사회통계연습
임계값과 유의확률

김현우, PhD[1]

[1] 충북대학교 사회학과 조교수

October 15, 2021

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

진행 순서

1. 지난 주 리뷰
2. 가설검정의 기본 논리
3. 임계값을 이용한 가설검정
4. 유의확률을 이용한 가설검정

지난 주 리뷰

지난 주 리뷰

퀴즈 #6 코멘트

- **문제 1.** 채림은 파이어족에 합류하기 위해 먼저 증권거래소에 상장된 수많은 주식회사의 작년 한 해 주가변동에 관한 데이터를 수집하였다. 그 결과 연평균 상승폭이 8,000원이고 표준편차는 10,000원 임을 확인하였다. 채림은 이들 중 30개의 주식을 랜덤하게 구입하고 1년을 기다렸다. 채림의 주식 가운데 주당 평균 11,000원 이상 가격상승이 있었던 주식은 얼마나 될까?
- NORM.DIST() 함수를 사용할때 자꾸 FALSE와 TRUE를 혼동하는 경우가 있었다. 차이를 기억할 것. 똑같은 잘못을 2번 문제에서도 저지르는 경우가 있었다.
- 표준화할 때, 분모가 σ만 들어가는게 아니라 $\sigma/\sqrt{n}$ 임을 잊으면 안된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

지난 주 리뷰

- **문제 3.** 서현 메뉴펙쳐링은 압도적 품질과 뛰어난 고객서비스로 국제적인 명성을 얻고 있다. 이 회사의 사장 서현이 불시에 네오청주 공장을 방문하였다. 지난 한 달 동안 제조된 아이스크림 껍질 중에서 200개를 뽑아서 두께를 재어보니 평균이 0.824cm이고 표준편차는 0.042cm였다. 서현은 표준편차가 이렇게 크다는 사실에 극대노하며 부사장 정후를 불렀다. "이봐, 내 표본을 가지고 전체 아이스크림 껍데기 두께의 평균을 추정할 수 있겠나? 자네도 알다시피 나는 관대하다네. 99.99%까지 신뢰구간까지 봐줄 의향이 있어." 정후는 자리로 돌아와 답을 쓸지 사표를 쓸지 머리를 싸매고 고민하고 있다. 정후를 위한 답은 무엇인가? (사표말고)
- 어떤 학생은 과반수 이상이라는 점에 집착해서 확률변수를 0.51부터 쟀다: $P(p > .51) = 1 - P(p < .51)$. 이렇게 하면 답이 꽤 크게 틀리는데 $p = .5$ 부근인 50%와 51% 사이에는 꽤 큰 확률들이 들어가 있기 때문이다(이항분포의 원리). 단순히 $P(p > .5) = 1 - P(p < .5)$를 보고하면 된다. 우리는 설령 비율 문제라도 그 표본평균의 표집분포가 (이산확률분포가 아닌) 연속확률분포라고 상정함에 주의할 것.
- 99.99% 신뢰구간 중 가장 확률이 높은 구간은 **NORM.INV(0.9999, 0, 1)**와 **NORM.INV(0.0001, 0, 1)**이 아니다. 반씩 나눠가져야 하므로 1-.9999=.0001의 반인 .0001/2, 즉 .00005다. 그러므로 **NORM.INV(0.99995, 0, 1)**와 **NORM.INV(0.00005, 0, 1)**가 맞다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

지난 주 리뷰

퀴즈 총평

- 대부분 모두 사소한 실수에 지나지 않았다! 변별력을 위해서라도 중간시험의 난이도를 올려야 겠다는 결심을 했다.
- 계산에 사용한 함수를 지우지 말 것. 답만 엑셀에 써놓지 마시오. 지우면 틀릴때 부분점수마저 잃게 되어 자기만 불이익이다(시험도 마찬가지).
- 숙제에서 cell reference를 적극 활용하자. 여러 함수를 연이어 사용할때 예전에 구한 답을 곧바로 reference하면 소숫점에 집착할 필요가 없어진다.
- 숙제 답안은 엑셀 파일 하나로만 주세요. 숙제 안에도 이름은 써주세요. 숙제 파일 이름에 주차와 학번은 써주세요(e.g., 3주차 20201234.xlsx). 도대체 이 말 몇 번째냐...

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

가설검정의 기본 논리

오늘 수업까지 포함하여 3주 동안 추리통계학의 기초를 다루게 되었다.

- 이 부분은 학부 수준 사회통계학의 중급 레벨로 나아가기 위한 기초로 사실 가장 어렵고 지루한 파트라고 할 수 있다.
- 5주차에서 "확률과 확률분포"로 추리통계의 기초를 배웠다.
- 6주차에서 "표집분포와 추정"으로 추리통계의 기초 나머지와 더불어 첫번째 추리통계 기법인 추정(estimation)을 배웠다.
- 7주차가 오늘 할 부분인데 바로 "가설검정(hypothesis test)"이 그 주제이다.

통계적 추리(statistical inference)에는 "추정"과 "가설검정"이 있다고 언급하였다.

- 가설검정 역시 추정과 사실상 같은 아이디어인데 다만 접근의 순서가 다른 탓에 다소 혼동스러울 수 있다.
- 중요한 것은 논리의 흐름, 그 다음이 계산의 정확성이다. 복습할 때는 논리의 흐름을 먼저 명확하게 이해해야 한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

먼저 영가설과 대립가설을 바르게 세울 수 있어야 한다!

- 우리는 연구에 앞서 모집단의 평균(population mean)에 대해 영가설(null hypothesis)을 세운다. 예컨대 (-5에서 5사이로 측정된 만족도 점수에서) "네오청주의 평균 거주만족도는 0이다" 또는 $H_0 : \mu = 0$.
- 여기서 주목해야 하는 부분은 내가 지금 세운 영가설이 "표본에 관한 것($\bar{X}$)"이 아니라 "모집단에 관한 것(μ)"이라는 점이다.
- 대립가설(alternative hypothesis)은 영가설을 기각하면 받아들이게 되는 가설이다. 예컨대, "네오청주의 평균 거주만족도는 0이 아니다" 또는 $H_a : \mu \neq 0$.
- 영가설과 대립가설 사이에는 놓치는 가능성이 없어야 한다. 예컨대 "네오청주의 평균 거주만족도는 0이다"라고 영가설을 세웠는데 대립가설은 "네오청주의 평균 거주만족도는 0보다 적을 것이다"라면 틀린 가설 설정이 된다(Why?)

