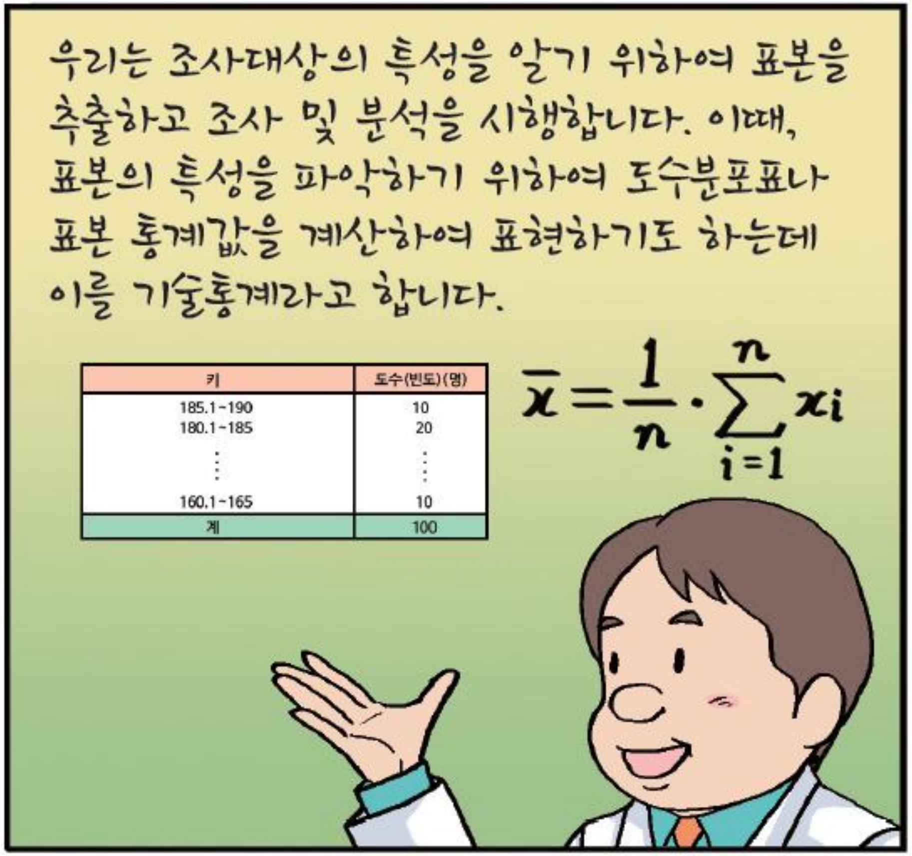
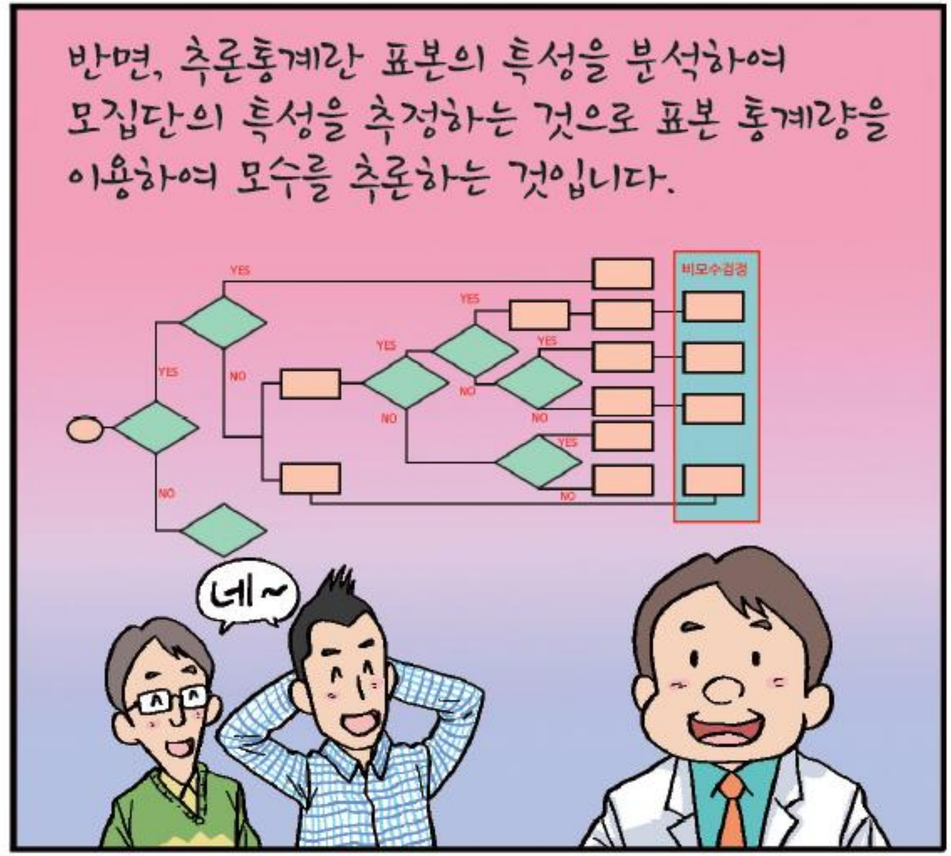
CHAPTER

5

기술통계와 추론통계


인표야, 논문은 잘 돼가? 장애인의 직업결정요인에 대해 연구 중이라며?
응, 직업재활시설 근로자들을 표본으로 해서 통계를 내봤어.
통계검증은 해봤고?
통계검증? 그게 뭐지?
안 가르쳐줘~
하하하
통계의 검증은 측정한 결과가 가설을 지지할 수 있을지 없을지 결정하기 위해 필요한 것입니다.
통계검증은 기술통계와 추리통계로 설명될 수 있습니다.
기술통계
추리통계


우리는 조사대상의 특성을 알기 위하여 표본을
추출하고 조사 및 분석을 시행합니다. 이때,
표본의 특성을 파악하기 위하여 도수분포표나
표본 통계값을 계산하여 표현하기도 하는데
이를 기술통계라고 합니다.
키 | 도수(빈도)(명)
185.1~190 | 10
180.1~185 | 20
⋮ | ⋮
160.1~165 | 10
계 | 100
$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$


반면, 추론통계란 표본의 특성을 분석하여
모집단의 특성을 추정하는 것으로 표본 통계량을
이용하여 모수를 추론하는 것입니다.
비모수검정
YES
NO
네~

