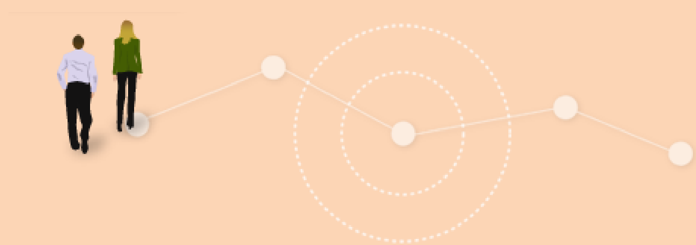
7 장

목 차

01 가설검정과 유의수준

:: Keywords 가설 | 유의수준 | 통계적 오류



가설검정

■ 가설(hypothesis)

주어진 사실 혹은 조사하고자 하는 사실이 어떠하다는 주장이나 추측

→ 모수를 추정할 때, 모수가 어떠하다(혹은 어떠할 것이다)는 조사자의 주장이나 추측을 일컬음

가설검정

■ 귀무가설(null hypothesis)

귀무가설(歸無假說, null hypothesis)은

일반적으로 믿어왔던 사실을 가설로 설정한 것으로,

영가설(零假說)이라고도 하며, 영(零, 0)이라는 의미로 H_0로 표기

→ 귀무가설에 대한 조사는 의미가 없다고 볼 수 있다.

∵ 연구를 하더라도 일반적으로 모두가 인정하고 받아들이는 사실이므로 어떤 의미를 찾아내기 어렵기 때문

Ex. 스포츠 이온음료의 용량이 제품에 표기된 300ml가 맞는지에 대한 조사

H_0 : 스포츠 이온음료의 용량은 제품에 표기된 300ml가 맞다.

가설검정

■ 대립 가설(antihypothesis)

대립가설(對立假說, antihypothesis)은 공공연하게 사실로 받아들여진 현상에 대립되는 가설로, 일반적으로 연구를 통한 대립가설의 조사는 의미가 있다고 받아들인다.

대립가설은 연구가설(硏究假說, research hypothesis)2이라고도 하며, 영(零, 0)에 반대가 된다는 의미로 H_1로 표기

Ex. 스포츠 이온음료의 용량이 제품에 표기된 300ml가 맞는지에 대한 조사

→ 귀무가설과 대립하여, 스포츠 이온음료의 용량이 300ml라고 표기된 것이 사실이 아니라고 설정하면 됨.

H_0 : 스포츠 이온음료의 용량은 제품에 표기된 300ml가 아니다.

양측검정과 단측검정

■ 기각의 판단 기준은 양측검정과 단측검정으로 구분

→ 조사 결과가 유의수준 α에 포함되면 기각, 포함되지 않으면 채택

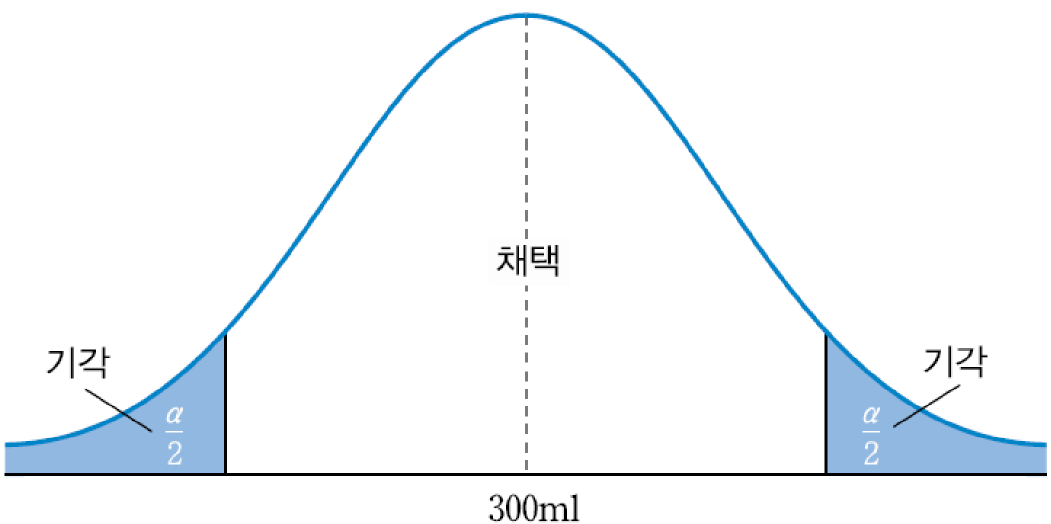
양측검정과 단측검정

■ 양측검정(two-sided test)

