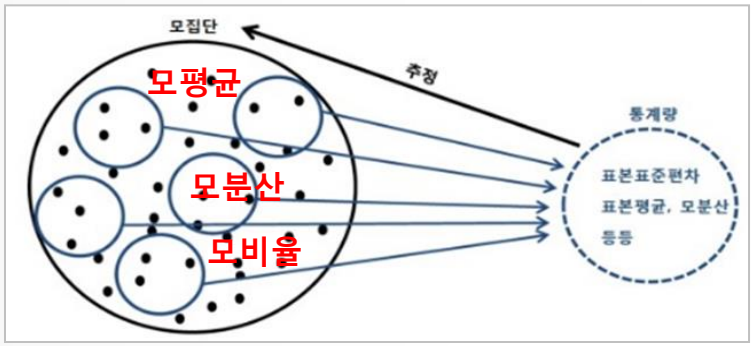
Chapter13-2.
통계분석(가설검정)

목차

6. 가설검정

- **가설검정(Hypothesis Testing)?**
  - ✓ 모집단의 모수를 검증하기 위해서 예상되는 가설(해답)을 제시하고, 표본(Sample)으로 이를 증명하는 과정
  - ✓ 표본(통계량)을 이용하여 가설이 맞는지 아니면 틀리는지 통계적 실험으로 얻은 증거로 가설을 지지하는 확률이 높으면 가설을 채택(지지)아니면 기각(부정)



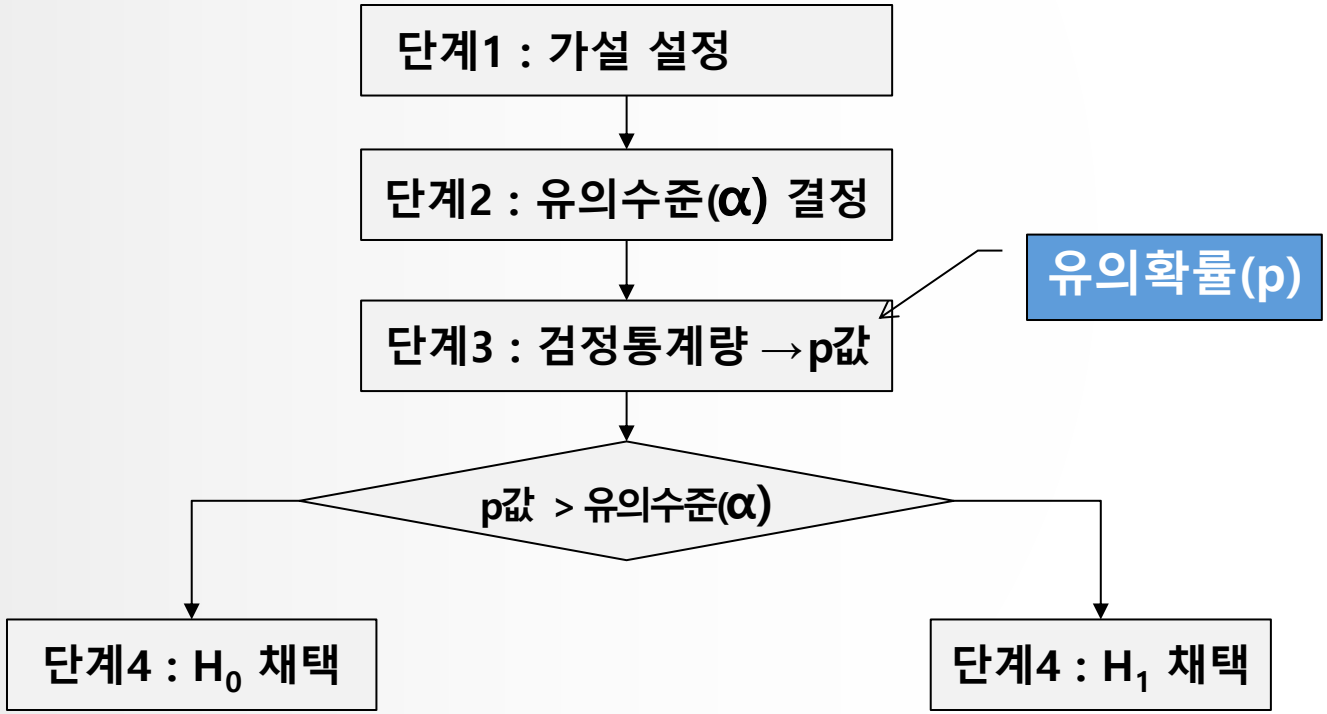
모수 vs 통계량

- **모수와 통계량 표현**

구분	모수(모집단)	통계량(표본)
의미	모집단의 특성을 나타내는 수치	표본의 특성을 나타내는 수치
표기	그리스, 로마자	영문 알파벳
평균	**μ(모평균)**	**$\bar{X}$(표본의 평균)**
표준편차	**σ(모표준편차)**	**S (표본의 표준편차)**
분산	**σ^2(모분산)**	**S^2 (표본의 분산)**
대상 수	**N(사례수)**	**n(표본수)**

가설검정 절차

<가설검정 절차>



가설 이란?

- **가설(假說, Hypothesis)?**
  - 이미 알려진 상황을 설명하기 위한 가정
    - ✓ 가정 : 어떤 명제를 사실이라고 추론하는 행위

    예) 대한민국 수도는 서울이다.
  - 어떤 문제를 검증하기 위해서 미리 세운 결론
  - 주어진 연구 문제에 대한 예측적 해답(잠정적 진술)
  - 통계분석을 통해서 채택 또는 기각(통계적 가설검정)

※ 추론 통계분석에서 가설의 설정은 매우 중요하다.

가설 유형

1. 귀무가설(영가설)

' 두 변수간의 관계가 없다.' 또는 '차이가 없다.'('효과가 없다.')

✓ 부정적 형태 진술, 사실과 같다.

예1) H_0 : 교육수준에 따라서 만족도에 차이가 없다.

예2) H_0 : 2024년도 고3 남학생의 평균키는 175cm이다.

2. 대립가설(연구가설)

' 두 변수간의 관계가 있다.', '차이가 있다.'('효과가 있다.')

✓ 긍정적 형태 진술, 사실과 다르다.

예1) H_1: 교육수준에 따라서 만족도에 차이가 있다.

예2) H_1 : 2024년도 고3 남학생의 평균키는 175cm가 아니다.

※ 항상 귀무가설을 기준으로 가설검정을 수행한다.

단계1. 가설 설정

● 가설 설정

- 어떤 사실이나 현상에 내재되어 있는 법칙이나 결과를 검증하기 위한 가설검정의 첫번째 단계

[가설 예] 2024년도 고등학교 3학년 남학생의 평균키는 175cm이다.

