

제4차 유럽조사 연구학회(ESRA) 참가 결과 보고
The 4th Conference of the European Survey Research Association

2011. 8. 1.

통 계 개 발 원
조 사 연 구 실

I. 참가 개요

□ 참가회의 : 제4차 유럽 조사연구 학술대회 (ESRA*)

○ 대회기간 : 2011. 7. 18(월)-7. 22(금) / 5일

○ 개최지 : 스위스, 로잔(로잔대학)

* The conference of the European Survey Research Association은 유럽국가를 중심으로 어려운 조사환경을 극복하여 통계품질을 향상시키기 위한 전문학술회의

○ 대회규모

- 5일 동안 약 150여개 세션, 약 550여개 논문 발표

- Short course 2개 주제(조사 무응답, 다수준 모형 추정)

- Keynote 2개 주제(웹조사, 복합조사 자료 분석)

→ 학계, 국가통계기관 실무자들이 참여하여 발표한 세션 중 대부분이 국가통계 신뢰성 향상을 위한 방법론 연구가 핵심

□ 출장기간 : 2011. 7. 17(일) ~ 7. 24(일) / 6박8일

□ 출 장 자 : 조사연구실 김서영 사무관

□ 대회목적

○ 유럽지역 국가들을 중심으로 조사통계 연구 전문가들이 모여 조사통계를 작성하는 과정에서 발생하는 문제해결을 위해 학문적, 실질적 방법들에 대한 연구결과를 발표.토의

○ 어려운 조사환경은 조사통계 품질에 직접적인 영향을 미친다는 점을 공감하고 이를 극복하기 위한 다양한 아이디어를 공유

- 특히, 학자들과 통계 이용자들이 이론 또는 현실적 측면을 활발하게 논의함으로써 조사통계의 현실적인 대안을 모색하고, 이를 위한 객관적인 방법을 공유한다는 데 학회의 의미가 있음

□ 출장목적

- 통계개발원 연구 결과를 국제학회에서 발표함으로써 통계품질 향상과 국가통계발전을 위한 통계청의 성과를 국제사회에 홍보함과 동시에 통계청의 위상 강화에 기여
- 질 좋은 조사통계 작성을 위한 국제적 연구동향을 파악하여 통계개발원과 통계청 연구결과의 질적 향상을 도모하고,
- 특히, 응답률과 통계품질 향상 측면에서 접근할 수 있는 자료수집방법에 관한 대처방안을 모색하기 위한 선진통계기법을 공유함으로써 통계청의 미래 대응전략을 세우는 데 도움을 얻고자 함

II. 주요내용

1. 통계개발원 연구결과 발표

- 다양한 형태의 표본조사자료와 등록자료 및 전수자료를 복합적으로 이용할 수 있는 추정기법으로,
 - 기존의 추정 방법에 비해 좋은 표본조사자료를 이용할 수 있다는 점과 행정자료를 커버리지 오차를 제어할 수 있다는 점에서 실용성과 학술적 측면에서 매우 고무적인 방법
- 주제 및 내용

(발표세션) Combining and Enhancing data II (7월 18일)

(주제) Small area estimation combining information from several sources

(다양한 보조정보를 이용한 소지역 추정)

(내용 요약) Small area estimation improves the direct estimators using auxiliary variables. The auxiliary variables are often obtained from census or from other external sources. A statistical model is used to link the direct estimator and the auxiliary variable. Other external sources can include other independent surveys and results from other administrative data

2. 본 학회의 이슈

□ ESRA 2011 주요 세션 토픽

- Sampling and sample design
- Combining and Enhancing data
- Mixed mode surveys
- Using paradata to improve survey data quality
- Nonresponse-Unit nonresponse and attrition
- Indicators of survey data quality
- Linking survey and administrative data
- Web surveys
- The quality of Longitudinal survey data
- Survey analysis techniques
- Question testing and piloting

□ 주요 세션 내용(참가 이슈별)

- 어려운 조사환경을 극복하고 질 좋은 통계 생산을 위한 수단으로
 혼합조사 모드 활용을 위한 연구 활동 활발

- 혼합조사의 경우, 한 조사에서 다양한 자료수집방법을 사용하기 때문에 응답률 향상과 대면조사 대비 비용감소 측면에서 많은 장점이 있음
- 반면에 측정오차, 선택편향, 모드 효과, 비용효과 등 다양한 문제는 자료품질에 좋지 않은 영향을 주는 것으로 판단
- 이를 극복하기 위한 다양한 연구가 진행
- 혼합조사 모드 활용이 통계자료에 미치는 영향
- 혼합조사에서 비용과 자료의 품질과의 관계
- 효과적인 혼합 조사표 설계

