

14강. 비표본오차

◆ 담당교수 : 이기재 교수

■ 주요용어

용어	해설
표본오차(sampling error)	조사대상의 일부만을 선정하여 조사한 내용으로 전체 조사대상자의 특성을 추정하기 때문에 발생한 오차
비표본오차(nonsampling error)	조사기획에서 결과의 공표과정까지 단계에서 부주의 등으로 인해 발생하는 오차
무응답 오차(nonresponse error)	조사단위에서 데이터를 수집하지 못한 데서 발생하는 오차
응답오차(response error)	조사대상자가 응답한 내용이 참값과 다른 데서 발생하는 오차
단위 무응답(unit nonresponse)	조사단위로부터 관찰값을 확보하지 못한 경우를 말함
항목 무응답(item nonresponse)	조사단위가 응답을 하였으나 일부 문항 중에서 대답하지 않은 경우를 말함
평균대체법	항목 무응답을 보완하는 방법으로 조사단위의 특성이 유사한 것끼리 층을 구분한 후에 동일한 층에서 응답자들의 평균을 항목 무응답의 대체값으로 이용하는 방법
최근방대체법	전체 표본을 유사한 특성별로 구분하여 대체 층을 구성한 후 각 층 내에서는 지역별, 특성별로 기준을 정하여 응답자를 정렬한 후 항목 결측값이 발생한 경우에는 해당 응답자의 바로 앞 단위의 항목 응답값으로 대체하는 방법
회귀대체법	항목 무응답 있는 문항을 종속변수로 간주하여 그것과 연관성이 큰 문항들을 독립변수로 하여 선형회귀모형을 조사 데이터에 대해 적합하여 얻은 예측값으로 항목 무응답을 대체하는 방법
핫덱대체법	무응답 대체 층 내에서 응답한 조사결과 중에서 랜덤하게 선정하여 항목 결측값의 대체값으로 이용하는 방법

■ 연습문제

1. 다음은 표본조사에서 발생하는 오차에 대한 설명이다. 옳지 않은 것은?

- ① 표본조사의 오차는 우연오차(variable error)와 편향(bias)으로 구성된다.
- ② 표본조사의 오차는 발생 원인에 따라 크게 표본오차(sampling error)와 비표본 오차(non-sampling error)로 구분된다.
- ③ 편향은 비확률변수적 특성으로 방향성을 갖는다.
- ④ 표본오차는 무응답오차, 응답오차, 자료처리과정 중의 오차(종합표 작성 오차)로 구분되며 통제할 수 없다.

정답 : ④

해설 : 무응답오차, 응답오차, 자료처리과정 중의 오차는 비표본오차의 종류이다.

2. 일관성 점검, 사후표본조사 점검, 중복부차표본 기법, 외부기록 점검 등은 무엇을 위한 것인가?

- ① 오차를 줄이는 방법
- ② 추정량의 표본오차를 계산하는 방법
- ③ 비표본오차를 측정하는 방법
- ④ 종합표작성오차를 줄이는 방법

정답 : ③

해설 : 일관성 점검, 사후표본조사 점검, 중복부차표본 기법, 외부기록 점검 등은 비표본오차를 측정하기 위한 방법이다.

3. 항목 무응답의 대체 방법 중에서 표본분포가 모집단의 분포에 비해서 왜곡될 수 있기 때문에 주의해야 할 대체법은 어느 것인가?

- ① 평균대체
- ② 최근방대체
- ③ 회귀대체
- ④ 핫덱대체

정답 : ①

해설 : 평균대체법은 한 층에 여러 개의 결측값이 있을 경우에 동일한 표본평균값이 대체되기 때문에 표본분포가 모집단의 분포에 비교해서 심각하게 왜곡될 수 있다.

4. 사후 표본조사 점검을 실시하는 목적을 설명한 것이다. 잘못 설명한 것은 어느 것인가?

- ① 본 조사 후 2-3주 이내에 실시하여 기억오차를 줄여야 한다.
- ② 전체 응답자의 10% 정도를 랜덤하게 선정하여 경험이 많은 조사원이 응답내용

을 확인하는 조사이다.

- ③ 전수조사인 센서스 등에서는 사후 표본조사 점검을 실시하지 않는다.
- ④ 조사원의 사전 교육에서 사후 표본조사 점검에 대해 알리면 조사과정에서 오류를 줄일 수 있다.

정답 : ③

해설 : 전수조사인 센서스 등에서도 사후 표본조사 점검을 실시하여 비표본오차가 어느 정도 되는지 살펴보게 된다.

5. 다음은 통계조사에서 발생하는 편향(bias)에 대한 설명이다. 옳지 않은 것은?

- ① 모수 추정에 사용된 통계량의 기대값과 모수 참값의 차이를 말한다.
- ② 전수조사에서는 발생하지 않는다.
- ③ 발생 원인이 복잡하고 체계적인 측정이 어렵다.
- ④ 편향을 줄이기 위해서는 표본조사 전체 과정에 대한 세심한 관리가 필요하다.

정답 : ②

해설 : 편향은 그 발생 원인이 복잡하고 체계적인 측정이 어려울 뿐만 아니라 전수조사에서도 발생할 수 있다.

6. 다음의 설명 중에서 옳지 않은 것은?

- ① 핫덱방법(hot deck method)이나 콜덱방법(cold deck method)은 무응답 대체 방법 중 일종이다.
- ② 일관성 점검, 표본점검, 중복부차표본기법 등은 무응답오차를 줄이는 방법이다.
- ③ 비표본오차가 발생하는 요인으로는 무응답, 응답오차, 종합표작성의 오차 등이 있다.
- ④ 조사결과에 대한 표본오차는 추정량의 표준오차나 변동계수를 통해서 파악할 수 있다.

정답 : ②

해설 : 일관성 점검, 표본점검, 중복부차표본기법 등은 비표본오차를 측정하는 방법이다.

■ 정리하기

- 표본조사의 오차는 모수의 참값과 표본조사에서 추정된 값과의 차이를 말하며 오차는 편향과 우연오차로 구분된다.
- 편향(bias)은 표본통계량의 기대값과 모수의 참값의 차이를 뜻하며 비표본편

향과 표본편향으로 구분할 수 있고, 표본편향을 발생 원인에 따라 추출틀 편향, 일치성 표본 편향과 상수성 통계적 편향으로 구분할 수 있다.

- 비표본오차는 표본조사뿐만 아니라 전수조사에서도 발생하며, 종류는 표본 조사 준비 단계에서 생기는 무응답오차, 자료수집과정 중에 발생하는 응답오차와 자료분석 및 공표과정에서 생기는 종합표 작성의 오차 등이다.
- 비표본오차의 측정방법으로 일관성 점검, 표본점검, 사후 표본조사점검, 외부 기록 점검, 품질관리기법, 중복부차표본 방법 등이 있다.
- 중복부차표본은 연결부차표본과 독립적 부차표본으로 구분되며, 조사원의 업무의 질을 평가하는데 이용할 수 있는 기법이다.
- 통계조사에서 조사단위로부터 아무런 관찰값을 얻어내지 못한 경우를 단위 무응답이라 하고, 조사항목 중에서 일부는 응답을 구했으나 몇 항목에 대해서는 응답을 얻어내지 못한 경우를 항목 무응답이라 한다. 항목 무응답에 대한 대체 방법은 평균대체, 최근방대체, 회귀대체, 핫덱대체 등이 있다.

■ 참고문헌

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