

2015년

제60차 ISI 세계통계대회 출장결과
The 60th World Statistics Congress of Int'l Statistical Institute

2015. 8.



기 획 조 정 관
국 제 협 력 담 당 관



차 례



I. 출장 개요	3
II. 주요 활동	3
III. 세션 참가 요약	7
IV. 맺음말	22
【참고】 세션별 참가 결과 보고 자료	24

I 출장 개요

- 회의장소 : 브라질 리우데자네이로
- 회의기간 : '15. 7. 26.(일) ~ 7. 31.(금) ※ 출장기간 : '15. 7.24.(금) ~ 8.3.(월)
- 참석자 : 대표단장(허남덕 기획조정관)외 5명

구분	성명	소속(직급)	업무수행내역
대표단 총괄 및 단장 보좌	허남덕	기획조정관 (국장)	.제60차 ISI참가 대표단장 .대표단 대외협력활동 업무 총괄
	김병현	국제협력담당관 (6급)	.ISI 본회의 준비 및 참가 .단장 보좌, 자료 수집, 결과보고
논문 발표	박운영	국제협력담당관 (5급)	.발표주제 : “한국의 국제기구 자료 제공 및 활용 관리 현황과 개선계획” .타 세션 참가 및 자료 수집
	김두만	통계조정과 (5급)	.발표주제 : “한국의 통계역량 강화 경험으로 본 The Post-2015(SDGs) 이행체계 개선 방향” .타 세션 참가 및 자료 수집
	김현기	품질관리과 (5급)	.발표주제 : “한국 공식통계의 품질관리시스템” .타 세션 참가 및 자료 수집
	손희경	통계기준과 (7급)	.발표주제 : “노동시간과 소비에 영향을 미치는 거시 경제 요인 분석” .타 세션 참가 및 자료 수집

II 주요 활동

1. 우리 청 참가 논문 발표 및 질의 응답 실시

[논문 발표 일정]

구분	발표자	발표일	소속(직급)	발표주제
초청	김현기	7.28	품질관리과(5급)	한국 공식통계의 품질관리시스템
기고	손희경	7.29	통계기준과(7급)	노동시간과 소비에 영향을 미치는 거시경제 요인 분석
기고	박운영	7.30	국제협력담당관실(5급)	한국의 국제기구 자료 제공 및 활용 관리 현황과 개선계획
기고	김두만	7.30	통계조정과(5급)	한국의 통계역량 강화 경험으로 본 The Post-2015(SDGs) 이행체계 개선 방향

2. 주요행사 및 전문세션 참가를 통한 정보수집

- (주요행사) ISI 현재 활동 점검 및 방향설정을 위한 종합 행사 등

- **총회***, 환영식, 환송식, 시상식, 폐회식, ISI 회장 특강, ISI Family 회장단 회의, IAOS 및 IASS 주관 행사 참가

* 총회(7. 30일 개최) 안건으로 상정된 활동보고서, 재무보고서, ISI 산하 협회 활동보고서, 세계통계대회 개최 및 유치(2015 Rio, 2017 Marrakech, 2019 Kuala Lumpur)에 대해 심의 및 승인 등의

- (전문세션) 국내외적으로 부각되고 있는 **Big Data** 등 관심 세션 참가

- Big Data 관련 **주요세션***에 참가하여 최신동향을 파악

* (IPS 074) 빅데이터 통한 공간분석의 과제와 기회
 (IPS 064) 빅데이터 회귀분석에서 희소성과 레버리징
 (IPS 044) 위성영상분석에 적용한 빅데이터 추론과 공식통계 모델링
 (IPS 026) 국가통계기관의 새로운 데이터 소스 및 자료수집방법활용 방안
 (IPS 078) 빅데이터와 공식통계-국가통계시스템의 빅데이터 과제
 (IPS 146) 개인정보보호, 비밀정보보호, 빅데이터