01 기술통계

1) 기술통계의 이해

- 수집한 자료를 요약, 조직, 단순화해서 변수들의 기본적인 정보와 변수 간의 관계에 대해 분석하는 통계 방법
- 도수분포표, 백분율, 평균, 표준편차 등에 대한 분석이 포함
- 기술통계를 통해 연구대상의 특성을 파악

2) 기술통계의 이해

(1) 도수분포 및 백분율

도수분포(빈도분포)란 자료를 일정한 수의 범위로 나누어 분류하고, 각 범위별로 수량의 분포를 보는 것

키	도수(빈도)(명)	백분율(%)
185.1~190	10	10
180.1~185	20	20
:	:	:
:	:	:
160.1~165	10	10
계	100	100

2) 기술통계의 이해

(2) 중심경향치

- 측정 자료들을 빈도곡선으로 나타내었을 때 어디에 집중되어 있는가를 파악하기 위한 통계량
- 평균, 중앙값, 최빈값

1 평균값

- 가장 자주 사용되는 중심경향치
- 평균값은 어디서나 존재
- 자료의 모든 값을 고려
- 극단치에서 변동의 폭이 심함

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$ 즉 (평균)=(자료 전체의 합)÷(지료의 개수)

$$= \frac{(\text{자료 전체의 합})}{(\text{자료의 개수})}$$

2 중앙값

- 자주 사용하지만 평균값만큼은 아님
- 중앙값은 어떤 경우에나 존재함
- 자료의 모든 값 자체를 고려하지는 않고, 몇 개의 값이 있는지만 고려
- 중앙값은 극단적인 자료에서도 크게 변하지 않음
- 극단치가 있으면, 평균값보다 중앙값이 자료의 대표성을 띈다고 말할 수 있음

예시

표 5–2. 중앙값의 예

변수	평균	중앙값
1 2 3 4 5 6 7 8 9 10	5	5
1 2 2 4 5 6 6 7 10	4.8	5
1 2 2 4 5 6 6 7 10 12	5.5	5.5

* 홀수이면 중앙에 위치하는 값, 짝수이면 가운데 좌우 자료의 평균이 중앙값이 됨

3 최빈값

- 아주 드물게 사용
- 특히(이름으로 표현하는)명목변수나, (제한된 수의 실제 값만을 갖는)이산변수를 기술할 때 사용
- 최빈값은 하나 이상일 수 있지만, 전혀 없을 때도 있음
- 최빈값은 자료의 모든 값을 반영하지는 않음

예시

표 5-3. 최빈값의 예

연령군	빈도
1 : 10대	12명
2 : 20대	17명
3 : 30대	23명
4 : 40대	19명
5 : 50대	8명

최빈값은 3 : 30대이다.

(3) 산포도

- 측정된 자료에서 관측값들이 얼마만큼 퍼져있는를 측정하는 척도로 변동성의 척도, 분산도라고 함
- 산포도에는 범위, 분산, 표준편차, 사분위수, 변동계수

1 범위

- 연속형 변수에서 사용
- 변수의 최고값에서 최소값을 뺀 것
- 연속적 변인인 경우 계산법(R=최고값−최소값+1 : 1을 더하는 이유는 최고값 상한계에서 최소값 하한계까지의 거리가 범위가 되기 때문.

예를 들어, 3, 6, 8, 9의 네 점수가 있는 경우 범위는 R=9−3+1=7)

- 장점은 계산하기에 간편하고 쉽게 이해할 수 있다는 것
- 단점은 최고값과 최소값에 의해서만 범위가 결정되므로 그 사이에 존재하는 값들이 퍼져있는 정도를 알 수 없음
- 극단치가 있을 때는 영향을 크게 받아 변동을 표현하는 데 적합하지 않음

예시

* 연령 변수의 범위

범위: 최대값 87, 최소값 44 87−44=43

표 5-4. 산포도의 예

성별	연령	성별	연령	성별	연령	성별	연령
1	67	2	62	2	87	1	47
2	84	2	78	1	57	1	80
1	44	1	45	2	60	2	78

2 분산과 표준편차

분산(sample variance)

- 측정값의 편차를 제곱하여 계산
- 모분산이나 표본분산의 측정단위는 관측값들의 측정단위와 일치하지 않음

표준편차(standard deviation; SD)

- 분산의 양의 제곱근(square root)
- 측정값의 측정단위와 일치시키기 위함

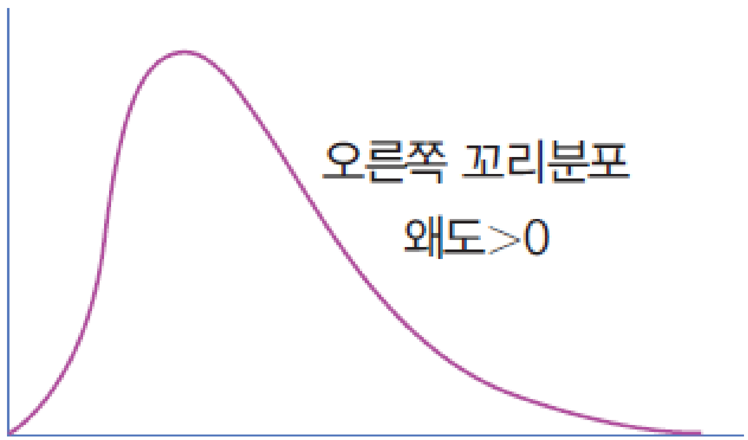
표준편차	내 용
큼	자료가 평균값을 중심으로 광범위하게 분포
작음	평균값을 중심으로 밀접하게 분포