가설검정 예

- **통계적 가설검정 예**
  - 표본으로 얻은 정보로 귀무가설이 옳고(채택), 그른지(기각) 결정

1. 귀무가설(H_0) 2024년도 고등학교 3학년 남학생의 평균키는 175cm이다.

↓

2. [표본 & 통계량] 주요 10개 도시를 대상으로 100명씩 표본으로 선정하여 평균키를 계산한다.

↓

3. [가설검정] 가설을 지지하는 확률에 따라서 채택 or 기각

가설검정 원리

- **가설검정의 원리**

  가설A가 맞는지 아니면 틀리는지 확인하기 위하여 관찰 또는 실험을 통해서 증거를 얻고 이 증거가 가설 A가 맞다는 가정 하에서 발생할 가능성이 어느 정도 크면 (통상 5% 보다 큰 경우) 가설 A를 부정할 수 없고, 가능성이 어느 정도 작으면 부정할 수 있다.

가설 설정 방법

- **귀무가설 = 영가설**

  영가설(null hypothesis)은 "관계가 없다" 또는 "차이가 없다" 등과 같이 주로 부정적 형태로 **반증** 가능성이 있도록 진술한다. 통계적 가설검정에서 "영가설을 채택한다"라고 표현하지 않고, "**영가설을 기각할 수 없다**" 라고 말한다.

  귀무가설(영가설)은 뚜렷한 증거를 제시하지 못하면 기각할 수 없다.

  [예] 법정 형사 재판에서 피고인에 대한 영가설은 다음 중 어느 것이 합리적일까?

  H_0 : 피고인은 유죄다.(1번)

  H_0 : 피고인은 무죄다.(2번)

  정답) 2번은 가설을 반증하기 위해서 뚜렷한 증거를 제시할 수 있다.

  ❖ 무죄라고 생각하기 어려울 정도의 증거가 있으면 영가설을 기각할 수 있다.

가설 설정 방법

- **가설의 부등호 규칙**

  귀무가설(영가설)은 등호가 반드시 포함되어야 하고, 대립가설은 절대로 등호가 포함되지 않아야 한다. 또한 모집단의 모수(μ)로 표현한다.

  대립가설 H_1: $\mu < 5kg$일 때

  귀무가설 H_0: H_0: $\mu = 5kg$ 또는 $\mu \geq 5kg$로 표현(전자 표현 권장)

  대립가설 H_1: $\mu > 161cm$ 일 때

  귀무가설 H_0: H_0: $\mu = 161cm$ 또는 $\mu \leq 161cm$로 표현(전자 표현 권장)

가설 설정 방법

[사례] A회사 '단백질보충제'에 단백질 함량이 **45g**으로 광고되어 있다고 가정할 때 다음과 같은 방법으로 양측검정과 단측검정을 위한 가설을 세울 수 있다.

- 양측 검정(모평균을 모르는 경우 표본으로 모평균을 추정하기 위한 가설)
  H_0 : $\mu = 45g$ (모평균:μ = 가설에서 주장하는 평균:μ_0)
  H_1 : $\mu \neq 45g$ (방향성이 포함되지 하지 않는 가설)

- 단측검정(광고 내용의 정확성 확인을 알기 위한 가설 : 소비자 입장)
  H_0 : $\mu = 45g$ 또는 $\mu >= 45g$
  H_1 : $\mu < 45g$ (방향성이 포함된 가설)
  (회사는 45g이라고 주장하지만 고객들은 45g이 안된다고 주장하는 경우)

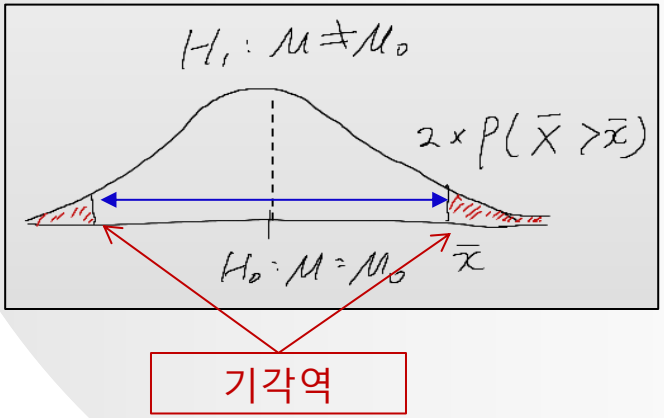
검정 방법 선택

- **양측검정과 단측검정 : 대립가설 기준으로 결정**
  - ✓ **양측검정(two-side test)** :, 대립가설(H_1)에 방향성이 없는 경우

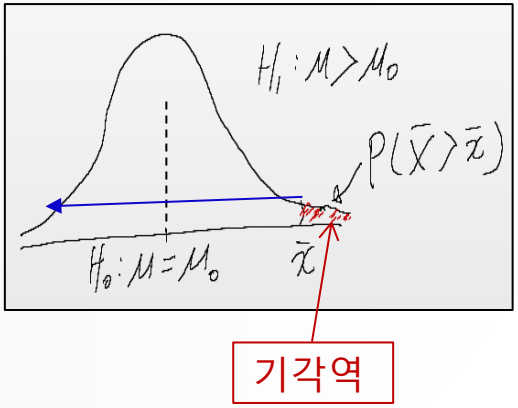
    예) H_1: $\mu \neq 50kg$ (H_0: $\mu = 50kg$)
  - ✓ **단측검정(one-side test)** : 대립가설(H_1)에 방향성이 포함되는 경우

    예) H_1: $\mu > 50kg$ 또는 H_1: $\mu < 50kg$ (H_0: $\mu = 50kg$)

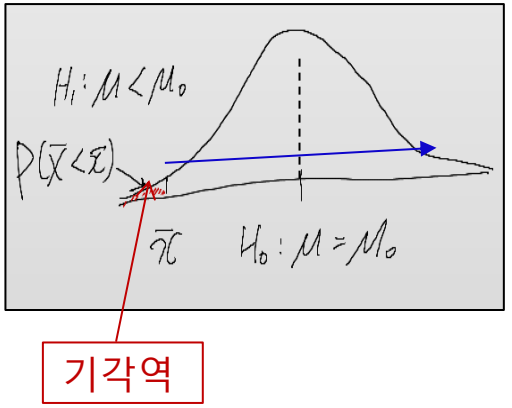
양측검정(모평균:μ, 비교평균:μ_0)



단측검정(우측)



단측검정(좌측)

