

# CHAPTER 1

---

## 설문 조사 개요

### 1.1. 설문 조사

#### 1.1.1. 정의

A survey is a system for collecting information to describe, compare, or explain knowledge, attitudes, and behavior. –Survey Kit, Sage Publication-

설문 조사는 조사자가 관심을 갖는 특성이 의견이나 행동 양식과 같이 **측정이 불가능한** 경우 적절한 설문 문항을 작성하여 원하는 정보를 얻어 내는 일련의 과정이다.

조사자가 관심을 갖는 사회 현상, 특별히 사람이 주체가 되어 만들어내는 사회 현상에 대한 정보를 얻기 위하여 자료를 수집, 분석하는 것을 설문 조사라 한다.

설문 조사는 사전 지식이나 선향 연구가 없을 때 사회 현상을 탐색하거나, 사회 현상을 기술하고 설명하기 위하여 모집단(관심 집단)이나 표본 집단을 대상으로 연구 조사를 시행하는 것을 의미한다.

사회 조사라고도 불리는 설문 조사는 19 세기 후반부터 사회 과학자들에 의해 사용되기 시작하였다. Karl Marx(1880)는 1880 년, 2 만 5 천명의 프랑스 노동자를 대상으로 그들의 정치적 태도와 성향에 대한 우편 조사를 시행하였고 Max Weber 는 직접 관찰과 조사 연구를 통하여 노동자들의 심리, 태도, 직업관 등을 연구하였다. 20 세기에 들어오면서 미국 사회학자

들에 의해 사회 조사 방법론 연구가 활발하게 진행되었으며, 미국 통계국(Bureau of Census)은 표본 추출 방법과 자료 수집 방법에 대한 연구에 집중하였고, Gallup 과 Roper 와 같은 여론 조사 기관은 방법론을 발전시켰다.

### 1.1.2. 설문 조사자 자격

설문 조사를 실시하려는 사람은 다음의 능력을 갖고 있어야 한다.

(1)관심을 갖는 사회 현상에 대해 폭넓은 지식이 있어야 한다. 우리 나라에 가장 적절한 대학입시 방법에 대해 국민들의 의견을 조사하고자 한다고 가정하자. 이에 대한 조사 연구를 위해서는 교육에 대한 전반적인 이해가 있어야 하고, 조사자가 필요하다면 교육 전문가, 대학 신입생, 고등학생들의 견해를 듣는 노력이 필요하다.

(2)설문 조사 전 과정 이해

계획 수립부터 보고서 작성 단계까지 일련의 과정을 직접 수행하거나 지시할 수 있는 능력이 있어야 한다.

(3)설문지 작성 능력

응답자가 사람이므로 사람들에 대한 이해가 필요하므로 심리학, 조사 관련 분야에 대한 지식을 이용하여 조사 목적에 맞는 적절한 설문지를 작성할 수 있어야 한다.

(4)표본 추출 방법에 대한 지식

표본 설문 조사의 경우 모집단의 축소판인 표본을 추출할 수 있는 능력이 있어야 한다.

(5)통계 소프트웨어 활용 및 분석 결과 해석 능력

자료 코딩 방법, 적절한 분석 방법인지, 분석 결과에 대한 해석 능력이 있어야 한다.

좋은 설문 조사가 되려면 다음 사항에 유의해야 한다.

(1)Specific objectives

조사 목적을 명확하게 기술하여야 한다.

(2)Straight-forward questions

응답자의 생각을 정확하게 조사할 수 있도록 설문 문항을 적절히 구성해야 한다.

**(3)Sound research design**

적합한 조사 방법 및 표본 추출 방법을 선택해야 한다.

**(4)Reliable and valid survey instruments**

적절한 자료 분석 도구 및 결과 표현 기법을 이용해야 한다.

**(5)Reasonable resources**

조사 비용 예상, 조사원 선발 및 교육, 조사 가능성 진단을 할 수 있어야 한다.

**1.1.3. 설문 조사 예제****(1) 여론 조사 (poll)**

(예제)당신이 선택할 대통령/국회의원은?

(예제)OO 법안에 대한 국민 의견 수렴: 원전 핵 폐기물 시설 입지 선정 (cf.) NIMBY,  
PIMBY

**(2) 기업 경영 관련 조사**

(예제)소비자의 제품 선호도나 만족도

(예제)소비자 구매 패턴 조사

(예제)OLTP (On-Line Transaction Process) → Data mining, CRM

**(3) 학문적 연구 관심: 측정이 불가능 항목에 대한 계량적 연구**

(예제)대학 시설에 대한 학내 구성원 만족도

(예제)아동 복지 시설 현황 조사

설문 조사에서 조사되는 항목은 다양하지만 다음과 같이 정리될 수 있다.

**(1)인구학적 특성**

성별, 연령, 학년, 부서, 거주 지역, 가족

(예제)학년별/성별 학교 만족도의 차이가 있는가?

**(2)응답자 의견을 보기 항목이나 주관적 기술로 응답하게 한다.**

(예제)당신의 소득은?

▪ 개방형 문항: ( )만원

▪ 폐쇄형 문항: (1)200만원 미만 (2)200-300만원 미만 (3)300만원 이상

(3)선호도

보기 문항에 대해 응답자의 선호하는 순서를 조사한다.

(예제)다음 장소 중 신혼 여행 후보지로 마음에 드는 순서대로 우선 순위를 매기시오.

