

# 전화조사 오차요인과 개선방안

김 지 연\*

## 1. 문제 제기

여론조사에서, 특히 선거여론조사에서 전화조사가 차지하는 비중은 절대적이다. 일반적으로 선거여론조사는 크게 정당에서 공천이나 경선 목적으로 실시하는 여론조사<sup>1)</sup>, 후보자나 캠프 등에서 선거전략수립을 위해 실시하는 여론조사, 언론사에서 보도를 목적으로 실시하는 여론조사, 그리고 방송사 선거예측조사로 구분해 볼 수 있다. 이중 선거전략 수립조사에서 대인면접이나 좌담회(FGI), Ballot조사가, 그리고 선거예측조사에서 출구조사(Exit Poll)가 실시되고 있으나, 아직까지 주류는 전화조사라 할 수 있다. 한편, 2000년을 전후로 전화조사의 몇가지 문제점을 극복하고자 하는 취지에서 온라인조사, 모바일조사, ARS조사 등에 대한 연구가 활발히 진행되고 있지만, 아직 전화조사를 대체할 만한 수준은 아니라는 평가가 일반적이다.

본 글은 전화조사에서 발생할 수 있는 주요 오차요인을 살펴 보고 그 개선방안은 무엇인지 살펴보고자 한다. 대안이 아니라 개선방안이라고 설정한 것은 실제 조사과정에서 시간과 비용이라는 현실적인 제한요건을 고려할 수 밖에 없기 때문이다.

## 2. 표본추출 과정에서의 오차

표본추출 과정이 엄격하지 못해서 발생하는 오차중에서 가장 대표적인 3가지가 표본추출방

---

\*밀워드브라운미디어리서치

1) 2000년 16대 총선까지는 주로 중앙당의 공천조사가 주류를 이루었으며, 2004년 17대 총선에서는 중앙당 공천조사와 시/도당, 지구당 단위의 경선조사가 비슷한 수준에서 이루어 졌다. 2006년 4회 지방선거에서는 경선조사가 대부분을 차지하였다. 여기서 선거인단 모집 조사는 제외하였다.

법, 표집틀(Sampling Frame), 그리고 조사시간/요일 등과 관련된 것이다. 표본추출방법은 확률 표집을 하느냐의 문제이며, 모집단 커버리지의 문제는 표집틀과 조사시간대의 문제로 각각 전화번호부 등재율과, 특정계층 재택율과 연결되는 문제이다.

## 1. 표본추출방법

현재 여론조사회사에서 가장 대표적으로 사용하고 있는 방법이 성·연령·지역별 비례할당 추출법이다. 표본이 충분히 크고, 모집단이 전국이거나 광역단체일 경우는 이러한 3-Way Quota를 사용하나 적은 표본으로 소규모 지역을 조사하는 경우는 2-Way Quota를 사용하기도 한다. 조사목적에 따라 성, 연령, 지역이외에 다른 할당변수를 사용하기도 하지만 일반적이지는 않다. 다만, 선거여론조사에서 매우 중요한 변수인 연령의 경우, 기존의 10세 간격이 아니라 5세간격, 혹은 세대변수로 대체함으로써 보다 정확한 결과를 산출해 낼 수 있기도 하다. 50대와 60세이상 연령층을 합치거나 분리하여 할당을 부여하는 것도 모집단 특성에 따라 다르기 때문에 주의하여야 한다.

이러한 할당표집은 시간대별 재택율 등의 문제와 연결되어 표본의 대표성 유지에 어려움이 따른다는 문제가 있다. 당연히, 확률표집이 그 대안으로 언급되고 있다.

전화여론조사에서는 CNU나 Kish Grid보다는 상대적으로 간단한 birthday 방법이 선호된다.