- 아울러, 인구주택총조사 등 업무 **유관세션***에 참가하여 정보수집

* (IPS 016) Post 2015 환경하에 발전, 지속가능성장과 웰빙 측정
 (STS 012) 국가통계기관의 품질관리
 (IPS 022) 인구주택총조사의 변화: 진화에서 혁명으로
 (IPS 025) 거버넌스와 혁신/통계 품질 관리 시스템: 구축, 이행, 모니터링
 (STS 034) 공식통계에서 새로이 부상하는 주제 및 기술
 (STS 068) 공식통계에서 행정자료의 역할
 (IPS 101) 통계 및 공간지리정보의 통합을 향하여
 (STS 071) 공식통계 개선을 위한 현대화 전략
 (IPS 116) 통계역량구축: 최근 경험, 수요, 도전
 (IPS 050) 방법론과 기술이 변화하는 시대의 센서스
 (IPS 147) 인터넷 조사-과제 및 기회
 (IPS 105) 표본조사: 과거, 현재와 미래 방향

3. 양자협력 실시

[아부다비 통계청]

- 대표단은 7. 27일 아부다비 통계청 전시관을 방문하고 현장에서 국제과장(Abubaker)일행과 주요 관심사에 대해 논의 실시

- 아부다비측은 등록센서스 지식공유를 위한 직원파견, 2016 IAOS 참가 요청, 양 기관 MOU 재개정 등에 대하여 의견을 타진
- 우리측은 파견허락시 관리 실무부서인 조사관리국에 문의 후 조율할 것을, IAOS 참가는 행사계획이 수립되면 참가방안을 검토하고, MOU 재개정보다는 공식서한 등을 통한 협력관계를 강화함으로써 형식보다는 실질적인 협력을 강화할 것을 주문

[독일통계청]

- 대표단은 7. 30일 독일 통계청 대표단(단장 : Mr. D. Sarreither 부청장)일행과 오찬을 가지고 주요 관심사에 대해 논의 실시
- 독일 측은 통계청장의 후임으로 부청장이 내정(두 번째 내부승진)된 반가운 소식을 전달, 아울러 2016년 한독 양자협력에 청장을 단장으로 5월 중순경 방한 의사를 타진하고, 관심 의제로 인구센서스를 제안 및 한국측 관심사항 문의
- 우리대표단은 청장내정 사실을 축하하고, 내년 5월 중 방한에 대해 긍정적인 검토와, 추가의제로 행정자료 활용(법적이슈, 조직재설계, 자료입수 등)을 제시하고, 방한관련 내용은 귀국 후 양측이 추가로 협의해 가기로 함

4. 이외 국제 교류 활동

[모로코통계청]

- 대표단은 7. 28일 ISI 2017을 유치한 모로코 대표단 Hasnae Fdhil(국제과장)을 만나 유치 경험에 대해 질의하고
- 모로코측은 2010년 ISI 회장 및 집행이사회 주요임원을 모로코로 초청하여 대회장 소개, 유치계획 등 사전 설명회를 실시하고, 2013년 ISI 총회에서 유치지 승인을 받고, 2015년 동 행사에서 공식초청(2017 ISI PPT 및 동영상) 실시 예정 설명

[ISI 회장 등 면담]

- 대표단은 7. 30일 ISI 회장 등 주요임원*들과 면담을 가지고 관심 사항을 토의
 - * Pedro Luis Do Nascimento Silva(ISI 회장), Helen MacGillivray(ISI 차기 회장 내정자), Ada Van Krimpen(ISI 사무국장)
- 대표단은 2001 ISI 개최이후 상당기간이 지남에 따라 재유치 가능성 문의(적정 시점과 방법 등 자문) 및 유치 후 20년 경과와 대륙순환 규정의 실효 여부 질의
- ISI측은 한국의 성공적 유치경험을 기억하고, 2021년 이후 개최지는 공식적으로 결정된 바 없음(단, 일부국가 비공식적으로 관심표명 중)을 전달하고, 20년 경과 및 대륙순환 규정은 가이드라인이므로 강행 규정이 아님을 설명
- 아울러 재유치는 희망의사, 희망시기, 희망장소 등 을 명기한 유치 희망공식서한을 작성하여 ISI측에 전달하고, 특정연도 유치 희망시는 유치명분을 상세화하면 검토 시 유리함을 강조

[빅 데이터 세션]

