사회통계연습
단측검정 및 가설검정 연습

김현우, PhD[1]

[1] 충북대학교 사회학과 조교수

October 15, 2021

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

진행 순서

1. 단측검정
2. 가설검정 연습
3. 마지막 코멘트

단측검정

단측검정

단측검정이 아까 배운 양측검정보다 오히려 더 직관적이고 쉽다(아님 말고).

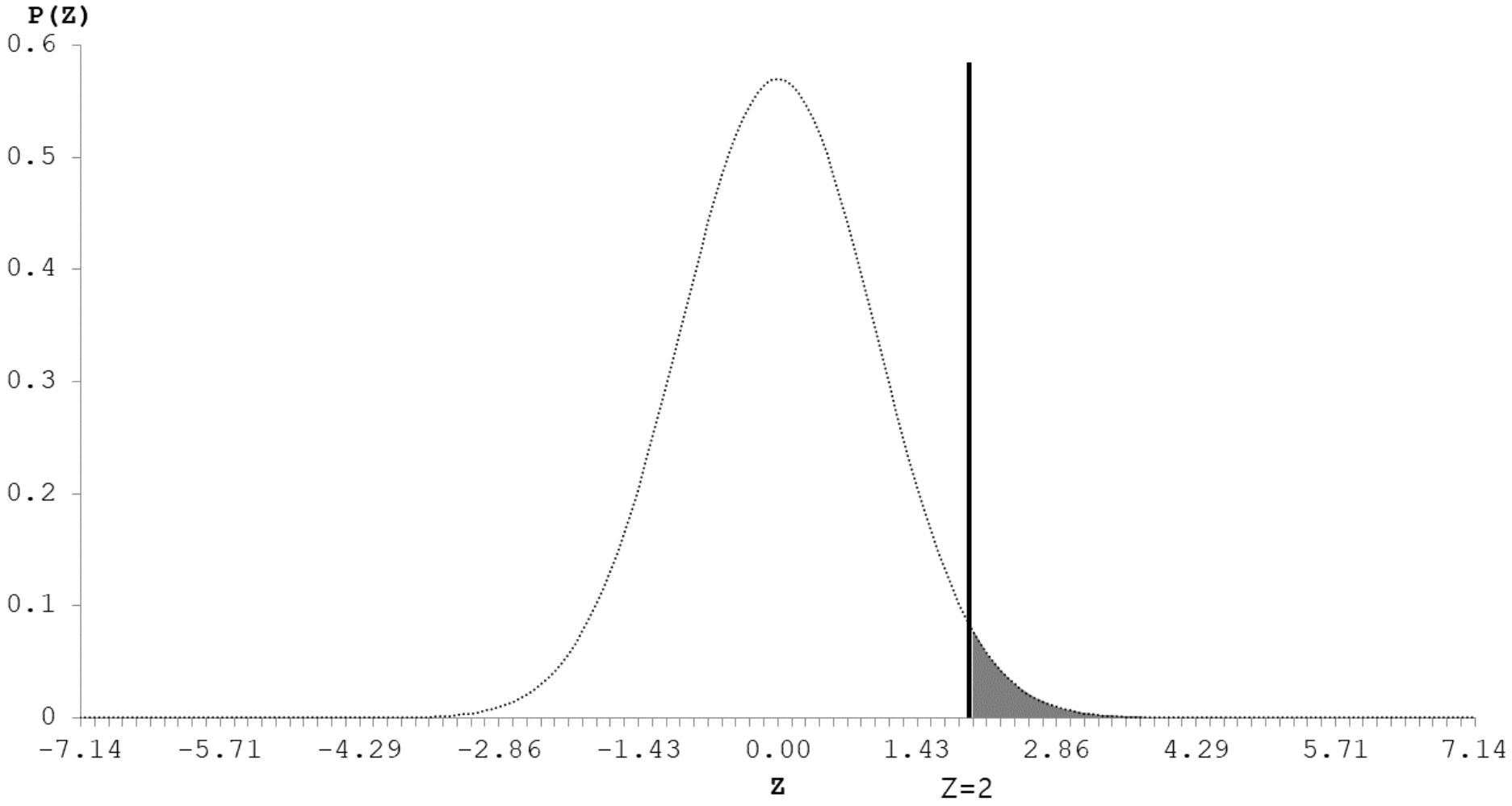
- 아까 다룬 **예제 1**과 **예제 2**는 모두 양측검정이다. "임계값을 사용하는 방식/유의확률을 사용하는 방식"와 "단측검정/양측검정"은 별개의 문제이다. 다만 앞으로는 유의확률을 사용하는 방식만 다룬다.
- 어느 표본의 평균은 54, 표준편차가 16, 표본수가 64라고 하자. 그리고 영가설과 대립가설이 아래와 같이 주어졌다. 이 경우 "영가설이 옳은데 기각할 확률", 즉 유의확률(p-value)은 어떤 식으로 그려질지 상상해 보자.