【표1】 확률추출 실험결과

구분	최초 Call 결과	2회 Call Back 결과
조사시점	11일 11:30-22:00	12일 오후1차, 저녁 2차
접촉가구수	2,508개	2,508개
성공	11.0	19.8
통화중/비수신	44.3	16.2
응답자부재	14.9	3.2
가구단위거절	14.9	34.1
응답자거절	3.0	12.6
결번	9.6	9.6
기타	2.3	4.5

※ 조사지역 : 광주,  
 조사기간 : 2002년 5월 11일 - 12일,  
 표본추출방법 : Random Sampling(Birthday Method)  
 조사기관 : 미디어리서치

표1은 유효표본 500개를 목표로 미디어리서치에서 실시한 확률추출 실험결과이다. 2회 콜백을 실시한 결과 성공률이 약 9%정도 높아졌으며, 통화중/비수신 비율이 28%정도 낮아졌다. 반면, 유효표본을 채우기 위해서는 일반적인 할당표집 500명 조사의 경우 보다 2.5배 정도의 비용과 기간이 더 소요되었다. 정확성과 효율성의 문제가 제기되는 대목이다. 매우 정확한 결과가 요구되는 선거예측조사나 시간과 비용에 상대적으로 제한을 덜 받는 연구목적의 조사에는 이러한 확률표집이 적합할 수 있으나, 현재 가장 많이 실시되고 있는 정당이나 후보자, 언론사 조사 등에서는 사용하기 힘든 측면이 있다.

## 2. 표집틀(Sampling Frame)

현재 사용하고 있는 전화조사 표집틀은 전화번호부이다. 과거에는 Paper 전화번호부를 사용했기 때문에 법정동 중심의 분류체계로 인한 지역할당의 어려움과 면접원에게 엄격하게 Skip interval을 지키도록 통제하는 것이 어려웠으나, 현재는 대부분의 여론조사회사에서 전화번호 DB를 이용하기 때문에 응답가구 추출단계까지의 무작위성은 확보가 가능한 상황이다.<sup>2)</sup>

문제는 전화번호 등재율이다. 정확한 통계가 나오지는 않으나 한국통신과 전화번호부주식회사 관계자에 따르면 전화번호 등재율이 70%를 밑도는 수준이라고 한다. 여기에다 이동전화 보급률이 높아지면서 1~2인 가구의 유선전화 보급률이 떨어지고 있는 추세임을 고려할 때, 실제 전화조사 모집단 커버리지는 60%를 밑돌 가능성도 존재한다.

이러한 전화번호 미등재 문제는 RDD방법을 도입하면 해결될 수가 있다. 현재 우리나라 전화번호부 체계상의 문제로 인해 몇가지 어려운 점<sup>3)</sup>이 있기는 하지만, 장기적으로는 RDD방법 도입이 필수적일 수 밖에 없는 것으로 판단된다. 전화번호 DB에서 인명편 가구의 국번을 추출하고, 업종/상호편 전화번호를 제거한 후 남은 숫자를 가지고 랜덤하게 전화번호를 생성하여 추출하는 방법이 현재 연구중이다. 컴퓨터가 직접 전화를 걸어 연결되는 전화번호만 면접원에게 배분하는 Auto-dialing 시스템을 도입하면 효율성이 더욱 높아질 수 있다. 문제는 이 역시 비용과 시간의 문제이다. RDD 도입 초기에는 표본손실이 일반조사에 비해 10배이상 높을 것으로 추정하고 있다. 실제 조사과정을 통해 결번 전화번호에 대한 지속적인 제거 작업을 수행할 경우, 2~3년 후에는 효율성이 상당히 높아질 것으로 예상된다.

---

2) 2004년 총선을 전후로 주요 여론조사회사를 중심으로 CATI가 도입됨으로써 전화조사 응답가구추출의 엄격성이 높아졌다.

3) 국내 전화번호 숫자에 (소)지역과 업종/인명에 대한 정보가 정확하게 들어 있지 않다.

