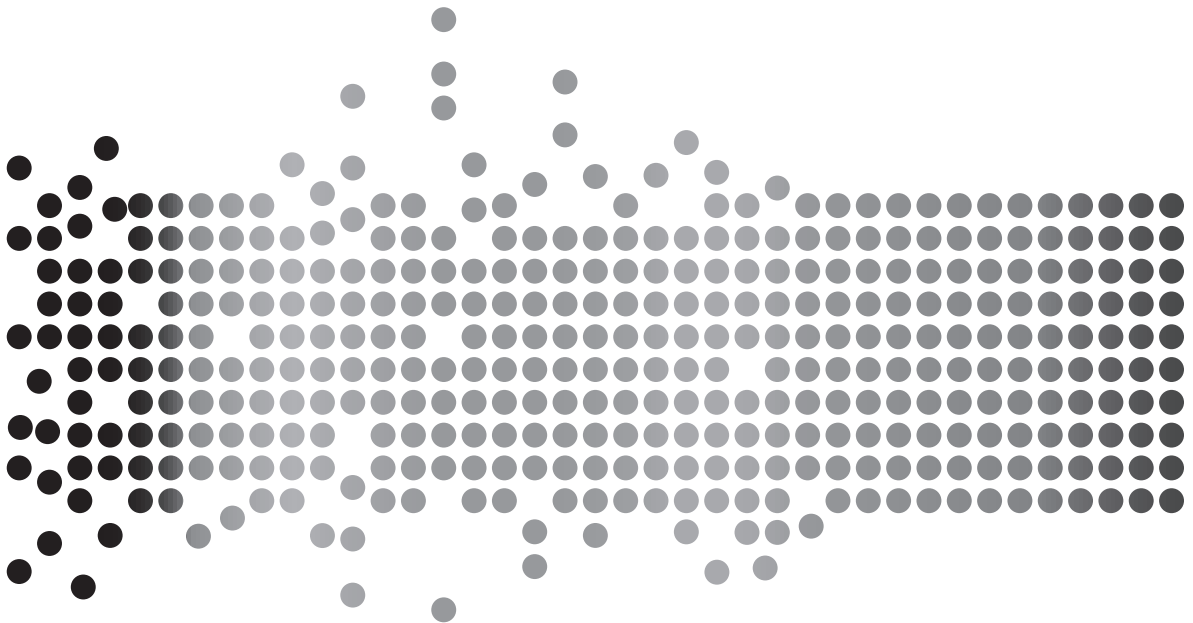


보건복지분야 조사통계 선진화 방안 연구

Study of Developing Strategies of Survey Statistics for
Health & Welfare In Korea

손창균 · 김은주 · 류제복 · 이기성



연구보고서 2012-46

보건복지분야 조사통계 선진화 방안 연구

발 행 일 2012년
저 자 손 창 균 외
발 행 인 최 병 호
발 행 처 한국보건사회연구원
주 소 서울특별시 은평구 진흥로 235(우: 122-705)
전 화 대표전화: 02) 380-8000
홈페이지 <http://www.kihasa.re.kr>
등 록 1994년 7월 1일 (제8-142호)
인 쇄 처 범신사
가 격 6,000원

© 한국보건사회연구원 2012

ISBN 978-89-8187-974-7 93510



머리말

보건복지분야는 과거에 비해 매우 광범위하고 근원적인 통계자료를 해마다 생산하고, 이를 기반으로 복지 정책 및 보건의료정책을 수립하며, 기존의 정책을 수정 보완하는 근거 자료로 폭넓게 활용되고 있다. 전국민의 건강보험 가입과 1인 이상 사업장 근로자의 국민연금 가입 등으로 관련 정책 자료가 지속적으로 축적되고 있다. 이러한 정책 영역적 특성과 더불어 국민을 대상으로 한 다양한 조사통계자료 또한 지속적으로 축적되어 타 분야에 비해 상대적으로 근거기반 정책수립의 토양을 건실하게 다져가고 있다. 그럼에도 불구하고, 통계작성 환경은 시간이 갈수록 점점 나빠지고 있는 것이 현실이다. 가구 또는 개인, 의료기관 등에 대한 조사를 통한 보건 및 복지분야의 실태를 파악하기 위해서는 매우 비싼 비용이 소요되며, 이와 더불어 조사대상들의 참여 거부 및 불성실한 응답 등으로 조사자료의 신뢰성에 많은 도전이 되고 있다.

최근에 조사원에 의한 면대면 조사 방법이 IT 기술의 발전과 전화 조사 기법의 발전 등으로 다양한 조사 기법으로 대체되고 있으며, 국내외적으로 통계적 추정기법 등의 개발을 통해 통계학적 측면에서 확률 표본 추출 방법에 대한 여러 가지 주장들이 제기되고 있다.

또한 보건복지분야의 특성에 따라 대량의 DB가 축적되어 관련 분야에 대한 다양한 통계자료가 생산되고 있고, 이러한 DB의 활용성에 많은 학자 및 연구자들의 관심이 점점 높아지고 있다.

이러한 관점에서 본 보고서는 국민의 삶과 직접적으로 관련이 있는 보건 및 복지분야 조사 통계에 대한 다양한 관점에서 현황을 파악하고, 향후 이 분야 통계 생산의 효율적인 방법을 모색하고, 보다 신뢰성 있는 국가 통계생산을 위한 방안을 제안하고, 정책적 시사점을 제안하고자 한다.

본 보고서는 현재 동국대학교 정보통계학과에 재직중인 손창균 교수의 연구책임하에 청주대학교 통계학과 류제복 교수와 우석대 아동복지학과 이기성 교수, 보건사회연구원의 김은주 연구원이 연구진으로 참여하였다. 또한 보건사회연구원의 도세록 연구위원과 정영철 연구위원께서 본 보고서의 검독을 위해 노고를 아끼지 않았으며, 외부검독위원으로 한신대 한근식 교수, 서울시립대 김규성 교수, 그리고 동신대 홍기학 교수께서 좋은 조언을 제안하였다.

2012년 12월

한국보건사회연구원장

최 병 호

Abstract	1
요 약	3
제1장 서론	9
제1절 연구의 배경 및 목적	9
제2절 연구방법	10
제3절 연구내용	11
제2장 국내 보건복지분야 통계생산 현황	15
제1절 보건복지분야 통계 현황	15
제2절 보건복지분야 통계별 주요사항	17
제3절 보건 및 복지분야 통계 생산 관리 현황	22
제3장 자료수집 방법의 현황과 미래 그리고 시사점	27
제1절 데이터 수집 방법에 관한 고찰	27
제2절 자료수집 방법의 차원	31
제3절 자료수집 방법간의 차이 연구	34
제4절 자료수집 방법의 향후 전망	41
제5절 조사방식의 미래와 미래의 조사방식	44
제6절 정책적 시사점	51
제4장 표본조사의 현황과 미래 그리고 시사점	57
제1절 표본조사의 배경	57

제2절 통계적 추론 및 비용 그리고 표본추출	60
제3절 불확실성의 미래	64
제4절 확률 표본의 향후 전망	71
제5절 정책적 시사점	72
제5장 해외 주요 패널조사의 현황과 시사점	77
제1절 해외 패널조사의 현황	77
제2절 정책적 시사점	88
제6장 조사자료의 신뢰도 향상을 위한 방안과 시사점	93
제1절 조사 응답률 산정	93
제2절 조사자료 에디팅 방안	109
제3절 정책적 시사점	119
제7장 결론 및 정책적 시사점	123
참고문헌	127
부록	135

표 목차

〈표 2-1〉 보건분야 조사통계 현황 및 주요 통계	15
〈표 2-2〉 복지분야 조사통계 현황 및 주요 통계	16
〈표 2-3〉 보건의료분야 조사통계의 조사대상별 분류	18
〈표 2-4〉 보건의료분야 조사통계의 조사방법별 분류	19
〈표 2-5〉 보건의료분야 조사통계의 표본추출방법별 분류	20
〈표 2-6〉 복지분야 조사통계의 조사대상별 분류	21
〈표 5-1〉 PSID의 조사방법	84
〈표 5-2〉 주요 패널간 비교	87

그림 목차

[그림 1] 보건의료분야 통계의 생산 및 관리 체계	22
[그림 2] SLID의 패널 구성 형태	86

An abstract graphic consisting of a central grey cloud-like shape with several smaller, lighter grey shapes floating around it, resembling a splash or a cluster of particles.

Abstract

Study of Developing Strategies of Survey Statistics for Health & Welfare In Korea

- Status of health and welfare survey statistics in domestic
 - Survey statistics in the field of health
 - 9 kinds of health policy field, 3 kinds of prevention of diseases policy and 10 kinds of other field.
 - Survey statistics in the field of welfare
 - 4 kinds of children policy, 3 kinds of disability, 3 kinds of elderly care policy and 8 kinds of other field.
- Status of aboard panel survey and its policy implications
 - Many countries use multi-survey methodology such as telephone and mail survey as well as face to face survey.
 - Provide the alternate methods of reducing of response burden to the survey participation
 - rotation survey design
 - register-based data of respondents
- The paradigm shift of data collection mode
 - Changes of survey method and data collection modes

2 • 보건복지분야 조사통계 선진화 방안 연구

- Need to diversification of survey modes besides of the face-to-face survey or household survey.
 - Need to consider feasibility of the internet, telephone or mail survey
 - Need to consider various data collection mode based on IT
- Change from PAPI to CAPI or CAWI
 - Need to use the computer based survey mode.
- Development of data collection mode by social media as SNS.
- Improvement of adaptability for survey environment using the mixed survey mode
- The paradigm shift for probability sampling
 - It should consider that decreasing of dependency of the probability sampling due to the change of demographic structure and the derement of population coverage.
 - Provide the availability of model-based estimation and non-probability sampling
- Some suggestions to assurance of reliability of survey data
 - Need to provide the response rate of survey data
 - Need to build the editing system of survey process
- Consider the application of the public data base system in order to overcome the limitation of probability sampling.
- Construction of IT infra structure to use the real time data
- Need to study of big data in the field of health and welfare

요약

I. 연구의 배경 및 목적

1. 연구 배경

□ 국가통계의 분류

- 조사통계, 보고통계 및 가공통계로 구분
 - 통계자료의 수집 과정이 조사를 통해 이루어졌는지, 행정 절차를 통해 수집되었는지, 또는 기존의 통계자료를 2차 가공하여 수집되었는지에 따라 구분

□ 통계분류별 특징

- 조사통계 : 정책수립과정에서 요구되는 통계를 시의 적절하게 생산할 수 있기 때문에 근거기반 정책수립에서 매우 중요하게 활용
- 보고통계 및 가공통계 : 법률이나 규정에 따라 반드시 이루어져야 하는 일종의 행정절차를 통해 생산되기 때문에 비용측면에서는 효과적일 수 있으나, 시의성이나 관련성 측면에서 조사통계에 비해 유용성이 떨어지는 특성이 있으며, 통계작성 보다는 행정 절차에 따르기 때문에 조사통계에 비해 통계생산에 대한 목적성이 약함.

□ 조사환경의 변화

- 통계생산과정에서 노력과 비용은 시간이 지날수록 점점 커지고 있으며, 거대한 비용과 시간이 소요되고 있음.
- 응답자들의 응답 거부, 부재, 거짓응답 등으로 조사환경은 점점 나빠지고 있는 실정임.
- 시의적절하고 신뢰성 있는 보건 및 복지 분야의 통계자료 생산이 절실함.

2. 연구 목적

□ 보건 및 복지분야의 조사통계를 중심으로 통계작성 현황과 각 통계별 특징을 파악

□ 조사방법 및 자료수집방법의 개선을 통해 시의성과 관련성을 적절히 보완할 수 있는 방향 제시와 효과적인 국가통계생산을 위한 정책을 제안

II. 주요 연구결과

□ 국내 보건 및 복지분야 조사통계현황

- 보건 분야 조사통계
 - 건강정책분야 9종, 질병정책분야 3종과 기타 10종의 통계가 생산
- 복지분야 조사통계
 - 복지분야 조사통계는 아동분야 4종과, 장애인분야 3종, 노인분

야 3종, 기타분야 8종으로 총 17종의 조사통계가 생산

□ 해외 패널 조사의 현황과 그에 시사점 제시

- 방문 면접조사 이외에 전화조사 및 우편조사 등 다양한 조사 방법을 활용함.
- 패널조사 참여자의 응답 부담 경감을 위한 대안의 제시
 - 순환 표본설계를 통한 응답자 부담 경감 모색
 - 외부 등록자료의 적극적 활용

□ 자료수집 방법의 패러다임 전환

- 조사방법 또는 자료 수집 방법의 전환
 - 개별 면접 조사 또는 가구 방문을 통한 각 또는 개인 대상 조사방법의 다양화 필요
 - 인터넷, 전화, 우편 조사 등을 활용한 면접조사의 대체 또는 부가적 활용 가능
 - IT 기반 조사방법의 다양성 모색
- 종이 설문지(PAPI: paper and pencil) 에서 컴퓨터를 이용한 조사 방법(CAPI 또는 CAWI)으로의 전환
 - 조사원 면접 방식과 더불어 응답자 기입식 조사 방법을 컴퓨터 기반 조사 방법을 도입하여 다양한 자료 수집 가능성 모색
- 소셜 미디어의 활용을 통한 자료 수집 방법 개선 방안 고려
- 혼합조사 방식을 통한 조사환경에 대한 적응성 개선

□ 확률 표본 조사에 대한 패러다임 전환

- 인구 구조의 변화와 모집단의 포괄성 감소에 따른 확률 표본의 의존성 약화

- 모형기반 추정과 비확률 표본 추출의 가능성 제시
- 조사 자료의 신뢰성 확보를 위한 제언
 - 응답률 산정 방법의 제시와 이를 통한 자료의 신뢰성 확보 방안 제시
 - 자료에 대한 에디팅 시스템 구축을 통한 자료의 정확도 개선 방안 제시

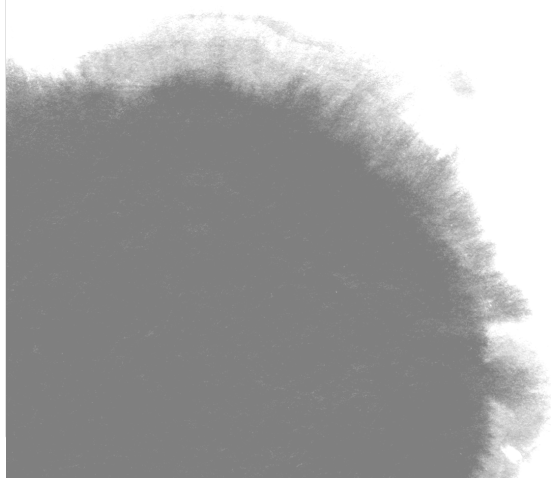
III. 결론 및 시사점

- 자료수집 방법의 다양화 모색
- 확률 표본의 한계점 극복을 위한 등록 DB활용에 대한 연구 필요
- 자료의 신뢰성 확보를 위한 응답률 산정 및 에디팅 시스템 구축을 위한 제언 제시
- 실시간 빅데이터 구축 및 활용을 위한 인프라 구축 및 연구환경 조성
- IT기반 환경에서의 통계자료 활용성 모색

*주요용어: 보건, 복지, 조사, 통계

1장

서론



제1 장 서론

제1 절 연구의 배경 및 목적

국가통계는 통상적으로 조사통계와 보고통계 및 가공통계로 구분하고 있다. 이러한 통계의 분류 기준은 통계자료의 수집 과정이 조사를 통해 이루어졌는지, 행정 절차를 통해 수집되었는지, 또는 기존의 통계자료를 2차 가공하여 수집되었는지에 따라 구분된다.

자료 수집과정에서 소요되는 다양한 자원의 규모를 고려할 때, 조사통계가 가장 큰 비중을 차지한다고 볼 수 있다. 또한 정책수립과정에서 요구되는 통계를 시의 적절하게 생산할 수 있기 때문에 근거기반 정책수립에서 매우 중요하게 활용된다. 이와는 다르게 보고통계 및 가공통계는 법률이나 규정에 따라 반드시 이루어져야 하는 일종의 행정절차를 통해 생산되기 때문에 비용측면에서는 효과적일 수 있으나, 시의성이나 관련성 측면에서 조사통계에 비해 유동성이 떨어지는 특성이 있으며, 통계작성 보다는 행정 절차에 따르기 때문에 조사통계에 비해 통계생산에 대한 목적성이 약하다.

최근에는 통계청을 중심으로 등록자료(register-based data)에 대한 활용도를 높이기 위한 다양한 통계작성 방법과 추정방법들을 연구하였고, 국내 여러 통계학자들에 의해 가능성을 타진하고 있다.

조사통계 생산과정에 소요되는 노력과 비용은 시간이 지날수록 점점

커지고 있으며, 거대한 비용과 시간이 소요되고 있음에도 조사환경은 점점 나빠지고 있는 실정이다. 그럼에도 불구하고 거대한 국가 예산이 투입되고 있는 대다수의 조사통계의 경우 과거에 수립된 예산을 현재 까지도 그대로 적용하고 있어, 조사환경의 변화를 적절히 반영하지 못하고 있다. 이러한 측면에서 국민의 건강한 삶에 대한 욕구 증대 및 사회 각 분야에서의 복지에 대한 필요성 증대에 따라 국가 재정의 효율적인 배분과 집행이 요구되고 있기 때문에 시의적절하고 신뢰성 있는 보건 및 복지분야의 통계자료 생산이 절실하다. 그럼에도 불구하고, 전술한 바와 같이 조사통계 생산 현장은 이러한 요구를 반영하기에는 많은 도전이 따르고 있다.

따라서 본 연구에서는 국민의 삶의 질과 밀접한 연관이 있는 보건 및 복지분야의 조사통계를 중심으로 통계작성 현황과 각 통계별 특징을 파악하고, 조사방법 및 자료수집방법의 개선을 통해 시의성과 관련성을 적절히 보완할 수 있는 방향을 제시하고, 효과적인 국가통계생산을 위한 정책을 제안하고자 한다.

제2절 연구방법

본 연구는 문헌 연구를 중심으로 관련 분야의 연구자 및 공무원, 학자 등에 대해 심층면접을 통해 향후 조사통계의 방향성을 모색하며, 이를 통해 합리적인 통계생산 정책수립에 관한 제언을 하고자 한다.

보다 구체적으로 보건복지분야 통계의 다양성을 고려하여 가구 및 의료기관, 시설 등에 대한 조사 통계를 중심으로 보고 및 행정통계와의 연계를 통한 향후 통계생산의 효율성 제고를 위한 정책 방향을 수립하고자 한다. 국내의 보건 및 복지분야 통계의 현황을 파악하고, 다음으로

해외 선진국의 보건 및 복지분야 통계생산 및 관리 현황과 통계정책을 중심으로 고찰하고, 특히 최근 급격히 증가하고 있는 패널조사의 조사 현황 및 자료 관리를 중심으로 다루고자 한다. 이와 더불어 해외의 자료 수집방법의 향후 발전 방향을 소개하고, 국내 조사환경 변화에 따른 조사 통계의 발전방향에 대해 논의하였다.

제3절 연구내용

첫째 보건 및 복지분야의 조사통계를 중심으로 국가통계로 승인된 각종 통계의 현황을 파악하며, 둘째 자료수집 방법에 관한 통계학적 변화를 파악하며, 셋째 최근 많은 연구기관에서 생산되고 있는 조사통계 중에서 패널조사에 관해 해외 주요 패널 조사에 관한 생산 절차를 고찰하고, 마지막으로 보건 및 복지분야 조사통계의 생산과정을 개선하고, 신뢰성 있는 통계생산을 위한 정책적 제언을 하고자 한다.

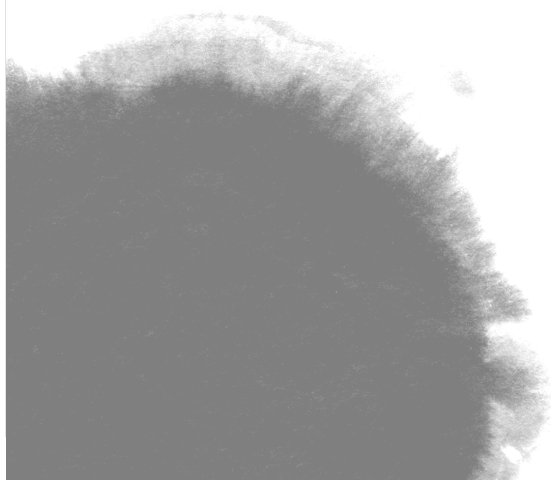
이를 위해 다음과 같은 주제를 중심으로 연구를 수행하고자 한다.

- 보건복지분야 조사통계의 현황
- 데이터 수집방법의 현황과 미래
- 표본조사 방법의 현황과 미래
- 패널조사의 합리적 생산 방안
- 조사자료의 신뢰성 증진 방안
- 조사자료의 에디팅 방법



2장

국내 보건복지분야 통계생산 현황



제2장 국내 보건복지분야 통계생산 현황

제1 절 보건복지분야 통계 현황

2011년 현재 보건 및 복지분야의 통계현황을 살펴보면 보건의료분야는 22종의 통계가 생산되고 있고, 이중 건강정책분야 9종, 질병정책분야 3종과 기타 10종의 통계가 생산되고 있다.

〈표 2-1〉 보건분야 조사통계 현황 및 주요 통계

통계유형	통계명칭	주요통계단위
조사 통계 (22종)	국민보건의료실태조사 국민구강건강실태조사 청소년건강행태온라인조사 정신질환실태조사 국민건강영양조사 환자조사 퇴원손상 심층조사 전국장내기생충실태조사 의료기기화장품제조유통실태조사 선천성이상아통계조사 영아모성사망조사 지역사회건강조사 의약품소비량 및 판매액통계조사 인구공통전염병위험군의감염실태조사 건강보험환자진료비실태조사 한국의료패널조사 전국출산력및가족보건복지실태조사 시민보건지표조사 병원경영실태조사 한국인인체지수조사 근로환경조사 한방의료이용 및 한약소비 실태조사	의료기관 인력,장비,시설 개인의 치아 우식률 (성인, 청소년)흡연율, 중고등학교 재학생 건강상태 유병률 및 위험요인 개인, 가구의 질병 질환별 환자수 퇴원환자의 질환특성 개인의 기생충감염률 업체수, 판매액, 판매량 선천성이상아 유병률 모성사망비 지역 및 개인의 건강상태 의약품 소비량 및 판매액 병원체보유율 퇴원 및 외래 환자진료비 개인 및 가구의 의료비 15~59세 기혼여성 출산력 의료이용수준 및 유병률 병원매출액, 종사자수 개인의 신체특성 작업환경, 산업재해 한약소비량, 소비액

보건의료분야 조사의 경우 가구 및 의료기관 및 기업체 등 조사 대상이 다양한 것이 특징이며, 의료기관대상조사의 경우 의료기관 현황과 환자 기록부를 조사단위로 하고 있는 것이 타 조사와는 구별되는 특징이다. 한국 의료패널조사의 경우 가구대상 조사로서 동일 가구를 지속적으로 추적조사하여 가구의 의료비를 조사하며, 조사결과를 건강보험DB와 연계하여 분석하는 것이 특징이다.

한편 복지분야 조사통계는 아동분야 4종과, 장애인분야 3종, 노인분야 3종, 기타분야 8종으로 총 17종의 조사통계가 생산되고 있다.

〈표 2-2〉 복지분야 조사통계 현황 및 주요 통계

통계유형	통계명칭	주요통계단위
조사 통계 (18)	아동종합실태조사	소득계층별 아동현황(수)
	장애인생활체육실태조사	시설 및 이용자 현황
	노인실태조사	65세 이상 노인수
	유망사회서비스수요실태조사	사업체수 및 종사자, 매출액
	보육실태조사	보육아동수 및 시설
	장애인실태조사	장애유형별 만족도
	생명보험성향조사	개인의 보험가입률
	한국복지패널조사	가구및 가구원의 소득 지출, 빈곤
	국민노후보장패널조사	중고령자의 노후보장실태
	서울시복지패널조사	서울시거주 가구및 가구원의 복지실태
	가정폭력실태조사	가구내 폭력발생률
	기업및공공기관의가족친화수준조사	가족친화지수
	장애인편의시설설치현황조사	대상시설별, 종류별 현황, 시설수
	최저생계비계측조사	가구규모별 최저생계비
	고령화연구패널조사	고령인구의 일자리수
	성폭력실태조사	성폭력발생건수
	아동패널조사	아동의 복지수요율

복지분야 조사통계 중에서 국민노후보장패널조사와 서울시 복지패널조사, 한국복지패널조사 및 아동패널조사는 동일가구를 지속적으로 추적조사하는 종단면 조사이며, 노인실태조사는 횡단면과 종단면 모두를 조사 대상에 포함시켜 조사가 되고 있다.

보건의료분야와 복지분야는 보건의료 정책별로 정책수요에 대응하기 위해 통계를 생산하며, 복지분야의 경우 복지 정책 수요 대상자를 중심으로 조사가 수행된다. 두 분야 모두 통계생산 기관이 매우 다양하며, 보건복지부의 정책 담당 부서를 중심으로 관련 통계를 작성하고 있다.

통계청의 국가통계품질진단 자료로부터 대다수의 통계가 통계청으로부터 통계승인을 받은 국가통계로서 관련 모집단에 대한 대표성과 확률 표본추출방법을 적용하고 있으며, 일정수준의 품질을 유지하고 있다. 정책수요에 따라 통계가 산출되고 있기 때문에 일부 통계를 제외하고 이용자 관련성(relevance)이 낮은 것으로 파악되었고, 반면에 정확성(accuracy)은 높게 나타났다. 실제로 보건복지의 각 부서별로 산재된 통계자료를 보건복지통계포털(<http://stat.mw.go.kr/>)에서 검색이 가능하도록 함으로서 관련성을 높이기 위해 노력하고 있다.

한편 통계품질 향상을 위해 통계청의 품질관리 사이트와 연계하여 보건복지부 관련 통계에 대한 품질정보와 품질진단과정을 상세히 검색할 수 있도록 함으로서 부서의 통계담당자들의 품질제고에 기여하고 있다.

타 분야에 비해 조사 통계의 비중이 가장 높은 분야가 보건 및 복지분야이며, 국민의 생활과 직결된 정책 수립을 위한 기초자료가 되는 만큼 정확성, 관련성 등 품질지표의 향상에 많은 노력을 경주해야 할 것이다.

제2절 보건복지분야 통계별 주요사항

1. 보건의료분야 조사통계 현황

보건의료분야의 조사통계는 의료기관 및 사업체 대상 조사가 11종, 가구 및 개인조사 대상이 11종으로 파악되었다. 의료기관 대상 조사의

경우 주로 환자의 진료 레코드를 대상으로 조사되는 경우와 의료기관의 현황을 파악하기 위한 조사로 분류된다. 가구 및 개인단위 조사의 경우 개인의 의료기관 이용 이력, 개인의 현재 건강상태 등을 파악하기 위한 조사가 다수를 이루고 있다.

〈표 2-3〉 보건의료분야 조사통계의 조사대상별 분류

조사대상	통계명
의료기관 또는 의료장비 사업체 (11종)	국민보건의료실태조사 환자조사 의약품소비량 및 판매액통계 의료기기, 화장품제조유통 실태 조사 퇴원손상심층조사 영아모성사망조사 건강보험환자 진료비 실태조사 병원경영실태조사 근로환경조사 한방의료이용 및 한약소비 실태조사 선천성이상아 통계
가구 또는 개인 (11종)	국민구강건강실태조사 정신질환자실태조사 청소년건강행태온라인조사 국민건강영양조사 지역사회건강조사 전국장내기생충실태조사 인수공통전염병위험군의 감염실태조사 한국의료패널조사 한국인 인체치수조사 전국출산력 및 가족보건복지실태조사 시민보건지표조사

한편 각 통계별로 조사 방법에 따라 분류한 경우 다음과 같이 분류가 가능하다. 22종의 조사통계를 조사 방법에 따라 분류하면, 방문 면접조사 방법과 자기기입식 방법으로 구분할 수 있으며, 주로 가구 및 개인대상 조사의 경우 방문 면접조사가 많고, 기관 대상 조사의 경우 자기기입

식 조사가 많은 것으로 파악되었다.

〈표 2-4〉 보건의료분야 조사통계의 조사방법별 분류

조사방법	통계명
웹조사, 전화조사, 자기기입식조사	국민보건의료실태조사 환자조사 의약품소비량 및 판매액통계 의료기기, 화장품제조유통 실태 조사 퇴원손상심층조사 영아모성사망조사 건강보험환자 진료비 실태조사 병원경영실태조사 한방의료이용 및 한약소비 실태조사 선천성이상아 통계 청소년건강행태온라인조사
방문 면접, 검사 또는 검진조사	국민구강건강실태조사 정신질환자실태조사 국민건강영양조사 지역사회건강조사 전국장내기생충실태조사 인수공통전염병위험군의 감염실태조사 한국의료패널조사 한국인 인체치수조사 전국출산력 및 가족보건복지실태조사 시민보건지표조사 근로환경조사

통계별 표본추출방법에 따라 분류해보면 <표 2-5>와 같이 표본조사 통계의 경우 대다수가 층화 추출 또는 층화 집락 추출방법을 적용하고 있으며, 특히 가구 및 개인 대상 조사의 경우 층화집락 추출을 사용하는 것으로 파악되었다. 전수조사로 통계를 작성하는 경우는 총 4종으로서 의료기관 대상 조사로서 “국민보건의료실태조사”, “영아모성사망조사”, “병원경영실태조사”, “선천성이상아 통계”등으로 파악되었다.

〈표 2-5〉 보건의료분야 조사통계의 표본추출방법별 분류

통계명	표본추출 방법
국민보건의료실태조사	전수조사
환자조사	표본조사-층화 표본추출
의약품소비량 및 판매액통계	표본조사-층화 표본추출
의료기기, 화장품제조·유통 실태 조사	표본조사
퇴원손상심층조사	표본조사
영아모성사망조사	전수조사
건강보험환자 진료비 실태조사	표본조사
병원경영실태조사	전수조사-1,600개 의료기관
한방의료이용 및 한약소비 실태조사	표본조사
선천성이상아 통계	전수조사
청소년건강행태온라인조사	표본조사
국민구강건강실태조사	표본조사-층화집락추출
정신질환자실태조사	표본조사
국민건강영양조사	표본조사-층화집락추출
지역사회건강조사	표본조사-층화집락추출
전국장내기생충실태조사	표본조사-층화집락추출
인수공통전염병위험군의 감염실태조사	표본조사
한국의료패널조사	표본조사-층화집락추출
한국인 인체치수조사	표본조사
전국출산력 및 가족보건복지실태조사	표본조사-층화집락추출
시민보건지표조사	표본조사-층화집락추출
근로환경조사	표본조사-층화추출

2. 복지분야 조사통계 현황

복지분야 조사통계의 경우 조사 대상에 따라 분류하면, <표 2-6>과 같이 사업체 또는 기관 대상 조사통계가 총 4종이며, 그 외는 모두 가구 또는 개인 대상의 조사인 것으로 나타났다.

〈표 2-6〉 복지분야 조사통계의 조사대상별 분류

조사대상	통계명
기관 또는 사업체 (4종)	유망사회서비스수요실태조사 보육실태조사(시설조사) 기업 및 공공기관의 가족친화지수 측정 장애인편의시설설치현황조사
가구 또는 개인 (15종)	아동종합실태조사 장애인생활체육실태조사 노인실태조사 보육실태조사(가구조사) 장애인실태조사 생명보험성향조사 한국복지패널조사 국민노후보장패널조사 서울시복지패널조사 가정폭력실태조사 최저생계비계측조사 고령화연구패널조사 성폭력실태조사 아동 패널조사 차상위 실태조사

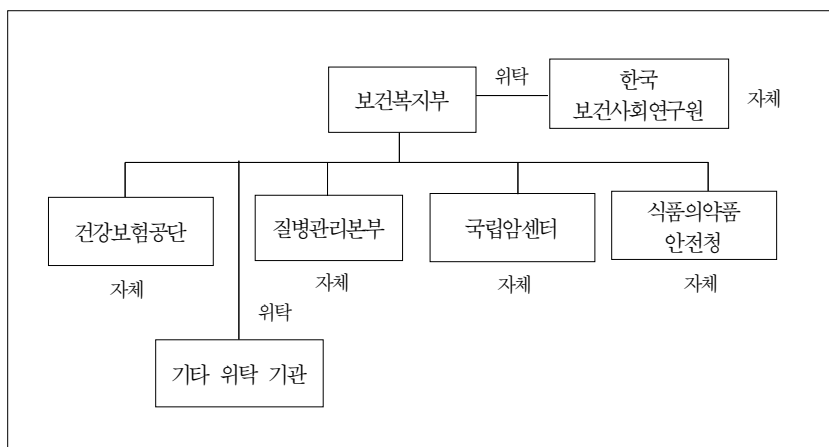
보건의료분야와는 다르게, 복지분야의 조사통계는 모두 표본조사 방법으로 통계가 작성되었으며, 총화 추출을 기반으로 총화 집락 추출을 활용하여 표본을 추출한 것으로 파악되었다.