$$H_0 : \mu \leq 50, H_a : \mu > 50$$

- 표준오차($SE_{\bar{X}}$)는 2이다(Why?). 그러므로 표본평균의 Z값은 2이다(Why?). 여기서 $Z = 2$보다 오른쪽에 있는 끄트머리 면적이 곧 (모집단의 평균이 50보다 작다는) "영가설이 옳을 확률"을 의미한다.
- 물론 영가설이 옳다는 전제 아래 여전히 $Z = 2$보다 큰 값이 표본으로 추출되었을 수 있다. 다만 확률이 낮을 뿐이다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정



충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정

- 만일 $Z = 2$보다 오른쪽에 있는 음영의 면적이 0.05보다 작다면 그것은 무엇을 의미할까? "영가설이 옳은데 기각할 확률", 즉 유의확률(p-value)이 0.05보다 작다는 의미다.
- 영가설이 옳을 확률이 그렇게나 작다면 좀 자신있게(confidently) 그 영가설을 기각할 수 있을 것 같다. 이때 나는 95% 신뢰수준(confidence level)에서 영가설을 기각할 수 있다.
- 조금 더 일반화하여 말하자면, 이 구간에 속하는 Z값들이 나올 확률이 너무 낮아서 만일 α 보다도 낮다면 α 유의수준 또는 $(1-\alpha)$의 신뢰수준으로 그 영가설을 기각할 수 있다.
- 만일 유의확률이 심지어 0.01보다 작다면 어떨까? 물론 그 경우에도 나는 95% 신뢰수준(confidence level)에서 영가설을 기각할 수 있다. 하지만 기왕이면 나는 99% 신뢰수준(confidence level)에서 영가설을 기각할 수 있다고 말하는 편이 더 좋다. 더 엄격하고 보수적인 테스트를 통과했다는 의미이기 때문이다.
- 참고로 여기서 $Z = 2$보다 왼쪽에 있는 면적은 무엇을 의미할까? 그것은 (1) 영가설이 옳다는 전제 아래 (2) 표본평균이 2보다 작은 값이 나올 확률을 의미한다. 그림에서 보면 당연히 엄청 큰 부분이다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정

예제 3. 노바제천에서 고교생을 대상으로 교사에 대한 태도 검사를 실시한 결과 평균값은 178이었고 표준편차는 34였다. 그 결과를 전해듣고 네오청주 교육감 채린은 네오청주에서 자신만만하게 36명의 임의표본을 추출하여 동일한 검사를 실시하였다. 그 표본의 평균값은 170이었고 표준편차는 24였다. 채린은 풀이 죽어 노바제천보다 정말로 네오청주의 고교생이 교사를 덜 존중하는가를 살펴보려고 한다. 5% 유의수준으로 결과에 대해 논하라.

- 채린의 영가설은 $\mathrm{H}_0 : \mu \geq 178$ 이고 대립가설은 $\mathrm{H}_a : \mu < 178$이다(Why?)
- 평균이 178만원인 모집단에서 무한히 많은 표본평균들을 구해 (가상적인) 표집분포를 그린다면, 그것의 표준오차(standard error)는 $\sigma/\sqrt{\mathrm{N}} = 24/\sqrt{36} = 4$이다.
- 이제 민경이 확보한 표본평균을 표준화하면 다음과 같다:

$$\mathrm{Z} = \frac{\bar{\mathrm{X}} - \mu}{\mathrm{S}/\sqrt{\mathrm{n}}} = \frac{170 - 178}{24/\sqrt{36}} = -2$$