### 3. 재택율

그 다음은 계층별로 재택률의 차이에 의해 발생하는 문제이다. 특히 조사가 완료되는 밤 10시 까지도 귀가하지 않는 계층들은 원천적으로 조사대상에서 제외되는 문제가 발생한다.<sup>4)</sup> 현재 대부분의 조사 회사에서 시행하고 할당추출법이 이러한 문제를 더 심화시키고 있다. 할당추출은 성, 연령, 지역 할당을 채우는 방식이므로 응답자 대체 문제는 별로 민감하게 고려하지 않는다. 따라서 직장인 20, 30대가 채워지지 않으면, 집에 있는 20, 30대를 조사하면 된다. 저녁 18시부터 조사를 진행하거나 주말에 조사를 진행하면 이러한 문제는 일정 부분 해결될 수 있으나, 이 역시 효율적이지 못하다.<sup>5)</sup>

이러한 문제들을 해결하기 위해 확률추출에서 사용하는 Call-back을 고려해 볼 필요가 있다. 즉, 기존의 할당추출방법을 유지한 채, 전화번호 배부 체계를 개선함으로써 이러한 문제점을 개선할 수 있을 것으로 판단된다. 이 방법은 고려대학교 통계학과 허명희 교수가 처음 제안한 것으로 이번 제 4회 KBS-SBS 지방선거 예측조사에서 일부 사용된 방법이다..

【표2】 Call-back 실시와 미실시의 응답자 특성

대전시장 선거	미실시			Call-back 실시		
	7시29분이전	7시30분이후	계	7시29분이전	7시30분이후	계
19~29세	0.22	0.18	0.21	0.19	0.40	0.23
30~39세	0.21	0.28	0.23	0.23	0.34	0.25
40~49세	0.24	0.21	0.23	0.24	0.21	0.23
50~59세	0.15	0.16	0.15	0.17	0.02	0.14
60세 이상	0.18	0.17	0.18	0.18	0.03	0.15
계	1.00	1.00	1.00	1.00	1.00	1.00

이 방법은 낮 시간대에 부재중/통화중인 전화번호를 Call-back하는 것이다. 오후 13시~19시 경까지 사용한 전화번호중 부재중/통화중 전화번호<sup>6)</sup>만을 따로 보관하였다가 저녁 19시30분 이후 Call-back을 실시함으로써 낮 시간대에 집에 없는 직장인들에 대한 접촉이 개선될 수 있을 것이라는 가설이 그 출발점이었다.

4) 주로 직장여성, 학생, 20대, 30대 남자 등의 조사시간대(13시~22시) 재택율이 떨어진다.  
 5) 1차적으로 비용증가의 문제가 발생한다. 또한, 주5일근무 이후 국민들의 라이프스타일이 변화되면서 야외활동하는 계층이 체계적으로 빠지는 문제가 발생한다.  
 6) 2006년 미디어리서치 실험 결과, 낮시간 부재중 비율은 조사지역, 시간 등에 따라 다르지만 대체적으로 25% 전후 나타난다.

표2는 2006년 5월초 동일한 시점, 동일한 지역(대전)을 대상으로 밀워드브라운미디어리서치가 Call-back을 실시한 결과이다. 20대와 30대의 응답율이 높아지는 특징을 보였다. 다만, 이러한 가설이 검증되어 일반화되기 위해서는 보다 많은 조사경험이 축적되어야 할 것으로 판단된다.

### 3. 조사과정에서의 오차

조사과정에서의 오차는 흔히 말하는 비표집오차와 관련되어 있다. 조사방법 설계, 할당표 작성, 질문지 설계, 분석 설계, 조사도구 준비, 면접원 교육, Field work, 검증, 에디팅/코딩, Data Punch, Data Cleaning, Programming, 가중치 부여, 보고서 작성 등 조사 전과정에 걸쳐 발생할 수 있다. 이중 가장 대표적인 2가지가 질문지 작성과 관련한 오차와 면접원과 관련한 오차라 할 수 있다.