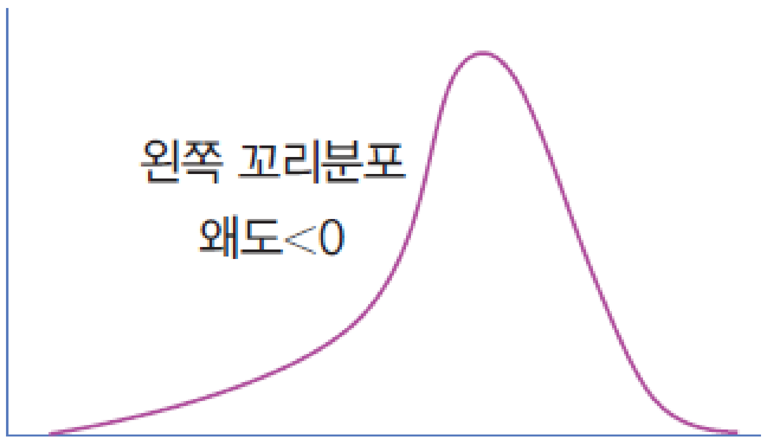
(4) 비대칭도

- 분포의 모양이 중앙 위치에서 상하좌우로 치우진 정도
- 왜도와 첨도

1 왜도

연속형 자료에서 좌우대칭으로

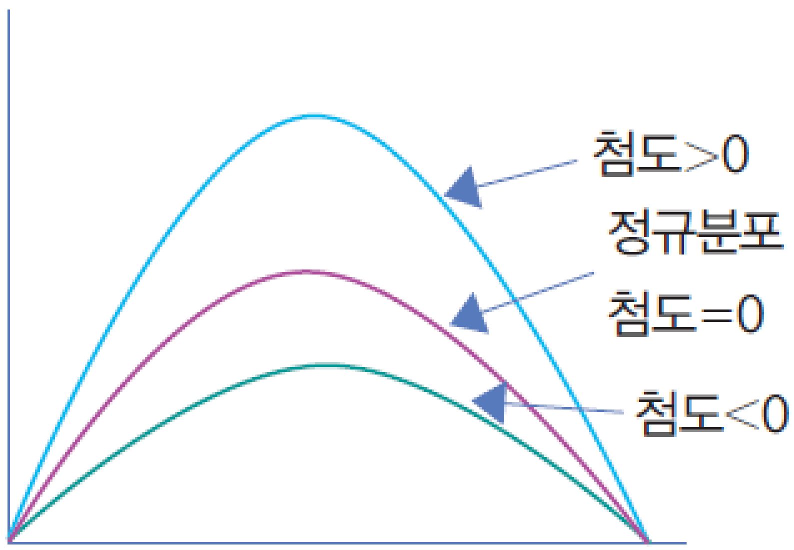




2

첨도

분포가 얼마나 뾰족한지 아니면 편평한지를 나타내주는 통계량



01 기술통계

2) 자료의 분포

(1) 정규성

- 많은 연구에서 사용된 자료는 정규분포를 가정하고 사용
- 자료가 정규분포를 가정한다면 이것을 모수 검정법
- 30개 이상의 자료는 정규분포를 이룬다고 가정
- 정규성 검정은 어떤 분석을 하기 전에 모수적 통계법을 사용할 것인지 아니면 비모수적 통계법을 사용할 것인지를 판단할 때도 사용

(2) 동질성

- 집단이 동질하지 않으면 처치 효과를 명확히 알 수 없음.
- 집단의 동질성을 확인하는 것은 실험 전에 고려

예시 **집단의 동질성 확인**

학부 3, 4학년 때 해당 전공별 실습을 나가게 된다.

실습 후 학생들은 과연 자신의 전공에 대한 진로선택 수준이 향상 되었을까?

전공실습을 이수한 학생(실험집단)과 아직 이수하지 못한 학생(통제집단)의 진로선택에 대한 생각을 묻는 사전-사후비교로 실험하였다.

실습을 이수한 집단과 이수하지 않은 집단은 동질집단이며, 과연 실습은 이수한 학생들의 진로선택에 영향을 주었는가를 통계적으로 검증하라.

표 5-5. 실습 여부에 따른 집단 차이

	실습을 마친 학생들	실습을 마치지 않은 학생들
실습 전 진로선택 수준	80	80
전공선택에 대한 만족도	낮음	높음
봉사활동 경험	많음	적음
학업수준(학점)	낮음	높음
실습 후 진로선택 수준	90	75

공변량분석

시간, 환경, 개인 요인 등 여러 가지 실험 결과에 영향을 미치는 요인들을 통계적인 방법으로 통제하고, 독립변수들 간 영향이 있는지를 보는 것

02 추론통계

- 표본의 특성을 분석하여 모집단의 특성을 추정하는 것으로 표본통계량을 이용하여 모수를 추론하는 것.
- 가설을 검증할 때 모집단으로부터 추출된 표본을 통하여 추정을 하게 되는 데 이런 과정에서 추론의 근거를 뒷받침해줄 수 있는 가장 중심이 되는 것이 바로 모집단의 함수 형태 즉 분포에 관한 모집단의 분포가 정규분포이면 모수검정을 실시하고 정규분포를 가정하지 못하면 비모수 검정을 실시