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

유의확률을 이용한 가설검정

- $Z = -2$보다 왼쪽 끄트머리 음영의 면적이 가지고 있는 의미를 잘 고민해보면, 이것은 (영가설이 옳다는 전제 아래) Z값이 -2보다 작은 값이 나올 확률을 의미한다. 이것이 크다면 영가설을 기각할 수 없다(Why?).
- 영가설이 옳은데 표본평균이 -2보다 작은 값이 나올 확률은 NORM.DIST(-2,0,1,TRUE)로 구할 수 있다. 그 값은 0.023 정도로 5% 유의확률보다 작다.
- 계산된 0.023이라는 p-value는 실제 의미로 생각해 볼 때 "영가설이 옳은데 기각할 확률"이 무려 약 2.3% 정도라는 것으로 100번 중 2.3번 정도만 오류를 저지른다는 것이다.
- 채린은 5% 유의수준에서 또는 95% 신뢰수준에서 통계적으로 유의하게 "네오청주의 고교생은 노바제천의 고교생보다 교사를 같은 수준으로 또는 더 존경한다"라는 영가설을 기각한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정

예제 4. 굿데이 타이어에서 생산중인 타이어의 평균 수명은 37,000km 이고 표준편차는 5,000km인 것으로 알려져 있다. 이 회사의 연구원 재현은 기존 공정을 혁신적으로 뒤바꾸는 쾌거를 달성하였다. 그의 비법에 따라 생산된 타이어 100개의 표본을 뽑아 조사한 결과 평균 수명은 무려 38,000km였고 표준편차도 동일하게 유지되는 것으로 알려졌다. 그의 혁신을 평가하기 위해 적절한 가설을 제시하고 5% 유의수준에서 테스트하시오.

- 재현의 영가설은 $\mathrm{H}_0 : \mu \leq 37,000$ 이고 대립가설은 $\mathrm{H}_a : \mu > 37,000$이다(Why?)
- 평균이 37,000인 모집단에서 무한히 많은 표본평균들을 구해 (가상적인) 표집분포를 그린다면, 그것의 표준오차(standard error)는 $\sigma/\sqrt{\mathrm{N}} = 5000/\sqrt{100} = 500$이다.
- 이제 민경이 확보한 표본평균을 표준화하면 다음과 같다:

$$\mathrm{Z} = \frac{\bar{\mathrm{X}} - \mu}{\mathrm{S}/\sqrt{\mathrm{n}}} = \frac{38000 - 37000}{5000/\sqrt{100}} = 2$$

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정

- $Z = 2$보다 오른쪽 끄트머리 음영의 면적이 가지고 있는 의미를 잘 고민해보면, 이것은 (영가설이 옳다는 전제 아래) Z값이 2보다 큰 값이 나올 확률 $P(Z \geq 2)$를 의미한다. 이것이 유의확률보다 크다면 영가설을 기각할 수 없다.
- 영가설이 옳은데 표본평균이 2보다 큰 값이 나올 확률은 1-NORM.DIST(2,0,1,TRUE)로 구할 수 있다. 그 값은 0.02275로 5% 유의확률보다 작은 값이다.
- 2.275%라는 유의확률은 실제 의미로 해석할 때 "영가설이 옳은데 기각할 확률"이 겨우 2.275% 정도라는 의미이므로 확신을 가지고 영가설을 기각할 수 있다.
- 재현은 5% 유의수준에서 또는 95% 신뢰수준에서 통계적으로 유의하게 새로운 공정이 기존의 공정과 같거나 그보다 못하다는 영가설을 기각할 수 있다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정

우리는 오늘 하루종일 가설검정을 배우면서 "평균에 관한 문제"만 다루었을 뿐 아직도 "비율에 관한 문제"를 다루지 않았다.

