

15강. 표본조사 사례

◆ 담당교수 : 이기재 교수

■ 주요용어

용어	해설
확률화 응답기법	개인적으로 민감한 내용을 묻는 경우에 민감한 내용의 질문에 대한 긍정과 부정으로 표현하는 두 종류의 형태로 구성된 후에 응답자는 긍정 문장 형태와 부정 문장 형태 중에서 랜덤하게 선정하여 "예" 또는 "아니오"라고 응답하게 하는 조사방법. 질문형태를 랜덤하게 선택하는 방법과 응답결과인 "예" 또는 "아니오"의 비율만을 이용하여 조사대상자들의 민감속성의 비율을 추정할 수 있음
표본설계의 주요 절차	조사목적 명시, 모집단 정의, 목표 정도의 선택, 자료수집방법의 결정, 표본추출방법 결정, 실사관리방안, 자료분석 및 통계산출 등의 과정으로 요약할 수 있음
확률비례계통추출법	조사구 추출에서 조사구내에 포함된 가구 수에 비례하도록 추출확률을 주어 표본조사구를 추출하는 방법. 가구 수가 많은 조사구가 추출될 확률이 커지므로 대표성이 높아짐
사후층화 (post-stratification)	표본조사를 완료한 후에 모집단의 구조와 표본의 구조를 맞추기 위해서 층화하는 방법. 모수추정에서 편향을 줄이는 효과가 있음
제공근비례 배분법	층의 크기에서 많은 차이가 있을 경우에 층 단위의 통계생산을 안정적으로 하기 위해서 층 크기의 제공근에 비례하도록 층별로 표본을 배분하는 방법

■ 연습문제

1. 마약 복용 여부, 시험 중 부정행위 경험여부, 세금탈루여부 등과 같이 개인적으로 민감한 사안에 대한 조사에서 직접 질문하는 대신에 확률화 응답기법을 사용하는 이유를 가장 정확하게 설명한 것은?
 - ① 직접 질문에 비해서 조사비용과 시간을 절약할 수 있다.
 - ② 직접 질문과 비교해서 모수추정과 추정오차의 계산이 간편하다.
 - ③ 직접 질문에 비해서 표본오차를 줄일 수 있는 방법이다.
 - ④ 직접 질문에 비해서 거짓응답을 줄임으로써 비표본오차를 줄일 수 있다.

정답 : ④

해설 : 확률화 응답기법은 직접 질문을 하는 경우에 발생하는 거짓응답으로 인한 비표본오차를 줄일 수 있다.

2. 세금탈루 여부, 시험 중 부정행위 여부, 마약복용 여부 등과 같은 민감한 사안에 대해서 조사를 진행하고자 한다. 직접적으로 질문하는 경우에 발생하는 가장 심각한 문제는 무엇인가?

- ① 거짓 응답으로 인하여 비표본오차가 커져서 문제다.
- ② 직접적으로 질문하는 경우에는 추정량의 표준오차가 커져서 문제다.
- ③ 결과를 분석하여 표를 작성하는데 어려움이 많다
- ④ 모수 추정방법에 복잡해져서 문제이다.

정답 : ①

해설 : 세금탈루 여부, 시험 중 부정행위 여부, 마약복용 여부 등과 같은 민감한 사안에 대해서 조사에서 직접 질문하게 되면 응답자들은 거짓 응답으로 실제와는 다른 응답을 하기 쉽다. 이에 따라 비표본오차가 크게 발생할 수 있다.

3. 층화임의추출법에서 층별로 표본배분법의 특징을 설명한 것이다. 가장 정확하게 설명한 것은 어느 것인가?

- ① 균등배분법은 층의 크기가 다를 경우에도 적용할 수 있다.
- ② 추정량의 분산을 가장 작게 하는 배분법은 비례배분법이다.
- ③ 층별로 통계를 생산하고자 할 경우에는 비례배분법이 가장 효과적이다.
- ④ 전체 모집단 단위의 통계생산에서는 제곱근비례배분법이 효과적이다.

정답 : ①

해설 : 균등배분법은 층별 추정이 목적인 경우에 주로 사용되며, 층의 크기가 다를 경우에도 적용할 수 있는 방법이다.

4. 주거실태조사의 표본설계에서 적용한 표본추출법에 대한 설명이다. 가장 정확하게 설명한 것은?

- ① 조사구마다 포함하고 있는 가구 수가 상이하므로 조사구의 표본추출확률을 가구 수에 비례하게 하기 위해서 확률비례계통추출법을 사용하여 조사구를 선정하였다.
- ② 조사구의 특성을 표본추출에 반영하기 위해서 특·광역시와 도 지역에서 공통적으로 공동주택비율, 자가주택비율, 가구주연령과 가구주 학력을 기준으로 정렬한 후에 계통추출법으로 조사구를 추출하였다.
- ③ 표본조사구의 특성이 변경되었을지라도 조사현장에서 조사가 가능한 경우에는 표본조사구의 교체 없이 조사를 진행해야만 표본의 대표성을 유지할 수 있다.

- ④ 조사구 내의 조사대상가구 선정은 조사원이 현장에서 조사편의성을 고려하여 조사에 적극 협조하는 가구를 우선적으로 조사하였다.

정답 : ①

해설 : 주거실태조사에서는 조사구 내 가구수에 비례하는 확률비례계통추출법을 사용하여 조사구를 선정하였다.

■ 정리하기

- 항목 무응답이 있는 경우에 응답자의 응답내용과 기타 보조정보를 이용하여 무응답 항목의 값을 대체값으로 채워 넣는 것을 항목 무응답 대체(imputation)이라 하는데 항목 무응답 대체 방법으로는 평균대체, 최근방대체, 회귀대체, 핫덱대체 등이 있다.
- 개인적으로 민감한 사안 조사에서 직접질문을 할 때는 응답거부 또는 거짓응답으로 비표본오차가 발생하기 때문에 이에 대한 대책방안으로 확률화 응답 기법이라는 간접질문방법을 사용한다.
- 주거실태조사는 지역별·계층별 주거실태 파악을 통해 정부정책수립 및 관련 분야 연구를 위한 기초자료를 제공하는 것을 목적으로 한다. 각 시도별 통계생산을 위해서 7개 특·광역시와 9개 도 지역을 1차로 층화하고, 9개 도 지역에 대해서는 동부와 읍·면부로 2차 층화하였으며, 서울은 4개 권역으로 세분하여 통계작성을 정확성을 높이고자 하였다.
- 표본크기 결정을 위해서 3개 관심변수(아파트 비율, 자가주택 비율, 사용면적)에 대한 목표상대표준오차를 지역별로 차등적용 하였으며, 전체적으로 표본크기는 3,300개 조사구에 33,000가구로 하였다. 16개 시도에서 세부 층별 표본배분은 네이만배분법을 이용하였다.
- 표본추출은 조사구를 1차 추출단위로 하였고, 가구를 2차 추출단위는 하였는데 조사구는 조사구 내의 가구 수에 비례하도록 확률비례계통추출법을 적용하였다. 가구는 선정된 조사구 내의 가구명부에서 계통추출 또는 단순임의추출법으로 추출하였다. 모수추정은 표본가중추정법을 적용하였는데, 가중치는 설계가중치, 응답률 보정, 사후층화보정의 과정을 거쳐서 산출하였다.

■ 참고문헌

- 이계오, 박진우, 이기재, 표본조사론, 한국방송통대학교출판부, 2013. 제1장
- 통계청 홈페이지 : <http://www.nso.go.kr>