조사하고자 하는 대립가설, 즉 '사실이 아니다'라는 것을 검정하여 귀무가설을 기각하고 대립가설을 채택하고자 하는 것

$$H_0 : \mu = 300\,\mathrm{ml}$$

$$H_1 : \mu \neq 300\,\mathrm{ml}$$



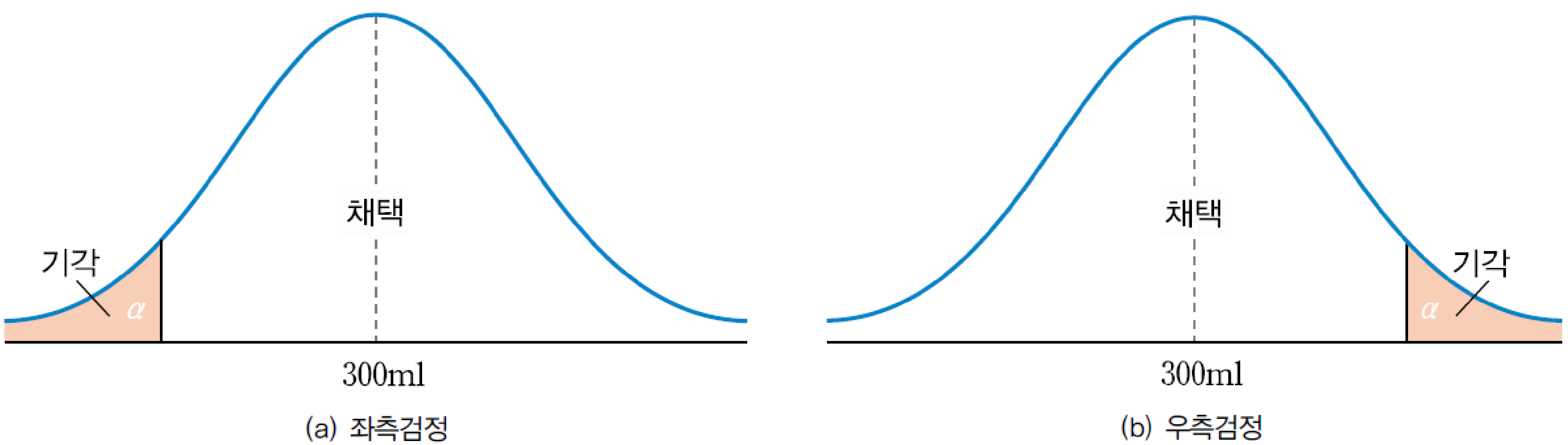
양측검정과 단측검정

■ 단측검정(one-sided test)

조사의 목적에 따라 대립가설을 스포츠 이온음료의 용량이 300ml보다 적다고 수립하거나, 혹은 300ml보다 많다고 수립하여 한 쪽만 살펴보는 것

$$H_0: \mu = 300\,\mathrm{ml} \qquad\qquad H_0: \mu = 300\,\mathrm{ml}$$

$$H_1: \mu < 300\,\mathrm{ml} \quad \text{혹은} \quad H_1: \mu > 300\,\mathrm{ml}$$



유의수준과 통계적 오류

■ 통계적 판단

→ 모수를 추정한다는 의미

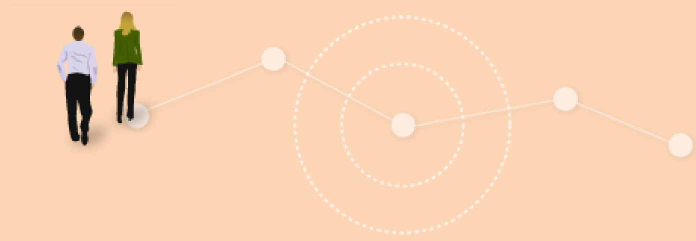
■ 추정은 틀릴 가능성을 내포

→ 모수를 추정할 때는 항상 오류 가능성(확률)을 제시

- 통계학에서는 모수의 추정이 맞을 확률을 1-α로 표시
  α는 유의수준(significance level)으로,
  확률(prpbability)로 표시되므로 약자를 사용하여 p값$(p - value)$으로 표시