①미국/캐나다( )    ②유럽( )    ③호주/뉴질랜드( )

④동남 아시아( )    ⑤국내( )

(4)만족도

응답자가 만족하는 정도를 조사한다. 각 보기는 내용에 따라 점수화 하게 된다.

(예제)리커드(Likert) 척도 문항

## 1.2. 설문 조사 순서

(1)조사 목적 및 계획 수립

(2)표본 프레임 및 표본 응답자 추출 방법

(3)조사 방법 선택

(4)설문지 작성

(5)사전 조사(pilot survey, preliminary survey, pretest) 실시

(6)조사원 모집 및 교육

(7)설문 조사 실시

(8)수집 설문지 정리 및 자료 코딩

(9)설문 분석 및 결과 보고서 작성

다음의 절에서 각 순서를 상세히 알아보자.

### 1.2.1. 조사 목적 및 계획 수립

(1)누구를 대상으로 (2)어떤 조사 방법으로 (3)어떤 내용을 설문할 것인지 포함하고 있어야 한다. 이것은 설문 조사의 첫 단계이자 가장 핵심적인 단계이다.

어떤 집단으로부터 어떤 정보를 얻을 것인가?	▶ 조사 대상: 표본 프레임 ▶ 무엇을 물을 것인가: 설문 항목 결정
-----------------------------	---

조사 목적에는 연구 가설(research hypothesis) 설정도 포함되어 있어야 하며 이에 의해 설문지가 작성된다. 설문 조사 분석 결과를 이용하여 연구 가설을 역으로 설정하거나 조사 목적 수립 시 결론을 미리 내놓는 것은 좋은 설문 조사가 아니며 조사가 왜곡될 가능성이 있으므로 피하기 바란다.

조사 목적이 설정되면 focus group(설문에 이해 관계가 있는 사람)이나 consensus panel(설문 조사 관련 전문가 집단)을 운영하여 좋은 설문지가 작성 될 수 있게 힘써야 한다. 그리고 조사 하려는 분야와 유사한 기존 설문 조사에 대한 문헌 연구도 필요하다.

### 1.2.2. 표본 프레임 및 표본 추출 방법

설문 조사 대상이 되는 전체 응답자의 리스트를 표본 프레임(sample frame)이라 한다. 즉, 조사 대상이 되는 응답자 전체와 그 응답자를 구별할 수 있는 ID(예: 주민 번호, 학번, 사원 번호)를 리스화 한 목록을 표본 프레임이라 (sampling frame) 한다. 조사 대상에 따라 표본 프레임이 규정될 수도 있고 그렇지 않을 수도 있다. 표본 프레임 작성 시 제외된 응답자는 표본 추출 대상에서 제외되어 조사가 왜곡될 가능성이 있으므로 적절하고 완전한 표본 프레임 작성은 매우 중요하다. 조사 목적에 의해 응답 대상자가 결정되지만 현실을 고려하여 결정해야 한다.

(예제)OO 회사 직원 만족도 조사: 회사 직원 리스트가 표본 프레임이다.

(예제)OO 백화점 고객 만족도 조사: 백화점 카드 소지자 명부를 표본 프레임으로 이용할 수 있다.

조사 목적에 의해 표본 프레임이 결정되나 현실적으로는 표본 프레임을 규정하기 어려울 때가 있다. 이처럼 응답 가능 대상자 리스트가 현실적으로 불가능 할 때 이를 “conceptual” 응답자 집단이라 한다. 예를 들어 가정 폭력 피해자 실태 조사나 AIDS 관련 연구자의 연구

실태 조사의 경우 응답 가능 대상 규정이 모호하다. 이런 경우 표본 프레임 작성은 1) snowball 방법에 의지하거나 (상당소에 상당하거나 경찰에 신고하는 가정 폭력 피해자를 대상으로 설문 조사) 2) 일부 제한된 정보를 이용하게 (예: 학회에 등록된 연구원 대상) 된다.

고객을 백화점 카드 소지자로만 한정하지 않는 경우 표본 프레임이 명확하게 규정될 수 없다. 이런 경우 표본 프레임은 조사 일에 백화점을 찾는 고객이 표본 프레임이다.

(예제)한남대학교 학생들의 학교 만족도 조사: 응답 가능 대상자는 2003년 3월 현재 등록된 학생으로 규정할 수 있다. 이 경우 표본 프레임은 학생들의 학번이다.

(예제)대통령 선거 당선 후보 예측 여론 조사: 유권자 명부가 표본 프레임이다.

학교 만족도 조사나 대통령 선거 여론 조사의 경우 표본 프레임은 존재하지만 설문 응답자로 추출된 응답자를 찾아 설문을 실시하는 것은 시간적 경제적 효율성이 떨어지거나 실제 불가능한 경우가 발생한다. 이런 경우 표본 프레임을 이용하기 보다는 적절한 표본을 추출할 방법을 찾아야 한다. 예를 들어 여론 조사의 경우 전화 번호 명부를 표본 프레임으로 이용한다. 설문 응답자 표본 추출 방법에 대해서는 3장에서 다루기로 한다.

### 1.2.3. 조사 방법 선택

설문 조사 방법은 면접 조사(일대일, 그룹), 우편 조사, 전화 조사, 컴퓨터(CASI: Computer-Aid-Self-administrated Interview) 조사, 인터넷 조사 등이 있다. 조사 방법은 조사의 긴급성(시간), 비용, 회수율, 조사 중요도 등을 고려하여 결정하게 된다. 예를 들어 신속한 조사를 원하면 전화 조사를, 설문지 문항이 다소 전문적인 경우는 응답자가 충분한 시간을 가질 수 있는 우편 조사 방법을 선택하게 된다. 조사 방법에 대해서도 3장에서 자세히 다루기로 한다.

