

- 학습자 중심 실전 문제 -

금융정보통계학과 20173204 광명빈

1. 다음은 통계학의 연구대상에 대한 Fisher의 이야기를 요약한 것이다. 빈칸을 순서대로 채우면 순서대로 괄호의 내용을 쓰시오.

Statistics may be regarded (i) as the study of (populations), (ii) as the study of (variation), (iii) as the study of the methods of the (reduction of data).

2. H대학에서는 최근 학령인구가 저하되어, 입학생의 수요가 매우 어려워짐에 따라, 신입생 장학금 제도를 개선하고자 한다. 기본에는 전교수석(장학금 100원, 1명), 단대수석(장학금 70원, 5명), 입학전형 수석(장학금 30원, 수시 6명, 정시 6명)의 3가지 종류가 있었다. 새롭게 설계되는 장학금의 총량은 1000원이고, 입학생 50명에게 골고루 지급하고자 한다. 어떻게 장학제도를 설계할 것인가? 그 기준을 명확히 설명하고, 50명의 학생들의 사례를 직접 들어서, 구체적인 설계 장학금 수혜액을 제시하시오.

H 대학의 문제는 학령인구 저하에 따른 입학생 수의 감소이다. 이를 해결 하려고 장학금의 총량을 810원에서 1000원으로 늘리고 장학금수혜 학생을 18명에서 50명으로 늘리려는 계획을 하고 있다.

	기존		변경	
전교수석	12.3%	1명	10%	1명
단대수석	8.64%	5명	6%	5명
입학전형 수석	3.7%	12명	2.6%	10명
그 외	0%	·	1%	34명
합계	-	18명	-	50명

기존에는 장학금 810원을 18명에게 위와 같은 비율로 분배를 하였다. 표를 살펴보면 1명에게 12.3% 5명에게 8.64% 12명에게 3.7% 로 차이를 비교해보면 전교수석과 단대수석은 3.6%p 단대 수석과 입학전형 수석은 4.94%p 차이를 보인다. 변경 후에는 더 많은 사람에게 장학금을 줘야하기 때문에 기존의 수혜학생들에 비해 더 적은 장학금을 받을 수밖에 없다.

하지만 기존의 장학금을 비율을 적극 반영하여 표와같은 계획을 제시하였다.

이렇게 되면 기존의 학생들과 변경후 학생들의 장학금 수혜액은 총 장학금량이 늘었기 때문에 큰 차이가 없다. 전교수석은 그대로 100원, 단대수석은 70원->60원, 입학전형수석은 30원에서 26원을 받게된다. 나머지는 10원씩 34명이 받을수 있다.

3. (주) 뽀로로 자동차는 신차 홍보전략으로 연비효율을 제시하고 있다. 영업본부장은 실제로 방송사와 협의하여 연비효율의 직접적인 사례를 보여주기 위하여 자사의 구형 자동차와의 연비효율 비교를 진행하기로 하였다. 김 부장은 구형 A,B,C모델과 뽀로로 신형 D모델의 차량연비를 비교하는 실험설계를 직접 팀원들에게 맡기고, 가장 새로운 방법을 제안하는 제안서를 채택하여, 업무실적에 반영기로 하였다. 직접 방안을 설계하는 제안서를 만드시오. 여기서, 연비란 동일거리 대비 휘발유 소모량이 적을수록 연비가 높다고 말함.

실험조건

- 동일한 인물이 주행
- 동일한 장소 주행
- 구형모델은 중고가 아닌 신차를 사용

* ·은 휘발유 소모량임

	평균시속50km	평균시속70km	평균시속100km	평균시속110km
A	·	·	·	·
B	·	·	·	·
C	·	·	·	·
D	·	·	·	·

사용 휘발유양이 D가 가장 낮음을 강조한다

4. 다음의 기사 내용에 답하시오.

[영 네이처지] "신부 나이 적을수록 첫아들 출산" , '첫 아들을 갖고 싶은 남자는 나이 어린 여자와 결혼하라.'

