- 평균제곱(MSB, MSW):
  - MSB = SSB / dfB = 203.35 / 2 = 101.68
  - MSW = SSW / dfW = 40 / 12 = 3.33

---

5. F 검정통계량

$$F = \frac{MSB}{MSW} = \frac{101.68}{3.33} \approx 30.5$$

---

6. 기각역 결정

- 자유도 (2, 12), α=0.05
- F 임계값 ≈ 3.89

---

7. 판정

- 계산된 F = 30.5 > 3.89
- ⇒ 귀무가설 기각

📌 결론

유의수준 5%에서 **세 가지 학습법의 평균 점수는 동일하지 않다.**
즉, **적어도 한 학습법은 다른 학습법과 평균 효과가 유의하게 다르다.**

✅ 정리

- H_0: μA=μB=μC → 기각됨
- F=30.5, p<0.05
- 결론: 학습법 간 점수 차이가 있다

- 두 공정의 분산 비교

· 공정 1과 공정 2의 제품 길이 변동성이 같은지 확인.

· 두 집단의 분산비를 비교 → F분포 이용.

📌 문제

어느 공장에서 두 가지 공정(공정 1, 공정 2)을 통해 동일한 부품을 생산하였다. 두 공정의 제품 길이 변동성(분산)이 동일한지 검정하기 위해 각각의 표본을 추출한 결과는 다음과 같다.

- 공정 1: n_1=10, 표본분산 $s_1^2 = 4.0$
- 공정 2: n_2=12, 표본분산 $s_2^2 = 2.0$

유의수준 0.05에서 두 공정의 분산이 동일하다고 할 수 있는지 검정하시오.

📌 풀이

1. 가설 설정

- H_0 (귀무가설): 두 집단의 분산이 같다 ($\sigma_1^2 = \sigma_2^2$)
- H_1 (대립가설): 두 집단의 분산이 다르다 ($\sigma_1^2 \neq \sigma_2^2$)

---

2. 검정통계량

F 통계량은 **큰 분산 ÷ 작은 분산**으로 정의.

$$F = \frac{s_1^2}{s_2^2} = \frac{4.0}{2.0} = 2.0$$

---

3. 자유도

- 분자 자유도: $df_1 = n_1 - 1 = 9$
- 분모 자유도: $df_2 = n_2 - 1 = 11$

4. 기각역 결정

- 유의수준 α=0.05, 양측검정
- F 분포표 (df_1=9, df_2=11) 참고
  - 상한 임계값: $F_{0.975,9,11} \approx 3.59$
  - 하한 임계값: $F_{0.025,9,11} = \frac{1}{F_{0.975,11,9}} \approx \frac{1}{3.68} \approx 0.272$

즉, 기각역은 **F < 0.272 또는 F > 3.59**

---

5. 판정

- 계산된 F = 2.0
- 기각역(0.272 ~ 3.59) 안에 포함됨

---

6. 결론

유의수준 5%에서 **두 공정의 분산은 동일하다고 볼 수 있다.**

즉, 공정 1과 공정 2의 제품 길이 변동성에는 유의한 차이가 없다.

√ 정리표

분포	주로 쓰는 상황
Z분포	모분산 알고 있거나, 표본 크기 충분히 클 때 평균·비율 검정
t분포	모분산 모를 때 평균 검정, 작은 표본, 회귀계수 검정
카이제곱(χ^2)	분산 검정, 분할표 독립성 검정, 적합도 검정
F분포	두 분산 비교, 분산분석(ANOVA), 회귀모형 전체 검정

√ 요약표

분포	사례 1	사례 2
Z분포	공장 불량률 ≤ 5% 판정	전국 평균 키와 표본 평균 비교
t분포	신약 복용 환자 혈압 평균 vs 기준	두 반의 시험 성적 평균 차이
χ^2분포	주사위가 공정한지 적합도 검정	흡연과 질병 발생 독립성 검정
F분포	3개 학습법 점수 차이 (ANOVA)	두 공정 분산이 같은지 비교

□ 가설검정에서 쓰이는 대표적인 분포들

1. 정규분포 (Normal distribution)

- 용도: 많은 검정이 정규분포를 기반으로 파생됨.
- 사례:
  · Z검정의 근거 분포
  · 표본평균의 분포 (중심극한정리)

2. 이항분포 (Binomial distribution)

- 용도: 성공/실패 같은 이산 확률 상황에서 검정
- 사례:
  · 동전이 공정한지(앞면 확률 p=0.5 검정)
  · 제품 불량률이 5%인지 검정

3. 포아송분포 (Poisson distribution)

- 용도: 희귀 사건 발생 횟수에 대한 가설검정
- 사례:
  · 시간당 교통사고 발생 횟수가 일정한지 검정
  · 특정 구역에서 하루 평균 사건 발생 건수 비교

4. 지수분포 (Exponential distribution)

- 용도: 대기시간, 수명분석에서 검정
- 사례:
  · 부품 고장 시간 분포가 평균 λ^{-1}인지 검정
  · 대기시간 분포가 특정 모형과 일치하는지

5. t, F, χ^2 의 파생분포들

  - 스튜던트화된 범위분포(Studentized range, q분포)
    · 용도: 다중비교 (Tukey HSD 검정 등)

 - Hotelling's T^2 분포
   · 용도: 다변량 평균 검정 (여러 변수 동시 평균 비교)

6. 비모수 검정에서의 분포

 - 순위 기반 분포 (Rank-based distributions)
  · Wilcoxon 검정, Mann-Whitney U 검정 → 특정 분포 가정 없음
 - 부트스트랩 분포 (Bootstrap distribution)
  · 컴퓨터 시뮬레이션 기반으로 가설검정

○ 정리

분포	쓰임새
정규분포	많은 검정의 기반, 평균 관련 검정
이항분포	비율·성공/실패 검정
포아송분포	희귀 사건 발생 횟수 검정
지수분포	대기시간, 생존 분석 검정
t, F, χ^2 파생분포	다변량, 다중비교
비모수 분포(순위/부트스트랩)	분포 가정을 하지 않는 검정

○ 가설검정 분포 선택 흐름도

```
시작
 │
 ├─▶ 1. 자료 형태는?
 │       ├─ 이산형(성공/실패, 횟수) → [이항분포, 포아송분포]
 │       │       ├─ 성공/실패 비율 검정 → 이항분포
 │       │       └─ 희귀 사건 발생 횟수 검정 → 포아송분포
 │       │
 │       └─ 연속형(평균, 분산 등) → 다음 단계
 │
 ├─▶ 2. 모수는?
 │       ├─ 평균 → Z분포, t분포
 │       │       ├─ 모분산(σ²) 알고 있거나 n≥30 → Z검정
 │       │       └─ 모분산 모름, n 작음 → t검정
 │       │
 │       ├─ 분산 → χ²분포 (단일 분산 검정), F분포 (분산 비교/ANOVA)
 │       │
 │       └─ 회귀/상관 → t분포(계수), F분포(모형 전체)
 │
 ├─▶ 3. 변수 수는?
 │       ├─ 단일 변수 → 위 분포들 그대로 사용
 │       └─ 다변량 (여러 변수 동시에) → Hotelling T², MANOVA(F분포 변형)
 │
 ├─▶ 4. 분포 가정 가능한가?
 │       ├─ 가능 → 위의 모수적 검정
 │       └─ 불가능 → 비모수 검정 (순위분포, Mann-Whitney, Wilcoxon 등)
 │
끝
```