### 1.2.4. 설문지 작성

인쇄물을 이용하거나 컴퓨터 (Computer Aid Self-administrator Interview) 혹은 인터넷 조사를 사용할 수 있다. 설문지는 측정 도구이므로 설문지 작성에 오류가 있거나 응답자가 솔직하게 응답을 할 수 없는 설문 문항이 작성되면 응답자의 성향이나 의견을 정확하게 측정할 수 없다. 그러므로 좋은 설문지 작성에 가장 많은 시간을 할애하며 여러 전문가의 의견을 수렴하여야 좋은 설문 조사를 할 수 있다.

설문 목적에 따라 작성된 리스트를 이용하여 적절한 설문 문항을 작성한다. 관심 그룹, 전문가 그룹 등을 이용하면 보다 나은 설문지를 얻을 수 있고 문장, 어휘 선택은 응답 대상 속성, 조사 방법에 따라 선택해야 한다. 설문지는 설문 조사의 측정 도구이므로 정확한 응답을 얻어 낼 수 있도록 작성되어야 하며 조사 목적에 합당한 질문 문항들로만 구성되어야 한다. 설문 조사 관련 기존의 신뢰할 만한 설문지가 있으면 참고하는 것이 좋으나 외국에서 사용된 설문지는 문화적 차이를 고려하여 사용할 때 신중을 기해야 한다. 좋은 설문지 작성에 대해서는 2장에서 다루기로 한다.

### 1.2.5. 사전 조사(pilot survey, preliminary survey, pretest) 실시

설문지를 작성 하는데 아무리 많은 시간과 노력을 들여도 작성된 설문지에 오자나 탈자가 있거나 응답자가 오해하거나 답하기 어려워하는 부분이 있을 수 있다. 이런 오류를 발견하기 위하여 조사 대상과 유사한 집단이되 조사자가 신뢰할 만 하거나 응답에 매우 호의적인 사람들(적게는 20 명, 많게는 50 명)을 대상으로 사전 조사를 실시한다. 설문 조사 연구자가 직접 실시하는 것이 좋으며 그룹 응답 보다는 일대일 면접 조사를 하는 것이 좋다. 사전 조사 응답자들이 응답 중 의문점이나 제안 사항이 있으면 언제 자신들의 의견을 말할 수 있는 분위기가 갖추어져야 하며, 조사자는 글들의 의견을 반영하여 설문지를 수정하게 된다. 사전 조사를 통해 수정된 설문지가 최종 설문지가 되며 본 조사에 이용하게 된다. 다음 사항을 알아보기 위하여 사전 조사를 실시하게 된다.

- (1)응답 시간은 어느 정도 걸리는가?
- (2)설문 질문 & 문항 보기를 다 읽은 후 답하는가?
- (3)문항이 너무 전문적이거나 이해가 어려운 질문은 없는가?
- (4)응답을 꺼리거나 거짓으로 답변할 가능성이 있는 문항은 없는가?
- (5)질문 응답자에 따라 자의적으로 해석될 가능성은 없는가?
- (6)예상하지 못한 보기가 존재하지는 않는가? 기타 문항을 사용하여 괄호 안을 적게 함.
- (7)문법적 오류나 오자가 없는가?
- (8)응답자가 설문 응답하는 분위기 파악

사전 조사는 설문 조사지의 신뢰성을 높이기 위해 실시하므로 사전 조사 응답자 선정 시에는 표본 추출 방법에는 신경 쓸 필요가 없다. 설문 응답 대상자들과 유사한 성향을 가진 사람들 중 조사자가 신뢰할 만한 사람을 대상으로 실시하면 된다.

### 1.2.6. 조사원 모집 및 교육

설문지가 측정 도구라면 조사원은 설문지라는 측정 도구를 이용하여 사람들의 생각을 측정하는 사람이다. 아무리 좋은 측정 도구라도 측정자가 무성의 하다면 정확한 측정이 되지 않는 것처럼 조사원의 역할도 이와 같다. 또한 조사원의 외모나 억양이 응답자의 응답에 영향을 미치므로 신뢰할만한 조사원 선정 및 조사 방법에 대한 철저한 교육이 필요하다.

설문 응답 시 응답자(respondent)들은 조사원의 외모, 말씨, 태도 등에 영향을 받으므로 조사원의 영향을 최소화 하고 조사원 간의 편차(bias)가 생기지 않도록 하기 위한 조사원 선발과 교육이 필요하다. 조사원이 여럿인 경우는 각 조사원의 영향을 최소화 할 수 있도록 조사원을 선발해야 하며 각 조사원의 설문 지역 배정도 주의해야 한다. 예를 들어 ○○대와 ○○대의 학생 학교 만족도 조사 시 A 조사원 팀은 ○○대, B 조사원 팀은 ○○대만 조사할 것이 아니라 두 조사 팀이 각 학교를 모두 조사할 수 있도록 응답 인원을 반씩 배정하는 것이 좋다.