○ 무응답 처리 문제 연구

- 단위 무응답과 항목 무응답 특성 연구
- 단위무응답의 가중치 조정과 항목 무응답의 복합대체 방법 이슈

○ 자료 혼합과 강화

- 각기 다른 목적으로 조사된 또는 등록된 자료의 혼합을 통한 새로운 정보 도출 가능
- 다양한 전산적 또는 통계적 방법 (링킹, 모형) 개발에 관심

○ 다양한 자료의 복합적 활용을 위한 자료 링킹 방법

- 행정자료간의 링킹
- 조사자료와 행정자료간의 링킹

○ 조사자료의 품질 측정 방법

(1) 혼합조사 모드

- 웹조사를 위한 주소기반 표본추출과 mail contact
 - 랜덤샘플 조사에서 email 주소 이용이 어려운 문제
 - 혼합조사로써 이 문제를 해결
 - 가구 주소기반 표본추출과 우편에 의한 웹조사 응답여부 조사
 - 웹에 응답하고자 하는 가구는 웹조사를 실시, 그렇지 않은 가구는 우편조사 실시

- **혼합조사에서 측정오차** : 자료수집방법과 질문지 형태
 - 혼합조사의 목적은 다른 소스 자료를 combine 하기 위함
 - 이 경우 자료는 잘 혼합될 수 있고, 측정은 동일하다고 가정
 - 자료수집방법이 다르면 자료는 다를 수 있음. 왜냐하면 ① 모드 자체가 다른 응답프로세스를 갖고, ② 모드에 따라 다른 질문지가 사용될 수 있기 때문
- **혼합조사에서 무응답 편향 조정 반드시 필요**
 - 프랑스 가구 여행자 조사는 응답률 감소에 대한 노력으로 혼합조사 방법을 연구
 - 자료수집방법에 따른 모드 효과와 선택 편향을 분리해서 추정
 - 모드에 따른 무응답 가중치 조정 반드시 필요→편향 감소
- 네덜란드 사례
 - Non western minorities (SIM2010)
 - 이전까지는 대면조사만 사용
 - Sequential mixed-mode 조사

(2) 무응답 처리 문제 - 단위 무응답과 항목 무응답

- 무응답 왜 문제인가 ?
 - 표본크기의 감소로 인한 결과의 신뢰성에 영향
 - 이것만 문제라면 초기 표본을 많이 추출함으로써 대처
 - 무응답 문제는 표본크기보다는 편향에 의한 오차 발생에 있음
 - 무응답 편향은 랜덤이 아닌 선택적 무응답에 있음. 이는 응답 선택확률에 영향을 받지만, 이 선택확률은 아무도 모름. 따라서 정확한 추정치 계산은 더 이상 불가능
- 무응답이 있는 조사는 편향된 추정량을 도출
 - 대부분 모든 조사에서 무응답은 발생하고, 이를 반드시 조정함으로써 편향을 보정할 필요가 있음
- 무응답은 자료수집 모드에 영향을 받음
 - 대면조사(높은 응답률) vs. 자기기입식(낮은 응답률)

- Multiple imputation 방법
 - 다중 대체 방법은 다양한 자료 특성을 고려한 방법으로서 최근 대중적으로 사용되고 있음
 - 이산형 또는 연속형 자료를 겨냥한 기법 개발
 - 특히, 패널 서베이의 경우 다중 대체 기법이 유용
- 연구사례들
 - 조사에 참여한 그룹과 참여하지 않은 그룹간 특성 비교(벨기에)
 - 온라인 서베이에서 왜 응답자들이 탈락하는가? (독일)
 - German Soci-Economic panel study

(3) 자료 혼합

- 자료수집 방법이 다른 자료들간의 연계시 선택 편향 문제
- 개인 수준 매칭에서 무응답이 있는 자료들의 통계적 매칭 연구
- 조사자료와 행정자료의 연계 방법
- 개인 레코드 매칭에서 근사적 string(문자열) 매칭 알고리즘
- 통계적 모형을 이용한 자료 혼합 방법 (합성추정, 복합추정)

(4) 행정자료와 조사자료 혼합 : 방법론적 전망

- 행정자료와 조사자료를 연계했을 때 무응답 처리
- 개인 레코드 단위 자료 연계시 응답자, 조사원 동의 문제
 - 개인 레코드 기록은 동의 없이는 사용할 수 없기 때문에 이 경우 비동의(non-consent)에 따른 편향 추정은 매우 복잡
- 유일한 키변수가 없는 경우 통계적 매칭방법은 다른 자료 소스를 혼합하는 데 널리 응용되고 있음
- 아일랜드 사례 : 행정자료들간 연계 자료의 시험 분석