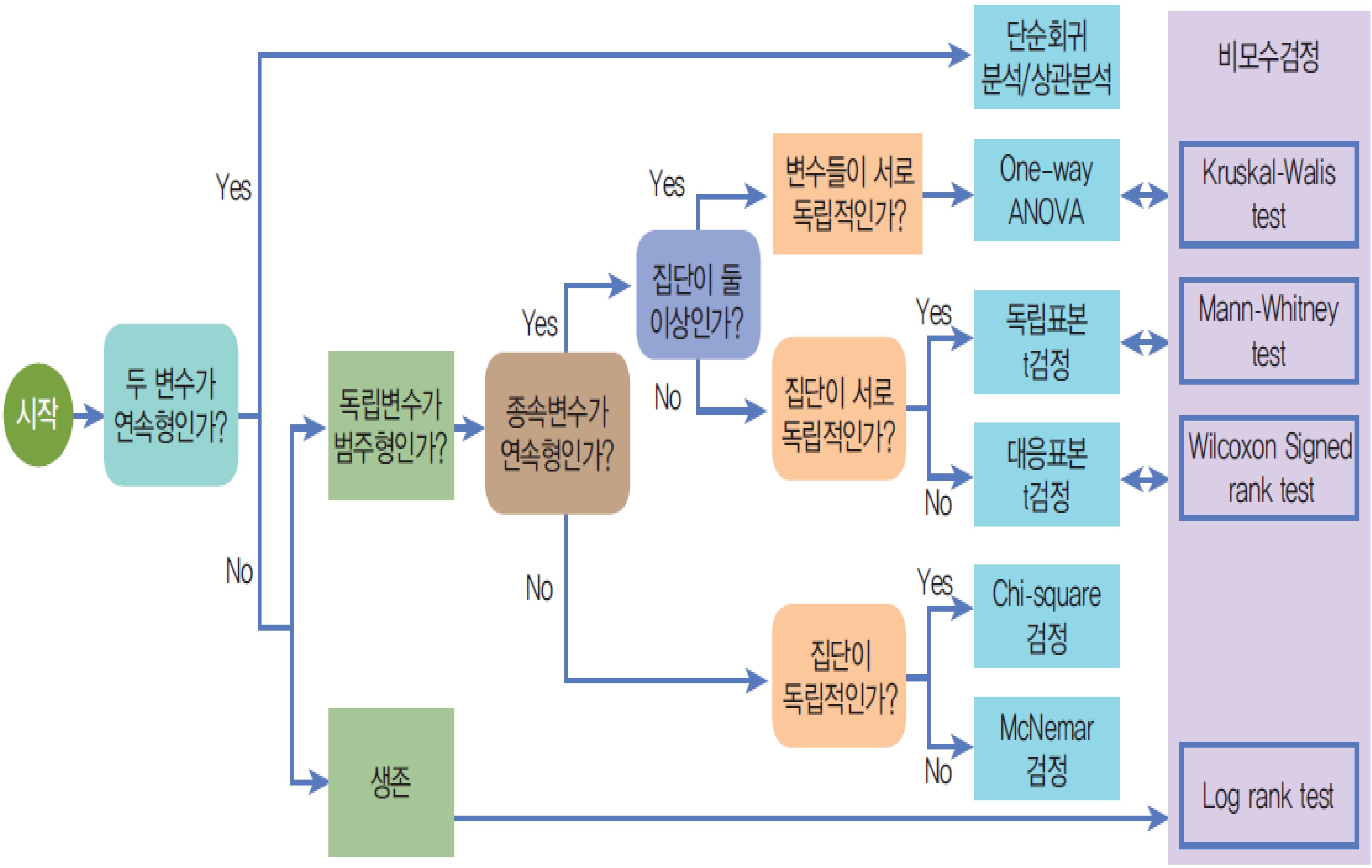


그림 5-4. 변수의 형태와 종류에 따른 분석모형

1) 모수적 방법

- 두 기간에서의 변화를 분석할 때에는 **대응표본 t검정**
- 하나의 기간에서 서로 다른 2개의 집단을 분석할 때에는 **독립표본 t검정**
- 하나의 기간에서 3개 이상의 집단을 분석할 때에는 **일원배치 분산분석**
- 하나의 집단에서 3개 이상의 기간에서의 변화를 분석하고자 할 때에는 **반복측정 분산분석**

2) 비모수적 방법

- 집단이 하나이고 두 기간에서의 변화를 분석할 때에는 **윌콕슨 부호순위 검정(Wilcoxon signed rank test)**

- 하나의 기간에서 서로 다른 2개의 집단을 분석할 때에는 **맨휘트니 검정(Mann-Whitneytest)**

- 하나의 기간에서 3개 이상의 집단을 분석할 때에는 **크러스컬-월리스 검정(Kruskal-Walis test)**

- 하나의 집단에서 3개 이상의 기간에서의 변화를 분석하고자 할 때에는 **후리드만 검정(Friedman test)**

3) 가설검정

- 모집단의 특성에 대한 가설을 설정하고 표본에서 얻은 정보와 비교하여 검토한 후 타당성을 입증해 가는 통계적 추론의 방법

(1) 통계학적 가설 수립

- 모집단의 특성에 대한 자료를 바탕으로 평균을 우선 비교하며 영가설과 대립가설 두 가지로 수립할 수 있다

1 영가설 = 귀무가설

두 집단의 모수 간에 차이가 없다는 가설
귀무가설에서 귀무란 없던것으로 돌아간다는 뜻!
무효화하기 위해 설정하는 모수의 값. H_0
가설을 기각할 것을 전제로 한 가설.

예를 들면 대퇴사두근의 근력운동을 한 집단과 하지 않은 집단 간
대퇴사두근(넙다리네갈래근)의 근력향상에 차이가 없다는 의미의
가설로 영가설(null hypothesis; Ho)은 기각 되기 위해 설정되는 가설

② 대립가설

모수 간에 차이가 있다고 주장을 하고자 설정

귀무가설이 기각된다면 자동적으로 대립가설이 채택. H_A or H_1

두 집단 간의 평균에 차이가 있다($\mu_1 = \mu_2$)는 의미로 채택할 것을 전제로 한 가설

Tip

제1종 오류(α오류) 영가설이 옳은데도 틀렸다고 기각하는 오류

제2종 오류(β오류) 영가설이 틀렸는데도 옳다고 수용하는 오류

검정의 오차

- 연구에서 표본 집단을 가지고 모집단을 추정하기에 오류가 나고 표본 통계량을 가지고 가설을 단정해야 하기에 충분한 오차가 존재한다.
- 통계를 갖고 실상은 아니지만 잘못 판단할 가능성의 2가지 오류가 존재.
- 귀무가설이 옳은데 통계치에 의해 기각되고 대립가설이 채택되는 오류를 **1종 오류**이고 표현은 유의 수준 α 이는 정규분포곡선에서 전체면적에서 신뢰구간의 면적을 뺀 것.
- **2종 오류**는 반대로 귀무가설이 거짓인데 채택한 경우이고 β라 표

1종 오류와 2종 오류

	귀무가설 참일 때	귀무가설 거짓일 때
귀무가설 기각	1종 오류(α)	올바른 의사결정1 − β
귀무가설 채택	올바른 의사결정1 − α	2종 오류(β)

EX)

- **기존에 알려져 있기에 남자들의 키가 여자들보다 크다고 알려져 있습니다. 그런데 한 연구자가 최근 통계청 결과를 살펴보니 이제는 여자의 키가 더 큰 것 같다는 사실을 발견했습니다. 이럴 때 분석을 실시한다면~**

귀무가설 : 남자의 키가 여자보다 크다

대립가설 : 여자의 키가 남자보다 크다

- **기존에 알려진 사실 또는 분석자가 기각하고자 하는 내용이 귀무가설, 그 반대내용이 대립가설이지요?**
- **만약에 유의수준을 0.05로 정했다면 p-value(유의확률)이 0.05보다 작으면**
- **귀무가설을 기각 대립가설 채택, 0.05보다 크다면 귀무가설을 기각 할 수 없게됩니다.**

연구 : 어느 모집단의 우리나라 1학년 성적 평균이 70점인가?

귀무가설 H0은 다음과 같이 설정.

H_0 : μ = 70점 ; 평균이 70점이다.

대립가설 HA 는 다음과 같이 표시 H_A: μ = 70 ; 평균이 70점이 아니다.

연구 사람은 동물인가?

귀무가설 H0 : 사람 = 동물; 사람은 동물이 아니다.

대립가설 HA : 사람 = 동물; 사람은 동물이다.