#### 1. 질문지 작성

질문지 작성시 가장 빈번하게 발생하는 문제는 질문항목의 배열, 질문 워딩, 보기 구성 등 일반적으로 생각되는 것이 아니라 질문이 너무 많거나 어려운 문제이다. 물론, 질문 순서나 워딩 등은 매우 중요하며 Bias가 들어갈 경우, 심각한 결과왜곡을 가져다 준다. 하지만 일반적인 선거 여론조사에서 이러한 오류를 범하는 경우는 드물다. 그것 보다는 응답자가 도저히 대답할 수 없는 난이도 높은 질문을 하거나 거의 면접조사 분량의 질문을 전화조사를 통해 해결하고자 하는 시도 등이 문제이다. 문제는 어려운 질문을 하더라도 응답이 나온다는데 있다. 국가 정책과 관련하여 전문가들이 응답하기도 힘든 질문에 일반 국민들도 쉽게 응답을 한다. ‘대충 응답하기’ 또는 ‘지지정당 따라가기’가 빈번하게 나타나는 것이다. 처음 통화하는 낯선 여자에게 ‘잘 모르겠다’라고 응답하면 체면이 안선다고 생각하기 때문인 것으로 추정된다. 면접시간이 긴 질문은 더욱 문제이다. 중요한 질문은 앞으로 배치하고 덜 중요한 질문은 뒷부분에 배치하면 된다는 시각은 잘못된 것이다. 면접시간이 길면, 조사중단이 빈번하게 일어나고, 조사원들은 더욱 빨리 질문을 읽거나 건너뛰게 된다는데에 문제가 있다. 부득이하게 긴 질문을 소화해야 할 경우, 응답자에게 인센티브를 제공하는 것이 필수적이다.

질문을 구성할때는 다음과 같은 원칙이 준수되어야 한다.

- 꼭 필요한 내용인가?,
- 응답자가 필요한 정보를 알고 있는가

- 응답자가 그 정보를 제공해 줄 수 있는가?
- 쉽고 의미가 명확하게 구분되는 단어를 사용했는가?
- 포괄성과 상호배타성의 원칙이 지켜졌는가?
- 하나의 질문으로 두가지 정보를 얻으려 하지는 않는가?
- 응답자가 대답하기 곤란한 사항에 대해 직접적으로 질문을 하지는 않는가?
- 유도성 질문은 아닌가?
- 질문순서나 배치는 적당한가?

## 2. 면접원 선발 및 관리

실제 실사과정을 책임지는 면접원의 중요성은 아무리 강조해도 지나치지 않는다. 조사원 모집 및 관리 시스템, PAY 및 인센티브 제도, 불량면접원 제거 시스템, 검증 방법/주체, 면접원 연령/학력/출신지역/정치적 성향 구성비율, 면접원별 무응답률 관리 방법, 할당 배부 및 채우는 방식, 면접과정 통제(비표준화된 면접, 면접원의 임의성 개입, 응답유도, 면접원별 probing 강도의 차이, 래포형성, 면접원 역량 등) 등에서 오차가 발생할 수 있다.

면접원과 관련된 비표집오차는 적절한 수준의 PAY체계, 상시적 교육과 모니터링, 검증, 평가 등으로 줄일 수 있다. 그중에서도 상시적·반복적 교육을 통한 면접과정의 표준화, 면접과정 및 응답결과에 대한 엄격한 평가를 통한 상벌제도 확립 등이 효율적이다.

## 4. 응답자 오차

오차의 3대 주체, 즉 연구원, 면접원, 응답자 중 응답자 오차가 특히 중요한 것이 바로 선거여론조사이다. 일반적인 여론조사와 달리 최근의 선거여론조사에서는 일부 응답자들이 자신들의 응답을 통해 선거결과에 영향을 미치기를 시도한다는 점에서 더욱 그러하다. 과거 군사정부 시절이나 YS, DJ 집권시절까지만 하더라도 자신의 정치적 입장을 밝히기 꺼려하는 현상이, 혹은 자신에게 피해가 오지 않을까하는 우려에 기반한 ‘소극적 불성실 응답’ 차원이라고 한다면, 2002년 16대 대선 당시 盧-鄭 후보단일화 과정에서 제기된 역선택 논쟁은 응답자들의 ‘적극적 불성실 응답’ 현상의 단초를 제공하였다. 이후 2004년 탄핵국면에서 노인평화발언으로 연결되는 17대 총선정국에서도 이러한 현상이 두드러지게 나타났다. 이러한 현상은 조사참여 및 거절 상황과 주요 질문에 대한 응답과정에서 두드러지게 나타나고 있다.<sup>7)</sup>