조사원이 조사에 성실히 임할 수 있도록 충분한 시간과 비용을 배정해야 하며 전화 조사 시에는 조사원의 말투와 조사 하나에도 영향을 받으므로 설문 문항은 물론 인사말까지 작성해 주고 말 토씨 하나까지 그대로 읽도록 하고 늘 같은 음성으로 전화할 수 있도록 한다. 예를 들어, 대전 시민 공공 교통에 대한 만족도 조사를 이틀 동안 대전 시내 지역에서 시민을 대상으로 실시했다고 하자. 첫 날과는 달리 둘째 날 ○○단체 시위로 인해 교통 체증이 심했다는 사실을 조사원이 기록하지 않거나 알려주지 않았다면 왜곡된 결과를 발표하게 될 것은 자명한 일이다.

설문 조사 분석 결과를 해석하는데 조사 과정에서 발생한 모든 일들은 매우 귀중한 정보이므로 조사원들에게 조사 환경(기후, 장소, 분위기), 응답자가 불성실한 경우 설문지에 표시하는 등 조사 과정에서 발생한 일들을 적게 하여 분석 결과를 해석하는데 이용하여야 한다.

### 1.2.7. 설문 조사 실시

언제 조사할 것인가 (시간) 어디서 조사할 것인가 (장소) 결정하는 문제 또한 중요하다. 무슨 요일에 조사하느냐, 하루 중 언제 조사하느냐에 따라 설문 회수율, 응답 성향 등이 달라



진다. 학교 만족도 조사를 학교 교정, 도서관, 길거리 장소에 하느냐에 따라 응답 결과도 달라질 것이다. 장소 선택 문제는 응답 대상자 전체를 대표하는 적절한 표본 응답자를 얻을 수 있느냐가 중요한 변수로 작용한다. 예를 들면 ○○대학생 의식 조사를 위해 응답 장소를 도서관에서 하느냐 정문에서 하느냐 휴게실에서 하느냐? 이는 조사 목적에 따라 다르다. 면학 분위기 관련 조사는 도서관에서, 학교 편의 시설 만족도 조사는 휴게 공간에서, 학교 만족도 조사는 정문이나 각 단과 대학 건물 앞에서 실시하는 것이 적절하다.

실제 조사를 나서는 예상하지도 못한 문제가 조사를 방해한다. 응답자들의 답변 회피, 좋지 않은 날씨, 부적합한 설문 장소 등 조사에 영향을 미치는 상황이 너무 많이 발생한다. 어찌 하랴. 조사는 이미 시작되었고. 게다가 이런 어려움으로 인하여 조사원들도 조사에 불성실하게 되고. 참으로 어렵다. 그러기에 이런 문제를 최소화 할 수 있도록 조사 설계를 치밀하게 수립해야 한다. 환경적 요인의 영향을 최소화 하기 위해서는 가능하면 조사를 하루에 끝내는 것이 좋다. 이를 이상 조사하면 요일 별 영향, 기후의 변화로 인하여 동일한 의견이나 성향을 가진 응답자라도 응답이 달라질 수 있기 때문이다. 여기서 언급되지 않은 문제들 포함 조사에 영향을 미치는 사건들이 실제 설문 조사 과정에서 발생하지만 그 문제점들은 조사자가 알고 물어 버리거나 부지간에 묻혀버린다. 이미 해 버린 조사인데 어찌랴. 그러나 설문 조사 과정에서 발생한 사건들은(환경적, 기후적 변화) 조사 결과를 해석하는데 중요한 역할을 하므로 조사원들에게 조사 과정에서 발생한 사건들을 반드시 적게 한다. 예를 들어 보자. A 지역에서 설문 조사하는데 모 회사에서 홍보 차원의 선물을 나누어 주고 있었다. 이로 인하여 사람들은 기분이 좋아 설문 응답에 호의적이었고 만족도 문항들에 대해 실제보다 높게 표시하였고, 이 지역에 응답자 중 여자 비율이 다른 지역에 비해 높았다. 그리고 설문 분석 결과 남녀별 OO 만족도에 차이가 있었다면 우리는 남녀 만족도 차이가 있다고 결론 지을 수 있는가? 이에 대한 올바른 해석을 하려면 조사 결과에 영향을 미치는 사건들에 대한 기록이 있어야 한다.

설문 조사 시 기념품은 응답률을 높일 수 있지만 응답 정확성을 높이지는 않는다는 것을 주의하기 바란다.

### 1.2.8. 수집 설문 자료 코딩

설문 조사가 완료되면 불성실하게 응답된 설문지(조사원이 체크)나 응답하지 않는 문항이 많은 설문지는 제외하고(응답 문항 비율이 75% 미만) 각 설문지에 일련 번호를 매긴 후 자료를 코딩 하게 된다. 행은 응답자, 열은 문항 응답으로 하여 사용하게 될 통계 소프트웨어에 자료 입력(코딩)을 하면 된다. 대부분의 통계 소프트웨어들이 엑셀 자료나 텍스트 자료

(메모장에서 입력)을 불러오는 기능이 있으므로 분석자가 편한 방법을 선택하면 된다. 설문 조사 자료는 메모장에서 입력하되 열 사이 공백 없이 입력하는 것이 코딩 오류가 적고 빨리 입력할 수 있다. 이 부분에 대해서는 4장에서 다루기로 한다.

### 1.2.9. 설문 분석 및 결과 보고서 작성

코딩 자료를 이용하여 연구 목적에 적합한 설문 분석 방법을 적용하고 분석 결과를 정리, 요약하여 보고서나 연구 결과를 발표한다. 이를 위해 필요한 도구를 정리하면 다음과 같다..