(2) 검정통계량 선정

- 표본으로부터 계산되는 표본통계량
- 연구의 가설 및 설계방법, 변수의 측정척도에 따라 결정되는 것으로 검정통계량의 결정은 곧 분석방법을 결정하는 것

(3) 유의수준 결정

- 유의수준(level of significance; a)은 통계학적 검정에서 사용하는 판단의 기준

- 제1종 오류를 범할 확률의 최대 허용한계로서 보통 0.05(95% 신뢰)와 0.01(99% 신뢰)을 많이 사용

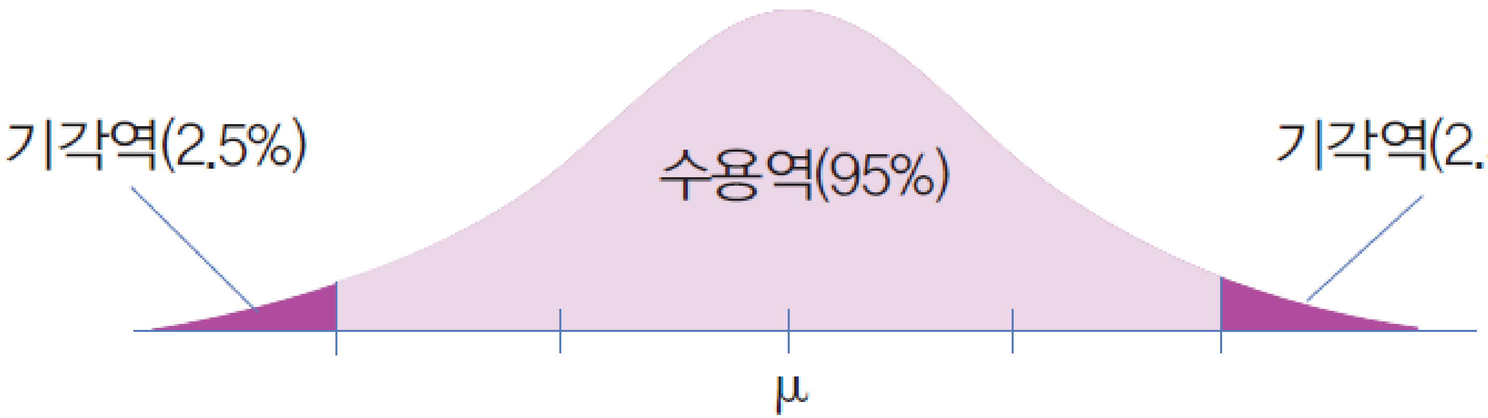


그림 5-5. 정규분포에서 가설검정의 수용역과 기각역

(4) 검정통계량 계산

- 표본자료로부터 검정통계량을 계산하고 정규분포에서 기각역인 범위에 위치하는 값일 때 영가설을 기각하고 수용역에 위치할 때는 영가설을 수용

(5) 판정

- 유의확률(p값)이 이미 설정한 유의수준(α)보다 작으면 영가설을 기각하고 대립가설을 받아들여 ‘두 모집단의 평균은 유의한 차이가 있다’라고 판정
- p값이 유의수준보다 크면 영가설을 기각하지 못하고 ‘두 모집단의 평균은 유의한 차이가 없다’라고 판정

Tip

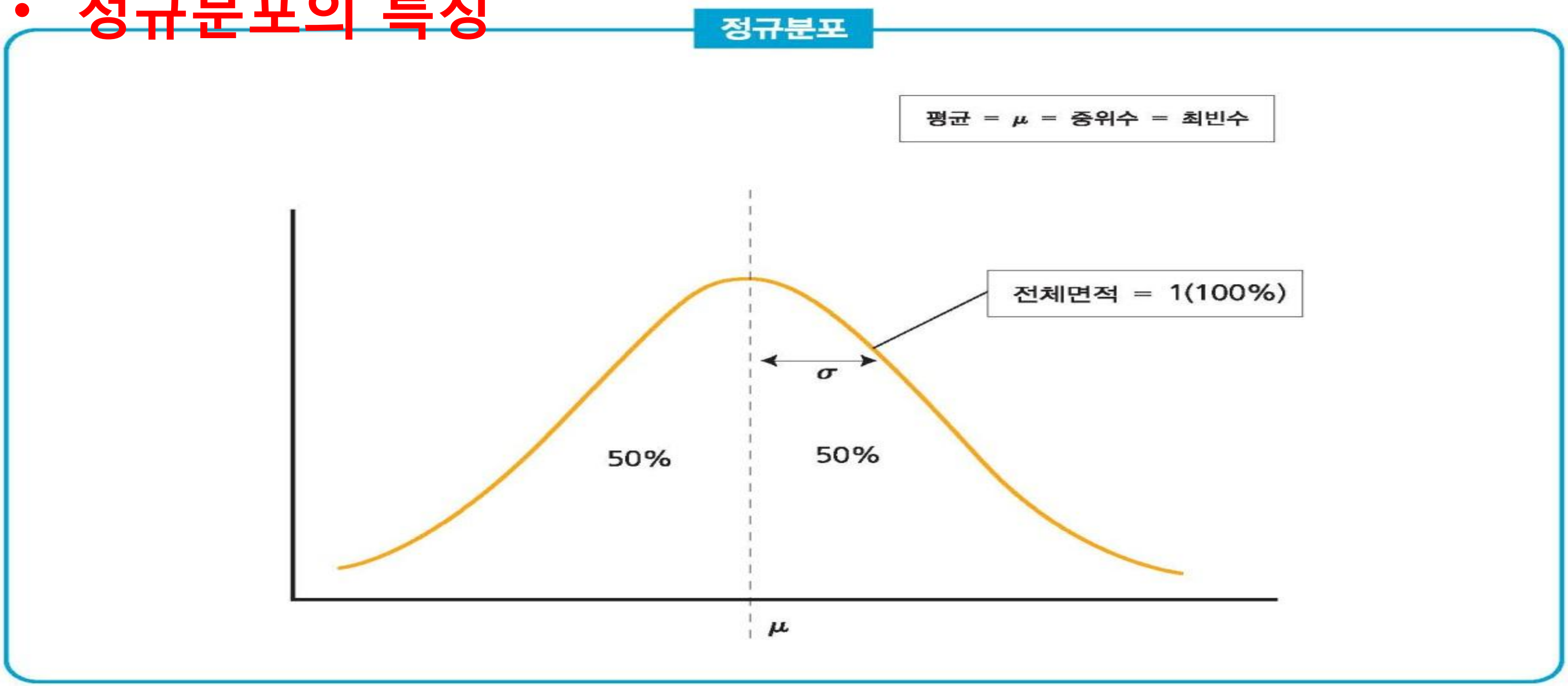
판정을 내릴 때 기각역이 평균을 중심으로 한쪽에 있는 경우(단측 검정)와 양쪽에 있는 경우(양측 검정)가 있는데 이는 결과에 영향을 미치는 요인의 작용방향이 확실치 않으면 양측검정을 실시하고, 요인의 작용방향이 어느 쪽인지 확실하면 단측검정을 실시한다. 예를 들어 대조군을 기준으로 실험군의 결과가 어느 쪽 방향으로 나타날지 모른다면 양측검정을 실시하고, 실험군의 결과가 대조군에 대하여 나타나는 결과의 방향이 명확하다면 이때는 단측 검정을 실시한다.

