훈련결과보고서 요약서

훈련자	소속	직급	성명
	통계청 통계개발원 통계방법연구실	통계사무관	최 재 혁
	통계청 조사기획과	통계주사	허 창 수
	통계청 조사시스템관리과	전산주사보	김 상 우
훈련국	스페인, 크로아티아	**훈련기간**	7.7.~7.21.
훈련기관	스페인 폼푸파브라대학교 조사방법론 연구전문센터, 유럽조사연구학회 2019 컨퍼런스	**훈련구분**	단기 (부서맞춤형)
훈련과제명	조사효율화를 위한 파라데이터 수집 및 활용방법 관련 선진사례 연구		

훈련내용

1. 조사방법론 전문 교육프로그램

□ 사회과학에서의 기계학습

○ 파라데이터를 기초적으로 분석하여 활용할 수 있는 것은 조사과정에 발생하는 자료를 토대로 응답자를 어떻게 접촉해야 응답률을 높일 수 있는지 의사결정을 하는 과정으로 기계학습으로 분석할 수 있음

- 선형모형과 최근방 K 이웃방법
- 능형 회귀분석과 라소 방법
- 의사결정나무방법(회귀나무 방법과 분류나무 방법)
- 배깅, 무작위 숲, 부스터 방법

○ 기계학습은 관리되는 방법과 관리되지 않는 방법으로 구분됨

- 관리되는 방법은 기존의 기본방법을 사용
- 관리되지 않는 방법으로는 주성분분석, 군집분석(k-평균 군집분석 방법과 계층적 군집분석 방법)
- 추가적으로 유클리드 거리가 아닌 상관관계 기반으로 분석도 가능

□ 고품질 웹조사의 구현

○ 웹 설문 조사는 시각적 측면이 매우 중요(조사표 설계와 밀접한 관계)

- 시각적 배치가 잘 된 응답자 친화적 설문지는 자기기입식 설문지에서 무응답 감소 효과를 낼 수 있는 매우 중요한 방법
- 각 페이지의 동일한 위치에서 시작하는 웹 질문을 삽입하여 응답 공간을 유지하는 것이 매우 중요

○ 웹 조사는 비확률표본의 일환으로 추정에 대한 정도 확보가 어렵다는 단점을 가지고 있어 전통적인 표본의 정보를 이용하여 웹 조사의 추정을 위한 연구가 필요
- 두 표본 설정에서 웹 반응성향모형의 모수를 추정하기 위한 암묵적 로지스틱 회귀분석방법
- 추정된 웹반응성향에 근거하여 모집단 평균 추정 방법

2. 유럽조사연구학회 2019 컨퍼런스

□ 단기과정

○ 조사방법론의 본질
- 조사방법론에서 활용되는 이론으로 총조사오차 패러다임, 인지적 측면에 대한 응답과정모델, 활용성 이론, 사회교환이론 등이 있음
- 총조사오차 패러다임은 전체 조사진행 과정을 세분화하고, 각 단계에서 조사결과의 정확성에 영향을 미치는 원인에 관심을 두는 이론

○ 스마트폰 : 조사부터 센서까지
- 많은 국가에서 스마트폰 사용자가 많아지고 있어 현재 통계조사는 웹 조사에서 모바일 조사로의 패러다임 변화가 진행 중임
- RDD, Email, QR code, Text-SMS, App, 위치기반 표지(GPS chip) 등 응답자들은 다양한 방식으로 모바일조사에 접근 가능
- 회상오류가 감소하고 응답부담이 줄어들 수 있으며, 고품질의 자료를 더 자주 수집할 수 있다는 장점이 존재

□ 발표세션

○ 스마트폰을 통한 자료수집
- 최근 연구에서는 모바일 앱 자료수집 참여를 늘리는 다양한 방법이 검토
- 스마트폰 센서자료는 응답부담 감소, 정확도 증가 등 유용하지만 응답자

