

2-2 사회조사 및 실례



2-2-1. 사회조사

사회통계조사 (social survey)

- 한 지역이나 국가의 경제, 또는 사회적 현상 등을 측정
- 통계청에서 조사해 월별, 분기별, 연도별로 발표하는 각종 경제, 사회지표
- 보건복지부에서 조사해 발표하는 인구 및 후생복지 관련 지표
- 농수산부에서 발표하는 농업, 임업, 어업 관련 통계

여론조사 (poll)

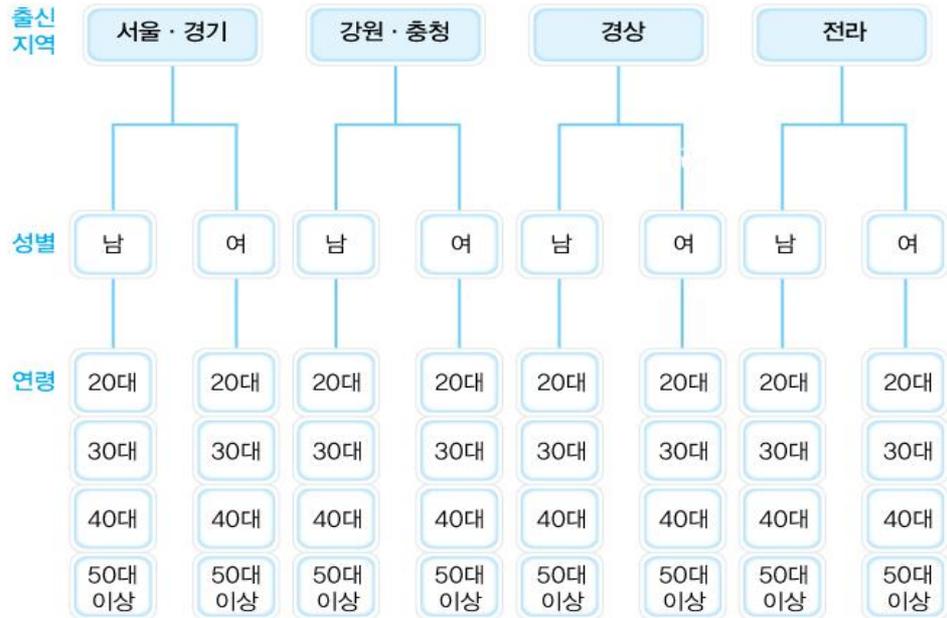
- 어떤 특정한 사회·정치적 현상에 대해 지역주민 또는 국민들의 의견을 조사
- 선거에 앞서서 실시되는 선거 관련 여론 조사



조사목적의 설정	'무엇'에 대한 조사를 실시할 것인가?
조사 대상 집단의 정의	조사목적에 적절한 조사대상집단 설정
설문지의 작성	객관적, 설문지의 내용이 분명하게 정의
표본설계	사용 가능한 조사비용, 대상집단의 목록
표본추출	확률 추출법, 비확률 추출법
조사실시	면접 조사, 전화 조사, 우편 조사
조사결과의 정리 및 발표	정확하게 정리 요약하여 발표(예: 조사기관,시점,방법 등)



그림 3-3 다단계 층화추출법과 할당추출법의 특성조합 예



- 전국 성인 남녀 1000명 조사시 서울은?
 - 전국의 성별/연령별/지역별 구성비 계산
 - 2017년 서울의 인구는 전국 인구의 19.5%, 즉 20%로 보고 200명 조사
 - 전국 남녀 구성비 50%:50%, 서울의 남녀 구성비 49%:51%, 즉 서울 50%:50%로 하여 남녀 각각 100명씩 조사
 - 서울의 연령대별 남녀 비율은 다음과 같음

구분	20대	30대	40대	50대 이상
남	18%	10%	20%	52%
여	18%	19%	19%	44%

- 서울 20대 여성은 18명을 랜덤으로 조사(다단계 층화추출법(확률추출법))
- 서울 20대 여성을 예를 들어 25명으로 지정해 주고 자의적으로 조사(할당추출법 (비확률추출법))