외국의 경우 보건의료 및 복지분야 통계의 주된 조사방법은 전화조사이며, 일부 복지분야의 경우 개인 면접과 전화조사 및 우편조사를 병행하고 있는 것으로 파악되어, 국내의 자료 수집방법과는 차이가 나타났다.

제3절 보건 및 복지분야 통계 생산 관리 현황

보건 및 복지분야 조사통계는 통계생산 주체가 보건복지부이며, 그 외 산하 단체 및 공공연구기관의 주관 하에 통계가 생산되고 있다. 즉, 보건의료분야 통계는 건강 보험공단, 질병관리본부, 식품의약품안전청, 국립암센터 및 한국보건사회연구원, 한국보건산업진흥원 등 다양한 기관에서 보건 및 의료분야 통계가 생산되고 있다. 이들 기관 중 대다수는 조사기관을 통해 조사가 실시되며, 조사관리 및 자료 산출을 조사기관에 의뢰하여 수행하는 것으로 나타났다. 의료기관 대상 조사의 경우 앞에서 다룬 바와 같이 전화 또는 웹을 이용한 조사를 실시하고 있으며, 면접조사는 주로 기관 보다는 가구 및 개인 대상의 조사가 주를 이루고 있다.

[그림 1] 보건의료분야 통계의 생산 및 관리 체계



조사통계는 표본설계 및 표본 추출, 조사, 조사원관리, 자료입력, 추정, 결과공표 등 전 분야에 걸쳐 매우 복잡하고, 전문적인 단계를 거쳐 통계가 산출되기 때문에 조사와 자료 공표가 이원화 되어 있는 것이 현실이

다. 보건 및 복지분야 조사통계에 대한 몇가지 이슈를 살펴보면 다음과 같다.

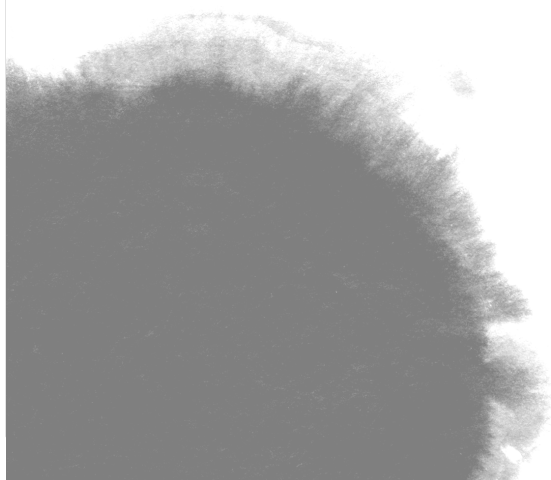
- 확률 표본추출 방법으로 표본추출과 추정이 수행되었음.
- 전통적인 가구와 개인 대상 조사통계가 대다수임.
- 위탁 수행 및 관리가 이원화됨.
- 보건통계는 주로 질환통계가 대다수이며, 복지통계는 각 부서별로 나뉘어 생산되고 있음.
- 정책통계담당관실과 각 부서별로 이원화되어 관리되고 있음.
- 생산 통계자료 관리가 주관 부서에서 이루어지지 않고 있음.

따라서 보건 및 복지분야 통계는 조사자료가 주를 이루고 있으며, 이에 대한 관리가 체계적으로 이루어 져야 할 것이다. 최근 국가통계포털과 나라통계 등 국가통계를 일괄적으로 관리하고 유지하는 노력이 통계청을 중심으로 진행되고 있다. 즉, 통계의 생산은 각 부서에서 주관하여 수행하고, 관리 및 유지는 통계청을 중심으로 수행되고 있다. 이러한 생산과 관리의 이원화는 국내 국가통계 현실을 반영한 결과로 볼 수 있으나, 주로 유럽의 사례에서는 각 부서에서 생산된 통계는 해당부서에서 관리하고 있는 것이 특징이다.

현재 통계청과 고용노동부를 제외하고 대다수 정부 부처 및 관련 연구기관의 통계관련 부서를 살펴보면 매우 제한적으로 운영되고 있다. 즉, 관련 부서에 통계담당 직원의 수가 상대적으로 적고, 전문성이 매우 낮은 상태이기 때문에 통계자료를 전문적으로 관리하고, 갱신할 수 없기 때문이다. 이러한 측면에서 보건복지분야의 통계생산과 관리의 전문화를 위해 보다 많은 투자와 인재 발굴이 요구된다.

3장

자료수집 방법의 현황과 미래 그리고 시사점



제3장 자료수집 방법의 현황과 미래¹⁾ 그리고 시사점

제1 절 데이터 수집 방법에 관한 고찰

자료수집 방법에 관한 개념은 비록 “조사 방법”과 관련된 용어가 1970년대 후반까지 애매한 형태로 나타나긴 했지만 조사 연구에서 오랜 역사를 가진다. 예를 들어 오차의 원천과 관련하여 Deming(1940)은 우편, 전화, 전보 그리고 직접면접 등을 언급하면서 “서로 다른 유형의 캔버스(different kinds of canvass)”라고 하였다. 이와 유사하게 Hochstim(1967)은 자료수집의 “전략”과 “방법”을 전통적인 방법과 비교하였다. “자료수집방법(mode)”에 관한 용어의 근원은 단순한 원천으로부터 발견하기는 어렵고 정의되지 않은 것처럼 보이지만, Groves and Kahn(1979)는 면대면 조사와 전화조사를 비교하는 과정에서 이 용어를 사용하였고, Groves(1979)와 O'Neil(1979)는 이들의 논문에서 이와 같은 용어를 사용하였다. 대체로 이때 처음으로 논문에서 “자료수집방법(mode)”이라는 용어를 사용한 것으로 추측된다.

그 이후 정확한 용어는 아닐지라도 조사 연구에서 빈번히 연구되는 분야가 되었다. 예를 들어 비록 기술이 변화함에 따른 차이가 있기는 하지만, 종이조사(paper and pencil: PAPI)와 컴퓨터를 이용한 조사

1) 본 장은 M. Couper(2011)의 “The Future of Modes of Data Collection” 논문을 편역하여 수록한 내용임.

(computer-assisted interviewing: CAI)에 관한 실험은 조사 유형의 비교이지 자료수집방법의 비교는 아니다. 그렇지만, 이와 같이 광범위한 용어의 사용은 조사자료 수집 방법을 다양하게 변화시켜 왔다.

최근의 경향은 “자료수집 방법”에 관한 논의가 외형적 방법으로 변화한 것이다. 적어도 조사에 관한 단순한 서술어로 용어를 사용하고 있다. 과거에 어떤 조사는 전화를 통해 많은 정보를 수집하였다. 그러나 “전국 면대면 조사” 또는 “RDD 전화조사”와 같은 단순한 기술로서는 충분하지 않다. 왜냐하면 여기에는 다음과 같은 세 가지 추세와 연관되어 있기 때문이다.

- 1) 자료수집방법의 확산
- 2) 자료수집방법의 복잡성 증가
- 3) 혼합적인 방법의 증가

1. 자료수집방법의 확산

최근 10년간 다양한 방법들이 활용되고 있으며 실제로 이들 방법들 간의 많은 변화가 있어 왔다. 우편조사와 면대면 조사가 1940년대부터 1970년 까지 주(主)를 이루며 발전되어 왔다(Lyberg and Kasprzyk, 1991). 1970년대 미국에서 전화조사가 급격히 확산되었고, 그 이후 유럽과 그 밖의 국가들에서 같은 현상이 일어났다. 전화조사의 성장은 그 당시 전화 보급률의 급격한 상승, 면대면 조사에 비해 상대적인 전화조사비용, 전화조사의 수행 속도, 면대면 조사와 비교할 때 조사 자료의 유사한 품질 등을 포함하여 다양한 요인들에 의해 고무되었다. 1990년대 들어서 비용과 속도의 측면에서 인터넷 조사가 전화조사의 영역을 위협하기 시작 하였다(Couper, 2000).

전화조사의 무응답 증가, RDD 전화조사의 포괄성에 영향을 주는 휴

대폰 이용의 증가, 주소기반 표본추출의 개발과 함께 우편조사에 관한 관심이 최근 조사방법에 관한 연구가 부활된 원인이다.

컴퓨터를 이용한 면접방법의 개발로 인해 CATI(Computer-Assisted Telephone Interviewing)와 CAPI(Computer-Assisted Personal Interviewing)같은 새로운 방법이 도입되었다. 이와 더불어 많은 연구에서 민감한 문항에 대한 자기기입식 설문 효과의 입증이 되어 면접자 기입 방식의 조사에서 새로운 자기기입식 방법의 발전이 이루어 졌다. 따라서 응답자 기입방식의 설문(self-administrative questionnaire :SAQ)은 컴퓨터를 이용한 자기 기입식 면접(CASI)을 포괄하며, 다양한 응용(음성-CASI)과 전화면접(interactive voice response: IVR 또는 T-CASI)이 개발되었다.

2. 복잡성 증가

다양하게 이용 가능한 자료수집 방법의 증가와 함께, 조사방법이 점점 복잡해지고 있다. 예를 들어 컴퓨터를 이용한 자기기입식 면접은 문자, 음성 또는 비디오를 통해 질문을 제시하거나 반응장치를 이용할 수 있다. 인터넷 조사 또한 비디오-CASI 등과 같은 유사한 유형의 질문을 포함할 수 있다. 인터넷을 통해 비디오-CASI로 측정 오차의 관점에서 면대면조사의 한 부분으로서 비디오-CASI와 같은 방법으로 다양한 속성을 공유할 수 있는 반면, 두 가지의 서로 다른 유형의 자료수집 형태(면대면과 인터넷)는 표본추출오차, 포괄성오차, 무응답오차와 같은 비관측 오차를 크게 할 수 있다.

인터넷을 이용한 웹 조사를 예로 들면 웹 조사는 하나의 단순한 자료수집 방법으로 볼 수 있지만, 조사를 위한 표본추출은 다양한 방법이 있을 수 있다. 여기에는 비확률 추출과 확률 추출에 기반한 방법을 포함할

수 있으며, 이러한 방법은 온라인 접근 패널, 중도 조사, 리스트에 기반한 표본과 같은 것을 포함한다. 다른 표본설계의 유형들(영역확률 표본, 임의 전화걸기표본)은 면대면 조사 또는 전화걸기 등과 같은 특정한 유형을 내포한다. 이러한 모든 접근 방법은 조사자료를 이용한 추론에 대해 매우 다른 방법을 사용하게 된다. 더욱이 웹조사를 다양하게 수행할 수 있는데, 즉, 측정의 관점에서 면접자 기입방식조사와 유사하며, 우편 조사를 반복적으로 수행하는 것과 같을 수 있다. 다시 말해서 인터넷은 매우 유용한 방법이지만, 인터넷 조사를 수행함으로써 조사의 품질(quality)이나 산출된 추정치의 품질을 평가하는데 불충분한 정보를 제공하게 된다.

이러한 사실로 부터 일련의 설문을 응답자에게 질문하는 면접자를 포함하여 단순한 면대면 조사의 개념은 갈수록 진부한 방법으로 간주되었다. 그러한 조사는 자기기입식 요소를 포함하고, 면접자에 의한 관찰과 신체 측정, 보기카드 사용 등을 포함하여 수집방법의 복잡성이 모든 면에서 증가하고 있다.

3. 혼합조사 방법

마지막으로 자료수집방법의 복잡성 설명과 연관된 경향은 혼합된 방법을 사용하여 자료를 수집하는 경향이 증가하고 있다는 것이다. 자료수집방법의 혼합된 형태에 관한 개념은 긴 역사를 갖는다. 예를 들어 1964년에 Stanley Payne는 “기본적인 조사 방법을 결합하여 사용하는 유용한 방법에 눈을 뜰 필요가 있다. 우리는 단순한 조사의 보충적인 부분으로 그러한 방법들을 관찰하기 위한 배타적인 대체성으로 바라봄으로서 그러한 방법에 관해 무지하다”고 언급한 바 있다.

이와 유사하게 Lyberg and Kasprzyk(1991)은 “실제로 일부 연구는

오로지 하나의 자료수집 방법에 의존하고 있다. 비용문제, 불완전한 포괄성, 무응답 및 측정오차들은 자료를 수집하는데 있어 2가지 이상의 방법에 항상 기인한다” 고 언급하였다. 그러나 과거 몇 년간은 부분적으로 앞에서 언급한 바와 같은 경향 때문에 즉, 주소기반 표본추출의 활용성과 우편조사에 관한 관심의 부활, 인터넷조사에 의해 요구되는 측정의 기회 등 수 십년 간 주요한 여론조사 방법의 하나로 여겨지던 전화조사에 대한 위협으로 혼합된 자료수집방법에 대한 관심이 상당히 높았다. 이와 더불어 이러한 세 가지 경향은 조사자료 수집 과정의 기법으로서 “자료수집방법”의 효용성의 한계를 극복하고 있다. 따라서 “자료수집방법”은 더 이상 기술적인 영역으로서 고려해서는 안되며, 다차원적인 구성으로 고려되어야 한다.

제2절 자료수집 방법의 차원

자료수집 방법은 조사에 대한 단순한 기술로서의 조사방법에 관한 고려보다는 조사 오차와 비용에 자료수집방법이 영향을 줄 수 있다는 측면에서 매우 수평적으로 여러 가지가 연관될 수 있지만, 여기서는 주요 차원에 대해서만 언급하기로 한다.

다양한 수집방법들과 수평적인 차원들 중 하나는 조사에 참여하는 면접자의 수준이다. 어떤 방법들(예를 들어 전화조사와 면대면 조사)은 면접자가 응답자에게 설문을 읽어주고 종이나 컴퓨터에 응답을 기록하는 면접자가 완전하게 관리한다. 어떤 방법은 면접자가 있는 상태에서 응답자가 응답을 기록하는 방법이다. 이러한 방법들은 컴퓨터를 이용한 자기 기입식 면접이나 종이 설문지에 직접 기록하는 방법이 이에 해당된다. 집단조사(학교와 같은 현장에서)는 이러한 유사한 특성을 공유한다. 더

육이 만일 응답자가 설문을 작성하는데 도움이 필요하면 지속적으로 설문지를 내려 받거나 면접원이 면접도구를 나누어 주거나 자동 시스템을 제공하는 것을 포함하여 외부 IVR/T-ACASI를 사용할 수 있지만, 설문지를 작성하는 동안 이러한 것들은 제시되지 않는다. 한편 비연속적인 방법들은 완전한 자기기입식 방법으로서 우편조사, 웹조사 그리고 내부 IVR등이다. 이 경우 설문작성에 참여한 면접자의 수준은 비용뿐만 아니라 비관측 (표본추출 및 무응답)오차와 측정오차(민감한 질문의 효과, 동기, 탐색 및 지원 능력)를 발생시킨다.

자료수집방법에 관한 두 번째 차원은 응답자와의 접촉정도가 포함될 수 있다. 다시 말해서 면대면 조사는 연속적으로 서로 다른 지점에서 면접자가 응답자를 직접 접촉하게 되는데, 이 경우 사용하는 언어적 및 비언어적 힌트가 허용된다. 또한 이러한 직접적인 접촉은 협조를 얻기 위해 신분증 및 편지와 같은 문구적인 사물을 사용할 수 있게 하며, 다양한 육체적, 생물학적인 측정을 하고 면접자가 관찰값을 수집하는 기회와 함께 응답을 가능하게 하기위해 카드 등을 보여주기도 한다.

전화조사는 조사자와 응답자간에 직접적인 접촉이 덜 빈번하게 되며 우편 및 웹조사는 보다 더 간접적이다. 설문지를 읽어주는 그림이나 비디오를 포함한 방법 또는 심지어 면접원을 대신하는 아바타 또는 애니메이션을 이용한 형식은 이러한 연속체의 중간에 위치한다.

서로 다른 자료 수집방법을 규정할 수 있는 세 번째 차원은 의사소통의 방식에 있다. 여기서는 조사원이 응답자에게로 또는 응답자에서 조사원으로부터와 같은 양방향을 고려할 필요가 있다. 전형적으로 전화조사 (그리고 IVR)는 단방향의 청각적 의사소통이 이루어진다. 반면에 우편 조사는 유사하게 단일한 시각적 의사소통으로 제한된다. 면대면 조사는 이러한 측면에서 보다 유연성을 갖는다. 즉, 조사 설문과 응답에 대한 기술적인 전달이 가능하며, 시각적인 도구를 보충적으로 활용할 수 있다.

또한 면대면 조사의 경우 조사 설문의 전달과 응답 선택은 동일한 의사소통 방법을 사용하지 않아도 된다. 예를 들어 응답카드를 사용할 경우 동일한 형태의 의사소통은 아니다. 즉, 면대면 조사에서 종이 조사표를 활용한 응답자 기입식 조사는 시각적인 자료를 활용할 수 있기 때문이다. 오디오-CASI는 청각과 시각적인 의사소통이 동시에 모두 가능하며, 따라서 웹조사의 경우 의사소통의 관점에서 다른 조사 방법에 비해 유연성을 갖는다. 대부분의 웹조사에서는 구문(문장)으로 설문을 제시하거나 오디오나 비디오로 설문을 제공하기도 하기 때문이다.

자료 수집 방법들은 또한 응답자를 통제하는 수단이 되기도 한다(Couper, 2008). 우편조사에서는 응답자가 설문에 응답하는 방법을 선택함에 있어 설문에 응답할지, 또는 언제 응답할지, 무엇에 응답할지를 완전히 통제한다. 면접자 기입 방식의 조사에서 조사에 대한 통제는 면접원에 의해 응답자의 보조와 흐름을 결정한다. 한편 컴퓨터를 이용한 조사에서 이러한 통제는 응답이 허용되는 응답이 무엇인지, 다음 순서에 묻는 질문이 무엇인지를 지시하는 조사 소프트웨어(프로그램)로 가능하다. 웹조사는 종이조사와 유사하게 보다 융통성 있게 또는 CAI 조사에서 유사한 방법으로 응답자에 대해 더 나은 통제력을 발휘할 수 있게 설계가 가능하다.

자료 수집방법의 다양성 측면에서 또 다른 차원은 응답자의 비밀보호 정도이다. 면접이 종료되거나 중단되었을 때 응답자가 다른 사람들과 같이 있을 때 면접하는 경우에 낮은 수준에서 비밀이 보호된다. 면접자 기입방식의 조사는 면접내용에 대해 다른 가구원이 들을 수 있는 가능성 때문에 중간 수준의 비밀이 보호된다. 이동전화를 사용하는 전화조사는 면접 상황에 의존하는 유선전화를 사용하는 전화조사 보다 다른 수준의 개인정보를 포함할 수 있다. 높은 수준의 비밀을 보호하는 자료 수집 방법은 CASI, 우편조사 그리고 웹조사인데 설문지 설계와 조사를 종료하

기위해 응답자가 어떻게 선택하느냐에 따라 정도가 다양하다.

마지막으로 자료수집 방법은 이용된 컴퓨터 기술의 정도에 따라 다양하다. 이러한 영역은 응답자 기입방식인지 또는 면접자 기입방식인지, 종이기반 조사로부터 단지 면접자가 사용한 기술(CATI 또는 CAPI)에 따라 다양하다. 더욱이 응답자가 조사기관이 제공하는 기술(CASI)을 사용하는 환경, 응답자가 그들의 기술을 이용하여 응답하는 경우(웹조사) 등 다양하다. 이 차원은 포함률뿐만 아니라 응답오차와 측정오차에 영향을 받을 수 있다.

제3절 자료수집 방법간의 차이 연구

조사방식간의 차이에 관한 연구들은 이미 많이 소개되어 있으며, 계속해서 그러한 연구들이 진행되고 있다. 예를 들어 AAPOR(the American Association for Public Opinion Research) 학회에서 발표된 조사방법론 관련 논문에서의 다수는 조사방식간 비교와 최근에는 혼합된 조사방식간 비교가 주를 이룬다. 여기에서 이러한 모든 연구 자료를 이해하는데 한계가 있기 때문에 몇가지 연구 결과만을 제시하고자 한다. 다양한 방법으로 조사방법의 효과를 연구하였는데, 어떤 연구들은 결과와 자료수집 전체 시스템을 연구하는데 초점을 맞추고 있다. 예를 들어 RDD 전화조사를 대면조사와 비교하였는데, 포괄성, 표본추출, 무응답, 측정 또는 처리오차 등에 기인한 차이점의 근원을 분리하지 않고 결과적인 추정치를 비교하였다. 이러한 접근 방식은 동일한 조건을 고려하지 않고 각 방법에 관한 최상의 설계가 되도록 한다. 다른 연구들은 인터넷조사와 전화조사간의 포함률 차이 같은 특정한 오류의 근원이나, 응답순서 효과 혹은 사회적으로 바람직함의 편향과 같은 차이를 만들어

내는 특수한 메커니즘에 초점을 맞추고 있다(Chang and Krosnick, 2009). 이러한 접근은 효과들을 독립시키기 위한 관심 이외의 모든 측면에서 가능한 한 유사한 조사방식을 수립하기 위한 설계가 되도록 한 것이다. 비록 두 방법들 모두 유용하지만, 여러 가지 이유로 조사방식의 효과에 대한 광범위하게 관련 연구를 종합하는 것은 다음과 같은 이유로 쉽지 않다. 첫 번째로 이러한 연구에 대한 초점은 매우 광범위하거나 매우 초점이 협소하기 때문에 다양한 오류의 근원들에 따라 동시에 여러 조사방법들을 비교하는 것이 어렵다. 두 번째로, 두 가지 이상의 조사방법에 대한 비교는 상대적으로 드물다. 그래서 다수의 대응비교에서 나온 조사방법 효과들에 관한 완벽한 결합을 찾는 것이 향후 연구과제가 될 수 있다. 추가적으로 몇몇 조사방법간의 비교에서는 면대면 조사방법과 웹조사 방법 간 비교 보다는 전화조사와 면대면 또는 웹조사와 이메일방법의 비교가 더 많이 이루어졌다. 세 번째로, 연구의 설계가 매우 다양하게 제시됨으로서(인구연구, 표집틀, 질문과 답변, 자료수집규모의 실행 등), 사례연구에서 다른 조사환경과의 비교로 일반화하기는 어렵다. 마지막으로, 어떻게 각각의 조사방법들을 적용하는지에 대한 상세한 설명들이 제시되어지지 않고, 어떻게 이런 조사설계가 이루어졌는지 대해 불완전한 하기 때문에 단순한 비교가 불가능하다. 이러한 경우의 한 예는 PAPI와 CATI의 비교에서 설문지의 레이아웃 차이가 조사방법의 차이가 아니라는 결과를 얻기도 했다.

현재 많은 비공식적 또는 공식적적인 메타분석이 이루어지고 있으며, 이러한 분석이 조사오류와 비용에서 보다 더 전체적으로 조사방법들의 차이를 설명하는데 도움을 주고 있으나, 이러한 비교가 여전히 특별한 경우에 주목하는 경향이 있다. 예를 들어 두가지 조사 방법간의 메타분석은 현재의 웹조사와 다른 조사방식들(주로 이메일)에서 응답률 차이를 분석하였다. 비록 지금 웹조사는 일반적으로 이메일 조사보다 낮은 응답

를 볼 때는 그것이 확실하지만, 아직 이러한 경우의 원인이나 어떠한 상황 하에서 응답률이 더 높아지거나 낮아지는지에 대한 차이는 명확하지 않다. 게다가 무응답률에 대해서는 상당히 많은 연구가 수행되었으나, 이러한 두 조사방법들 간의 무응답 편향 차이에 대해서는 거의 연구된 것이 없다.

이와 유사하게 많은 연구들이 조사방법의 측정오류 특성에 주목하고 있다. 그리고 다양한 조사방법들 간의 차이의 원인들은 알려져 있다. 그러나 이러한 연구의 대다수는 특정한 질문(응답범위) 혹은 효과의 종류(사회적 바람직함, 응답순서효과)에 한정되어 있다. 예를 들어 Holbrook, Green, Krosnick(2003)은 사회적 바람직함의 편향과 만족에 주목했다. 반면에 Ye, Fulton, Tourangeau(2011)는 응답범위에서 확실한 것과 극단편향을 연구하였다. 조사방법들 간에 측정오류율 차이의 범위와 이러한 효과를 제거하기 위해 무엇이 필요한지에 대한 몇몇 상충된 의견들이 여전히 존재한다. 몇몇 연구에서는 각각 다른 조사방법에 대해 사람들이 질문에 답하는 방법에 따라 차이가 있다고 주장하기도 하였고, 다른 연구에서는 이러한 차이가 작거나 어떤 경우에는 존재하지 않는다고 주장하기도 하였다. 이러한 차이가 있다고 주장하는 부분은 비교하고자 하는 조사방법들, 연구 질문에 관한 부분, 조사방법 효과를 연구하기 위해 사용되었던 방법들에 대한 부분들로 한정되는 경우가 대다수였다. 게다가 몇몇의 연구에서 발견된 조사방법간의 차이점은 응답자가 새로운 기술에 더 편안하게 느끼거나 조사 설계자들이 조사방법 간의 차이 효과를 완화하도록 학습함으로써 시간의 흐름에 따라 감소시킬 수 있다는 사실들이었다.

Groves와 Kahn(1979)의 연구에서는 전화조사와 대면조사간의 응답자의 반응에 차이가 거의 없다는 것을 보였으나, 차이점이 분명하게 발견되었다. 그리고 이러한 차이는 일부 홍보 전단을 사용하거나 조사에서

의 민감한 질문에 대한 사용하는가의 여부에 따라 차이가 나타났다. 이와 유사하게 많은 연구들이 이메일 조사와 웹 조사에 대한 차이를 연구하였다. 특히 이들은 대부분 같은 조사방법으로서 메일조사와 웹조사 둘 다 자기기입식 조사 방법이며, 따라서 조사방법간의 차이가 없을 수 있다. 그러나 차이가 큰 조사방식(즉, 전화조사와 웹조사)을 비교해 보면 더 큰 차이가 나타날 수 있다. 작업 기억용량, 사회적 바람직성 효과, 인터뷰의 속도, 좋아하는 것 같은 요소들이 모두 동일하지 않다면, 이 두 조사방법 간의 큰 차이가 있을 수 있다.

질문의 유형에 관한 측면에서 몇몇 유형의 질문에서 조사방식의 차이를 기대할 수 있다. 사실에 기반을 둔 질문들이나 “예/아니오”로 표시된 민감하지 않은 질문들은 조사자에 따라 영향을 받지 않으며, 일반적으로 시각이나 청각자료의 제시에 영향을 받지 않는 것이다(물론 질문 자체가 복잡하지 않다는 것을 가정하고). 반면에 민감한 질문의 경우 조사방법에 따라 차이가 있다는 것은 인지하고 있으며, 따라서 어떤 조사방법은 질문의 민감성에 영향을 받을 수 있고, 그렇지 않을 수 있다.

세 번째는 조사방법들과 무관하게 측정 차이를 비교하기 위해 일반적인 조사 설계에서 표본을 하나의 조사방법과 그 외 다른 조사방법에 임의적으로 할당하였다. 이러한 비교는 측정 차이에서 선택 효과(특히 포함률과 무응답)에 대한 구별을 어렵게 만든다. 다시 말해서 측정 차이는 조사방식의 측정비율 보다는 응답자 그룹에서 선택 차이 때문에 생기기 때문이다. 이것을 해결하기 위해서는 연구자들이 표본선정단계에서 자원의 피실험자들의 대조를 제한할 수 있다. 다른 방법으로 두 조사방법의 모든 항목을 같은 주제로 반복하여 측정하는 것이다. 이러한 환경 하에서는 하나는 응답자들이 먼저 응답한 내용을 회상하여 유사한 형식을 두 번 응답하게 유도할 수 있는 조건을 만들 수 있다. 측정오류에 대한 차이 연구에서는 조사방법의 차이에 대해 질문의 유형에 따라 측정을

어렵게 하기도 한다.

최근 10년간 비록 조사연구분야에서 조사방법 간의 측정차이의 원인을 이해하는데 많은 진보가 있었지만, 상대적으로 여러 개의 조사방법간 효과를 체계적으로 연구하거나, 다른 조사방법들 간의 이론적인 통합에 대한 시도는 de Leeuw와 van der Zouwen(1988), Tourangeau, Rips, Ransinski(2000), Jackle et al(2010) 의 연구를 제외하고 거의 없었다.

이와 같이 측정 모형들의 지속적인 발전은 기존의 조사방식에서 아직 시험해보지 않은 조사방법 혹은 심지어 이러한 조사방식의 특유의 특징에 기반하는 아직 개발되지 않은 조사방법으로 추론이 가능하게 할 수 있다. 다시 말해서 단순히 더 많은 조사방식에 대한 비교 연구가 아니라 조사방식간의 차이가 발생하는 경향의 원인과 시기를 더 잘 이해할 수 있는 연구가 필요하다고 할 수 있다.

Deming(1944,p362)이 지적한 바에 따르면, “문제는 차이가 있는지 없는지가 아니라 차이가 얼마나 큰 가, 왜 차이가 존재하는가, 차이들이 자료로 구성된 사용에서 어떤 영향을 미칠 것인가?” 이다. 조사방식을 선택하는 것은 오류의 서로 다른 원인 간의 차이를 평가하는 것을 의미한다. 비록 오차(응답률, 측정오류)의 원인에 대한 조사방식의 그럴듯한 효과에 대해 많이 알려져 있지만, 서로 다른 오차의 원인에 대한 상대적인 효과에 대해서 아직 충분히 인지하지 못하고 있다.

혼합 조사방식의 자료수집에 대해 이 주제에 대한 AAPOR 학회에서 발표된 논문의 수가 증명하는 것처럼, 많은 관심을 받고 있는 분야이다. 조사방식을 혼합하는 것은 누구도 개별 성분의 비율을 알 때 까지는 성분 혼합에 대해 배울 수가 없기 때문에 요리와 매우 유사하다고 할 수 있다. 자료수집을 위해 조사방식을 혼합하는 것에는 많은 서로 다른 방법이 있다. 그리고 서로 다른 조사방식을 혼합하는 방법은 조사의 오차와 비용의 모든 원인에 영향을 줄 수 있다. 또한 조사방식들은 비용절약,

응답률을 상승시키는 것에서부터 포함편향, 특정한 하위집단을 목표로 삼는 것, 측정을 향상시키는 것까지를 포함하여 여러 가지 이유로 조사 방식을 혼합하여 사용하고 있다.

De Leeuw(2005)는 자료수집에서 혼합된 조사방식에 다양한 방법들을 상세하게 구분하였다(Dillman, Smyth, Christian 2009도 유사한 분류를 한바 있다). 이러한 구분의 핵심은 서로 다른 조사방식이 접촉단계, 응답단계, 아니면 후속단계에서 사용되는지 여부이다. 사전통지나 접촉에 대해 조사방식을 혼합하는 것(전화조사를 위한 사전 편지, 웹조사를 위한 초대장 메일을 보내는 것)은 한 동안 관행처럼 행해졌다. 이와 유사하게 설문지의 다른 부분에서 서로 다른 조사방식을 사용하는 것(CAPI에 귀로 듣는 CASI를 더한 것처럼)도 또한 일반적이다. 요즘의 이러한 관심의 증가는 일부 표본에 대해 면접을 하거나, 한가지 조사방식에 대해 자료를 제공하고, 다른 표본은 다른 조사방식을 사용하는 혼합된 조사방식설계에 초점이 있다. 이러한 종류의 혼합된 조사방식의 조사는 스스로 여러 가지 형태가 될 수 있다. 초기의 예는 전화조사를 통해 표본의 일부가 조사되고, 그 표본들 이외에는 사람에게서는 지역표본을 통해 조사하는 이중 표집틀 설계이다.

또 다른 예는 조사방식이 어디에서 사용되었는지가 처음에서 결정되는 것이 아니라 자료수집과정에서 결정하는 것이다. 예를 들어 전체 표본이 처음에 하나의 조사방식을 사용해 조사되지만, 무응답자들에 대해서는 다른 조사 방법을 사용하는 것이다. 즉, 확률표본으로 부터 방문 면접조사를 수행하고, 무응답자들에 대해서는 전화나 우편조사를 수행하는 것이 사례가 될 수 있다.

비록 이러한 구분이 많은 변수들을 완벽하게 포함하지는 못하지만 동시에 다른 조사 방법을 사용하거나, 또는 순차적으로 혼합된 조사방식을 사용하게 된다. 예를 들어, 응답자에 대해 조사방식 선호도(응답자가 어

면 자료수집방법을 더 선호하는지에 대해 응답)와 조사방식을 선택할 수 있게 하는 경우이다. 후자는 응답자가 두 개 이상의 조사방식이나 그들이 선택한 특정한 조사방식의 선택할 수 있는 상황을 의미할 수 있다.