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

가설검정 방식에는 양측검정(two-tailed test)과 단측검정(one-tailed test) 두 가지가 있다.

- 양측검정에서 영가설과 대립가설의 구조는 다음과 같다.

$$H_0 : \mu = x, H_a : \mu \neq x$$

- 이와는 달리 단측검정은 두 가지 형태 중 하나의 영가설과 대립가설의 구조를 갖는다.

$$H_0 : \mu \geq x, H_a : \mu < x$$

또는

$$H_0 : \mu \leq x, H_a : \mu > x$$

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

이름이 이미 암시하고 있듯, 영가설은 보통 기각(reject)하기 위한 목적으로 세워진다.

- 영가설의 다른 이름인 귀무가설(歸無假說)은 "무로 되돌린다"는 의미를 가지고 있다.
- 예컨대 어느 방향제 제조회사에서 생산하는 방향제의 유효시간은 평균 75일이였다. 이 회사 소속 조향사는 자신이 개발한 새로운 첨가제를 넣으면 평균 유효시간이 늘어날 것이라고 주장하고 있다.
- 이 경우 이 회사에서 수행되는 통계분석의 영가설은 $\mathrm{H}_0 : \mu \leq 75$이고 대체가설은 $\mathrm{H}_a : \mu > 75$이다.
- 내심 의도하고 있는 것은 대체가설 쪽이고, 오히려 영가설은 짐짓 틀리면 좋겠는데 하는 의도를 담고 있다. 그러므로 영가설을 기각하지 못한다면 연구자 입장에서는 아쉬운 일이다.
- 다만 앞서 설명하였듯, 양측검정의 경우에는 영가설이 $\mathrm{H}_0 : \mu = \mathrm{x}$이고 $\mathrm{H}_a : \mu \neq \mathrm{x}$인 구조가 그대로 유지된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

가설검정(hypothesis test)이란 결국 영가설(null hypothesis)이 표본에 의해 지지되는가를 테스트(test)하는 것이다.

- "네오청주의 평균 거주만족도는 0이다" 또는 $H_0 : \mu = 0$ 라는 가설을 세웠다. 모집단으로부터 충분히 큰($n > 30$) 표본을 랜덤하게 추출하였다고 하자.
- 만일 이 표본에서 구한 평균이 영가설에서 설정된 모집단의 속성(e.g., 네오청주의 평균 거주만족도)과 비슷하다면(e.g., $\bar{X} = 0.1$) 이는 무엇을 시사할까? "아마도 영가설로 세운 모집단의 속성은 표본에 의해서 지지되므로 영가설은 옳다" 라고 해야 할 것이다.
- 또는 (영가설은 보통 기각하기 위한 목적으로 세워지기 때문에) "아마도 영가설을 기각하는데 실패했다" 라고 할 수 있다.
- 반대로, 만일 이 표본에서 구한 평균이 영가설에서 설정된 모집단의 속성과 크게 동떨어져 있다면(e.g., $\bar{X} = -3$) 무엇을 시사할까? "아마도 영가설에서 설정된 모집단의 속성은 표본에 의해 지지되지 않으므로 영가설이 틀리다." 라고 해야 할 것이다.
- 또는 "아마도 영가설을 기각했다" 라고 할 수 있다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

왜 "아마도" 라는 단서가 계속해서 붙어있나?

- 왜냐하면 설령 표본의 평균($\bar{X} = 0$)이 영가설로 세워진 모집단의 평균($\mu = 0$)과 정확히 일치하더라도, 그런 표본이 (영가설과 전혀 다른 모집단으로부터) "우연히" 뽑혔을 가능성을 완전히 배제할 수는 없기 때문이다.
- 다만 표본의 평균이 영가설 평균과 가까우면 가까울수록 아무래도 영가설을 지지한다고 볼 수 있다. "그런 값이 뽑힐 확률이 정말로 높긴 높기 때문이다!" 반대로 표본의 평균이 영가설 평균과 다르면 다를수록 영가설은 점점 의심받게 되어 어느 순간 기각되어야만 한다.

이것이 오늘 수업의 가장 중요한 논리적 핵심이다.

- 표본의 속성이 영가설에서 제시한 속성과 제법 가깝다면 "(1) 영가설이 옳다는 가정 아래 (2) 그런 표본이 나올 확률이 제법 높으니 (3) 영가설이 틀리지 않음"을 의미한다.
- 반대로 표본의 속성이 영가설에서 제시한 속성으로부터 너무 동떨어져 있다면 "(1) 영가설이 옳다는 가정 아래 (2) 그런 표본이 나올 확률이 너무 낮으니 (3) 영가설이 틀림"을 의미한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