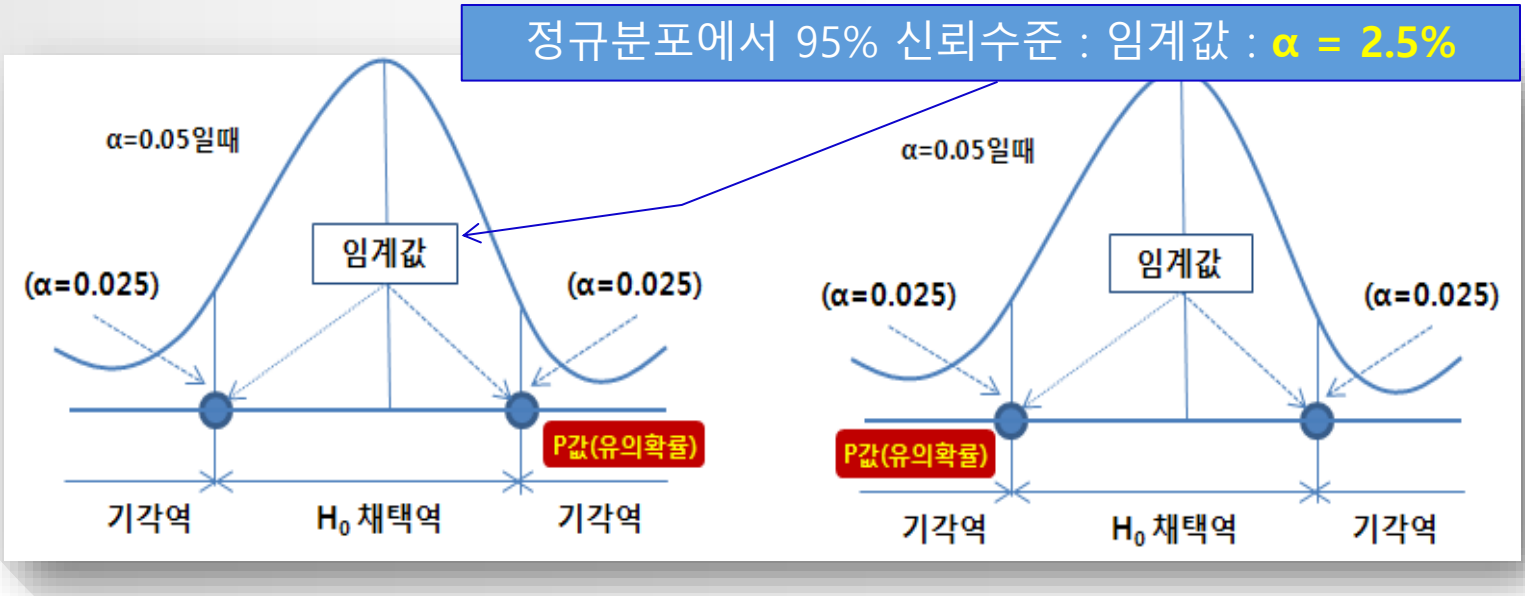


양측검정과 임계값

- **양측검정(2-sided test) : 임계값 2개**

H_0 : 성별에 따라 만족도에 차이가 없다.(남=여)

H_1 : 성별에 따라 만족도에 차이가 있다.(남≠여)



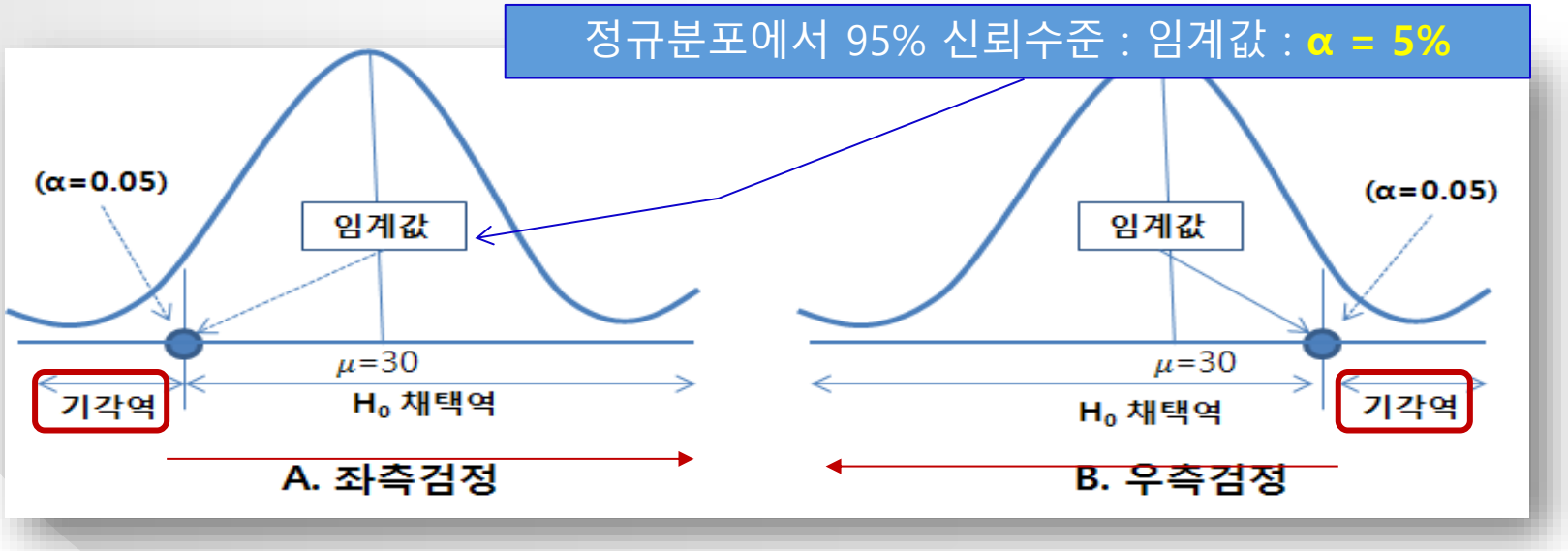
단측검정과 임계값

- **단측검정(1-sided test) : 임계값 1개**

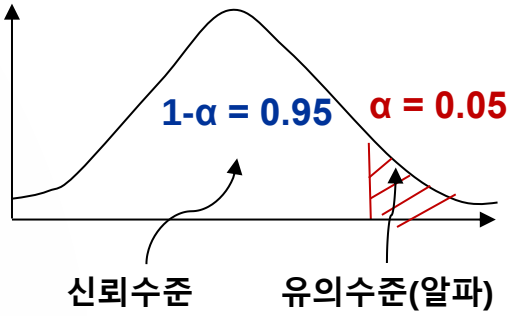
H_0 : 1일 생산되는 불량품의 개수는 평균 30개 이다.(μ=30)