과학전문잡지 '네이처(9월25일자)'에 보도된 영국 리버풀대학의 연구 결과에 따르면 부모의 나이 차이가 첫 출산아의 성 결정에 큰 영향을 끼친다고 한다. 이 연구진이 영국 웨일스와 잉글랜드 지방에서 1911~1952년까지 태어난 사람들을 대상으로 3백여 가계를 분석한 결과, 아버지가 어머니의 나이보다 많을수록 첫번째 출산에서 아들을, 적을수록 딸을 많이 낳았다. 연구진은 우수한 남성일수록 자신보다 나이가 어린 여성과 결혼하는 경향이 있으며, 이 경우 첫째를 남자를 낳을 가능성이 많았다고 밝혔다. 그러나 출산당시의 부모연령은 남녀 성비와는 무관했다. 이 연구진은 다만 이번 분석이 1-2차 세계대전 때와 전쟁 후에 해당, 부모 나이 차 외 다른 요인들을 고려할 수 있다고 밝혔다. 즉 전쟁중이나 전후의 특수한 행동습성, 물리적 환경변화, 심리적 중압감등으로 인해 성비의 편향요인으로 작용했을 가능성은 배제할 수 없다고 말했다. < 차병학기자 >

(1) 위의 기사내용의 결과를 적용할 모집단은 무엇인가? 만일 모집단을 언급하기 어렵다면, 어떻게 기사내용과 조사방법을 수정해야 하는지 제안하시오.

모집단은 1911~1952년까지 출산이 가능한 여성이다. 기사는 영국 웨일스와 잉글랜드 지방에서 1911~1952년까지 태어난 사람들을 대상 이 부분을 영국 웨일스와 잉글랜드 지방에 거주하는 여성으로 바꿔야 한다.

(2) 첫 아들의 출산여부에 신부의 나이가 영향을 미친다고 말할 수 있는가? 답변에 대한 근거를 제시하시오.

나이에는 영향을 미치지 않는다. 위 300백여 가구의 표본은 나이가 영향을 미친다고 할 정도의 표본의 개수가 떨어지며 신뢰도가 떨어진다. 신부의 나이에 영향을 미친다고 하려면 좀 더 많은 표본을 가지고 봐야할 것이다.

(3) 위 내용의 기사 주장에 반박하기 위하여, 어떤 문제제기를 통하여 데이터를 조사할 것이며, 발생될 수 있는 제 3의 영향요인은 어떻게 배제할 것인지 논하시오.

1911~1952년은 전쟁중이었기 때문에, 남자아이를 더 선호했을 가능성이 커서 남자아이가 아니면 낙태를 했을 가능성이 존재한다. 또한 주민등록상 등록이 되지않은 아이들도 많았을 것이기 때문에 제 3의 영향요인을 배제 하기 위해서는 임신하였을 때 아이의 유전자 검사를 통하여 남성인지, 여성인지 확인해봐야 할 것이다.

5. 평균과 중앙값의 차이와 장단점을 기술하시오.

	평균	중앙값
차이	평균은 모든 자료의 평균값을 나타낸 것이고 중앙값은 자료 중앙에 있는 값이다.	
장점	모든 자료의 값을 사용한다.	자료의 값이 크거나 작을 때 자료의 대표로 사용할수 있다.
단점	편차가 큰 자료의 값을 사용할 때 대푯값이 왜곡될 수 있다.	자료의 값 모두를 사용하지 않고 중앙값만 나타낸다.

6. 완전랜덤계획법과 블록랜덤계획법을 각각 설명하고, 두 설계의 차이점을 사례를 들어 설명하시오.

	완전랜덤 계획법	블록랜덤계획법
정의	모든 표본의 의 순서, 실험단위가 완전한 랜덤으로 실행되는 실험	비슷한 조건의 표본을 블록으로 묶은다음 진행하는 실험
차이	임상실험에서 완전 랜덤계획법은 랜덤하게 실험대상에게 주는 반면 블록랜덤계획법은 각 사람마다 성별 나이 체질이 다르기 때문에 비슷한 조건인 성별, 나이에 별로 묶어서 실험하는 방법이다.	

7. 비교실험 설계에 반드시 필요한 조건을 언급하고, 비교실험에서 발생 가능하여 교락되는 제 3의 영향요인을 통제하는 방안을 대표적인 실험설계를 사례로 들어 설명하시오.

반드시 필요한 조건

- 모든 조건이 동일 해야한다.
- 외생변수의 영향이 적어야한다.
- 랜덤배치 해야한다.(랜덤화)

실험군과 대조군을 비교한다

어떤 약이 더 효과적인가?에 대한 방안을 예로 들면, 실험군, 대조군을 두어 무작위로 약을 준다. 또한 신뢰도를 높이기위해 표본수를 크게 늘리면 된다. 이렇게 되면 많은 표본안에서 교락되는 제 3의 영향요인을 통제할수 있고, 어떤 약이 더 효과적인이 알 수 있다.