
2013 북미통계학회 참가 결과보고

2013 The Joint Statistical Meeting(JSM)

2013. 8.

통 계 개 발 원
동 향 분 석 실

1

출장 개요

□ 참가회의 : 북미통계학회(JSM* 2013)

JSM; Joint Statistical Meetings

○ 회의기간 : '13. 8. 3.(토) - 8. 8.(목) / 6일간

○ 개최지 : 캐나다 몬트리올 컨벤션센터

(Palais des congrès de Montréal)

○ 대회규모

- 699개의 세션, 약 4000여개의 논문 발표

- 약 6,000명 참여

- 단기 교육과정 프로그램(Short course) 39개 개최

○ 주관 및 협조기관

- American Statistical Association

- International Biometric Society,

- Institute of Mathematical Statistics,

- Statistical Society of Canada

- International Chinese Statistical Association

- International Indian Statistical Association

- International Society for Bayesian Analysis

- Korean International Statistical Society

□ 출장기간 : '13. 8. 3.(토) - 8. 9.(금) / 5박 7일

□ 출장자 : 통계개발원 윤연옥 동향분석실장

□ 출장목적

○ JSM에서는 통계와 관련된 다양한 주제가 다루어지며, 특히 우리 청에서 관심이 많은 가중치, 개인비밀보호기법, 통계조사기법, 자료수집, Big data 등과 관련된 최근의 동향과 다양한 사례 등이 소개되고 있음

- 다양한 측면의 통계적 방법론과 연구동향을 파악하여 우리청에 활용 가능한 부분을 습득하고 연구에 적용하고자 함
 - 선진통계국가들이 현재 어떤 고민을 하고 있고, 어떻게 해결책을 찾고 있는지를 이해함으로써 개발원 연구에 접목하고자 함

2 주요 내용

□ 정책과 통계와의 관계

- 공식통계가 정책에 영향을 받는 여러 가지 경우들에 대해 발표

[중국 사례]

- 최근 중국 공식통계의 문제점
 - 중국통계청은 정부에 유리한 자료를 공표하도록 요구되고 있으며
 - 경제조사는 이례적으로 아주 빠르게 조사(조사의 신뢰성이 의심됨) 되었으며, 종합적인 실업률보다는 협의적인 실업률을 공표하고, 매월 공표되는 GDP의 구성요소들은 연간자료의 구성요소들과 일치하지 않음
 - 또한 석탄생산량은 공표하지 않고 은행의 품질 관련 자료에 대한 독립보고서는 거부당하고 있음을 언급함
- 이러한 문제점에 대한 외부의 의식을 고려하여 중국정부는 여러 방향으로 정책을 변경함

[아르헨티나 사례]

- 아르헨티나 물가지수의 문제점에 대한 내용
 - 라틴아메리카 5개 국가들의 큰 슈퍼마켓으로 부터 수집된 자료 (2007년 10월 ~ 2011년 3월)를 분석하여 물가지수와의 관계를 분석
 - 분석결과, 브라질, 칠레, 콜롬비아와 베네주엘라의 온라인 물가지

수는 공식 인플레이션의 변화와 어느 정도 근접하게 추정된 반면, 아르헨티나의 온라인 인플레이션율은 공식통계보다 3배정도 높았음을 보여줌

[한국 사례]

- 한계레 신문에서 발표된 지니계수와 소비자물가지수에 대해 언급
 - 지니계수가 2012년 예정된 날짜에 공표를 하지 않는 것과 소비자물가지수에서 금(Gold)이 제외된 것에 대해 언급
 - 신문기사에 난 내용에 국한하여 설명하였으며, 한계레 신문이 진보의 성격이 강하다는 것과 중국이나 아르헨티나의 심각한 문제에 비해 한국은 비교적 덜 심각한 문제로 언급함

