

## 품질관리표본(Quality Control Sample)의 리인터뷰에 의한 사업체조사의 응답오차 측정

김설희<sup>1)</sup>, 박현영<sup>2)</sup>

### 요 약

최근 경제 사회의 급속한 발전에 따라 개인의 활동분야가 다양해지고 개인비밀보호 인식이 커지면서 응답자들이 통계조사에 잘 협조하지 않는 경우가 늘어나고 있다. 따라서 대부분의 통계를 현장조사에 의존하여 생산하고 있는 통계청에서는 현장조사 결과물의 품질상태에 좀더 관심을 가져야만 하게 되었다. 이러한 현장조사에 대한 품질관리의 일환으로 현재 통계청에서는 통계별 조사대상으로부터 품질관리표본(Quality Control Sample)을 추출하고 이를 대상으로 리인터뷰를 실시함으로써 응답오차측정, 조사현장의 실태파악, 응답자 의견수렴 등에 활용하고 있다. 리인터뷰는 조사직원의 고의적인 자료조작 또는 보충교육 필요성 등 현장조사업무를 평가하거나 응답분산(simple response variation), 응답편의(response bias) 등을 산출하고, 이를 분석하는 모델을 이용하여 응답결과의 신뢰도를 분석하는데 목적을 두고 있다. 본 연구에서는 품질관리표본(QC Sample) 설계 및 추출, 리인터뷰 시나리오개발, CATI(Computer Assisted Telephone Interviewing)를 이용한 리인터뷰 기법 등을 통계청 사업체조사 모니터링 사례를 중심으로 설명하고 조사직원 특성별 응답오차 측정 및 비교, 정확성 항목에 대한 차이분석 등 결과에 대하여 논하고자 한다.

주요용어 : 품질관리표본(Quality Control Sample) 리인터뷰(reinterview) 응답오차(response error)  
CATI(Computer Assisted Telephone Interviewing) 불일치지수(index of inconsistency)

### 1. 개요

주로 면접조사 방법에 의하여 자료를 수집해야하는 통계조사에서는 조사직원의 현장 활동이 조사결과 자료의 품질에 큰 영향을 미칠 수 있다. 특히 최근 경제·사회의 급격한 변화와 함께 개인의 생활이 다양해짐에 따라 응답자부재 현상이 늘어나고 사생활 비밀보호 등을 이유로 응답을 거부하는 등 현장조사의 환경이 열악해지고 있다. 이제는 과거와 같은 조사응답률과 응답자들의 적극적인 협조를 기대하기 어렵게 되어가고 있다. 따라서 조사현장에서의 질 좋은 자료를 수집하기 위해서는 인터뷰기술 개발은 물론 현장조사의 정확도, 조사직원의 숙련도 등에 대한 체계적인 관리 및 조사직원 교육을 강화시킬 필요가 있다.

본 연구에서는 품질관리표본(Quality Control Sample), 리인터뷰 및 응답오차측정 등에 관한 이론 및 사례를 살펴보았다. 아울러 통계청 사업체조사에 대한 현장조사 모니터링 사례를 중심으로 CATI(Computer Assisted Telephone Interviewing)를 이용한 리인터뷰 조사모델을 구성하고, 조사직원 특성별 응답오차 측정 및 비교, 정확성 항목에 대한 차이분석 등을 통한 현장조사의 신뢰성 분석에 대하여 논하고자 한다.

### 2. 조사모델과 리인터뷰

---

1) 이학박사, 통계청 품질관리팀장, 대전 서구 둔산동 920 정부대전청사 3동 1401호

2) 심리학석사, 통계청 품질관리팀 모니터링담당, 대전 서구 둔산동 920 정부대전청사 3동 1401호

품질관리표본의 리인터뷰에 의한 사업체조사의 응답오차 측정

통상 조사모델에서는 결과자료의 품질에 영향을 미치는 오차를 통합하여 평균제곱오차로 정의하고 이를 표본분산, 단순응답편의, 교차응답편의, 편차의 공분산, 제곱편의 등으로 나누어 분석하고 있다. 이 중 응답오차를 측정하기 위해서는 응답분산의 상관성분을 측정하는 상호관입표본(interpenetrated subsamples)방법과 단순 응답분산과 응답편의를 추정하기 위해서 동일대상에 대한 반복측정(Replicated measurement on the same units) 등 주로 두 가지가 활용되고 있는데 후자를 면접조사에서 리인터뷰(reinterview)라고 한다. 리인터뷰는 주로 현장조사 업무를 평가하고 조사모델에서의 오차요소를 추정하는데 목적을 두고 있다.

