

## 한국의 여론조사, 문제점과 대책은? (2)

### 휴대전화를 포함하면 조사의 품질이 향상되는가?

*조사회사가 보유한 휴대전화 패널은 통계학적으로 의미 없는 조사 결과 생산*

여론조사에서 가장 중요한 기본은 표본을 어떻게 추출(sampling)하느냐다. 여론조사가 통계학적으로 의미를 갖기 위해서는 모집단(population)을 대표할 수 있는 표본(sample)을 추출해야 한다. 즉 모집단에 포함된 모든 개인, 선거 여론조사의 경우에 해당 선거구의 유권자가 모두 같은 확률로 표본에 포함될 수 있어야 한다. 이러한 확률 표집(probability sampling)이 담보되어야만 통계적 추론이 가능하다. 반대로 확률 표집을 하지 않으면, 그 다음에 어떤 노력(예: 응답률 제고)을 하더라도 통계학적으로는 아무런 의미 없는 작업이 되어 버린다.

임의번호 걸기(random digital dialing, RDD)에 의한 표본 추출은 대표적인 확률 표집이다. 과거 표집에서 사용하던 전화번호부에 포함되지 않은 가구번호가 다수 존재하고, 집 전화를 보유하지 않은 가구·개인이 증가하면서 여론조사에 활용되는 빈도가 늘고 있다. 집 전화만을 이용한 여론조사의 표본 대표성(representativeness)에 문제가 제기되면서 휴대전화를 이용한 조사가 보편화되고 있는 것이다. 문제는 휴대전화 이용 표본을 어떻게 표집하느냐다.

미국과 우리나라는 휴대전화 번호 체계가 다르기 때문에 직접비교가 불가능하다. 그러나 미국 여론조사의 표본 추출 과정을 보면 집 전화만 있는 가정, 집 전화와 휴대전화가 모두 있는 가정(보다 구체적으로 집 전화와 휴대전화 중 주로 휴대전화 또는 집 전화를 이용하는 가정을 구분), 휴대전화만 이용하는 가정의 비율이 알려져 그에 따른 표집을 하고 있다. 전국단위 조사뿐 아니라, 주 단위 조사에서도 이 비율에 따라 조사에 가중치를 적용하여 결과를 발표한다.

반면 우리나라는 휴대전화 번호 체계가 지역을 포함하지 않기 때문에 지역단위 선거에서 휴대전화를 이용한 조사가 어렵다. 그럼에도 총선 전 여론조사를 보면 지역구 여론조사에서 휴대전화를 20~30 퍼센트 정도 포함한 조사 결과들이 보도되었다. 대부분은 조사회사에서 자체 모집한 패널의 휴대전화 번호를 이용한 것이다. 현재의 휴대전화 번호 체계로는 특정 지역에 살고 있는 유권자의 휴대전화 번호를 알 수 없기 때문에 쓰고 있는 방법이다. 이러한 패널들은 대부분 본인의 선택에 의해 패널에 참여한 경우다. 대표적인 비확률 표본이다.

위의 방식으로 모집된 패널과 RDD 전화조사 표본을 합하는 것은 비확률 표본과 확률 표본을 단순 결합시키는 것이다. 이는 여론조사의 기본 원칙을 전면적으로 위배하고 있기 때문에 그 어떠한 통계적 추론도 불가능하다. 정확한 조사를 위한다는 명목으로 조사를 더 악화시키고 있다. 혹시라도, 이러한 조사의 결과가 실제 결과와 유사했다고 하더라도 그 조사가 왜 정확하게 예측했는지를 설명할 수 없으며, 향후 조사에서 이 방법을 이용하더라도 결과가 반복될 가능성은 희박하다. 전국단위 여론조사에서는 휴대전화 RDD 조사가 가능하나, 지역단위 여론조사에서 패널을 이용한 휴대전화 표본을 포함하는 것은 오히려 조사의 품질을 저해할 수 있다.

휴대전화를 조사에 포함시키자는 논의는 휴대전화만을 이용하는 가구·개인에 대한 고려뿐 아니라 집 전화를 이용한 조사에서 젊은층이 조사에 잘 잡히지 않는 이유도 있다. 그러나 우리나라의 현재 조사관행을 개선하면 집 전화를 통한 조사에서도 젊은층을 충분히 확보할 수 있다.

휴대전화 조사의 편의를 위해 안심번호 도입을 논의하기 이전에 여론조사 과정에서 확률 표집의 원칙을 준수하도록 하는 것이 더 중요하다.

결론적으로 확률 표집이 이뤄지지 않는 상황에서 휴대전화를 조사에 포함하는 것을 논하는 것은 조사의 품질 향상과 아무런 관련이 없다.

\* 본 블로그의 내용은 연구진들의 개인적인 견해이며, 아산정책연구원의 공식적인 의견이 아닙니다.