□ 모집단과 관련 이슈들

- 미국 센서스국의 가구조사 모집단인 MAF(Master Address File)의 2010 센서스에서의 오차에 대해 언급하고, 특히 오류 수정이 많이 요구된 지역의 특성에 대해 설명
 - 이러한 지역의 MAF 관리를 잘 할 필요가 있음을 설명
 - 또한 행정자료를 활용하여 MAF 자료를 갱신함을 언급
- 시설단위조사구 중 대학기숙사 학생들이 모집단에서 잘 업데이트 되지 않는 경향이 있음(P.O.Box를 통해 우편물을 받을 수 있어서 새로 생긴 대학기숙사 학생들의 정보가 누락되는 경우가 많음)
 - 행정자료의 활용은 물론이고, 사전 전화조사를 활용하여 최신의 정보로 업데이트 하는 방법에 대해 토론
 - ※ 대학기숙사학생들은 시설조사 전체 인구의 66.7%를 차지
- 통계적인 모형을 활용하여 누락된 자료추정, MAF의 정확도 점검, 무응답자 추적 모형 등 통계적인 기법을 활용한 연구결과 발표

□ 북한에 대한 과학(조사통계) 외교

Organizer(s): Asaph Young Chun, US Census Bureau

Panelists: Norman Neureiter, American Association for the Advancement of Science
Chan-Mo Park, Pyongyang University of Science and Technology
Justin Fisher, Government Accountability Office
Elena Zafarana, Swiss Federal Statistical Office

- 과학을 통한 외교의 한 분야로서 조사통계를 북한에 전파하자는 내용으로 2012년 평양 과기대 여름학기* 에서 조사통계관련 통계자원자들의 강의가 있었고, 이에 대한 내용을 발표함
 - 지속적인 관심과 많은 통계전문가들의 자발적인 지원을 기대함
 - * 2012 the Pyongyang Summer Institute in Survey Science and Quantitative Methodology (PSI)
 - Floor에서 언급
 - 지속적으로 계속 유지되고 지원되어야 한다는 언급
 - 현재의 평양과기대는 남학생만 있는데, 왜 여학생이 없는지?
 - ↳ 남녀가 함께 기숙사를 사용할 수 없어서 그러하나(Fund의 제약으로 기숙사 1개 동만 있음), 내년에는 의학 관련 학과를 생길 예정이어서 여학생(간호학과 등)도 있을 예정임.
 - 이러한 일은 너무 빨리 진도를 나가는 것 보다는 천천히 지속적으로 나가면서 외부적인 환경도 고려하는 것이 중요함



(Chan-Mo Park, Norman Neureiter, Justin Fisher, Elena Zafarana, Asaph Young Chun)

□ 가족관계 질문의 심층분석

- 사회의 변화에 따라 가족관계를 묻은 질문의 변화 필요
 - 동성부부. 다양한 부모-자식 관계의 정립, 법적인 관계 파악 등
 - 기존자료 분석, Focus groups, 인지면접 및 시험조사를 통하여 측정
 - 기존자료분석: 이성부부의 성별표시의 오류로 동성부부 과대추정
 - Focus Groups: 응답자들은 통계조사가 일반적으로 법적인 관계를 묻는 것으로 간주
 - 인지면접: 부정적인 반응 없음, 응답자들은 의미있는 분류를 원함
 - 시험조사: 다양한 통계조사에서 시험조사 실시
(NHES, ACS-ADT, AHS, SIPP-EHC, ACS, 2020 Census)

※ 개선된 내용

- 배우자(spouse) 추가 : 남편/아내/**배우자**
- 결혼하지 않는 동거자(unmarried partner)를 문항의 위에 배치
- 이성 및 동성부부로 분리(**opposite- , same-sex**)
- "mother", "father" 로부터 “parent 1”, "parent 2" 로 변경 등

■ 개선된 질문지

How is this person related to Person 1? Mark (X) ONE box.