H_1 : 1일 생산되는 불량품의 개수는 평균 30개 이하이다.(μ<30) ▶ 좌측(왼쪽)검정

1일 생산되는 불량품의 개수는 평균 30개 이상이다.(μ>30) ▶ 우측(오른쪽)검정



단계2. 유의수준 결정



- **유의수준(Significant level) = 알파(α)**
  - 귀무가설 채택 또는 기각 기준(임계값)
  - 알파(α) : 가설을 신뢰할 수 없는 확률(**통상 5%**)
  - 신뢰수준(1- α) : 가설을 신뢰할 수 있는 확률(**통상 95%**)
  - 유의수준 이상 ➔ 귀무가설 채택(신뢰할 수 있는 확률 5% 이상)
  - 유의수준 미만 ➔ 귀무가설 기각(신뢰할 수 있는 확률 5% 미만)
  - 알파(α) vs 신뢰수준(confidence level) : 서로 반비례 관계

유의수준 결정

H_0 = '신약A는 A암 치료에 효과가 없다.'

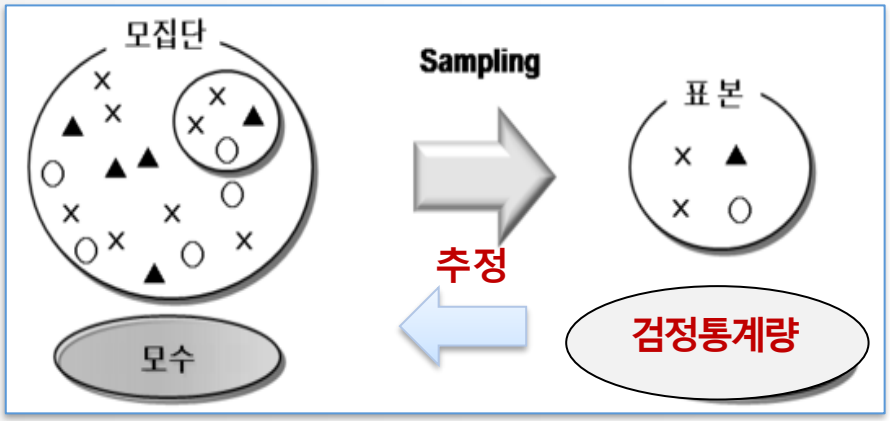
- **유의수준 결정**
  - ✓ 일반 사회과학 분야 : **α=0.05(5%)**
    - ➢ 표본의 통계가 모수를 나타내는 허용 오차 5%(신뢰수준 95%)
    - ➢ 알파(α) : 귀무가설을 신뢰할 수 없는 확률(나오기 어려운 확률)

    예) 100마리 중에서 5마리 이하로 치료 효과가 없는 경우(H_0기각)
  - ✓ 의.생명분야 : **α=0.01(1%)**
    - ➢ 허용 오차 1%(신뢰수준 99%)

    예) 100마리 중에서 1마리 이하로 치료 효과가 없는 경우(H_0기각)

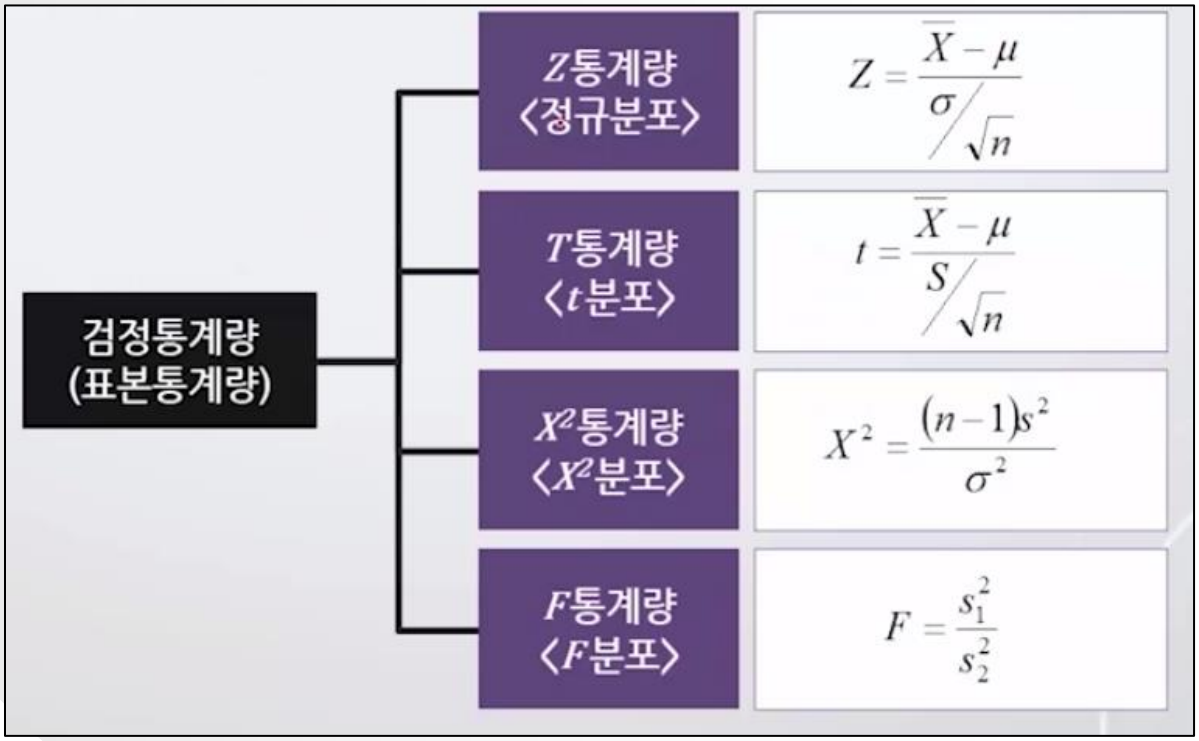
단계3 : 검정통계량과 유의확률

- **검정통계량(Test statistic)**
  - ✓ 표본의 특성을 설명하는 값(추정량 : 모수 추정에 이용)
  - ✓ 가설검정을 위해 수집된 표본 자료로 계산한 통계량
  - ✓ 유의수준(**α**)과 비교하여 가설검정에 사용되는 통계량
  - ✓ 모수 추정과 가설검정(귀무가설 채택/기각)에 이용되는 통계량



검정통계량

- **검정통계량 : 표본으로 모수(평균, 분산 등)를 검정에 이용되는 통계**
  - ✓ 표본분포의 유형에 따라서 검정통계량의 계산식 다름



모집단 정규분포,
모분산 알려진 경우
모평균 검정(z검정)

모집단 정규분포,
모분산 모르는 경우
모평균 검정(t검정)

모집단의 분산과
표본의 분산비
모분산 검정(X^2검정)

두 모집단의
분산차이 검정
모분산 검정(F검정)

단계3 : 검정통계량과 유의확률

[가설]

귀무가설(H_0) : '신약A는 A암 치료에 효과가 없다.'