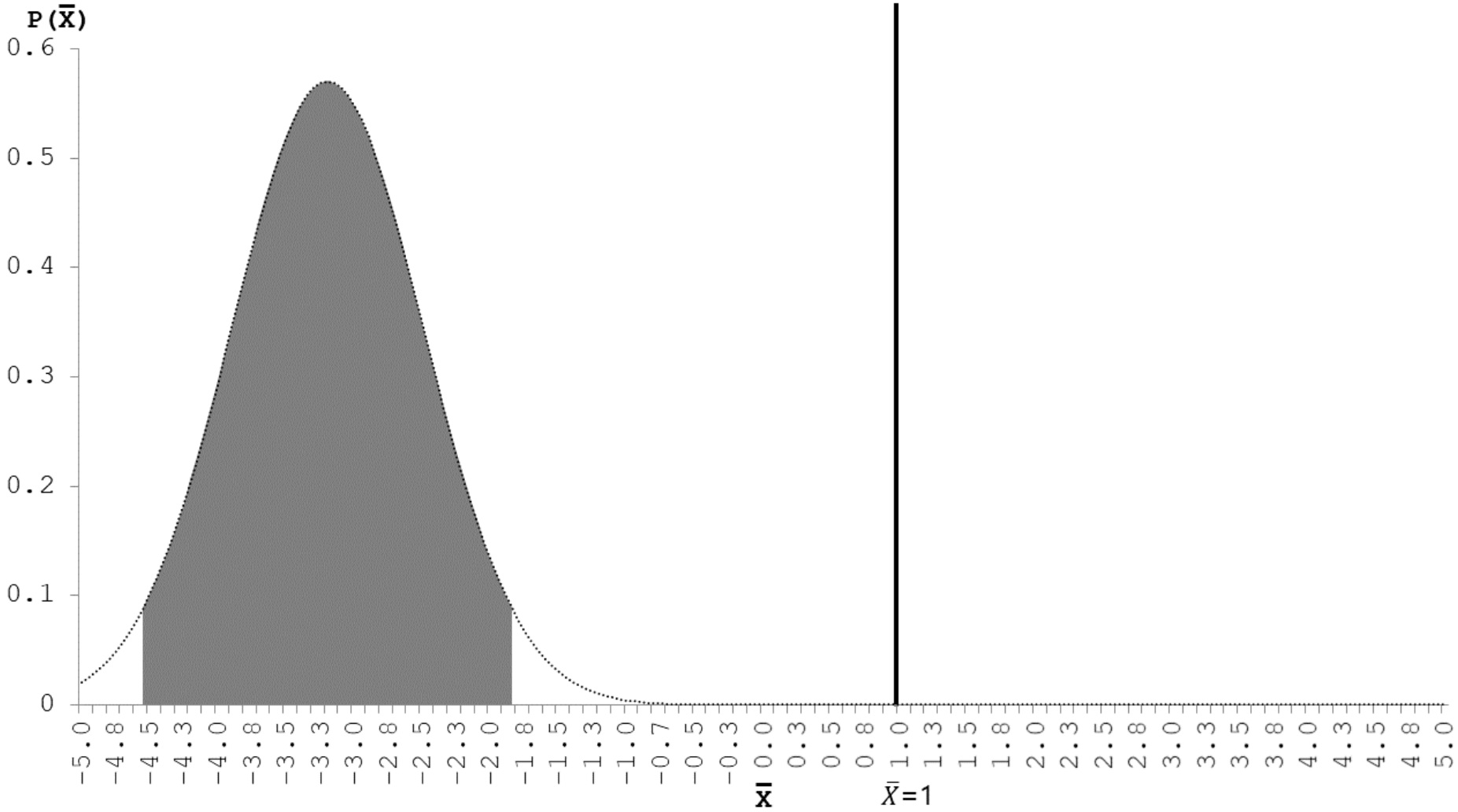
가설검정의 기본 논리

유의성 수준(significance level)을 통해 "아마도" 라는 단서를 떼어낸다.

- 네오청주시 시민을 대상으로 랜덤하게 표본을 통해 확인한 결과, $\bar{X} = 1$을 얻었다.
- 만일 시예진이 $\mu = -3.2$라고 영가설을 세웠다면, $\mu = -3.2$인 모집단에서 무한히 많은 표본평균들을 추출해내고 다시 표본평균들로 이 영가설에 따른 (가상적인) 표집분포를 그릴 수 있다.
- 확인해보면 표본평균($\bar{X} = 1$)은 영가설로 세운 (가상적인) 표집분포의 95% 신뢰구간 밖에 놓여있다.
- 이것을 해석해보자면, (1) 영가설인 $\mu = -3.2$가 옳다고 전제했을 때 (2) 95%의 확률로 나와주었어야 할 평균들인 [-4.55, -1.85] 바깥에서 표본의 평균 1이 발견되었으므로, (3) 시예진의 영가설은 95% 신뢰수준에서 기각된다.
- 그리고 -4.55와 -1.85는 각각 좌우의 임계값(critical value)이 되며, 이 값보다 크냐 작으냐에 따라 영가설이 유지되기도 하고 기각되기도 한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리



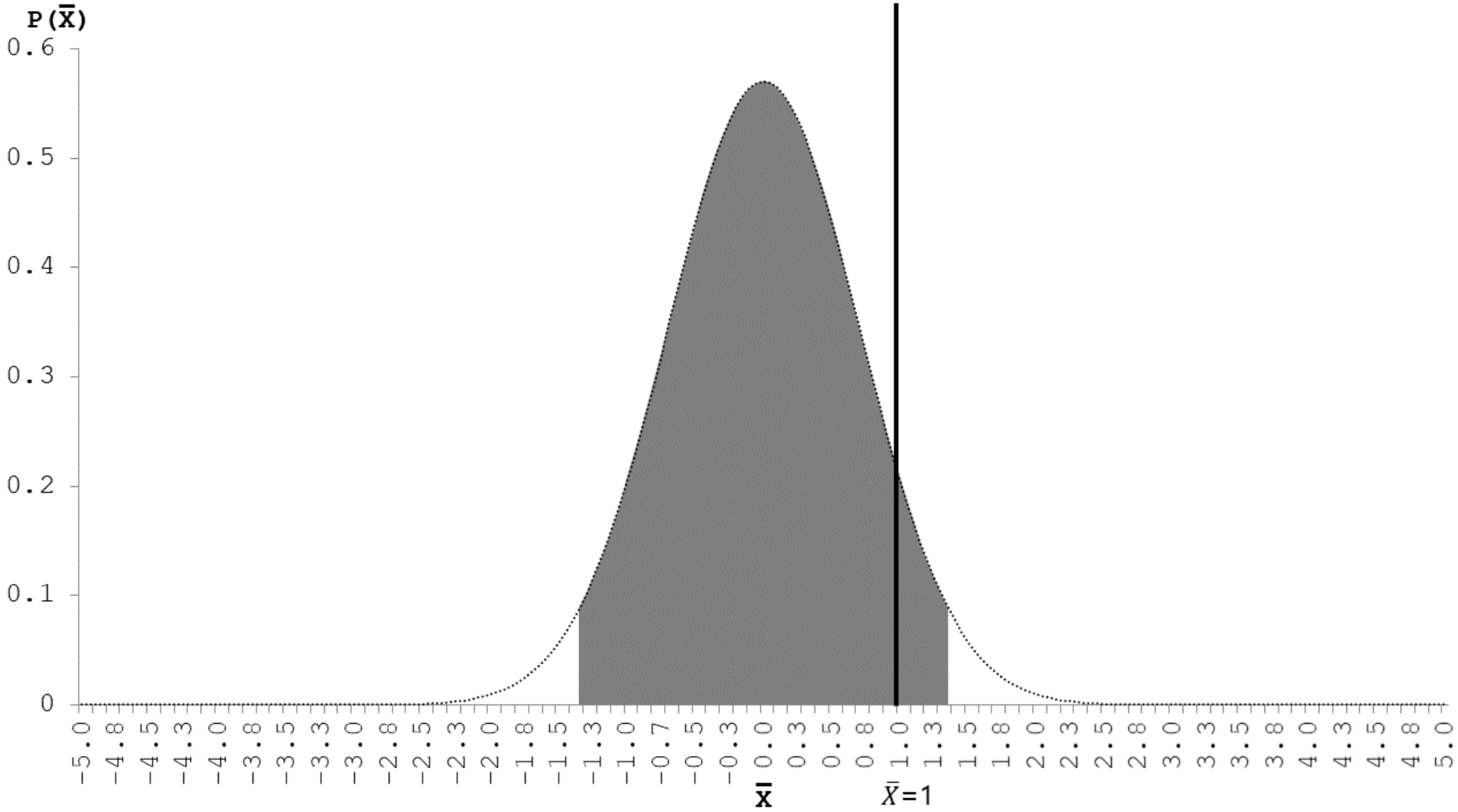
충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