리인터뷰를 위한 표본은 응답오차표본(Response error sample)과 품질관리표본(QC sample)으로 나누어 검토할 수 있다. 일반적으로 품질관리표본은 응답자, 사업체 또는 집락별로 1단 표본추출을 한다. 그러나 조사원의 현장업무 평가를 위한 리인터뷰인 경우 2단 표본추출 즉 조사원을 1차로, 조사원이 조사하는 응답자들을 2차로 하여 표본을 선정할 수 있다. 리인터뷰의 응답자는 본조사에 응하였던 응답자로 선정하여 가구원 중에서 선정하는 경우 본조사와 같은 방법을 적용하되 가구 내에서 가장 식견이 있거나 응답가능한 사람이어야 한다.

조사업무를 평가하려는 목적으로 실시할 경우에는 리인터뷰는 본조사 인터뷰로부터 가능한 빠른 시간 내에 시행되어야 한다. 시간이 너무 많이 경과되면 응답자가 실제 인터뷰가 일어난 것을 잊게 될 수가 있으며, 조사원들에게 업무수행결과를 피드백하기도 어렵게 된다. 또한 리인터뷰에서의 인터뷰를 전화조사로 할 것인지 면접조사로 할 것인지가 하나의 이슈가 될 수 있는데 미국 센서스국에서는 비용절감을 이유로 전화조사를 주로 실시하고 있다(Forsman 1999).

통계청에서는 통계조사에 대한 모니터링을 위하여 엑셀의 비주얼베이식을 사용한 대화식 조사표를 개발하여 리인터뷰를 실시하던 중, 리인터뷰 응답오차를 최대한 줄이고 효율적인 시스템 운영을 위하여 금년부터 CATI를 도입하여 활용하고 있다. CATI방식은 일반 전화조사 보다 몇 가지 장점을 가지고 있는데, 랜덤하게 표본을 추출하고, 전화통화에 실패한 경우 원인분석 및 다음 대상처로 전화걸기 알람기능과 통화예약처리 등을 자동화함으로써 전화성공률을 높이고 리인터뷰의 결과를 신속하게 처리할 수 있다. 또한 통화된 내용이 녹음되어 사후 검증하는데 활용할 수도 있으며, 슈퍼바이저가 통화내용을 실시간 또는 VOD(Video On Demand)형태로 청취하여 리인터뷰의 응답오차 점검과 미숙한 모니터요원의 재교육 등 리인터뷰과정 전반을 관리할 수가 있다.

미국 센서스국의 경우 CPS조사의 노동력자료를 위한 단순응답분산을 측정하기 위하여 정기적인 리인터뷰를 시행하고 있으며 동시에 CATI에 의한 리인터뷰 연구가 수행되었다(Schreiner 1989). 리인터뷰 표본은 원래 조사대상 중 집락을 구성하지 않고 추출하였으며 본조사의 정규조사원이 현장조사 직후 2-3일후에 본조사의 조사지침대로 응답자와 통화하였다. 표1은 CPS 정규리인터뷰(n=1,454)와 CATI 리인터뷰(n=8,217) 결과 매치된 응답자에 대한 결과의 불일치지수를 비교하고 있다.