대립가설(H_1) : '신약A는 A암 치료에 효과가 있다.'

귀무가설 지지 확률

[가정]

생쥐 100마리를 대상으로 신약A를 투약한 결과 검정통계량의 **유의확률**(P=0.03)이 나왔다.

- 이때 귀무가설은 기각되는가?

- **사회과학분야 임계값 : α=0.05(P < 5%미만)**
  - 적어도 95마리 이상 효과 → **H_1 채택**
- **의.생명분야 임계값 : α=0.01(P < 1%미만)**
  - 적어도 99마리 이상 효과 → **H_1 기각**

검정통계량과 가설검정 예

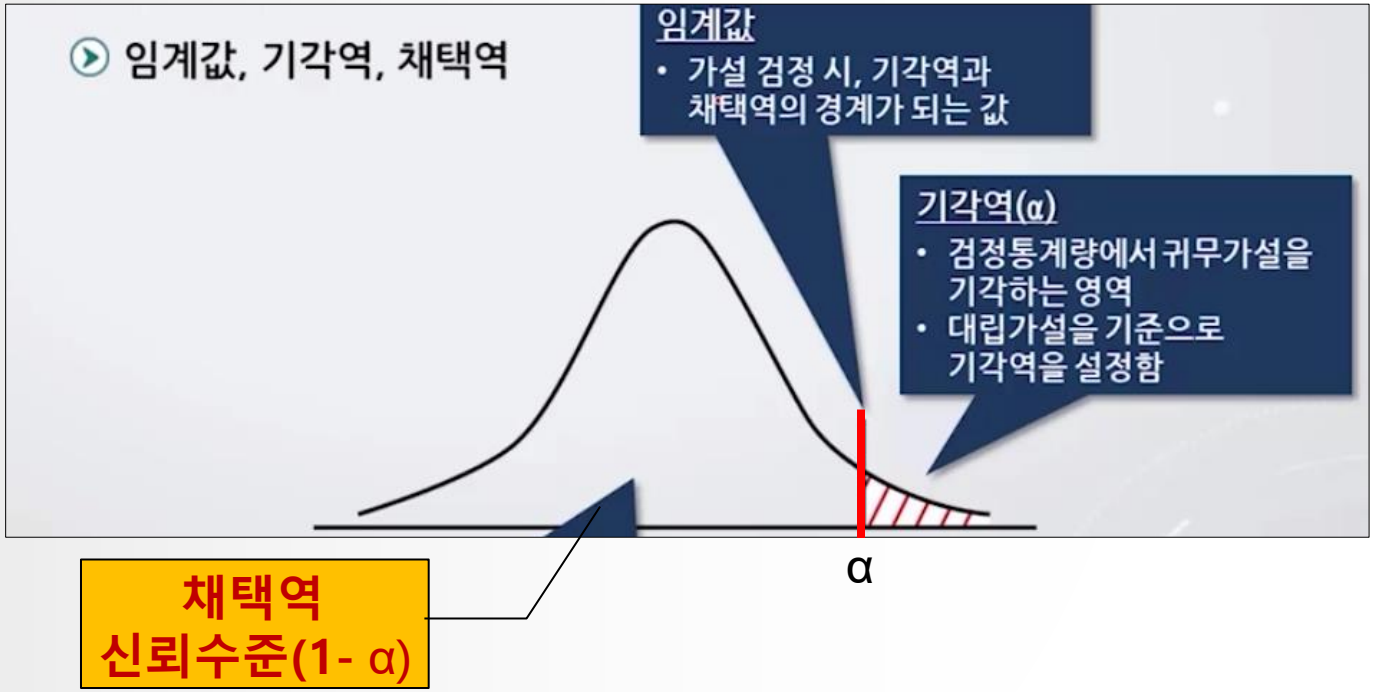
귀무가설(H_0) : '학력수준에 따라 제품만족도에 차이가 없다.'를 검정하기 위해서 **t검정**을 수행한 결과 검정통계량 t값=10.652, 유의확률 p값=0.012가 나왔다. 이때 유의수준 α=0.05 일 때 귀무가설은 기각 되는가 또는 채택 되는가?

검정통계량 t=10.652값에 의해서 유의확률 p=0.012가 검정 결과로 나온 경우 유의확률이 유의수준 α=0.05 보다 낮기 때문에 귀무가설은 기각된다.

[결론] '학력수준에 따라 제품만족도에 유의미한 차이가 있다'라고 볼 수 있다.