2-2-3. 사회조사결과의 이해

조사결과를 읽을 때 관심을 가져야 할 사항

1. 조사기관과 조사의뢰기관
2. 표본의 구성
3. 조사방법과 조사기간
4. 무응답/응답거부가 어떻게 처리되었는가
5. 숨겨진 자료는 없는가
6. 비교는 정확하게 이루어졌는가
7. 조사항목에 대한 개념이 명확하게 정의되었는가
8. 조사대상집단과 발표대상집단이 동일한가
9. 발표된 통계가 상식적으로 받아들여질 수 있나



2-2-4. 사회조사의 실례

표 3-2 제17대 대선 후보별 예상 득표율과 실제 득표율 (단위 : %)

구분	이명박	정동영	이회창	문국현	권영길
실제 득표율	48.7	26.1	15.1	5.8	3.0
KBS · MBC	50.3	26.0	12.5	6.1	2.9
SBS	51.3	25.0	13.8	5.8	3.0
YTN	49.0	25.3	12.7	8.0	3.6

표 3-3 제17대 대선 예상 투표율과 실제 투표율 (단위 : %)

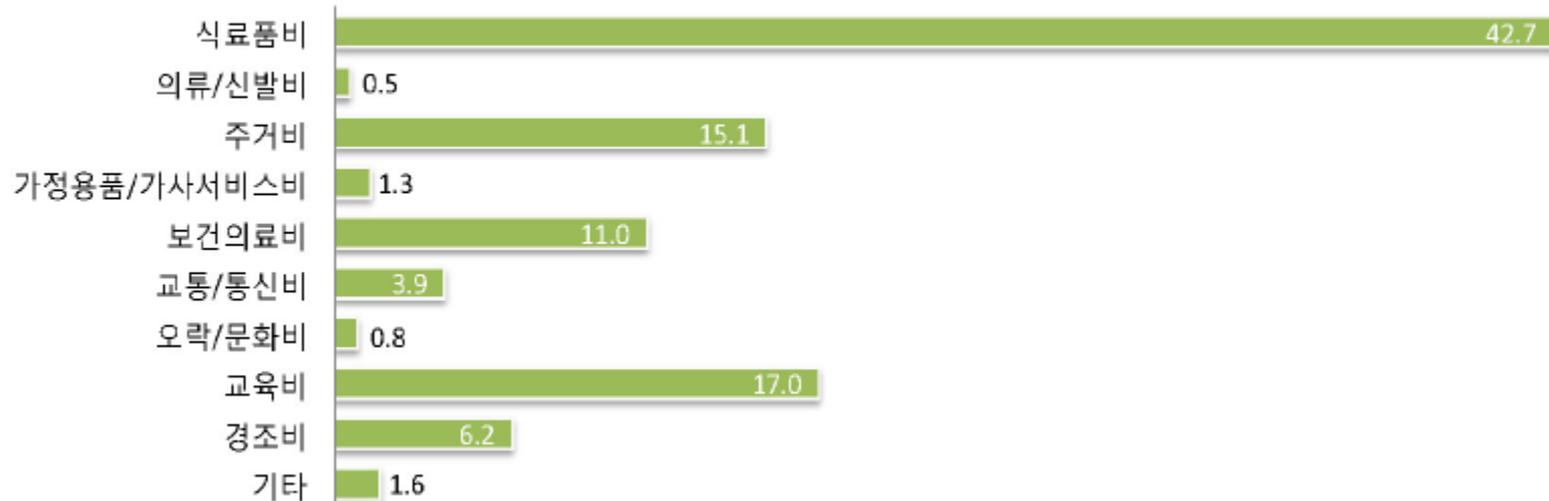
	전국	서울	인천 경기	강원	대전 충청	광주 전라	대구 경북	부산 울산 경남	제주
예상 투표율	76.6	78.3	79.9	87.7	82.0	71.2	75.3	74.2	81.0
실제 투표율	63.0	62.9	61.1	62.6	61.1	65.5	67.7	63.2	60.9

자료 : 갤럽, 조선일보



생활비 지출 (1순위 기준)

(단위 : %)