1. (IPS 074) 빅 데이터 통한 공간분석의 과제와 기회

○ 세션요지

- 계층적 공간과정모형은 통계 기법들을 효율적으로 이용하여 공간변수의 복잡성을 잘 설명하고 있으나, 이모형을 추정하기 위해서는 복잡한 매트릭스 분해가 필수 불가결했음을 설명
- 아울러, 최근접 가우스 과정(Nearest-Neighbor Gaussian Process)을 소개하고, 미국 산림통계 데이터베이스의 자료를 이용하여 생물체량(biomass)의 공간분포에 대한 추론 사례 소개
- 온라인 사회관계망은 지진이나 전염병과 같은 실시간 사건을 모니터링하기 위한 중요한 정보원이며 특히 이들의 행동 특성은 공간위치 파악에 중요한 정보를 제공해 줄 수 있으므로, 트위터 사용자의 공간위치를 추정할 수 있는 방법에 대해 논의하고, GPS에 의한 위치정보를 이용하여 트위터 이용자들의 대용량 자료속에서 모형의 정확성을 시험하였음

○ 세션시사점

- 공간확률과정 두 접근법 모두 많은 공간위치들 속에서 선형을 이루는 부동 소수점 연산(floating point operations)을 가진 마코브 체인 몬테카를로 알 고리즘(Markov chain Monte Carlo algorithms)에 수렴함
- 트위터 글과 지리적 표시를 결합하여 모형을 만든 확률적 접근법은 정보의 두 가지 유형 중 하나를 사용하는 two baseline algorithms(Naive Bayes Model and Integrated Data Aggregation)에 비교되는데, 결과적 측면에서 이 방법이 baseline methods보다는 확연히 더 우수함

2. (IPS 064) 빅 데이터 회귀분석에서 희소성과 레버리징

○ 세션요지

- 계수 모형을 다양화하는 것은 학술분야의 많은 영역에서 적용을 할 수 있으며, 공변량 동적 충격을 모형화하는데 매우 유용함

- 빅데이터에서 유용한 정보와 지식을 추출하는 것은 과학부문 뿐 아니라 전체 사회 측면에서도 부담이 되는 과제이며, 이를 위해서는 빅데이터 분석에 있어 효율적인 계산방법 및 통계적 접근방법 측면에서 괄목할 만한 성과가 필요
- 이러한 과제를 해결하는데 있어 중요한 기여를 하는 레버리징 알고리즘들을 소개하고, 컴퓨터로 조정되는 알고리즘들을 사용하여 아주 작은 대표적인 하위 데이터셋을 샘플링함으로써, 작은 하위 데이터 셋에서 방대한 데이터셋 관련 정보를 효율적으로 추출할 수 있으며, 이러한 알고리즘들은 빅데이터로 확장가능

○ 세션시사점

- 선행변수 선택절차는 기존 방법들과는 달리 선택일관성을 검증하기 위한 추가 확인단계가 필요 없으며, 비규칙적인 패턴 조건일지라도 다소 덜 제한적인 구조라면 효과적인 확인 절차로써 유용함
- 레버리징 알고리즘은 효율적인 데이터 마이닝을 가능케 하므로, 일반 사용자들이 태블릿컴퓨터를 이용하여 빅데이터 분석을 할 수 있는 등 빅데이터 분석 대중화에 기여할 것임

3. (IPS 044) 위성영상분석에 적용한 빅 데이터 추론과 공식통계 모델링

○ 세션요지

- 주로 빅데이터는 포괄범위 편의, 자기선택 편의, 대표도 편의, 응답편의 또는 측정오차 문제에 봉착해 있는데, 이러한 비표본 오차 문제를 해결하기 위해 빅데이터 추론을 위한 베이지언 구조-빅데이터 측정에 적용된 ①개념 변형, ②샘플링 및 ③검증과정에 대해 논의
- 고품질 및 다양성을 갖춘 데이터가 요구되는 현대사회에서 예산부담 및 응답부담 관련 비용은 조사 범위를 제한하므로, 공식통계에서 빅데이터 이용과 새로운 데이터 원천을 활용하는 것은 정보와 지식의 충분한 공급원이 될 수 있음
- 이탈리아 통계청은 웹자료수집과 문자 마이닝 기법을 이용한 통계작성을 위해 “기업 정보통신기술조사”를 실시하고 빅데이터 잠재력을 평가하기 위한 통합 연구프로젝트에서, 노동시장 지표와 구글 트렌드 지표를 추정하여 작성했으며, 현재 진행 중인 실험은 “사람과 지역”으로, 핸드폰 자료사용을 근거로 하여 지방자치단체 공간 단위로, 업무용과 학습용 일일 이동성의 기원/출발 매트릭스작성을 목표로 함