마지막으로 조사방식의 할당은 표집틀 정보, 응답자의 의식적 선호도, 임의적 할당 등에 기반해야 한다. 다시 말해서 누가 혼합된 조사방식에서 응답했는지, 어떤 선택사항이 누구의 통제 하에 있는지를 이해하는 것이 중요하다. 이러한 변동들은 혼합된 조사방식의 맥락에서 조사방식 효과의 논의를 더 복잡하게 만들기도 한다. 또한 혼합된 조사방식의 설계는 시간(추적 연구)이나 공간(국제적 연구)이 확장될 수 있다. 예를 들어 Martin(2011)은 유럽의 사회 조사의 본문에서 다국적으로 반복되는 측정설계인 혼합된 조사방식설계의 세가지 확장을 구체화 시켰다. (1) 국가경계를 넘는 혼합된 조사방식(ACMM), (2) 국가경계 안의 조사방식(WCMM), (3) 시간경계를 넘는 혼합된 조사방식이다. 결과적으로 요점은 혼합된 조사방식설계의 채택이 특히 조사방식의 혼합과 각 조사방식에서 사용하는 응답자의 비율이 역동적이었을 때 시공간의 비교를 복잡하게 만든다는 것이다. 특히 자기기입식을 포함하는 혼합된 조사방식설계의 적용 가능성은 한 조사방식의 단점이 다른 조사방식의 장점으로 보충될 수 있다는 것이다. 즉, 한 개의 조사방식보다 더 많이 사용하는 것은 단일 조사방식에 비하여 응답률은 증가할 것이며, 단일 조사방식보다 다른 그룹의 응답자들의 조사에 들어오기 때문에 잠재적으로 무응답 오차가 줄어든다.

따라서 지금까지의 논의로 볼 때 조사자의 관점과 응답자의 관점에서 혼합방식이 적절한 것으로 판단된다. 예를 들어 Link와 Mokdad(2006)는 그들이 전화방식에서 이메일과 웹을 추가함으로써 등록되어 있는 전화번호에 대한 응답률을 높일 수 있다고 주장하였다. 그러나 전화조사에서 과도하게 대표된 같은 그룹들은 또한 대체 조사방식을 사용하면 이

역시 과도하게 대표된다. 한편 혼합된 조사방식은 인구학적인 편향을 증가시킨다. 이와 관련하여 많은 사람들이 보다 더 비용이 적게 소요되는 조사방식을 이용하도록 장려함으로써, 조사 자원을 비용이 많이 소요되는 조사방식에 특수한 집단들 목표로 투입할 수 있다. 이것은 10년간의 인구조사와 미국 사회조사를 위한 미국 인구조사국의 혼합된 조사방식 전략에서 나온 주장이다. 이러한 주장은 조사방식 간의 측정오류편향이 크지 않다는 것이다(혹은 적어도 혼합된 조사방식의 이점을 무시할 만큼 크지 않다). 동시 혼합 조사방식 설계에서 조사방식에 대해 응답자가 선택하도록 하는 경우 여러 연구들에서 나타났듯이 응답률을 향상시키지 못했다는 사실이다. 그러나 표본으로 추출된 사람들에게 하나의 조사방식과 다른 조사방식을 제시하는 순차 혼합 조사방식이 더 유용한 것으로 나타났다.

제4절 자료수집 방법의 향후 전망

비록 혼합된 조사방식이 종종 비관측의 오류를 줄이는 것을 목표로 하고 있지만 - 인터넷 사용이나 전화 접근과 관련된 포괄성 문제나 이메일을 사용할 때 읽고 쓸 줄 아는 능력과 관련된 무응답 편향의 잠재성을 보완하는 것 - 조사방식을 혼합하는 것은 측정 오류의 관점에서 또 다른 문제를 야기할 수 있다. 가령 조사방식 간에 무시할 수 없는 오차라면, 자료수집과정에서 조사방법을 혼합할 때 서로 다른 측정오류의 문제를 설명하는 두 가지 주된 접근이 있다. “예방”전략으로도 볼 수 있는, 한가지 접근은 조사방식간 측정 오차를 최소화하려 시도한다. 설문지 설계에서 Martin et al.(2007)의 “전체적인 제시”와 Dillman(2007)의 “단일방식 설계” 접근들은 이러한 전략의 전형적인 예가 된다. “수

장” 또는 “조정” 된 접근 이라고 부를 수 있는 대체적인 방법은 측정 오차가 조사방식의 근본적인 특징이고, 측정오차와 조사방식을 분리하여 설계할 수 없다. 이러한 접근은 예방전략을 특성화할 수 있는 최소한의 공통분모를 도출하기위해 타협하기 보다는 각각의 조사방법의 설계이익 (혹은 각 조사방식의 최선의 설계)을 최대화 하도록 해야 한다. 그러면 측정차이는 사용된 조사방식에 대해 비교 가능한 측정값을 산출하기 위하여 통계적으로 수정되어야 한다. 이러한 연구가 얼마나 쉬운지 또는 어려운지는 이용된 조사방식에 좌우될 것이다. 예를 들어 자료수집을 위해 다른 방식과 비교하여 경쟁력이 있는 웹조사는 하나의 특별한 다용도의 조사방식이다. 예를 들어 웹조사를 스크롤링 하는 것은 많은 측면에서 메일 조사와 비슷하다. 반면에 페이징 하는 것은 전화조사와 더 비슷한 양상을 갖는다. 이와 유사하게 항목결측값은 메일조사나 면접자 기입방식의 조사와 유사한 방법에서 발생할 수 있다. 또한 오디오와 비디오는 웹 조사에서 그 조사의 근본적인 성격을 바꾸어 추가될 수 있다. 또한 대면조사의 경우 민감한 질문들에 대한 자기기입을 허용하고, 쇼카드의 사용, 조사자 관측, 물리적 측정, 다른 조사의 광범위한 개선과 같은 많은 다양성을 제공할 수 있다. 전통적인 전화조사는 음성대화만을 허용하고, 메일 조사는 시각적인 소통에는 한계가 있다. 이런 의미에서 이러한 두가지 조사방식들은 가장 제한적이라 할 수 있다. 결과적으로 이러한 조사방식들을 포괄하는 혼합 조사방식설계는 하나의 조사에서 동일한 효과를 나타내도록 하는 조사방식의 동등성(equivalence)을 달성하는 것이 더욱 어려울 수 있다.

혼합조사방식 설계의 또 다른 유형은 조사방식들이 응답자그룹 보다는 질문내용을 특정한 그룹에 대해 조사하도록 목표를 둔 것으로서 측정오류 분석에 초점이 있다. 어떤 특정 유형의 질문에 특정한 유형의 응답자들에 관한 조사방식의 특징과 효과에 관해 더 많이 알수록, 측정의

질을 개선하기 위해 유연한 조사방식을 더 많이 개발할 수 있다.

새로운 조사방법들을 개발하고 그들의 특성을 이해함으로써, 일단의 설문을 규정하여 특정한 방법으로 목표를 달성할 수 있을 것이다. 이것은 이미 민감한 질문들에 대한 CAPI 면접에서 CASI 요소를 이용하여 실행되고 있다. 한편 웹조사에서 다른 조사 방식 설계를 고려할 수 있는 한가지는 같은 조사방식에서 다른 방식과 보다 표준화된 설문으로 대화식 면접을 활용하는 것이다.

위에서 기술한 최근 경향의 중요한 영향중 하나는 조사방식에 대해 어떻게 응답하는지와 관련이 있다. 조사에서 활용되는 자료수집방법을 더 잘 설명하는 방법이 필요하다. 예를 들어 혼합된 조사방식 연구결과를 공표할 때, 단일 응답률을 제공하는 것은 불충분하다. 이 경우 각각의 조사방식의 응답률을 밝히는 것이 필요하다. 이와 유사하게 사용된 협약의 세부사항들 또한 조사의 가치나 그 결과를 판단하는데 필요하다. 또한 사용자나 분석자를 위한 보다 많은 문서화가 필요하다. 예를 들어 특수한 항목이 면대면 자기기입식으로 이루어졌는지 아닌지를 아는 것은 중요하다. 그러나 이것만으로 부족할지도 모르지만, 모든 사람들이 의도한대로 자기기입식 조사 방법의 경우 응답내용이 불완전할 수 있기 때문이다. 이와 유사하게 혼합된 조사방식 설계에서 어떤 조사방식이 특정한 항목에서 특정한 응답자들이 반응하는지에 대해 분석자가 아는 것은 중요하다. 무응답편향의 증가가 중요한 통계의 특성이라는 인식과 함께, 조사방식의 효과는 최고의 조사항목과 응답자의 수준에서 이해될 수 있다. 특정 문항에 대해 조사방법이 시간에 따라 응답자들 간에 변화할 수도 있는 이러한 이슈는 종단 연구자들에게는 중요한 문제이다. 이러한 측면에서 조사과정의 문서화는 매우 중요하며, 조사항목의 변화와 응답자의 태도변화를 연구하는데 유용한 자료가 된다.

제5절 조사방식의 미래와 미래의 조사방식

미래에 대한 예측은 항상 위험하지만, 조사방식에 관해서는 적어도 확실한 것은 조사방식은 계속해서 진화할 것이고, 사람들이 어떻게 의사소통하는지에 대한 사회적 변화에 대응하고, 기술적인 발전에 대응하는 방법은 지속적으로 진화 할 것이라는 것이다.

향후에 어떤 조사방식들이 사라지거나 무용지물이 될 것인가? 에 대해서는 잘대로 그렇지 않다는 것이다. 인터넷이라는 개념이 도입된 초창기에 인터넷 조사가 전화조사를 대체하고 모든 가능한 면접조사방식을 대체하게 될 것이라 주장했다(Couper, 2005). 그러나 이런 예측은 아직 일어나지 않았다. 이와 유사하게 CATI와 CAPI는 모든 상황과 모든 조사를 위해 PAPI의 등가로 대체되지 않았다. 비록 컴퓨터를 이용하는 조사방식이 현재 다수를 차지하고 있지만, 아직 펜과 종이를 이용하여 자료를 수집하는 조사들이 많이 존재한다. 비록 부분적으로 새로운 조사방식이 측정의 향상, 비용의 감소 등의 몇몇 문제점을 해소하고 있지만, 포함률과 무응답과 같은 문제점은 아직 해결하지 못하고 있으며, 따라서 새로운 조사방식이 기존의 조사방식을 대체하기보다 기존의 방식을 보완하는 방향으로 진화할 것이다. 다시 말해서, 모든 연구문제에 대해 모든 것을 해결할 수 있는 조사방식은 없다는 것이다. 다중 조사방식, 혼합 조사방식은 향후 연구조사를 위한 피할 수 없는 현실로서 받아들여질 것이다.

면접자가 더이상 필요 없을 것이라는 예측은 조만간 발생하지는 않을 것이다. 비록 조사자가 비용이 비싸고 오류를 가져올 수도 있지만 그들의 가치는 -표본설계 실행(지역표본 추출 설계에서 가구를 목록화하고 추출하며, 응답자들을 가구내에서 추출하는 것), 무응답 최소화, 물리적 측정 관리, 관측 설계, 설득, 분석, 면접 도중에 응답자를 이끌어 내는

것 등 무시할 수 없다. 이러한 측면에서 면접자의 역할은 더 진화할 것이다. 아마도 면접자들은 조사에 대한 협조를 얻거나 물리적 측정과 같이 추가업무를 관리하는 중요한 역할을 담당하고, 측정과정에서의 실제 질문/답변 부분에서 역할은 비중이 감소할 수 있을 것이다.

상호음성응답(IVR: interactive voice response)과 인터넷 조사와 같은 자기기입식 조사방식들이 저렴하고 조사설계가 용이할 수 있다. 또한 언제 그들이 조사가 끝나는지의 측면에서 응답자에게 어떤 편의를 제공할 수 있다. 그러나 우선 어떤것이 응답자로 하여금 자발적으로 응답하도록 동기를 부여할 것인가에 대한 고려가 필요하다. 응답자의 결단력이 요구되는 조사방식에 의존하는 것은 선택적인 편향을 야기하거나 무응답편향이 발생할 수 있다. 게다가 인터넷조사나 메일조사와 같이 어느 정도 읽고 쓸 줄 아는 능력이 요구되는 조사방식에 의존하는 것은 특정한 유형의 연구에 대해 일반화 가능성을 제한 할 수 있다. 이러한 이유로 면접자기입방식의 조사의 역할이 점점 제한될지라도 조사연구의 도구들 중 중요한 부분을 차지할 것으로 사료된다. 면접자 기입방식의 조사가 지속될 것이라는 또 다른 이유는 인터넷 조사(특히) 조사의 수를 크게 증가시킨다는 것이다.

시간이 지날수록 조사에 대한 응답률은 떨어지고 그에 따라 조사를 통해 정보를 얻는 것이 점점 어려워지기 때문에 조사기관에서 더 많은 시간과 노력을 통해 조사방식에 대해 잠재적인 응답자들에게 더 많은 홍보와 노력이 요구된다. 텔레마케팅이 전화조사에 직면하고 있는 문제 때문에 지탄을 받을 수 있는 것과 유사하게 인터넷 조사를 채택함으로써 전형적인 연구노력을 감소시킬 수 있다. 조사 연구자들이 직면한 도전은 일정한 홍보를 통해 각 조사의 중요성을 인지하도록 교육해야 한다. 이와 같은 맥락에서 조사가 대중들의 시간과 태도에 가치 있고 사회에 대한 기여라기보다는 조사에 참여하는 것을 즐길 수 있는 환경이

되도록 해야 한다. 조사 자료수집에서 자동화된 방법(river sampling, IVR, robo-polling, 그리고 심지어 온라인 접근 패널같은)의 중요한 전제는 응답자가 제한 없고 재생가능한 자원으로 구성되고, 응답자의 패턴이 유사하다는 것이다. 이 경우 각각의 잠재적인 응답자에게 접촉하는 비용은 거의 0에 가까워 질것이지만, 응답률은 낮아질 것이라 예상된다.

응답자의 목표숫자에 도달하기위해 얼마나 많은 재접촉이 필요한지는 별로 문제가 안된다. 그러나 자발적으로 응답할 수 있는 응답자들이 점점 줄어들고 있다는 것을 인식할 필요가 있다. 이러한 근거는 다수의 조사에 참여하는 특정의 “전문적인 응답자” 부족에 대한 우려에서 찾을 수 있다.

미래의 조사방식은 조사질문에 대해 비용이 저렴하고, 쉽게 응답할 수 있는 새로운 방법을 찾는 방향으로 간다는 것이다. 잠재적인 응답자를 구별하고, 그들의 참여를 요청하는 방법을 개선할 수 없었다면, 이러한 접근은 오직 제한된 부분에서만 성공했을 것이다. 새로운 매체를 이용한 조사의 초기의 흥미가 사라져버린다면, 또는 연구자 또는 시장 경영자들이 시류에 편승한다면, 새로운 방법으로부터 발생한 초기의 이익은 쉽게 사라져버렸을 것이다.

이것은 조사 연구자가 응답자들로 부터 얻는 정보가 매우 소중하다는 기본적인 관점을 상기시킨다. 조사참여는 결코 이타적인 시민의 의무나 다른 한편으로 경제적 교환으로 보지 않고, 사회적인 보답으로 응답한 것으로 간주되었다. 오로지 조사연구자로서 이익 측면에서만 응답자를 바라보는 것은 점점 응답자에게 중요할 수도 있는 사실들을 점점 무시하고 있다는 것을 의미한다. 조사에 참여하도록 잠재적인 응답자들을 납득시키려는 시간과 노력이 드는 바로 그 행위는 자동화된 방법으로 해결되지 못하는 조사의 가치를 그들에게 알려주는데 도움이 된다.

미래에 자동화된 전화 면접(IVR)을 위해 향후에 해야 할 것은 무엇인

가? IVR에는 여러 가지 장점이 있다. 다른 방식(예를 들어 인쇄된 수령증이나 메일)을 사용하여 응답자를 초대하는 Inbound IVR은 짧은 설문과 반복된 측정으로 몇몇 사업체 조사와 고객만족도 조사에서 인기가 있었다. 면접자를 포함한 Recruit-and-switch IVR(또는 telephone audio-CASI)은 민감한 주제에 대한 조사에 이용된다. 그러나 RDD표본을 바탕으로 하고 면접자가 개입되지 않는 자동화된 outbound IVR(robo-polls)은 정치여론조사에서 인기를 얻을 것으로 보인다.

최근 2008년의 선거전의 여론조사에서의 AAPOR 위원회는 “컴퓨터를 이용한 전화면접 기술과 상호 목소리 응답 기술 둘 중 하나를 사용했을때 추정치의 정확성에 차이가 없다”고 결론지었다. 이 결론은 robo-pollsters를 더욱 장려할 수 있지만, 여론과 사회조사의 다른 유형으로 사전선거 여론조사와 같은 협소한 주제로 부터 일반화하는 것은 아직 이른 것으로 보인다. 게다가 Survey Practice의 최근 기사는 live-interviewer 조사에 관한 robo-polls에 참여했던 응답자들 간의 차이점이 발견되었다. 확실한 것은 IVR의 미래에 대한 결정을 내리기 전에 더 많은 증거들이 광범위한 주제에 걸쳐 수집되어야 한다는 것이다. 그러나 무응답편향에 대해 질문과 특히 휴대폰에서의 자동화된 전화의 합법성에 관한 이슈는 전화조사의 대안의 성장을 제한할 수 있다.

주요 조사방식에 기반을 둔 정보와 소통 기술들은 계속 진화한다. 오늘날의 보다 대중화된 조사방식에 대한 타당성의 두 가지 예는 전화조사방식과 인터넷에서의 발전에 있다. 전화조사방식에 관해서 이미 일반 유선전화에 기반한 전화에서 휴대폰으로 전환되고, 표본추출, 포함, 무응답, 측정오류들의 영향을 처리하는 과정에 있다. 그러나 현재 조사연구자들은 모바일 기기들을 어떻게 사용하는가에 대한 연구에 큰 변화를 진행하고 있다. 휴대폰은 목소리 기반 통신기기보다 더 많은 기능을 가진다. 다양한 형태의 단문 메시지 서비스는 휴대폰의 웹기반 서비스들처

럼 입지를 넓히고 있다. 모바일 기기에 대한 통신 방식의 범위는 급격하게 넓어지고 있다. 스마트폰은 이미 목소리, 글자, 시각적 물질, 공간적인 위치, 움직임, 추가하드웨어나 응용소프트웨어를 통해 가능하게하는 다른 입력과 출력 방식들의 호스트들을 처리할 수 있다. 어플리케이션들은 이미 음성메시지를 텍스트로 바꾸고, 그 반대 기능도 가능하다. 스마트폰은 정말 다용도의 기기로서, 스마트폰을 조사에 어떻게 사용할지가 다음 몇 년간의 큰 도전이 될 것이다.

스카이프와 같은 VoIP서비스의 사용은 초기 비디오폰을 실용화 하여 인터넷을 통해 “대면” 면접을 하는 것을 가능하게 만든다. 이러한 통신의 새로운 방법을 조사 연구자에게 효율적인지 혹은 우리가 찾고 있는 응답자들에게 효율적인지 그리고 어떻게 효율적인지는 연구해야 한다.

또한 인터넷은 텍스트 기반의 정보 매체에서 소셜네트워킹을 위한 멀티미디어로 빠르게 진화한다. 페이스북과 트위터같은 소셜 미디어 사이트의 부상은 인터넷이 어떻게 사용되는지에 대한 큰 변화의 전조를 보인다. 연구자들은 이미 소셜미디어가 사회에서 조사를 하거나 주요 트렌드를 측정하는데 어떻게 사용되는지 연구하고 있다. 그들의 삶을 온라인에서 점점 상세하게 디지털 흔적으로 남기는 소셜미디어의 사용자로서, 넓은 분야의 현상에 간접적 관측 기회가 생길 수도 있다. 그러나 이러한 접근은 응답자들의 동의 문제를 공지하고 전체 인구집단의 포괄성에 대한 고려가 반드시 필요하다. 따라서 우리가 점점 조사로부터 수집하는 정보의 형태를 보충하기위해 이러한 자료의 소스들이 증가할 것이라는 것에는 동의할 수 있다.

한편 최근 소셜미디어의 폭발적인 증가에 대해 조사 연구자들은 깊은 관심을 가지고 있다. 2011년 9월 페이스북은 등록된 사용자가 7억 5천 만명을 넘어섰다고 보고되었다. 그것은 중국, 인도 다음으로 세계에서 세 번째로 큰 나라를 만들 수 있을 정도이다. 게다가 전하는 바에 따르

면 2억 5천만명 이상의 사용자들이 모바일 기기로 페이스북을 접속한다. 이와 유사하게 트위터는 2011년 4월에 등록된 사용자가 2억명 이상이라고 보고되었다. 이러한 소셜 미디어의 트렌드를 무시하기는 어려운 것이며, 따라서 연구자들은 이러한 사이트의 대표성에 많은 관심을 가지고 있다. 그러나 페이스북과 같은 사이트를 사용하여 조사하는 것은 조사 연구자들이 직면하는 일상적인 추론적인 문제들 때문에 많은 도전이 된다. 먼저 대규모의 이용자 수에도 불구하고, 페이스북 사용자들이 특정한 나라의 전체 인구를 대표하는 페이스북 사용자들은 거의 없고, 그들의 권리에만 관심있는 것이 사실이다.

두 번째로, 페이스북 사용자들의 집합은 표본추출 틀처럼 외부사회에서 구할 수 없다. 연구원들은 페이스북 사용자들을 연구하기 위해 다양한 비확률적 방법(snowball sampling: 눈덩이 추출법)을 사용하는 것에 한정된다.

세 번째로 페이스북 사용자들은 점점 보안문제에 대한 중요성을 인식하고 있다. 그리고 새로운 기기들은 사용자가 어떤 콘텐츠에 접근하는 누군가를 제어할 수 있도록 개발되고 있다. 이것은 페이스북 사용자들의 선택적인 부분집합에서 점점 사용자에서 정보를 공개적으로 공유하려는 사람들까지의 자료에 대한 간접적인 수집 까지도 제한한다. 이러한 풍부한 데이터 소스들은 여론에 대한 이해의 폭을 넓히는데 도움이 되지만, 가까운 미래에 기존의 전통적인 조사방법을 대체하진 못할 것이다.

새로운 기술들은 통신의 새로운 수단을 제공하고, 또한 그것들은 점점 새로운 기술로 소통하는 누군가를 제어하는 수단으로 사용자를 제공하며, 알려진 관련인들의 제한된 공간에 접근을 제한한다. 이것은 사람을 표본추출하는 것과 그들로부터 협력을 얻어내는 것, 그리고 조사방식의 선택과 조사 설계가 어떻게 그 결정에 영향을 끼치는지에 대한 기초적인 도전들이 조사작업에 주요 이슈로 남는다는 것을 의미한다. 이와 유

사한 방법으로 만약 우리가 인터넷에서의 광대한 양의 공공 정보를 사용하면, 우리는 더 넓은 사회에서 우리가 정보를 공유하려는 사람과 그렇지 않은 사람이 어떻게 다른지, 그리고 어떤 주제가 선택 편향들에 더 혹은 덜 민감한지에 대한 연구가 더욱 필요하다. 따라서 조사와 관련된 주된 연구주제로서는 사람들이 왜 조사에 참여하고자 하는지 - 또는 왜 그들의 정보를 온라인에 공공연하게 공유하는지 - 에 초점을 맞추는 것이 필요하다.

이미 진행중인 다른 트렌드는 통신기술이 점점 동시에 존재하지 않는다는 것이다. - 음성메일, 단문 메시지, 기타 같은 종류의 것들은 동시에 연결된 두 집단이 소통하는데 더 이상 필요로 하지 않는다.

조사원 기입방식의 조사는 두 집단이 동시에 소통할 수 있는 동시에 발생하는 모델에 크게 기반을 두고 있다. 이러한 통신기술의 변화가 조사활동에 어떻게 영향을 끼칠지는 두고 봐야 하지만 조사들은 언제, 어디서, 어떻게 질문들이 답해질지에 대한 더 큰 선택권을 가지고 있는 응답자들로 이루어진 더 조그만 덩어리들로 부서지는 것을 의미할 수도 있다.

마지막으로, 조사 설계의 모든 다른 요소와 같은 조사방식들은 발전하고 있다. 다양한 조사방식들과 조사방식 조합물들은 조사 연구자들이 사회적, 기술적 변화에 맞춤으로서 확장이 지속될 것이라는 지배적인 견해이다. 시장점유 또는 응시관찰이든, 혁신이나 새로운 방법을 만드는 것에 대한 필요와 주요 종적인 추정과 분석으로 시간에 걸친 비교가능성을 유지하는 것의 중요성 사이의 지속적인 긴장이 있다. 이러한 긴장은 건전한 것이며, 이러한 접근들 모든 연구에 가치가 있다. 조사연구에서 하나의 상수가 변화하는 것으로 보인다. 그리고 외면적인 변화와 조사를 향상시키는 방법론적인 연구 모두에 응하여 현존하는 방식을 적용시키는 것과 새로운 방식을 만드는 방법을 찾아야 한다.

제6절 정책적 시사점

현재 보건복지분야의 조사통계는 가구와 시설을 중심으로 면접조사가 대다수를 이루고 있다. 국가통계 생산의 측면에서 신뢰성 있고, 고품질의 통계생산을 위해 현재 국내 현실에서 면접조사방식을 채택하고 있는 것은 피할 수 없는 현실이기도 하다. 해외 연구사례를 통해 살펴본 조사 방법의 현실과 미래는 국민의 삶의 질 향상에 필요한 정책통계를 생산하고 관리하고 있는 정부 부처와 관계기관에서는 조사통계 생산 현실을 직시하고, 이를 개선하기 위한 정책을 강구해야 한다.

해외의 경우 초기에 전화조사를 기반으로 다양한 조사 방법이 개발되고 실제 조사현장에 적용하여 통계이론과 조사방법의 발전이 동시에 이루어졌다. 신뢰성 보다는 비용과 응답자의 측면에서 다양한 조사방법 또는 자료수집 방법의 개선이 있었으며, 통계 및 사회학자들을 중심으로 다양한 방법들이 적용되었다.

전화 조사의 응답률을 높이기 위해 전화조사와 면접조사 또는 웹조사를 병행한 방법이나, 반대로 웹조사의 참여율을 높이기 위해 전화조사나 면접조사를 병행하는 혼합방식의 사례는 현재 국내 고비용 조사환경에 대해 의미하는 바가 크다고 할 수 있다. 국내의 경우 국가통계 작성에서 전화조사와 면접조사, 웹조사를 혼합하는 방식은 전혀 고려된바가 없다. 또한 이와 관련된 통계적 이론의 개발도 일부 학자들에 의해 연구된 바는 있지만, 그리 유의미한 결과를 얻지 못하였다. 또한 최근 이슈화 되고 있는 사회연결망을 통한 조사방식과 영상음성인식 장비를 통한 조사 역시 연구단계 머무르고 있는 실정이다.

국내 보건 및 복지분야 통계는 앞서 언급한 바와 같이 국민의 삶의 질을 개선하고 향상시키기 위한 중요한 정책 통계로서의 역할을 수행하고 있다. 전통적인 조사방법의 고수는 지속적으로 고비용 저효율의 문제

를 야기시킨다. 이러한 측면에서 향후 조사 방식의 개선은 지속적으로 연구해야할 문제이다. 한편 조사방식의 개선은 표본 추출과 선정의 문제와 직결된다. 확률표본추출과 면접조사로 수집되는 정보는 현재 국내 현 시에서 가장 신뢰성을 갖는다고 볼 수 있다. 최근에 각 연구기관이나 조사기관에서 컴퓨터를 이용한 조사 방식을 채택하여 운용하고 있다. 이러한 자료수집방법의 변환은 단시간에 이루어졌고, 그로인한 조사정보의 손실이 상존하는 것이 현실이다. 이러한 문제를 개선하기 위해서는 조사과정에서의 다양한 정보를 동시에 수집할 필요가 있다. 즉, 조사보조정보의 수집에도 많은 노력을 기울여야 한다.

결과적으로 조사방식의 개선은 피할 수 없는 현실적인 문제이다. 그렇다면, 어떠한 방향으로 개선할 것인가의 문제는 표본선정의 문제와 같이 고려되어야 한다. 혼합조사 방식이 좋은 대안이 될 수 있으나, 상이한 조사방식을 적용할 때 발생할 수 있는 조사자료간의 이질성을 극복할 수 있어야 한다. 또한 조사비용을 간과할 수 없다. 대부분의 정책통계는 국민의 세금으로 작성된다. 따라서 효율적인 자원의 배분이 필요하다. 이러한 관점에서 몇가지 대안을 제시하고자 한다.

첫째, 조사방식의 개선은 통계이용자의 측면과 응답자의 측면을 고려하여 저비용 방법을 신축적으로 운용하여야 한다.

둘째, 조사항목의 대폭적인 개선이 필요하다. 보건 및 복지분야의 통계들의 조사표의 양은 타 분야 조사항목에 비해 월등히 많은 편이다. 이를 개선하여 중요 항목에 대한 조사에 집중하고, 나머지 항목은 행정정보 자료 등을 활용하도록 하는 것이 바람직하다.

셋째, 조사방식의 개선과 더불어 표본설계의 방향 또한 조사방식과 연동하도록 개선해야 한다. 다목적 표집틀 활용과 행정 정보의 활용으로 추정 가능하도록 설계방식을 개선할 필요가 있다.

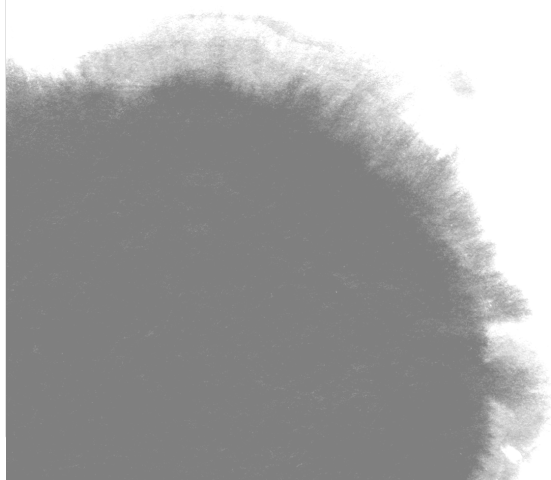
넷째, 사회적으로 민감한 조사, 또는 개인의 사생활이 보호되어야할

조사의 경우 확률화 응답 기법과 같은 조사 방식을 도입하는 것 또한 고려할 수 있다. 개인의 정보 보호에 대한 인식이 점점 높아지고 있는 상황에서 정확한 응답을 이끌어 내기란 갈수록 어려워지며, 이로 인한 응답회피 및 무응답이 증가하고 있기 때문이다.

마지막으로 소셜미디어의 등장으로 현재 인터넷 상으로 수많은 실시간 정보가 유통되고 있다. 이러한 정보를 적절히 가공하고, 분석할 수 있는 big데이터 기반의 통계적 분석 기법의 개발이 IT 기술 기반으로 확대 되고 있다. 조사방식의 혼합과 더불어 이러한 정보를 활용할 수 있다면, 매우 적절한 추론이 가능할 수 있을 것이다.

4장

표본조사의 현황과 미래 그리고 시사점



제4장 표본조사의 현황과 미래²⁾ 그리고 시사점

제1 절 표본조사의 배경

Frankel과 Frankel(1987)은 Public Opinion Quarterly의 50주년 기념호에서 표본조사의 역사 대해 고찰한 바 있다. 이들의 논문은 여전히 인용되고 있는데, 표본추출의 역사에 관해 거시적으로 관심이 있는 독자들은 이들의 고찰을 참고할 만하다. Frankel 과 Frankel(1987)은 두 단계로 표본추출의 역사를 고찰하였는데, 첫 번째 단계에서는 확률 추출방법으로 패러다임이 전환된 Neyma(1934)의 논문에 따라 기본적인 확률추출방법의 수립이다. 두 번째 단계는 기본적인 방법의 개선과 전화 표본추출과 컴퓨터화 정도와 같은 새로운 기술에 대한 확장으로 특성화된다. 이러한 이분법적인 분류는 1978년 이후 표본추출에 많은 발전을 가져왔다. 예를 들어 선거여론조사에서 등록-기반 표본추출(Green 과 Gerber, 2006), 이동전화에 대한 개인 표본추출(AAPOR, 2010) 등은 표준적인 표본추출기법의 혁신이다. 또 다른 기술적인 혁신인 웹조사는 집중적인 연구에서 사용되는 중요한 방법이 되고 있다. Couper(2000)가 언급한 바와 같이 이 시점에서 수립된 웹으로 부터의 표본추출방법은 일반적으로 받아들일 수 없는 방법이며, 이러한 사실은 10년이 지난 현재까지 사실로 받아들여지고 있다. 이러한 표본추출의 문제는 일련의 가

2) M. Bricks(2012)의 "The Future of Survey Sampling" 논문을 편역하여 수록한 내용임.

능한 해법들을 양산하고 있다.