<input type="checkbox"/> <u>Opposite-sex husband/wife/spouse</u>	<input type="checkbox"/> Grandchild
<input type="checkbox"/> <u>Opposite-sex unmarried partner</u>	<input type="checkbox"/> Parent-in-law
<input type="checkbox"/> <u>Same-sex husband/wife/spouse</u>	<input type="checkbox"/> Son-in-law or daughter-in-law
<input type="checkbox"/> <u>Same-sex unmarried partner</u>	<input type="checkbox"/> Other relative
<input type="checkbox"/> Biological son or daughter	<input type="checkbox"/> Roomer or boarder
<input type="checkbox"/> Adopted son or daughter	<input type="checkbox"/> Housemate or roommate
<input type="checkbox"/> Stepson or stepdaughter	<input type="checkbox"/> Foster child
<input type="checkbox"/> Brother or sister	<input type="checkbox"/> Other nonrelative
<input type="checkbox"/> Father or mother	

■ ACS 현재 질문지

2 How is this person related to Person 1? Mark (X) ONE box.

<input type="checkbox"/> Husband or wife	<input type="checkbox"/> Son-in-law or daughter-in-law
<input type="checkbox"/> Biological son or daughter	<input type="checkbox"/> Other relative
<input type="checkbox"/> Adopted son or daughter	<input type="checkbox"/> Roomer or boarder
<input type="checkbox"/> Stepson or stepdaughter	<input type="checkbox"/> Housemate or roommate
<input type="checkbox"/> Brother or sister	<input type="checkbox"/> Unmarried partner
<input type="checkbox"/> Father or mother	<input type="checkbox"/> Foster child
<input type="checkbox"/> Grandchild	<input type="checkbox"/> Other nonrelative
<input type="checkbox"/> Parent-in-law	

- 관계측정은 매우 복잡하나 연구결과에 의하면 개선은 가능한 것으로 나타남. 그러나 개선하기 전에 더 많은 검증절차를 거쳐서 사용되어야 함을 언급함

□ 사업체 자료의 매크로 자료 에디팅에 관한 다양한 사례들

- 미국 센서스국의 Quarterly Financial Report(QFR) 프로그램은 사업체 소득 및 대차 대조표의 자료를 매크로 단위에서 체크하고 있음. 일반적인 방법은 정해진 상한치를 초과할 때 점검이 되는데, 보다 더 근본적인 변화를 체크 하기위해 과정점검방법(Process control methodology)인 P-chart, stair-step chart를 도입하여 분석한 결과를 발표
- 캐나다 통계청에서는 사업체조사에 선택적 에디팅(selective editing)을 실시하고 있는데, 방법은 자료 수집이 되고 있는 와중에 품질체크 점수를 계산해서 각 개별 자료의 영향을 평가하여 에디팅에 이용함
 - 표본, 자료수집 및 선택적 에디팅을 통합한 방법임
- 미국 센서스국에서는 2002년부터 경제센서스의 지역통계작성에서 이상치 발견을 위해 Hidiroglou-Berthelot(HB) 에디팅 방법을 8개 무역분야 중 5개 영역에 적용해왔음
 - 광업, 제조업과 건설업 자료에 확장 적용해 본 결과, 서비스분야와는 다른 양상이 나타났음을 보여줌

□ 소수 및 접근이 어려운 그룹에 대한 표본 전략

- 조사하기 어려운 그룹(강간이나 성추행 피해자)에 대한 통계조사에서 (1) 일반 가구표본들과 (2) 복합표본들(일반 가구표본들+경찰 보고서를 바탕으로 구축된 틀)을 사용하여 조사한 결과 비용측면에서 복합표본들을 사용하였을 때 더 절감됨

- 부자들에 대한 표본틀 구축방법
 - Survey of Economically Successful Americans(SESAs)는 상위 1%의 부자들을 대상으로 조사가 되는데, 조사대상의 명부를 구하기가 어려운 문제를 안고 있음.
 - 명부 구축을 위해 시장연구조사에 활용되는 비즈니스 CEO와 부유한 개인에 대한 외부정보를 추가한 복합명부를 구축하여 대표성 있는 통계조사 실시함