(1)통계 분석: 통계 소프트웨어

(SAS, SPSS, Minitab) 등, MS-Excel (기본적인 분석)

(2)결과 요약을 위한 그래프 그리기

통계 소프트웨어, MS-Excel

(3)결과 보고서 작성:

MS-Word, 아래 한글

## 1.3. 시청률 조사

91년 MSK(Media Service Korea)에 의해 시청률 조사는 처음 시작되었고 99년도에는 TNS Media Korea(미국 조사 기관)가 가세하였다. MSK는 99년 9월에 영국 AC Nielsen에 의해 인수되었다. 시청률은 광고비 책정에 핵심 자료가 되므로 정확한 조사가 생명이다. 조사 방법에는 Passive Household Meter 방법(시청 여부와 프로그램만 조사)과 Set-top Box(People meter: 시청자 인적 사항까지 조사 가능)이 있다.

### 1.3.1. TNS Media Korea

98년 11월에 출범하여 99년 6월말에 처음으로 서울지역 300가구에 대한 시청률 조사 자료를 제공하기 시작한 후 인천 경기, 부산, 대구, 광주, 대전 등의 지역까지 확대하였다. 조사 가구수는 총 1,000가구로 표본 크기는 서울과 수도권을 중심으로 309가구, 인천·경기 권에 191가구, 부산 200가구, 대구 100가구, 광주 100가구, 대전 100가구 등으로 구성되어 있다.

패널 가구 선정을 살펴보면 우선 전국 24,000가구를 대상으로 무작위 전화조사를 실시해서 8,000가구에 해당하는 인구통계학적 자료를 추출한다. 이를 기초로 4,000가구의 sub-

표본을 작성하게 되는데 시청시간과 TV 수상기 수, 세대 구성원 수, 유선방송 가입 여부, 응답자 연령, 가구 수입 등의 변수들이 고려된다. 이를 통해서 1,000 가구의 패널을 추출하고 직접 방문 후 면접 조사를 통하여 최종 확정하였다.

패널 면접 조사에서는 지역, 성별, 연령, 학력, 직업, 수입 형태, 혼인 여부, 가구주, 주 소득자, 가사 담당자, 가사 담당자의 연령과 직업, 세대원 수, 자녀 수, 자녀 연령, 개인 소득, 가구 소득, TV 수상기 수, 케이블 TV 가입 여부, TV 시청시간, 거주 형태, 주택 보유, 승용차 보유 여부 등의 항목을 조사하였다.

패널 선택에 신중을 기했음에도 불구하고 특정 패널이 집중되게 나타난다면 이것은 전체 시청률의 대표성과 신뢰성을 해치게 되기 때문에 TNS 는 시청률이 편중되는 것을 방지하기 위해서 Rim Weight 를 통해서 패널을 관리하고 있다. 이는 실제 TV 를 시청한 개인의 시청 상황과 모집단 분포와의 차이를 최소화하고 각각의 범주 별 실제 표본 분포를 매일 확인하여 정확한 표본 관리와 대체가 가능한 시스템이다.

### 1.3.2. AC Nielsen

99 년 기존의 MSK 를 인수한 AC Nielsen 은 매년 11 월에 TV 환경조사를 실시하고 있다. 표본 추출 방법으로는 시, 구, 동, 가구 단위의 multi-stage random sampling 방법을 적용하였으며 설문지를 이용한 1 대 1 면접조사 후에 100% 검증하여 자료를 수집하고 있다. TV 환경변화를 분석하여 샘플링에 반영하고 대표성 확인과 유지, 설치 예비가구 확보 작업도 병행하고 있다. 또한 TV 환경조사 결과와 통계청의 인구정책 조사 자료를 기초해서 TV 수상기 보유 수와 가족 구성원수, 월평균 가구소득 등을 주 통제 변수로 사용하고 있다.

일단 패널로 선정된 가구에는 직접 방문해서 피플미터를 설치하고 사용방법에 대한 교육을 실시한 후 PCTIPS 와 기타 패널 관리 프로그램 등으로 패널에 대한 질적 수준 분석과 한국형 패널 관리 프로그램 등을 동원해서 패널 Quality 분석을 하고 있다. 연 평균 패널 교체율은 20% 정도이다.

### 1.3.3. TNS 와 AC Nielsen 차이점

TNS 은 디지털 방식의 Picture-matching people meter, AC Nielsen 은 아날로그 방식의 people meter 를 사용한다. People meter 는 기본적으로 TV 의 켜고 끈 상태와 채널 변경 사항 등을 자동적으로 체크하고 개인 시청 행태는 패널로 구성된 가구 구성원 들이 각자 본인에게 부여된 번호를 handset 을 이용해서 선택하도록 되어 있다. AC Nielsen 의 미터기가

채널의 탐지를 Probe 라는 주파수 탐지기가 하는데 비해 TNS 의 미터기는 TV 화면의 그림 일부를 저장해서 채널을 검사하는 Picture-matching 방법을 적용하고 있다는 점이다. Picture-matching 방식의 장점은 위성방송과 공중파 방송, 케이블 TV 방송 등의 디지털화가 이루어 지더라도 각 채널을 손쉽게 구분해 낼 수 있는 장점을 보유하고 있다.

