서베이통계의 비표본오차 축소방안 연구

통계생산 과정에서 발생하는 오류는 통계의 정확성뿐만 아니라 질적 신뢰성에 중대한 영향을 미치고 나아가서는 이를 근거로 수립한 정책의 효과도 저하시킬 수 있다. 서베이통계는 일부대상만을 조사하여 모집단 전체에 대하여 추론하므로 추정값과 참값간에 다소의 차이가 발생할 수 있으며 이 차이는 표본오차와 비표본오차로 구분할 수 있다. 표본오차는 조사대상 전체의 일부만을 선택함에 따른 확률적 오차로 표본이론을 적용하여 그 규모를 제어할 수 있는 반면 비표본오차는 조사 기획 및 수행의 전 과정에서 발생할 수 있으며 측정 및 제어가 용이하지 않다.

본 연구는 한국은행의 서베이통계와 관련하여 발생할 수 있는 비표본오차의 특성들을 총조사오차의 관점에서 살펴보고자 하였다. 비표본오차와 관련한 기존 연구를 검토하고 한국은행 서베이통계와 관련한 논점에서 살펴보았다. 관심대상에 대한 표본추출틀의 포함오차, 표본개체들이 조사에 참여하지 않음에 따른 무응답오차, 기타 비표본오차들(명세오차, 측정오차, 처리오차)에 대해 각각 살펴보았으며, 비표본오차 축소를 위한 체크리스트 예시와 기관차원의 서베이통계에 대한 지속적 품질향상을 가능하게 하기 위한 고려사항을 논의하였다.

최근 들어, 전 세계적으로 정부나 중앙은행이 편제하는 공식통계의 품질제고가 요구되는 상황에서, 한국은행의 서베이통계에서 발생하는 비표본오차에 대한 원인 및 개선방향을 모색하는 것은 매우 의미 있는 것으로 판단된다.

Ⅰ. 서론
1. 연구배경 및 필요성
2. 총조사절차·총조사오차
3. 내용구성

Ⅱ. 포함오차
1. 개요
2. 기술통계 관련
3. 분석통계 관련

Ⅲ. 무응답오차
1. 개요
2. 무응답
3. 경기심리조사 무응답오차 평가

Ⅳ. 기타 비표본오차
1. 명세오차
2. 측정오차
3. 처리오차

Ⅳ. 추가논의
1. 비표본오차 프로필
2. 조직운영 차원에서의 지속적 품질향상 노력

* 본고는 부경대학교 통계학과 박인호 교수, 한국은행 전경배 부국장(인사경영국), 주성제 과장(통계조사팀)이 작성한 것으로, 본고의 내용은 집필자의 개인의견으로서 한국은행의 공식견해를 나타내는 것은 아님.

Ⅰ. 서 론

1. 연구배경 및 필요성

한국은행은 경기진단과 통화신용정책의 수행에 필요한 객관적 정보의 수집을 위해 소비자동향지수, 기업경기지수, 물가통계 등의 다양한 서베이통계를 작성하고 있다. 특히 한국은행이 작성하는 통계는 정책판단, 경제관련 의사결정 등에 폭넓게 영향력을 행사하므로 국민계정을 포함하여 기업경기조사, 생산자물가지수 등 주요 경제통계는 통계의 작성방법, 공표일정 등에 있어서 객관적인 국제 통계품질 기준을 준수해야 하며 IMF는 그 준수여부를 정기적으로 점검하고 있다. 한국은행의 통계자료는 경제통계시스템(Economic Statistical System, ECOS)을 통해 정책입안자는 물론 일반인들에게도 신속하고 편리하게 제공되고 있다.

통계작성방법의 전문성과 통계품질에 대한 지속적 관리 및 개선 노력은 서베이통계의 정확성은 물론 질적 신뢰성에 중요한 영향을 미친다. 통계생산과정에서 발생하는 오류는 곧바로 정책 수집의 오류로 연결됨으로 통계작성기관은 통계생산과정의 엄밀성에 최선의 노력을 아끼지 않게 된다. 유럽과 북미 국가들은 국제기구들과 함께 서베이통계의 품질관리를 위한 표준적인 관리지표를 지속적으로 개발하고 사용하고 있다. 예로, 유럽통계위원회, 국제연합(United Nations, UN), 유럽통계청(Eurostat) 등은 1982년에 서베이통계 품질 관련한 보고에 포함시켜야 할 최소한의 요건 제정 및 지표에 대한 지침을 <표 1-1>과 같이 마련하였다. 또한 미국 경제분석국(U.S. Bureau of Economic Analysis, BEA)은 기관 홈페이지를 통해 기관의 임무와 비전은 물론 무결성(integrity), 품질(quality), 탁월함(excellence), 혁신(innovation) 등과 같은 기관이 추구하는 핵심가치를 공표하고 있다. 이는 해당기관이 얼마나 생산통계에 대한 정확성과 신뢰성을 중시하며 생산통계의 품질에 대해 열정과 노력를 기하는지의 면모를 잘 보여주고 있다고 할 수 있다.

서베이통계는 일부대상만을 조사하여 모집단 전체에 대하여 추론하므로 추정값(sample estimate)과 참값(true population value)간에 다소의 차이가 발생한다. 이러한 차이는 총조사오차(total survey error, TSE)로 <표 1-1>에 나타난 서베이통계 품질관리를 위한 표준적 관리지표에 명시된 요인들에 의해 발생할 수 있으며, 단순화시켜서 표본오차(sampling error)와 비표본오차(nonsampling error)의 두 가지로 구분할 수 있다.

표본오차는 조사대상 전체의 일부만을 선택함에 따른 확률적 오차(random error)로 표본

이론(예, Cochran, 1977)을 적용하여 그 규모를 제어할 수 있다. 반면 비표본오차는 조사 기획 및 수행의 전 과정에서 발생할 수 있다. 각 과정별로 선택되는 방법과 절차는 비표본오차의 잠재적 발생의 원인이 될 수 있는데, 이는 비확률적이고 비의도적인 오류 혹은 시스템적 과실에 해당하며 이에 대한 제어는 용이하지 않다. 최근 들어, 전 세계적으로 정부나 중앙은행이 편제하는 공식통계의 품질제고가 요구되는 상황에서, 한국은행의 서베이통계에서 발생하는 비표본오차에 대한 원인 및 개선방향을 모색하는 것은 매우 의미 있는 것으로 판단된다.

〈표 1-1〉 **서베이통계 품질관련보고 포함요건**

(1) 자료원(data source), 정의, 분류 등과 관련된 기본정보
(2) 조사모집단 정의, 표본틀 적정성 등 조사자료 포괄범위
(3) 표본선택 및 추정방법 기술
(4) 응답률(정의포함)
(5) 표본오차의 추정방법 및 관련측도 해석방법
(6) 발생 가능한 주요오차의 크기 및 방향은 물론 이들의 상대적 중요성 및 통계량에 미치는 영향 언급
(7) 통계량의 시계열적 비교에 영향을 미칠 조사절차의 주요변화 및 요인에 대한 언급
(8) 동일 주제에 대한 타 통계량과의 비교언급
(9) 자세하고 구체적 기술적 절차에 대한 보고 및 참고자료 언급

2. 총조사품질·총조사오차

경제협력개발기구(Organization for Economic Cooperation and Development, OECD)는 기업경기조사를 위한 지침서(OECD, 2003, 21쪽)에서 조사품질의 결정요인으로 다음의 5가지를 제시하고 있다. 먼저, 조사결과가 모수를 얼마나 정확히 측정하는지를 나타내는 신뢰성(reliability), 조사결과가 정책 결정자와 분석가에게 유용하게 적절한 시점에 발표되는지를 나타내는 적시성(timeliness), 조사결과가 표본의 범위, 분류 혹은 정의 등에 의한 단절이 없이 시계열적으로 일관성이 있는지를 뜻하는 비교가능성(comparability), 정보 이용자가 조사 실시방법 및 결과 도출 방법에 대해 알 수 있는지를 나타내는 투명성(transparency), 정보 이용자가 조사결과를 인쇄물 또는 전자파일 형태로 쉽게 얻을 수 있는지의 여부인 접근성(accessibility) 등이다.

이 중 조사품질을 결정하는 가장 중요한 요인은 신뢰성으로 총조사오차의 개념으로 평가될 수 있다. 총조사오차란 표본추정량과 모수 간의 제곱거리의 기댓값인 평균제곱오차(mean square error, MSE)로 정의된다.

$$MSE(\hat{\theta}) = E(\hat{\theta}-\theta)^2 = \sigma^2(\hat{\theta}) + bias^2(\hat{\theta})$$

여기서 기댓값 $E(\bullet)$는 반복적인 표본추출로 인해 얻게 되는 추정치의 평균값, $\sigma^2(\hat{\theta}) = Var(\hat{\theta})$은 추정량 $\hat{\theta}$의 분산, $bias(\hat{\theta}) = E(\hat{\theta}) - \theta$는 편향(bias)을 각각 나타낸다.

총조사오차는 조사설계, 자료수집, 자료처리 및 자료분석 등의 일련의 과정에서 발생할 수 있는 오류들을 모두 합하여져 누적된 오차(Biemer, 2010)를 뜻한다. 이는 조사품질을 결정하는 하나의 구성요소로서 조사절차의 운영적 측면에 주안점을 두는 개념이라고 할 수 있다. 조사주기별로 발생할 수 있는 잠재적인 오차 요소들을 식별할 수 있다면 주어진 예산 하에서 총조사오차 관점에서 최적의 조사품질을 달성할 수 있는 오차 제어 프레임을 제시할 수 있을 것이다. 총조사오차는 일반적으로 대표성(representation) 오류와 측정(measurement) 관련 오류의 두 영역으로 구분될 수 있다. <그림 1-1>은 조사주기(life cycle of survey) 관점에서 두 영역별 절차 간에 발생할 수 있는 오차를 시각화하여 나타내고 있다(Groves et al., 2009, 48쪽).

대표성 오류는 표본대상을 명시하고 표본추출틀로부터 개별 개체를 추출하는 절차와 관련된 표본오차의 측면이라면, 측정관련 오류는 조사특성에 대한 개념 설정을 바탕으로 조사문항을 구체화하고 이를 설문으로 측정하는 절차와 관련된 비표본오차, 즉 시스템적 오류을 뜻한다. 총조사오차를 줄임에 있어서 품질기준의 만족과 조사비용의 경감이라는 두 가지 측면을 적절히 균형을 맞추어 고려하여야 한다.