## 1. 투표율의 차이

여론조사에서 투표의사는 전형적으로 ‘사회적 요망 효과(Social Desirability Effect)’가 반영된 질문이다. 투표의사 조사결과와 실제 투표율의 상관관계가 생각만큼 높지는 않기 때문에 선거 예측에서 후보지지도 무응답 분류 보다 투표예정자(Likely Voters)를 파악해 내는 것이 더 어렵다.

【표3】 역대 선거 투표의사(적극투표의사층)와 실제투표율

		D-60	D-30	D-20	실제투표율
총선	2000년		57.3	65.3	57.2
	2004년	47.4	51.1	72.9	60.6
대선	1997년	74.9	79.4	82.5	80.7
	2002년	69.9	74.1	78.1	70.8
지방선거	1998년		63.4	62.5	52.7
	2002년	45.1		57.5	48.8
	2006년	51.9	59.8	62.8	51.6

※ 출처 : 미디어리서치 선거솔루션

표3에서 살펴볼 수 있듯이 선거의 성격 및 종류에 따라 다소 차이는 있지만, 대체적으로 ‘적극 투표의사층’ 비율이 실제 투표율을 보다 5%~15% 정도 높게 나타나는 것이 일반적이다. 현재와 같이 연령대별 지지성향의 차이가 뚜렷한 상황에서 특정 연령층의 투표참여/기권이 전체 선거 판세에 영향을 미치기 때문에 이러한 실제 투표율 추정작업이 매우 중요하게 된다. 저연령층에서 고연령층으로 갈수록 투표율이 높아지는 현상은 현재까지 유효한 설명력을 갖는다. 2002년 대선당시의 미군장갑차사건, 월드컵, 2004년 총선당시의 탄핵, 노인평화발언 등 전국을 강타하는 이벤트가 발생할 경우, 특히 이러한 이벤트가 지역 보다는 연령에 더 자극적인 이슈일 경우, 약간 달라질 수 있으나, 이 경우에도 이러한 원칙이 크게 벗어나지는 않았다.<sup>8)</sup>

투표의사의 ‘사회적 요망 효과’ 문제를 해결하기 위해서는 다음과 같은 두가지 방법이 가장 대표적으로 사용된다. 먼저, 질문항목을 통해 가려내는 것이다. 투표의사 질문이외에 선거관심

7) 후보지지도와 투표의사 질문에 대한 응답과정에서 거짓응답을 하거나 응답을 하지 않는 현상을 말한다.

8) 중앙선거관리위원회 17대 총선 투표율 분석자료에 의하면, 50대가 가장 높은 74.8%, 그 다음이 60대 이상 71.5%, 40대 66.0%, 30대 후반 59.8%, 30대 전반 53.2%, 20대 전반 46.0%, 20대 후반 43.3% 순으로 나타났다.

도를 측정할 수 있는 관여도 질문을 삽입하고, 과거 투표여부 질문과 인구통계학적 변수간의 조합을 통해 투표예상자를 추출하는 방법이다.

다음으로, 중앙선관위가 발표하는 과거 선거에서의 성별·연령별·지역별 투표율 자료를 이용하여 가중치를 부여하는 방법이다. 선거종류와 당시 선거 진행상황에 따라 이 두가지 방법중 하나를 사용하기도, 아니면 두 개를 한꺼번에 사용하기도 한다.