소프트웨어 측면에서 TNS 는 효과분석 시스템인 Info TV 와 효과예측 시스템인 Plan TV 를 공급하고 있는 반면 AC Nielsen 은 프로그램 위주의 분석 시스템인 Telescope 와 집행 광고 에 대한 효과 분석 시스템인 Ad watch 를 공급 중에 있다. 광고 모니터링 방법상에서 AC Nielsen 이 KADD 의 광고 모니터링 자료를 공급 받아서 내부 시청률 자료와 접목시키는 반면 TNS 는 자체 모니터링 시스템을 구축하여 직접 작업을 수행하는 점에서 서로 차이가 있다.

## 1.4. 여론 조사 (poll)

### 1.4.1. 개요

사회적 관심의 대상이 되는 사항에 대한 사람들의 의견을 조사하는 것을 여론 조사(poll)라 하며 선거에 앞서 유권자들의 투표율, 지지율 조사가 대표적인 예이다. 특히 주민에게 불편을 주는 시설(예: 원전 핵 폐기물 시설), 국민이 참여하게 대립하는 법안 (예: 호주제), 정부 예산을 많이 투자하는 시설(예: 대전 지하철)에 대한 사회 여론 조사의 경우 국민(주민)의 의견이 온전히 수렴될 수 있도록 충분한 시간과 비용을 갖고 여론 조사를 할 필요가 있다. 후보자들은 이를 근거로 선거 전략을 수립하게 되고 유권자들은 자신이 지지하는 후보의 지지 정도를 미리 알 수 있다. 그리고 지지 후보를 정하지 못한 부동표들의 행방을 결정하는 주요 요인이 되기도 한다. 후보자의 예상 득표율에 대한 예측과는 달리 social desirability 인하여 투표율에 대한 예측은 어렵다(미국 민주당 지지자들의 성향). 패널(Panel)을 구성하여 유권자들의 지지 성향의 변화 추이(trend)를 살펴 볼 수 있으나 패널을 구성하고 유지하는데 비용이 많이 들고 패널 구성원의 이탈 현상이 많아 거의 사용하지 않는다.

미국에는 8 대 조사 기관이 있다. CBS (CBS, New York Time), Chilton (ABC and Washington Post), Gallup, Louis Harrison and Associate, Media General, ICR (AP) Survey Research, Roper Starch Worldwide, Yankelovich Partners 이 있으며, 국내 여론 조사기관으로는 한국 갤럽, 리서치&리서치, 코리아리서치, 한국리서치 등이 있다.

### 1.4.2. 여론 조사 방법

여론 조사 초기에는 조사원이 직접 방문하여 조사하였으나 응답자들에 대한 인적 사항에 대한 사전 정보가 필요하며 시간과 비용이 많이 든다는 단점이 있었다. (표본 추출 방법은 **Multi-stage Stratified Sampling**) 그러다가 전화 보급률이 높아지면서 요즈음은 전화 여론 조사를 실시한다. (표본 추출 방법: **Multi-stage Stratified and Systematic Sampling**)

### 1.4.3. 전화 여론 조사 3대 구성 요소

#### (1) 전화 번호 선택

표본 프레임으로 전화 번호부를 이용하면 된다. 그러나 이사, 등재되지 않은 번호, 그리고 사업용 주거용 번호의 혼재(미국의 경우는 주거용, 사업용이 어느 정도 집단화 되어 있다)로 인하여 전화 번호부는 효과적인 표본 프레임이 아니다. 이에 해결책으로 **RDD** 방법을 사용한다.

**RDD(Random Digit Dialing)**는 시스템이 전화를 걸고 연결이 되면 조사자 설문 조사를 실시하게 된다. **RDD** 는 전 지역을 지역별로(우리나라의 경우에는 도별) 층화 한 후 먼저 국번을 임의로 선택한 후 전화 통화가 되면 그 후 한동안은 뒤 4 자리의 전화 번호를 **systematic** 하게 선택하여 전화를 건다. **RDD** 방법으로는 **Mitofsky-Waksberg** 방법, **Donnelly Quality Index** 방법 등이 있다.

**Mitofsky-Waksberg** 방법은 마지막 4 자리 숫자를 임의로 다이얼링 한다. 그 중 어떤 번호가 거주용 번호로 (예: **7622**) 밝혀지면 일단 **x-7000~x-7999** 일차 추출 단위로 (**sampling unit**) 간주하고 거기서 임의로 전화 번호를 선택. 일반적으로 미국의 경우와 같이 거주용 번호는 몰려 있는 지역에서 주로 사용한다. **Donnelly Quality Index** 는 거주 가구용 전화 번호 데이터 베이스 사용하여 시스템이 임의 추출하여 전화를 거는 방식이다.

통화 중이거나 부재 중인 경우에는 재시도 하게 된다. **CBS**(10 분 이내 재 통화, 4 일간 6-7 회) **Chilton** 과 **ICR**(3 회까지 재통화), **Gallup**(최소 2 회 이상), **Harrison**(3 일간 서로 다른 시간 대에 3 회 이상 재통화, 통화 중인 경우 15 분 이내 재 시도), **Media General** (5-6 회 재통화 시도, 통화 중인 경우 15 분 이내 재시도), **Yankelovich**(3 회 이상 재통화) 등 조사 기관마다 자체 재시도 규정을 가지고 있다.

조사자가 거절한 경우: CBS, Harris, Media General 은 1 번 더 시도, Gallup 은 2 번 더 시도하는 등 이에 대해서도 조사 기관마다 차이가 있다.