- 만일 현서 $\mu = 0$이라고 영가설을 세웠다면, $\mu = 0$인 모집단에서 무한히 많은 표본평균들을 추출해내고 다시 표본평균들로 이 영가설에 따른 (가상적인) 표집분포를 그릴 수 있다.
- 확인해보면 표본평균($\bar{X} = 1$)은 영가설로 세운 (가상적인) 표집분포의 95% 신뢰구간 안에 놓여있다.
- 이것을 해석해보자면, (1) 영가설인 $\mu = 0$이 옳다고 전제했을 때 (2) 95%의 확률로 나와주었어야 할 평균들인 [-1.35, 1.35] 안쪽에서 표본의 평균 1이 발견되었으므로, (3) 현서의 영가설은 95% 신뢰수준에서 기각될 수 없다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리



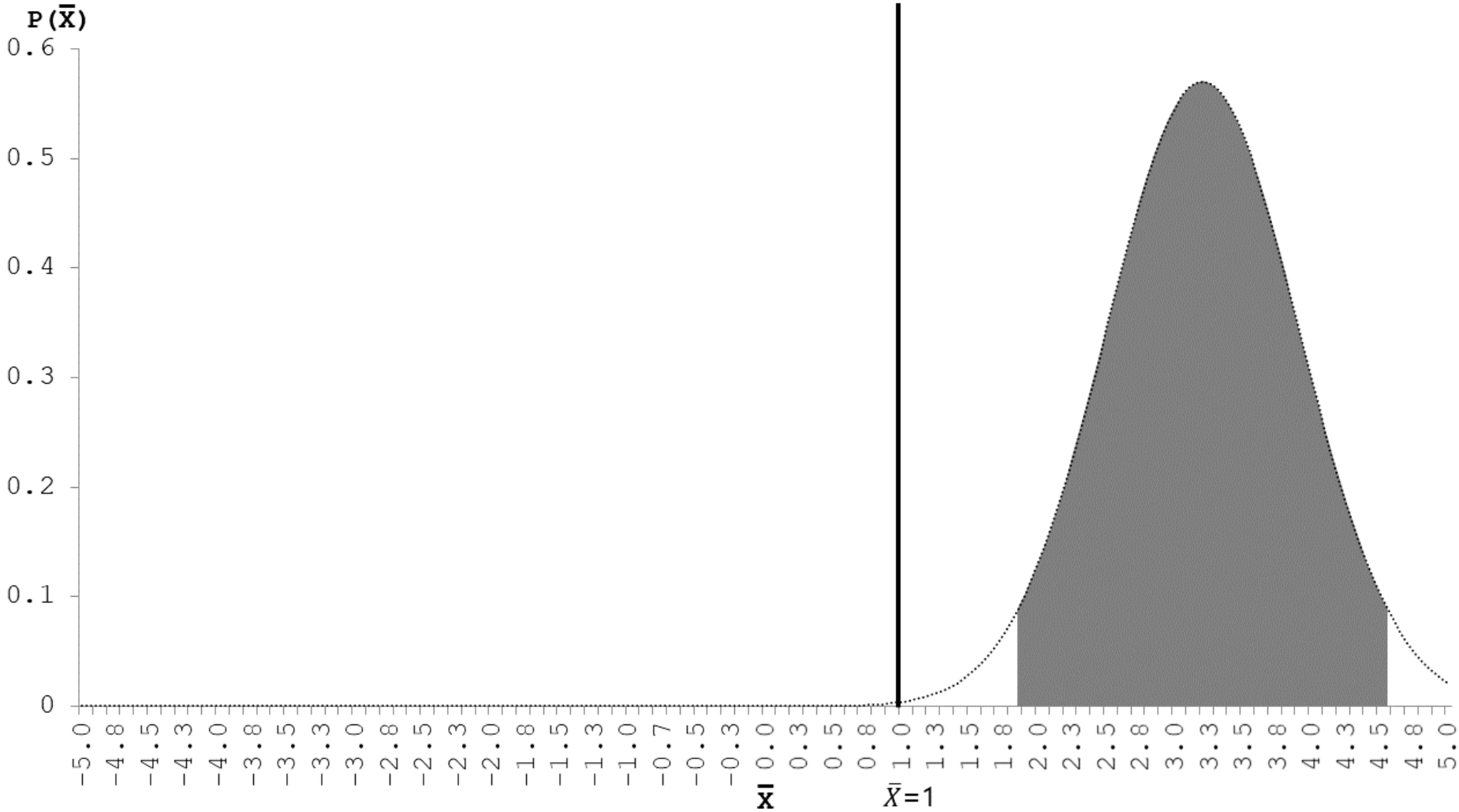
충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리