먼저 표본오차는 표본크기와 표본설계의 선택을 통해 최소한 주요 관심변수와 관련한 모수추정의 정확성으로 조절할 수 있다. 예를 들어, 한국은행의 대표적 서베이통계의 하나인 소비자동향조사(CSI)의 (응답)표본크기는 약 2,200가구이며 매월 조사되고 있다. 2012년 3월~6월 기간 중 추정한 CSI 주요 조사항목의 상대표준오차는 0.69~1.93%(표준오차는 0.82~1.65)의 범위를 갖고 있어 매우 안정적임을 알 수 있다(박인호·황희진, 2012). 정부승인통계의 경우 일반적으로 3%내외의 상대표준오차(혹은 변동계수)를 목표정도로 설정하고 있는 점을 감안할 때 표본크기 2,200가구를 대상으로 조사하는 소비자동향조사는 표본오차의 관점에서 볼 때는 충분한 정확성을 확보하고 있다고 평가할 수 있다.[1)]

반면 비표본오차는 체계적 오류(systematic error)로 측정이 용이하지 않고 제어가 되지 않

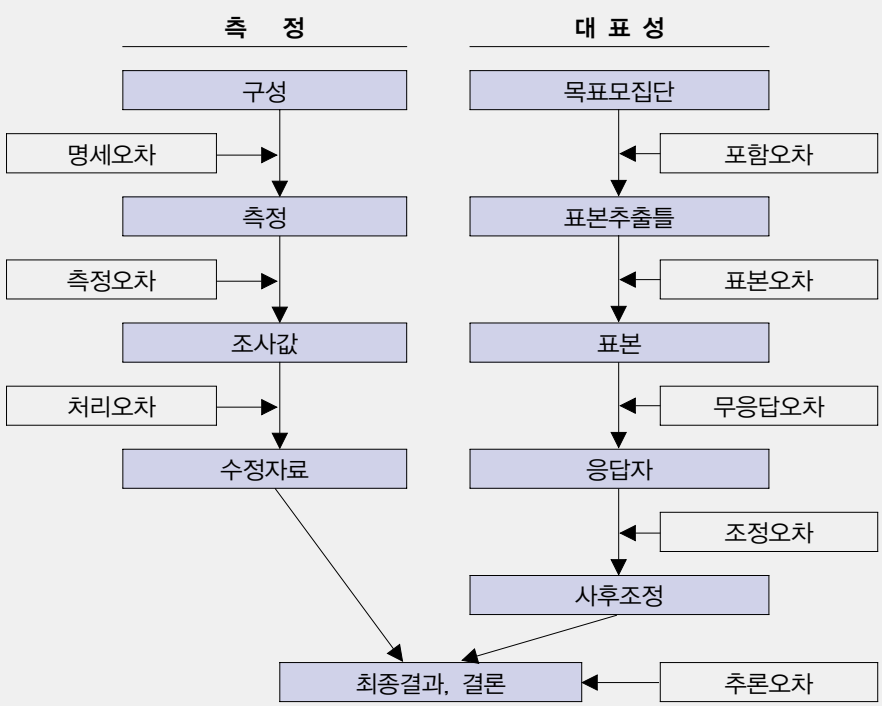
1) 대부분 국가에서 수행하는 소비자동향조사와 기업경기조사의 표본크기는 1,000~2,000개 범위를 갖는다(UN, 2014). 표본수를 늘이면 상대표준오차가 감소하지만 표본수의 제곱근에 반비례하게 된다. 예로, 모비율 추정치가 단순추출에서 갖는 상대표준오차는 $cv(\hat{p}) = \sqrt{(1-p)/(np)}$로 표본수를 2,000개에서 4,000개로 두 배로 늘린다 하더라도 상대표준오차는 30%정도만 감소하게 된다. 물론 지역별, 직업별, 업종 등과 같이 세부영역의 표본오차는 다소 큰 값을 가질 수도 있다. 정확성과 표본규모와의 관계에 대한 상세한 내용은 Cochran(1977)이나 UN(2014, 5장)을 참고할 수 있다.

았을 때 특정한 방향성을 갖는 매우 큰 편향(bias)를 가질 수 있어(Scali et al., 2005) 소비자동향조사 등 한국은행의 주요 서베이통계에서도 이에 대한 꾸준한 원인 규명과 평가가 필요하다.

3. 내용구성

본 연구는 한국은행의 서베이통계와 관련하여 발생할 수 있는 비표본오차의 특성들을 총조사오차의 관점에서 살펴보고자 하였다. 비표본오차와 관련한 기존 연구를 검토하고 한국은행 서베이통계와 관련한 논점에서 살펴보고자 하였다. 2장은 관심대상에 대한 표본추출틀의 포함오차를 다루며, 3장에서는 표본개체들이 조사에 참여하지 않음에 따른 무응답오차를 다룬다. 4장은 기타 비표본오차들로 명세오차, 측정오차, 처리오차에 대해 각각 간단히 살펴본다. 5장에서는 비표본오차 축소를 위한 체크리스트 예시와 기관차원의 서베이통계에 대한 지속적 품질향상을 가능하게 하기 위한 고려사항을 논의하였다.

〈그림 1-1〉 조사주기 관점의 총조사오차 구조



Ⅱ. 포함오차

1. 개요

조사자료(survey data)는 응답개체들에 대한 내용을 포함하지만, 일반적으로 보다 큰 포괄 범위를 갖는 개체들의 모임인 모집단을 연구하는데 사용된다. 조사표본이란 일반적으로 '잘 정의된 개체들의 모임'과 같이 교과서에 기술된 (예, 도시인구, 학생, 사업체 등의) 모집단으로부터 추출되는 확률적 표본이다. 하지만 실제 조사에서는 이보다 복잡한 문제에 당면하게 되며, 특히 모집단과 관련하여 조사대상자 포함과 관련한 오차가 존재한다.

포함오차는 연구 관심대상인 모집단과 연결하여 논의되어야 한다. Kish(1979)의 분류에 따른 모집단 종류를 살펴보면 다음과 같다. 먼저, 추정모집단(inferential population)이란 연구 대상 개체들을 일컬으며 연구 관점의 시간 측면에서 명백한 범위를 갖게 된다. 예로, 특정 기간 동안의 소비자 행태에 관심이 있다면 해당 기간 중 무수히 많은 시점에서 생존한 소비자들의 행태를 측정한다고 할 것이다. 이러한 경우 추정모집단은 관심변수들이 갖는 모형적 관계로 표현될 수 있는데, 조사연구 분야에서는 이를 초모집단(super-population)이라고 한다. 목표모집단(target popula- tion)은 조사 관심 대상으로 조사 자료를 통해 추론하고자 하는 모든 기본단위들의 집합으로 이는 유한개의 단위로 구성되고 시간적 제약을 가지며 관측 혹은 접근이 가능해야 한다. 예로, 소비자동향조사는 매월 15일이 포함된 일주일 동안 조사대상 소비자들에 대해 경제인식 및 전망 등을 조사한다. 조사모집단(survey population)은 목표모집단에 대해 접근 가능하도록 구축 혹은 선택할 수 있는 표본추출틀 혹은 명단으로 포함되는 단위들의 집합을 뜻한다. 예로 생성 및 소멸이 빈번한 사업체들을 대상으로 하는 조사에서 창업 이후 일정기간이 지난 이후에야 사업체명단에 포함된다. 추출틀모집단(frame population)은 표본이 실제로 추출되는 단위들의 목록이다. 대부분 개인, 가구, 지역지도, 사업체 등의 목록이며, 그 형태는 문서철이나 전자파일 등이다.

표본추출틀의 불안정성 종류에 따른 포함오차는 다음과 같이 구분할 수 있다. 먼저, 과소포함(under-coverage)은 목표모집단의 일부 단위들만 표본추출틀에 포함되고 제외된 개체에 대해서는 접근할 수 없는 경우에 발생한다. 과소포함은 발견하기 어려운데, 예를 들면, 전화조사에서 전화가 없는 개인은 전화번호부에 미등재되어 조사가 불가능하고 이러한 대상이 얼마나 많은지 파악하기는 일반적으로 쉽지 않다. 과대포함(over-coverage)은 목표모집

단에 속하지 않은 단위들이 표본추출틀에 포함되어 있는 경우에 발생한다. 예로, 사망자 혹은 폐업한 사업체가 여전히 전화번호부에 등재되어 있어 조사단위 목록에는 존재하지만 실제 존재하지 않는 경우이다. 중복기재(duplicate listing)란 목표모집단의 단위가 한번 이상 표본추출틀에 존재하는 경우이다. 예로, 한 사람이 여러 개의 전화번호를 갖고 있어 마치 여러 사람이 존재하는 것처럼 나타난다. 부정확한 보조정보는 표본추출틀에 제공되는 보조정보가 부정확한 경우이다. 예를 들면, 등록자료에 있는 사업체의 규모가 잘못 조사 혹은 기재되어 있는 경우이다.

2장에서는 포함오차를 다루며 상세한 논의순서는 다음과 같다. 2.2절에서는 목표모집단의 특성에 대해 추정하는 기술통계(descriptive statistic)와 관련하여 표본추출틀의 과소 혹은 과대 포함에 따른 포함오차와 일부 개체들의 중복포함에 따른 비효율성을 논의하고 이에 대한 축소방안을 논한다. 2.3절에서는 추정모집단 혹은 연구모형에 대한 분석통계에 있어서 포함오차와 관련된 오차에 대해 논의한다.

2. 기술통계 관련

가. 기본정의

포함오차의 기술을 위해 다음의 기본적인 기호와 정의를 사용하고자 한다.

U_T = 목표모집단

U_F = 추출틀모집단

s_F = 추출틀로부터의 확률표본

N_T = 목표모집단의 크기

N_F = 추출틀모집단의 크기

$\overline{Y}_T$ = 목표모집단 내 관심특성 y-평균

$\overline{Y}_F$ = 표본추출틀 내 관심특성 y-평균

나. 과소포함

표본추출틀이 목표모집단 전체를 포괄하지 못할 때 과소포함에 의한 포함오차가 발생한

다. 과소포함은 비포함 편향(non-coverage bias)의 형태로 모수추정에 대한 영향력을 평가할 수 있다.

논의를 단순화하기 위해 $U_F \subset U_T$, 즉 표본추출틀이 목표모집단의 부분집합인 $N_T > N_F$의 경우를 가정한다. 목표모집단의 관심특성 y-평균은 표본추출틀에 포함된 개체들과 포함되지 않은 개체들이 갖는 평균들의 가중평균으로 다음과 같이 분해되어 표현될 수 있다.

$$\overline{Y}_T = W_F \overline{Y}_F + (1 - W_F)\overline{Y}_{NF} \tag{2-1}$$

이때 표본추출틀의 포함률(coverage rate)은 $W_F = N_F / N_T$으로 정의되고 $\overline{Y}_{NF}$는 표본틀에 포함되지 않는 개체들의 평균을 나타낸다. 비포함 편향은 표본추출틀에 포함된 개체들과 포함되지 않은 개체들이 갖는 관심특성 y-평균 간의 차이에 비포함률을 곱한 형태로 다음과 같이 표현된다.

$$B_{NF} := \overline{Y}_F - \overline{Y}_T = (1 - W_F)\left(\overline{Y}_F - \overline{Y}_{NF}\right) \tag{2-2}$$

예를 들면, 평균 가구소득에 대해 조사한다면 표본추출틀의 비포함률$(1 - W_F)$이 20%이고 표본추출틀에 포함된 가구들이 그렇지 않은 가구들에 비해 평균소득이 1천만원이 많다면, 포함오차로 발생하게 되는 비포함 편향은 200만원(=20%×1천만원)이 된다. 따라서 불완전한 표본추출틀로부터 표본을 선택함에 따라 평균 가구소득이 200만원이나 과대하게 추정된다.