Mann-Whitney test
Wilcoxon signed rank test
Kruskal-Wallis test
Spearman의 순위상관분석

독립표본 T 검정
대응표본 T 검정
일원배치 분산분석
Pearson의 상관분석

정 규 분 포

- 연속확률변수에 속하고 모분포, 가우스 분포라 하여 이상적인 모형.
- 평균이 (모집단평균)이고 분산(모집단분산)인 곡선이고 평균과 표준편차에 의해정의되고 결정되는 대칭 곡선.
- 정규분포의 특징



- **중심 극한정리**

정규분포를 이루는 모집단에서 추출한 표본집단의 평균과 표본분포는 정규분포를 갖는다.

만약 모집단이 정규분포가 아니라도 표본의 크기가 증가함에 따라 모집단에 관계없이 표본집단 평균은 정규분포에 가까워지는데 이를 중심 극한정리라 한다.

보통 대상자가 30명 이상일 때 모집단에 상관없이 표본집단은 정규분포를 이룬다고 하고 대상자수가 많을수록 정규성에 가까워진다.

양측검정과 단측검정

- **양측검정** (two-tailed test) 은 차이가있다라는 가설에 대한 검정으로 표본 집단이 모집단과 차이가 나는가를 따질 때 사용.

(표본의 평균이 모집단의 평균과 같은가 또는 다른가를 판정하는 가설과 검정방법)

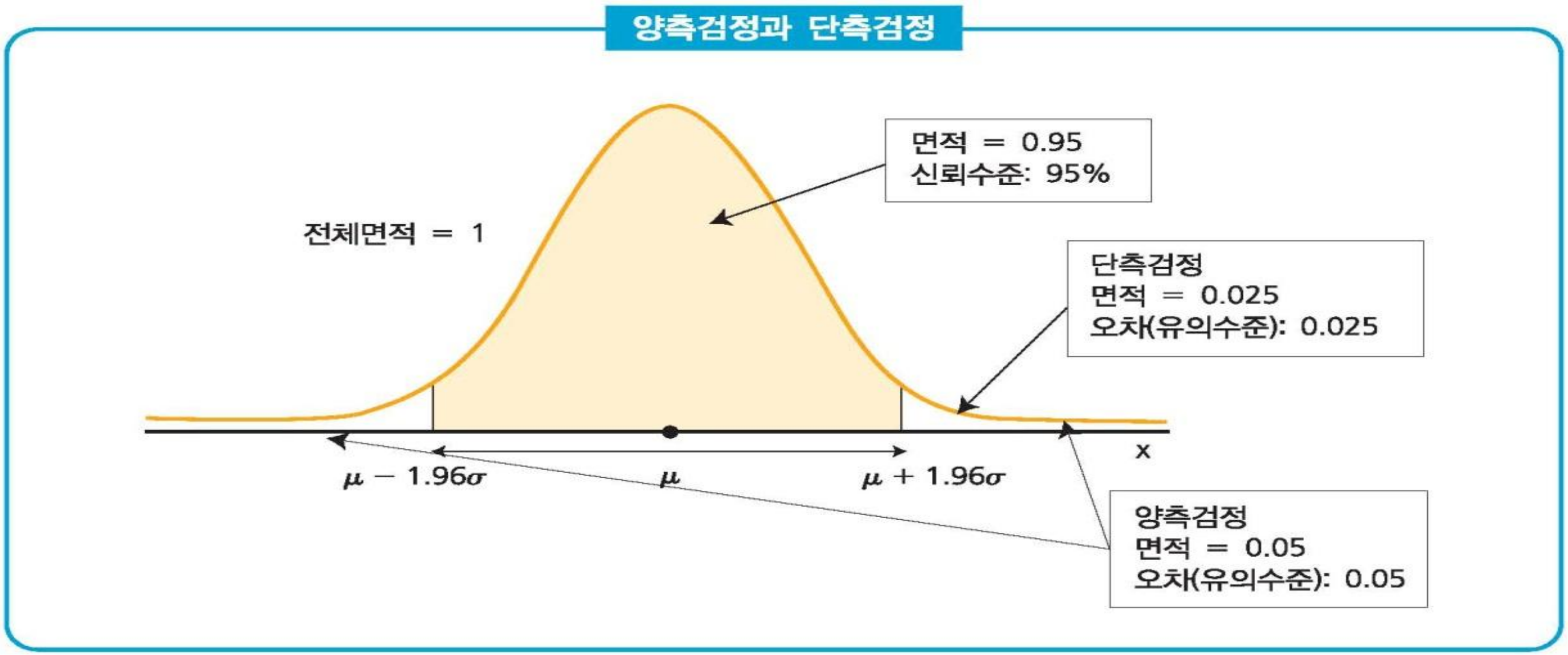
- **단측검정**은 효과가 좋다라는 가설에 대한 검정으로 우열을 가리거나 생산성의 증감을 따지거나 할 때 사용.

(표본의 평균이 모집단의 평균값보다 큰지 작은지 검정하는 경우의 가설과 검정방법)

- 양측검정은 유의 수준 5%의 경우 표준정규분포곡선의 기각되는 부분의 양쪽 끝 부분을 합한 면적이 0.05이므로 양쪽에 0.025씩 두개가 발생하고 표준화된 확률 변수가 1.96일 때 채택되는 부분의 확률이 0.95다.

- 가설검정을 하기 전에 사전 정보에 의하여 보이고자 싶어하는 대립가설이 한쪽 방향인 경우 단측검정을 사용하고, 그런 사전 정보가 없거나 양쪽을 다 검정하고 싶은 경우 양측검정을 사용한다.

  그리고 일단 단측검정이던 양측검정이던 하나를 선택하면, 유의수준은 5%를 사용하도록 설명되어 있다.



((연 구))

우리나라 1학년 성적 평균이 70점보다 높다고 할 수 있는가?

- **H_0 : $\mu \le$ 70점 ; 평균성적보다 낮다.**
- **H_A: $\mu >$ 70점 ; 평균성적보다 높다.**

우리나라 1학년 성적 평균이 70점보다 낮다라고 할 수 있는가?