표 1 미국 CPS조사의 CATI 및 정규 리인터뷰의 불일치지수 비교

구분	리인터뷰 방법	불일치 지수	90%신뢰구간
취업	CATI	7	5 - 9
	정규	8	7 - 9
실업	CATI	38	29 - 50
	정규	38	34 - 43
비경제활동인구	CATI	9	7 - 11
	정규	10	9 - 11

동일한 대상에 대한 조사를 반복하는 리인터뷰는 응답태도의 변이 즉 측정분산(measurement variance), 응답분산(response variance) 또는 신뢰도(reliability)를 파악하는데 활용된다. 매번 측정시마다 참값(true value)이 변하지 않는다고 가정할 때 각각 응답결과는 참값+오차로 표현 가능하다. 즉 본조사 응답 = 참값 + 본조사오차이며 리인터뷰 응답 = 참값 + 리인터뷰오차 즉,  $y_{1i} = X_i + \epsilon_{1i}$ ,  $y_{2i} = X_i + \epsilon_{2i}$ 로 표현된다. 여기서  $E(\epsilon_{1i}) = E(\epsilon_{2i}) = 0$ 을 가정하기 때문에 참값은 상수로 유지가능하다. 따라서 응답의 신뢰도를 분석하기 위하여 불일치지수(Index of inconsistency)를 이용할 수 있다. 불일치지수는 조사의 전체분산에 대한 응답오차 분산의 비율로서  $I = \frac{E(\epsilon_{1i} - \epsilon_{2i})^2/n}{\sigma_y^2}$ 이며 여기서  $I$ 는 불일치지수,  $n$ 은 리인터뷰 건수이며  $\sigma_y^2$ 는 조사결과  $y$ 에 대한 측정값의 개인별 분산이다. 따라서 불일치지수는 측정 총 분산에 대한 응답오차분산의 비율이라고 할 수 있으며, 보통 조사모델에서 0.0과 +1.0사이의 값을 가지게 된다. 여기서 Groves(1989)가 제시한 하나의 가정 즉 반복된 조사에 걸친 오차간의 공분산  $Cov(\epsilon_{1i}, \epsilon_{2i}) = 0$ 이다. 따라서 분모의  $E(y_{1i} - y_{2i})^2 = E(\epsilon_{1i} - \epsilon_{2i})^2$ 로 나타나게 되며 측정치 차이의 제곱은 오차분산의 추정값이 된다.

### 3. 통계청의 통계조사 리인터뷰 사례연구

본 연구에서는 두 가지의 사업체통계조사를 대상으로 선정하였는데, A조사와 B조사 각각 표본의 수가 18000개, 담당조사직원어 1200여명 정도 되는 조사이다. A, B 두 조사는 서로 다른 산업분야에 대한 조사이지만 조사모델이 유사하고 거의 동일한 기간동안 이루어지는 조사이기 때문에 사례분석 결과에 대한 상호 검증도 가능하여 연구대상으로 삼았다.

리인터뷰 프로그램은 여러 과정을 거쳐 설계되는데 우선 대상통계가 선정되면 모니터요원들과 수퍼바이저는 담당과로부터 조사지침 교육을 받고 조사현장에 조사원과 동행하여 이론과 실습을 겸한 기본지식을 습득한다. 이때 수퍼바이저는 리인터뷰를 위한 조사표와 시나리오를 설계하는데, 시나리오는 모니터요원들이 응답자에게 조사협조를 구하는 도입단계부터 조사마무리까지 응답자에게 부담을 주지 않는 정교화된 2-3분 내외의 과정별 구체적인 멘트, 질문순서, 응답자의 FAQ 등을 설계한 것이다. 시나리오 초안이 완성되면 시나리오 분석과 역할연기연습(role-playing)을 통한 반복적인 가상통화를 시행하고 요원들 각각의 전화응대에 대한 태도와 장단점을 비교하여 발생가능한 조사원오차를 줄이고 실제상황에 적절한 대응방법 강구 등 준비 과정을 갖는다. 리인터뷰 조사표는 조사실태를 파악하는 부분과 정확성을 점검하는 부분으로 구분하고, 정확성 점검항목은 본조사 항목 중에서 조사직원이 실제조사 없이 추측하여 작성하기 어려운 항목을 중심으로 선정한 것이다.