비포함 상대편향(non-coverage relative bias)도 유용한 측도가 되는데, 이는 비포함 편향에 대해 모평균을 기준으로 상대적인 값으로 다음과 같이 편향을 정의한다.

$$RB_{NF} := \frac{\overline{Y}_F - \overline{Y}_T}{\overline{Y}_T} = (1 - W_F)\left(\frac{\overline{Y}_F - \overline{Y}_{NF}}{\overline{Y}_T}\right) = \frac{\overline{Y_F}}{\overline{Y}_T} - 1 \tag{2-3}$$

앞선 예에서 목표모집단 평균값은 4천만원으로 (전년도 조사에서) 알려져 있다면, 비포함 상대편향은 5%(=2백만원/4천만원)이 된다.

간단히 정리하면, 비포함 편향은 다음의 두 요소, 즉 (ⅰ) 비포함률 $1 - W_F$와 (ⅱ) 표본틀 포함여부에 따른 개체간 특성차이 $\overline{Y}_F - \overline{Y}_{NF}$에 의하여 결정된다.

다. 과대포함

표본추출틀이 목표모집단은 물론이고 그 이외의 개체들도 포함하는 경우에 과대포함오차가 발생한다. 과대포함이 모수추정에 미치는 영향력은 표본조사로 파악한 부적격율로 추측할 수 있다. 과대포함, 즉 $U_F \supset U_T$ 경우, 표본추출틀 (혹은 추출틀모집단)은 목표모집단과 그 외 부분의 두 영역(domain)으로 다음과 같이 구분되어 나뉜다.

$$U_F = U_T \cup U_{F/T} \tag{2-4}$$

여기서 $U_{F/T} = U_F - U_T := U_{IE}$는 표본추출틀의 과대포함 부분, 즉, 표본추출틀 내에서 목표모집단에 속하지 않는 조사 부적격(ineligible, IE) 개체를 나타낸다. 만약 표본추출틀이 (2-4)와 같이 구성된다면, 표본추출틀 내 목표모집단 이외의 개체들의 상대적 비중은 부적격률(ineligible rate)로 불리며 다음과 같이 정의된다.

$$R_{IE} = \frac{N_F - N_T}{N_F} = 1 - \frac{N_T}{N_F} = \frac{N_{IE}}{N_F} \tag{2-5}$$

여기서 $N_{IE} = N_F - N_T$는 표본추출틀 내 부적격 개체수를 나타낸다. 따라서 서베이통계의 부적격률은 표본추출틀로부터 추출된 확률표본 s_F를 사용하여 영역추정(domain estimation)의 형태로 추정할 수 있다.

라. 중복포함

표본추출틀에는 때로는 특정 개체가 중복되어 포함될 수도 있다. 표본추출틀 내 중복개체들에 대한 제거 작업(unduplication) 이후에도 남아 있는 중복포함의 문제는 가중치 조정(weight adjustment)을 통해 표본에 추가적으로 반영될 수 있다.

가장 적절한 방법으로는 중복된 개체가 표본에 포함될 수 있는 결합추출확률(joint selection probability)을 산출하여 표본가중치 산출에 직접적으로 고려하는 것이다. 예로, i번째 표본개체는 표본추출틀 상에서 총 m_i개의 나열 개체들과 연관된다고 가정한다. 만약, 개별개체들의 표본추출은 독립적이고 π_{ij}이 표본개체 i에 대한 j번째 나열 개체가 갖는 추출확률(selection probability)이라면, 표본개체 i의 추출확률은 다음과 같이 계산될 수 있다.

$$\pi_i = 1 - \prod_{j=1}^{m_i} (1 - \pi_{ij}) \tag{2-6}$$

표본개체 i의 추출확률이 π_i일 때, 해당 개체 추출확률의 역수인 $w_i = 1/\pi_i$은 설계가중치(design weight)가 되며 이는 개체 i가 대표하는 모집단 내 유사한 개체수에 해당한다. 예로, 표본추출틀에서 표본개체별로 추출확률이 $1/10$으로 추출되었지만 i번째 개체만 추출틀에 두 번 나열되었다고 하자. 중복포함을 무시한 설계가중치는 $w_i = 10$이지만 중복포함을 반영하면 설계가중치는 $w_i = 100/19 = 5.26$이 된다. 이는 식 (2-6)에 따라 개체 포함확률은 $\pi_i = 1 - (1 - 1/10)^2 = 19/100$이기 때문이다. 따라서 중복포함을 무시할 때 해당 개체의 조사값이 갖는 대표성은 약 2배 정도로 과대하게 계상됨을 알 수 있다.

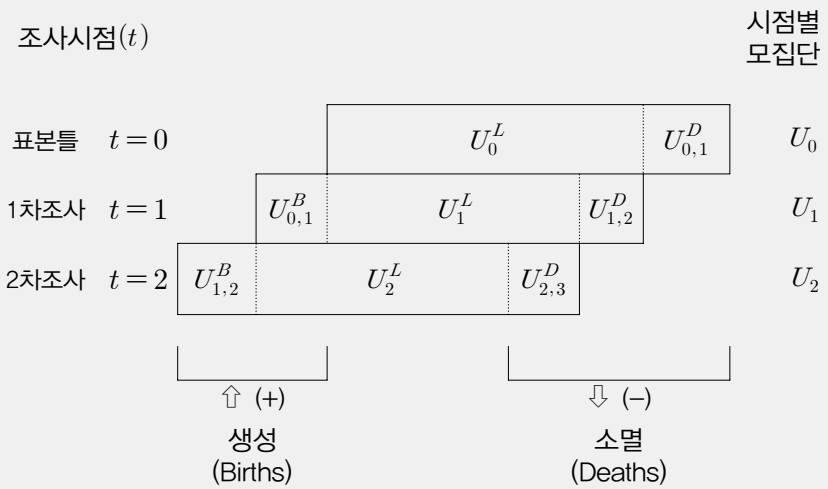
마. 반복조사의 포함오차 누적

모집단 내 개체의 구성과 특성은 지속적으로 변화한다. 특히 일반적인 사회조사 대상에 비해 사업체 모집단에서 이러한 변화는 매우 두드러지게 나타나기도 한다. 예로, 경기상황에 따라 폐업, 휴업, 합병 등의 이유로 인한 소멸(deaths)이 발생하며 새로운 사업체들이 생성(births)되고, 고용규모의 확대 혹은 축소 등의 규모변동도 발생하거나 업종변경 등의 특성 변화도 발생한다.

시간흐름에 따른 모집단 변동이 많음에도 불구하고 표본추출틀로부터 추출한 표본개체에 대해 여러 시점에 걸쳐 반복적으로 조사를 수행한다면 표본추출틀의 노후화로 인해 누적된 변동은 심화될 수 있다.

<그림 2-1>은 시점 간 생멸 누적에 따른 포함오차의 변동을 보여주고 있다(박인호, 2015). 만약 조사시점 이전($t = 0$)에 구성된 표본추출틀 U_0으로부터 최초년도 표본 s_1을 추출하고 이를 T개의 시점에 걸쳐 반복조사를 수행한다면 표본추출틀의 노후화가 진행됨에 따라 시점별로 $U_{0,1}^D$, $U_{1,2}^D$, $U_{2,3}^D$의 소멸개체로 인한 과대포함의 영역을 포함하게 되며 $U_{0,1}^B$, $U_{0,1}^B$ 등의 생성개체로 인한 과소포함의 영역들이 각각 누적된다. 따라서 특정시점의 표본추출틀로부터 표본추출 후 표본추출틀의 갱신을 반영하지 못하고 고정된 표본에 대해서 반복적으로 조사를 진행한다면 모집단 내 생성과 소멸에 의한 과소·과대포함으로 오차가 누적될 위험성을 내포함을 알 수 있다.

〈그림 2-1〉 생멸 누적에 따른 포함오차 변화



바. 포함오차 축소방안

다수의 조사를 수행하거나 동일 조사를 여러 시점에 걸쳐 반복적으로 수행하는 경우라면 양질의 표본추출틀을 구축하고 이에 대해 지속적으로 유지하고 보수하는 일은 매우 중요하다. 생성과 소멸, 규모와 주요활동의 변동이 심한 사업체 대상의 조사를 위한 표본추출틀의 유지 및 보수는 특히 중요하다. 표본추출틀의 구성과 유지에 대한 상세한 논의는 Wright and Tsao(1983)을 참고할 수 있다. 생멸과 규모 등의 정보관리가 잘 되는 경우에는 영구난수(permanent random number, PRN) 혹은 조정난수(collocated random number, CRN) 방법을 활용한다면, (i) 표본설계를 통해 모집단 변화를 잘 반영할 수 있게 되며, (ii) 반복조사 혹은 중복조사 등의 기관 내 조사수행에 따른 사업체 응답경감 및 추정효율성 제고 등의 다양한 이점을 살릴 수 있게 된다. 이에 대한 상세내용은 Hidirogolou and Lavellee(2009)를 참조할 수 있다.

또한 다중추출틀(multiple-frame) 기법을 고려하면 개별 표본추출틀이 갖는 과소포함의 문제를 줄일 수 있다. 다중추출틀 기법은 표본설계에 적용되며 적절한 추정방법과 함께 고려되어야 한다. 상세한 내용은 Lohr(2010, 514-516쪽)을 참고할 수 있다.

사. 포함오차 평가에 대한 해외사례

영국통계청(ONS)은 품질점검표의 <표 2-1>와 같은 포함오차 관련 문항 진단을 통해 포함오차의 축소를 지향하고 있다.

〈표 2-1〉 **영국통계청의 포함오차 관련 품질점검표 문항 예시**

조사범위	
목표모집단	
	조사의 목표모집단은 무엇인가?
	모집단의 조사단위를 파악하여 분류하는데 근거가 되는 기준은 마련되어 있는가?
	목표모집단에 대한 표준분류가 있다면, 이에 쉽게 접근할 수 있는 방법은 무엇인가?
조사모집단	
	조사의 실제 대상모집단은 무엇인가?
	목표모집단과 조사모집단이 어느 정도로 근접한가?
추출틀	
	표본조사의 경우 표본추출에 사용된 추출틀은 무엇인가?
	시간 경과에 따라 추출틀에 변동이 있는가? 변동이 있다면 어떻게 변동되는가?
	조사모집단 내부에서 발생, 소멸등과 관련하여 추출틀이 갱신되는가?
	보고서에 추출틀의 주요변수에 대한 요약된 통계표가 수록되는가?

참고: 한국보건사회연구원 (2007) 자체품질진단 매뉴얼 (20-21쪽)

3. 분석통계 관련

가. 연구모형과 분석통계

서베이자료를 이용하여 자료분석을 수행할 때 연구자가 흔히 간과하는 것은 표본설계가 유한한 모집단인 표본추출틀로부터 추출되었다는 점이다. 연구자는 서베이자료에 대해서도 조사대상의 특성간 관계에 대한 확률과정(stochastic process)에 의한 관측치로 여긴다. 따라서 조사자료 분석을 통해 모형(model)의 모수를 추정하고자 한다.

나. 포함오차·조정 - 예시 1

본 절에서는 Groves(1989, 3.3.3절)가 논의한 계량경제학적 연구결과[2]를 응용한 분석모형

과 관련된 포함오차에 대해 살펴보고자 한다. 참고로 계량경제학(Econometrics)에서는 목표모집단과 조사모집단은 각각 원모집단(original population)과 선택모집단(selection population)[3]을 뜻하며 선택편향(selection bias)은 이들 간의 차이가 된다.