- **H_0 : $\mu \ge$ 70점 ; 평균성적보다 높다.**
- **H_A: $\mu <$ 70점 ; 평균성적보다 낮다.**

((연구))

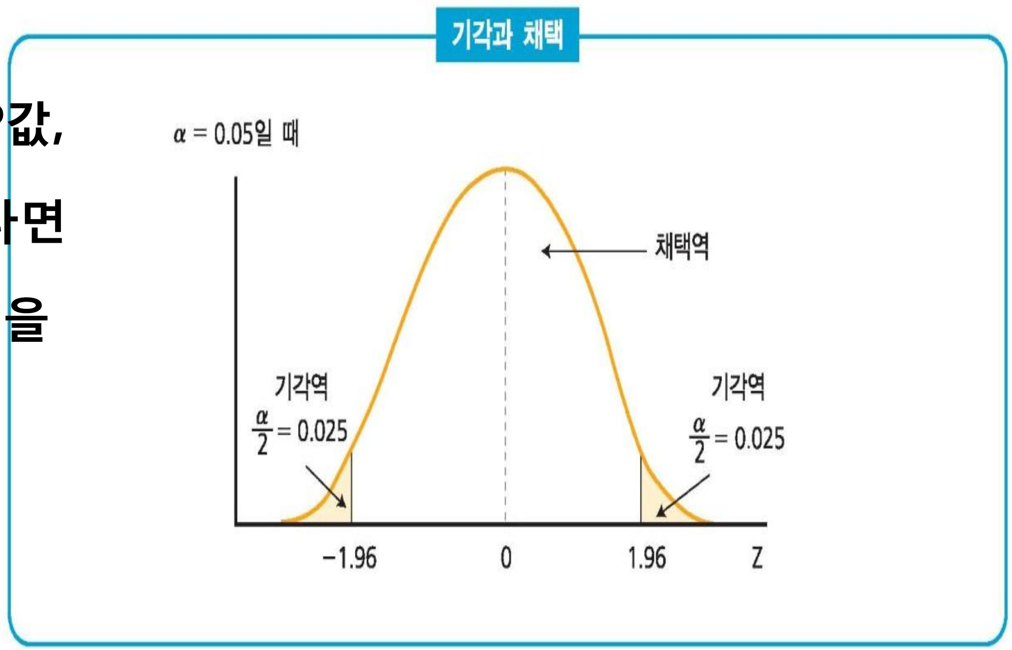
어느 모집단의 우리나라 1학년 성적 평균이 70점인가?

귀무가설 하나에 대립가설이 3개가 나올 수 있다.

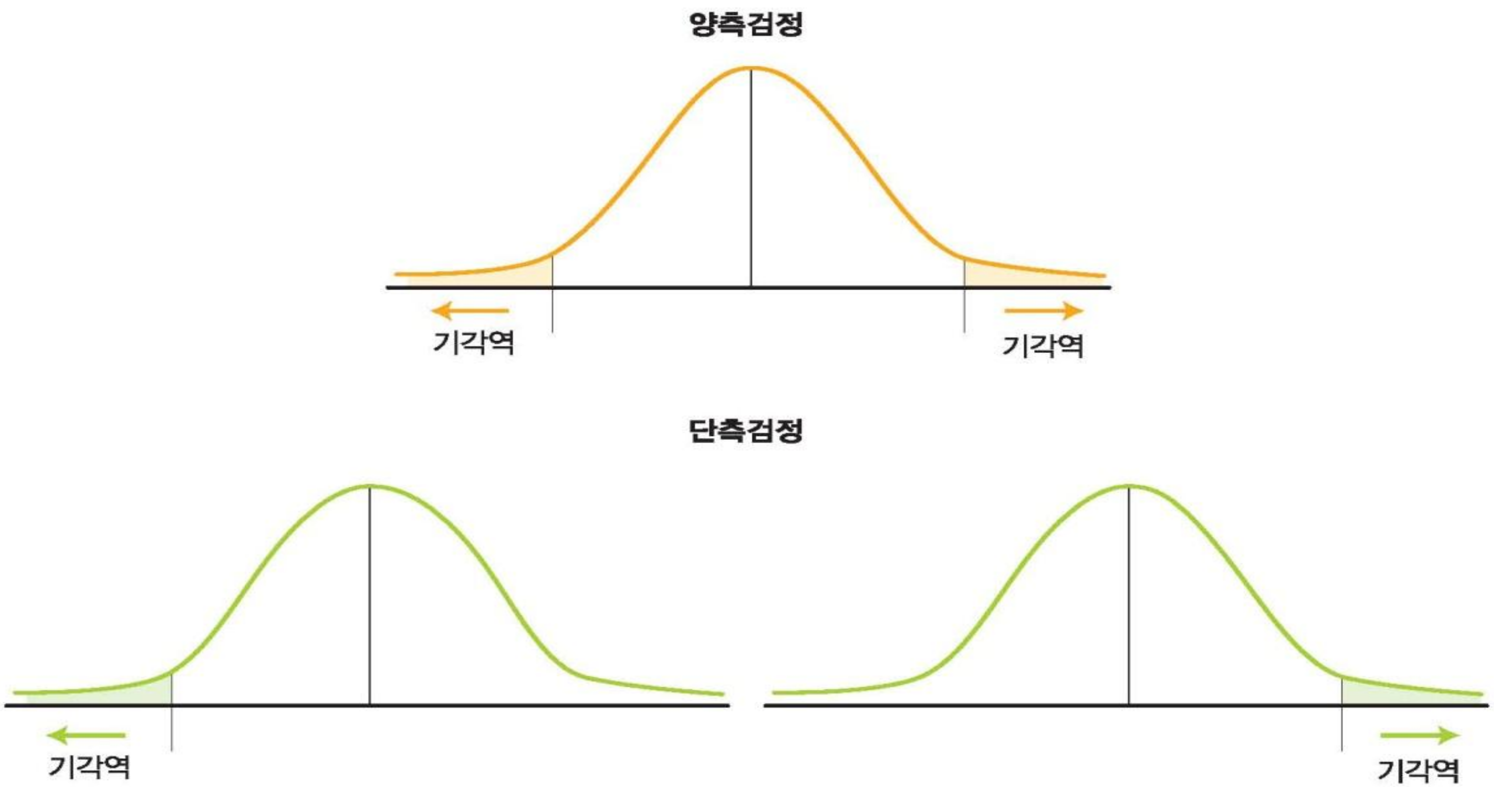
- H0 : μ = 70점 ; 평균이 70점이다.
- HA: μ ≠ 70점 ; 평균이 70점이 아니다. 양측검정
- HA: μ < 70점 ; 평균 성적보다 낮다. 단측검정
- HA: μ > 70점 ; 평균성적보다 높다. 단측검정

