

표본론 2023년 중간시험 (월, 수 10:30~11:45)

[1] 전화를 이용한 선거예측조사에서 무선전화면접 방식과 무선전화 자동응답시스템(ARS) 방식 중에서 어떤 방식으로 조사했느냐에 따라 조사결과가 달라진다는 주장이 있다. 이를 확인하기 위해 2022년 2월 교통방송(TBS)의 의뢰를 받아 한국사회여론연구소(KSOI)가 동일한 시점에 동일한 설문지를 이용해 두 가지 다른 방식으로 조사했다. 아래 자료를 보고 물음에 답하시오.

	무선전화면접	무선ARS
표본크기	1005	1000
이재명 후보의 지지율	43.8%	43.2%
윤석열 후보의 지지율	36.1%	45.0%

[1-1] 무선전화면접 조사에서 이재명 후보와 윤석열 후보의 지지율 차이에 대한 95% 신뢰구간을 구하시오. 1005명의 표본을 단순확률추출 했다고 가정하시오.

[1-2] 두 조사방식에서 윤석열 후보의 지지율에 유의한 차이가 있다고 할 수 있는가? 이때 두 조사에서 각각 단순확률표본을 추출했다고 가정하시오. 유의수준은 5%이다.

[2] 어떤 기업의 인사부 담당자가 지난 한 해 동안 기업 소유의 46개 생산공장에서 이루어진 병가 일수의 총합을 추정하고자 한다. 46개 공장을 20개의 작은 공장들로 이루어진 층, 26개의 큰 공장 층으로 나눌 수 있다고 한다. 작은 공장에서 한 해 동안 병가일수는 대략 0에서 100 사이이고, 큰 공장에서 병가일수는 대략 10에서 200 사이라고 한다. 한 해 동안 총 병가일수를 오차의 한계 100으로 추정하고 싶다고 할 때

[2-1] 표본크기를 두 층에 어떤 비율로 할당하는 것이 좋은가?

[2-2] 적절한 표본크기를 구하시오.

[3] 단순확률표본(simple random sample) y_1, \dots, y_n 의 표본평균을 \bar{y} , 표본분산을 s^2 이라고 할 때

[3-1] $V(\bar{y}) = \frac{\sigma^2}{n} \frac{N-n}{N-1}$ 임을 증명하시오. 증명 과정에서 $i \neq j$ 일 때 $cov(y_i, y_j) = -\frac{\sigma^2}{N-1}$ 임을 증명없이 이용해도 된다.

[3-2] $E(s^2) = \frac{N}{N-1} \sigma^2$ 임을 증명하시오. 이때 $V(\bar{y}) = \frac{\sigma^2}{n} \frac{N-n}{N-1}$ 을 이용해도 된다.

[4] “무응답층을 보정하기 위해 사후층화(poststratification)를 할 수 있으나 무응답층과 응답층이 다른 특성을 가지는 경우 무응답층으로 인한 편향은 여전히 문제가 된다”라는 주장이 무엇을 의미하는지를 구체적 수치를 들어 설명하시오.

[5] 중앙선거조사여론심의위원회 홈페이지에 다음과 같은 내용이 공지되어 있다.

※ 현재 선거여론조사 등록제의 취지를 제고하기 위해 '접촉률'을 공개하고 있으며, 접촉률과 응답률(협조율)을 통해 국제기준(미국여론조사협회)의 응답률을 산출할 수 있습니다.

$$\text{접촉률} \times \text{응답률(협조율)} = \text{국제 기준(AAPOR) 응답률}$$

전화조사에서 **협조율**은 접촉(전화연결)이 이루어진 총 사례수, 즉 조사거절, 중도이탈, 응답완료를 모두 합한 사례수에서 응답완료 사례수가 차지하는 비율이다. **접촉률**은 전화를 걸었을 때 접촉이 이루어진 사례수가 차지하는 비율로서 통화중이거나 전화를 받지 않은 사례가 많으면 접촉률은 낮아진다. (이때 결번으로 인한 접촉 실패는 접촉률 계산에서 고려하지 않는다.) 우리나라에서 응답률은 협조율이며 국제기준 응답률은 이 협조율에 접촉률을 곱한 값으로서 접촉률이 낮을수록 국제기준 응답률은 낮아지게 된다. 우리나라 응답률은 접촉률을 고려하지 않는데 접촉률을 고려하지 않아도 된다고 생각하는가? 국제기준 응답률과 같이 접촉률을 고려해야 한다면 그 이유는 무엇인지 설명하시오.