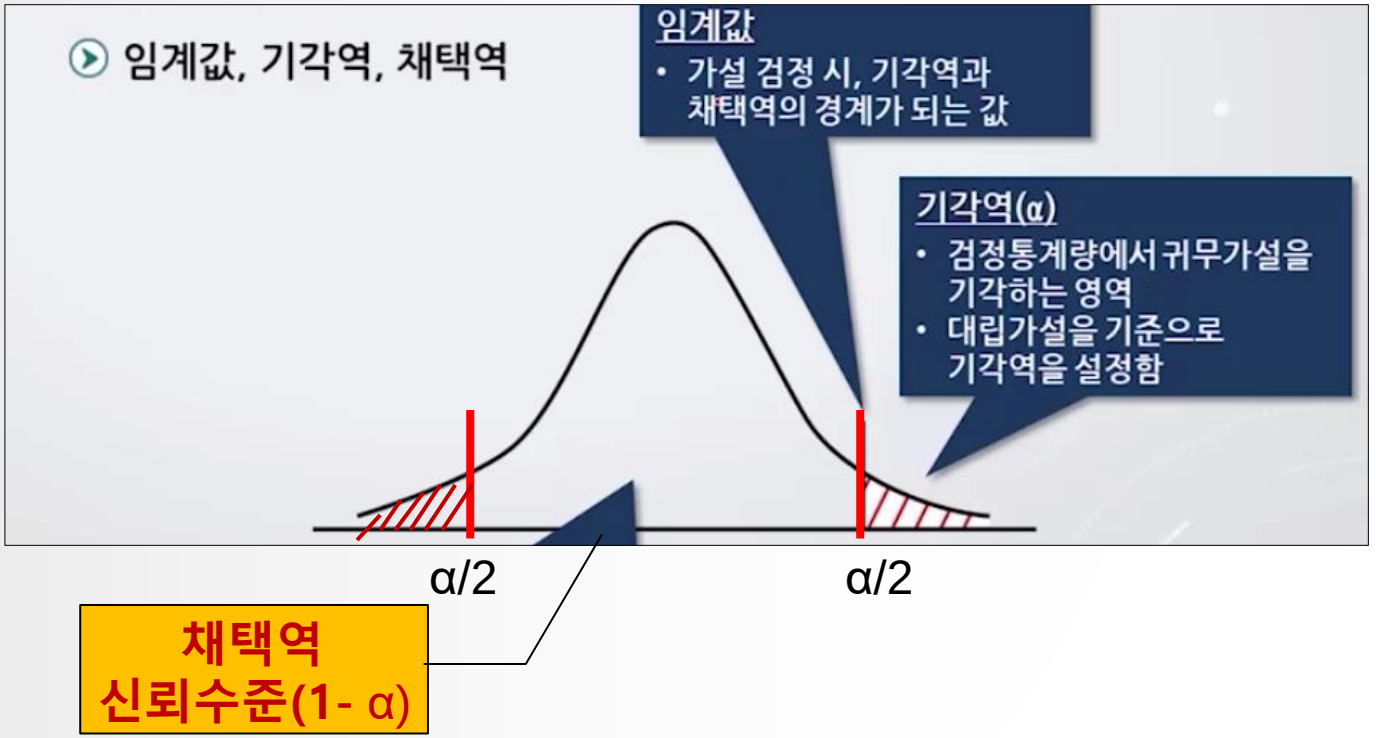
단계4 : 가설 기각과 채택

- 단측검정 : 한 쪽 임계값을 갖는다.



단계4 : 가설 기각과 채택

- 양측검정 : 양 쪽 임계값을 갖는다.



가설 기각 사례

- **유의확률(p) < 유의수준(α) 사례**
  - H_0 : '영양소별 효과의 차이는 없다'에서 유의수준(α=0.05) 일 때 유의확률(p-value) 0.04가 나왔다면 p(0.04) < α(0.05) ➔ 귀무가설 기각
  - 영양소별 효과의 차이가 없을 확률이 낮기 때문에 귀무가설 기각
  - 이때 '통계적으로 유의하다.'라고 해석, p<0.01이면 매우 유의하다.
  - p < 0.05 수준이면 통계적으로 유의적인 차이를 보인다.

7. 모평균 가설검정

- **숫자형 자료의 추론**

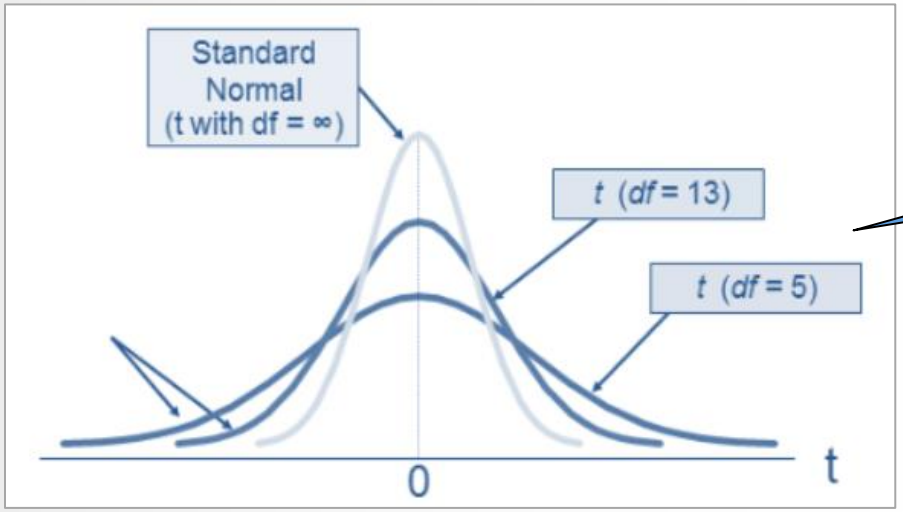
  모수가 수치형 자료(연속형, 이산형)인 경우 모평균의 가설을 검증하는 과정

  ✓ **모평균의 가설검정 유형**

  1) 모평균 검정(단일표본 t검정) : 모평균과의 차이 검정
  2) 두집단 평균 차이검정(독립표본 t검정) : 두 모집단의 평균 차이 검정
  3) 대응두집단 평균 차이검정(대응표본 t검정) : 한 모집단의 전과 후 차이 검정
  4) 분산분석 : 분산을 이용한 3집단 이상 평균차이 검정

Z분포 vs T분포

- **Z분포 vs T분포**
  - ✓ Z분포 : 표준정규분포(μ=0,σ=1), 표본수 충분히 큰 경우(n>30개 이상)
  - ✓ T분포 : 표본수 작은 경우(n<30개 미만) z분포 대신 사용하는 확률분포
  - ✓ 정규분포 가정 : 표본 크기가 클 수록 정규분포 모양 비슷함



자유도(df)가 클 수록
정규분포와 비슷해짐

Z 분포 vs T분포

Z-검정 vs T-검정

- **Z-검정**
  - ✓ **모집단 정규분포이고, 모집단의 분산(표준편차)이 알려진 경우**
  - ✓ 표본의 평균과 모집단의 표준편차를 이용하여 모평균 추정/검정(Z분포)
  - ✓ 기본가정 : 정규분포
  - ✓ 기본가설 : 모평균과 차이가 없다.

Z통계량 〈정규분포〉	$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$

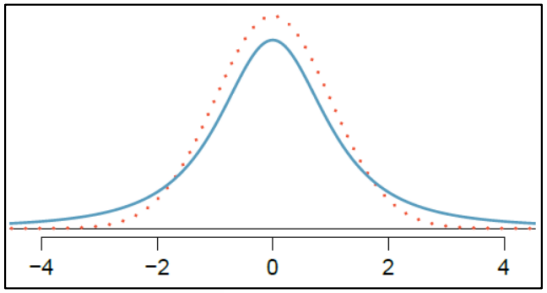
- **T-검정**
  - ✓ **모집단 정규분포이고, 모집단의 분산(표준편차)이 알려지지 않은 경우**
  - ✓ 표본의 평균과 표준편차 이용하여 모평균 추정/검정(T분포)
  - ✓ 기본가정 : 정규분포
  - ✓ 기본가설 : 모평균과 차이가 없다.