- 이번에는 명보가 $\mu = 3.2$라고 영가설을 세웠다면, $\mu = 3.2$인 모집단에서 무한히 많은 표본평균들을 추출해내고 다시 표본평균들로 이 영가설에 따른 (가상적인) 표집분포를 그릴 수 있다.
- 확인해보면 표본평균($\bar{X} = 1$)은 영가설로 세운 (가상적인) 표집분포의 95% 신뢰구간 밖에 놓여있다.
- 이것을 해석해보자면, (1) 영가설인 $\mu = 0$이 옳다고 전제했을 때 (2) 95%의 확률로 나와주었어야 할 평균들인 [1.85, 4.55] 바깥에서 표본의 평균 1이 발견되었으므로, (3) 명보의 영가설은 95% 신뢰수준에서 기각된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정의 기본 논리



충북대학교
CHUNGBUK NATIONAL UNIVERSITY

임계값을 이용한 가설검정

임계값을 이용한 가설검정

신뢰구간 바깥에 표본의 평균이 놓인다면 영가설을 기각한다.

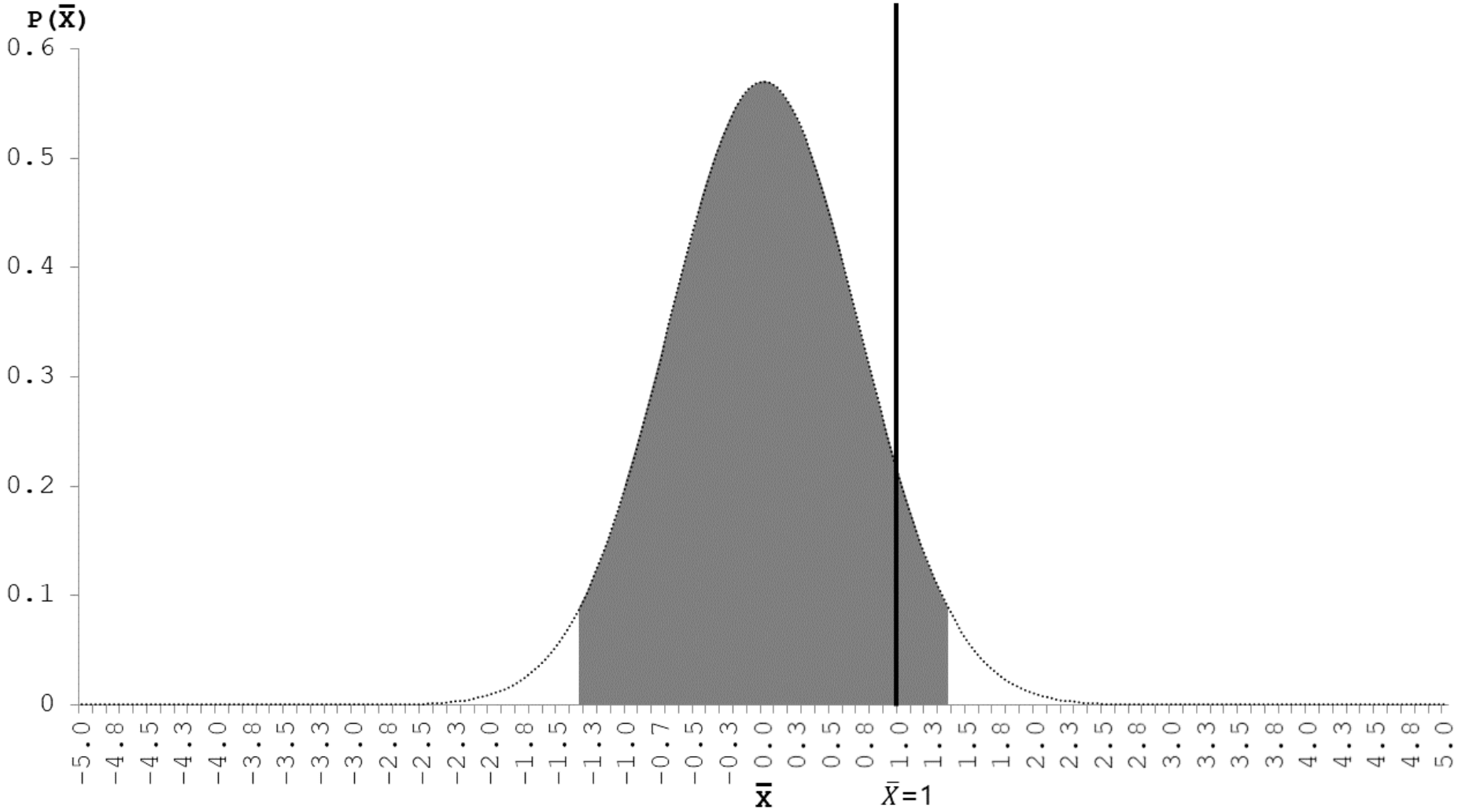
- (아까와는 조금 달리) 이번엔 원점수의 정규분포를 Z-점수로 표준화하여 표준정규분포(standard normal distribution)를 만들고 "가장 확률이 높은 부분으로만" "곡선 아래 95% 면적"을 색칠해보자.
- 그 구간은 표준정규분포에서 $Z = [1.96, -1.96]$에 대응하는데 이는 엑셀에서 **NORM.INV(0.025, 0, 1)**와 **NORM.INV(0.975, 0, 1)**를 통해 알 수 있다. 이때 **NORM.INV(0.95, 0, 1)**가 아님에 주의할 것(Why?).
- (Z-점수로 나타낸) 표준정규분포에서 [1.96, -1.96] 사이의 "곡선 아래 95% 면적"은 수식으로 이렇게 표현된다.

$$P(1.96 \leq Z < -1.96) = P(Z \leq 1.96) - P(Z \leq -1.96) = 0.95$$

- 다시 말해, 이 "곡선 아래 95% 면적"에 대응하는 Z값들은 영가설이 옳다면 "(표본에서) 95%의 확률로 나와줄 수 있는 값들"을 보여준다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

임계값을 이용한 가설검정



충북대학교
CHUNGBUK NATIONAL UNIVERSITY

임계값을 이용한 가설검정

가설검정에는 조금 특별한 전문용어가 사용된다.