#### 3.1 품질관리표본 추출 및 리인터뷰 대상응답자선정

A, B 두 조사의 품질관리 표본은 조사 완료된 대상처 목록을 표본프레임으로 하여 2단계 표본추출하며 사무소별 조사원을 1단계 표본추출단위(psu, primary sampling unit)로 하고, psu안에서 응답자를 선정하였다. 즉, 1단계에서는 각 조사별, 사무소별로 조사를 담당하는 조사직원을 추출하고 2단계에서는 조사직원이 담당하는 대상처 중에서 무작위로 추출하였다. 사무소별로 대상처규모에 따라 리인터뷰의 표본비율이 정해지면 담당조사직원별로 리인터뷰의 대상처 수를 랜덤하게 추출하며, 일부 조사직원에게만 치우치지 않도록 고르게 표본추출 하였다. 2단계에서는 조사직원이 담당하는 대상처 중에서 일정한 비율을 랜덤하게 추출하였다. A조사와 B조사의 리인터뷰를 위한 품질관리표본은 A조사의 전체 대상처 약 14%인 2,589개와 B조사의 13%인

품질관리표본의 리인터뷰에 의한 사업체조사의 응답오차 측정

2,276개를 추출하였다. 이 중 A조사의 경우 리인터뷰에 성공한 대상처는 모두 750개로 응답성공률이 29%에 해당하며 표집오차는 신뢰수준 95%에서  $\pm 1.75\%$ 가 되고, B조사의 리인터뷰 성공 대상처는 모두 748개로 응답성공률이 33%이며 표집오차는 신뢰수준 95%에서  $\pm 1.81\%$ 가 되었다.

리인터뷰 대상 응답자는 본조사의 응답자와 동일하게 선정하였으며, 따라서 리인터뷰에서는 본조사에 응답해 준 응답자와 통화를 하기 위해 실제로 여러 번의 통화를 시도하고 있다. 또한 대상처에서 본조사에 응답해 준 응답자를 찾을 수 없을 경우에는 전화실패 사례로 처리하여 분석에서는 제외하였다.

3.2 CATI에 의한 리인터뷰 통화결과 분석

추출된 표본을 대상으로 본조사의 응답자와 리인터뷰하기 위하여 동일한 표본대상처에 통화를 시도한 횟수를 살펴보면 A조사의 경우 리인터뷰를 위한 4292통화시도 중 전체의 18.1%인 777통화가, B조사의 경우는 3121통화 시도 중 23.6%인 738통화가 리인터뷰에 성공하였다. 1회 시도 시 성공한 경우는 A, B 각각 전체의 50.5%와 55.3%로 가장 높은 비율을 보였고 3회까지의 성공 횟수는 전체의 87.8%, 91.6%를 보였다. A조사와 B조사 모두 7회까지 전화통화를 시도했을 때 98.1%, 99.2%의 성공률을 보인 반면 8회부터 15회의 전화통화 시도는 시도한 횟수에 비해 성공률이 낮아 사실상 거의 시도할 필요가 없는 것으로 판단된다(표 2).

표 2 리인터뷰를 위한 통화시도 횟수

구분 횟수	A사업체조사			B사업체조사		
	빈도	퍼센트	누적퍼센트	빈도	퍼센트	누적퍼센트
1	392	50.5	50.5	408	55.3	55.3
2	202	26	76.4	186	25.2	80.5
3	88	11.3	87.8	82	11.1	91.6
4	34	4.4	92.1	31	4.2	95.8
5	23	3	95.1	14	1.9	97.7
6	13	1.7	96.8	4	0.5	98.2
7	10	1.3	98.1	7	0.9	99.2
8	5	0.6	98.7	3	0.4	99.6
9	5	0.6	99.4	2	0.3	99.9
10	3	0.4	99.7	1	0.1	100.0
12	1	0.1	99.9	-	-	-
15	1	0.1	100.0	-	-	-
합계	777	100.0		738	100.0	

요일별 전체 통화 중 리인터뷰 성공의 빈도가 높은 요일은 목요일로 A, B 각각 26.6%와 23.3%를 보였고, 이외의 다른 요일은 A, B조사 모두에게서 유사한 결과를 보이지 않으므로 리인터뷰를 위한 통화성공은 요일에 따라서는 긍정적인 효과를 보이지 않는 것으로 판단되었다(표 3).