### (2) 응답자 선정

일반적으로 전화를 받는 사람은 주로 청소년이나 주부이므로 유권자 인적 사항(성별, 나이, 직업 등)을 맞추기 위하여 조사 기관은 자체 순서(예를 들어 최근 생일이 지난 사람, 조사자가 인구학적 비율을 맞추어 감)를 정해 놓고 있다. 이렇게 하더라도 모집단 전체 유권자의 인구학적 비율과 맞추기는 불가능하므로 다음과 같이 가중치를 고려하여 조사된 비율을 조정한다.

### (3) 가중치

가중치를 어떻게 고려하는지 예를 들어 설명해 보자. 유권자 모집단 성비가 55(남자):45(여자)이다. 전화 여론 조사 결과 남자 60 명, 여자 60 명이 조사되었고 ○○ 후보에 대해 남자는 30 명, 여자는 20 명이 지지한다고 조사되었다.

① 가중치를 사용하지 않는 경우 →  $50/120=0.417$

② 가중치 사용하는 경우 →  $(30 \times 0.55/0.5 + 20 \times 0.45/0.5)/120=0.425$

○○ 후보의 지지율 경우 여자에 비해 남자 지지율이 더 높게 나타났다. 그러나 표본 추출 결과 남자 비율이 모집단 남자 비율에 비해 적게 조사되었으므로 가중치 사용하면 지지율이 상승한다. 이와 같은 방법으로 모든 인구학적 비율을 동시에 고려하여 재계산한다.

미국 Gallup 경우 성, 인종(백인, 흑인, 기타), 지역(동부, 서부, 중서부, 남부), 나이(18-24, 25-34, 35-44, 45-54, 55+), 교육(고졸 미만, 고졸, 대학 미만, 대졸) 분포를 맞추기 위해 인구학적 분포에 의한 가중치를 이용하여 1 차 보정한다. 2 차 보정은 투표 의향과 관련된 문항의 응답 결과를 토대로 투표 의향 점수를 (7 점-0 점) 계산하여 가중치로 사용한다. 이와 같은 방법으로 지지율을 보정한 결과 1988 년 대통령 선거에서 Bush 예측 지지율은 조사된 지지율보다 보정한 경우가 8% 높게 나왔었는데 실제 조사 기관은 보정된 지지율을 예측 지지율로 발표하게 된다. sampling error 가 2.5%인 것을 감안하면 8% 가중치의 보정 효과는 매우 큼을 알 수 있다.

#### 1.4.4. 전화 여론 조사 실시

전화 여론 조사의 경우 설문 문항은 응답자의 인구학적 항목(성별, 직업, 나이, 소득 정도, 현주소 거주 기간, 교육)들과 조사의 주 문항(지지 후보는? 정책 찬성 여부?)을 묻는 항목으로 구성한다. 전화 조사의 경우 다음 사항에 유의해야 한다.

- (1)응답자들에게 응답을 유도하는 **leading statement**(유도 지문) 영향이 매우 크므로 유의해야 한다. “농촌 경제가 살아나고 있습니다. 당신은 현 정부의 경제 정책에 대해 얼마나 만족합니까?” 이런 유도 지문 문항은 실제 국민들이 느끼는 만족 정도보다 높게 조사되게 하여 조사 결과를 왜곡한다.
- (2)문항이나 보기가 길거나 많으면 응답자가 기억할 수 없으므로 가능하면 간단하게 구성해야 한다.
- (3)전화 여론 조사는 **Tele-marketing** 이 아니므로 대답을 강요하거나 불필요한 친절은 조사 피해야 한다.
- (4)응답자들이 자신의 신분이 드러났다는 생각으로 인하여, 혹은 조사 기관이나 조사원에 대한 불신으로 인하여 무응답 혹은 자신의 의견과 다른 응답을 하는 경우가 많으므로 응답자에게 신뢰를 줄 수 있어야 한다
- (5)무응답을 어떻게 처리할 것인가? 인구학적 특성에 의한 **Discriminant Analysis**(판별 분석)에 의해 무 응답자들의 성향을 분석한다.

전화 조사의 경우 몇 명을 조사할 것인가? 국민 전체가 대상일 경우 모집단의 크기가 너무 크므로 **SRS** 표본 오차 공식( $z_{\alpha/2} \sqrt{\frac{0.5^2}{n}}$ )을 이용하면 된다. 표본 오차가 2.5%이고 신뢰 수준이 95%인 경우( $z_{\alpha/2} = 1.96$ )이므로  $n$ (응답자 수)은 1537 명이다. 표본 오차가 2%(신뢰 수준 95%)인 경우  $n = 2401$  명, 3%인 경우에는  $n = 1067$  명 이다. 일반적으로 여론 조사는 표본 오차 2.5%, 신뢰 수준 95%에서 실시하게 되므로 표본 응답자 수는 1500~1600 명 이다. 표본 오차를 2.5%에서 2%로 0.5% 줄이기 위해 850 명 더 추출해야 하므로 비용 증가에 비해 오차 줄어드는 비율이 상대적으로 적다.

## 1.5. Comment

### 1.5.1. Census vs. sample survey

census 는 조사 대상인 사람들(모집단) 모두를 조사하는 것이고 sample survey 는 전체 대상자 중 일부만 조사하는 방법이다.