- 빅데이터는 공식통계작성기관에 있어 도전과 위협이 되는 특별 과제이고, 통계 이용자들은 그 어느 때보다도 보다 신속한 통계, 보다 다양한 통계를 요구하고 있는 실정이며, 빅데이터가 그 해답이 될 수도 있지만, 정확성 문제 소지
- 품질 측면에서 지속적인 단기예측을 위한 벤치마킹의 기초를 제시하므로, 공식 통계를 위한 패러다임으로써 “REVEAL”을 제안하며, 그 과정은 인터넷과 행정 자료와 같이 잘 알려진 자료원천으로부터 정보를 재수집하는 것, 사전 조사나 총조사 같은 기준점으로부터 추정하는 것, file operation을 가지고 추정치를 확인하는 것, 측정 차이를 설명하고 평가하는 것, 오차 원인을 분석하고, 향후 이러한 자료 원천을 사용할지 여부를 결정하는 것, 예측 방법론을 개선하기 위해 품질측정과 추론을 재 연계하는 것, 기존 조사들을 검토하기 위해 추정치를 사용하기보다, 조사는 추정치의 적합성을 확인할 수 있는 것이어야 함

○ 세션시사점

- 공식통계로서의 빅데이터의 가능성과 함께, 일반적인 논의사항(품질관리, 정확성 및 적시성, 자료 연계, 프로파일링 방법 및 시각화 도구) 뿐 아니라, 최근 대두되고 있는 프라이버시 문제(법적 구조, 윤리적 가이드라인 및 사생활 보호를 위한 technology solution 등) 그리고 IT 이슈(cloud computing 및 machine learning)는 시급히 선결해야할 과제임
- 공식통계작성기관에서 빅데이터를 통계작성 방법으로 사용하기 위해서는 기존 조사의 장점들을 적극 활용하여 새로운 통계작성 패러다임으로의 변화를 도모해야 함

4. (IPS 026) 국가통계기관의 새로운 데이터 소스 및 자료수집방법활용 방안

○ 세션요지

- 정보통신기술의 발전과 함께 새로운 자료 원천이 많은데, 새로운 자료를 사용하고자 하는 에스토니아 통계청의 포부를 설명하고, 자료 접근 및 자료처리를 위한 전제조건들과 법적 측면을 제시한 후, 관광통계 작성을 위해 위치이동자료사용에 중점을 두었던 국제 컨소시엄 결과를 간략히 소개
- 2008년, 에스토니아 은행은 위치이동자료 연구 경험이 있는 연구원들과 공동으로 새로운 자료수집 방법론을 개발하기 시작했고, 그 결과 개발된 방법론은 에스토니아 국경을 건너 해외로 여행하는 사람들과 에스토니아로 입국하는 사람들을 대상으로 작성한 통계에서 신뢰성을 입증받았으며, 2012년 이후 에스토

니아 은행은 공식적으로 국경 관광 통계를 작성하여 공표하고 있음.