Neyman(1934)에 의해 발전된 확률 표본 추출과 그 이후 거의 모든 조사 표본추출 저서들은 추론에 있어서 매우 구체적인 체계를 가지고 있다. 모집단에 있는 모든 단위들의 프레임을 구성할 수 있고 그 프레임에 있는 모든 단위들이 표본으로 추출될 양의 확률을 가지며, 이러한 확률은 각각의 표본단위들에 대해 측정이 가능하다고 가정한다. 일단 표본으로 추출된 모든 단위들의 특성 값들은 정확하게 측정이 가능하다. 추정치들은 모든 가능한 표본들의 분포를 사용하여 추정된다. 다단계 표본추출과 확률 표본 추출방법과 같은 다양한 방법들이 이러한 일반적인 체계를 만족한다. 무응답, 모집단의 불완전한 포괄성, 측정오차 등은 이러한 순수한 확률 추출의 가정에 위배되는 문제들이다.

최근에 다른 표본추출방법들이 제안되고 있으며, 이들 중 하나는 확률추출 체계내에서 정확하게 일치하지 않는 방법들이다. 이들 중 하나는 응답자-기반 표본추출(response-driven sampling; Heckman, 1997) 방법으로 매우 드문 모집단이나 접근이 어려운 모집단에 대해 사용되는 방법이다. 이 방법은 네트워크(Network) 또는 스노우볼(Snowball) 표본추출로서 조사에 적합한 사람들을 알고 있는 이들에게 드문 집단의 구성원들을 질문하는 방법이다. 이러한 추출방법에 대해 수학적인 모형을 자료 분석에 이용하였다. Volz and Heckman(2008)은 관측된 표본이 특정한 조건하에서 확률 표본과 같음을 보였으나, 이러한 조건들은 현실적으로 매우 드물게 만족한다.

또 다른 두 가지 새로운 표본추출방법은 전통적인 표본추출방법과 매우 근접한 표본추출방법으로 응답자 네트워크를 사용한 적응 표본추출(adaptive sampling; Thompson and Seber, 1996)과 간접표본추출(indirect sampling; Lavalley, 2007)이다. 전형적으로 적응설계는 인구 집단에 대해서는 사용되지 않고, 자연과학에서 드물게 발생하고 군집을

이루고 있는 모집단에 대해 밀도를 추정하는 데 이용된다.

간접표본 추출은 가구나 개인조사에서 사용되었다. 간접표본추출은 목표모집단이 아닌 개인이나 단위의 확률 표본에서 시작되었다. 그 다음 목표모집단을 표본단위와 연계하면 목표모집단의 추출확률을 계산할 수 있다. Kalton 과 Brick(1995)는 이러한 접근 방법을 가구의 종단면 조사에서 첫 번째 웨이브 이후 표본가구로 진입한 개인의 가중치를 할당 하는데 이용하였다. 이러한 접근 방법은 충분히 발전 가능성이 있는 방법이다.

이와는 다르게 추론의 관점에서 표본추출이론을 살펴보면, Godambe (1966)과 Royall(1970)은 설계 기반 접근방법으로서 추출된 표본단위의 확률에 기반 한 모집단의 추론에 대해 언급하였으며, 이들 이외의 연구자들은 다른 통계분야에서 사용되는 보다 전형적인 모형 가정을 이용하였다(모형기반 또는 베이지안 접근 방법으로 불리워짐). Holt와 Smith(1979)는 설계기반 표본추출이 조건부 원리에 부합되지 않을 수 있음을 지적하였고, Hansen, Madow 그리고 Tepping(1983)은 설계기반의 활용성에 대해 주장하였고, 모형기반 접근의 가정에서의 작은 오차가 대규모 표본을 추출했을 때 추정치의 오차에 심각한 편향을 가질 수 있음을 보였다. 근본적인 문제점들이 해결되지 않았음에도 여전히 설계 기반 접근 방법은 표본추출에서 중요한 방법으로 존재한다.

대부분의 조사 자료를 이용한 추론은 설계기반에 의존하고 있지만 소지역 추정과 같은 학문영역에서는 모형기반 방법이 이용되고 있다. 일부 학자들은 확률표본의 시대는 저물고 있다고까지 말하고 있다. 이는 자료를 수집하는 데 소요되는 비용과 무응답 문제, 그리고 비포괄성 등의 이유로 설계기반 추론의 비효율성에 기인한다. 낮은 응답률과 포괄성을 갖는 확률 표본은 비확률표본이나 자발적 표본에 비해 좋지 못하다는 것이다. 이러한 문제는 표본추출 방법의 개발이 필요하고, 향후 직면할 문

제이다. 즉, 자료 수집과 추론을 위해 계속 확률표본추출을 고집할 것인지 아니면 다른 방법을 사용할 것인지에 대한 문제이다.

제2절 통계적 추론 및 비용 그리고 표본추출

앞에서 표본조사의 초창기를 살펴보면, Kaier(1895-1896), Bowley(1926), Neyman(1934)들은 전수조사의 종속성을 표본추출로 바꾼 것이 중요한 점으로 설명할 수 있다. 이러한 변화는 매우 중요하다. 예를 들어 Duncan and Shelton(1978)은 1900년도부터 1940년까지 하나의 혁명으로서 이러한 표본조사의 선구자들의 노력의 결과로 미국 정부의 자료 수집활동에 변화의 효과를 설명하였다. 이는 일찍이 중요한 패러다임의 전환이었다. 이러한 역사적으로 관심이 가는 측면중의 하나는 표본추출에서 혁명은 정부와 공식통계의 수요에 기인한 점이라는 것이다.

표본조사를 채택한 동기는 비록 비용이 중요하긴 했지만, 단순한 비용 효과의 측면이 아니었다. 초기 연구단계에서 비용은 거의 명시적으로 언급되지 않았다. 1960년대 표본추출기법들이 발전한 것과 침예하게 대비되는데 이는 비용 절감이 전화에 의해 수행된 조사와 관련이 있어 시장을 유도하였고, 기타 비정부 단체는 면대면 가구조사를 실시하지 않았다. 이러한 변화 과정에서 정부 연구들은 초창기 연구에 전화조사를 이용하지 않았으며, 최근에 부분적으로 받아들여졌다. 왜냐하면 전화번호의 효과적인 추출이론에 관해서 아직 개발 중에 있기 때문이다.

미국 정부는 1980년대 까지 전화조사 표본추출을 사용하지 않고, 전화조사에 대한 통계적 이론과 표준적인 사례가 개발된 이후 다른 분야에서 정립되었다. 전화표본추출이 앞서 언급한 확률추출의 서두와 같이 하나의 패러다임 전환은 아니었던 한편 두 가지 발전은 이론과 실제의

선제적으로 어떠한 실마리를 제공한 측면이 있다. 두 경우 모두 통계추론 이론의 발전이 선행되지 않았으며, 오히려 통계이론은 기존의 사례들을 지지하기 위해 개발되었다. 최근의 조사환경에서 비용은 아마도 새로운 표본추출을 찾는 가장 중요한 요인일 것이다. 낮은 비용으로 웹상에서 자료를 수집하는 것이 조사관련 요소들에 우선하며, 특히 시장 조사에서는 광범위하게 채택된 통계적인 이론 없이 이러한 방법이 널리 확산될 것으로 본다. 이는 1960년대에 전화표본추출을 비용의 문제로 채택한 원인이 된다. 그러나 통계이론과 자료 수집은 전화 표본추출의 타당성을 뒷받침 해주는 중요한 사항이다. 이론과 실제의 겹침은 현재 웹을 이용한 조사의 낮은 비용에 대해 효과적이기 위해 새로운 표본추출 및 조사방법에 대한 도전으로 남아 있다.

조사방법을 선도하는 기술의 새로운 혁명의 역할은 표본조사의 발전에서 기술의 영향에 관하여 연구한 Frankel과 Frankel(1987)의 논문에서 나타나 있다. Bellhouse(2000)는 조사 표본추출과 기술적 혁신간의 상호작용을 표본추출에 관한 그의 업적에서 초점으로 고려하였다. 몇 년 동안 기술과 표본조사가 서로 연관되어 있고, 각각 서로 다른 영역에서 새로운 발전을 선도하고 있다는 데에는 이의가 없다.

표본추출과 기술에서의 상호 발전이 미래에는 지속되지 않을 것이라 상상하기 힘들다. 사람들이 대화하는 방법에서 엄청난 변화는 많은 양의 표본조사를 하게 만든 기술적인 발전이다. 한 예로 미국의 경우 유선전화에서 무선이나 휴대폰으로의 전환이 있었다. 2009년에 성인의 25%가 일반전화가 없는 집에 살았다. 이러한 발전과 동반하는 것은 대단히 많은 과거 5년간의 휴대폰 표본추출에 관한 논문들이다. 그러나 그 당시 웹조사에 관해서는 거의 연구가 없었다. 그 이후 연구자들은 또한 문자 메시지, 스마트폰, 사회 네트워크에 잠재적으로 적용할 수 있는 조사를 찾기 시작했다. 또한 일반전화 조사보다 더 비싼 휴대폰의 경우를 제외

하고는 이러한 기술들의 대부분은 적어도 부분적으로는 그것들의 적은 비용 때문에 사용된다. 기술의 진화에 따라 이러한 변화들은 증가하고 있으며 패러다임의 이동은 아니라는 것을 인식하는 것이 중요하다.

또 다른 저비용의 대체방법 행정등록 자료를 사용하는 것이다. 비록 이러한 접근은 새롭지 않거나 새로운 기술로 평가되지 않았다. 이러한 예상에도 불구하고 행정등록 자료를 사용하는 것은 그것의 잠재력에 부응하지 못한다. 적어도 미국이 아닌 경우에는, Jabine and Scheuren(1985)은 미국정부가 10년간의 표본추출에서 행정등록 자료의 사용을 늘리기 위해 설정한 여섯가지 목표를 설명하였다. 대략 25년 후 많은 목표들이 대부분 실현되지 못한 것으로 개요를 서술했다. 행정등록 자료로부터 표본을 추출하는 것은 정부 통계자료로 활용하는데 한계가 있을 것이다. 예를 들어 선거 조사의 경우 선거등록명부에서 표본추출하여 조사한다. 하지만 이러한 방법은 행정자료의 불완정성, 시간정확성의 부족, 응답자의 위치나 접촉 정보부족, 기록의 유지 실패 때문에 조사적용 상에 문제가 발생한다. 이러한 문제는 행정등록 자료의 목적이 표본추출하는 목적이 요구하는 수준에 맞지 않기 때문이다. 이러한 문제들 중 몇몇은 국가조사협회의 주 선거인 등록 데이터베이스 검토에서 잘 나타나고 있다. 이러한 사실로 부터 왜 행정등록 자료가 한 특정한 분야에서만 사용가능한지를 확실히 보여준다. 이는 행정등록 자료의 활용에 대한 미래의 도전이 될 것이다.

표본추출이론에 의한 현재의 연구 중 다수는 무응답, 범위 오류, 결측값을 계산하기위한 추정방법들 이다. 이러한 접근은 저비용의 대체 방법이 된다. 왜냐하면 분석적인 방법이나 가중치 조정으로 값비싼 데이터를 대체할 수 있기 때문이다. 예를 들어 가중치 조정은 본질적으로 무응답자로부터 응답자로 가중치를 확대함으로써 무응답자의 가중치를 기본으로 만들 수 있다. 이 분야에 대한 중대한 기여는 calibration estimator

이라 불리는 계층적 추정방법을 만든 Deville과 Sarndal(1992)이다. 그리고 calibration estimators는 무응답과 미포함 항목을 처리하는데 빠르게 일반화되었다. 조사과정에서 결함은 직접적으로 이론에 기반한 설계에서 다루어진다. Rubin(1976)은 대체된 데이터를 위한 분산추정의 방법인 다중대체법을 제안했다. 대체법(imputation)은 일반적으로 미완성된 항목에 채워 넣거나 대체 응답된 것을 사용하는 기법이다. 결측값을 위한 대체법은 수십년간 있었지만, 다중대체법은 새로운 연구로 주목받고 있다.

추정방법으로 사용되는 확률표본추출에서의 데이터수집의 결점을 보완하기 위한 다른 표본추출방법들에 의해 연구되었다. 온라인 조사는 참여자들의 임의적이지 않은 선택을 보완하기 위해 대체적인 추정 방법으로 사용되는데 특히 가중치가 있는 방법에 사용되었으며, 이에 대한 지속적인 연구가 이루어지고 있는 분야이기도 하다.

조사환경에서 비용이 중요하지만 그것이 확률표본추출의 시대의 시작에서 표본추출의 혁신을 선도하지는 않았다. 비록 변화가 있음에도 불구하고 비용은 표본추출방법의 변화에 주요한 대체적 요인으로 간주된다.

최근 10년간 일어나고 있는 조사비용의 증가는 조사 참여가 감소하는 근본적인 현상이다. 전화 자료수집과 대면 자료수집 모두 표본으로 참여하려는 사람이 20~30년 전보다 상당히 저렴하다.

조사 연구에서 이러한 조사에의 낮은 참여율에 따른 영향은 사회적 변화로 기인한 것으로 종종 간주된다. 비록 분명히 사회가 변화하고 있지만, 조사연구에서 조사의 품질을 높이고 비용을 줄이는데 효과적일 만한 수단이 있다. 조사의 분석보다는 응답자에 초점을 맞춘 미래 연구가 이러한 두 가지 목적을 성취하는데 도움을 줄거라 믿는다.

응답자 중심의 방법론은 Dillman(1978)이 제안하였다. 응답자 중심 방법은 조사요청에 대한 응답자의 참여를 높이려는 데 있다. 이렇게 하

여 많은 비용이 드는 데이터 수집의 후속 조치와 가중치 조정에 소요되는 자원을 감소시킬 수 있다. 이러한 근본적인 개념은 응답자가 조사참여의 부담보다 조사 참여의 혜택이 더 크다는 것을 인지하게 하는 방법에 대한 연구가 필요하며, 이것이 가능해진다면 저렴한 비용과 높은 품질을 달성할 수 있다.

제3절 불확실성의 미래

과거와 현재에 대한 이러한 관점은 표본추출이 미래에 나아갈 방향을 암시한다. 가까운 미래에는 Frankel과 Frankel(1987)이 언급한 새로운 기술을 적용하기 위해 기본적 방법과 확장이 지속적으로 이루어 질것으로 예상된다. 하나의 예로써 발전 가능성이 있는 분야는 다중추출틀의 활용이다. 다중추출틀을 사용할 경우 포함률이나 효율을 높이기 위해 두 개, 혹은 더 많은 추출틀에서 개체가 선택된다. 예를 들어 사업체 조사에서 한 개의 추출틀은 완벽하지 않지만, 사업체 목록에 접근하기 쉽고 다른 추출틀은 사업체가 기록된 곳이나 추출된 지리적 위치가 있는 추출틀이 될 수 있다. 두 개의 틀을 사용하는 것은 구할 수 있는 자료를 이용하면서 효율성 높은 표본을 제공한다. 다중표본틀은 행정등록 자료들을 사용함으로써 비용을 줄일 수 있기 때문에 더욱 효과적이다.

추가적으로 다중 추출틀 이론은 현재 휴대폰 설문조사에서도 적용되는데, 이 연구결과는 나중에 설명될 표본추출에서의 몇가지 근본적인 결점을 드러냈다. Brick(2010)은 전체 무응답을 설계단계에서 서로 다른 표본에 대해 최적화당과 다른 가중치로 접근하는 것에 문제가 있음을 지적하였다.

표본추출이론의 전제는 여전히 추출틀 내에서 전체 인구를 완전히 포

괄한다고 가정하고, 추출된 표본들로부터 정확한 응답을 얻을 것이라고 가정한다. 추출들의 불완전 포괄성은 대부분 전체적으로 추정단계에서 대체될 수 있다.

미래의 표본추출의 중요한 새로운 방향은 어렵게 하나의 추정치를 작성하는 효과보다 표본설계와 자료수집 과정에서의 비표본오차의 효과를 고려하는 것이다. 이것은 표본에서 응답자들의 특성을 더 일관성 있게 만들기 위해 노력함으로써 자료수집단계에서 무응답편향을 줄이기 위한 설계를 고려하는 것이 더 좋은 방법이 될 것이다. 관심변수와 강한 상관성이 있는 보조자료들은 무응답편향을 줄이는데 필수적인 요소로 간주되고 있다. 또한 표본설계가 앞에서 언급한 자료들이 이용가능할 때 유용하기 때문에, 표본추출 단계에서 이러한 자료들을 결합하는 방법들에 관한 연구가 새로운 기저가 될 것이다.

자료수집비용은 조사원들에게 저렴한 비용으로 자료수집을 강요하는 요인이 된다. 현재 조사원에 투입되는 조사비용은 전체 조사비용의 약 70%를 상회하고 있으며, 웹상에서 수집된 비확률 조사자료를 보완해주는 통계학적 이론이 반드시 필요하다. 그러나 아직 미해결된 문제는 저비용의 고품질 조사자료의 수집이 확률표본추출 이론에 기반한 설계로서는 성공하기 어렵다는 것이다.

1. 비확률 표본(nonprobability sampling)

확률표본을 대체할 새로운 이론의 성질에 대한 추측은 매우 어려운 일이지만 비확률표본이나 자원자 표본으로부터 나온 추론을 만들기 위해 조사되는 방식에 대한 비판적인 견해는 당연하다. 관측된 자료에 대한 분석은 웹 지원자 표본들의 분석과 가장 밀접하게 관련된 통계적 분야이다. 사실 웹 표본을 이용한 연구는 관측연구를 위해 가장 주목할만한

성향 점수화로 본래부터 발전시킨 기술을 사용해왔다. 이러한 이유 때문에 우리는 성향 점수화방법을 논의를 시작한다.

성향점수화는 관측자료나 임의추출에 기반하지 않게 선택된 자료의 묶음을 분석하려고 제안된 기술이다. 기본적인 생각은 개인에게 수행이 임의적으로 배정되지 않았을 때 수행에 대한 인과관계 효과를 평가하기 위해 “수행한” 개인의 그룹과 “수행하지 않은” 개인의 그룹의 결과를 비교해보는 것이다. 성향 점수화는 조건부 확률이다. 개체들은 모형안에서 수행되었고 보조변수의 모음이 주어진다. 만약 연구원들이 성향점수화에서 수행 그룹과 비수행 그룹의 균형을 맞출 수 있다면, 그리고 만약 모델이 맞다면, 편견 없이 그 수행의 인과관계는 추정할 수 있다.

이러한 지원자 표본에 대한 장점이 있으나 관측연구와 조사 간에는 중요한 차이점이 있다. 중요한 차이점 중 하나는 대부분의 조사가 반대의 연구 대상의 집단을 설명하기 위한 것이다. 비록 실험들과 같은 특정한 목표를 위해 조사의 설계를 할 수 있는 반면에 관측연구는 대부분 처리 효과를 결정하는 주요한 요인을 위해 작은 양의 조사에 응답하게 된다.

관측연구를 위해 성향점수화를 사용하는 연구는 조사에 자발적 참여로 이끄는 방법에 대한 모형 설계에 도움이 된다. 이러한 노력은 관측자료 연구에서 잠재적으로 수행 효과에 관련한 모형을 위한 변수들을 확인하는 것이 필요하다. 조사 계획에서 자발적 참여로 계산된 점수성향화 모델의 발전은 더욱 어렵다. 왜냐하면 범위가 단일처리나 결과 측정이 아니기 때문이다. 자발적 참여 변수들과 다중 결과 변수와 관련된 변수들은 편향을 감소하여 모델에 적절히 포함되어야만 한다.

성향점수화는 무응답과 미포함을 보완하기 위해 확률표본을 사용한다. 적합한 보조 변수가 이러한 모형에서 이용가능할 때 성향점수화(또는 다른 무응답 조정 방법)는 상당히 많이 편향을 감소시키는 것으로 나타났다.

다. 그러나 이러한 방법들은 작은 편향을 제거하는데는 효과적이지 않다. 이러한 모형화로 부터 무응답 편향을 줄이지 못하는 두가지 이유는 주로 주요한 변수들을 제외하거나 변수들 사이의 상호작용을 설명하지 못하는 모형오류와 임의 측정 오류 때문이다. 회귀분석과 같이 모형 오류는 예측변수를 모형에서 제외하거나 변수들간의 상호작용을 모형에 포함하지 못할때 추정에서 편향을 갖게 된다. Steiner, Cook, Shadish(2011)는 편향이 임의측정오류에 기인하여 발생하며, 이 결과를 직접적으로 지원자 표본에 적용하였다. 그들은 비록 모형의 특성화가 정확하다 하더라도 보조변수의 측정오류가 처리효과의 추정에서 상당한 편향을 발생시킬 가능성이 있다고 보았다. 그러므로 성향점수화와 또 다른 추정 방법은 모형을 정확하게 규정하는 것 뿐만 아니라, 적은 측정오류를 가진 보조 정보를 이용해야 한다.

또한 확률 표본들에서 사용되는 무응답조정방법은 비확률적 표본이나 지원자 표본에 대해 사용할때 큰 이점을 가진다. 표본조사에서 추출된 모든 개체들이 조사에 참여하도록 권유를 받는다. 그러므로 성향점수화 모형을 모형화하거나 특성화하기 위해서 추출된 개체들의 조사와 관련된 전체 보조변수가 필요하다. 그러나 이러한 작업은 상당히 복잡하다. 예를 들어 비접촉과 거절의 경우는 이외는 다른 관계가 있을 수 있다. Lin과 Schaeffer(1995)는 모형규정을 어렵게 하는 조사접촉과 거절이 반대 방향으로 편향이 발생되는 것을 발견하였다.

확률표본에서 가중치 조정방법의 또 다른 이점은 조사에서 무응답에 관한 많은 연구가 있다는 것이다. Groves와 Couper(1998)는 응답과 추정량의 성질과의 관계를 규명하였다. 이러한 정보는 가중치 조정이나 성향모형을 만드는데 매우 유용할 수 있다.

온라인 지원자 표본의 문제들 중 하나는 목표 인구의 전체 구성원들이 온라인 조사에 동일한 확률로 참여하지 않는다는 것이다. 일부 인구

집단의 경우 온라인에 접근할 수 없기 때문에 조사에 참여할 수 없을 수도 있다. 반면에 조사를 홍보하는 웹페이지에 절대 방문하지 않거나, 조사에 대해 개인의 노출을 피하기 위해 방화벽을 설정한 경우와 같이 다른 구성원에게 접근하지 못할 수도 있다. 또한 다른 구성원들은 웹사이트에 방문할 수 있지만 조사의 무응답문제와 비슷한 여러 가지 이유로 조사 참여에 흥미가 없을 수도 있다. 지금까지 웹조사에서 편향에 대한 다수의 모형화는 단순히 이런 온라인상에 접근이 안되거나 방문을 꺼리는 사람들과 같은 웹사이트의 노출부족에만 초점이 맞춰져 있었다.

비록 할당표집 방법은 확률표본추출을 대체할 수 있는 한 가지 방법이다. 할당표집은 확률표본보다 저렴하고 합리적인 추정치를 제공하기 때문에 할당표집 방법을 활발히 사용하였으며, 여러 나라에서 아직 까지 많이 사용되고 있다. Stephan과 McCarthy(1958)는 실증적인 증거를 검사하고, 할당표집이 비록 할당으로 인한 편향을 상당히 증가시키는 상향이 있지만 편향의 측면에서 확률표본추출보다 지속적으로 나쁘지 않다는 사실을 제시하고 있다. 오늘날 할당표집은 미국에 지방정부에서 여전히 사용된다. 그러나 다수의 할당 표집은 조사 추정치를 산출할때 조정과 같은 처리가 적용되지 않은 경우 통상적인 자료로 받아들여지지 못한다. 미국의 경우 할당표집에 대한 활용이 감소하는 경우는 선거에서 득표율의 차이와 승자 예측에 대해 편향이 발생하고, 이러한 원인이 결과적으로 할당표집에 따른 것이라고 규명되었기 때문이다. 선거여론 조사에서 다수의 편향 발생이 할당 표집에 기인한다고 볼 수 있다.

주요한 할당표집의 단점은 조사자에 의한 할당그룹 내에서 응답자를 선택하는 과정에서 편향이 발생한다는 것이다. 만약 성향 점수화모형과 할당표집을 결합하려 한다면, 응답 설계는 개별 조사자에 의해 통제되도록 응답자의 선택과정을 모형화하여야 한다. 만약 이 선택을 조사자와는 별 관련이 없다면, 확률표본추출과 관련된 이익이나 비용절감은 거의 없다.

온라인 조사에 대한 연구로서 Rivers(2007)은 할당표집과 유사하지만 사례대조연구와 더 유사하게 분석하였다. Rivers는 가급적 확률표본을 사용하여 명부나 추출틀에서 선택한 목표 표본으로 표본 매칭을 제안하였다. 예를 들어 목표 표본이 미국 사회조사나 현재 인구조사 같은 센스 표본에서 임의적으로 선택하고, 각각의 목표표본 개체의 성질들은 웹 패널들의 구성원들과 비교하게 된다. 그러면 웹 패널의 가능한 응답자의 그룹에서 가장 가깝게 대응된 표본에 대해 조사를 수행하도록 하는 것이다. 본질적으로 웹 응답자들은 임의적으로 추출되어 대응된 개체들로 대체가 된다. Rivers의 방법은 면접 표본의 품질을 향상시키기 위해 목표표본과 웹패널 모두를 위한 적절한 크기의 보조정보가 요구된다. 할당표집에서와 같이 웹패널조사에서는 응답자를 선택하는 확률은 계산할 수 없다. 그러나 할당표집과 다른 점은 표본설계와 일치된 개체를 선택하기 위한 기준이 더 세분화되고, 편향을 증가시키는 주관적인 방법을 쓰지 않는다는 점이다. 이 방법의 단점은 웹패널의 구성원을 목표 표본에 연계하는 모형 가정에 너무 의존한다는 것이다. 이러한 방법은 또한 사례대조 연구에서도 주요한 약점이 된다. 만약 결과 변수가 대응그룹에서 동질성을 갖는다면, 이 결과들에 대한 추론은 로버스트해야 한다. 이러한 가정은 할당표본이 편향되지 않도록 하는 동일한 조건이지만, 이 경우 보조정보는 이용가능한 할당 그룹에 대한 비교를 위해 추가적인 동질성을 제공한다.

Rivers의 대응 접근에 방법의 또 다른 측면은 효율적인 표본추출 기술이다. 대응은 좋은 층화이고 상대적으로 정확한 추정을 할 수 있다. 그러나 웹 패널로부터의 표본은 지원자의 표본 결과 값들이 보조변수에 대해 조건부로 목표 모집단의 특성과 동일해야만 목표 모집단에 대해 비편향 추정치를 얻기 위해 사용될 수 있다. 이러한 과정은 대응그룹을 구성하기 위한 강한 보조변수에 좌우되며, 관심변수와 보조변수간의 관계

를 잘 이해해야 가능하다. 이러한 관계들은 모형화에서 반드시 고려되어야 하는 것이다. 2010년 선거에서 Rivers의 방법이 적절히 적용되었고, 투표자들의 성향과 밀접하지 않은 다른 특성치들에 대한 추정에서는 큰 편향이 발생하였다. 따라서 다양한 결과변수들에 대해 이 방법을 적용하는데에는 일정 정도 한계성이 있다.

2. 확률표본추출에 대한 이론(異論)

일부 연구자들은 낮은 응답률을 가진 확률표본 자체가 지원자 표본이라고 주장하고 있다. 그러므로 비확률적 표본을 넘어서 어떤 이점도 가지고 있지 않으며, 낮은 응답률을 갖는 잘 구성된 확률표본은 지원자 표본보다 고품질의 표본(넓은 범위의 추정치에 대해 보다 작은 평균 편향)될 수 있다고 주장한다.

응답률이 표본추출 방법 보다 일반적으로 문제가 더 적을 수 있다는 가설은 쉽게 검증된 것이 아니다. 경험적 연구결과로 부터 단지 광범위하게 성립되는지를 제안할 수 있다. 이러한 가설의 기저는 표본으로 추출된 모든 개체들이 적극적으로 모집되고, 참여하도록 하였기 때문에 훨씬 더 균질한 기반을 가진다는 것이다. 더욱이 큰 규모의 무응답 편향은 일반적으로 인지된 것 보다 실제로 덜 빈번히 발생한다는 조건하에서 나타난다. 예를 들어 Groves, Presser, Dipko(2004)는 무응답 편향을 유발하기위해 계획된 실험을 가정하여 무응답 편향의 성질을 연구하였다. 무응답 편향을 발생시킬 수 있는 상황에 대한 연구에도 불구하고 큰 무응답 편향은 발생하지 않았지만, 그들이 가정했던 것만큼 규칙적이지 않았다. 심지어 무응답 편향이 발생할 때, 조사로 부터 추정치들의 작은 부분에서 큰 편향을 있었다.

심각한 포괄성 문제가 있는 추출틀에서 얻은 확률표본이 같은 이점이

있는지는 확실치 않다. 목표모집단이 상당한 과소포함 편향이 있다면 지원자 표본들의 편향과 여러 가지 측면에서 매우 비슷하다. 과소포함은 무응답 보다 더 은밀하게 추정치에 영향을 준다. Van de Kerckhove et al(2009)는 2007년의 약 40%이나 낮은 응답률로 실시한 RDD조사에서 추정치에 무응답편향이 없다고 언급하였다. 그러나 같은 조사에서 일반 전화가 없는 가정을 제외함으로서 과소포함 편향이 발생하였다. 이때는 포함률이 80% 상회하였다.

과소포함을 가진 표본 조사에서의 편향에 대한 우려는 미래의 표본추출을 위한 행정등록 자료의 활용에 영향을 미칠 수도 있다. 행정등록 자료는 종종 불완전하거나 또는 낡은 자료가 된다. 그리고 전형적으로 자료들이 무작위로 누락되지 않는다. 불완전한 행정등록 자료로부터 표본추출은 과소포함 편향을 일으킬 수 있다. 심지어 행정등록자료 데이터베이스가 좋은 보조 변수를 포함하고 있을때, 편향은 통계학 범주에서 지속적으로 추정되어야 한다. 그 이유는 포함 오류 때문에 표본으로 추출되지 못한 목표 모집단의 단위들에 대해 추론과정에서 쉽게 고려되지 못하기 때문이며, 왜냐하면 결측을 일으키는 매커니즘이 알려져 있지 않고 다면적이기 때문이다. 이러한 이유로 다중추출틀을 활용한 조사가 표본추출의 향후 방향임을 의미한다.

제4절 확률 표본의 향후 전망

표본추출의 역사에 관한 조망은 현재까지의 표본추출방법이 매우 변화하였다는 것과 향후 지속적으로 변화할 것이라는 점이다. 확률표본 추출의 과학적 측면과 비확률 표본추출의 현실적인 측면을 고려하고, 이를 체계화 할 수 있는 통계적 이론의 개발이 요구된다.

두 가지 요소가 표본추출의 역사에서 중요하고, 미래와 관련있을 수도 있다. 하나는 사회에서 과학적 생각의 풍요로움이다. 꼭 표본조사방법이나 통계학에서 오지 않더라도 20세기 초는 과학적 그리고 통계학적 생각의 엄청난 발전의 시기였다. 1934년에 낸 논문으로 표본추출조사의 새로운 패러다임을 만든 Neyman은 표본추출조사에 관해 오직 두가지 기여를 했다. 그러나 그는 그의 많은 다른 기여들과 함께 일반적인 통계학적 이론에 엄청난 영향을 주었다. 이러한 풍요로운 사회환경 속에서 기법들의 성장과 변화는 그럴듯하다.

두 번째 요소는 정보에 대한 사회의 요구이다. 표본조사는 부분적으로는 확대될 수 있다. 왜냐하면 20세기 사회는 광범위한 주제에 관해 알기를 원하고, 이러한 조사에 자원을 투입하고자 하기 때문이다.

이러한 요구에 부합하는 결과를 얻기 위해 하나의 모집단 전체 구성원들을 조사하는 것은 너무 번거롭고 비용이 많이 든다. 만약 확률표본 추출 패러다임이 도입되지 않았다면, 관측 자료를 이용한 추론을 위한 방법이나 절차가 이러한 요구들을 충족할 수 있도록 발전될 수 없을 것이라 짐작할 수 있다.

제5절 정책적 시사점

해외 학자들의 연구에 따르면, 조사통계에서 확률표본 추출의 기반은 지속될 것이지만, 이에 보완적인 방법으로 다양한 비확률 추출기법의 개발이 지속될 것으로 예상하고 있다. 확률추출의 대체는 여러가지 측면에서 비과학적이다. 그러나 목표 모집단의 포괄성과 응답률을 감안할때 그에 대한 보완적인 방법으로 다양한 추정 방법의 개발이 요구된다.