- 만일 표본평균의 표준점수 Z가 (가설적으로 그려진 표준정규분포)의 95% "곡선 아래 면적" 바깥에 위치해 있다면 "95%의 신뢰수준(=5%의 유의수준)에서 통계적으로 유의(statistically significant)하게 영가설을 기각했다" 라고 말한다.
- 반대로 표본평균의 표준점수 Z가 (가설적으로 그려진 표준정규분포)의 95% "곡선 아래 면적" 안에 위치해 있다면 "95%의 신뢰수준(=5%의 유의수준)에서 통계적으로 유의하게 영가설을 기각하는데 실패했다" 라고 말한다.

신뢰수준은 반드시 95%여야 하는 것은 아니고 90%나 99% 혹은 다른 것일수도 있다.

- 하지만 관습에 따라 흔히 90%, 95%, 99%를 많이 본다.
- 당연히 신뢰수준을 높일수록 가설 기각은 어려워지므로 추정 결과를 받아들일 때는 보수적이 된다(Why?).

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

임계값을 이용한 가설검정

예제 1. 재범은 네오청주의 C대 재학생의 월 지출액 평균이 뭐 대충 100만원일 것이라는 영가설을 세웠다. 그는 열심히 발로 뛰어 C대 재학생으로 구성된 100명의 임의 표본을 확보하였다. 표본에서 밝혀진 평균은 97만원이고 표준편차는 10만원이였다. 재범의 영가설을 95% 신뢰수준에서 테스트하라.

- 재범의 영가설은 $\mathrm{H}_0 : \mu = 100$ 이다. 대립가설은 $\mathrm{H}_a : \mu \neq 100$이다.
- 평균이 100만원인 모집단에서 무한히 많은 표본평균들을 구해 (가상적인) 표집분포를 그린다면, 그것의 표준오차(standard error)는 $\sigma/\sqrt{\mathrm{N}} = 10/\sqrt{100} = 1$이다.
- 이제 표본평균을 표준화하면 다음과 같다:

$$\mathrm{Z} = \frac{\bar{\mathrm{X}} - \mu}{\mathrm{S}/\sqrt{\mathrm{n}}} = \frac{97 - 100}{10/\sqrt{100}} = -3$$

주의! 여기서 $\bar{\mathrm{X}}$와 μ에 넣을 값들을 서로 혼동해서는 안된다!

충북대학교 CHUNGBUK NATIONAL UNIVERSITY

임계값을 이용한 가설검정

- 표준화된 95% 신뢰구간 또는 채택영역(acceptance area)은 엑셀에서 간단히 계산할 수 있다. 좌측 끄트머리 Z값은 **NORM.INV(0.025, 0, 1)**이고 우측 끄트머리 Z값은 **NORM.INV(0.975, 0, 1)**이다. 각각 [-1.96, 1.96]이다.
- ± 1.96이 바로 좌우의 임계값이 된다. 물론 두 임계값 사이에 채택영역이 존재하고, 그 바깥은 기각영역(rejection area)라고 부른다.
- 만일 영가설이 옳다면 채택영역 [-1.96, 1.96]안에 뽑힐 수 있는 표본평균의 95%가 들어있어야 한다!
- 명백히 표본평균의 Z값인 -3은 95% 채택영역(± 1.96) 바깥에 놓여있다. 다시 말해, 재범의 "영가설은 95% 신뢰수준에서 통계적으로 유의하게 기각된다."

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

유의확률을 이용한 가설검정

다른 접근방식을 통해 가설검정을 할 수도 있다.

- 앞서 임계값 또는 신뢰구간 개념을 통한 가설검정을 설명했지만, 그와는 달리 유의확률(p-value) 개념을 통해 가설검정을 할 수도 있다.
- 핵심은 이것이다: 만일 (1) "영가설이 옳을 확률"이 (2) "충분히 작다면" 확신을 가지고 영가설을 기각할 수 있다.
- 여기 (1) "영가설이 옳은데 이를 기각할 확률"을 유의확률(p-value)이라고 한다.
- 유의확률이 (2) "충분히 작다"는 것은 우리가 옳은 영가설을 버릴 위험이 낮다는 의미로 해석된다.
- 충분히 작다는 것은 얼마만큼 작아야 하나? 관례적으로 10%, 5%, 1% 유의수준(significance level) 셋 중 하나를 많이 쓴다. 각각에 대응한 것이 90%, 95%, 99% 신뢰수준(confidence level)이다.
- 1에서 유의수준을 빼면 신뢰수준이 된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

유의확률(p-value)이 충분히 작으면 자신있게 영가설을 기각한다.