T통계량 〈t분포〉	$t = \frac{\overline{X} - \mu}{S / \sqrt{n}}$

❖ t검정을 주로 이용 : 모집단의 실제 모평균과 모분산이 알려지지 않은 경우가 대부분으로 표본을 통해 이를 추정 및 검정한다

1) 모평균 t검정

- **모평균 t 검정(단일표본 t검정)**
  - ✓ 모집단의 모평균(μ)과 표본의 평균 간의 차이가 있는지를 검정
  - ✓ 기본가정 : 확률변수(x)는 정규분포이다.(모수검정)
  - ✓ 기본가설 : 모평균과 차이가 없다.
  - ✓ 귀무가설 : 모집단의 평균과 표본 평균은 같다.



❖ 모수검정 : 확률변수가 정규분포라고 가정될 때 사용할 수 있는 검정 방법

모평균 t검정 사례

- **모평균에 대한 t 검정**

✓모집단이 근사적인 정규분포이고 표본의 크기가 30보다 작을 때 사용하는 검정 방법

[사례] 2001년에 성인여자(19~24세)의 50m 달리기 평균 기록은 9.6초이다. 2015년 무작위로 성인 여자 25명을 뽑아 50m 달리기 기록을 조사하였더니 평균 9.85초, 표준편차는 0.33초이다. 유의수준 5%에서 성인여자의 50m 달리기 평균 기록이 9.6초에서 변했다고 할 수 있는지 검증하시오.

- 가설 설정

귀무가설(H_0) : $\mu = 9.6$

대립가설(H_1) : $\mu \neq 9.6$

$$t = \frac{9.85 - 9.6}{0.33/sqrt(25)}$$

단일표본 t검정 python 코드

```
ttest_1samp = stats.ttest_1samp(sample, 9.6,
                                alternative='two-sided') # 양측검정
print('t검정 통계량 = %.3f, pvalue = %.5f'%(ttest_1samp))
```

t검정 통계량 = 3.664, pvalue = 0.00123

[결과 해석]

p-value 0.00123은 유의수준 0.05보다 작으므로 영가설을 **기각할 수 있다**. 따라서 유의수준 5%에서 성인여자의 50m 달리기 평균 기록이 9.6초에서 변했다고 할 수 있다.

2) 두집단 평균 차이 t검정

- **두집단 평균차이 t검정(독립표본 t검정)**
  - ✓ 서로 독립된 모집단으로 부터 추출된 표본의 평균 차이 검정
  - ✓ 기본 가정 : 두 집단의 분포는 동일하다.(등분산성 검정)
  - ✓ 기본 가설 : 두 집단간 평균의 차이는 없다.
  - ✓ 예) A음료수에 대한 남.녀간의 만족도에 차이가 있는지 or 없는지

두집단 평균 차이 t검정 사례

- **독립인 두 모집단에서 평균 차이 t 검정**
  - ✓ 두 모집단은 서로 독립이며 근사적인 정규분포이나 표본 크기가 30보다 작을 때

[사례] 10대와 20대의 스마트폰 이용 시간을 조사한 결과는 다음 표와 같다. 유의수준 5%에서 10대보다 20대가 스마트폰 평균 이용 시간이 많다고 할 수 있는지 검증하시오. 단, 스마트폰 이용 시간은 근사적인 정규분포를 따른다고 가정한다.

구분	n	$\bar{x}$	s
10대	24	185분	47분
20대	28	223분	58분

- 가설 설정

귀무가설(H_0) : $\mu_1 - \mu_2 = 0$

대립가설(H_1) : $\mu_1 - \mu_2 < 0$

μ_1 : 10대 이용 시간 모평균

μ_2 : 20대 이용 시간 모평균

$$t = \frac{185 - 223}{\sqrt{\frac{47^2}{24} + \frac{58^2}{28}}} = -2.61$$

독립표본 t검정 python 코드

```
two_sample = stats.ttest_ind(teenager, twenty,
                             alternative='less') # 좌측검정
print( ' 검정통계량 = %.3f, pvalue = %.3f'%(two_sample))
```

t검정 통계량 = -2.072, pvalue = 0.022

[결과 해석]

p-value 0.022은 유의수준 0.05보다 크므로 영가설을 **기각**할 수 있다. 따라서 유의수준 5%에서 10대보다 20대가 스마트폰 평균 이용 시간이 많다고 할 수 있다.

3) 대응표본 t검정

- 대응표본 T 검정(동일한 모집단)
  - ✓ 동일한 모집단 대상 두 번 반복 측정하여 전과 후 평균 차이 검정
  - ✓ **짝 자료(paired data)에 차이에 대한 검정**
  - ✓ 기본 가정 : 정규분포(정규성 검정)
  - ✓ 기본 가설 : 전과 후 평균의 차이가 없다.

예) A다이어트식품 복용 전과 후 몸무게에 차이가 있는지 or 없는지

3) 대응표본 t검정

[예제] 1주일 다이어트 프로그램에 참여한 73명의 평균 체중 차이(=참가 후 - 참가 전)는 -0.4kg이고 표준편차는 1.4kg이다. 유의수준 1%에서 다이어트 프로그램은 효과가 있다고 할 수 있을까?

[풀이] 1주일 다이어트 프로그램에 참여했을 때 체중 차이의 평균을 μd라고 표기할 때

H_0 : μd = 0 (참가 후 – 참가 전 = 0)
H_1 : μd < 0 (참가 후 – 참가 전 < 0)

검증통계량 값 $z = \frac{\bar{x}_d - \mu_0}{s/\sqrt{n}} =$ -0.4/0.163858 = -2.44

t검정 통계량 = -2.44, pvalue = 0.0073

[결과 해석]
유의수준 1%에서 1주일 다이어트 프로그램은 효과가 있다고 할 수 있다.
(즉 참가 후보다 참가 전 체중이 많다고 할 수 있다.)

3) 대응표본 t검정

[문제] 2013년 프로 축구 K리그 클래식 총 경기 수는 267 경기이다. 각 경기에서 홈 팀 골수와 방문 팀의 골수 차이의 평균을 계산하니 0.24이고 표준편차는 1.5이다. 유의수준 1%에서 홈 팀이 유리하다고 할 수 있는지 검정하시오.