보건 및 복지통계 생산과정에서 발생하는 무응답편향의 문제는 지속

적으로 발생할 것이다. 이러한 측면에서 다양한 추정기법의 개발과 더불어 확률표본추출기법의 보완책을 마련해야 한다. 국내의 가구대상 조사의 경우 표본대체를 허용하지 않을 경우 평균 70% 이상을 상회하기 어렵다. 이러한 응답률 감소현상은 사회적 문제로 돌리기보다는 무응답 편향 해소에 필요한 다양한 기법을 적용하는 것이 보다 현실적인 방법이라 할 것이다. 보건분야의 경우 또한 이와 유사한 현상을 보인다. 또한 복지분야의 경우 노인, 장애인 등 사회적으로 소외된 계층에 대한 조사에서의 무응답 성향은 일반인 대상의 조사와는 다르다. 그러므로 확률표본 추출방법과 함께 자원자를 대상으로 한 조사방식 또한 고려할 수 있을 것이다. 소외계층의 행동양식은 대체로 유사한 경향을 갖기 때문에 이를 보완하기 위한 방법이 될 수 있을 것이다. 물론 속성이 다른 두 집단을 연계할 수 있는 추정 방법도 같이 개발되어야 할 것이다.

한편 조사통계의 대부분은 순수한 조사 결과만을 이용하고 있다. 즉, 고도로 발전된 국내 행정전산망을 활용할 경우 응답률이 낮은 조사에서 이러한 정보가 무응답을 대체 할 수 있을 것이다.

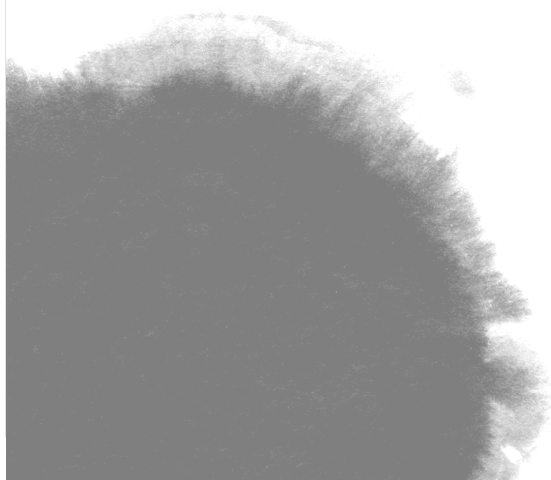
추정과 관련된 다양한 이론들은 이미 국내외적으로 상당히 개발되어 있다. 그러나, 현실적으로 이를 적용하는 데는 한계점이 있다. 특히 통계생산의 주체인 정부의 각 부처에서 이를 적극적으로 활용할 수 있는 기반을 마련해 줄 필요가 있다. 미국의 경우 확률 표본 추출, 비확률 표본 추출 등 다양한 추출기법이 공공부문에서 먼저 개발되고, 연구되어 활용된 사례가 의미하듯이 통계가 필요한 분야에서 먼저 이를 수용하여야 하며, 이와 더불어 많은 정보를 공개하여 발전적인 방향으로의 개선이 요구된다.

결과적으로 확률표본 추출은 기본사항이다. 그러나 이 방법만으로 통계가 생산되기에는 한계가 있다. 점차적으로 다양한 기법을 적용할 수 있는 기반 조성이 필요하며, 이에 대한 정부 부처의 역할이 필요할 것이다.



5장

해외 주요패널조사의 현황과 시사점



제5장 해외 주요 패널조사의 현황과 시사점

제1 절 해외 패널조사의 현황

1. 호주의 HILDA

가. 개요

HILDA(Household, Income and Labour Dynamics in Australia) Survey는 2001년부터 시작된 가구단위의 패널 조사로서 1차년도 조사 대상 가구에 대한 추적조사로 원 가구에 출생 혹은 가구구성의 변화에 따라 진입하는 신규가구를 포함하여 조사가 수행되었다. HILDA 는 Department of Family and Community Services의 Commonwealth Government에서 주도하고 재원을 조달하고, 조사 설계와 관리는 멜번 대학의 경제사회연구소(Melbourne Institute of Applied Economic and Social Research), 교육연구회(Australian Council for Educational Research) 및 가족연구소(Australian Institute of Family Studies) 등에서 담당하며, 자료 수집은 민간 리서치 회사인 AC. Nielsen에서 수행하고 있다.

HILDA의 1차년도 패널은 7,682가구, 19,914명의 표본으로 구성되었고, 각 연도별로 특정 주제를 선정하여 부가조사를 실시하고 있으며, 조사방법은 조사대상 가구의 모든 성인에 대해 면접조사를 실시하며, 패널

가구에 대해서는 지속적으로 추적조사하는 것을 원칙으로 하고 있다.

HILDA의 조사 시점은 1차년도인 경우 2001년 8월~2002년 1월 이었으며, 그 이후 동일한 시점에서 조사를 실시하고 있다.

나. 표본추출

HILDA의 기준 모집단으로서는 1차년도 대상은 호주에 거주하는 모든 사람으로 타국 외교관, 외국인(1년이하 체류자), 외국인 주둔군인, 인구밀도가 희박한 지역주민은 제외하고, 기준모집단 내에서 추출된 확률은 동일하며, 원거주지가 아닌 곳에 머무는 사람의 경우 제외하였다. 표본추출방법은 다단계 층화집락 추출방법으로 1단계에서는 약 200~250가구로 구성된 PSU에서 전국 488개 Census Collection Districts(CDs)를 추출하고, 2단계에서는 각각의 표본 CD에서 그 지역의 기대 응답율과 부재율을 고려하여 22~34거처(dwelling)를 추출한 후 마지막으로 3단계에서는 표본 거처에서 3가구씩을 표본으로 선정하였다. 최종적으로 총 12,252개 거처가 추출되었는데, 이중 803개 거처는 사람이 살지 않거나 비거주용 혹은 외국인으로 이루어져 있어 제외되고, 추출된 거처 중 다가구의 경우 조정된 후의 총 표본은 11,693가구였다.

다. 조사방법

HILDA는 본 조사에 앞서 조사대상에 대한 직접 방문하기 전에 조사 실시에 대한 안내 편지 발송하고, 그 이후 조사과정에서 조사대상자에 대한 면접이 불가능한 경우 지속적으로 재방문을 하며, 조사 이후 감사 편지를 발송하고, 패널가구에 신규로 진입한 가구원에 대해서는 신규조사용 브로셔를 가구 방문 1주일 전에 발송한다. HILDA의 설문지는 총

4종으로 가구설문과 가구원 설문, 개인용 설문과 자기기입식 설문(유치용조사표)으로 구성된다. 개인용 설문은 2차년도 이후 지속가구원에 대한 설문과 신규가구원용 설문으로 구분하여 응답을 받고 있다. 개인설문에 대해 응답 대상자가 모두 완료한 가구에 대해서는 사례비로 50달러를 지급하고, 그렇지 않고 최소한 1명 이상이 응답해준 경우에는 20달러의 사례비를 지급한다. 2차년도 이후 패널가구에 대한 추적 방법은 주소와 호주의 우편번호 레코드를 대조하여 주소를 확인하고, 만일 패널가구에 지속적으로 거주하고 있는 가구원이 있는 경우 가구원에게 추적대상 가구원에 대한 정보를 수집하고, 만일 패널가구가 이사한 경우 해당 주소지에 이사온 사람들로부터 추적대상 가구에 대한 정보를 얻는다. 이와 더불어 1차년도 면접시 얻은 모든 접촉 가능한 전화번호와 주소지의 주변사람들로부터 정보를 얻고, 전화번호를 검색해 표본 가구원의 신규 전화번호에 대한 정보를 얻는다. 또한 1차년도 조사에서 이메일 주소를 받아놓은 경우에는 이메일을 보내거나, 이전 조사에서 정보를 받은 친척이나 친구를 접촉하여 이사 가구에 대한 정보를 얻게 된다. 면대면 조사를 원칙으로 하지만, 방문면접 조사가 불가능한 경우 일부 전화조사를 허용하고 있다.

2. 영국의 BHPS

가. 개요

BHPS(The British Household Panel Survey)는 1991년 9월에 시작한 영국의 가구패널조사로서 사회·경제구조의 변화 추세를 관찰하고 이에 대한 정보를 수집 및 분석하여 21세기 영국의 정치, 사회, 경제 구조를 예측함으로써 각종 정책 수립의 기초 자료로 활용하고자 하는 목적

으로 고안되었다. BHPS는 매년 조사가 실시되며, 2004년 9월 현재 14차년도 조사(wave 14)가 수행되었다. 영국의 경제사회 연구회에서 예산을 지원 받으며, 북아일랜드의 조사는 Eurostat으로부터 예산을 지원받고, 11~15세 조사는 보건교육국에서 예산을 지원한다.

BHPS의 조사 주체는 엑세스(Essex) 대학의 사회경제연구소(Institute for Social and Economic Research: ISER)³⁾ 내의 ESRC Research Centre on Micro-social Change와 UK Longitudinal Studies Centre에서 주관하고, 실사는 민간조사기관인 NOP(National Opinion Poll)Market Research 사⁴⁾에서 실시한다.

나. 표본추출

패널조사 대상 가구의 초기 추출은 2단계집락확률설계(two-stage clustered probability design)와 계통표집(systematic sampling)을 사용하였다. 이때, 표본 단위 선정을 위한 추출틀은 Caledonian 운하(즉 북아일랜드는 제외)의 남쪽 영국에 해당하는 PAF(Postcode Address File, 우편번호주소록)만 포함하도록 하였다. 1991년 1차년도에 표본으로 추출된 8,167가구를 대상으로 면접에 성공한 가구는 5,130가구로 가구구성원 중 16세 이상의 모든 가구원이 포함되도록 조사하였다.

1차년도 이후 표본은 첫째, 원표본구성원(OSMs; Original Sample Members)은 1차년도 조사 가구의 모든 가구원을 포함하며 이와 더불어 조사 이후에 출생한 신생아들도 포함하였다. 둘째, 일시적 표본구성원(TSMs; Temporary Sample Members)은 연구를 시작한 이후에 원표

3) Institute for Social and Economic Research(ISER) 내에는 ㉠ ESRC Research Centre on Micro-social Change, ㉡ UK Longitudinal Studies Centre, ㉢ European Centre for Analysis in the Social Sciences (ECASS) 등의 Centre들이 있음.

4) 북아일랜드에서의 현장 조사는 NISRA에 의해 수행되고 있음.

본구성원(OSMs)과 가구를 형성한 사람을 말하는데, 이들은 원표본구성원과 함께 사는 동안만 조사대상에 포함하도록 하였다. 셋째, 지속 표본구성원(PSMs; Permanent Sample Members)으로 일시적 표본구성원이 원표본구성원과 충분히 강하고 지속적인 유대를 갖는다고 간주된다면 PSMs로 전환하여 조사가 이루어지도록 하였다.

다. 질문방법

자료 수집을 위한 설문 응답 방식은 자기기입식 설문지와 대리 응답 설문지 및 아동 설문지로 구성되며, 자기기입식 질문지는 개인 질문에 응답한 모든 가구원을 대상으로 한다. 한편 대리 질문지는 현장조사 기간 동안 가구 구성원이 부재하거나 혹은 면접을 완료하기에 너무 나이가 많거나 병약한 경우에 그 사람의 정보를 모으기 위해서, 배우자 또는 성인 아동 등 다른 가구 구성원을 통해 작성하도록 하였고 부가적으로 면접의 유연성을 위해 전화조사도 가능하도록 하였다. 마지막으로 아동 설문지는 조사 대상 가구 내에 포함되는 11세~ 15세 사이의 아동에게 아동들의 건강, 가족, 희망과 관심 등에 대해 조사하되 직접 기입하는 방식으로 진행되었다.

라. 조사방법

1991년부터 1998년까지는 종이 질문지를 이용하여 면접자가 대면 면접을 하여 자료를 수집하였고, 9차년도 조사에서 컴퓨터지원 개인면접(CAPI; Computer Assisted Personal Interviewing)으로 전환하였다. 4월~ 5월까지 1,000명을 표본으로 예비조사(pilot survey)가 실시하여 질문 내용들을 검토하고, 본 조사는 9월에서 12월까지의 4개월에 걸쳐 실

시되었다. 응답사례비로는 10파운드 (18000원정도)에 해당하는 상품권 (gift voucher)을 지급하고 있으며, 표본 유지율을 높이기 위해 패널가구원들을 국내거주자에 한해 추적한다. 패널조사원들은 1차년도 조사에 표본의 250 지역을 담당하기 위해 243명의 조사원이 고용되었고, BHPS에 요구되는 성격 때문에 평균 이상 수준의 경험과 능력의 조사원이 고용되었고, 이들 상당수는 수년간 BHPS 조사에 관여하고 있다. 현장 조사 과정에서는 패널가구를 방문하기에 앞서 조사원은 모든 표본 주소로 센터명의로 조사시작 편지를 발송하는데 이때 조사의 목적을 설명한 간단한 ‘리플릿’을 함께 동봉한다. 조사원은 조사시작 편지를 발송한 후 일주일 안에 연락하고, 패널에 참여하는 모든 가구는 나중에 보다 자세한 ‘브로셔’를 받는데 이는 조사에 대한 자세한 정보를 담고 있으며 더불어 조사 참여자에 대한 감사를 담고 있다. 조사원은 사전 편지를 보낸 뒤 기억이 생생한 며칠 이내로 전화를 하되, 여러 가지 질문에 대비하여 조사 관련 내용을 명확히 숙지하고 있어야 하고, 여분의 상품권을 패널에 신규로 진입한 사람이나 이전 조사 후에 받지 못했다는 사람에게 제공한다. 패널 조사의 품질을 보장하기 위하여, 모든 조사원들은 면접 초반에 조사지도원(supervisor)과 동행하여 지도를 받는데, 조사지도원은 2주간의 현장 과정 모니터를 통해 조사원을 평가하고 있다. 조사의 질을 통제하기 위해 완전하게 조사가 완료된 패널 가구 중에서 약 10%에 우편 연락을 하거나 전화로 직접 확인하는 것으로 파악되었다.

3. 미국의 PSID

가. 배경

OEO(Office of Economic Opportunity)가 국민의 경제적 복지수준

향상에 대한 평가를 위해 센서스국에서 조사를 실시하게 되었다. 조사 초기에는 SEO(Survey of Economic Opportunity)라는 명칭으로 1966년에 3만 가구에 대한 조사가 실시되었다. 그 이후 OEO는 미시간대학 부설 조사연구센터(Survey Research Center : SRC)에 SEO 표본으로부터 저소득 가구 약 2,000가구를 부차표본으로 추출하여 계속 조사를 수행하도록 하였다. SRC에서는 빈곤층과 중산층 및 부유층까지도 포함하여 전체 미국인들을 대표할 수 있는 연구로 확장하였고, 새로운 표본 설계 하에서 SRC의 표본으로부터 새로운 표본을 추출하여 기존의 SEO 표본에 추가하고, 원가구를 떠나 신규가구를 생성하거나, 학업으로 원가구로부터 이주한 경우에도 추적 조사하도록 하였다. 1968년부터 조사를 시작한 PSID는 1차년도에 40개주에서 저소득가구 1,872가구를 포함하여 4,802가구에 대해 조사를 완료하였고, 가족들의 확대와 이사로 1991년에는 50개주와 다른 나라까지 지역 확대하여 추적 조사하였다.

나. 개요

PSID는 1968년에 조사가 시작되었으며, 1년 주기로 매년 조사를 실시하고 있다. 1968년 18,000명의 개인이 1988년에는 37,500명으로 늘었으며, 5,000명 이하로 시작한 가구는 라틴계 가구를 제외하고도 현재 7,000가구까지 확대가 된 상태이다. 초기 연구에서는 1968년 표본에 저소득층을 과대표집하였고, 1968년 이후 이민자탈락으로 인한 표본의 대표성을 조정하기 위해 1990년에 라틴계 가구(멕시코계, 쿠바계, 푸에르토리코계) 2,043가구를 추가하였다. 조사방식은 1968년~1972년까지는 면대면 방식을 취했으나, 1973년부터는 비용절감을 위해 전화면접을 전환하여 조사가 실시되었다. 1973년 이후에는 전화가 없거나(매년 약 500명) 특별한 이유로 전화면접이 불가능한 경우에는 조사원이 직접 방

문하여 조사를 실시하였고, 면접시간을 감소시키고, 비용절감을 목적으로 조사표를 간소화 하였다.

〈표 5-1〉 PSID의 조사방법

연도	조사방법
1968~72	직접 면접조사(기장조사)
1973~90	전화조사 및 면접조사
1991~93	CATI ¹⁾ 로 전환
1994~현재	CATI 와 CAPI ²⁾ 로 조사(면접조사는 CAPI로 실시)

주: 1) CATI: 컴퓨터에 미리 입력된 조사대상자들의 전화번호를 이용해 컴퓨터가 임의로 전화를 걸고 전화통화를 통해 들은 응답결과를 면접원이 컴퓨터에 입력하는 방법

2) CAPI: 조사대상자가 컴퓨터를 이용해 직접 응답하는 방법

다. 조사관리

실사감독관 즉 슈퍼바이저와 조사원들이 이 데이터 수집에 참여하며, 실사감독관들이 면접원들을 선발하여 교육시키고 면접과정을 감독, 면접의 질과 완결성, 정확성 등을 검토, 면접원과 응답자 사이에서 발생하는 제반문제를 해결한다. SRC 사무국은 감독관들과 면접원들의 실사 수행에 대한 모의실험을 통하여 실제의 면접이 원활하게 수행되도록 하며, 실사과정에서 감독관들 및 면접원들의 문의 사항에 답변하며, 면접자료들의 배포 및 수거, 감독관들 및 면접원들로부터 접수되는 설문지와 표지들을 기록, 관리하기위한 표본통제관리시스템 유지한다.

PSID 스태프들은 면접원들의 선발 및 교육을 도와주고, 응답자들의 선정 및 면접에 관한 실사 사무국으로부터의 질문에 답하며, 모의 전화면접 수행 도움, 또한 면접원 교육과정을 참관하고 필요시 감독관 및 면접원들과 의견을 교환하며, 실사사무국으로부터의 표본통제관리 시스템에

대한 보고를 수행한다.

1968년의 PSID 표본은 200개 이상의 1차 표본지역(primary sampling area)들에 집락(cluster)을 구성하고 있는데, 이는 데이터 수집의 용이함과 이에 따른 비용절감을 위한 것이었다.

PSID 초기에는 면접원들을 지리적으로 인접한 응답자들에게 할당하였으며, 직접면접은 지금도 이 원칙에 의해 수행하나, 1985년 장거리 전화요금을 절약하기 위해 전화면접원들 재배치하였고, 가능한 한 응답자들에게 고정적인 면접원들을 접속시키는 방향으로 노력하였다.

4. 캐나다의 SLID

가. 개요

1980년경부터 캐나다 통계청은 개인 및 가족의 노동시장 활동에 관한 패널데이터를 수집하고 있다. 1980년대 후반의 노동시장조사(Labor Market Activity Survey: LMAS)가 SLID의 기원이다. LMAS는 1986~87년의 제1패널과 1988-90년의 제2패널로 구성되었으며 이를 발전시킨 것이 SLID이다. SLID 1차 조사는 1993년 시작되었고, 1차 조사에서는 응답자의 부담경감과 면접의 용이성을 위해 인구학적, 개인적 배경에 관한 질문이 주요 항목이었다.

나. 조사방법

SILD는 매년 2회에 걸쳐 인터뷰를 실시하는데, 1월에는 전년도 노동시장 활동에 관해 조사를 실시하고, 다음으로 전년도 세무신고가 끝난 5월에 소득조사 실시하였다. 이때, 가능한 한 동일한 면접원을 배치

〈표 5-2〉 주요 패널간 비교

	PSID	NLSY97	GSOEP	BHPS	SLID
주관	미시건 대학 SRC	오하이오 주립대학 CHRR	DIW	Essex 대학의 ESRC와 ULSC	캐나다 통계청
실사	SRC 실사사무국	사카고 대학 NORC	Infratest(민간)	NOP(민간)	
표본	- 원표본(1968) • SRC표본 2,930가구 • SEO표본 1,872가구 (저소득) - 부가표본: 라티노표본	1996년 12월 31일 기준으로 12~16세 청소년	- West-German 4,528가구(1984) - Foreigner 1,393가구(1984) - East-German 2,179가구(1990) - Immigrant 522가구(1994/95) - Refreshment sample 1,067가구(1998) - Innovation sample 6,052가구(2000) - High Income sample 1,224가구(2002)	- 원표본 Essex 표본 5,000가구 - ECHP 표본 wave7 1,500가구 - 스코틀랜드 표본 wave9 1,500가구 - 웨일즈 표본 wave9 - 북아일랜드 표본 wave11 1,978가구	LES표본에서 추출한 15,000가구 6년마다 표본교체 1차: 1993~ 2000 2차: 1997~ 2003 3차: 2000~ 2007
초기 표본규모	4,802가구 (1968)	8,984명 (1997)	4,528가구 (1984)	5,130가구 (1991) 13,840명	15,000가구 (1993)
현재 표본규모	7,823가구 (2003)	7,883명 (2002)	12,000가구 24,000명(2003)	12,234명 (2001)	-
시작연도	1968년	1997년	1984년	1991년	1993년
사례비	55달러 (57,000원 상당)	20달러(2002) (20,000원 상당)	5~10DM 상당의 선물지급	10파운드 의 선물증서 (18,000원 상당)	없음. - 동기유발 자료제공과 면접원과의 인간적 관계 중심
인터뷰	CATI와 CAPI 병행	CAPI	CAPI와 PAPI 병행	CAPI	CATI

제2절 정책적 시사점

해외 패널 조사를 고찰한 결과 조사 방법들 간의 몇 가지 유사성을 발견하였다. 첫째, 가구 방문 면접과 여타의 조사방법을 함께 혼용하여 자료를 수집하고 있다는 점이다. 특히 컴퓨터의 발전과 더불어 컴퓨터를 이용한 전화조사(CASI)방법이 널리 활용되고 있음에 주목해야 할 것이다. 이는 향후 국내 보건복지분야 패널 조사의 지속성과도 매우 관련이 있기 때문에 이에 대한 연구가 국내에서도 반드시 필요할 것으로 사료된다.

다음으로 패널 조사원의 선발과 관리의 측면에서 각국의 조사환경에 따라 다양하게 운용되고 있다는 점이다. 국내 패널조사의 경우 조사기관에서 일괄적으로 운용하고 있으며, 이들에 대한 관리와 선발과정에서 연구자들의 개입이 거의 이루어지지 않고 있음에 주목할 필요가 있다. 즉, 슈퍼바이저 및 조사원의 선발과정 및 관리단계에서 해외사례의 경우 매우 밀접하게 관여되고 있는 반면, 국내의 경우는 거의 그렇지 못하다는 점이다. 조사원의 조사에 대한 이해도가 조사의 성패를 좌우한다고 볼 때, 보건 및 복지분야의 패널조사에서는 향후 조사원 선발과 관리 방법에 많은 노력이 필요할 것이다.

셋째, 각국의 패널 조사의 경우 가구의 소득과 관련된 자료는 응답자가 개인적인 조사응답과 더불어 국세청 세무자료의 활용을 선택할 수 있도록 한다는 점이다. 소득관련 자료는 응답자의 측면에서 매우 민감한 항목이다. 현재 국내의 다양한 조사에서 가구 또는 개인의 소득자료는 거의 응답자의 회상에 근거하여 수집되는 자료이다. 이러한 측면에서 향후 소득자료의 정확성을 위해 정책적으로 국세청과 협의하여 국세 자료의 활용을 적극 검토할 필요가 있다. 특히 복지분야의 소득자료는 국내 복지정책의 수립과 정책의 지속가능성을 뒷받침 할 수 있는 매우 중요

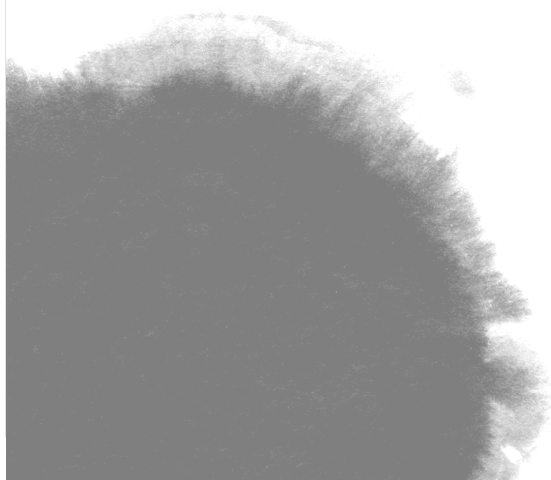
한 자료가 되기 때문에 소득자료 수집과정에서 보다 유연한 자료수집방법을 고려할 필요가 있다.

셋째, 패널자료의 관리와 운용 측면에서 매우 전문적인 인력이 투입되어 운영되고 있는 것이 특징이다. 통계, 사회, 심리, 복지, 경제 등 다양한 분야의 전문가가 연구 그룹을 형성하여 패널자료의 생산과 분석에 많은 노력을 투입하고 있다는 것이다. 조사문항의 개발과 패널의 설계 및 분석, 관리, IT 기술의 적용 등 매우 전문성이 요구되는 분야에 해당 전문가의 책임하에 운용되고 있어 신뢰성 높은 자료가 생산되고 있다. 이러한 측면은 국내 패널조사 관련 연구자들이 신중히 고려해야 할 부분으로 사료된다.

마지막으로 응답자 부담 경감을 위해 패널설계를 유연하게 수립한 것이 특징이다. 연간 단위로 동일한 응답자를 추적조사하고 있기 때문에 한번 패널조사 대상자로 선택되면, 거의 수년간 동일한 문항에 대해 응답해야하는 부담을 갖게 된다. 이러한 응답 부담은 결과적으로 응답자의 타성적인 응답만을 받게 되며, 패널 이탈의 주된 원인이 되기도 한다. 따라서 향후 보건 및 복지분야 패널 조사의 경우 보다 유연한 패널 설계를 통해 응답자 부담을 경감시킬 수 있는 방안을 모색할 필요가 있다.

6장

조사자료의 신뢰도 향상을 위한 방안과 시사점



제6장 조사자료의 신뢰도 향상을 위한 방안⁵⁾과 시사점

제1 절 조사 응답률 산정

1. 조사무응답 산정 배경 및 관련 연구동향

통상적으로 조사를 통한 통계를 작성할 경우 다양한 정보를 이용자에게 제공함으로써 통계의 신뢰성을 향상시킬 수 있다. 조사과정에서 수집된 다양한 보조정보를 “para-data”라고 하며, 이러한 보조정보는 사후 분석에서 매우 유용하게 활용될 수 있다. 최근 조사통계분야서 통계품질의 향상을 위해 이와 같은 보조정보의 수집을 권장하고 있으나, 국내 조사현실은 그렇지 못한 실정이다. 이러한 맥락에서 조사 응답률의 산정과 그 결과의 제시는 이용자 뿐만 아니라, 관련 전문가들의 사후 분석에 매우 유용하게 사용될 수 있는 정보가 될 것이다.

조사무응답(또는 응답)에 관한 정보는 모든 조사과정에 관한 중요한 실행 척도를 제공할 수 있는데, 이는 무응답 (또는 응답)에 관한 척도는 다음과 같은 전반적인 영역에 정보를 제공하게 된다. 즉, 조사기관 (survey agency)의 관점에서는 조사지역, 조사항목, 조사에 활용된 설문지의 개선에 좋은 정보가 될 수 있으며, 통계자료를 이용하는 자료 이용

5) 본 고는 2012년 통계개발원에서 주관한 “국가통계 신뢰도 향상 방안”의 워크샵에서 발표한 내용을 기반으로 재 편집한 내용임.

자(data user)의 측면에서는 단위 및 항목 무응답에 관한 정보를 얻고 이를 통해 가중값 분석이나, 통계자료의 변동 등을 파악할 수 있다. 또한 현장실사관리자(filed manager)측면에서는 조사에 투입된 조사원의 조사수행능력에 관한 정보를 얻음으로서 향후 관련 분야 조사원의 선발과 교육의 참고자료가 될 수 있다.

우선 조사과정의 무응답 (또는 응답) 자료로부터 무응답률(또는 응답률)을 산출할 수 있는데, 이를 위한 기본적인 원칙은 다음과 같다.

- 주어진 척도를 포괄할 수 있는 개념과 정의의 일관성 유지
- “적격(eligible)” 표본단위는 무엇인가?
- “대상외 표본(out-of-scope)”은 어떻게 정의하는가?

미국의 센서스국(통계청)내부에서 기업체(또는 사업체)조사에 관한 전문가 그룹이 결성되어 무응답 관련 연구를 진행한 바 있으며, 이 그룹을 “IGEN98(Interagency Group on Establishment Nonresponse)”으로 명명하였다. 이들은 BLS, Census Bureau, National Agricultural Statistics Service, National Center for Education Statistics 등의 관련 전문가가 참여하여 관련 통계조사의 무응답과 관련된 연구를 수행하였다. 한편 이와 유사한 연구를 수행한 위원회를 파악해 보면, CASRO(Council of American Survey Organizations)(2000)와 OMB(Office of Management and Budget)(2000), APPOR(2011) 등이 있으며, 주로 이들의 연구결과에서는 응답률과 조사완료율에 관한 완벽한 설명을 포함하는 문서와 무응답 척도에 관한 성실하고 객관적인 산출방법을 서술할 것을 권고하고 있다. 또한 영국의 통계청인 ONS(Office for National Statistics, UK)은 주요 항목에 대해 항목무응답 뿐만 아니라 비접촉 및 응답거절 등에 기인한 무응답을 나타내주

는 테이블을 제시할 것을 권고하고, 최근 사회사에서 응답률 추정에 관한 표준화방법 제시(2000)하고 있다.

2. 조사응답률에 관한 정의들

가. 사업체 조사에서의 응답률 정의

1) IGEN98의 응답률 정의

여기서는 IGEN98(Interagency Group on Establishment Nonresponse)에서 제시한 조사응답률을 다루었다. 국내 보건 및 복지분야 조사통계의 경우 의료기관이나, 보건 산업종사 기업체를 대상으로 조사할 경우에 조사응답률을 산정하는데 도움이 될 것이다.

□ 비가중 응답률: 조사에 참여한 적격 단위에 대한 응답단위의 비율을 의미한다.

$$\text{Unweighted Response Rate} = \frac{\# \text{ of responding eligible reporting units}}{\# \text{ of eligible units in survey}}$$

□ 가중 응답률: 적격단위의 총합의 추정치와 응답단위의 총합 추정치의 비율을 의미한다.

$$\text{Weighted Response Rate} = \frac{\text{Total weighed quantity for responding reporting units}}{\text{Total estimated quantity for all eligible units}}$$

비가중 응답률은 기업체(또는 사업체)조사의 경우 대기업의 비율이 상대적으로 적기 때문에 가중응답률을 산정하는 것이 적절하다. 특히 의료

기관을 대상으로 하는 조사의 경우 종합병원과 일반 의원간의 차이가 크기 때문에 가중응답률을 산정하는 것이 필요하다.

2) BLS의 조사응답률 정의

다음으로 미국의 노동 통계국에서 활용하고 있는 응답률 정의에 대해 살펴보기로 한다. 응답률 산정에 필요한 요소가 되는 변수들을 각각 정의하고 있다.

□ 비가중 응답률: 적격단위와 적격성이 결정되지 않은 단위에 대해 응답단위의 비율을 나타낸다.

$$\text{Unweighted Response Rate} = \frac{\# \text{ of responding reporting units}}{\# \text{ of eligible units} + \# \text{ of sample units with eligibility not determined units}} \times 100$$

□ 가중 응답률 : 적격단위와 적격성이 결정되지 않은 단위에 대한 가중 총합에 대해 응답단위 가중총합 비율을 나타낸다.

$$\text{Weighted Response Rate} = \frac{\text{Total weighed quantity for responding units}}{\text{Total estimated quantity for all eligible units or eligibility not determined units}} \times 100$$

○ Initial Data collection Response Rate (초기 자료수집응답률)

- 개별 조사를 위해 초기에 접촉한 기업체의 수에 근거하여 산정한다.
- 표본으로 선정된 기업체에 기반하며, 자료수집요구에 응한 회사가 요구된 모든 자료를 제공하는지 않은지에 따라 조사에 협조한 기업체를 고려한다.

- Update Collection Response Rate (갱신된 수집자료 응답률)
 - 가장 최근에 갱신된 기업체에 대한 응답률로서 갱신기간동안 수집된 자료에 대해 추정에 이용가능한 자료를 제공한 기업체의 비율로 나타낸다. 특히 조사초기에 협조한 기업체가 몇 년간 정기적으로 갱신된 자료를 수집하여 제공하는지에 대한 성과지표를 산출하는데 적용이 가능하다.
- Update Estimation Response Rate(갱신된 추정치들의 응답률)
 - 실제 조사 추정치를 산출하는데 충분한 자료를 제공한 기업체의 비율로서 초기 접촉한 기업체들의 품목수가 분모가 된다.
- Total Survey Response Rate(전체 조사응답률)
 - 분자 항에는 이용가능한 모든 표본단위의 수가 되며, 분모항은 조사범위 내에 있는 모든 표본단위가 된다.