조사자료를 이용하여 다음의 단순회귀모형 분석을 한다고 가정하자.

$$y_i = \alpha + \beta x_i + \epsilon_i \tag{2-7}$$

여기서 y_i와 x_i는 각각 조사모집단 내 i번째 개체의 종속변수(외생변수)와 독립변수(내생변수) 값이고, α와 β는 각각 선형회귀식의 절편과 기울기이며, ϵ_i는 오차항을 나타낸다. <그림 2-2>은 추출틀모집단을 구성할 때 목표모집단의 일부인 종속변수가 특정값 미만 $(y < y_{cut})$인 경우를 체계적으로 제외할 때 발생할 수 있는 포함오차를 회계계수 추정에 대한 영향력 측면에서 나타내고 있다. 목표모집단에 대해 선형회귀적합식은 검정색 직선으로 표시되어 있고, 추출틀모집단에 대해 선형회귀적합식은 파란색 점선으로 표시된다. <그림 2-2>의 예시를 위해 $N = 1,000$개의 두 변수 (x, y) 값을 평균 (100,100), 분산 (90,90), 공분산 65를 갖는 이변량정규분포로부터 확률난수를 생성하였다. 목표모집단과 추출틀모집단의 선형회귀적합식은 각각 $\widehat{y_T} = 31.22 + 0.6890\,x$과 $\widehat{y_F} = 39.92 + 0.6087\,x$이며 절사점 85에 의한 제외로 인해 선형회귀식의 기울기가 약 88.3% 정도 과소하게 추정되는 것을 나타낸다. 절사점에 의한 표본추출틀의 포함률은 $W_F = 95.9\%$이고 표본추출틀의 포함여부에 따른 평균차이는 $\overline{Y_F} - \overline{Y_T} = 101.5728 - 100.5108 = 1.0621$이며 상대차이는 1.057%이다. 즉, 평균추정에 대한 포함오차는 그다지 두드러지지 않음을 알 수 있다. 두 변수 x와 y가 모두 정규분포를 따를 때 Goldberger(1981)는 목표모집단의 회귀계수 β_T와 추출틀모집단의 회귀계수 β_F간의 관계식을 다음과 같이 유도하였다.

$$\beta_F = A_{y\zeta}\,\beta_T \tag{2-8}$$

여기서 $A_{y\zeta} = \dfrac{\sigma_{yF}^2/\sigma_{yT}^2}{1 - \rho_{xyT}^2\left(1 - \sigma_{yF}^2/\sigma_{yT}^2\right)}$

σ_{yF}^2 = 추출틀모집단의 y-변수 분산

2) Tobin(1958), Heckman(1979), Goldberger(1981), Berk and Ray(1982)

3) 계량경제학에서는 무응답 혹은 조사비협조 등으로 인한 측정실패 또한 선택편향을 야기하는 대상자로 인식됨

σ^2_{yT} = 목표모집단의 y-변수 분산

ρ^2_{xyT} = 목표모집단에서 회귀식 결정계수

관계식 (2-8)에서 알 수 있듯이, 모평균 기준의 포함편향과는 달리 목표모집단과 추출틀 모집단의 상대분산과 목표모집단에서의 모형적합정도에 영향을 받는다. 과소포함으로 인한 분석모형의 적합편향(fitting-bias)은 식 (2-8)이 보여주듯이 목표모집단의 모형적합이 좋을수록 작아진다. <표 2-2>는 두 변수 간 상관관계 ρ_{xyF}의 정도에 따른 β_F, β_T, β_F/β_T 과 평균값에 대한 비포함 상대편향을 각각 보여준다. 결정계수값이 커질수록 β_F/β_T의 값이 1로 편향이 작아지지만 평균값에 대한 비포함 상대편향은 일정한 값을 유지하고 있다.

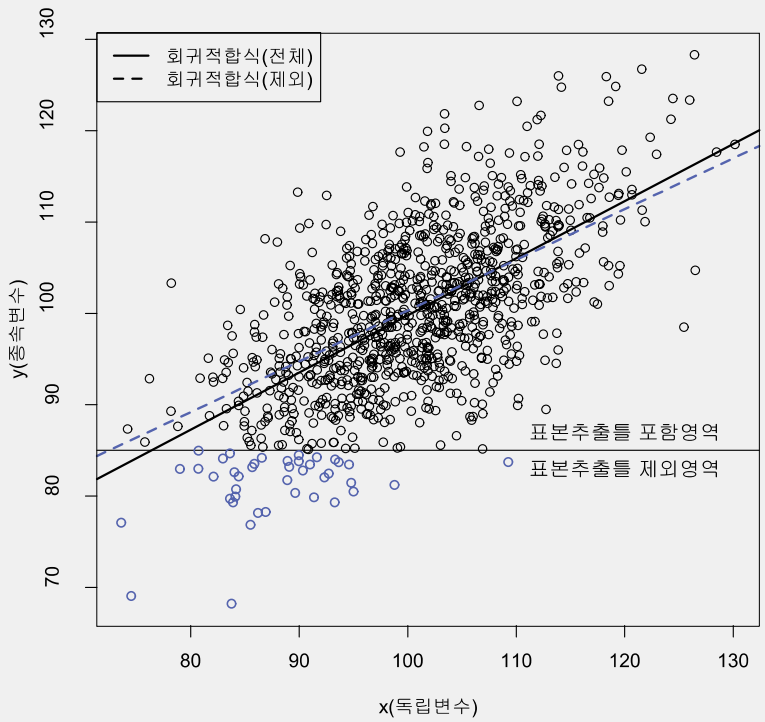
〈표 2-2〉 **모형적합정도에 따른 선형회귀계수와 평균값의 비포함 상대편향**

모수	예시값				
ρ_{xyF}	0.1006	0.2949	0.5093	0.7045	0.9102
β_F	0.2403	0.4344	0.6087	0.7544	0.9138
β_T	0.3024	0.5207	0.6891	0.8168	0.9398
$\beta_F/\beta_T(=A_{y\zeta})$	0.7948	0.8343	0.8833	0.9236	0.9724
RB_{NF}	0.0104	0.0109	0.0106	0.0102	0.0103

다. 포함오차·조정 - 예시 2

앞서 살펴본 예시보다 조금 더 현실적인 비포함 사례들은 종속변수의 값에 의한 다기보다는 다른 과정의 결과로 이루어질 수 있다. 예로, 명부작성(listing)의 기준으로 상시근로자 30인 이상과 같은 것이 고려될 수 있다.

〈그림 2-2〉 절사추출에 의한 회귀추정계수의 영향력



Heckman(1979)은 이러한 비포함을 부수적 선택(incidental selection)이라 칭하였는데, 이를 반영하기 위해 추출틀 구성과 모형분석을 위한 각각의 선형모형을 다음과 같이 고려하였다. 목표모집단이 주어질 때 이를 연구하기 위한 표본추출틀의 구성을 위한 종속변수는 u_i이고 z_i는 관련된 독립변수이며, 분석모형의 종속변수와 독립변수는 각각 y_i와 x_i이다.

$$u_i = \gamma + \delta z_i + \upsilon_i \qquad \text{(구성변수)} \tag{2-9}$$
$$y_i = \alpha + \beta x_i + \epsilon_i \qquad \text{(분석모형)}$$

물론 부수적 선택모형에 의한 비포함 편향은 식 (2-9)의 편향요소는 물론 두 오차변수 υ_i와 ϵ_i의 공분산 구조에 영향을 받게 된다. 논의의 단순화를 위해 예시는 생략하였고 상세한 내용은 Berk and Ray(1982)를 참조할 수 있다.

Ⅲ. 무응답오차

1. 개요

무응답은 추출된 표본개체 중 일부로부터 자료를 얻지 못할 때 발생하며 이로 인한 오차를 무응답오차라 한다. 무응답에는 표본개체가 조사에 전혀 응하지 않는 단위무응답(unit nonresponse)과 일부 항목만 응답하는 항목무응답(item non- response)으로 나뉠 수 있는데 본 연구에서는 단위무응답에 대해서만 다루고자 한다.

무응답이 발생하면 이로 인해 (i) 응답수가 줄게 되어 추정분산의 증가를 초래하여 표본오차가 커지며, (ii) 응답과 무응답 개체 간에 존재하는 평균적 특성 차이는 잠재적인 편향(bias)인 비표본오차의 증가를 가져올 수도 있다.

3장에서는 무응답과 관련한 비표본오차를 논의하고자 하며 그 순서는 다음과 같다. 3.2절에서는 단위무응답과 관련하여 발생원인, 조사단계별 단위무응답의 축소방안, 단위무응답 축소를 위한 조사설계, 단위무응답으로 인한 편향, 무응답 메카니즘, 조정, 그리고 평가에 대해서 차례로 다룬다. 3.2절의 내용은 박인호(2013, 96-106쪽)를 참조하였다. 3.3절에서는 기업들에 대한 한국은행의 기업경기조사와 관련한 무응답오차의 평가 및 처리, 모형기반의 무응답 편향의 평가를 차례로 다룬다.

2. 무응답

가. 발생원인

(단위)무응답의 발생요인은 일반적으로 크게 두 가지로 구분할 수 있다. 첫째, 표본개체에 대해 접촉에 실패하였거나 접근이 불가한 경우이다. 예로, 조사원이 방문하거나 혹은 전화를 걸었을 때 표본가구원이 집에 없거나, 조사기간의 제약으로 인해 더 이상 조사 시도를 할 수 없을 때, 혹은 주소나 전화번호 등의 정보가 부정확하거나 오래되어 쓸모없게 된 경우 등으로 발생한다. 둘째, 표본개체에 대한 조사 설득이 실패하는 경우이다. 예로, 조사대상자가 응답할 의사가 없거나, 주택출입관리자가 조사원의 접근을 막거나, 전화조사에서 최초 수신자가 선택된 표본가구 내 가구원을 바꾸어주지 않는 경우 등에 해당한다.

나. 조사단계별 단위무응답 축소를 위한 방안

단위무응답을 줄일 수 있는 조사단계별 도구목록(Groves et al., 2009)은 <표 3-1>와 같다.

〈표 3-1〉 **단위무응답 축소를 위한 조사단계별 도구목록**

단 계	도 구 목 록
접촉 (contact)	• 방문횟수와 시기 (number and timing of calls) • 자료수집기간 (length of data collection period) • 조사원 업무부담 (interviewer workload) • 조사원 관측 (interviewer observations)
초기결정 (initial decision)	• 조사원 행동 (interviewer behavior) • 후원 (sponsorship) • 사전공지 (pre-notification) • 사례 (incentives) • 부담 (burden) • 응답규칙 (respondent rule) • 가구/면접원 매치 (householder/interviewer match)
최종결정 (final decision)	• 이중추출기법의 사용 (two-phase sampling) • 면접원교체 (interviewer switch) • 조사모드교체 (mode switch) • 조사설득편지 (persuation letters) • 사후조정 (postsurvey adjustment)

다. 재조사·이중추출

무응답 대상에 대한 추적조사를 위해 이중추출법(double sampling) 혹은 이상추출법(two-phase sampling)으로 불리는 표본설계의 한 방법의 사용을 고려할 수 있다(Hansen and Hurwitz, 1946). 이 방법에서는 일상(first-phase)이라고 칭하는 일차 추출 및 조사에서 선택된 n개의 표본개체를 n_R개의 응답개체와 n_{NR}개의 무응답개체로 구분하고, 이상(second-phase) 혹은 이차 추출 및 조사에서 무응답 일부인 $100\nu\%$를 부차 혹은 부가표본(subsample)으로 추출하여 상대적으로 많은 비용이 소요되더라도 완벽한 응답, 즉 무응답 추적을 달성할 수 있도록 수행하는 표본조사 방식이다. 이중추출법을 적용할 때는 주로 일차조사에서는 우편조사와 같이 저비용의 조사모드(survey mode)를 적용하고, 무응답 대상에 대한 추적조사를 수행하는 이차조사에서는 면접원 방문에 의한 면접조사와 같은 많은 비용이 소요되는 조사모드를 적용하게 된다(Lohr, 2010). 이러한 접근은 다음의 세 가지의 가정으로 (i) 결정적 무

응답, (ii) 무응답에 대한 완전한 추적, (iii) 모드 간 측정차이의 부재 등을 전제로 하게 된다.

