

여론조사에서 응답률의 영향

□ 대통령 A에 대한 지지도를 전화로 조사하고자 한다. 다음을 가정한다.

- 조사에 대한 응답은 “지지한다”와 “지지하지 않는다” 둘 중 하나를 선택하는 것이다.
- 총 n 명의 응답을 받고자 한다. 전화조사의 응답률은 α 라고 하자. 단, $0 \leq \alpha \leq 1$ 이다.

– 그러므로 총 n 명의 응답을 받기 위해서는 $\frac{n}{\alpha}$ 명에게 전화를 돌려야 한다.

- “실제로는 알 수 없지만” 모집단(국민 전체)의 A에 대한 지지도는 $0 \leq p \leq 1$ 라고 하자.
- 무응답자 중 A를 지지하는 사람의 비율을 q 라고 하자. 단, $0 \leq q \leq p$ 이다. A를 지지하지 않는 사람은 A를 지지하는 사람보다 조사에 끝까지 응답하지 않을 가능성이 더 높을 수 있기 때문이다.

□ 응답률에 따른 응답자의 A에 대한 지지도는 다음과 같이 계산될 수 있다.

○ 무응답자 수: $\frac{n}{\alpha} - n = \frac{(1-\alpha)n}{\alpha}$

○ 무응답자 중 A를 지지하는 사람 수: $\frac{(1-\alpha)n}{\alpha} \times q$

○ 전화조사 대상자 중 A를 지지하는 사람 수: $\frac{n}{\alpha} \times p$

– 전화조사 대상자 즉 표본틀(sample frame)은 모집단을 잘 대표한다고 가정한다.

○ 응답자 중 A를 지지하는 사람 수: $\frac{n}{\alpha} \times p - \frac{(1-\alpha)n}{\alpha} \times q = \frac{np - (1-\alpha)nq}{\alpha}$

○ 응답자 중 A를 지지하는 비율(A 지지도 조사결과): $R = \frac{np - (1-\alpha)nq}{n} = \frac{p - (1-\alpha)q}{\alpha}$

구분	응답자	무응답자	합계
합계	n	$\frac{(1-\alpha)n}{\alpha}$	$\frac{n}{\alpha}$
A를 지지한다	$\frac{np - (1-\alpha)nq}{\alpha}$	$\frac{(1-\alpha)n}{\alpha} \times q$	$\frac{n}{\alpha} \times p$
A를 지지하지 않는다	$n - \frac{np - (1-\alpha)nq}{\alpha}$	$\frac{(1-\alpha)n(1-q)}{\alpha}$	$\frac{n(1-p)}{\alpha}$

– 만약 $q = p$ 이면 즉 지지도가 응답률에 영향을 미치지 않는다면 $R = p$ 이다. 즉 조사결과는 왜곡되지 않는다.

– 만약 $\alpha = 1$ 이면 즉 응답률이 100%라면 $R = p$ 이다. 즉 조사결과는 왜곡되지 않는다.

○ 만약 $\alpha < 1$ 이고 $q < p$ 이면 즉 응답률이 100%가 아니고 지지도가 응답률에 영향을 미치면 $R > p$

이다[증명: $R - p = \frac{p - (1-\alpha)q}{\alpha} - p = \frac{(1-\alpha)(p-q)}{\alpha} > 0$]. 즉 조사결과는 왜곡된다. 특히 응답률

α 가 낮을수록 (분모는 작아지고 분자는 커지는 것이므로) 왜곡은 더 심해진다.