가 자료제공 의향이 필요

○ 면접원과 응답자 간 상호작용

- 개방형 질문에 대한 응답을 폐쇄형 응답범주로 기록하는 과정에서 오류 발생 가능성 있으며, 이 원인 중 하나는 면접원이 긴 응답범주를 확인하기 어렵기 때문임
- 질문 응답하는 과정에서 CAPI보다 CATI에서 더 많은 단어를 사용하며, 응답자는 CATI에서 응답하는 것을 더 어려워 함

○ 파라데이터 분석 및 활용

- 모든 현장 조사 팀원이 동일한 업무윤리를 가지고 있는 것은 아니므로 모니터링 시스템(파라데이터 수집)은 정직한 업무에 보상하고, 신뢰할 수 없는 조사자에게 할당된 업무를 줄어야 함
- GPS-파라데이터 품질 향상과 의심스러운 인터뷰 하위 수준과 연결될 수 있는 CAPI를 사용하면서 면접자 교육에 초점을 맞춰야 함
- 불확실성이 조사 설계에 이슈가 되고 있어 파라데이터와 같은 현장에서 들어오는 데이터를 사용하여 적응적·반응적 조사설계는 이러한 불확실성을 해결하는 방법으로 활용
- 적응적·반응적 조사설계는 비용 및 오차 지표를 개발하고 비용 증가 또는 오차 안정화, 오차 증가에 대해 계획된 개입을 통해 진행

시사점 및 정책제언

□ 국가통계생산은 지속적인 조사환경 변화의 시대에 직면

○ 통계생산 비용과 품질 측면 최적화를 위해 유연한 조사 관리와 설계가 필요한 시점

○ 통계조사 효율화를 위한 현장조사 지침 및 조사방법 개선의 일환으로 조사과정 중에 수집되는 파라데이터 활용 필요성이 강하게 대두되고 있어 파라데이터 수집 및 활용방안 연구를 통해 국가통계 품질, 응답률, 비용 효율화 등 조사효율화를 위한 최신기법 검토가 필요

□ 외국에서는 자동화된 시스템에서 파라데이터를 수집하고 이를 조사환경 개선 및 조사설계 단계에서 활용

○ 파라데이터 수집을 통해 허위·부실조사를 방지하고 조사원 평가 기준 수

립, 조사 기술 향상 개발 및 교육 등에 활용하여 조사효율화에 기여

○ 파라데이터 수집 및 활용에 대한 유럽의 선진사례를 벤치마킹하여 조사환경 개선 및 조사방법 선진화에 적용 필요

□ **현 시점에서는 고품질 및 저비용의 통계 생산을 위하여 응답자 특성을 반영한 적응적·반응적 조사설계 방안 연구 및 도입 검토 필요**

○ 적응적·반응적 조사설계란 통계조사에서 오차와 비용은 상충관계에 있으므로 오차 최소화 목적 하에서 주로 파라데이터와 보조자료를 사용하여 최적 비용 수준의 조사를 설계하는 방법

○ ① 현장조사 지침 개선, ② 맞춤형 응답자 접촉전략 수립, ③ 설계조정방법, 품질 및 비용의 최적화 전략 연구, ④ 적응적·반응적 조사설계 현장 적용 방안 연구, ⑤ 현장 유용성 평가 등이 있으며, 이에 대한 지속적인 연구와 습득이 필요

□ **조사원 및 응답자의 부담감소를 위해 파라데이터를 시의 적절하게 활용하기 위해서 시스템을 통해 바로 수집하고 확인할 수 있는 여건 마련 필요**

○ 업무 부담을 완화하고 효율적인 현장조사 운영을 위해서는 조사원이 개별적으로 결정하기보다 조사원 위치, 업무장소, 최선의 접촉 시간을 고려해서 조사원들에게 배정하는 방법에 대한 고민 필요

○ 다만 조사환경은 각 나라마다 다르고 변화하는 모습도 달라서 다른 나라의 우수한 연구사례를 무조건적으로 받아들여서는 곤란하고 우리나라 실정에 맞는 방법을 검토하고 시험조사를 통해 지속적인 검토 필요