(5) 웹조사의 측정오차

- 장점과 단점
 - 응답자의 응답을 위한 시각적 도구 사용 가능 : 대면조사나 전화조사의 인식 부담 감소

- 실시간으로 응답자의 반응 수용
- 너무 지나친 시각적 강조에 따른 응답 편향
- 국가통계작성시 웹조사 사용: 적절한 디자인에 대한 연구 필요

(6) 파라미터를 이용한 무응답 편향 평가 및 보정 방법

- 조사 추정치에서 무응답 편향은 무응답률과 간접적인 관계
 - 무응답 편향 측정을 위한 보조변수를 찾는 것이 중요
- 성향점수 방법을 통한 무응답 편향 조정
- 칼리브레이션에 의한 무응답 조정을 위해 파라미터 사용
- 무응답 편향 조정을 위해 방문기록 자료 확보

Ⅲ. 참가소감 및 시사점

- 유럽조사연구학회는 유럽사회를 측정할 수 있는 통계와 관련하여 다양한 분야의 조사통계 활용자 및 연구자들의 토론의 장
 - 통계가 국제사회 흐름을 측정하는 지표가 된 만큼, 객관적인 통계작성은 더 이상 어느 한 국가만의 문제는 아님
 - 기본적으로 통계에 대한 객관성과 신뢰성 보장을 위한 다양한 통계작성 방법론과 이렇게 작성된 통계자료를 통해 경제사회 현상을 정확하게 해석할 수 있는 분석연구가 중심
 - 통계는 객관적으로 만드는 것이 가장 중요하고, 더불어 정확하게 활용할 수 있을 때 그 가치가 있다고 보임
- 어려운 조사환경을 극복하기 위해서는 다양한 통계적 방법론 연구는 필수적이며, 이에 대한 적극적인 투자가 중요
 - 유럽은 조사를 통한 통계작성의 한계에 직면하고 있는 듯한 느낌
 - 서베이로 극복할 수 없는 부분들은 행정자료나 통계적 방법을 통해 정보 창출을 하고자 노력하는 분위기

- 유럽 등 통계선진국가들처럼 과제 또는 사안에 따라 보다 원시안적인 시각으로 국가 프로젝트를 운영하는 것이 절실한 시점
 - 참고로, 최소한 단위 무응답에 관한 한 통계청의 관심도와 연구수준은 유럽사회에 비해 상당한 뒤져 있는 모습
 - 통계선진국의 전문성과 기술적 수준을 충분히 이해할 필요
- 실무 적용까지 충분한 연구기간을 거침으로써 실적위주의 즉흥적 개편에 따른 시행착오를 줄일 수 있을 것임
- 국가통계가 그 나라의 경제사회 현상을 파악하는 척도인 만큼, 그 사회의 풍습, 관습, 국민의식수준 등을 반영할 수 있는 통계 작성이 중요
 - 새로운 통계정책을 시행하기 전에 충분한 기초연구와 실험연구, 시험작성을 거쳐서 새로운 통계가 탄생한다는 느낌
 - 통계조사가 힘든 유럽 국가들은 다양한 행정자료 활용이나 행정자료와 조사 자료의 혼합을 통해 새로운 정보 창출에 노력
 - 이러한 일련의 과정들이 학계와의 연계를 통해 다양한 가능성을 전제로 연구를 거치고 있다는 느낌
 - 통계를 만들고, 바꾸고, 없애는 작업이 매우 중요한 만큼 단시간적인 고민에 의한 결정이 아닌, 좀 더 치밀한 계획하에 검토와 연구를 거듭하여 통계정책에 반영하는 것이 중요
 - * 충분한 검증과 연구가 밑받침되지 못한 방법을 국가통계에 적용함으로써 실패한 경우들도 많은 다른 문헌과 학술대회를 통해 보고되고 있음
- 우리 상황에 맞는 연구와 선진통계기법 벤치마킹이 중요
- 선진통계국과 우리와 통계 상황이 다른 경우가 많고, 연구현황 및 실무통계에 대한 방법론 고민도 다름
 - 유럽은 사회복지정책이 잘 갖추어진 나라로서 이에 걸맞는 행정자료들이 존재. 이를 근거로 통계를 작성하는 것은 우리에게 비해 훨씬 수월할 수도 있음

- 유럽은 서베이 현실이 매우 열악한 상황이라 할 수 있음. 논문발표자 및 실무자들의 발표에 따르면 조사통계의 응답률을 50%이상 기대하기가 매우 어려운 현실
 - 따라서 어떤 식으로든 국가적 차원의 정보를 생성하기 위해서는 모든 수단이 동원될 수밖에 없음
 - 연구에 있어서도 정보의 정확성에 비해 현실성에 더 치중하는 느낌을 받았음
 - 우리는 충분한 검토와 연구 없이 흔히 ‘선진통계기법’이라고 하는 것들을 흡수하는 경향이 있음

- 통계의 질적 향상을 위해서는 우리 통계가 처한 현실을 직시하고 통계선진국의 연구방향을 주시하면서 통계청을 업그레이드하는 것이 중요하다고 생각됨