3) Census Bureau의 조사응답률 정의

미국의 인구 통계국에서 월간 소매판매조사(Monthly Retail Sales), 월간 소매재고조사(Monthly Retail Inventories), 월간 도매조사(Monthly wholesale)에서 조사응답률을 아래와 같은 기준에 의해 조사응답률을 산정하고 있다. 이들 3종의 조사는 주로 우편조사와 전화추적조사를 병행하고 있으며, 월간 소매판매조사는 12,000개 사업체를 대상으로 하며, 월간 소매재고 조사는 4000개 업체, 월간 도매조사는 4000개 업체를 대상으로 한다.

- 응답률 1 : $RR1=R/M$

여기서 R : 응답 단위의 수, M : 응답을 얻기 위해 접촉한 단위의

수이다.

응답률 1은 먼저 최종까지 전화접촉 사업체를 포함한 모든 유형의 응답을 포함하며, 분모 및 분자항에 부분응답도 포함한다. 만일 수집된 자료가 행정자료인 경우는 응답률 산정에서 제외되며, 무응답대체가 수행된 경우 분자에는 포함하고, 분모에서는 제외하고, 우편이 반송된 경우는 분자와 분모 모두에서 제외한다. 또한 응답거절은 분자에서는 제외하고, 분모는 포함한다.

이렇게 계산된 응답률은 특정시점에서 수집된 설문지의 비율이 얼마인지를 알 수 있기 때문에 조사관리자에게 유용하지만, 모집단이 왜도된 경우 잘못 해석될 수 있는 단점이 있다.

$$\text{○ 응답률2 : } RR2 = \sum_{i=1}^R W_i t_i / T$$

여기서 W_i 는 무응답조정 전의 i 번째 단위의 설계가중값이고, t_i 는 응답률 산정을 위해 i 번째 단위에 대해 조사된 값, 그리고 T 는 변수 t 에 대해 추정된 총합을 의미한다.

무응답대체된 값은 분자에는 제외하고 분모에는 포함하고, 행정자료를 이용하여 수집된 경우 분자와 분모 모두에서 제외된다. 특히 이 비율은 동일조사에서 변수가 다른 경우 변화될 수 있으며, 분모에 무응답대체된 값이 포함되기 때문에 대체 방법에 따라 다를 수 있다. 한편 대체율은 $1-RR2$ 로 계산이 가능하다.

4) StPES 측도들

StPES은 미국 센서스국내에서 구성된 연구 그룹으로서 다음과 같은 측도들을 제시하였다.

- Check in Rate = # checked-in cases/ # attempted cases
- Delinquent Rate = 1-Cheched Rate
- Extention Rate = # deliquent cases granted extensions to respond /# attempted cases
- Recieved Rate = # cases with respondent Data / # active cases
- Edit Rate = # edited cases / # actives cases

$$\text{○ Imputaion Rate} = \frac{\sum_{Imp} W_i y_i + \sum_{Adj} (W'_i - W_i) y_i}{\sum_{notAdj} W_i y_i + \sum_{Adj} W'_i y_i}$$

여기서 W_i : 무응답 조정이 않된 i 번째 단위의 추출가중값

W'_i : 무응답 조정이 된 i 번째 단위의 추출가중값

y_i : i 번째 단위의 분석변수

Imp : 무응답 대체된 cases

Adj : 단위무응답 조정에 사용된 sub-domain의 응답자

notAdj : 무응답 조정 가중값을 사용한 sub-domain에 속하지 않은 active case.

$$\text{○ Response Rate} = \frac{s}{e+u} \times 100 = \frac{s}{(n-a-i)} \times 100$$

$$\text{○ Quantity Response Rate} = \sum_{i=1}^S W_i t_i / T \times 100$$

$$\text{○ Total Quantity Response Rate} = \sum_{i=1}^{(S+A)} W_i t_i / T \times 100$$

여기서 S : 자료수집에 성공한 적격 케이스 집합

s : S 에 있는 단위의 수

e : 자료수집을 시도한 적격 단위의 수(자료수집에 실패하거나 목표모집단의 원소인지에 관한 정보가 없는)

u : 적격성 여부가 알려져 있지 않은 단위의 수

n : 표본 또는 센서스에 있는 단위의 수

A : 응답자로부터 자료를 수집하기로 한 목표모집단에 속한 단위의 집합

a : 거절 등의 사유로 자료수집하지 못한 목표모집단에 속한 단위의 수

i : 비적격 단위의 수

W_i : i 번째 단위의 추출가중값

t_i : i 번째 단위의 주요 변수의 값

T : 변수 t 의 총합 추정값

나. 가구 조사(IHSNG: Interagency Household Survey Nonresponse Group)

가구조사에서 무응답을 산정하고, 표준측도를 연구하기위해 미국의 센서스 국을 중심으로 구성된 연구 그룹인 IHSNG에서는 Current Population Survey(CPS), Consumer Expenditure Quarterly(QED), Consumer Expenditure Diary(CED), National Crime Victimization Survey(NCVS), Survey of Income and Program Participation(SIPP), National Health Interview Survey(NHIS) 조사에 관한 무응답 척도를 제시하였다.

1) 단위무응답 척도에 대한 주요 용어

IHSNG에서는 가구 조사에서 발생하는 단위 무응답을 정의하기위해 필요한 주요 변수들을 다음과 같이 정의 하였다.

- 초기 면접응답률: The Initial Interview Response Rate(IIRR)
- 초기 비면접률: The Initial Noninterview Rate(INR)
- 초기 언어문제로 면접하지 못한 비율: The Initial Noninterview Language Problem Rate(INLR)
- 초기 일시 부재로 인해 면접하지 못한 비율: Initial Noninterview Temporary Absence Rate(INTAR)
- 초기 가구 부재로 인해 면접하지 못한 비율: The Initial Noninterview Not at Home Rate(INNHR)
- 초기 응답거절로 인해 면접하지 못한 비율: The Initial Noninterview Refusal Rate(INREFR)
- 초기에 위의 이유 이외에 기타 원인으로 면접하지 못한 비율 :
The Initial Noninterview Other Reasons Rate(INOR)

2) 무응답률 계산에 필요한 주요 변수들

한편 무응답률을 계산하기 위해 필요한 주요 변수들을 다음과 같이 정의하고 있다.

- NEU (# of Eligible Units)
 - 표본에서 면접에 적격일 것으로 생각되는 단위의 수
 - 계산 방법 : 모든 적격 단위들의 합으로 계산
- NINT (# of Interviewed Units)
 - 사전에 정의된 면접이 완료된 모든 단위의 수
 - 계산 방법 : 면접이 완료된 모든 단위의 수
- NENIU (# of Eligible Non-Interviewed Housholds)
 - 언어, 부재, 일시부재, 응답 거절 및 기타 사유로 면접이 진행

되지 못한 적격 단위의 합.

— 계산 방법 : 면접이 수행되지 못한 모든 적격단위의 합.

○ NEUE(# of Eligible non-interviewed Units Excluding refusal, language problems, no one at home, temporarily absent)

— 면접에 적격인 면접단위로서 응답거절, 언어문제, 부재 또는 일시적 부재 등을 제외한 경우

— 계산 방법 : 면접이 수행되지 못한 모든 적격단위의 합(위의 이유를 제외한).

○ NREF (# of Eligible Units Refusing to be Interviewed)

— 응답을 거절한 적격 단위의 수

— 계산 방법 : NREF 로 분류된 단위들의 합

○ NEUNH (# of Eligible Units not interviewed due to No One Home)

— 부재중인 적격 단위의 수

— 계산 방법 : NEUNH 로 분류된 단위들의 합

○ NEUTA (# of Eligible Units not interviewed due to Temporary Absence)

— 일시 부재로 인해 면접이 수행되지 못한 적격 단위의 수

— 계산 방법 : NEUTA 로 분류된 단위들의 합

○ NEUL (# of Eligible Units not interviewed due to Language Problem)

— 언어문제로 면접이 수행되지 못한 적격단위의 수

— 계산 방법 : NEUL 로 분류된 단위들의 합

3) 무응답률 척도

위에서 정의된 변수들을 기반으로 다음과 같은 무응답 척도를 계산할 수 있다.

- IIRR : 초기 응답률 = $(NINT/NEU)*100$
- INR : 초기 비면접률 = $(NENIU/NEU)*100$
- INLR : 언어 문제로 초기 면접이 이루어지지 않은 비율
= $(NEUL/NEU)*100$
- INTAR : 일시적 부재로 인해 초기면접이 이루어지지 않은 비율
= $(NEUTA/NEU)*100$
- INNHR : 부재로 인해 초기면접이 이루어지지 않은 비율
= $(NEUNH/NEU)*100$
- INREFR : 거부로 인해 초기면접이 이루어지지 않은 비율
= $(NREF/NEU)*100$
- INOAR : 기타의 사유로 인해 초기면접이 이루어지지 않은 비율
= $(NEUE/NEU)*100$

4) 특정한 조사 무응답률 척도

또한 특정 조사에서 무응답률 계산에 필요한 주요 변수들은 다음과 같다.

- NEU (# of Eligible Units)
 - 표본에서 면접에 적격일 것으로 생각되는 단위의 수
 - 계산 방법 : 모든 적격 단위들의 합으로 계산
- NINT (# of Interviewed Units)
 - 사전에 정의된 면접이 완료된 모든 단위의 수

- 계산 방법 : 면접이 완료된 모든 단위의 수
- NENIU (# of Eligible Non-Interviewed Housholds)
 - 언어, 부재, 일시부재, 응답 거절 및 기타 사유로 면접이 진행되지 못한 적격 단위의 합.
 - 계산 방법 : 면접이 수행되지 못한 모든 적격단위의 합.
- NINP (# of Interviewed Persons)
 - 현재 웨이브에서 사전에 정해진 규칙에 따라 응답이 완료된 모든 면접자수
 - 계산 방법 : 면접이 완료된 모든 개인의 합
- NPNI (# of Interviewed Persons Not Interviewed)
 - 1회 이상의 면접 시도에서 면접이 이루어지지 않은 사람 수
 - 계산 방법 : 면접이 이루어지지 않은 개인들의 합
- NEULU (# of Eligible Unable to Locate Units)
 - 주소파악이 안되는 지역이거나 PSU에서 보다 10마일 이상 떨어진 지역에 이주하고, 전화조사가 이루어지지 않은 단위의 수
 - 계산 방법 : 위치파악이 안되는 적격단위들의 합

5) 무응답률 척도

개인단위의 조사에서 무응답률에 관한 척도들은 위에서 제시한 변수들을 바탕으로 다음과 같이 정의하였다.

- PNR : 개인 무응답률

$$= (NPNI)/(NINP+NPNI)*100$$
- ULMR : 타지역으로 이주하여 면접 불가능한 비율

$$= (NEULU)/(NINT+NENIU+NEULU)*100$$

- INLR : 언어 문제로 초기 면접이 이루어지지 않은 비율

$$= (NEUL/NEU)*100$$
- SLR: 표본소실율

$$= ((NENIU+NEULU)*\text{승수})/((NENIU+NEULU)*\text{승수}*NINT)$$
- INNHR : 부재로 인해 초기면접이 이루어지지 않은 비율

$$= (NEUNH/NEU)*100$$
- INREFR : 거부로 인해 초기면접이 이루어지지 않은 비율

$$= (NREF/NEU)*100$$
- INOAR : 기타의 사유로 인해 초기면접이 이루어지지 않은 비율

$$= (NEUE/NEU)*100$$

다. BCS 응답률 정의(Bethelhem, Cobben & Schouten, 2011)

덴마크의 통계학자들 Bethelhem, Cobben & Schouten 은 가구조사에서 자기기입식조사와 면접원 기입방식에 따라 응답률 산정이 다르다고 주장하였으며, 이러한 응답상황에 따라 응답률 산정이 적절히 이루어져야 함을 제시하였다. 이들의 연구에 의하면 가구 단위 조사에서 응답률은 다음과 같이 제시하고 있다.

□ 초기표본규모 : $n_I = n_{NC} + n_{OC} + n_{RF} + n_{NA} + n_R$

여기서 n_{NC} : 비접촉 표본수, n_{OC} : 접촉불가능 표본수, n_{RF} : 응답 거절표본수, n_{NA} : 응답불가 표본수, n_R : 응답자 수를 나타낸다.

□ 응답률(1) : $RR = \frac{n_R}{n_E}$

이 응답률의 산정에서 반드시 필요한 변수가 적격응답자(n_E)를 계산해야하는 것이다. 예를 들어 연령기준에 따라 응답 여부가 달라지는 경우 적격응답자와 비적격 응답자를 구분할 수 있다. 한편 비접촉자(no contactor)는 적격비접촉자와 부적격 비접촉자가 포함될 수 있으나, 비접촉자중 적격자의 규모를 모르기 때문에 응답률 산정이 불가능한 경우가 발생한다.

$$\square \text{ 응답률(2) : } RR = \frac{n_R}{n_{NC} + n_{RF} + n_{NA} + n_R}$$

이 응답률은 응답거절자수, 비접촉자수, 접근 불가능한지 등에 대한 정보가 있어야 구할 수 있는 응답률이다. 다음의 응답률(3)과 비교하여 거의 유사한 응답률을 얻을 수 있고, 비교적 산정이 간편하다.

□ 응답률(3) :

$$RR = \frac{n_R}{n_{NC} \left(\frac{n_{RF} + n_{NA} + n_R}{n_{OC} + n_{RF} + n_{NA} + n_R} \right) + (n_{RF} + n_{NA} + n_R)}$$

이 응답률은 비접촉자 중 적격자의 비율이 접촉자 중 적격자의 비율과 동일하다는 가정하에서 산정된 응답률로서 주로 가구당 1명(가구주)을 조사하는 경우 산정이 가능한 응답률이다.

$$\square \text{ 응답률(4) : } RR = \frac{n_R}{n_{NR} + N_R}$$

가구조사에서 면접원에 의한 조사가 아니라, 응답자가 직접 조사표를 작성하는 우편조사 및 웹조사와 같은 자기기입식 조사에서의 응답률(적격 응답에 대한 통제 불가능) 산정에 적합한 산식이다.

라. OMB 응답률 정의

OMB의 조사관련 매뉴얼에 제시되어 있는 응답율은 다음과 같다. 주로 가구 조사에서 적용가능한 응답률이며, 특별히 종단면조사에 대한 응답률 산정식을 제시하고 있는 것이 특징이다.

$$\square \text{ 응답률(1)} : RRU = \frac{C}{C + R + NC + O + e(U)}$$

여기서 C : 완전하게 설문지에 응답하거나 충분히 부분적으로 응답한 케이스 수이며, R : 응답을 거절한 케이스 수, NC : 적격자 중에서 접촉하지 못한 표본단위의 수, O : 거절 이외의 이유로 응답하지 않은 적격 표본단위의 수, U : 적격성을 알 수 없거나 완전하지 않은 표본단위의 수, e : 적격성을 알 수 없는 표본단위들이 적격으로 추정된 비율을 나타낸다.

$$\square \text{ 가중 응답률(1)} : RRW = \frac{\sum w_i C_i}{\sum w_i (C_i + R_i + NC_i + O_i + e(U_i))}$$

여기서 $C_i = 1$, 만일 i 번째 단위가 완전하게 설문지에 응답한 경우 (또는 충분히 부분적으로 응답한 경우); $= 0$, 그외; $R_i = 1$, 만일 i 번째 단위가 거절한 경우; $= 0$, 그외; $NC_i = 1$, 만일 i 번째 단위가 적격자중에서 접촉하지 못한 경우; $= 0$, 그외; $O_i = 1$, 만일 i 번째 단위가 거절 이외의 이유로 응답하지 않은 경우; $= 0$, 그외; $U_i = 1$, 만일 i 번째 단위가 적격성을 알 수 없거나 완전하지 않은 경우; $= 0$, 그외; e : 적격성을 알 수 없는 표본단위들이 적격으로 추정된 비율, w_i : i 번째 단위의 추출가중값을 나타낸다.

□ 관심 변수 y 에 대한 가중 단위 응답률(2)

$$RRW = \frac{\sum w_i y_i C_i}{\sum w_i y_i (C_i + R_i + NC_i + O_i + e(U_i))}$$

조사변수 중에서 주요분석변수 y 에 대한 응답률을 산출할 경우 추출 확률의 역수인 초기 설계 가중치를 적용하여 응답률을 산출한다. 주요변수의 응답률을 통해 조사의 신뢰성을 판단할 수 있다.

□ 종단 조사에서의 응답률(1)

$$RRO^L = \prod_{k=1}^K \frac{I_k^L}{I_k^1 + R_k^1 + NC_k^1 + e_k(U_k^1)}$$

여기서 K : 분석에 사용된 최종 자료 수집 단계, I^L : 모든 웨이브에 응답한 자료수, R_k^1 : 단계 k 에서 1차웨이브 표본의 거절 수,

$I_k^1 + R_k^1 + NC_k^1 + e_k(U_k^1)$: 1차 웨이브에 들어온 전체 표본수를 나타낸다.

패널조사와 같은 종단 조사의 응답률은 초기에 구축된 패널 자료를 기반으로 지속적으로 패널이 유지되어야 하기 때문에 1차 자료의 수집 과정에서 얻은 정보를 기반으로 k 차 패널의 응답률을 산정하여야 한다. 따라서 위에서 제시된 응답률은 현재 “패널 유지율”로 공표되고 있는 통계를 대체할 수 있다.

제2절 조사자료 에디팅 방안

1. 에디팅의 필요성

조사 자료에 대한 에디팅(editing)은 자료의 오류 수정과 더불어 조사 변수들의 무응답 대체 및 가중치 조정을 모두 포괄한 의미이다. 통계이용자들은 사회, 인구, 산업, 경제, 금융, 정치, 문화 전분야에서 높은 품질의 통계자료를 요구하고 있고, 그에 따라 국가통계기관은 이러한 이용자의 요구에 맞는 통계생산을 위해 지속적인 투자를 경주하고 있다. 대다수의 자료에는 오차(error)를 수반하고 있고, 자료수집단계에서는 응답자의 잘못된 응답, 응답자가 응답을 하지 않는 경우, 설문지에 기재된 자료를 컴퓨터로 변환하는 과정의 오차잠재적인 오류 등을 포함한 다양한 오차가 존재한다. 따라서 이와 같이 수집된 자료에 대한 오류검증(error checking)에는 광범위한 노력이 수반되며, 이러한 검증 및 수정과정을 통칭하여 통계적 자료에디팅(statistical data editing : SDE)이라 한다.

자료를 점검하고 수정하는 과정은 먼저 오류가 있는 레코드와 해당 레코드의 필드에 대한 오류의 위치를 탐색(error localization)하여야 하며, 다음으로 그러한 오류에 대한 대체(imputation)를 통해 오류를 수정하게 된다. 전통적인 조사과정에서 SDE는 전체 조사자료에 대해 쌍방향(interactive activity)으로 모든 데이터 레코드에 대해 점검한다.

이와 관련하여 Granquist(1984), Granquist(1997), Granquist and Kovar(1997), Pannekoek and De Waal(2005) 등은 “신뢰할만한 자료를 얻기 위해 데이터세트에 있는 모든 오류를 제거할 필요는 없다”고 언급하였는데, 즉, 표본을 통해 얻어지는 통계자료의 경우 일정 정도의 오류(표본오차)를 내포하고 있고 그러한 오류들은 일정정도 허용하여도 무방하다고 주장하였다. 즉, 통상적으로 데이터에서 가장 영향력 있는

오류(influential error)를 제거하는 것으로 충분하다는 의미이다.

과거에는 오류를 수정하기 위해 과도한 노력을 기울였으며, 궁극적으로 이러한 노력이 실제 공표자료에 거의 미미한 영향을 미치기 때문에 불필요한 예산의 낭비, 시간의 낭비(자료수집시점과 공표시점이 불필요하게 길어짐)를 초래하는 과대 에디팅 (over-editing)은 불필요 하다. 따라서 에디팅의 효율을 개선하기 위해 선택적에디팅(selective editing), 자동에디팅(automatic editing), 매크로에디팅(macro editing)이 실무현장에서 많이 수행되는 방법이다.

에디팅 적용 규칙은 기본은 하나의 레코드가 일관성을 가지는 지를 결정하기 위해 사용되어야 한다는 것이다. 하드에디팅과 소프트에디팅으로 오류가 발견되지 않는 의심스러운 레코드를 검사하고 처리해야 하며, 이때 모든 에디팅 규칙을 만족하는 일관된 레코드는 별도의 에디팅은 불필요하며, 이상치(outliers)는 의심스러운 값으로 에디팅 과정에서 반드시 점검할 필요가 있다.

이와 같이 에디팅은 자료의 수집 단계에서부터 공표까지 전과정에서 신뢰할만한 통계자료 생산을 위해 반드시 수행되어야 하지만, 전체 자료가 모두 일관성을 갖거나, 오류가 없이 수정한다는 것은 현실적으로 어렵기 때문에 앞에서 언급한 바와 같이 전체 일정을 고려하고, 자료중에서 반드시 검토해야 할 변수만을 지정하여 선택적으로 에디팅을 수행하는것이 바람직 할 것이다.

2. 쌍방향 에디팅(interactive editing)

쌍방향 에디팅은 조사과정에서 오류를 수정함과 동시에 조사표상의 오류를 응답자 재접촉을 통해 수정하고, 전문가가 참여하여 수정하는 전반적인 과정을 의미한다.

에디팅 과정에서 컴퓨터의 사용은 이미 오래 전 부터 지속되어 왔고, 최근에는 이러한 역할이 제한적으로 수행되는데, 오류가 발생한 모든 레코드에 대해 목록을 만들어 해당 설문지를 직접 수정하는 방식으로 에디팅을 수행하고 있다. 에디팅에 참여한 인력의 전문성은 각 주제별로 다를 수 있기 때문에 주제별 전문가가 직접적으로 레코드를 수정하는 것이 바람직하며, 자료의 오류가 발생한 경우 응답지를 재 접촉하여 수정하거나, 응답자와 유사한 응답과 비교하여 수정하는 방법도 고려할 수 있다. 이와 같이 오류를 수정한 다음 컴퓨터에 재입력하여, 다시 한번 에디팅 규칙을 만족하는지 점검하며, 모든 레코드가 에디팅 규칙을 만족할 때 까지 반복적으로 실시한다.

이러한 쌍방향 에디팅은 수작업을 통해 수정되는 동안 일관성 체크가 되지 않기 때문에 결과적으로 수정된 레코드는 지속적으로 한 번이상 에디팅 규칙에 위배될 수 있으며, 시간과 비용이 많이 소요 되는 단점이 있다(Federal Committee on Statistical Methodology, 1990; Granquist and Kover, 1997). 또한 주제별 전문가는 해당분야에 대한 광범위한 지식을 겸비해야 하며, 이러한 지식을 가능한 사용하여 오류를 점검해야 한다. 최근에 블레이즈(Blaise), CSPro가 대표적인 조사처리시스템으로 전문가에 의해 작성된 조사프로그램으로서 조사자료의 입력과정에서 자료를 점검하도록 하고 있다. 이러한 방식을 쌍방향 에디팅 또는 컴퓨터에 의한 에디팅 이라고 하며, 쌍방향 에디팅은 수치자료와 범주형 자료에 모두 적용 가능하다.

따라서 조사관리자는 고품질의 통계자료를 제공하기 위해 쌍방향으로 에디팅 한 자료를 고려해야하지만, 각 레코드가 일단 에디팅 되었다 할지라도 여전히 모든 레코드가 에디팅 되어야 하기 때문에 과대 에디팅의 문제를 초래할 수 있다. 이러한 과대 에디팅을 방지하기 위해 모든 레코드에 대한 에디팅 대신 영향력 오류에 대한 에디팅을 고려한 선택

적 에디팅이 바람직하다.

3. 자료수집단계에서의 에디팅

조사현장에서 직접적으로 에디팅이 가능하도록 하기 위해서는 컴퓨터를 이용하여 직접적인 에디팅이 가능해야 한다. 앞서 언급한 블레이즈 시스템은 에디팅 규칙뿐만 아니라, 라우팅 규칙까지 적용이 가능한 소프트웨어이다. 이 소프트웨어는 CAPI, CATI, CASI, CAWI시스템이 이러한 기능을 수행할 수 있다.

CAPI 와 CATI는 우선 시스템 구축비용이 비싸다는 단점이 있고, 면접과정에서 응답자가 질문에 응답할 수 있어야 하며, 특히 기업체조사와 같이 응답가능자가 다수일 경우 응답자를 구별할 수 없다는 단점이 있어 기업체 조사보다는 주로 가구조사에 많이 활용하고 있고, 현재 국내에서도 다수의 조사기업체에서 이러한 시스템을 도입하여 활용하고 있다.

한편 CASI 와 CAWI는 빠른 자료수집과 자료전송의 보안성이 보장된다는 장점이 있으나, 웹조사에서 대표성 결여와 같은 통계적 문제 수반하고 있다. 통계적으로 고품질의 데이터가 수집될 것처럼 보이지만, 실제로 종이설문보다 품질이 떨어질 수 있고, 응답자가 이미 수립된 에디팅 규칙을 준수하도록 해야 하며, 자동계산이 가능하도록 하여 균형 에디팅을 피할 수 있도록 해야 한다. 즉, 수집된 자료에 대해 에디팅 규칙에 위배되는 자료가 적을수록 고품질의 데이터를 보장할 수 있고, 국가통계기관은 지속적으로 자료수집형태를 복합적으로 구현하려고 하고 있다.

4. 선택적 에디팅(Selective Editing)

선택적(유의적)에디팅은 데이터세트에서 영향력 오류 즉, 공표자료에

지속적으로 영향을 끼치는 오류를 식별하기 위한 여러 가지 방법에 대한 포괄적인 용어로서 Hidioglou and Berthelot(1986), Latouche and Berthelot(1992), Lawrence and McDavitt(1994), Lawrence and McKenzie(2000), Farwell and Raine(2000), Hoogland(2002), Hedlin(2003) 등이 이러한 에디팅 방법을 제안하였다. 선택적 에디팅의 목표는 데이터를 2가지 stream으로 분리(critical and non-critical streams)하는 것이며, 첫째 critical stream은 영향력 오류가 있음직한 레코드들로 구성하며, 직접적 에디팅이 요구되는 데이터 집합이다. 한편 non-critical stream은 영향력 오류들이 없는 레코드들로 구성되며, 에디팅이 불필요하거나 자동에디팅을 수행하여 오류를 점검한다.

이러한 선택적 에디팅은 주로 사업체 규모에 따라 응답 자료의 중요도가 다르기 때문에 사업체 자료로 제한되어 있으나, 점차적으로 사업체 자료에서 선택적 에디팅이 확산되고 있으며, 국가통계기관에서도 점점 활용이 증대되고 있다.

선택적 에디팅 방법은 상식적인 수준으로 매우 간단하며, 가장 기본적인 개념은 점수함수(score function)를 이용하는 것이다(Hidioglou and Berthelot, 1986; Van de Poland Molenarr, 1995). 점수함수를 전역함수와 지역함수로 구분하는데, 선택적 에디팅의 단계는 다음과 같은 순서로 진행된다.

- 1단계: 원시자료에 있는 레코드의 전역점수를 계산하고, 이 점수의 순서에 따라 레코드를 순위를 매긴다.
- 2단계: 여러 개의 비율 (p) 값에 따라 처음 p%만큼의 레코드를 에디팅하고, 나머지는 에디팅 하지 않는다.
- 3단계: 우선순위대로 순위가 부여된 원시자료에 있는 p%만큼의 레코드를 정제된 값으로 대체하면 종료한다.

선택적 에디팅을 위한 접근 방법은 크게 edit-related approach(Hedlin, 2003)와 experimental approach(네덜란드 통계청)으로 구분할 수 있다. 전자는 영향력이 있는 오류들이 에디팅 규칙을 위반한다는 기본개념으로 출발하여 하나의 레코드에 대해 얼마나 많은 에디팅이 실패했는지를 측정하고, 균형에디팅의 경우 관찰된 총합계와 관찰된 요소들의 합계간의 절대편차로 측정된다. 이방법의 장점은 하나의 목적변수에 초점이 있는 것이 아니며, 범주형 변수에 적용가능하다는 것이다.

한편 후자의 방법은 서로 다른 변수들간의 복잡한 상호작용을 검출하기 위한 방법으로서 로지스틱회귀 방법으로 영향력 오류를 포함하고 있는 레코드의 확률을 추정하거나 영향력오류를 포함하고 있는 그 레코드에 있는 특정변수의 확률을 추정한다. 이경우 영향력 오류를 포함한 레코드나 변수는 상호 에디팅이 필요하다고 판단하게 된다. 이때 영향력 오류를 포함한 특정 변수의 확률 π 를 추정하기 위해 이전 주기로 부터 원시자료와 정제된 자료가 포함된 훈련자료(training data set)가 필요하며, 원시자료와 정제된 자료간의 차이가 많이 나면, 큰 확률이 배정되고, 그렇지 않으면 낮은 확률이 해당 레코드에 배정되는 원리를 적용한 방법이다.

5. 자동에디팅(Automatic editing)

사람의 간섭 없이 오로지 컴퓨터에 의해 에디팅 하는 경우를 지칭하며, 자동에디팅은 수작업으로 하는 에디팅의 반대 개념이다. 이 방법은 수치변수와 범주형 변수 모두에 적용가능하며, 통계적 자료 에디팅의 두 가지 단계인 오류체크와 대체가 자동으로 이루어진다. 이때 오류는 체계적 오류(systematic error)와 확률적 오류(statistical error)로 구분되며, 체계적 오류는 응답자에 의해 일관되게 보고된 오류(설문 문항의 잘

못된 이해: 단위 오류)를 말하고, 확률적 오류는 우연적으로 발생한 오류(응답자의 편향오류)를 말한다. 체계적 오류는 이전 값을 비교, 전문가의 직감, 등록자료와 비교하여 탐색이 가능하며, 특히 반올림오류(rounding error)가 대표적인 사례이다. 또한 이러한 체계적 오류는 포함된 변수값이 약간의 변화를 만족하는 균형 에디팅이 실패했는지를 테스트함으로써 탐지가 가능하다. 한편 확률적 오류를 탐색하는 방법은 두 가지로 대별할 수 있는데, 결정적 방법과 통계적 방법으로서 통계적 모델링에 의해 이상치 탐색이나 신경망에 의한 방법이 대표적인 에디팅 방법이다. 결정적 탐색 방법은 어떤 레코드가 에디팅 규칙에 위배되었을 때 오류변수로 검사하는 방법으로서 방법의 적용면에서 투명하고 간단하다는 장점이 있으나, 매우 상세하게 점검해야 하는 단점이 있고, 체계적인 확률오류를 수정하기 위해 개인적 편향이 개입될 수 있다.

확률적 오류에 대한 자동오류검색 문제는 먼저 Freund & Hartley(1967)가 수학적 최적화문제로서 관측자료와 수정된 자료간의 거리의 합을 최소화 하는 문제로 고려하였고, 다음으로 Casado & Valera(1996)에 의해 제안된 식으로서 수정된 데이터는 모든 에디팅 규칙을 만족한다는 제약조건하에서 관측자료와 수정자료간의 2차 함수를 최소화하는 방법을 제안하였다. 세번째는 결측값과 수많은 도너(donor) 레코드를 이용하여 핫데트대체로서 불일치한 레코드에 대해 잠재적 오류값을 대체하는 방법이 제안되기도 하였다.

확률적 오류 검색문제에 대한 가장 유용한 방법으로 Fellegi-Holt(1976) 방법 있으며, 이들은 자동에디팅과 대체의 3가지 원칙을 제시하였다.

첫째, 레코드에 있는 데이터는 가능한 최소한의 항목(필드)의 변화로 모든 에디팅 규칙을 만족해야 한다.

둘째, 데이터파일의 빈도(도수)구조는 가능한 한 원 상태를 유지하여야 한다.

셋째, 대체규칙은 명시적 규정을 갖지 않는 에디팅 규칙으로부터 도출되어야 한다.

대표적인 자동에디팅 소프트웨어 패키지로는 GEIS, SPER, AGGIES, SAS, CherryPi, DIA, SCIA, DISCRETE, CANCEIS, SLICE 등이 있다.

6. 매크로에디팅

매크로에디팅은 개별적인 레코드에 있는 의심스러운 자료를 식별하기 위해 추정치에 대한 잠재적인 영향을 검사하는 기법이다. 매크로에디팅은 선택적에디팅이나 자동에디팅에서 감지하지 못한 오류를 발견할 수 있으며, 마이크로에디팅과 매크로에디팅은 상호 보완적인 관계를 갖는다. 앞서 언급한 선택적에디팅은 에디팅할 자료가 지속적으로 수집되는 동안 SDE과정에서 적용할 수 있는 반면, 매크로에디팅은 자료가 이미 수집된 후 SDE 종반부에 에디팅을 적용하며, 선택적에디팅은 기본적으로 에디팅할 레코드를 개별적으로 분리하여 처리한 반면, 매크로에디팅은 한꺼번에 에디팅할 데이터세트를 처리한다. 또한 선택적 에디팅은 각각의 레코드가 에디팅 해야할지를 점검하는 반면에 매크로에디팅은 전체 데이터세트를 에디팅 해야할지를 점검한다.