2-2-5. 비율추정과 최대허용 오차

'95% 신뢰수준 하에서의 최대허용오차'의 의미

(추정비율) \pm (최대허용오차) 사이에 실제 비율이 있을 가능성을 95%정도 신뢰할 수 있다는 주장

예

95% 신뢰수준 하에서 모비율 p 에 대한 추정의 허용오차

$$\pm 1.96 \sqrt{\frac{p(1-p)}{n}} \rightarrow p \text{를 알 수 없음}$$

$\Rightarrow 0 < p < 1$ 에서 $1.96 \sqrt{\frac{p(1-p)}{n}}$ 의 최대값을 구한다.

$$\therefore \pm 1.96 \sqrt{\frac{0.25}{n}} \quad : 95\% \text{ 신뢰수준 하에서 최대허용오차}$$



$$p(1-p)$$

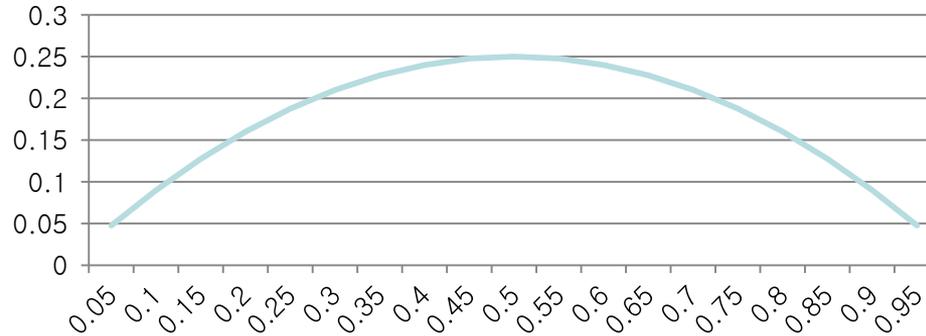


표 3-4

95% 신뢰수준 하에서의 최대허용오차

표본의 크기	최대허용오차(%)
100	9.8
500	4.4
1,000	3.1
1,500	2.5



사례 1

한국갤럽이 8월27일부터 31일까지 전국 1,555명을 대상으로 휴대전화 임의번호 걸기 및 전화조사원 인터뷰 방식으로 18대 대선후보에 대한 여론조사를 실시한 결과 A후보 45%, B후보 40%였다.
표본오차는 95% 신뢰수준에 $\pm 2.5\%$ 포인트이다.

- A후보 : $45\% \pm 2.5\%p \rightarrow 42.5\% \sim 47.5\%$
- B후보 : $40\% \pm 2.5\%p \rightarrow 37.5\% \sim 42.5\%$

▶ 결론적으로 통계적으로 유의한 차이가 없음

사례 2

제주 제2공항건설 찬반 여론조사가 2개 여론조사전문기관(A기관, B기관)에 의해 도민 2000명에게 '제2공항건설'에 대한 찬성-반대 의견을 묻는 방식으로 이뤄졌다. A기관의 조사결과, '반대' 의견이 51.1%로, '찬성'(43.8%)보다 7.3%p 높았고 95% 신뢰수준에서 오차범위는 $\pm 2.19\%$ p였다. B기관의 조사결과도 '반대' 47.0%, '찬성' 44.1%로 반대 의견이 2.9%p 높았고 95% 신뢰수준에서 오차범위는 $\pm 2.2\%$ p였다. "A기관 및 B기관의 조사결과 모두 제2공항건설에 대한 반대 의견이 찬성 의견보다 높게 나타났으므로 제2공항건설을 중단해야 한다"라는 주장의 진위를 판단하라.

•A기관 : (반대) $51.1\% \pm 2.19\%$ p → 48.91%~53.29%

(찬성) $43.8\% \pm 2.19\%$ p → 41.61%~45.99%

•B기관 : (반대) $47.0\% \pm 2.19\%$ p → 44.81%~49.19%

(찬성) $44.1\% \pm 2.19\%$ p → 41.91%~46.29%

→ A기관의 조사결과는 '반대'와 '찬성'이 95% 신뢰수준에서 통계적으로 유의함
B기관의 조사결과는 '반대'와 '찬성'이 95% 신뢰수준에서 통계적으로 유의하지 않음