- 말레이시아는 공식통계시스템하에서 지식기반경제지표를 전반적으로 분석할 필요가 있음을 설명함
- 지금의 공식통계시스템은 농업 및 산업현장 그리고 정보화 시대 성장을 설명하는데 있어 타당성 및 포괄성 부족에 근거를 두고 있음. 인터넷 등 광범위한 커뮤니케이션은 지식기반경제의 규모를 계속 증대시키고 있으며, 사·공간적 제약없이 디지털 프로세스와 콘텐츠를 강화하고 있으므로, 지식기반경제를 측정할 수 있는 새로운 통계 조사들을 수용하는 것을 포함하여, 現 공식통계시스템을 면밀히 검토하고, 재정비해야함을 말레이시아 지식기반경제 통계 경험을 통해 설명함
- 새로운 자료원천의 대두는 공식통계를 위한 많은 전략적·방법론적 논의를 제기했음. 지식, 전문성 그리고 기법과 방법론을 공유할 때, 효율적인 협업은 공식통계산업이 빠르고 효과적으로 현대화하는데 있어 매우 중요함. 통계생산 및 서비스를 위한 고위급 그룹¹⁾(HLG)는 통계생산과정의 효율성과 사용자 니즈를 충족시키는 결과를 생산하는 능력을 향상하는 것을 목표로 하고 있으며, 그 핵심 사업²⁾ 중 하나가 공식통계에서의 빅데이터 프로젝트임
- 이용가능한 데이터의 증가와 이용자들의 요구 증가로, 국가통계작성기관들은 관련성과 적시성을 갖춘 통계를 생산해야 하는 난관에 봉착하였고, 통계가치를 극대화하기 위해 비전통적인 자료원천과 조사 자료원천을 통합함으로써 그리고 고급분석을 할 수 있는 새로운 기술을 활용함으로써, 빅데이터 특히 행정자료의 장점을 이용하기 위하여 업무 관행을 변경할 필요
- 호주통계청에서 의미기술(semantic technology)을 사용하여 개발하고 있는 시험용 GLIDE(prototype Graphically Linked Information Discovery Environment)가 바로 그 사례이며, 이 환경은 국세청 자료와 세부적인 미시 분석이 가능한 ABS 사업 등록 자료를 통합한 시험용 고용주-피고용주 연계 DB를 포함하고 있음

○ 세션시사점

- 급변하는 시대적 변화를 반영하고 이용자들의 니즈를 충족시키기 위해, 새로운 자료원을 공식통계 작성에 사용할 수 있는 방법 모색 필요

1) High-Level Group for the Modernization of Statistical Production and Services

2) Generic Statistical Business Process Model (GSBPM), a standard Generic Statistical Information Model (GSIM), a Common Statistical Production Architecture (CSPA) and project on Big Data in Official Statistics.

- 새로운 자료를 통계생산 및 서비스의 현대화에 활용하기 위해 공식통계작성 기관들은 서로 공조하여야 하며, 기존 통계와 새로운 자료와의 통합을 통해 보다 부가가치가 높은 통계를 생산해야 할 시점임

5. (IPS 078) 빅 데이터와 공식통계-국가통계시스템의 빅 데이터 과제

○ 세션요지

- 현대사회가 점점 복잡해지면서 보다 더 높은 수준의 관련성과 시의성을 갖춘 새로운 통계에 대한 수요가 증가하고 있으나, 전통 조사 방법론은 목적을 가지고 이미 만들어진 자료를 수집하는 것이므로 목적부합성은 높을지라도, 정부와 응답자의 부담은 상당
- 홍콩통계국은 정책입안을 뒷받침하기 위해 지역 의료서비스 케이스 스터디를 통해 행정자료의 적극적 활용의 실현가능성을 포괄범위의 완전성 및 신뢰성 측면에서 검토하였음
- 신속성, 다양성 그리고 대용량의 특성을 가진 자료 원천인 빅데이터의 이용가능성은 정보기술의 급속한 발전과 전자기기의 대중적인 보급과 함께 계속해서 증가추세이며, 이에 빅 데이터에서 의미 있는 정보를 찾기 위해 새로운 틀과 방법들이 개발되고 있음
- 공식통계를 위한 빅데이터 UN 작업반 그룹은 빅데이터에 대한 범세계적 업무 프로그램을 제공하고, 빅데이터 원천의 실질적 사용을 도모하며, 공식통계를 위한 민간의 빅데이터 사용에 공적 신뢰를 구축해야할 의무가 있기 때문에, 현재 이러한 작업³⁾이 진행 중임
- 빅데이터는 GPS 위치 추적 장치, 위성 이미지, 전자상거래, 핸드폰, 소셜 미디어, 스마트 계량기와 같은 광범위한 자료원천에서 비롯되는데, 이는 우리 일상생활의 거의 모든 분야에서 대량으로 발생하고 있으므로, 빅데이터는 공식통계 및 포스트 2015 개발 의제(post-2015 development agenda)를 모니터링하기 위한 의미있는 지표 전달에 사용될 상당한 잠재력을 가진 자료로 간주됨
- **주요 난제**(①데이터 접근성, ②프라이버시와 신뢰, ③빅데이터 품질관리 제도, ④빅데이터 프로젝트 이행에 있어서의 IT 요건 그리고 ⑤비용을 강조하면서, 이러한 난관을 해결할 수 있는 방법으로서 ①만-관합작(public-private partnerships)과