- 돌이켜보면 2주에 걸쳐 우리는 평균과 비율을 애써 구별하여 다루었다! 평균은 연속확률변수와 정규분포에, 비율은 이산확률변수와 이항분포에 대응시켰다.
- 그렇다면 비율의 가설검정은 아무 의미도 없기 때문일까? 설마 그럴리는 없다.
- 표본수(N)가 꽤 크다는 가정 아래 여러분은 그냥 비율 문제가 주어질 때 그냥 그것을 평균처럼 생각하고 풀면 된다. 어떻게 이것이 가능할까?
- 가장 근본적인 이유는 대규모 표본을 주로 분석하는 사회과학 분야에서 이항분포의 정규근사(normal approximation to the binomial)가 자연스럽게 활용될 수 있기 때문이다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

단측검정

- 앞서 우리는 6주차 A섹션에서 범주형 척도인 인종(1=백인; 2=흑인; ...)이나 종교(0=없음; 1=기독교; 2=불교; ...)의 평균에는 아무런 의미도 없다고 말했다. 하지만 여기에도 방법이 있다. 이런 범주형 척도도 일단 더미변수(dummy variable)로 만들면 비율의 의미가 생겨난다. 예컨대 "더미변수로서 백인"의 평균은 곧 백인의 비율을 반영하기 때문이다.
- 그렇기 때문에 (보건의료 통계학과는 별개로) 사회통계학에서는 비율에 관한 분석기법을 평균과 구분해서 배울 필요가 좀 적다.
- 물론 보건의료 통계학을 해야 하는 상황이라면 반드시, 그것도 꽤 잘 알아야 한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

가설검정 연습

예제 5. eClass에서 OVERWORK.xlsx을 다운받아 엑셀에서 열자. 이 자료는 네오청주와 노바제천 시민들 2272명을 대상으로 연령(AGE)과 주당 근로시간(WORKHOURS)을 조사한 표본이다. NEOCHUNGJU 변수는 더미변수로 1이면 네오청주 시민이고 0이면 노바제천 시민을 의미한다. 준필은 네오청주 시민의 평균 주당 근로시간이 39시간보다는 짧지 않을까 예상하고 있다. 적합한 가설을 세우고 이를 1% 유의수준에서 테스트하라.

- 먼저 네오청주 사례만을 골라내자.
- 평균, 중위값, 최빈치, 분산, 표준편차, 그리고 관찰값의 숫자를 보고하라. 표준편차를 사용할 때 우리는 지금껏 STDEV.P() 함수를 사용하였다. 그 밑에 STDEV.S() 함수도 사용해보자.
- 상단 [데이터] 메뉴에서 우측 끝의 [데이터 분석]을 클릭하여 "기술 통계법"을 선택한다. 입력 섹션과 출력 옵션 섹션으로 구분되어 있다. WORKHOURS 전체를 하이라이트한다. 입력을 어떻게 하는가에 따라 "첫째 행 이름표 사용"을 체크해야 한다. 출력 옵션에서는 "요약 통계량"을 반드시 체크해야 한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

- STDEV.P() 함수와 STDEV.S() 함수의 차이는 자유도(degree of freedom)에서 온다. 사례가 작을때 자유도는 더 정확한 추정을 위해 필수적이지만, 구체적으로 어떻게 계산하여야 하는가는 여러분이 구태여 알지 못해도 좋다.
- 준필의 영가설은 무엇인가?

$$\mathrm{H}_0 : \mu \geq 39, \mathrm{H}_a : \mu < 39$$

- 주어진 표본에서 평균의 표준화된 Z값을 계산하자. 답은 -1.954이다.
- 테스트할 유의확률은 0.01이므로 Z분포에서 왼쪽 임계값 끄트머리의 면적이 0.01보다 작으면 영가설을 기각할 수 있다(Why?).
- NORM.DIST(-1.954, 0, 1, TRUE)로 면적을 계산해본 결과 0.01보다 살짝 크므로 1% 유의수준에서 통계적으로 유의하게 영가설을 기각할 수 없다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

준필의 연구를 곰곰히 지켜본 동혁은 노바제천 시민들이야말로 평균 주당 근로시간은 37.5시간이 아닐까 생각하고 있다. 적절한 가설을 세우고 1% 유의확률로 테스트하라.

- 동혁의 가설은 무엇인가?
- 주어진 표본에서 평균의 표준화된 Z값을 계산하라.
- 테스트할 유의확률은 무엇인가? Z분포에서 왼쪽과 오른쪽 임계값 끄트머리의 면적은 각각 무엇인가? 합하면 무엇인가?
- 어떻게 엑셀 함수로 그 면적을 계산할 것인가?
- 1% 유의수준에서 통계적으로 유의하게 영가설을 기각할 수 있는가?
- 5% 유의수준에서라면 어떤가?