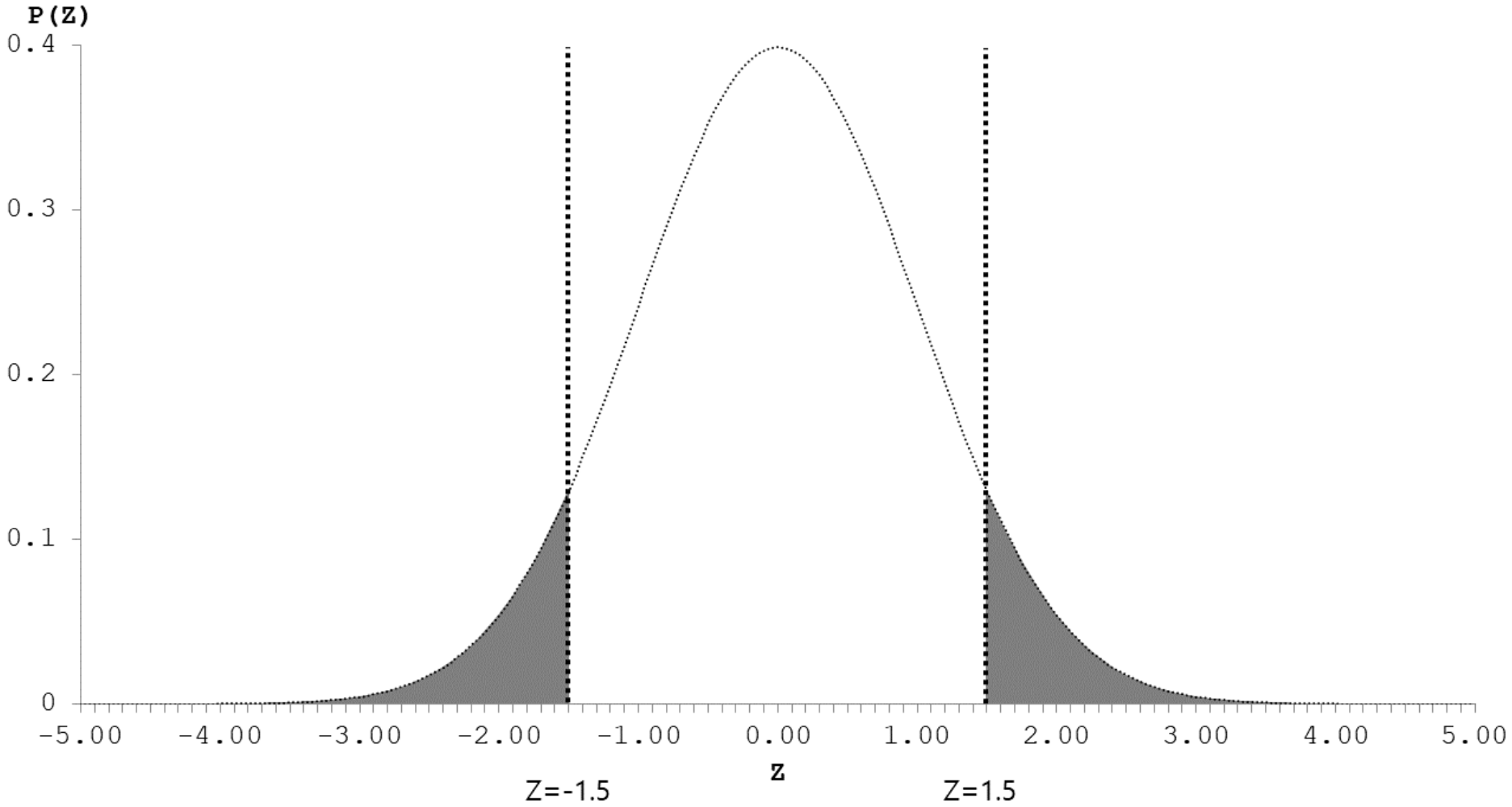
- **예제 1**로 되돌아가자($\bar{X} = 97, S = 10$). 이번엔 태준이 새로운 영가설 $H_0 : \mu = 98.5$를 제시하였다.
- 태준의 영가설 모평균을 표준화하면 다음과 같다:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} = \frac{97 - 98.5}{10/\sqrt{100}} = -1.5$$

- 아까 계산한 바에 따르면 표준화된 95% 신뢰구간, 즉 채택영역은 [-1.96, 1.96] 였다. 이때, -1.96보다 작은 Z값들과 1.96보다 큰 Z값들로 채워지는 나머지 5%도 여전히 (낮을지언정) 영가설이 옳을 확률을 보여주기는 한다.
- 이와 유사한 논리로 태준이 영가설로 제시한 표준화된 모평균 ± 1.5 바깥의 좌우측 끄트머리 면적도 영가설이 옳을 확률을 보여준다(Why?).
- 왜냐하면 이는 태준의 영가설이 다름아닌 $\mu = 98.5$ 였기 때문이다. 이 경우 표본의 평균이 0보다 지나치게 "크지도 작지도 않음"을 검정해야 하므로(지나치게 커도 기각되고 작아도 기각된다!), 이른바 양측검정(two-tailed test)이 요구된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정



충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

- 태준이 영가설로 제시한 표준화된 모평균 -1.5보다 작은 값들의 면적은 다음과 같이 엑셀 함수로 계산된다: NORM.DIST(-1.5, 0, 1, TRUE)=0.067
- (표준화된 모평균 -1.5과 중심축과의 거리를 거울반사한) Z값 1.5보다 큰 값들의 면적은 다음과 같이 엑셀 함수로 계산된다: 1-NORM.DIST(1.5, 0, 1, TRUE)=0.067
- 두 면적을 합한 값(=0.134)은 태준이 제시한 영가설이 옳을 확률을 보여준다.
- 이것은 0.05 유의수준, 즉 95% 신뢰수준보다 큰 폭이다. 그러므로 태준은 "5% 유의수준에서 또는 95% 신뢰수준에서 통계적으로 유의하게 영가설을 기각할 수 없다."
- 실제 의미를 해석해보면 "영가설이 옳은데 기각할 확률"이 무려 13.4%에 달하므로 영가설을 자신있게 기각하기에는 위험이 너무 컸던 것이다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