표 3 요일별 리인터뷰 통화성공빈도 (%)

구분 요일별	A사업체조사			B사업체조사		
	전체통화수	통화성공빈도	통화실패빈도	전체통화수	통화성공빈도	통화실패빈도
월	966	211(21.8)	755(78.2)	605	147(24.3)	458(75.7)
화	944	133(14.1)	811(85.9)	336	79(23.5)	257(76.5)
수	703	73(10.4)	630(89.6)	564	147(26.1)	417(73.9)
목	414	110(26.6)	304(73.4)	722	168(23.3)	554(76.7)
금	876	190(21.7)	686(78.3)	579	118(20.4)	461(79.6)
토	389	60(15.4)	329(84.6)	316	79(25.0)	237(75.0)
전체	4292	777(18.1)	3515(81.9)	3122	738(23.6)	2384(76.4)

시간대별 전체 통화 중 리인터뷰성공의 빈도가 높은 시간은 오후 1시에서 2시 사이로 A, B 각각 30.7%와 27.3%의 결과를 보였다. 비교적 낮은 성공률을 보인 시간대는 오후 5시에서 6시 사이로(16.6%, 17.0%), 그 외의 다른 시간대에는 통화의 차이를 보이지 않고 있다(표 4).

표 4 시간대별 리인터뷰 통화성공빈도 (%)

구분 시간대별	A사업체조사			B사업체조사		
	전체통화수	통화성공빈도	통화실패빈도	전체통화수	통화성공빈도	통화실패빈도
9-10	209	56(26.8)	153(73.2)	219	53(24.2)	166(75.8)
10-11	964	146(15.1)	818(84.9)	518	127(24.5)	391(75.5)
11-12	631	144(22.8)	487(77.2)	456	106(23.2)	350(76.8)
12-1	23	2(8.7)	21(91.3)	43	9(20.9)	34(79.1)
1-2	261	80(30.7)	181(69.3)	366	100(27.3)	266(72.7)
2-3	428	82(19.2)	346(80.8)	459	111(24.2)	348(75.8)
3-4	694	87(12.5)	607(87.5)	360	88(24.4)	272(75.6)
4-5	583	97(16.6)	486(83.4)	289	70(24.2)	219(75.8)
5-6	494	82(16.6)	412(83.4)	388	66(17.0)	322(83.0)
6-7	5	1(20.0)	4(80.0)	24	8(33.3)	16(66.7)
합계	4292	777(18.1)	3515(81.9)	3122	738(23.6)	2384(76.4)

### 3.3 응답오차 분석

A, B 두 통계조사의 항목별 응답신뢰도 분석을 위하여 각 한개 항목씩 리인터뷰에서 재조사하고 응답결과의 오차를 분석하였다. 조사직원의 성별에 따른 차이 발생을 살펴본 결과 남자조사직원의 경우 A, B 각각 21.4%, 21.9%의 불일치율, 여자 조사직원의 경우는 25.0%, 18.9%를 보여 불일치비율은 조사직원의 성별에 영향을 받지 않는 것으로 나타났다(표 5).

표 5 성별에 따른 응답오차 분석

구분 성별	A 사업체조사			B사업체 조사		
	전체	일치(%)	불일치(%)	전체	일치(%)	불일치(%)
남	430	338(78.6)	92(21.4)	351	274(78.1)	77(21.9)
여	320	240(75.0)	80(25.0)	397	322(81.1)	75(18.9)
전체	750	578(77.1)	172(22.9)	748	596(79.7)	152(20.3)

사업체조사에 투입되는 조사직원은 직렬에 따라 일반직, 별정직, 계약직, 임시직으로 구분된다. 직렬별 차이비율을 비교한 결과 두 조사에서 임시직의 불일치비율이 A조사의 경우 29.8%, B조사의 경우 23.8%로 다른 직렬에 비하여 높게 나타났으며, A조사는 계약직의 불일치율이 가장 낮고(18.0%), B조사는 별정직이 가장 낮다(17.8%). 이중 정기적인 업무성과 심사 후 재임용 조건으로 근무하는 계약직이 현장조사에 대한 성실도가 상대적으로 높은 것으로 나타났다(표 6).