라. 무응답 현상·무응답 편향

무응답이 가져오는 모수추정에 대한 영향력은 무응답오차 혹은 무응답편향의 형태로 평가될 수 있다. 이는 무응답 현상에 대한 표본이론을 접목하여 고려될 수 있다. 무응답 현상은 각각 결정적(deterministic) 현상과 확률적(stochastic) 현상으로 해석할 수 있다(Bethlehem et al., 2011, Valliant et al., 2013).

먼저, 결정적 무응답이란 표본선택과 관계없이 개별개체가 조사에 협조하여 응답하는 것은 이미 고정되어 정해져 있어서 응답개체 s_R와 무응답개체 s_{NR}으로 나뉜다고 가정한 것이고, 이는 고정응답모형(fixed response model, FRM)이라 표현된다. 예로, 모평균 $\overline{Y}$의 Hajek-평균추정량 $\hat{\overline{y}}_\pi$의 편향은 다음과 같이 정의될 수 있다.

$$bias_{FRM}\left(\hat{\overline{y}}_\pi\right)=\frac{N_{NR}}{N}\left(\overline{Y}_R-\overline{Y}_{NR}\right) \tag{3-1}$$

여기서 Hajek-평균추정량은 $\hat{\overline{y}}_\pi=\Sigma_{s_R}d_{0i}y_i/\Sigma_{s_R}d_{0i}$이고, d_{0i}는 표본추출률의 역수인 설계가중치(design weight)이고, $N=N_R+N_{NR}$은 모집단 크기, N_R와 N_{NR}은 각각 모집단 내 전체 응답과 무응답 개체수, $\overline{Y}_R$와 $\overline{Y}_{NR}$은 각각 응답 및 무응답 개체의 모평균을 나타낸다. 즉, 무응답이 결정적 현상이라면 평균추정량의 무응답편향은 (i) 응답자와 무응답 간의 평균값 간의 차이가 크거나, (ii) 무응답률이 클수록 커질 수 있다는 것이다. 무응답 편향을 나타내는 식 (3-1)은 포함오차를 나타내는 식 (2-1)와 유사한 개념이라고 할 수 있다. 다시 말해, 과소포함과 무응답은 모두 정보부재의 측면에서는 동일하지만, 과소포함이 표본으로의 선택이 불가능함을 의미한 데 반해 무응답은 선택된 표본개체의 조사 비협조로 인한 자료 수집이 불가능함이라는 점에서 차이가 난다.

반면, 확률적 무응답이란 조사요청 시 개체별 협조여부는 확률적으로 결정된다는 가정 하에 확률응답모형(random response model, RRM)으로 표현한다. 예로, I_i는 개체 i가 표본에 포함되었는가를 나타내는 표본(포함)지시자이고 R_i는 응답지시자를 나타낸다면, $\pi_i=P(I_i=1)$은 포함확률이 되며 표본포함에 따른 응답확률은 다음과 같이 정의될 수 있다.

$$\phi_i = P(R_i = 1 | I_i = 1)$$

Rosenbaum and Rubin(1983)은 위의 ϕ_i를 성향점수(propensity score)라고 칭하였는데, 확률적 응답모형 하에서 Hajek-평균추정량 $\hat{\overline{y}}_\pi$의 편향은 다음과 같이 유도된다(Kalton and Maligalig, 1991).

$$bias_{RRM}\left(\hat{\overline{y}}_\pi\right) \doteq \frac{1}{\overline{\phi} N} \sum_{i \in s} \left(y_i - \overline{Y}_U\right)\left(\phi_i - \overline{\phi}\right)$$

여기서 $\overline{\phi} = N^{-1} \Sigma_U \phi_i$는 응답확률의 모평균을 나타낸다. 무응답편향은 조사변수 y와 응답확률 ϕ의 공분산과 응답확률의 모평균 역수에 비례한다. 따라서 y와 ϕ 간에 상관관계가 없다면 무응답편차도 발생하지 않는다.

마. 무응답메카니즘

무응답오차에 대한 좀 더 깊은 이해를 위하여 표본지시자 I와 응답지시자 R로 정의하는 응답확률에 조사변수 Y와 보조변수 X의 영향력 여부를 함께 고려하는 Little and Rubin (2002)의 세 가지 무응답 체계(nonresponse mechanism)을 살펴보고자 한다.

첫째, 응답확률 ϕ는 조사변수 Y와 보조변수 X에 좌우되지 않는 형태로 이러한 무응답은 완전임의무응답(missing completely at random, MCAR)이라 부른다. MCAR은 마치 무응답자가 전체 표본으로부터 임의로 확률추출된 것으로 무응답의 영향력이 무시될 수 있는 완전한 임의적인 형태의 무응답 상황임을 나타낸다. 만약 무응답이 MCAR이라면 응답자 s_R이 전체표본 s을 잘 대표하게 된다. 따라서 MCAR의 무응답 상황이라고 한다면, 무응답은 단순히 무시될 수 있고 응답자들의 자료만을 갖고 분석하면 된다는 것을 의미한다.

둘째, 응답확률 ϕ는 조사변수 Y에 좌우되지 않지만 보조변수 X의 일부 혹은 전부에 좌우되는 형태로 이러한 무응답은 임의무응답(missing at random, MAR) 혹은 무시가능무응답(ignorable nonresponse)이라 부른다. MAR의 무응답 상황은 보조변수 X에 의존하는 응답모형을 통해 충분히 설명할 수 있게 된다. 응답여부에 관계없이 모든 표본개체의 보조변수 X를 알 수 있다면 응답확률은 $P(R|Y,X) = P(R|X)$으로 표현할 수 있다. 다시 말해, 보조변수 X가 주어지면, 조사변수와 응답지시자는 서로 독립인 $Y \perp R|X$의 관계가 성립된다. MAR의 무응답 상황에서는 이후 3.2.6~3.2.8절에서 다루게 될 응답성향점수, 무응답조정, R-

지표 등의 방법들이 고려될 수 있고 적용될 수 있게 된다.

셋째, 응답확률 ϕ이 하나 혹은 그 이상의 조사변수 Y에 좌우되는 형태로 이러한 무응답은 비임의무응답(not missing at random, NMAR)이라 부른다. 자료분석에서 비임의무응답을 정확히 평가할 수 있는 방법은 없는 것으로 알려지고 있다. 따라서 무응답 메커니즘에 대한 조건부확률은 더 이상 단순화되지 않게 된다. 하지만 NMAR의 응답상황이라면 3.3.2절에서 다루게 될 무응답 편향에 대한 평가를 고려할 수도 있을 것이다.

바. 응답성향점수

표본개체의 응답확률이 주어진 보조정보 X로 설명될 수 있다는 임의무응답의 상황이라면 무응답 혹은 응답에 대한 성향은 점수로 정의할 수 있다. Rosenbaum and Rubin(1983)은 모든 표본개체 i에 대해 보조변수 X_i가 주어질 때 다음과 같이 성향점수를 추정하는 기법을 소개하였다.

$$\phi(x_i) = P(R_i = 1 | I_i = 1, X_i = x_i)$$

성향점수는 종종 로지스틱회귀(logistic regression) 모형이나 프로빗(probit) 모형을 이용하여 다음과 같이 산출할 수 있게 된다.

$$\phi(x_i) = \frac{\exp(x_i'\beta)}{1+\exp(x_i'\beta)} \text{ 혹은 } \phi(x_i) = \Phi(x_i'\beta)$$

여기서 x_i는 p개의 값들로 이루어진 보조변수 벡터이고 β는 계수벡터이며, $\Phi(u) = \int_{-\infty}^{u} (2\pi)^{-1/2} \exp(-u^2/2) du$는 표본정규분포의 누적확률함수(cumulative probability function, cdf)을 나타낸다.

사. 무응답조정

무응답에 대한 조정은 무응답 메커니즘에 대한 이해를 바탕으로 적절한 방법을 선택하고 적용하는 것이 중요하다. 무응답조정에는 크게 가중치 계층조정방법과 응답성향점수에 의한 두 가지로 나뉠 수 있다.

결정적 무응답에 의한 편향의 축소 혹은 제거 방법으로는 대표적으로 가중치 계층조정(weighting class adjustment)을 고려할 수 있다. 응답자와 무응답자의 조사특성 y가 같은 $\overline{Y}_R \approx \overline{Y}_{NR}$을 만족할 수 있는 여러 개의 계급(class 혹은 cell)을 고려하여 표본개체를 나누고 조정계급별로 설계가중치를 사용한 가중응답률의 역수를 설계가중치에 곱하여 무응답을 조정하는 방법이다. 가중치 조정계층별로 가중응답률 rr_{wc}을 계산하고 이를 조정계층 내 응답개체의 설계가중치(혹은 조정 전 가중치)에 곱하여 다음과 같이 무응답을 조정한다.

$$w_i^{FRM} = \begin{cases} 1/(\pi_i \times rr_{wc}) & i \in s_R \\ 0 & i \in s_{NR} \end{cases}$$

여기서 조정계층 U_c별 가중응답률은 다음과 같이 정의된다.

$$rr_{wc} = \frac{\sum_{i \in s_R \cap U_c} (1/\pi_i)}{\sum_{i \in s_R \cap U_c} (1/\pi_i) + \sum_{i \in s_{NR} \cap U_c} (1/\pi_i)}$$

확률적 무응답에 의한 편향 축소 혹은 제거 방법은 응답개체 s_R의 조사결과에 기초하여 산출하게 되는 총합추정량인 표본가중합의 확률적 기댓값을 통해 이해할 수 있게 된다. 개별 응답개체 i에 대해 표본추출 I_i과 응답발생 R_i을 감안한 표본가중치 w_i^*를 부여하면 표본가중합 $\hat{T}_y$과 기댓값은 다음과 같이 표현될 수 있다(Valliant et al., 2013, 318쪽).

$$E_R E_I(\hat{T}_y) = E_R E_I\left(\sum_{i \in U} R_i I_i w_i^* y_i\right) = \sum_{i \in U} w_i^* y_i E_R E_I(R_i I_i) = \sum_{i \in U} w_i^* \pi_i \phi_i y_i$$

여기서 $\hat{T}_y = \Sigma_{s_R} w_i^* y_i$ 이고 $E_R E_I(R_i I_i) = E_I[I_i E_R(R_i | I_i)] = \pi_i \phi_i$ 이다. 따라서 표본가중합이 불편추정량(unbiased estimator)이 되기 위해서는 응답개체의 무응답조정가중치는 다음과 같이 유도될 수 있다.

$$w_i^{RRM} = 1/(\pi_i \phi_i)$$

무응답 조정을 위해 고려할 수 있는 보조변수 X는 다음과 같은 세 가지의 특성을 갖추

고 있는 것이 좋다. 첫째로 응답성향을 잘 설명할 수 있어야 하며, 둘째로 조사가 목표로 하는 주요 관심변수를 잘 설명할 수 있어야 한다. 마지막으로, 통계작성을 위해 고려하는 가장 중요한 영역을 잘 나타내 주어야 한다.