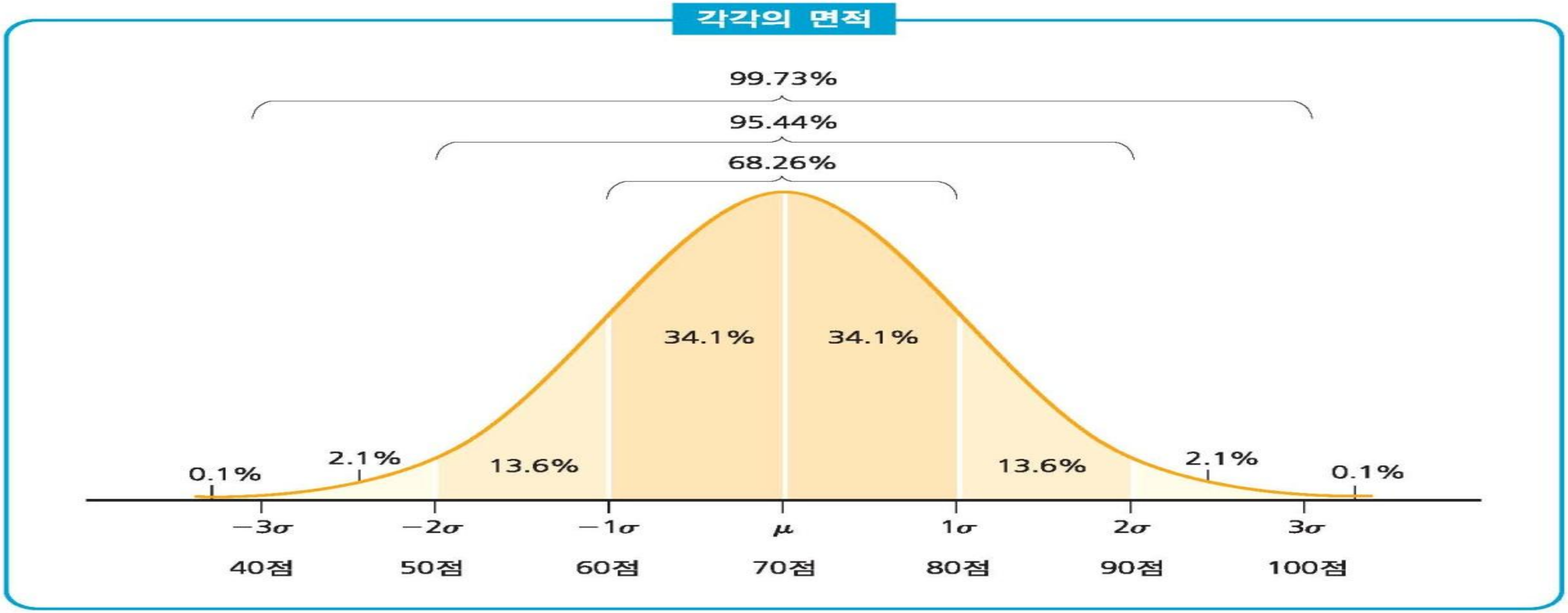
기각역 과 채택역

- 신뢰구간 95%영역은 모집단의 수가 포함될 확률이고 5%영역은 포함되지 않을 확률.
- 95%를 채택역이라 하고 5%영역을 기각역.
- 기각역은 귀무가설을 기각할 수 있는 통계량의 영역이고 채택역은 귀무가설을 채택할 수 있는 통계량의 영역을 말하며 그 경계를 임계역(유의 수준으로 결정)이라 한다.
- 유의수준을 0.05로 설정하면 P값, 유의확률을 보고 0.05보다 작다면 기각역에 포함되기에 귀무가설을 기각한다.



양측검정과 단측검정의 기각역과 채택역





전체 평균은 70점이고 표준편차는 10점이다. A집단은 평균점수가 80점이다.

이 집단은 전체의 평균점수와 가타고 볼 수 있는가?

신뢰수준 95%, 유의 수준(α)0.05이다.

- **H0 : μ(모집단 평균) = 70점 = 80점; 모집단과 A집단의 평균은 같다.**
- **HA: μ = 70점 ≠ 80점 ; 모집단과 A 집단의 평균은 다르다.**
- **HA: μ = 70점 < 80점 ; 모집단보다 A 집단의 평균이 높다.**

HA: μ = 70점 > 80점 ; 모집단보다 A 집단의 평균이 낮다.

연구) : 세 집단의 효과가 차이가 있는지에 대한 연구

- **H0 : 세 집단은 차이가 나지 않는다.**
- **HA : 세 집단은 차이가 난다.**
- **위의 가설을 설정하여 아래와 같은 결론을 얻었다.**

효과		값	F	가설 자유도	오차 자유도	유의확률
측정	Pillai의 트레이스	.328	8.607(a)	3.000	53.000	.000
	Wilks의 람다	**.672**	**8.607(a)**	**3.000**	**53.000**	**.000**
	Hotelling의 트레이스	.487	8.607(a)	3.000	53.000	.000
	Roy의 최대근	.487	8.607(a)	3.000	53.000	.000

유의 수준을 0.05로 정했다. 분산분석을 통해 위와 같은 연구결과를 얻었다.

4가지 결과에서 진한 부분의 결과를 해석하면 유의확률이 0.000이 나온다.

이 것은 0.05보다 작은 수치여서 위의 그림에서 0.000은 0.05보다 작은 면적에 위치함으로 기각역에 존재. 그래서 귀무가설을 기각하고 대립가설을 채택하여 세 집단간 차이가 난다고 할 수 있다.

- # 가설검정 방법

① 귀무가설 H0와 대립가설 H1를 설정한다. 양측검정인지 단측검정을 할 것인지 결정한다.

②유의수준 α(보통 실험연구에서는 α=0.05 또는 α=0.01)값을 설정하고 기각역과 채택역을 정한다.

③검정통계량을 계산한다.

④통계적 결정을 내린다.

⑤결론을 내린다.

계산된 유의 확률(p값)이 유의수준보다 작으면 귀무가설을 기각하고 대립가설을 채택하여 통계학적으로 유의한 결과를 얻었다고 하며, 계산된 유의확률(p)이 유의 수준보다 크면 귀무가설을 채택하고 대립가설을 기각하여 유의 수준 α에서 통계학적으로 유의하지 않다고 한다.