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

같은 데이터를 지켜보던 승연에게 사실 평균 주당 근로시간은 아무런 관심 대상도 아니다. 그보다 승연은 노바제천의 평균연령이 40세보다 적을 것이라고 예상하고 있다. 승연을 위한 적절한 가설을 세우고 1% 유의확률로 테스트하라.

- 승연의 가설은 무엇인가?
- 주어진 표본에서 평균의 표준화된 Z값을 계산하라.
- 테스트할 유의확률은 무엇인가? Z분포에서 왼쪽/오른쪽 임계값 끄트머리의 면적은 무엇인가?
- 어떻게 엑셀 함수로 그 면적을 계산할 것인가?
- 1% 유의수준에서 통계적으로 유의하게 영가설을 기각할 수 있는가?
- 결론은 무엇인가?

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

채은에게는 이제 선택의 여지가 남아있지 않다. 어차피 데이터에서 아직 분석되지 않은 건 네오청주의 평균연령 밖에 없다. 어찌되었든 채은은 네오청주의 평균 연령이야말로 39세보다 많을 것이라고 생각한다. 채은을 위한 적절한 가설을 세우고 1% 유의확률로 테스트하라.

- 채은의 가설은 무엇인가?
- 주어진 표본에서 평균의 표준화된 Z값을 계산하라.
- 테스트할 유의확률은 무엇인가? Z분포에서 왼쪽/오른쪽 임계값 끄트머리의 면적은 무엇인가?
- 어떻게 엑셀 함수로 그 면적을 계산할 것인가?
- 1% 유의수준에서 통계적으로 유의하게 영가설을 기각할 수 있는가?
- 결론은 무엇인가?

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

잘보면 영가설과 대립가설 중에 =은 항상 영가설쪽에 붙는다.

- 양측가설이건 단측가설이건 H_0 쪽에 =이 붙어있는데, (양측가설이야 그렇다 치더라도) 단측가설의 경우, (부등호의 방향이 다른) 대립가설과 비교되는 순간의 영가설은 결국 =이 나온 부분이기 때문이다.
- 예컨대 아래의 "모평균이 50보다 크다"는 대립가설을 채택할 것인가를 결정하기 위해 정말로 들여다보아야 하는 영가설은 "모평균이 50이 되는" 순간이고 "모평균이 50보다 작은" 부분들은 별 의미가 없다(Why?).

$$H_0 : \mu \leq 50, H_a : \mu > 50$$

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

왼쪽 꼬트머리인가 오른쪽 꼬트머리인가 고민할 필요는 실용적이지 않은 고민이다.

- 표집분포는 $N > 30$의 조건 아래 결국 정규분포하고 이는 대칭적이다. 다시 말해, 양쪽 꼬트머리의 면적은 같다.
- 문제를 성급히 풀려고 하기 전에 먼저 차분히 그림을 그리고 논리의 흐름을 되짚어 보자. 논리적인 흐름이 중요하기 때문에 이것을 구분했을 뿐, 계산 자체는 어느 쪽 꼬트머리를 사용하건 같다.
- 계산하여 나온 p-value가 "말이 되는지" 충분히 음미해 보아야 한다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

가설검정 연습

참고하는 교과서에 따라서는 이번 주 수업내용이 제대로 다루어지고 있지 않을수도 있다.

- 어떤 교과서에서는 이 파트에서 표준정규분포, 즉 Z 분포 대신에 t 분포(t distribution)와 단일표본 t 검정(one-sample t test)을 다루고 있을 수도 있다.
- 이 부분은 스스로 공부해도 좋고 넘어가도 좋다. 어차피 우리도 곧 따라가 공부할 내용이다.
- 왜 t 분포를 언급할까? t 분포는 사례가 적고 모집단의 표준편차를 모르는 경우에 유일한 대안이기 때문이다.
- 특히 옛날 스타일로 쓰여진 교과서에서는 자꾸 사례수(N)가 극단적으로 적은 상황과 결합하여 모집단의 표준편차를 아는지 모르는지를 구분한다. 이런 구분 자체가 다소 현실성이 없다. 모집단의 표준편차를 알 정도라면 표본을 구태여 뽑을 이유도 없다.
- 우리는 당연히 모집단의 표준편차(σ)를 모르기 때문에 이것의 불편추정량(unbiased estimator)인 표본의 표준편차(S)를 대신 사용하고 있다.
- 게다가 솔직히 말해 사례수가 30에도 미치지 않는 사회통계분석은 거의 없다(단 보건의료통계학은 예외!).