아. 무응답평가

단위무응답의 주요 발생은 접촉시도와 조사설득의 과정으로 나누어 살펴 볼 수 있고, 이와 관련한 지표들은 조사결과에 대한 조사품질기준으로 종종 고려된다. 조사결과에 대한 지표들은 <표 3-2>의 조사결과분류를 바탕으로 정의할 수 있는데, 많은 조사에서 흔히 인용하는 미국여론조사학회 지침서(AAPOR, 2011)의 기준으로 살펴보고자 한다.

〈표 3-2〉 **조사평가를 위한 조사결과분류**

구 분 자	분 류	조 사 적 격 성
I	조사완료	적격
P	부분완료	적격
R	거절/중단	적격
NC	접촉불가	적격
U	적격불명	적격불명
NE	부적격	부적격
O	기타	적격

접촉률(contact rate, CR)은 표본개체 중 접촉에 성공한 개체수의 비율로 다음과 같이 정의한다.

$$CR = \frac{\text{접촉성공적격개체수}}{\text{적격개체수}}$$

$$= \frac{I+P+R+O}{I+P+R+NC+U+O}$$

접촉률 CR은 AAPOR에서 제시하는 $CON1$이라는 비율에 해당한다. 이 외에도 두 종류의 추가적인 접촉률을 제시하는데 이는 적격불명 개체에 대한 처리방식에 따라 달리 정의된다.

응답률(response rate, RR)은 조사적격개체 중 응답개체의 비율로 다음과 같이 정의할 수 있다.

$$RR = \frac{\text{응답개체}}{\text{조사적격개체}} = \frac{I+P}{I+P+R+NC+O+(e\times U)}$$

여기서 e는 적격성 여부를 아는 개체들에 기초하여 추정한 비율로 적격성을 확인할 수 없는 개체들에 대한 적격률(eligibility rate)의 추정치가 되며, RR의 분모는 표본개체 중 추정된 조사적격수에 해당된다. 응답률 RR은 AAPOR의 네 번째 응답률 $RR4$로 CASRO[4] 응답률이라고도 불리는데 가장 널리 사용된다.

오랫동안 응답률과 무응답편향은 동일시 여겨지거나 응답률이 낮은 경우 무응답편향이 커질 수 있다고 인식하였다. 하지만 최근 응답률과 무응답편향 간의 상관관계가 약할 수 있음이 지적되었고 응답률을 대체할 무응답편향에 대한 지표가 새롭게 개발되고 있다. 무응답편향과 무응답률 간의 비례적 관계로 보았던 논리의 근거는 고정응답모형에 의한 표본평균의 무응답편향을 나타내는 식 (3-1)을 통해서도 알 수 있다. 따라서 가능한 무응답률을 줄여서 무응답편향을 축소시키고자 노력하였다. Schouten et al. (2009)는 응답률에 대체할 무응답편향에 대한 지료로 R-지표(R-indicator)를 다음과 같이 제시하였다.

$$R[\phi(x)] = 1 - 2S[\phi(x)]$$

여기서 $S[\phi(x)]$는 응답성향점수 $\phi(x)$의 표준오차이고 x는 모든 표본개체에 대해 알고 있는 보조정보를 나타낸다. $R[\phi(x)]$는 전체 표본개체와 비교하여 응답개체의 응답성향점수가 유사한가를 측정한다고 할 수 있고, 1의 값에 가까울수록 응답개체의 대표성이 양호함을 뜻하며 이는 무응답 편향이 작음을 의미한다.

4) CASRO란 The Council of American Survey Research Organizations의 줄임말인데, RR은 CASRO에서 추천하는 응답률 정의임

3. 경기심리조사 무응답오차 평가

가. OECD 무응답 평가 및 처리

본 절에서는 한국은행의 기업경기조사와 관련하여 OECD (2003, 22-24쪽)가 제시하는 기업경기조사와 관련한 무응답오차의 평가와 처리를 간략히 요약하고자 한다.

무응답률(non-response rate)의 측정은 서베이를 통한 자료수집 절차의 효율성을 검토하는데 유용하게 사용되며 가능한 측도로서 단순(비가중)무응답률, 개체별 표본가중치와 규모가중치를 고려한 가중무응답률을 각각 제시할 수 있다. 단순무응답률 (r_{M1})은 조사에 불응한 사업체가 표본에서 차지하는 비중을 산출하는 것이다. 이때 휴폐업 또는 업종 변경 등은 조사 부적격으로 무응답률의 계산에서 제외된다. r_{M1}은 조사 효율성의 관점이나 균등한 표본가중치를 갖는 경우에 사용한다.

$$r_{M1} = \frac{\sum_{i \in s_{NR}} 1}{\sum_{i \in s_R \cup s_{NR}} 1} \times 100(\%)$$

표본가중치를 고려한 가중무응답률 r_{M2}은 불균등한 표본포함확률을 사용한 표본설계에서 적절한 측도이다.

$$r_{M2} = \frac{\sum_{i \in s_{NR}} w_i}{\sum_{i \in s_R \cup s_{NR}} w_i} \times 100(\%)$$

여기서 w_i는 표본가중치를 나타내는 것으로 표본추출확률의 역수로 정의된다. 표본가중치와 더불어 표본개체의 규모(예, 매출액, 근로자수 등)를 함께 고려한 가중무응답률 r_{M3}은 불균등 표본포함확률과 더불어 모집단 전체에 대한 개체 기여도를 함께 고려한다.

$$r_{M3} = \frac{\sum_{i \in s_{NR}} w_i z_i}{\sum_{i \in s_R \cup s_{NR}} w_i z_i} \times 100(\%)$$

여기서 z_i는 규모가중치를 나타낸다. 세 가지의 무응답률은 3.2.8절에서 언급한 $RR4$ 나 3.2.7절의 (가중)응답률 rr_{wc}과 동일한 기준이다. 단, 무응답률은 100%에서 응답률을 차감된다.

무응답 처리와 관련해서는 OECD(2003)는 다음의 세 가지 규칙을 제시하고 있다. 먼저, "무응답을 허용 수준까지 축소시키기 위해 충분한 자원을 써야 함"을 권하고 있다. 이는 (단위)무응답 축소방안(3.2.2절)과 재조사 혹은 이중추출(3.2.3절)에 해당된다. 둘째, "무응답이 아닌 것(휴폐업 등으로 모집단에 포함되지 않은 단위)과 무응답을 구분하여야 함"을 권하고 있다. 이는 3.2.7절과 3.2.8절에 상세히 정의된 무응답률을 적용하는 것이다. 마지막으로 "무응답 중 특이성을 가지고 있는 보고단위(매우 성공한 기업 등)들이 과도하게 혹은 과소하게 포함되어 있는지 살펴볼 것"을 제시하고 있다. 이는 r_{M2}와 r_{M3}의 가중무응답률과 함께 고려해야할 사항으로 조사자료의 모집단 대표성과 무응답조정에 따른 지나친 영향력을 배제하고자 함이다.

나. 세일러의 모형기반 무응답 편향 평가

독일 Ifo 경제연구소의 Seiler(2013)는 기업경기조사의 균형통계(balance statistic)가 갖는 경기순환선(cycle function) 추정에서의 무응답 편향에 관한 연구를 수행하였다. 대부분의 연구에서는 주로 임의무응답(MAR)이 고려되는 반면, Seiler(2013)는 경기순환 관련한 질문에 대한 3점의 리커트척도(Likert scale)인 긍정(상승,+), 보통(동일,0), 부정(하락,-)의 수준별 응답확률인 비임의무응답(NMAR)을 가정하였다.

잠재변수모형(latent class variable model)을 이용하면 균형통계와 경기순환선의 관계를 관련지을 수 있게 된다. 예를 들어, 조사문항이 제시되면 비록 조사원이 직접 관측할 수는 없지만 응답자가 내재적으로 갖게 되는 연속형적인 반응은 $y_i^*(t)$이고 설문지에 제시되는 3점 척도(긍정(+1), 보통(0), 하락(-1))에 맞추어 $y_i(t)$의 값으로 응답한다고 가정할 수 있다. 기업 i의 연속형적 반응은 경기순환식 $\beta(t)$에 영향을 받게 되며 개별적 차이에 따른 오차항 ϵ_i이 더하여져 다음과 같이 표현될 수 있다.

$$y_i^*(t) = \beta(t) + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_\epsilon^2) \tag{3-2}$$

여기서 $i = 1, \ldots, n$과 $t = 1, \ldots, T$는 각각 기업과 시점을 나타내고 σ_ϵ^2는 오차분산을 나타낸다. 연속형 반응의 기댓값은 다음과 같이 경기순환식이 된다.

$$E[y_i^*(t)] = \beta(t) \tag{3-3}$$

조사값 $y_i(t)$은 잠재변수 $y_i^*(t)$에 대해 두 개의 한계점 $-\infty < \tau^- < \tau^+ < \infty$에 의해 다음과 같이 결정된다고 가정한다.

$$y_i(t) = \begin{cases} +1 & y_i^*(t) \in (\ \tau^+, \infty) \\ 0 & y_i^*(t) \in (\ \tau^-, \tau^+] \\ -1 & y_i^*(t) \in (-\infty, \tau^-] \end{cases} \tag{3-4}$$

또한 잠재변수모형의 가정 하에서 조사값이 갖는 기댓값은 다음과 같다.

$$\begin{aligned} E[y_i(t)] &= \sum_{y=+1,0,-1} y\, P[y_i(t) = y] \\ &= (1)\, P[y_i(t) = 1] + (-1) P[y_i(t) = -1] \\ &= P\left[\beta(t) + \epsilon_i > \tau^+\right] - P\left[\beta(t) + \epsilon_i \le \tau^-\right] \\ &= \left[1 - \Phi\left(\frac{\tau^+ - \beta(t)}{\sigma_\epsilon}\right)\right] - \Phi\left(\frac{\tau^- - \beta(t)}{\sigma_\epsilon}\right) \\ &= \Phi^+(t) - \Phi^-(t) \end{aligned} \tag{3-5}$$

여기서 $\Phi(\bullet)$는 표준정규분포의 분포함수를 나타내며 $\Phi^+(t) = 1 - \Phi((\tau^+ - \beta(t))/\sigma_\epsilon)$ 이고 $\Phi^-(t) = \Phi((\tau^- - \beta(t))/\sigma_\epsilon)$으로 표기된다. 연속형 측정값은 $\beta(t)$의 기댓값을 갖는 반면, 이산형 측정값은 정보축소로 인해 $\Phi^+(t) - \Phi^-(t)$을 기댓값으로 갖는다. 가정 (3-2)와 가정 (3-4)는 다소 제한적이지만 무응답 편향에 대한 이해에는 도움이 된다. 식 (3-5)의 $E[y_i(t)]$은 무응답이 없을 때 기대되는 조사값이 갖게 되는 기댓값이다.