### 1.5.2. 표본 오차(Sampling error)

조사 대상자 전체를 조사하지 못하여 발생하는 오차를 표본 오차(sampling error)라 한다. 표본 추출 방법에 따라 표본 오차 계산 식이 존재하지만 계산이 매우 복잡해 SRS(임의 추출 방법) 방법을 사용할 경우 발생하는 표본 오차를 여론 조사 표본 오차로 사용한다. 조사 발표 시에는 표본 오차를 오차 한계라고도 한다. 다음은 SRS 방법 사용 시 표본 오차 계산 식이다.

$$z_{\alpha/2} \sqrt{\frac{0.5^2}{n}}$$

95% 신뢰수준:  $z_{\alpha/2} = 1.96$ , 응답자 수  $n=1600$  ▶ 표본 오차=0.0245

여론 조사 시 국민 전체 의견이나 생각을 1600 명 정도를 정확하게 예측할 수 있을까? 라고 의심하고 믿지 않으려는 경향이 있다. 이로 인해 여론 조사 발표 시 응답자 수를 밝히기 보다는 신뢰 수준 95%, 표본 오차(오차의 한계) 2.5%라고 발표한다. 이는 응답자 수가 1600 명이라는 것과 동일하다.

### 1.5.3. 출구 조사(exit poll)

Bishop, G. and Fisher, B.(1995), "Secrete ballots and self-reports in an exit-poll experiments", Public Opinion Quarterly, Vol. 59, 568-588 출구 조사를 면접 조사와 투표 조사 두 방법을 실시하였으며, 대통령 후보, 상원 의원 후보, 그리고 노인 복지를 위한 세금 신설에 대해 물었다. 면접 조사의 경우 무응답/거절이 9.9%였고 투표 조사는 1.2%였다. 두 조사 모두 실제 선거 결과와 유사하였다. 그러나 노인 복지를 위한 세금 신설에 대해서는 투표 조사 결과가 면접 조사에 비해 반대가 많았다. 이는 Social desirability(사회적 규범: 도덕적으로 당연히 해야 할 일이라고 생각하기에 공공 장소에서는 찬성 하지만 비밀이 보장되면 행하지 않거



나 반대하는 사항)의 한 예이다. 투표율 예측의 경우에도 실제 투표율에 비해 예측 투표율이 높는데 이는 사람들이 투표는 국민의 의무라고 생각하기 때문에 사전 조사 시 투표할 것이라고 답하기 때문이다.

1996년 우리나라 4·11 총선거에서는 출구 조사가 실패하였다. 그 이유는 방송사들의 과열 보도 경쟁을 막기 위하여 유사 출구 조사만 허용하였기 때문이다. 선거법에서 투표소 반경 500m 이내에서는 출구 조사를 못하게 하여 투표를 마친 유권자들을 대상으로 전화 조사를 실시하였다. 자신의 신분이 드러나는 문제로 인하여 무 응답률이 약 20%에 이르고 자신의 의견이 일방적으로 드러난다는 사실로 실제로 선거한 사람과 다른 후보를 말하거나 선거를 하지 않았다고 응답하는 경우가 발생하였다. 출구 조사가 허용된 2002년 대통령 선거에서는 박빙의 승부였음에도 불구하고 당선 예상자와 지지율을 정확하게 맞추었다. 또한 투표가 진행되는 동안 실시된 출구 조사 결과가 집계되어 이것이 투표 전략으로 이용되기도 하였다.

#### 1.5.4. Warner 방법

답변이 곤란하거나 거짓말할 가능성이 많은 사항에 대해 설문 조사할 때 사용하는 방법으로 사회학자 Warner 에 의해 제안되었다. 단점으로는 1)응답자의 구성 비율을 알아야 하고 2)단 하나의 항목만 조사할 수 있다는 것이다.

통계학과 02 학번 40 명을 대상으로 cheating 의 경험이 있는가에 대해 알아보려고 한다. 이를 조사하기 위하여 우선 40 명 중 주민 등록 번호의 홀수, 짝수의 비율을 조사한다. 예를 들어 홀수:짝수인 학생의 비율이 0.4:0.6 이라 하자.

질문 문항을 2 개를 만든다. 1)당신의 주민 번호 뒷자리는 홀수인가? 2)당신은 cheating 경험이 있는가? 그리고 응답자들에게 동전을 가져가 자리로 돌아가 자신 스스로 동전을 던져 앞면이면 1)번 질문에 답을 뒷면이면 2)번 질문에 답을 하게 한다. 이 실험에서 가정은 동전의 앞면 나올 가능성(확률)은 1/2 이다. 40 명 학생 조사 결과 예라고 답한 사람의 비율이 0.3 이었다고 하자. 그러면 cheating 경험이 있는 학생의 비율은 다음과 같이 계산된다.

$$0.4 \times 0.5 + x \times 0.5 = 0.3 \quad \blacktriangleright \quad x = 0.2 : \text{cheating 경험이 있는 사람의 비율은 20\%이다.}$$

## [연습문제]

(1)다음 설문 조사의 경우 응답자를 몇 명이야 조사해야 하는지 계산하시오.

- 표본 오차 2.5%, 신뢰 수준 90%
- 표본 오차 2%, 신뢰 수준 90%

(2)최근 발표된 여론 조사나 설문 조사를 스크랩 하고 문제점이 있으면 적고 그 기사로부터 얻은 정보를 1 페이지 분량으로 요약하시오.

(3)NIMBY 현상 관련 여론 조사나 설문 조사를 스크랩하고 자신의 의견을 1 페이지 분량으로 적으시오.

(4)Warner 방법을 이용하여 설문 조사를 실시하고 조사 과정과 결과를 요약하시오.

(5)지난 주에 발표된 시청률 조사를 하나 스크랩하고 문제점이 있으면 기술하시오.