02 모집단 평균의 가설검정

:: Keywords 모분산을 아는 경우의 가설검정
모분산을 모르는 경우의 가설검정
p값을 이용한 가설검정



모분산을 아는 경우의 가설검정

■ 가설을 수립

귀무가설을 모평균이 계산된 특정 값과 동일하므로

$$H_0 : \mu = \mu_0$$

대립가설은 양측검정의 경우 모평균이 특정 값과 동일하지 않으므로

$$H_1 : \mu \neq \mu_0$$

단측검정의 경우에는 좌측검정: $H_1 : \mu < \mu_0$

우측검정: $H_1 : \mu > \mu_0$

모분산을 아는 경우의 가설검정

■ 유의수준의 결정

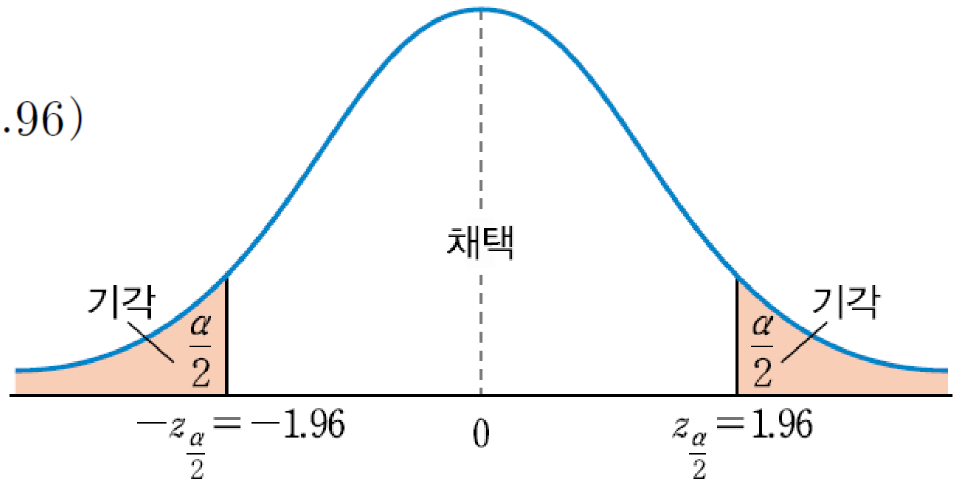
귀무가설과 대립가설에 대해 $\alpha = 0.05$에서의 가설검정

양측검정이면 $\pm z_{\frac{\alpha}{2}} = z_{0.025} = \pm 1.96$,

검정통계량은 $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

Z의 절대값이 1.96보다 크면 $(|z| > 1.96)$

귀무가설 H_0를 기각



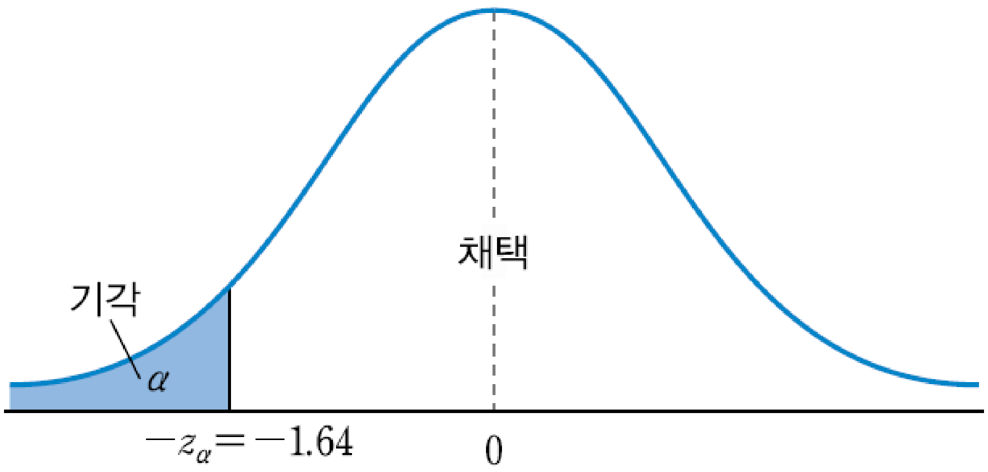
모분산을 아는 경우의 가설검정

■ 유의수준의 결정

귀무가설과 대립가설에 대해 $\alpha = 0.05$에서의 가설검정

좌측검정이면 $-z_\alpha = -z_{0.05} = -1.64$

검정통계량 z값이 -1.64보다 작으면 $(z < -1.64)$ 귀무가설 H_0를 기각



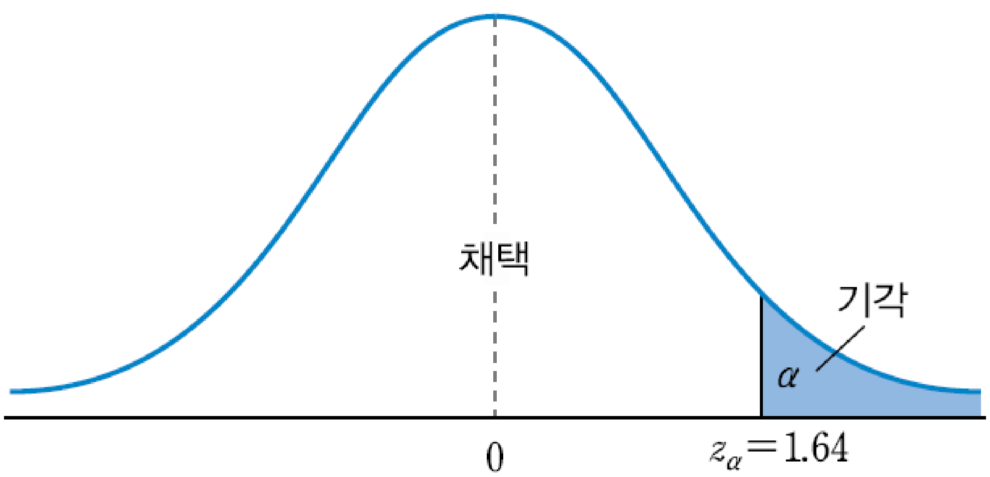
모분산을 아는 경우의 가설검정

■ 유의수준의 결정

귀무가설과 대립가설에 대해 $\boldsymbol{\alpha = 0.05}$에서의 가설검정