조사응답결정의 확률과정을 고려하면 무응답에 따른 균형통계의 무응답 편향을 도출할 수 있다. 확률응답모형(3.2.4절)으로 이항확률변수 $R_i(t)$을 가정하면, 응답에 대한 기댓값은 다음과 같이 유도될 수 있다.

$$E(y_i^{obs}) = \sum_y y\ P\left[y_i(t) = y \mid R_i(t) = 1\right] \tag{3-6}$$

$$= \sum_y y \frac{P\left[R_i(t) = 1 \mid y_i(t) = y\ \right] P\left[y_i(t) = y\right]}{\sum_y P\left[R_i(t) = 1 \mid y_i(t) = y\ \right] P\left[y_i(t) = y\right]}$$

$$= \frac{\pi^+(t)}{\bar{\pi}(t)} \Phi^+(t) - \frac{\pi^-(t)}{\bar{\pi}(t)} \Phi^-(t)$$

여기서 $\pi^y(t) = P\left[R_i(t) = 1 \mid y_i(t) = y\right]$은 y-수준의 조사대상자가 응답할 확률을 나타내며 $\bar{\pi}(t) = \Sigma_y \pi^y \Phi^y(t)$으로 정의된다. 응답값과 조사값 간의 차이의 기댓값으로 표현한 무응답 편향은 식 (3-5)와 (3-6)의 차이로 다음과 같이 표시할 수 있다.

$$E_t(y_i^{obs} - y_i) = \left(1 - \frac{\pi^+(t)}{\bar{\pi}(t)}\right)\Phi^+(t) - \left(1 - \frac{\pi^-(t)}{\bar{\pi}(t)}\right)\Phi^-(t)$$

위 식의 조건부 응답확률 $\pi^{sgn}(t)$에 대해 다양한 모형적 가정을 고려한다면 향후 현실적으로 발생할 수 있는 NMAR 무응답 상황에 대한 이해를 도울 수 있을 것으로 기대된다.

Ⅳ. 기타 비표본오차

1. 명세오차

명세오차(specification error)는 조사목적에 의한 개념이 조사항목에 함축된 개념과 다를 때 발생한다. 이러한 상황 하에서는 조사를 통해 잘못된 모수(wrong parameter)가 추정되고 이 추정값에 근거한 추론도 잘못될 것이다. 명세오류는 종종 연구자와 설문지 개발자 간의 잘못된 의사소통에 의해 발생된다.

설문지 개발자가 조사연구의 목적을 명쾌하게 이해하지 못하거나 혹은 조사 자료가 어떻게 분석되어질지를 모른다면 이로 인해 발생하는 명세오류의 존재 가능성도 인식하지 못하게 된다. 따라서 명세오류를 확인하기 위해서 연구자는 조사목적에 맞추어 조사항목들이 제대로 작성되었는지 철저히 검토하여야 한다. <표 4-1>과 같이 연구주제(research question)와 조사문항(survey question) 간의 부합성을 비교하여 연구목적들에 맞게 조사문항이 적절히 포함되었는지 또한 중복이 없는지를 점검함으로써 명세오류를 최소화할 수 있을 것이다.

〈표 4-1〉 **연구목적과 조사문항 간 부합성 점검 (예시)**

연구주제	조사문항							
	문항1	문항2	문항3	문항4	문항5	문항6	문항7	← 문항7은 불필요하므로 삭제가능
주제1	✓	✓						
주제2	✓		✓					
주제3				✓				
주제4					✓	✓		
주제5								← 주제5를 다루는 문항 부재

참고: Biemer and Lyberg (2003, 29쪽, 표 2.1) 인용

또한 사업체 조사와 관련한 명세오차 발생의 원인, 제어 그리고 측정에 대한 사례는 Freedman(1988)이 기술한 미국 에너지정보원(Energy Information Admini- stration, EIA)의 경험을 참조할 수 있다.

2. 측정오차

측정오차(measurement error)는 자료를 수집하는 과정에서 발생하며 특정변수에 대해 기록된 값과 실제 값 간의 차이를 뜻한다. 측정오차는 산출된 통계치의 편향과 변동성이 그 원인이 된다. 측정오차가 발생하게 되는 원인을 좀 더 세분하여 구분하면 다음과 같다.

① 조사표: 자료수집을 위해 사용한 조사표의 형태 또는 조사표, 측정도구 등이 잘못된 값을 기록하게 할 수 있음. 예로,
   (a) 질문의 뜻이 불분명하게 표현된 구절
   (b) 명확하지 못한 조사표의 구성
② 조사원: 응답자가 제시하는 응답에 조사원이 영향을 끼칠 수 있음
③ 응답자: 의식적이든 무의식적이든 간에 응답자가 잘못된 응답을 할 수 있음. 예로,
   (c) 정확한 응답을 위한 지식 부족
   (d) 정확히 응답하려는 동기 부족

예로, 한국은행의 기업경기조사에서 수주액, 재고수준, 업황 등의 분야에 대한 조사대상 기업들의 평가는 적정수준보다 다소 낮게 보고하는 경향이 있는 것으로 알려져 있다. 이는 독일 Ifo 연구소의 기업경기조사에서도 발견(Stnagl, 2008)되며 OECD(2003, 59쪽)의 기업경기조사 매뉴얼에서도 다음과 같이 언급하고 있다.

"응답자의 대부분은 그들의 수주액의 수준이 적정수준보다 더 낮다고 기록하는 경향이 있다. 스웨덴 기업경기조사와 관련된 ....을 보면 수주액의 밸런스[5] 평균이 몇 년간 계속 0 이하로 나타나고 있음을 알 수 있는데 이는 사실 불가능한 일이다. 또한 재고수준의 경우 스웨덴 기업체들이 시종일관 그들의 재고가 적정수준보다 높다고 대답하는 경향이 존재한다고 볼 수 있으며 이러한 경향은 나라마다 상이하다."

소비자 혹은 기업을 대상으로 경기심리를 파악하는 경기심리조사에서 이러한 측정오차는 주로 상대화 혹은 추세선을 통해 교정될 수 있는 것으로 판단한다. 예로, OECD(2003)의 매뉴얼은 다음과 같이 지적하고 있다.

5) 여기서 밸런스 통계(balance statistics)는 긍정과 부정의 응답비율 간의 차이로 정의된다.

"일반적으로 지나치게 낙관적 또는 비관적인 응답 행태는 0을 기준치로 밸런스 값을 평가하는 것보다 밸런스 값과 그것의 장기평균값을 비교함으로써 교정될 수 있다."

측정오차에 대한 평가로는 다음의 것들을 고려할 수 있다. 먼저, 반복측정에 의해 측정도구와 조사원의 영향을 측정할 수 있다. 예로, 서로 다른 조사도구나 조사원을 이용하여 반복적으로 측정하여 평가할 수 있다. 부차적으로 추출된 표본에 대해 조사도구나 조사원을 할당하여 수행할 수 있다. 특정개체 i가 조사항목에 대하여 갖는 참값을 Y_i로 표기하고 측정 혹은 응답값을 y_i로 표기하면, 조사항목의 측정에 따른 체계적 차이인 측정오차는 $y_i - Y_i$로 정의된다. 이러한 체계적 차이는 응답편의(response bias)라고도 부른다.

기업경기조사의 측정오차를 축소하는 방안으로는 앞서 언급한 세분화된 측정오차 발생원인별로도 고려가 가능할 것이다. OECD(2003, 25-26쪽)는 조사표로 인한 여러 가지 문제를 감소시킬 수 있는 방안을 제시하고 있다. 예로, 문제 (a)와 (b)는 많은 국가에서 오랜 기간 동안의 시험을 거친 이른바 "통일기준 조사표(harmonised questionnaire)"를 사용함으로 감소시킬 수 있는 것으로 보고 있다.

응답자로 인한 문제 (c)는 보고단위로 사업부[6]를 사용하고 최고경영자 혹은 이들과 밀접한 관련이 있는 협회로부터 정보를 수집함으로 줄일 수 있다고 지적한다. 더불어 조사대상으로부터 수용할 수 있는 이상의 정보를 얻으려 한다면 응답기업의 반발을 초래할 수 있다는 점을 언급하고 있다.

또한 응답자로 인한 문제 (d)를 최소화할 수 있는 가장 좋은 방법은 기업경기조사의 결과가 기업 스스로에게도 유용하다는 것을 주지시키는 것이다. 무엇보다 측정오차를 줄일 수 있는 가장 효율적인 전략은 조사 설계 단계부터 이러한 오차발생의 가능성을 제거하는 것일 것이다. 이러한 측면에서 예비 조사표, 응답자를 위한 작성지침 결과 처리 절차 등을 조사 이전에 시험해 보는 것이 무엇보다도 중요하다.

3. 처리오차

자료처리과정의 오차는 데이터 입력에서부터 편집과 도표작성 등에 이르는 일련 과정에서 발생할 수 있다. 이러한 오차는 전화조사, 우편조사, 인터넷조사 등과 같은 조사모드(survey mode)의 특성별로 체계적으로 발생할 수 있다.

6) 사업부단위(kind of activity unit)는 지역 및 업종과 관계없이 업종이 같은 사업장들의 결합으로 이루어짐(OECD, 2003, 14쪽)

처리오차와 관련하여 OECD(2003, 26~27쪽)는 다음과 같은 축소방안을 제시하고 있다. 자료수집별로 자료 입력단계에서 오차를 피할 수 있도록 하는 것이다. 예로, 전화 혹은 면접조사에는 가급적 컴퓨터의 도움을 받는 CAPI(computer assisted personal interview) 혹은 CATI(computer assisted telephone interview)의 프로그램을 적용하여 조사 수행 및 응답기록 과정에서 시스템적으로 논리적이고 일관된 관리가 가능하도록 할 수 있을 것이다.

반면, 자료 수집 방법과 관계없이 자료입력이 용이하도록 설계하는 것이다. 질문과 응답지문이 밀접하게 연관되도록 하여 질문-응답간의 애매함이 없도록 하여야 할 것이다. 조사표 설계는 완료 전에 데이터 입력을 담당하는 인력에 대한 테스트가 반드시 이루어져야 할 것이다.

한편 데이터 편집 과정에서 의외의 처리오차가 발생할 수도 있다. 자료 편집 과정은 기초 자료의 오류를 찾아내고 제거하는 것에 목적을 두지만 오히려 잘못된 조정으로 인해 오차를 생성할 수도 있게 된다. 따라서 논리적 통제(logical control)를 엄격히 하고 명백한 논리적 오차와 응답내용이 일관적이지 못한 부분을 판별함으로써 처리오차를 축소시킬 수 있다. 또한 편집과정에서 발견된 오류를 기록하여 남기는 것은 매우 중요하다. 이렇게 함으로 조사표에서 개선이 필요한 부분, 응답자에게 제공되어야 할 작성지침 등을 파악하는데 많은 도움이 될 것이다.