- National HIV Behavior Surveillance System(NHBS) 표본 디자인
 - 2003년부터 질병관리본부(Center for Disease Control and Prevention)에서 National HIV Behavior Surveillance System(NHBS)조사를 20개의 관할 구역에서 실시하고 있음
 - 에이즈의 감염이 높은 3개의 그룹(높은 위험군의 이성교제자, 남성 동성연애자, 마약사용자들)에서 면접을 실시하는데 표본대상자를 찾기가 매우 어려움
 - 각 관할 구역에서 매년 450~500명을 면접하는데, 표본을 추출하기위한 추출틀이 없음
 - 높은 위험군의 이성교제자와 마약사용자들 그룹에 대해서는 Respondent-driven 표본추출방식을, 남성동성연애자 그룹에 대해서는 Venue-based 표본추출방식을 적용하여 조사

○ MMP(Medical Monitoring Project) 표본 디자인

- NHBS는 에이즈 증상을 가진 모든 사람을 포함하고 있기는 하지만, 인구학적 및 병원치료에 대한 정보가 제한되어 있어서, 에이즈치료를 받고 있는 환자를 대상으로 전국 및 지역을 대표하는 확률표본을 설계하고 진료 및 행동결과를 관찰함(MMP)
- 전국을 대표하는 3단계(관찰지역-시설기관-환자) PPS 추출방식으로 표본설계함

○ Respondent-Driven 표본추출(RDS)의 민감성 분석

- 희귀한 그룹이나 접근하기 어려운 그룹에 대한 조사를 위해 RDS 추출방식을 많이 쓰고 있는데, 처음 선택하는 응답자(seed respondent) 네트워크의 수에 따른 결과물 비교 분석에 기초하여 초기 응답자 수의 민감성에 대해 논의

□ **빅 데이터**

- 빅 데이터에 관한 단기코스(Short Course)와 다양한 세션이 열렸으며, 빅데이터 분석을 위한 관련 컴퓨터 소프트웨어들에 대한 소개 및 특징들, 그래픽을 이용한 분석 등이 있었음
- 분야로는 의학, 암 연구, 환경, 생태학(Ecology) 등에서 활용한 사례가 많았음
- 대학 및 연구소에서 연구결과를 많이 발표함

3

참가소감 및 시사점

- 통계와 관련된 전 분야의 연구자들이 서로의 연구결과를 공유하고 토론하는 자리로써 통계잔치를 하는 것으로 느껴짐
 - 이론 통계에 국한하지 않고, 조사통계, 국가통계, 의학통계 등 다양한 분야의 연구흐름 파악과 더불어 서로 접목할 수 있는 기회를 제공하고 있음

- 미국 센서스국(Census Bureau), 노동통계국(Bureau of Labor Statistics), 캐나다 통계청 등에서 정부통계(Government Statistics)와 관련된 많은 발표가 있었음
 - 센서스 프레임, 노동력조사(CPS), 데이터 연결, 비밀보호, 자료분석, 무응답, 물가지수 등 다양한 분야의 연구결과를 발표
 - 외국 정부통계기관의 향후 방향 및 현재의 상황을 공유할 수 있는 기회제공 및 우리청에서 고민하고 있는 문제점에 해결방안 모색에 도움

- 해외연구진과 네트워크 형성 활성화 노력
 - KISS(Korean International Statistical Society)가 JSM의 공동 주관 기관으로 설정되어 있으며, 한국학자들의 역할이 과거보다는 훨씬 활동적으로 수행되고 있음(현재 약 50명으로 구성)
 - 통계청업무 발전을 위한 KISS 학자들과의 공동연구(또는 자문) 추진을 위한 네트워크 형성

- 통계 관련 다양한 영역의 내용을 포함하고 있어서, 연구부서(개발원)는 물론 통계청 각 분야에서 많이 참여하여 선진 통계의 변화에 접할 수 있는 기회가 제공되었으면 함