우측검정이면 $z_\alpha = z_{0.05} = 1.64$

검정통계량 z값이 -1.64보다 크면 $(z < 1.64)$ 귀무가설 H_0를 기각



모분산을 아는 경우의 가설검정

예제 7-1 **모분산을 아는 경우의 가설검정** 준비파일 | 7장_가설검정(모분산을 아는 경우).xlsx

스포츠 이온음료의 용량이 제품에 표시된 300ml보다 모자라는 것 같아서 직접 조사를 해보고자 한다. 표본 300개를 대상으로 용량을 측정한 결과, 평균이 244.65로 확인되었다. 표준편차가 20일 때, 가설을 수립하고 유의수준 0.05에서의 좌측검정을 실시하라.

모분산을 모르는 경우의 가설검정

■ 표본이 큰 경우의 가설검정

모수를 모르더라도 표본이 아주 큰 경우,

즉 $n \longrightarrow \infty$ 이면 모집단의 정규성과 상관없이 중심극한정리에 의해 s^2 은 σ^2 으로 수렴

그러므로 표본이 큰 경우의

검정통계량은 모분산이 주어진 것과 동일하게 $z = \dfrac{x - \mu}{\sigma / \sqrt{n}}$ 로 계산

■ 표본이 작은 경우의 가설검정

표본이 충분하지 못한 경우는

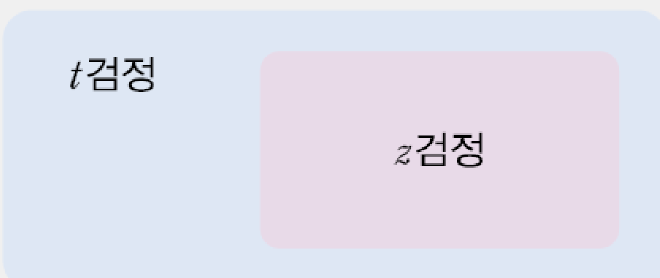
표본통계량 $t_{n-1} = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$ 는 자유도 $(n - 1)$ 에서 t분포를 따르게 된다.

모분산을 모르는 경우의 가설검정

Note **t검정과 z검정의 관계**

통계학을 학습할 때 검정 방법이 다양하여 혼란스러울 수 있다. t검정과 z검정의 관계만 보더라도 검정통계량을 구하는 공식은 비슷한데 도출 결과가 다르다. 이 두 가지 검정 방법은 중심극한정리에 의해 구분되는 것으로 판단하면 이해하기 쉽다.