매크로에디팅 방법에는 Aggregation 방법(Granquist, 1990 and 1995)과 Distribution 방법 등 두 가지 유형이 있다.

가. Aggregation 방법

이 방법은 자료공표 전에 있을 수 있는 오류를 검토하는 방법으로 『과거 테이블값과 현재 통계값을 비교, 또는 등록자료와 비교하거나, 다른 출처로부터 관련된 값을 비교한다. 만이 검토과정에서 이상치가 관측

되었다면, 마이크로에디팅 절차를 이 값에 영향을 받는 개별 레코드와 필드에 대해 적용한다. 이때 이상치의 검출기준은

$$\left| \frac{Y - \hat{Y}}{Y} \right| > p\%$$

여기서 Y 는 검토할 공표자료의 값이며, $\hat{Y}$ 는 이 공표값에 대한 예측값이고, p 는 특정 분율이다.

만일 위 식이 참이면 즉, Y 가 예측값과 비교하여 $p\%$ 만큼 차이가 있으며, 공표자료 Y 는 마이크로에디팅이 필요하게 된다. 가장영향력이 있는 관측값으로부터 시작해서 개별데이터를 직접적으로 점검하고 수정하며, 가중치를 조정하거나 레코드를 제거한다. 이와 같은 수정이 추정된 모집단 값에 무시할만한 영향을 미칠때 직접적인 에디팅과정을 중단하며, 수정 후 영향을 재추정한다.

나. Distribution 방법

이용 가능한 데이터는 변수의 분포를 특정 지워줄 수 있는 데이터세트이거나 참고 데이터세트이다. 모든 개별값들을 이 분포와 비교하여 전형적으로 위치척도와 크기척도를 비교한다. 이상값을 가진 레코드는 점검대상이 되며, 가능한 수정된다. 이때 활용되는 통계적 방법으로는 탐색적 자료분석(EDA: Exploratory Data Analysis) 방법이다. 이때, 상자그림이나 산포도, 또는 6-D 플롯으로 분포를 비교할 수 있다.

마이크로에디팅에 비해 매크로에디팅의 장점은 마이크로에디팅은 자동 에디팅 또는 상호 에디팅이든 에디팅 규칙이 요구된다는 점이다. 마이크로 에디팅의 경우 에디팅 규칙을 만족하도록 하는 것이 매우 어렵거나와 시간이 많이 소요되며, 데이터들이 전반적으로 수정할 부분이 거의

없거나, 오류가 극히 일부인 경우 마이크로 에디팅의 경우 과대 에디팅이 될 가능성이 있다. 에디팅 규칙을 구체화하지 않는 것이 매크로 에디팅의 장점중의 하나이며, 동시에 위험스러운 면이기도 하다. 매크로 에디팅을 적용하고 에디팅 규칙을 구체화 하지 않은 경우 해당 레코드가 의심스러운 것인지 또는 그렇지 않은지는 주제별 전문가가 결정한다. 자동에디팅과 대체는 결과를 재생산 하는 방법중의 하나이며, 동일한 데이터셋을 에디팅하고, 대체하면 동일한 결과를 얻게 된다. 상호에디팅과 매크로 에디팅은 부분적으로 주관적인 에디팅 방법이며, 부분적으로 전문가의 결정에 좌우되며, 데이터에 다량의 오류가 있을 경우 마이크로 에디팅은 매크로 에디팅 보다 효과적이다. 마이크로 에디팅이 매크로 에디팅에 비해 먼저 이루어지며, 이 경우 마이크로에디팅이 덜 비용적이며, 추정치를 개선하는 효과가 있다.

7. 통계적 에디팅 전략

수집된 자료의 신뢰성을 향상시키기 위해 에디팅할 필요가 있다고 가정할때, SDE 전략은 데이터가 양적자료인지, 명목자료인지에 따라 다르게 적용된다. 수치자료(numerical data)에 대한 SDE 전략은 다음과 같은 단계로 이루어진다.

첫째, 단위오류, 부호오류, 비용과 수익의 바뀜, 반올림오류와 같은 명백한(체계적)오류를 수정한다.

둘째, Critical stream과 non-critical stream으로 레코드를 분리하여 선택적 에디팅을 적용한다.

셋째, 자료를 에디팅 한다. critical stream에 속한 레코드를 상호 에디팅하고, non-critical stream에 속한 레코드는 자동 에디팅 한다.

넷째, 매크로 에디팅으로 공표자료에 타당성을 검증한다.

Granquist(1995)는 SDE의 주요 목적을 다음과 같이 정의하였다.

첫째, 전체 조사 프로세스에서 피드백을 제시하기 위한 오류의 원인을 식별하고, 둘째, 유입 및 유출 자료의 품질에 관한 정보 제공하며, 셋째, 개별 데이터에 있는 영향력 오류와 이상치를 식별하고 처리하고, 필요하다면, 완전하고 일관성 있는 자료를 제공하는 것이다.

데이터에 대한 에디팅 작업은 고품질의 국가통계 생산을 위해 반드시 필요하며, 이용자에 대한 에디팅 정보를 반드시 제공하여 분석 과정의 이해를 돕도록 해야 한다. 통상적으로 에디팅과 가중치 조정 및 대체는 동일한 차원이며, 국가통계 기관의 에디팅 전략은 반드시 명시적으로 보유하고 있어야 하며, 시스템으로 구축되고 관리되어야 한다. 한편 에디팅 작업은 통계생산 절차의 하나로서 그에 따른 적절한 비용과 인력이 투입되어야 가능하다. 결과적으로 조사자료 ⇒ 입력 ⇒ 통계가공 ⇒ 공표의 모든 단계에서 에디팅은 반드시 필수적이라 할 수 있다.

제3절 정책적 시사점

조사자료에 대한 응답률 산정과 에디팅 작업은 하나의 연속된 업무로 고려할 수 있다. 에디팅 과정에서 무응답과 결측치 및 이상치 등에 대한 처리와 이를 위한 다양한 자료가 요구되기 때문이다.

조사통계의 경우 자료의 신뢰도를 향상시키고, 이용자에게 완전한 자료를 제공한다는 측면에서 매우 중요한 업무이다. 현재 보건복지분야의 조사통계들은 각각 분야별로 전문가들이 낮은 수준의 에디팅을 수행하고 있으며, 조사 응답률보다는 목표 표본수에 대한 완료율을 제공하고 있다. 조사 응답률과 완료율은 근본적으로 다르다. 표본대상자로 선정된 단위들의 응답특성을 파악하기 위한 척도가 응답률이며, 조사완료율은 조사

과정에서 무응답이 발생하면, 초기 표본수를 맞추기 위해 대체가 수행되고 따라서 원래 초기에 선정된 표본과는 다른 대체 표본이 자료에 포함됨으로서 기본적인 가정에 위배된다. 따라서 이러한 측면에서 응답률은 조사의 신뢰도뿐만 아니라 조사과정의 엄밀성을 측정하는 척도로 이용된다. 특히 조사에 참여한 조사원들의 업무수행능력을 평가하는 척도로 이용가능하며, 관심변수에 대한 추정치의 정도를 평가하는 척도로서 활용이 가능하다.

한편 에디팅 작업은 조사자료의 일관성을 보장하고, 관심변수에 영향력이 큰 단위를 식별하여 수정하며, 조사과정에서 직접 에디팅이 가능하거나, 최종 자료의 공표전에 전체적인 자료의 일치성 확보를 위해 필요한 과정이다. 에디팅은 단순한 과정부터 복잡한 과정까지 다양한 수준으로 가능하며, 조사 전과정에서 에디팅이 수행될 수 있다.

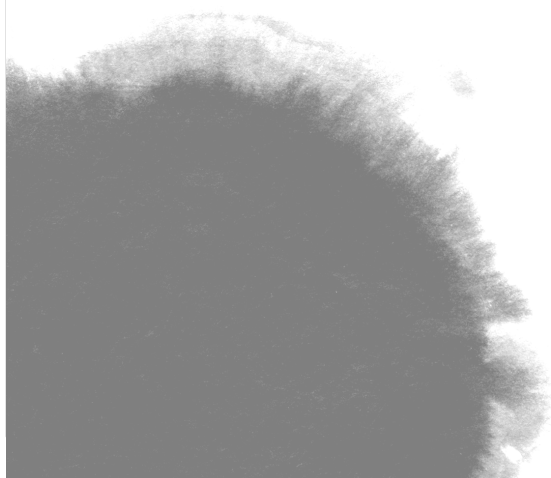
응답률 산정과 에디팅 작업을 위해 반드시 수반되어야 할 것은 조사 전과정에 대한 문서화 작업이다. 조사과정에서 발생하는 사소한 내용이 라하더라도, 향후 자료에 대한 점검시 참고자료로 활용될 수 있기 때문에 조사관리자는 반드시 이러한 내용을 문서화 하도록 해야 한다.

보건 및 복지분야의 조사통계의 경우 이와 같은 작업이 수행되어 자료를 공표하고 있다. 현재보다는 높은 수준의 에디팅 작업이 필요하다. 즉, 각 통계별로 전문가가 참여하는 에디팅 작업이 요구되며, 기본적으로 적용할 에디팅 분야를 지정하여 문서화함으로써 자료의 일관성을 유지하도록 하는 것이 필요하다.

또한 통계전문가의 참여가 반드시 필요하다. 특히 선택적 에디팅 작업을 수행할 경우에는 더욱 그러하다. 보건의료산업관련 조사의 경우 사업체 자료에 대한 에디팅을 위해 선택적 에디팅을 권고하고자 한다.

7장

결론 및 정책적 시사점



제7장 결론 및 정책적 시사점

국민의 삶의 질과 밀접하게 연관된 복지와 보건분야의 다양한 정책과 이를 평가하고, 개선하기위해 필연적으로 요구되는 것이 정책통계이다. 이러한 관점에서 보건 및 복지분야의 다양한 통계 생산 및 관리체계를 점검하고, 개선하는 것이 필요하다. 특히 국민 생활의 바탕이 되는 복지와 보건(또는 건강)에 대한 다양한 척도들은 통계자료로부터 발현되며, 따라서 정확하고 신뢰성 있는 통계 생산과 관리는 필요불가결한 과제일 것이다.

2012년 현재 보건 및 복지분야의 조사통계는 약 40여종에 이르며, 이를 생산하고 관리하는데 많은 노력과 비용이 소요된다. 그러나 투입되는 예산과 인력에 비해 과연 신뢰성 높은 통계자료가 생산되고 있는지는 의문이다. 이러한 측면에서 개별 통계에 대한 품질을 평가하고, 그에 대한 대안을 마련하기에는 물리적인 한계가 있다. 따라서 거시적 측면에서 보건 및 복지분야의 전체 조사통계들의 현황을 분석하고, 이에 따라 자료 수집과정에서 보다 정확한 통계생산을 위한 과제가 무엇이고, 향후 이를 개선할 수 있는 학문적 관점에서의 제언점을 제시함으로써 보건 및 복지분야 뿐만 아니라 타 분야에의 통계생산의 개선에 기여할 수 있는 방안을 제시하고자 하였다.

즉, 조사통계의 관점에서 자료 수집 방법에 대해 예산과 인력 소요 및 향후 지향해야 할 방향성을 모색하였으며, 다음으로 패널조사의 경우

해외 패널의 특징과 그에 따른 시사점들을 도출하였다. 조사통계의 경우 통계의 신뢰성 및 정확성 향상을 위해 응답률 산정 방법과 에디팅 방법을 소개하고, 이를 바탕으로 정확한 통계생산의 가능성을 모색하고자 하였다.

보건 및 복지분야 조사통계의 경우 가구 면접 방식의 자료 수집과정을 보다 다양화 할 필요가 있으며, 그에 따른 에디팅 시스템의 개발과 적절한 응답률 산정을 통해 이용자 측면에서 관련 통계에 대한 정확한 정보를 제공하도록 해야 할 것이다. 가구 또는 개인 대상 조사의 경우 사업체 대상 조사보다 더욱 복잡하고, 그에 따라 오류 발생 가능성이 상대적으로 높기 때문에 에디팅 시스템의 도입은 반드시 필요하며, 모든 조사에서 일괄적이고, 표준적인 에디팅 시스템을 도입하기는 어렵지만, 그럼에도 불구하고, 조사과정상에서 발생할 수 있는 오류를 감소시키기 위한 조사도구상의 에디팅 시스템 도입은 매우 좋은 방안으로 사료된다.

한편 조사 통계의 제한점을 극복하기 위해 현재 다양한 행정 DB가 구축되어 있기 때문에 이를 활용할 수 있는 방안을 적극적으로 모색해야 한다. 행정DB의 활용은 향후 조사통계분야의 당면 과제가 될 것으로 예상된다. DB가 정확한 정보를 제공할 수 있다면, 응답자 부담을 줄이고, 통계산출의 기간과 그에 따른 비용을 대폭 경감시킬 수 있기 때문에 해외 여러 선진국에서는 이에 대한 통계적 이론들이 개발되고 있다. 또한 자료 공유 측면에서 다양한 통계자료의 가공이 가능하기 때문에 보다 다양한 분야의 통계생산이 가능할 것으로 예상된다.

보건 및 복지분야 조사통계는 거의 대부분 확률 표본설계에 따라 표본이 추출되고, 조사된다. 앞서도 언급하였지만, 보건 및 복지분야의 조사대상은 매우 다양하다. 아동, 노인, 장애인 등 소외계층을 대상으로 한 조사통계가 다수를 차지하고 있기 때문에 확률 표본의 측면에서 이들 대상자를 선정하고 조사하는데에는 많은 노력이 요구된다. 따라서 현실

적인 측면에서 확률표본을 대신할 수 있는 표본설계의 도입도 고려하는 방안이 필요하다. 최근 SNS(social network service)가 보편화 되어 인터넷을 통해 매우 다양한 정보들이 서로 교환되고 있다. 또한 방대한 실시간 데이터들(빅데이터: big data)이 망을 통해 수집되고, 이를 적절한 통계적 방법으로 가공하고 분석하여 정보의 가치를 향상시키는 노력들이 지속되고 있다. 즉, 표본조사를 통해 얻지 못하는 광범위한 정보들이 정보망을 통해 상호 교환되고, 이를 통해 통계를 생산함으로써 국민의 요구와 취향을 분석하여 정책수립의 기초자료 생산이 가능한 환경에 이르고 있다. 따라서 이러한 관점에서 향후 정책통계의 방향성에 대한 패러다임을 확대 개편할 필요가 있을 것이다.

마지막으로 이러한 패러다임의 전환에는 반드시 그에 따른 인프라가 구축되어야 한다. 예산과 인력의 지원 그리고 관련 법률의 제정 및 개편을 통해 가능하다. 정확한 통계 또는 신뢰성 높은 통계자료는 어떤 한 분야의 개선을 통해 얻어지는 것이 아니며, 관련 분야의 협업과 활발한 연구를 통해 가능하다. IT 인프라 기반위에서 통계적 분석 모형의 개발과 이를 통한 다양한 정책과제에 대한 연구가 뒷받침 되어야 할 것이다.

참고문헌

- AAPOR Cell Phone Task Force Report. (2010). New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers.
- Baker, R, Stephen Blumberg, Brick, J. M., Couper, M., Courtright, M., Dennis, J. M., Dillman, D., et al. (2010). "Research Synthesis: AAPOR Report on Online Panels." Public Opinion Quarterly 74, pp.711-81.
- Bellhouse, D. (2000). "Survey Sampling Theory over the Twentieth Century and Its Relation to Computing Technology." Survey Methodology 26, pp.11-20.
- Beniger, J. R. (1998). "Survey and Market Research Confront Their Futures on the World Wide Web." Public Opinion Quarterly 62, pp.442-52.
- Blumberg, S., Luke. J. (2010). Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July - December 2009. Hyattsville, MD: U.S. Centers for Disease Control and Prevention, National Center for Health Statistics. <http://www.cdc.gov/nchs>.
- Bowley, A. (1926). "Measurement of the Precision Attained in

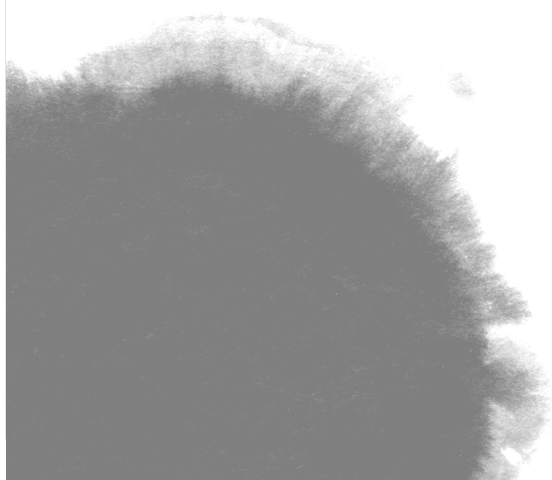
- Sampling." Bulletin of the International Statistical Institute 22: Supplement to Liv. 1, pp.1-62.
- Brick, J. M. (2010). "Allocation in Dual Frame Telephone Surveys with Nonsampling Errors." Paper presented at the Statistical Society of Canada Annual Meeting, Quebec City, Canada.
- Brick, J. M. (2012). "The Future of Survey Sampling", Public Opinion Quarterly, pp.872-888.
- Couper, M. (2000). "Web Surveys: A Review of Issues and Approaches." Public Opinion Quarterly 64, pp.464-94.
- Couper, M. P. (2011) "The Future of Modes of Data Data Collection" Public Opinion Quarterly 75(5), pp.889-908.
- de Leeuw, Edith, D. (1992). Data Quality in Mail, Telephone, and Face-to-Face Surveys. Amsterdam: TT-Publikaties.
- _____. 2005. "'To Mix or Not to Mix Data Collection Modes in Surveys." Journal of Official Statistics 21(2), pp.233-55.
- Deming, W. E. (1944). "On Errors in Surveys." American Sociological Review 9(4), pp.359-69
- Dillman, D. (2007). Mail and Internet Surveys; The Tailored Design Method. 2nd ed. New York: Wiley.
- Dillman, D., Jolene D. S., and Leah M. C. (2009). Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method. New York: Wiley.
- Duncan, J., and William, S. (1978). Revolution in United States Government Statistics, 1926 - 1976. Washington, DC: U.S. Department of Commerce, Office of Federal Statistical Policy and Standards.
- Frankel, M., and Frankel. L. (1987). "Fifty Years of Survey Sampling

- in the United States." *Public Opinion Quarterly* 51(2), pp.127-38.
- Godambe, V. (1966). "A New Approach to Sampling from Finite Populations. II. Distribution-Free Sufficiency." *Journal of the Royal Statistical Society Series B (Methodological)* 28, pp.320-328.
- Green, D., and Alan, G. (2006). "Can Registration-Based Sampling Improve the Accuracy of Midterm Election Forecasts?" *Public Opinion Quarterly*, 70, pp.197-223.
- Groves, R., and Couper, M. (1998). *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Groves, R, Presser, S., and Dipko. S. (2004). "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68, pp.2-31.
- Groves, R. (1979). "Actors and Questions in Telephone and Personal Interview Surveys." *Public Opinion Quarterly* 43(2), pp.190-205.
- Groves, R. (1989). *Survey Errors and Survey Costs*. New York: Wiley.
- Groves, R. M., Fowler Jr., Couper, M., Lepkowski, M., Singer, E., and Tourangeau, R. (2009). *Survey Methodology*. 2nd ed. New York: Wiley.
- Groves, R. M., and Kahn. R. L. (1979). *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- Groves, R. M., and Lepkowski. J. M. (1985). "Dual Frame, Mixed Mode Survey Designs." *Journal of Official Statistics*, 1(3), pp.263-86.

- Holt, D. T, and Smith. (1979). "Post-Stratification." *Journal of the Royal Statistical Society Series A*, 142, pp.33-46.
- Kaier, A. (1895-96). Observations et experiences concernant des de'nombrements representatives. *Bulletin of the International Statistical Institute* 9, pp.176-83.
- Kalton, G., and Brick. J. M. (1995). "Weighting Schemes for Household Surveys." *Survey Methodology* 21(1), pp.33-44.
- Keeter, S. (2006). "The Impact of Cell Phone Noncoverage Bias on Polling in the 2004 Presidential Election." *Public Opinion Quarterly* 70, pp.88-98.
- Kreuter, F., Presser, S., and Tourangeau. R. (2008). "Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity." *Public Opinion Quarterly* 72(5), pp.847-65.
- Kruskal, W., and Mosteller. F. (1980). "Representative Sampling IV: The History of the Concept in Statistics, 1895-1939." *International Statistical Review/Revue Internationale de Statistique*, 48, pp.169-95.
- Pierre. L. (2007). *Indirect Sampling*. New York: Springer.
- Lyberg, L. E., and Kasprzyk. D. (1991). "Data Collection Methods and Measurement Error: An Overview." In *Measurement Errors in Surveys*, edited by Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman New York: Wiley, 237-257.
- Martin, E. A., Childs, J. H., DeMaio, Th, Hill, J. Reiser, C., Gerber, E., Styles, K., and Dillman. D. (2007). *Guidelines for Designing Questionnaires for Administration in Different Modes*.

- Washington, DC: U.S. Census Bureau.
- Martin, P. (2011). "What Makes a Good Mix? Chances and Challenges of Mixed Mode Data Collection in the ESS." Centre for Comparative Social Surveys, City University London. Working Paper No. 02, February.
- Neyman, J. (1934). "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection." *Journal of the Royal Statistical Society* 97, pp.558-625.
- Thompson, S., Ramsey, F., and Seber. G. (1992). "An Adaptive Procedure for Sampling Animal Populations." *Biometrics*, 48, pp.1195-99.
- Thompson, S, and Seber. G. (1996). *Adaptive Sampling*. New York: Wiley.
- Tourangeau, R. (2007). "Incentives, Falling Response Rates, and the Respondent-Researcher Relationship." In *Proceedings of the Ninth Conference on Health Survey Research Methods*. Hyattsville, MD: NCHS, 244-53.
- Waksberg, J. (1978). "Sampling Methods for Random Digit Dialing." *Journal of the American Statistical Association*, 73, pp.40-46.

2017





부록

1. 보건의료분야 조사통계 현황

통계명	통계종류	조사대상	조사방법	조사체계
1. 국민 보건의료 실태조사	일반통계 조사통계	- 대상객체 • 의료기관 • 보건기관 및 약국 - 전체 보건의료기관 - 조사범위 • 전국 보건의료기관 및 약국 • 건강보험, 의료급여 대상자의 2% - 조사단위 • 의료지원 • 의료이용 - 조사지역: 전국	- 보건의료기관의 조사DB시스템 입력 - 시·도, 시·군·구 보건소 담당자 협조조사방법 안내, 조사기간 안내 등) - 건강보험심사평가원, 건강보험공단 등의 자료 제공 협조	의료기관(입력) →시·군·구→ 시·도→한국보건 사회연구원→ 보건복지부(의료 자원과)
2. 국민 구강건강 실태조사	일반통계	- 조사범위 • 초등학교: 1학년(만 6세), 3학년(만 8세), 5학년(만 10세) • 중학교: 1학년(만 12세) • 고등학교 1학년(만 15세)에서 학급을 2차 추출단위로 함. - 조사지역: 전국	- 구강검사, 설문조사 • 구강검진: 구강검진 기준은 세계보건기구(WHO)에서 권장하는 기준을 우리나라의 실정에 맞게 사용함. 구강검진 조사지침에 따라 구강검진 조사자를 교육하고 훈련시킨 후에, 예비조사를 실시함. 예비조사결과 나타난 문제점을 보완하여, 구강 검진 조사자가 처경과 치주탐침기를 이용하여 조사자의 사진을 통해 구강검진을 실시함.	조사원(차의대 교수팀)→ 보건복지부

통계명	통계종류	조사대상	조사방법	조사체계
3. 정신 질환자 실태조사	일반통계 조사통계	- 대상객체: 개인 - 조사범위: 전국 12,849 가구를 대상으로 함. - 조사단위: 만 18세 이상 64세 이하의 가구원 한 명 - 조사지역: 전국	- 표본 - 방문면접조사	조사원→서울대학교 의과대학→보건복지부
4. 청소년 건강행태 온라인조사	일반통계 조사통계	- 대상객체 • 전국 중 · 고등학교 재학생 - 조사범위 • 중 · 고등학교/각 400개교, 전국 800개)에서 학년별로 1개 학급 (75,643명, 2011년 조사기준) - 조사지역: 전국, 시 · 도	익명성 자기기입식 온라인조사 (http://yhs.cdc.go.kr 접속 → "설문참여" 버튼 클릭 → 참여번호로 로그인 → 설문참여)	표본학교→질병관리본부
5. 국민건강 영양조사	지정통계 조사통계	- 조사범위 • 조사대상기간에 표본가구 내에 상주하는 자 중 만 1세 이상인 자 (단, 현역군인 및 전투경찰, 교도소, 소년원 수감자 등 제외) - 조사단위: 가구 및 개인 - 조사지역: 전국	검진조사, 면접조사	- 조사기획, 결과분석 및 공표: 질병관리본부 건강영양조사과, 보건복지부 건강정책과 - 현장조사, 자료처리: 질병관리본부 건강영양조사과
6. 환자조사	일반통계 조사통계	- 대상객체 • 조사일 현재 개설중인 의료기관 및 이들 기관을 이용한 환자 - 조사범위: 전국의 모든 의료기관 및 보건기관 - 조사단위: 사업체 (의료기관)	시·도 및 보건소를 통하여 배부된 조사표를 의료기관의 장 또는 의료기관의 종사자(의무기록실장 등)가 직접 작성 (자계식 조사)	보건복지부→시·도→보건소→의료기관
7. 의약품 소비량 및 판매액 통계	일반통계 조사통계	- 조사범위 • 약국 및 의료기관 각 1,200개(2,400개) - 조사지역: 전국	조사관리시스템과 우편조사 병행 (응답자 기입방식)	약국 및 의료기관→한국보건사회연구원→보건복지부

통계명	통계종류	조사대상	조사방법	조사체계
8. 의료기기, 화장품 제조·유통 실태조사	일반통계 조사통계	- 조사범위 • 의료가기 제조업, 수입업 허가 업체 및 판매업 신고 업체, 화장품 제조업 신고 업체 등 - 조사지역: 전국	면접조사	사업체→한국 보건산업진흥원 →보건복지부
9. 퇴원손상 심층조사	일반통계 조사통계	- 대상객체 • 100병상 이상 일반병원 중 시·도별, 병상규모별 표본추출된 170개 의료기관의 퇴원환자. 단, 단과병원, 요양병원, 노인전문병원, 보훈병원, 국군병원, 재활병원 등은 제외 - 조사범위 • 퇴원요약자료 전산화병원(자체조사): 연간 총 퇴원환자수의 약 9% • 퇴원요약자료 비전산화병원(파견조사): 병상규모에 따라 연간 퇴원환자 차등적 추출 • 표본병원(전산화병원, 비전산화병원)의 표본환자 약 19만건 - 조사단위: 개인 - 조사지역 • 16개 지역의 100병상 이상 일반병원 중 시·도별, 병상규모별로 표본 추출된 170개 의료기관을 설정하여 조사 실시	- 의무기록조사 • 퇴원요약자료 전산화병원에는 표본병원의 의무기록부서에서 의무기록사가 직접 조사하고 비전산화병원은 질병관리본부 의무기록조사 전문연구원을 파견하여 의무기록 검토, 조사지를 작성하고 손상심층조사 및 질병, 수술·처치 및 검사에 대한 질병 및 수술분류번호 부여	표본병원→시·도 및 시·군·구 보건소→ 질병관리본부
10. 지역사회 건강조사	일반통계 조사통계	- 대상객체 • 시·군·구 보건소 당 평균 900명 조사범위 • 표본가구 가구원 중 19세 이상 성인 - 조사단위 • 1차 표본추출단위 : 통·반/리 • 2차 표본추출단위 : 가구 - 조사지역: 전국	- 표본 - 면접조사	- 질병관리본부에서 조사 표준화 - 시·군·구별 보건소에서 참여대학교에 위탁하여 조사수행

통계명	통계종류	조사대상	조사방법	조사체계
11. 전국 장·단기생충 실태조사	일반통계 조사통계	- 대상객체 • 개인 • 통계청 표본추출 기준에 의한 전국민 대상 - 조사범위 • 2000년 인구주택 총 조사의 10% 표본조사구 중 통계청에서 실시하는 각종조사(경제활동 인구조사, 도시가계조사 등)에 사용되고 있는 조사구를 제외한 23,536개 조사구 중 섬지역, 기숙사, 사회복지시설, 농촌 등 153개 조사구를 제외한 23,364개 아파트(8,178개) 및 일반(15,186개) 조사구 - 조사단위 • 표본 가구: 23,364개 아파트 및 일반조사구를 기준으로 조사구 모집단에서 300개 조사구 - 조사지역	- 표본 - 면접조사, 실측조사 • 대변검사: - 셀로판후충도말법, 포화식염수법으로 검사하여 결과를 합산/총란계산법은 가토·카츠법으로 수행 • 셀로테이프항문도말법: 대상자 전원에 대하여 1회 항문주위 도말법으로 총란검사수행 • 총란양성률의 사회의학적 분석: 설문조사에 의한 기생충 총란양성률의 추이분석/설문조사 항목은 (1)주거형태, (2)월수입, (3)식습관, (4)학력, (5)직업	표본설계작성 →기생충 조사항목확정 →검사담당자 →통계분석→ 보건복지부
12. 영아 모성사망 조사	일반통계 조사통계	- 대상객체 • 개인 • 기관 - 조사범위 • 모성사망 사인 동시분석, 영아사망, 사산의 사인 별도 분석 - 조사단위 • 영아사망, 모성사망, 사산아, 의료기관 - 조사지역: 전국	전수	조사대상의료 기관→시·도 보건소, 대한 의무기록협회 →보건복지부

통계명	통계종류	조사대상	조사방법	조사체계
13. 인수 공통전염병 위험군의 감염실태 조사	일반통계 조사통계	- 대상객체 • 전국의 젓소 사육 낙농업 종사자 - 조사범위 • 2010년 경기도 제외한 전국 젓소 사육 낙농업 농장 4,808농가 중 표본추출된 농가에서 낙농업 종사자 500명 이상 조사 - 조사단위 • 젓소 사육 낙농업 농가 - 조사지역: 경기도 제외한 전국	면접요원, 채혈을 위한 연구원 및 연구보조원으로 구성된 현장방문팀 구성 (팀당 4명으로 3개 팀을 구성) 하여 사전연락후 각 지역 보건소나 보건지소를 방문하는 연구대상자를 상대로 개발된 설문지를 이용해 개별 면접방식으로 이루어지고, 연구보조원이 각 지역 보건소나 보건지소를 방문하는 연구대상자에게 검사내용을 설명하고 연구대상자를 상대로 직접 채혈하고 질병관리본부로 검체를 송부하여 검사가 이루어지게 함.	연구용역기관→ 질병관리본부→ 통계청
14. 건강 보험환자 진료비 실태조사	일반통계 조사통계	- 대상객체: • 개인 • 기관 - 조사범위 • 표본 요양기관에 매년 12월내에 외래방문 및 퇴원한 환자 전수 (2010년 조사의 경우 표본기관 850개소, 분석건수 1,702,096건) - 조사단위 - 진료건, 의료기관 - 조사지역: 전국	증화계통추출법에 의한 표본조사	국민건강보험공단 (표본기관 선정)→ 보건복지부, 국민건강 보험공단(공문발송) →요양기관(자료 작성, 이메일송부) →국민건강보험공단 →자료 정비→분석 ※의료기관→국민 건강보험공단(공단 에서 조사와 관련한 조사표 등을 사전에 배포하고 의료기관 에서는 공단 홈페이지의 서식을 다운로드하여 조사 대상기간에 입력 후 이메일 송신