3) World POP, United Nations University에서의 Bigdata technology 교육, Global Land Project인 FutureEarth, 구글 등과의 업무협약 등

- ②데이터 접근을 위한 원칙 개발과 법규와 제도를 통한 프라이버시 보호 및 신뢰 유지, ③비통계작성 목적으로 작성된 빅데이터의 품질을 위해 새로운 방법론과 품질프레임워크(ex. UNECE의 sandbox)개발, ④ 기존 데이터와 빅데이터 통합을 위한 IT(cloud computing 및 big data analytics(Hadoop, R, SAS) 그리고 ④통계작성기관들 간의 비용 공동부담과 정부지원을 언급

○ 세션시사점

- 非통계작성 목적의 빅데이터를 공식통계 작성에 이용하고, 빅데이터의 유용성을 제고하기 위해서는 선결해야 할 많은 당면과제들이 있으며, 이러한 문제를 해결하기 위한 노력이 이미 국제기구를 중심으로 이루어지고 있기 때문에 각국의 공식통계작성 기관들의 적극적 참여가 요구됨

6. (IPS 146) 개인정보보호, 비밀정보보호, 빅 데이터

○ 세션요지

- 센서스 자료의 변천 소개 , 빅데이터로서의 다양한 자료의 결합, 개인정보보호 방안, 개인정보의 개념으로서 유용성, 투명성, 이중성에 대한 설명, 표준적인 정보보호와 개인보호방법, 차별적인 개인정보 방법으로서의 입실론(ϵ)-Differential Privacy, 베이시 안방법 및 복합 대체 활용 설명

○ 세션시사점

- 통계 이용자들은 지속적으로 마이크로 데이터에 대한 제공을 요청하고 있으나 통계작성 기관에서는 정보보호와 개인정보 노출 등의 문제로 제공에 대하여 부정적 의견이 있음
- 통계청도 개인정보 보호 및 정보노출을 줄이기 위한 다양한 방법(노이즈 추가, 데이터 결합 등)을 통해 마이크로데이터 제공을 위한 연구를 지속적으로 실시하여야 할 것과 아울러 빅 데이터시대를 준비하기 위해 자료의 정보보호와 개인정보 노출에 대한 선진국의 방법 연구 필요

[업무 유관 세션]

7. (IPS 016) Post 2015 환경하에 발전, 지속가능성장과 웰빙 측정

○ 세션요지

- 영국은 국가적인 웰빙을 어떻게 정의할 것인가, 어떻게 측정할 것인가, 궁극적인 목적은 무엇인가 등을 고민하고, 6개월에 한 번씩 자료를 업데이트 하고 있으며, "Life in UK" 라는 책자도 발간
- 호주는 2002년부터 2013년까지 호주의 진보 측정(Measure of Australia's Progress : MAP)내용을 발간, 웰빙의 영역을 사회, 경제 환경, 정부, 4가지로 구성하고 있음, MAP의 참고를 위한 주요 수단으로는 온라인 블로그, 비디오, ABS 페이스북 등을 활용하고 추가적으로는 컨퍼런스, 전통적인 미디어, 전문가 그룹, 워크숍, 공식회의 등을 활용하고 있음
- UNECE는 58개국의 회원국을 가지고 있으며, 8개국에 대하여 사전조사를 실시하였으며, 2016년에는 지속가능한 개발 목표의 측정을 다시 시작하려고 하고 있고, 통계청의 역할은 자료 수집보다는 국가통계 시스템의 관리일 것이며, 결론으로는 모니터링과 리포팅을 위한 관련 메카니즘을 정립하는 것이 필요함을 설명