홈팀 골 – 방문 팀 골 = 0.24
홈팀 골 수가 조금 많다는 의미

[풀이] 골수 차이의 평균을 μd라고 표기할 때 가설

$H_0 : \mu_d = 0$ (홈 팀 골수 – 방문 팀 골수 = 0)

$H_1 : \mu_d < 0$ or $\mu_d \neq 0$ (홈 팀 골수 – 방문 팀 골수 < 0)

검증통계량 값 $z = \dfrac{\bar{x}_d - \mu_0}{s/\sqrt{n}} =$ 0.24/1.5/$\sqrt{267}$= 0.24/0.09178 = 2.61

t검정 통계량 = 2.61, pvalue = 0.0045

[결과 해석]
p-value = 0.0045 은 유의수준 0.01보다 작으므로 영가설을 기각할 수 있다.
따라서 유의수준 1%에서 홈 팀이 유리하다고 볼 수 없다.
(즉 홈 팀 골수보다 방문 팀 골수가 많다고 할 수 있다.)

4) 분산분석

분산분석

3개 이상의 집단이 있을 때 평균을 비교하고 싶다. 예를 들어 집단 A, B, C가 있을 때 A와 B의 평균을 비교하고 B와 C의 평균을 비교한다. 마지막으로 C와 A의 평균을 비교한다. 그러나 이러한 전략은 신뢰할 수 없다. 집단이 10개만 되더라도 쌍으로 비교할 평균은 45 개나 된다. 이렇게 많은 비교를 하다 보면 집단들 간에 차이가 없더라도 결국에는 어떤 차이를 발견할 수 있게 된다.

분산분석(analysis of variance, ANOVA)은 여러 개의 평균을 비교할 때 사용하는 통계적 기법이다.

분산분석 절차를 수행하기 전에 반드시 다음 3가지 조건을 점검해야 한다.
- 관측값은 집단 내에서 독립이고 또 다른 집단의 관측값과도 독립이어야 다.
- 각 집단 내에서 관측값은 거의 정규분포를 따른다.
- 각 집단의 표준편차는 거의 같다.

분산을 분석해서 평균을 비교하는 기법이다.

4) 분산분석

분산을 분석해서 평균을 비교하는 기법이다.

- 분산분석
  - ✓ 서로 독립된 3개 이상 모집단 간의 평균 차이 검정
  - ✓ 집단 내의 분산, 평균, 각 집단의 평균 차이에 의해서 생긴 집단 간 분산의 비로 만들어진 F분포를 이용한 가설검정
  - ✓ 기본 가정 : 각 집단의 분포는 동일하다.(등분산성 검정)
  - ✓ 기본 가설 : 각 집단간 평균의 차이는 없다.
  - ✓ 대립 가설 : 적어도 한 집단 이상 평균의 차이가 있다.

  예) A음료수에 대한 연령별(20대,30대,40대) 만족도에 차이가 있는지 or 없는지

분산분석(analysis of variance, ANOVA)은 여러 개의 평균을 비교할 때 사용하는 통계적 기법

❖ 분산분석 절차를 수행하기 전에 반드시 다음 3가지 조건을 점검해야 한다.

1. 관측값은 집단 내에서 독립이고 또 다른 집단의 관측값과도 독립이어야 한다.
2. 각 집단 내에서 관측값은 거의 정규분포를 따른다.
3. 각 집단의 표준편차는 거의 같다.

분산분석 방법

- **분산분석 방법**
  - ✓ 방법 : 일원분산분석, 이원분산분석, 다원 변량 분산분석

종류	변수 개수	사례
일원 분산분석 (One-way ANOVA)	독립변수 : 1개 종속변수 : 1개	교육 방법에 따른 성적 비교 독립변수(범주형) : 방법1, 방법2, 방법3 종속변수(숫자형) : 성적
이원 분산분석 (Two-way ANOVA)	독립변수 : 2개 종속변수 : 1개	쇼핑몰 고객의 연령대(30,40,50대), 시간대(오전/오후)별 구매현황 분석 독립변수(범주형) : 연령대, 시간대 종속변수(숫자형) : 구매현황
다원 변량 분산분석	독립변수 : 1개, 2개 종속변수 : 2개	교육 방법에 따른 국어와 영어 성적 비교 독립변수(범주형) : 방법1, 방법2, 방법3 종속변수(숫자형) : 국어성적, 영어성적

4) 분산분석

분산분석 절차

통계학 시험을 치른 33명의 점수와 출신 학과를 기록한 것입니다. 점수는 $y_1, y_2, \cdots, y_{33}$로 표기한다. 출신 학과는 A, B, C인데 학과 별로 평균 점수가 다른 지를 검증하려고 한다.
따라서 영가설은 "학과별 평균 점수가 모두 같다"이고
대립가설은 "학과별 평균 점수가 모두 같지는 않다"

$$H_0 : \mu_A = \mu_B = \mu_C$$
$$H_1 : \text{최소한 하나는 다르다}$$

1. **총제곱합**(total sum of squares, SST) : 각 점수와 전체 평균 사이의 거리를 제곱하여 합한 값
   총제곱합은 세 학과의 평균이 같다고 가정할 때 데이터의 전체 변동을 보여준다

$$SST = \sum (y - \bar{y})^2 = 1113$$

2. 오차제곱합(error sum of squares, SSE) : 각 점수와 출신 학과 평균 사이의 거리를 제곱하여 모두 합한 값

$$SSE = \sum_{j=1}^{k} \sum (y - \bar{y}_j)^2 = 797.1$$

3. 그룹간 제곱합(sum of squares between groups) : 총제곱합에서 오차제곱합을 뺀 값

$$SSB = SST - SSE = 1113 - 797.1$$

=315.9

4. 평균 제곱합(MSB) : 각 제곱합을 자유도로 나눈 값을

5. 오차 평균제곱합(MSE)

6. F = MSB / MSE (두 개의 자유도 갖는 F분포를 다른다.)

4) 분산분석

변동	자유도	제곱합	평균제곱합	검증통계량
Between	$k-1$	SSB	MSB	$F = \frac{MSB}{MSE}$
Error	$n-k$	SSE	MSE	
Total	$n-1$	SST		

F 값은 두 개의 자유도(분자의 자유도 $k-1$, 분모의 자유도 $n-k$)를 가진 F 분포를 따릅니다. 집단별 평균이 거의 비슷하다면, 즉 영가설이 맞다면 F값이 크지 않습니다. 그러나 영가설이 틀리다면 F값이 커집니다. 따라서 유의확률은 F 분포에서 오른쪽 꼬리의 확률입니다.

출신 학과별 통계학 시험 평균 점수가 다른 가를 검증하는 분산분석 표는 다음과 같습니다.

변동	자유도	제곱합	평균제곱합	검증통계량
Between	2	315.9	157.93	$F = \frac{157.93}{26.57} = 5.944$
Error	30	797.1	26.57	
Total	32	1113		

자유도가 (2,30)인 F분포에서 5.944보다 클 확률은 0.0067이다. 유의수준 1%에서 학과별 평균 점수는 모두 같다 라고 볼 수 없다.