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

마지막 코멘트

마지막 코멘트

가설검정은 기초사회통계를 마치고 중급사회통계로 진입하는 관문과도 같다.

- 지금까지 걸어온 길을 반추해보자. 가설검정을 이해하기 위해서는 확률변수, 확률분포, 표집분포, 표준오차, 누적분포함수, 표준정규분포, 임계값, 유의수준을 모두 알아야만 했다!
- 이제 여러분은 가설검정을 이해하고 있다. 내가 겪은 바에 따르면 절대다수의 현업 사회조사 전문가는 이 이상의 통계 원리를 실무에서 사용하지 않는 것 같다(물론 미래에는 어떻게 될지 모른다).
- 이들의 노하우는 오히려 더 기본적인 것들을 잘 실천하는데 있다: 패널유지, 조사설계, 문항개발, 기본적인 표 만들기와 기술통계량 해석 등등!
- 믿지 못할지도 모르지만 여러분이 실무자로서 (어쩌면 연구자로서!) 가설검정을 사용해야 할 상황은 사실 굉장히 많다. 다만 스스로 그런 상황에 놓여있음을 눈치채지 못하는 것 뿐이다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

마지막 코멘트

그럼 사회통계연습의 남은 시간동안 우리는 무엇을 배울까?

- 가설검정을 할 수 있게 된 우리는 이제 그 원리를 활용하여 실천적인 기법을 연습한다.
- 두 개의 표본이 주어져 있을때 (1) 평균과비율 또는 (2) 분산을 비교한다. 몇 주 전에 살짝 맛보기 했던 (3) 교차표를 좀 더 공부한다. (4) 문항 신뢰도를 측정하기 위한 기법과 상관분석을 공부한다. (5) 회귀분석을 공부한다. (6) 회귀분석을 더 공부한다. 이렇게 남은 6주 동안 오로지 실천적인 기법만 연습한다.
- 가설검정은 물론 계속해서 쓰이게 될 것이다. 실천적인 기법 속에 가설검정이 어떻게 녹아들어가 있는지 알게 된다.

충북대학교
CHUNGBUK NATIONAL UNIVERSITY

마지막 코멘트

중간시험에 관한 코멘트(중요함)

1. 모든 과제를 전부 다 처음부터 풀어볼 것. 하나도 모르는 것은 없어야 한다. 단 하나도. 왜냐하면 모든 과제가 한꺼번에 나오기 때문이다(JASP와 Tableau 제외). 쉬지 않고 쭉 풀어보는 연습을 해서 "오~ 모든 과제들이 이런 식으로 연결되는구나" 하고 깨우쳐야 한다.
2. 시험시간 내내 반드시 카메라를 켜고 컴퓨터 앞에 앉아있어야 하며 얼굴을 비추어야 한다. 카메라가 계속 천장을 향하거나 꺼지면 결시로 간주한다. Zoom에서 튕기면 곧바로 다시 들어올 것.
3. 원한다면 마음껏 인터넷 검색을 하거나 교과서를 참고해도 좋다. 숙제를 다시 봐도 좋고 신문을 읽어도 좋다. 하지만 기억할 것: 시간은 1시간 30분 밖에 없다. 이 시간을 넘기면 감점한다.
4. 과제는 워드 파일로 주어지며 최선을 다해 모든 풀이과정과 논리적 흐름을 담아 작성할 것. 계산이 틀려도 논리가 맞으면 부분 점수가 인정된다.
5. (필요하다면) 이번 주에는 가능한 친구들과 직접 만나 토론하여 모르는 것을 서로 물어보고 가르쳐 줄 것. 그래도 모르겠으면 메일 또는 조교 선생님을 통해 약속을 잡아 내 연구실로 찾아올 것(다음 주 내내 오후에는 시간을 비워둘 예정). 짧은 내용이라면 메일로 질문해도 괜찮다.

충북대학교 CHUNGBUK NATIONAL UNIVERSITY

끝!

와~ 중간고사 전인데 숙제가 있네~

충북대학교
CHUNGBUK NATIONAL UNIVERSITY