즉 표본이 30보다 적으면 t분포를 이용하고, 그보다 많으면 z분포를 이용한다. 다만 t분포표를 보면 알 수 있듯이 t분포는 30보다 적은 표본의 분포를 나타내는 동시에 30보다 많은 표본의 분포도 포함한다. 다시 말해, z분포는 t분포에 포함된다고 할 수 있다.



[그림 7-9] **z검정과 t검정의 관계**

모분산을 모르는 경우의 가설검정

예제 7-2 모분산을 모르는 경우의 가설검정 준비파일 | 7장_가설검정(모분산을 모르는 경우).xlsx

[예제 7-1]의 스포츠 이온음료의 용량에 대한 조사를 다음의 표본 15개를 대상으로 살펴보고자 한다. 가설을 수립하고, 유의수준 0.05에서의 좌측검정을 실시하라. 단, 모집단의 분포는 정규분포라 가정한다.

번호	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
용량(ml)	308	302	290	292	327	290	320	285	315	285	295	288	310	325	300

■ p값을 이용한 가설검정

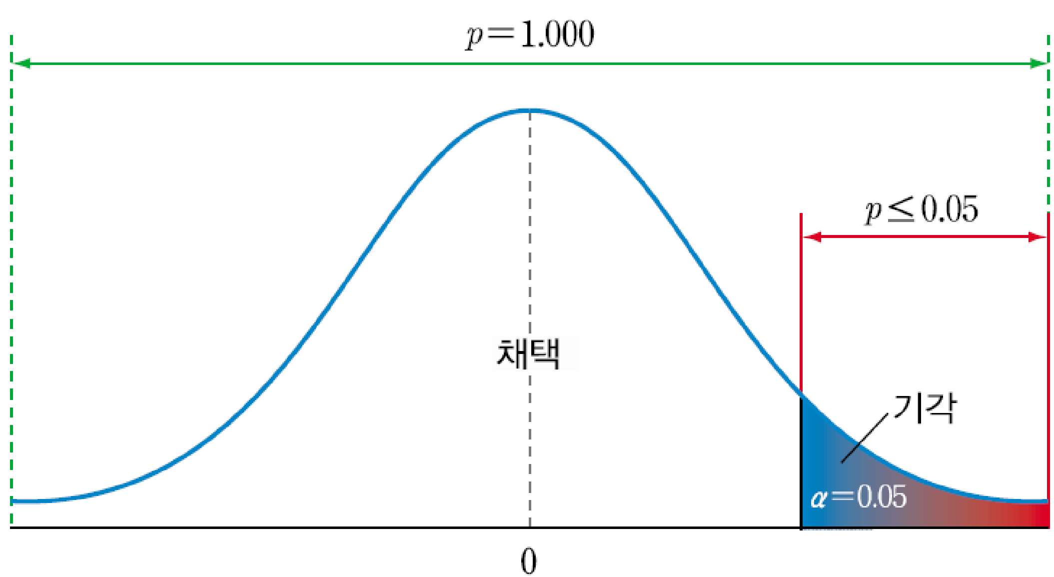
유의수준에 따라 채택/기각을 결정한 지금까지의 방법은 신뢰범위에 표함되는지 그렇지 않는지만을 제시하므로 채택/기각에 대한 강도를 표현하기에는 충분하지 않음

→ 이러한 단점을 보완하기 위해 채택/기각에 대한 기준을 확률 p로 나타내려는 방법

p값을 이용한 가설검정

■ P값을 이용한 가설검정

p값은 귀무가설을 기각하기 위한 최대한의 한계점을 나타내는데, 유의수준 α를 기준으로 보면 α로부터 멀리 떨어져 있는 확률을 나타낸다.



p값이 0.05보다 작거나 같다면 귀무가설을 기각

p값을 이용한 가설검정

예제 7-3 **p값을 이용한 가설검정** 준비파일 | 7장_가설검정(p값).xlsx

[예제 7-1]의 스포츠 이온음료의 용량에 대한 조사를 표본 300개를 대상으로 살펴보고자 한다. 용량을 직접 측정한 결과 평균이 294.65로 확인되었다. 표준편차가 45일 때, 가설을 수립하고, 유의수준이 0.05에서의 귀무가설의 기각 여부를 판단하라.