추가로 자료를 도표로 작성하는 과정에서의 처리오차는 잘못된 추정공식, 공식 내의 부정확한 매개변수, 잘못된 자료 처리 프로그램 등에 기인한다. 첫 번째 원인은 정확한 공식을 사용하고 자료 처리 프로그램의 전문가와 함께 작업함으로서 제거할 수 있게 된다. 두 번째 원인은 주의 깊은 교정(proof reading) 작업을 통해 제거할 수 있다. 마지막 원인은 실제 자료 처리 전에 시범 자료로 소프트웨어를 점검함으로써 이러한 오류의 제거가 가능할 것이다.

Ⅴ. 추가논의

1. 비표본오차 프로필

앞서 살펴본 바와 같이 비표본오차는 모든 조사단계에서 다양한 이유로 발생할 수 있다. 더불어 발생 가능한 모든 종류의 비표본오차를 해결할 수 있는 완벽하거나 특별한 이론이 존재하지 않는다. 다만 현재로서 고려할 수 있는 최선의 방법은 각각의 서베이에 대해 "오차프로필(error profile)", 즉 비표본오차의 주요 발생 원인들에 대해 기술하는 것부터 시작하는 것이다.

<표 5-1>은 총조사오차와 관련한 쟁점과 관련하여 Whitely (2014)이 고려한 오차 프로필의 예시를 나타내고 있다. 개별 조사별로 <표 5-1>에서와 같이 조사주기 단계별로 발생 가능한 오차들에 대해 점검 논의와 관련된 쟁점을 나열하고 이와 관련하여 서베이통계 품질에 대한 잠재적 영향력을 평가함으로써 개별 오차에 대한 제어가 가능하게 된다.

2. 조직운영 차원에서의 지속적 품질향상 노력

기관차원에서 서베이통계에 대한 지속적 품질향상이 가능하도록 하기 위해 고려해야 할 몇 가지의 필수요건(Marker and Morganstein, 2004)은 간단히 정리하면 다음과 같다.

첫째, 고위관리층부터 지속적인 품질향상(continuous quality improvement, CQI)에 대하여 업무상 우선순위에 두고 열정을 쏟아야 할 것이다. 이러한 관리자의 리더쉽이 구현되려면 품질향상 노력에 대한 평가 및 보상이 적절한 형태로 꾸준히 이루어지는 것이 바람직하다.

둘째, 조직 구조적 측면에서 서베이통계의 정확성을 향상시키기 위해서는 측정 및 무응답 오차와 관련된 이론적 훈련을 받은 전문가를 확보하여야 한다. 관련된 연구는 물론 통계업무상의 측정 및 무응답 오류를 이해하고 극복할 수 있는 능력을 배양하여야 하며, 다양한 형태와 특성의 조사오류에 대한 기관 안팎의 구성원들의 이해를 제고하여야 한다.

셋째, 지속적 품질개선에 대한 기관인식 혹은 문화는 반드시 자립적이고 자급적이어야 한다. 지도팀을 구성하여 지침 및 중간단계 교정 등을 제공하게 함으로 지속적 품질향상의 초기단계의 성공적 정착이 될 수 있도록 한다. 이를 위해 지도팀의 구성원은 최고관리층을 대표할 수 있는 자, 중간관리자, 업무수행자 등으로 구성하여 정기적 회의를 통해 품질향

상을 저해하는 장애물을 헐고, 또한 시의적절한 피드백을 받아 추가적인 지원, 훈련, 방침 이행이 필요한 지를 결정할 수 있어야 한다.

넷째, 지속적 품질향상은 조직 내 탁월한 의사소통이 전제되어야 한다. 현행최상방법론(current best methods CBM)과 같은 품질향상을 위한 최신의 방법론을 숙지하고 적용할 수 있도록 기관 간 업무 및 경험을 공유할 수 있어야 한다.

다섯째, 품질향상팀에 초급자부터 전문가까지를 아우르고 서베이통계의 세부업무별 수요자와 공급자를 포함할 수 있도록 한다. 이를 통해 조사절차에 대한 이해로부터 시작하여 업무상 투입물과 산출물은 물론 세부절차에 대한 지식을 갖는 사람들이 관여할 수 있도록 한다.

마지막으로 지속적 품질향상을 기관업무의 핵심가치로 두고 매 통계작성 과정 중에 경험하고 배운 교훈을 정리하고 후임 담당자들을 위한 편람으로 마련하고 전수하는 노력이 필요하다.

〈표 5-1〉 **서베이통계 총조사오차 및 잠재적 영향 파악을 위한 체크리스트 (예시)**

		총조사오차 점검 논의	관련 쟁점	자료품질에 대한 잠재적 영향
대표성 관련 오차				
	포함오차	조사범위 모집단이 어떻게 정의되었는가? 표본추출틀의 구성이 조사적격 모집단과 일치하는가? 표본추출틀에 부적격 혹은 중복된 개체들은 없는가?	조사범위 모집단이 부정확하거나 불량하게 정의됨 선택된 표본추출틀이 우리나라 기업체를 적절히 대표하지 못함 부적격 개체들이 표본으로 추출됨	높음
	표본오차	표본크기가 적절한가? 오차한계가 기대치보다 큰가?	표본크기가 부적절함 조사자료가 분석/보고 목적을 이루기 위해 불충분함	높음
	무응답 오차	단위무응답은 어떻게 정의되는가? 모집단 내 특정 하위집단의 무응답이 존재하는지? 항목무응답은 어느 정도인지?	무응답률이 높아 무응답 편향이 발생할 수 있음 특정 하위집단이 지나치게 많이 무응답으로 탈락함 특정 항목들에 대한 높은 항목무응답으로 편의 발생 가능	높음
	조정오차	표본가중치 산출방식이 적절한가?	가중치 산출방식이 다소 부정확함	낮음
측정 관련 오차				
	타당성	조사도구가 타당한지, 즉 신뢰할 수 있는지?	조사도구가 희망하는 개념을 측정하지 못하거나 일관적이지 못함	낮음
	측정오차	설문지 설계가 잘되었는가? 조사원의 무의식적/의식적으로 부정확한 정보를 제공가능?	잘못된 설계로 부정확하거나 불완전한 응답을 받게될 가능성 있음 조사원이 응답자로 하여금 응답값을 변경하거나 조작할 수 있음	중간
	처리오차	분석단위가 적절히 정의되었는지? 자료정제(data clean)이 어떻게 이루어지는지? 자료코드가 적절히 정의되었는지?	분석단위가 부적절히 정의됨 응답자료 타당성이 부적절히 확인 코딩 오류가 존재하거나 개방형 문항의 응답사항을 잘못 입력함	중간
	추정오차	조사자료가 바르게 분석되었나?	옳지 않은 분석기법이 사용됨 부정확한 추정 발생	낮음

참고문헌

· 박인호 (2015). 사업체패널조사의 시점 및 표본 대표성. 한국노동연구원 - 패널자료 품질 개선 연구[V], 3-22.

· 박인호·황희진 (2012). 소비자동향지수의 분산추정에 관한 연구. 국민계정리뷰, 4권, 1-19.

· 박인호 (2013). 조사모드의 효과적 활용 및 추정방법의 향후 과제. 통계개발원 정책연구용역 보고서「다양한 출처 자료 처리 및 통계 생산방안 연구」.

· AAPOR. (2011). Standard definitions: final despositions of case codes and outcome rates for surveys, 7th ed. Technical report. The Americal Association for Public Opinion Research, Deerfield, IL, URL http://www.aapor.org/pdfs/standarddefs_4.pdf.

· Berk, R. A. and Ray, S. C. (1982). "Selection biases in sociological data," *Social Science Research*, 11, pp. 352-398.

· Bethlehem, J., Cobben, F., Schouten, B. (2011). *Handbook of Nonresponse in Household Surveys*. New York: Wiley.

· Biemer, P. (2010). "Total Survey Error: Design, Implementation and Evaluation," *Public Opinion Quarterly*, 74, pp. 817-848.

· Biemer, P. and Lyberg, L. (2003), *Introduction to Survey Quality*. New York: Wiley.

· Cochran, W. (1977). *Sampling Techniques*. John Wiley & Sons, Inc. New York.

· Freedman, S. R. (1988). "Control and measurement of specification error in Establishment Surveys." ASA Proceedings, *Survey Research Methods Section*, pp. 304-308.

· Hansen, M. H. and Hurwitz, W. N. (1946). "The problem of non-response in sample surveys," *Journal of the American Statistical Association*, 41, pp. 517-529.

· Hidiroglou, M. A. and Lavelle, P. (2009). "Sampling and estimation in business surveys," *Sample Surveys: Design, Methods and Application*, Vol 29A, Elsevier.

· Goldberger, A. S. (1981). "Linear regression after selection," *Journal of Econometrics*, 15, pp. 357-366.

· Groves, R. M. (1989). *Survey errors and survey costs*. New York: John Wiley & Sons.

· Heckmann, J. J. (1979). "Sample selection bias as a specification error," *Econometrica*, 47, pp. 153-161.

· Kalton, G., Maligalig, D. S. (1991). "A comparison of methods of weighting adjustment for nonresponse," Proceedings of the U.S. Bureau of the Census Annual Research

Conference, pp. 409-428.

· Kish, L. (1979). "Populations for Survey Sampling." No. 1, February, pp. 14-15.

· Little, R. J. A., Rubin, D. B. (2002). *Statistical Analysis with missing data*. New Jersey: John Wiley & Sons.

· Lohr, S. (2010). *Sampling: design and analysis*, 2nd ed., Boston: Brooks & Cole.

· Marker, D. A. and Morganstein, D. R. (2004). "Keys to Successful Implementation of Continuous Quality Improvement in a Statistifcal Agency," *Journal of Official Statistics*, 20, pp. 125-136.

· OECD. (2003). *Business Tendency Surveys*. A Handbook. Paris.

· Rosenbaum, P. and Rubin, D. B. (1988). "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, pp. 41--45.

· Schouten, B., Cobben, F., Bethlehem, J. (2009). "Indicators for the representativeness of survey response." *Survey Methodology*, 35, pp. 101—113.

· Scali, J., Testa, V., Kahr, M., Strudler, M. (2005). "Measuring Nonsampling Error in the Statistics of Income Individual Tax Return Study," Proceedings of the Section on Survey Research Methods, American Statistical Association.

· Tobin, J. (1958). "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, pp. 24-36.

· United Nation. (2014). Economic Tendency Surveys – Handbook. http://unstats.un.org/unsd/nationalaccount/workshops/2014/Hangzhou/R-BD1.PDF

· Valliant, R. (2004). "The Effect of Multiple Weighting Steps on Variance Estimation", *Journal of Official Statistics*, 20, pp. 1—18.

· Valliant, R., Dever, J. A., Kreuter, F. (2013). *Practical Toos for Designing and weighting survey samples*. New York: Springer.

· Whitely, S. (2014). "Total Survey Error and Institutional Research: A Case Study of the University Experience Survey," presented at the Australasian Association for Instituional Reserach Forum, Melbourne, Australia.

· Wright, T. and Tsao, H. J. (1983). A frame on frames: an annotated bibliography. In: T. Wright (ed.), *Statistical Methods and the Improvement of Data Quality*. New York: Academic Press, pp. 25-72.