통계명	통계종류	조사대상	조사방법	조사체계
15. 한국 의료패널 조사	일반통계 조사통계	- 대상객체 • 개인 • 가구 - 조사범위 • 전국의 전체 가구 - 조사단위 • 가구별 의료 실태 - 조사지역: 전국	- 표본 - 면접조사	조사원교육→7개 조사팀 순회조사→ 한국보건사회연구원
16. 병원 경영실태 조사	일반통계 조사통계	- 대상객체: 기관 - 조사범위 • 병원급 이상 의료기관(일반병원, 정신병원, 전염성병원, 한방병원, 치과병원 등) 약 1,600개 의료기관 - 조사단위: 의료기관 - 조사지역: 전국	- 전수조사 - 인터넷조사	전국 병원급 이상 의료기관(일반병원, 정신병원, 전염성 병원, 한방병원, 치과병원 등에서 병원경영분석시스템에 직접 입력하고, 입력된 자료는 시스템 분석틀에 의하여 작성됨.
17. 한국인 인체치수 조사	일반통계 조사통계	- 대상객체: 개인 - 조사범위: 전국에 있는 0-90세의 국민 - 조사단위: 개인의 인체치수 - 조사지역: 전국	표본조사	조사원→각 권역별 팀장→한국표준과학연구원→기술표준원
18. 근로 환경조사	일반통계 조사통계	- 조사범위 • 조사전주 수업을 목적으로 1시간 이상 일한 만 15세 이상 취업자	가구방문 면접조사	조사전문기관에서 가구 방문 면접조사
19. 한방 의료이용 및 한약 소비 실태조사	일반통계 조사통계	- 대상객체 • 한방의료이용 실태조사: 표본으로 선정된 한방의료기관의 외래 및 입원 환자 • 한약소비 실태조사: 전국 한방병원, 한의원, 한약방, 한약조제약국, 한약국 - 조사범위 1. 대상범위 및 규모 • 한방의료이용 실태조사: 전국의 한의원 및 한방병원 외래 환자와 입원 환자 5,000명 • 한약소비실태조사: 전국의 한방병원, 한의원, 한약방, 한약조제약국, 한약 2,110개소 - 조사지역: 전국	면접조사	한국보건사회연구원 보건의료연구실→ 보건복지부 한약정책과

통계명	통계종류	조사대상	조사방법	조사체계
20. 선천성 이상아 통계	일반통계 조사통계	- 조사범위 • 출산아 선천성이상 진단 의료기관(전국의 종합병원, 병원, 의원 약 3000개소) - 전수	배포(유치)	의료기관→사업체 (용역업체: 의무기록사협회, 한국보건사회연구원) →보건복지부
21. 전국 출산력 및 가족보건 복지실태 조사	지정통계 조사통계	- 조사범위 • 16,000여 표본가구에 거주하는 15~59세의 기혼부인	직접면접조사	조사대상→조사원→한국보건사회연구원
22. 시민 보건지표 조사	일반통계	- 조사범위 • 서울에 거주하는 15,000세대	면접조사	대상가구→조사원 →용역기관(보건 사회연구원)→서울시

2. 복지분야 조사통계 현황

통계명	통계종류	조사대상	조사방법	조사체계
1. 아동 종합실태 조사	일반통계 조사통계	- 대상객체: 가구 - 조사범위 • 전국 거주 6,900가구의 양육자, 아동 및 청소년 - 조사단위 • 각 가구의 양육자, 아동 및 청소년 - 조사지역 • 16개 시·도(제주도 포함, 기타 섬지역은 제외)	- 표본 - 부모조사: 구조화된 설문지를 이용한 가구방문 면접조사 ※아동 및 청소년 조사에 한하여 자기기입식 병행	조사원→한국갤럽(조사기관)→용역기관(서울대학교 산학협력단)→보건복지가족부
2. 장애인 생활체육 실태조사	일반통계 조사통계	- 대상객체 • 대상객체: 조사 기준시점 현재 등록장애인 • 이용자료: 2009년말 현재 등록장애인(보건복지부) • 대상자료: 10~69세 등록장애인 ▶전국등록장애인(2009년말 현재 2,429,547명)중 10~69세 등록장애인 (2009년말 현재 1,820,943명) 중 1,500명 표본추출 - 조사범위 • 조사범위: 보건복지부 등록장애인중 1,500명 표본 추출 - 조사단위: 1인 - 조사지역: 전국 16개 시·도	- 구조화된 설문지를 이용한 전화면접조사를 기본으로 장애유형에 따른 개별대인 면접 등을 병행 조사 - 조사 방법: 전문 조사원이 장애인 가구에 전화하여 해당 장애인과 통화 (통화가 불가능한 경우는 보호자를 통한 통화)하여 조사를 실시함. - 전화면접조사가 불가능한 경우에는 조사원이 해당 가구에 방문하여 설문지에 대한 응답 내용을 기입하는 방문면접조사를 병행하여 실시함. - 조사 절차: 조사원의 가구 전화면접조사(또는 방문면접조사) ⇒ 지역별 실사 감독원의 관리 및 통제 ⇒ 설문지 집계 ⇒ 보완조사 및 재조사 ⇒ 최종 자료 검증	조사담당자→현장조사(16개 시·도)→조사결과→대한장애인체육회→문화체육관광부
3. 노인 실태조사	일반통계 조사통계	- 조사범위: 3,086개 조사구 약 15,000 (조사구당 평균 4.9명)	면접	조사원교육(연구원)→현재 15개 조사팀(1개팀 지도원 1명, 조사원 3명) 순회조사→보건복지가족부

통계명	통계종류	조사대상	조사방법	조사체계
4. 유망 사회서비스 수요실태 조사	일반통계 조사통계	- 대상객체 • 2011.8.17: 4개 전자바우처 사업 및 4개 유망사회서비스 이용경험자 및 비경험자 - 조사내용: 조사대당 8개 유망 사회서비스유형별 서비스 이용 경험자와 비경험자로 구분하여 조사항목을 구성 - 조사범위: 3,000 표본(8개 서비스별 유의할당 및 서비스 내 이용자/ 비경험자 비율 3:7로 조정) - 조사단위: 명 - 조사지역: 전국	면접조사	조사원→사업체 방문 면접조사→한국보건 사회연구원→ 보건복지부(통계청 조사대행과의 통계조사대행)
5. 보육 실태조사	일반통계 조사통계	- 대상객체 • 개인 • 가구 - 조사범위 • 전국의 초등학교 이하 아동가구 - 조사단위 • 아동의 보육 및 교육시설 - 조사지역: 전국	- 표본 - 면접조사	조사원→ 한국보건사회연구원
6. 장애인 실태조사	지정통계 조사통계	- 대상객체: 개인, 가구 - 조사범위 • 개별조사: 2008년 3월말 기준 보건복지가족부의 장애인 등록 DB에 기재된 등록장애인 • 사회복지시설: 총 1,068개의 사회복지시설(장애인, 아동, 노인, 모부자, 부랑인, 정신요양, 한센병, 결핵)내 등록장애인 - 조사단위: 장애인 개별조사, 사회복지시설 장애인 - 조사지역: 전국	- 표본 - 면접조사, 우편설문 조사	조사원 및 시설→ 한국보건사회연구원 →보건복지부
7. 생명 보험성향 조사	일반통계 조사통계	- 대상객체: 개인, 가구 - 조사범위: 74개 도시의 가구(2,000가구) - 조사단위: 전국 가구 - 조사지역: 전국	- 표본 - 면접조사	조사원→조사전문기관 (용역)→생명보험협회

통계명	통계종류	조사대상	조사방법	조사체계
8. 한국 복지패널 조사	일반통계 조사통계	- 대상객체: 개인 - 조사범위: 국내에 거주하는 모든 개인 - 조사단위: 개인의 복지실태와 욕구 - 조사지역: 전국	- 표본 - 면접조사 - 자기기입식조사 병행	조사원→ 한국보건사회연구원
9. 국민노후 보장패널 조사	일반통계 조사통계	- 대상객체 • 만 50세 이상 중고령자가 있는 가구와 이들 가구에 속한 만 50세 이상의 개인조사대상자 및 그 배우자 - 조사범위 • 2005년 조사시점에 만 50세 이상 중고령자가 있는 가구. 해당 가구에 속한 만 50세 이상의 가구원과 배우자는 개인조사대상자임. 배우자가 만 50세 미만일 경우라도 개인조사대상이 됨. • 약 5,000가구(50세 이상 가구원 10,000명) - 조사단위 • 표본가구: 2005년 1차 본조사에서 원표본으로 구축된 5,110가구 • 표본개인: 2005년 1차 본조사에서 원표본으로 구축된 개인조사대상자 8,689명 - 조사지역: 전국	조사업체의 조사원이 표본가구 및 분가가구를 방문하여 가구조사와 만 50세 이상 개인조사 대상자의 개인조사를 조사표에 직접 입력하는 면접타계방식	조사원→전담 조사기관→ 국민연금공단 (국민연금연구원)

통계명	통계종류	조사대상	조사방법	조사체계
10. 서울시 복지패널 조사	일반통계 조사통계	- 조사범위 • 서울시 3,000가구에 거주하는 만 15세 이상의 남녀 - 조사지역: 서울시	면접조사	전문조사기관용역→서울복지재단
11. 가정 폭력실태 조사	일반통계 조사통계	- 대상객체 • 2005년 현재 대한민국에 거주하는 19세 이상 성인 남녀 - 조사범위 • 가구내 거주하는 19세 이상의 가구주 또는 그 배우자, 또는 성인 가구원 - 조사단위 • 개인(19세 이상의 가구주 또는 그 배우자 또는 성인 가구원) - 조사지역: 전국	- 면접조사 • 조사원이 대상 가구를 방문하여 면접과 동시에 조사내용을 조사표에 입력	연세대학교산학협력단 →여성가족부
12. 기업 및 공공기관의 가족친화 지수 측정	일반통계 조사통계	- 조사범위: 코스피 및 코스닥 상장법인 - 조사단위: 기업체 - 조사지역: 전국	기타	조사표작성, 조사 실시, 기초데이터생성, 조사 분석 및 결과도출에 근거한 가족친화지수 구성, 최종결과 제출
13. 장애인 편의시설 설치현황 조사	일반통계 조사통계	- 대상객체 • 1998년 4월 이후 신·증·개축 및 용도변경 등 건축행위 발생 대상시설(공원, 공공건물 및 공공이용시설, 공동주택, 다만 읍·면·동 및 공공행정기관 청사, 종합병원 등은 '98년 이전 건물이라도 조사대상임) • 편의증진법 제8조 및 동법 시행령 제4조와 별표2에 의하여 위 대상 시설별 의무 설치해야 하는 주출입구 접근로, 장애인전용주차 구역 등 각종 편의시설	- 전수 - 실측조사(교육 받은 조사원(공무원 포함)이 시설주관기관이 선정한 대상 시설을 방문하여 현장 실사를 한 후 조사표에 설치 및 적정설치 여부의 조사 내용을 기록)	전수조사 조사표 개발(보건복지가족부, 한국장애인개발원)→대상시설 선정(시설주관기관)→조사원 교육(시설주관기관, 한국장애인개발원)→전수조사(현장실사) 실시 및 조사결과 보고(시설주관기관, 조사원)→전체 조사결과 보고서 제출(한국장애인개발원)→조사결과 발표(보건복지가족부) ※시설주관기관은 편의시설 설치에 관한 실태조사를 실시함(법제11조).

통계명	통계종류	조사대상	조사방법	조사체계
14. 최저 생계비 계측조사	일반통계 조사통계	- 대상객체: 가구 - 조사범위: 각 시·도의 섬 지역을 제외한 전국(제주도 포함)의 가구 - 조사단위: 가구의 생계비용 - 조사지역: 전국	- 표본 - 면접조사	- 기초조사: 생활실태조사 조사팀→한국보건사회연구원 - 2차 조사: 생활실태조사 및 가계부 기장 (자기기입식방법)→한국보건사회연구원 - 3차 조사: 생활실태조사 조사팀→한국보건사회연구원
15. 고령화 연구패널 조사	일반통계 조사통계	- 대상객체: 개인 - 조사범위: 전국 6,171개 가구(제주도 제외)에 거주하는 45세 이상 중고령자 10,255명 - 조사단위: 45세 이상 중고령자 - 조사지역: 전국	- 표본 - 면접조사, CAPI 조사	조사담당자→TNS Korea →노동연구원→노동부
16. 성폭력 실태조사	일반통계 조사통계	- 개인 - 조사범위: 19세 이상 성인 2,200명(200조사구) - 조사지역: 전국	면접조사	조사대상→연세대 사회복지대학원→여성가족부
17. 아동 패널조사	일반통계 조사통계	- 개인 - 조사범위: 아동 - 조사지역: 전국	면접조사	조사대상→조사기관→보건복지부
18. 차상위 실태조사	일반통계 조사통계	- 개인 및 가구 - 조사범위: 가구 및 개인 - 조사지역: 전국	면접조사	조사대상→조사기관→보건복지부

간행물회원제 안내

▶ 회원에 대한 특전

- 본 연구원이 발행하는 판매용 보고서는 물론 「보건복지포럼」, 「보건사회연구」도 무료로 받아보실 수 있으며 일반 서점에서 구입할 수 없는 비매용 간행물은 실비로 제공합니다.
- 가입기간 중 회비가 인상되는 경우라도 추가 부담이 없습니다.

▶ 회비

- 전체간행물회원 : 120,000원
- 보건분야 간행물회원 : 75,000원
- 사회분야 간행물회원 : 75,000원

▶ 가입방법

- 홈페이지 - 발간자료 - 간행물회원등록을 통해 가입
- 유선 및 이메일을 통해 가입

▶ 회비납부

- 신용카드 결제
- 온라인 입금 : 우리은행(019-219956-01-014) 예금주 : 한국보건사회연구원

▶ 문의처

- (122-705) 서울특별시 은평구 진흥로 235 한국보건사회연구원
간행물 담당자 (Tel: 02-380-8157)

Kihasa 도서 판매처

- | | |
|---|---|
| ■ 한국경제서적(총판) 737-7498 | ■ 교보문고(광화문점) 1544-1900 |
| ■ 영풍문고(종로점) 399-5600 | ■ 서울문고(종로점) 2198-2307 |
| ■ Yes24 http://www.yes24.com | ■ 알라딘 http://www.aladdin.co.kr |

2 · 간행물위원회 안내 및 발간목록



연구보고서 발간목록

발간번호	2011년도 보고서명	연구책임자
연구 2011-01	u-Health 현황과 정책과제	송태민
연구 2011-02	보건의료분야 여건변화에 따른 의료기관의 지출 및 수입구조 분석	조재국
연구 2011-03	천서민 건강관리서비스 확충을 위한 건강관리서비스제도 활성화 방안	이상영
연구 2011-04	약제비 지출의 효율화를 위한 고비용 의약품 관리방안	박실비아
연구 2011-05	식품안전 환경변화에 대응하기 위한 국가 아젠다 개발 등 추진전략 수립	정기혜
연구 2011-06	소바자 중심의 유기식품의 관리체계 및 개선방안 -유기가공식품 표시제 중심으로-	곽노성
연구 2011-07	저소득층 아동비만 및 저체중 문제의 진단과 대응방안	김혜련
연구 2011-08	치료에서 예방으로의 패러다임전환에 따른 건강증진정책 개선방안에 관한 연구	최은진
연구 2011-09	인구집단별 의료이용의 형평성 현황 및 형평성에 영향을 미치는 요인 분석	김동진
연구 2011-10	통일대비 북한 위기상황에 따른 보건복지 대응방안	황나미
연구 2011-11	건강보험 보험료 부담의 공정성 제고방안	신영석
연구 2011-12	노후준비 실패를 반영한 노후소득보장체계 구축방안: 노후소득보장제도와 관련 복지제도간 연관성을 중심으로	윤석명
연구 2011-13	사회보장재정과 재원조달에 관한 연구	최성은
연구 2011-14	보편적 복지와 선별적 복지의 조화적 발전방안에 관한 연구	유근준
연구 2011-15	장애연금제도 발전방안 연구 -장애·장해·장애인 연금간 효과적인 역할정립 중심으로	신화연
연구 2011-16-1	선진국의 아동사례관리체계비교연구: 영국, 미국, 뉴질랜드를 중심으로	김미숙
연구 2011-16-2	호주 사회보장체계 연구	여유진
연구 2011-17-1	정부의 복지재정지출 DB구축방안에 관한 연구(5차년도): 복지수요와 사회복지재정에 관한 연구	고경환
연구 2011-17-2	노인복지서비스 공급방식의 변화와 복지경영 -지방정부를 중심으로-	고경환
연구 2011-17-3	2011 사회예산분석	최성은
연구 2011-17-4	2011 보건복지재정의 정책과제	유근준
연구 2011-17-5	공적연금 재정평가 및 정책현안 분석	윤석명
연구 2011-17-6	사회복지 재정추계 모형개발 연구	원종욱
연구 2011-17-7	건강친화적 재정정책 구축을 위한 연구	정영호
연구 2011-18	공정사회를 위한 천서민정책 개선방안	이태진
연구 2011-19	한국인의 복지의식에 대한 연구: 사회통합을 위한 정책과제	노대명
연구 2011-20	계층구조 및 사회이동성 연구	여유진
연구 2011-21	한국복지패널 연계 질적패널 구축을 위한 기초연구 -저소득층 양적 & 질적 연계 패널조사-	최현수
연구 2011-22	기초생활보장제도 재정평가 및 재정추계 기본모형 개발연구	김태완
연구 2011-23	공공부조 정책 내용과 집행의 상호조응성 분석 -TANF의 배경과 그 집행의 특징-	이현주
연구 2011-24	2011 빈곤연계연보	김문길
연구 2011-25	사회복지제도 운영체계 국제비교 연구: 호주·뉴질랜드·캐나다 영국을 중심으로	강혜규
연구 2011-26	중산층가족의 복지제감도 증진방안 연구	김유경
연구 2011-27	다문화가족아동의 사회적응실태 및 아동복지서비스 지원방안 연구	김미숙
연구 2011-28	지역별 건강수명의 형평성 분석과 정책과제	변용찬

발간번호	2011년도 보고서명	연구책임자
연구 2011-29	장애인 서비스 연계방안 연구	김성희
연구 2011-30	장애인 복지지표를 통해 살펴 본 OECD 국가의 장애인정책 비교 연구	김성희
연구 2011-31	사회적기업의 사회복지서비스 제공 실태 및 운영 구조 연구	강해규
연구 2011-32	저출산·고령사회에서 외국인 유입의 파급효과 분석	이삼식
연구 2011-33	건강지표 산출을 위한 보건기관통합정보시스템 활용 및 제고방안	정영철
연구 2011-34	보건복지통계의 품질관리 표준화 방안 연구	손창균
연구 2011-35	사회복지 통계생산 효율화방안 연구	도세록
연구 2011-36	한국의 보건복지동향 2011	장영식
연구 2011-37-1	출산율예측모형개발	이삼식
연구 2011-37-2	저출산에 대한 한호의 영향과 정책과제	김태홍(외부)
연구 2011-37-3	출산관련 행태 변화에 따른 신생아건강 동향과 정책과제	최정수
연구 2011-37-4	소득계층별 출산·양육 행태 분석 및 정책방안	김은정
연구 2011-37-5	보육의 공공성 강화를 위한 정책방안	백선희(외부)
연구 2011-37-6	일가정양립정책과 보육정책간 연계방안 연구	이삼식
연구 2011-37-7	자방자치단체 저출산 대책의 효율적인 운영방안 연구	박종서
연구 2011-37-8	외국의 이민정책 변천과 사회경제적 영향	임정덕(외부)
연구 2011-37-9	베이비 부머의 삶의 다양성에 관한 연구	정경희
연구 2011-37-10	저출산고령화 시대의 노인인력 활용 패러다임 모색: 연금제도와 고령자 경제활동의 관계를 중심으로	이소정
연구 2011-37-11	노인장기요양보험제도의 형평성 평가	이윤경
연구 2011-37-12	노인장기요양보험의 재정지출 분석 및 정책방안	선우 덕
연구 2011-37-13	예방적 관점에서의 효과적인 노인건강관리서비스의 개발 연구 -M시 종적연구기반(I)	오영희
연구 2011-37-14	고령친화 여가산업 활성화 방안	김수봉
연구 2011-37-15	저출산·고령사회 대응 국민인식 연구	오영희
연구 2011-37-16	저출산대책 관련 국제동향 분석 -스페인·폴란드 판-	이삼식
연구 2011-37-17	선진국 고령사회 대응정책 동향	정경희
연구 2011-37-18	저출산고령사회 대응관련 쟁점연구 -저출산고령사회 정책 사각지대 분석-	이소정
연구 2011-37-19	출산행동의 동향분석을 위한 출산관련 조사자료DB구축	신창우
연구 2011-37-20	결혼이주여성의 성공적 정착과 농촌사회 지속가능한 다문화사회 구축방안 연구	김기홍(외부)
연구 2011-37-21	북한인구의 동태적 및 정태적 특징과 사회경제적 함의	정영철(외부)
연구 2011-37-22	저출산 시대 아동의 안전한 사회 환경 조성방안	이미정(외부)
연구 2011-38	보건의료자원배분의 효율성 증대를 위한 모니터링시스템 구축 및 운영 -2006-2010년 결과 보고서-	오영호
연구 2011-39-1	건강영향평가의 제도화 방안 연구	이상영
연구 2011-39-2	건강도시산업의 건강영향평가 및 기술지원	김동진
연구 2011-39-3	아태 지역 유럽 지역의 건강영향평가와 정책동향	최은진
연구 2011-39-4	건강영향평가 DB 구축	김동진
연구 2011-40-1	기후변화 관련 건강문제 적응대책에 대한 평가체계 개발	김남순
연구 2011-40-2	기후변화에 따른 식품안전사고 위기대응방안 연구	김정선
연구 2011-41-1	아시아 국가의 사회보장제도	홍석표
연구 2011-41-2	한국 보건의료분야 공적개발원조(ODA)의 효율적 운영방안 연구	홍석표

4 • 간행물위원회 안내 및 발간목록

발간번호	2011년도 보고서명	연구책임자
연구 2011-42	취약 위기 및 다문화가족의 예방맞춤형 복지체계 구축 및 통합사례 관리 연구 (2차년도)	김승권
연구 2011-43	친서민정책으로서의 사회서비스일자리 확충 전략 I 아동분야 사회서비스를 중심으로	김미숙
연구 2011-44-1	2011년 한국복지패널 기초분석 보고서	남상호
연구 2011-44-2	2011년 한국복지패널 자료를 통해 본 한국의 사회지표	강신욱
연구 2011-45	2009년 한국의료패널 기초분석보고서(Ⅱ)	정영호
연구 2011-46	2011년 인터넷 건강정보 게이트웨이 시스템 구축 및 운영	송태민
연구 2011-47	2011년 보건복지통계정보시스템구축 및 운영(3년차)	이연희

발간번호	2012년도 보고서명	연구책임자
연구 2012-01	주요국의 사회보장제도(12권)	정기혜
연구 2012-02	보건의료분야 시장개방 이슈와 대응방안 연구 -한미FTA중심으로	김대중
연구 2012-03	초·중·등 고등학교 교과서에 수록된 식품(안전) 내용에 관한 분석 및 개선방안 도출	김정선
연구 2012-04	식품안전분야 연구개발사업 효율화 방안에 관한 연구	곽노성
연구 2012-05	근거중심보건의료에 대한 정책분석과 개선방안	김남순
연구 2012-06	약제비 지출의 목표관리를 위한 예산제의 국가별 비교 연구	박설비아
연구 2012-07	제약산업 구조분석과 발전방향	윤강재
연구 2012-08	건강형평성 강화를 위한 의료서비스 전달체계 개선방안	신호성
연구 2012-09	건강증진서비스 전달체계 다원화 방안 연구	이상영
연구 2012-10	다문화가족 여성과 아동의 건강실태 및 건강서비스 지원방안 연구	김혜련
연구 2012-11	농어촌 지역 주민의 건강증진을 위한 지역사회 자원 조직화 방안	김동진
연구 2012-12	정신건강위험자 관리체계 정립방안에 관한 연구	정진욱
연구 2012-13	식품안전분야 인식조사 개선을 위한 조사시스템 구축방안	정기혜
연구 2012-14	건강보장체계의 New Paradigm 전환에 따른 기반 구축 연구	신영석
연구 2012-15	보험자 내부경쟁을 통한 효율화 방안 연구	김진수
연구 2012-16	국민연금 적정부담 수준에 관한 연구	윤석명
연구 2012-17	건강보험 노인의료비의 효율적 관리방안	신현웅
연구 2012-18	장애인소득보장제도간 급여의 형평성 제고방안 연구	신화연
연구 2012-19	사회정책목표의 실질적 달성을 위한 중장기 복지재정 운용방향	유근춘
연구 2012-20	사회환경에 따른 복지지출 수요와 경제주체별 재정부담능력에 관한 연구	원종욱
연구 2012-21	복지지출 수준에 따른 사회현상과 정책과제	고경환
연구 2012-22	중앙과 지방의 사회복지 재원분담	최성은
연구 2012-23	지방정부의 복지재정과 발전방안에 관한 연구-지방정부의 복지수준과 욕구의 대응성 분석	고경환
연구 2012-24	2012년 사회예산 분석	최성은
연구 2012-25	2012 보건복지재정의 정책과제	유근춘
연구 2012-26	공무원연금 등 직역연금 재정평가와 정책현안 분석: 군인연금과 노르딕 모델을 중심으로	윤석명·신화연
연구 2012-27	OECD 국가의 사회복지지출과 재정건전성 비교연구	원종욱
연구 2012-28	국민기초생활보장제도 개편에 따른 저소득층 소득지원제도 발전방향	강신욱
연구 2012-29	청년층 근로빈곤 실태 및 지원방안 연구	김태환
연구 2012-30	중고령자의 소득자산 분포와 노후빈곤 가능성 분석	남상호

발간번호	2012년도 보고서명	연구책임자
연구 2012-31	현세대 노인의 빈곤 실태 및 소득보장 방안 연구	김미곤
연구 2012-32	빈곤에 대한 대안적 접근: 욕구변주를 고려한 다차원성에 대한 분석	이현주
연구 2012-33	빈곤층 라이프스타일 분석 및 복합적 커뮤니티 케어 제공방안 연구	염주희
연구 2012-34	사회정책과 사회통합의 국가비교: 아시아 국가를 중심으로 I (터키)	이현주
연구 2012-35	인구구조변화가 불평등에 미치는 영향에 대한 연구	김문길
연구 2012-36	한국복지패널 연계 질적연구(2차): 빈곤층의 삶과 탈빈곤 노력을 중심으로	김미곤
연구 2012-37	2012년 빈곤통계연보	김문길
연구 2012-38	사회서비스 바우처사업의 정책효과 분석 연구	강혜규
연구 2012-39	아동복지지출실태 및 적정 아동복지지출 규모 추계	김미숙
연구 2012-40	수요자 중심 장애인복지정책 개발을 위한 연구: 2011년 장애인 실태조사 심층분석	김성희
연구 2012-41	다문화가족의 변화와 사회적 대응방안 연구	김유경
연구 2012-42	장애인의 소득보장과 사회서비스 연계동향 및 정책과제	박수지
연구 2012-43	보건복지부문의 소셜미디어 활용 현황 및 정책과제	정영철
연구 2012-44	한국의 보건복지 동향 2012	장영식
연구 2012-45	의료이용 통계생산 개선에 관한 연구	도세록
연구 2012-46	보건복지분야 조사통계 선진화 방안 연구	손창균
연구 2012-47-1	미래 성장을 위한 저출산부문의 국가책임 강화 방안	이삼식
연구 2012-47-2	국가 사회 정책으로서 통합적인 저출산 정책 추진 방안	신윤정
연구 2012-47-3	중앙정부와지방정부의저출산정책연계방안	이상립
연구 2012-47-4	여성근로자의 노동조건에 따른 출산수준 차이와 정책방안	김원식
연구 2012-47-5	친가족기업 지표개발과 적용방안: 가족친화인증제도의 성과점검과 향후과제	이철선
연구 2012-47-6	한국사회 결혼규범이 저출산에 미치는 영향 분석: 다출산 가정을 중심으로	염주희
연구 2012-47-7	주거행태와 결혼·출산 간 연관성 분석	이삼식
연구 2012-47-8	임신 및 출산을 위한 난임 시술비 지불보상 현황과 정책방향	황나미
연구 2012-47-9	신생아기 저출생체중아 사망영향요인과 관리방안	최정수
연구 2012-47-10	둘째자녀 출산제약 요인분석과 정책방안	정은희
연구 2012-47-11	저출산고령화에 따른 유산상속 동기변화 전망과 정책과제	김원식
연구 2012-47-12	고령화·저출산에 따른 지역별 인구분포와 변화요인 분석과 정책과제	(이현정)
연구 2012-47-13	남북한 통합시 인구가동 전망과 대응과제	이상립
연구 2012-47-14	2011년도 노인실태조사 심층분석	정경희
연구 2012-47-15	100세 시대 건강한 노화의 양상과 정책과제 - M시 종적연구(II)	오영희
연구 2012-47-16	노인 장기요양서비스 전달체계의 평가 및 개선방안: 재가서비스를 중심으로	선우덕
연구 2012-47-17	노인장기요양요구필요도측정방식개발	이윤경
연구 2012-47-18	고령화에 관한 마드리드 국제행동계획(MIPAA) 이행실태 및 평가	정경희
연구 2012-47-19	복지용구사업장규모추계와활성화방안	김대중
연구 2012-47-20	저출산현상의동태적분석을위한지역사례조사	박종서
연구 2012-47-21	백세시대 대응 고령화 지역 연구	이윤경
연구 2012-47-22	저출산대책 관련 국제동향분석: 미국·영국 편	이삼식
연구 2012-47-23	선진국의 고령사회정책 유럽각국의 활기찬고령화(active ageing)정책을 중심으로	선우덕
연구 2012-47-24	저출산·고령사회 대응 국민인식 연구(II)	오영희

6 • 간행물위원회 안내 및 발간목록

발간번호	2012년도 보고서명	연구책임자
연구 2012-47-25	가족구조 변화와 정책적 함의: 1인가구 증가와 생활실태를 중심으로	정경희
연구 2012-47-26	출산력시계열자료 구축 및 분석	신창우·이삼식
연구 2012-47-27	저출산고령화에 따른 사회복지공적 전달체계 개편 방안: 공적전달체계의 수직적편제와 수평적배열의 재구조화	정홍원
연구 2012-47-28	부모에 대한 경제적 의존과 자녀의 만혼화	(이만우)
연구 2012-47-29	저출산에 대응한 영유아 보육·교육 정책 방안	신윤정
연구 2012-47-30	농업인의 노후준비실태와 정책대안	(최경환)
연구 2012-47-31	저출산 고령화 대응 영세자영업자 생활실태 연구	박종서
연구 2012-48	보건의료자원배분의 효율성 증대를 위한 모니터링시스템 구축 및 운영 - 2006년 및 2011년 결과 보고서 -	오영호
연구 2012-49-1	중앙정부의 문화정책에 대한 건강영향평가	이상영
연구 2012-49-2	지방자치단체환경보건정책에 대한 건강영향평가	김동진
연구 2012-49-3	아태지역 및 유럽지역의 건강영향평가 동향 및 정책과제	최은진
연구 2012-49-4	건강증진서비스이용에 대한 개인의 사회적 자본의 영향 연구	최은진
연구 2012-50-1	지역사회 기후변화 관련 건강적응대책 발전방안	김남순
연구 2012-50-2	기후변화 대응을 위한 식품접객업소의 위생관리 개선 방안	김정선
연구 2012-51	아시아 국가의 사회정책 비교연구: 빈곤정책	홍석표
연구 2012-52	취약·위기가족 및 다문화가족의 예방맞춤형 복지체계 구축 및 통합사례 관리 연구 (3차년도)	김승권
연구 2012-53	친서민정책으로서의 사회서비스 일자리확충 전략II: 영유아돌봄 및 초등 방과후서비스를 중심으로	김미숙
연구 2012-54	2012년 전국 출산력 및 가족보건·복지실태조사	김승권
연구 2012-55-1	2012년한국복지패널기초분석:한국의복지실태	최현수
연구 2012-55-2	2012년한국복지패널심층분석:인구집단별생활실태와복지육구의동태분석	남상호
연구 2012-56-1	2012년 한국의료패널 기초분석보고서(I)	정영호
연구 2012-56-2	2012년 한국의료패널을 활용한 의료이용 심층연구	김대중
연구 2012-57	2012년 인터넷 건강정보평가시스템 구축 및 운영	송태민
연구 2012-58	보건복지통계정보시스템 구축 및 운영(4년차)	이연희
연구 2012-59	의료분쟁조정중재제도 활성화를 위한 정책과제와 대책	이상영
연구 2012-60	북한주민의 생활과 보건복지실태	황나미
연구 2012-61	사회보장 재정주계 방법론 개발을 위한 기초연구	원종욱
연구 2012-62	미래 보건복지 방향설정과 정책개발에 관한 연구	신영석
연구 2012-63	보건의료 분야 법령 현황과 주요 과제	윤강재
연구 2012-64	우리나라의 자살급증 원인과 자살예방을 위한 정책과제	이상영
연구 2012-65	복지정책의 지속가능성을 위한 조세·재정정책 정립 방향 - 스웨덴, 프랑스, 영국을 중심으로 -	고경환
연구 2012-66	OECD 보건통계로 본 한국의 보건의료 위상과 성과 및 함의	김혜련
연구 2012-67	보건복지 지표·지수 연구	남상호
연구 2012-68	2012년 지역복지개발평가센터 운영보고서	김승권
협동 2012-1	2012년 사회보건분야 기후변화 취약성 평가 및 적응역량 강화	신호성
협동 2012-2	2012년 비영리법인 제도의 개선방안에 관한 연구	오영호