## 2. 응답불성실 효과

응답자가 지지도 측정 질문과 관련하여 지지후보를 밝히지 않거나 거짓응답을 하는 경우, 전화조사의 오차를 크게 하는 요인이 된다. 이러한 불성실 응답자가 특정후보 지지층이라면, 이러한 편향은 더욱 커지게 된다. 응답불성실 효과와 관련된 편향은 선거종류, 조사시점, 프로빙 강도, 후보구도, 판세, 선거테마 등 다양한 원인에 의해 발생된다. 따라서 보다 정확한 조사를 위해서는 다양한 사례분석을 통해 불성실 효과의 '발생 여부'와 '크기', '방향'에 대한 경험적 모형을 개발하는 것이 필요하다.

이러한 응답불성실 효과를 측정하기 위한 방법중에 하나가 지지도 관련 질문을 두 개 정도 던진 후 그 일관성 정도를 체크하는 방법이다. 즉, 불일치 비율이 높을 수록 응답불성실 효과가 큰 것으로 간주하는 것이다.

## 3. 조사거절자 편향 효과

특정성향의 계층이 체계적으로 조사를 거절했을 때, 발생하는 편향을 말한다. '여당과대계상 편향', '보수주의자 회피효과', '침묵의 나선 효과' 등으로 나타난다.

【표4】 KBS-SBS 선거예측조사의 정당별 편파 평균<sup>9)</sup>

구분	17대 총선(2004년)	4회 지방선거(2006년)
한나라당 편파	-1.2(2위 정당)	+1.6(1위 정당)
열린우리당 편파	+1.4(1위 정당)	-0.7(2위 정당)

표4에서 살펴볼 수 있듯이 현재의 조사거절차 편향은 주로 침묵의 나선이론과 연결되고 있다.

9) 밀워드브라운미디어리서치와 TNS가 컨소시엄으로 예측조사를 수행하였으며, 17대 총선은 243개 지역구 결과를, 4회 지방선거는 16개 시/도 결과를 평균한 것이다.



【표5】 계층별 전화조사 거절률<sup>10)</sup>

성/연령	거절률	지역	거절률
남자	44.2	서울	64.7
여자	62.6	인천/경기	49.4
20대	35.4	강원	41.4
30대	53.7	충청	50.5
40대	58.4	호남	51.6
50대	55.6	TK	52.6
60세이상	63.8	PK	45.1

표5에서 살펴볼 수 있듯이 주로 여자, 60세 이상 고연령층, 서울에서 거절률이 높은 것으로 나타났다. 가급적 최초 접촉자를 계속 설득하여 조사를 성공하는 것이 바람직하지만, 실사 과정에서 이러한 문제를 해결할 수 없을 경우는 거절에 따른 계층별 가중 계수를 개발하여 보정하는 방법도 좋을 것이다. 이밖에 응답자에게 인센티브를 실시함으로써 응답성실성 향상과 함께 조사거절률을 낮추도록 유도할 수 있을 것이다.

## 5. 글을 맺으며

현재까지 15번의 방송사 선거예측조사가 진행되면서, 방법론적으로 많은 발전을 이루었다. 총선과 같이 당선자 예측에 실패를 해서 크게 비난을 받았던 경우도 있었고, 대선과 같이 매우 정확한 예측을 수행한 적도 있었지만, 전체적인 방향은 오차가 감소하는 쪽이었다.<sup>11)</sup> 이러한 선거예측조사를 수행하면서 방송사간, 조사회사간에 치열한 경쟁이 전체 여론조사의 질 향상에 큰 공헌을 했다는 것은 주지의 사실이다. 하지만 해결되지 않고 계속되는 문제점은 조사의 정확성과 비용/시간대비 효율성의 적절한 콤비네이션이 어느 수준인가라는 것이다. 방송사와 여론조사협회, 그리고 학계가 공동으로 힘을 모아 보다 정확한 조사를 할 수 있는 물적/인적 토대가 마련되기를 기대한다.

10) 미디어리서치가 2002년 12월 전국 2,000명 조사시 실험한 결과이다.

11) 방송사 선거 예측조사 당선자 예측오차 평균  
 - 총 선 : 15대 6.1% → 16대 4.4% → 17대 2.7%  
 - 지방선거 : 2회 3.9% → 3회 3.1% → 4회 2.7%