표 6 직렬에 따른 응답오차 분석

구분 직렬	A 사업체조사			B사업체 조사		
	전체	일치(%)	불일치(%)	전체	일치(%)	불일치(%)
일반직	75	59(78.7)	16(21.3)	75	59(78.7)	16(21.3)
별정직	428	332(77.6)	96(22.4)	320	263(82.2)	57(17.8)
계약직	116	95(81.9)	21(18.1)	151	120(79.5)	31(20.5)
임시직	131	92(70.2)	39(29.8)	202	154(76.2)	48(23.8)
전체	750	578(77.1)	172(22.9)	748	596(79.7)	152(20.3)

또한 조사직원의 근무년수별 차이발생 비율을 분석한 결과, 근무년수 1년 미만인 경우가 각각 26.2%, 24.1%로서 가장 높은 불일치를 보이고 있다.

표 7 근무년수에 따른 응답오차 분석

구분	A 사업체조사			B사업체 조사		
	전체	일치(%)	불일치(%)	전체	일치(%)	불일치(%)
근무년수						
1년 미만	183	135(73.8)	48(26.2)	257	195(75.9)	62(24.1)
10년 미만	129	106(82.2)	23(17.8)	150	126(84.0)	24(16.0)
20년 미만	253	198(78.3)	55(21.7)	220	182(82.7)	38(17.3)
20년 이상	185	139(75.1)	46(24.9)	121	93(76.9)	28(23.1)
전체	750	578(77.1)	172(22.9)	748	596(79.7)	152(20.3)

앞서 설명한 바와 같이 본 연구에서는 CATI에 의한 리인터뷰 결과를 분석하였으며 미국 센서스국 사례에서와 같이 다른 방법에 의한 리인터뷰는 적용하지 않았다. 또한 리인터뷰를 실시한 모니터링원들은 이미 조사된 본조사 결과를 알지 못한 채 독립적인 환경에서 반복 인터뷰를 실시하였고 차이가 발생한 경우에도 자료의 사후조정(reconciliation)은 하지 않았다. 리인터뷰의 대상조사 A, B의 결과로부터 Groves방법에 따라 불일치지수를 산출하여 결과를 비교하였다(표 8).

표 8 본조사와 리인터뷰 결과의 불일치 지수 비교

구분	측정치 차의 제곱	리인터뷰 개수	본조사결과 분산	불일치지수
A조사	914	750	2.164	≈0.563
B조사	984	748	2.082	≈0.632

#### 4. 결론

현장조사의 정확성에 대한 체계적, 과학적 관리는 통계품질의 수준향상 및 유지를 위하여 필수적인 과정으로 인식되고 있다. 본 연구에서는 자료수집 현장에서의 응답오차를 분석하고 결과물의 품질을 평가하기 위한 수단으로 CATI에 의한 리인터뷰를 품질관리표본에 적용하여 분석하였다. 리인터뷰 통화결과분석, 조사원특성별 오차분석 및 본조사와의 불일치지수 산출을 통해 종래의 문서행정 위주의 현장조사관리에 대한 개선안을 도출하게 되었다. 그러나 이 연구에서 리인터뷰 조사모델의 확대 및 리인터뷰 결과의 사후조정문제는 실험환경의 한계로 남게 되어 향후 유사한 연구에서 다루어 줄 것을 기대한다.

#### 참고문헌

- Robert Groves(1989), Survey Errors and Survey Costs, The University of Michigan, John Wiley & Sons Inc.
- Robert Groves(1991), Measurement error across the disciplines, Measurement Errors in Surveys, Eds Biemer, Groves, Lyberg, Mathiowetz, and Sudman, pp1-25, John Wiley & Sons Inc.
- Gosta Forsman & Irwi Schreiner(1991), The Design and Analysis of Reinterview: an Overview, Measurement Errors in Surveys, Eds. Biemer, Groves, Lyberg, Mathiowetz, and Sudman, pp279-301, John Wiley & Sons Inc.
- Paul Biemer & Lynne Stokes(1989), The Optimal Design of Quality Control Samples to Detect Interviewer Cheating, Journal of Official Statistics, Vol. 5, No. 1, pp 23-39
- Johnny Blair & Seymour Sudman(1996), The Survey Reinterview: Respondent Perceptions and Response Strategies, Journal of Official Statistics, Vol. 12, No. 4, pp 421-426
- Gosta Forsman (1989), Early Survey Models and Their Use in Survey Quality Work, Journal of Official Statistics, Vol. 5, No. 1, pp 41-55