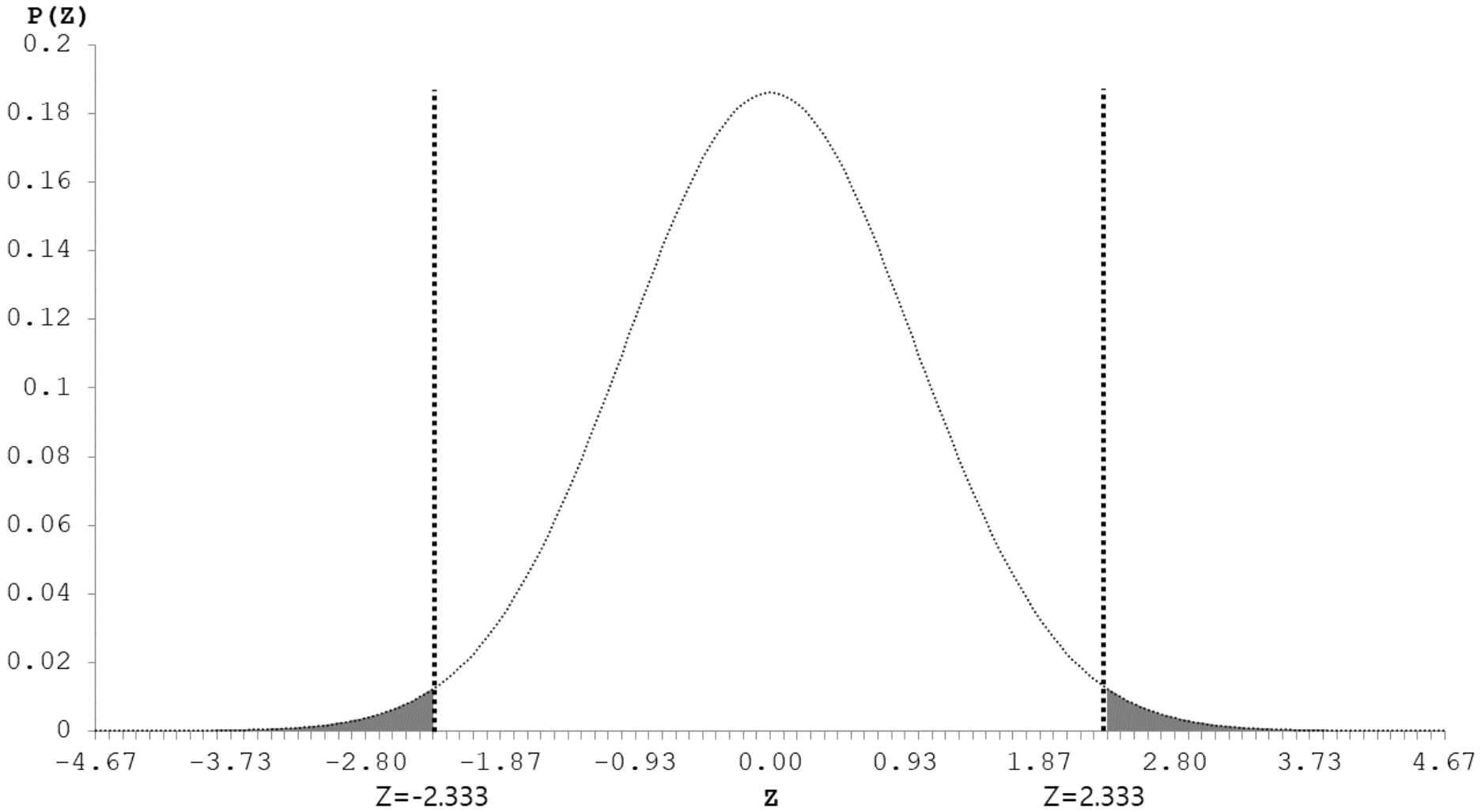
예제 2. 노바제천의 초등학교 1학년용 표준화 읽기검사 도구는 평균점수 100점, 편차는 15점을 갖도록 개발되어있다. 정후는 교육사회학적 관점에서 조기교육에 관한 석사 논문을 쓰고 있다. 그는 자녀교육에 열성인 부모를 가진 아이들이 입학 전 조기교육을 받기 마련이고, 그 때문에 읽기검사 도구의 당초 기대와는 다른 점수가 나오지 않을까 생각하고 있다. 정후는 이 연구가설을 테스트해보기 위해 조기교육을 1년 이상 이수한 초등학교 1학년생 49명을 표본으로 뽑아 표준화 읽기검사를 실시하였다. 그 결과 표본의 평균점수는 105점이였다. 정후를 위해 적절한 가설을 설정하고 5% 유의수준에서 테스트하라.

- 정후의 영가설은 $\mathrm{H}_0 : \mu = 100$ 이다. 대립가설은 $\mathrm{H}_\mathrm{a} : \mu \neq 100$이다(Why?).
- 평균이 100점인 모집단에서 무한히 많은 표본평균들을 구해 (가상적인) 표집분포를 그린다면, 그것의 표준오차(standard error)는 $\sigma/\sqrt{\mathrm{N}} = 15/\sqrt{49} = 2.143$이다.
- 이제 표본평균을 표준화하면 다음과 같다:

$$\mathrm{Z} = \frac{\bar{\mathrm{X}} - \mu}{\mathrm{S}/\sqrt{\mathrm{n}}} = \frac{105 - 100}{15/\sqrt{49}} = 2.333$$

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정



충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

- 영가설이 옳은데 표본평균이 -2.333보다 작은 값이 나올 확률은 NORM.DIST(-2.333,0,1,TRUE)로 구한다.
- 영가설이 옳은데 표본평균이 2.333보다 큰 값이 나올 확률은 1-NORM.DIST(2.333,0,1,TRUE)로 구한다.
- 두 면적은 각각 0.0098로 모두 더하면 0.0196다. 이는 (5%보다 낮은) 1.96%의 유의확률(p-value)을 가진다.
- 5% 유의확률은 실제 의미로 해석할 때 "영가설이 옳은데 기각할 확률"이 5%라는 의미이므로 기각할 때의 위험은 제법 낮은 것이다.
- 정후는 5% 유의수준에서 또는 95% 신뢰수준에서 통계적으로 유의하게 영가설을 기각하고, "조기교육을 받은 이들의 읽기점수는 그렇지 않은 이들의 읽기점수와 달랐다"는 결론에 도달하게 된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

일단 원칙적으로 임계값을 사용하는 방법과 유의확률을 사용하는 방법 둘 다 이해해야 한다.

- 하지만 실무에서나 연구에서는 압도적으로 유의확률(p-value)을 사용하는 방법이 선호된다.
- 컴퓨터 통계분석 패키지는 사용하면 자동적으로 p-value를 보고하고 여러분은 단지 그 값이 0.01 (10% 유의수준), 0.05 (5% 유의수준), 0.01 (1% 유의수준) 등만 쓱 훑어보면 유의성 여부를 알 수 있기 때문이다.
- 만약 컴퓨터를 사용할 수 없고 오로지 손으로 계산을 수행해야 한다면 신뢰구간을 사용하는 방법이 편할 수 있다.
- 그렇기 때문에 숙제나 시험에서는 어쩐지 유의확률을 사용하는 방법만 물어볼 것 같은 예감이 든다(아님 말고).

충북대학교
CHUNGBUK NATIONAL UNIVERSITY