○ 세션시사점

- 여러 국가나 국제기구에서 웰빙에 대하여 정의하고 있으며, 지속가능한 개발 목표를 향해 다양한 노력을 하고 있고, 영역과 지표의 기준을 다르게 설정하는 바 한국도 국가차원에서의 웰빙 정의 및 기준을 명확하게 설정하고 다양한 분야로의 지속적인 개발이 필요
- 궁극적으로 통일된 기준으로 지수를 나타내는 것은 한계가 있으며 사회적 연결, 주관적인 웰빙 등에 대해서는 국제기구.선진국 공조 등 다양한 시도가 필요

8. (STS 012) 국가통계기관의 품질관리

○ 세션요지

- 아일랜드는 Lean* Six-sigma를 활용하여 효율적이며 시의성(Timeless) 개선을 위해 증거 위주의 결정을 강조하고 분산 및 시간낭비를 줄이고 예측을 높이는 방안 제시
- 슬로베니아는 품질관리를 위해 통계청은 2007년부터 지속적인 개선을 실시하였고, 품질관리를 위해 타다운방식과 버튡업방식을 병행하여 실시, 타다운

방식에는 공통적인 평가 프레임(Common Assessment Framework:CAF), EU의 이행규정(Code of Practice)과 상세한 리뷰들로 구성되어 있으며, 버튬업 방식은 조사품질 보고와 DESAP(자체통계품질진단)으로 구성

- 브라질은 2013년에 2011년 아메리카 컨퍼런스에서 제시한 가이드라인, 원칙, 지표 및 우수사례 등을 기반으로 Cop를 만들고, 17개의 키 원칙과 80개의 지표로 코드를 구성하였으며, 사회통계 위원회를 2007년에 설립하였으며, 연합정부의 사회통계 생산자들로 구성되고, 경제통계 위원회구성을 위해 노력 중
- 호주는 통계작성과정에 대한 체크포인트와 결정 포인트를 마련하여 프로세스의 품질 또는 명확한 평가를 위해 활용하고 있으며, 그런 의미에서 품질 게이트체계를 구축하여 신호등(블루, 옐로우, 레드)과 같은 개념으로 관리하고 있으며, 지속적으로 개선

○ 세션시사점

- 통계작성과정의 효율성과 시의성 확보를 위해 Lean six-sigma를 도입하고, GSBPM과 같은 체계를 구축하여 통계작성과정을 표준화하고 각 국가별로 중요시하는 품질차원을 개선하려는 노력을 하고 있음
- 또한, 기존에 국가통계의 품질관리를 위해 활용하는 품질차원들을 고도화하고 보완하는 방안을 지속적으로 노력하고 있으며, 우리나라도 시의성과 효율성 등을 확보하기 위한 방안 검토가 필요

9. (IPS 022) 인구주택총조사의 변화: 진화에서 혁명으로

○ 세션요지

- 스페인의 경우, 2001년에는 약간의 등록자료를 활용하여 전수조사를 실시하였고, 2011년에는 등록 자료를 활용하고 10%를 조사에 의존, 2021년에는 현장조사 없이 기존의 조사결과와 등록 자료를 연계하여 활용할 계획임
- 2021년에는 인구등록자료, 인구와 관련된 다른 등록자료, 등록자료로 부터 얻어내기 어려운 자료에 대해 기존의 조사결과, 사전 센서스, 토지등록자료, 유동성에 대하여 모바일 폰으로 부터의 빅데이터 등을 통합하여 실시 예정
- 이를 위해 2014년과 2015년에는 인구등록자료와 관련자료(과거 센서스, 통계적 모델) 등을 활용하여 이주자과 가족관계 등의 자료를 구축하고, 현재는 이용자의 요구에 부합하는 다른 형태의 결과를 제공하기 위해 Big microdata file과 지리정보를 활용
- 아부다비도 행정자료 기반 2020 등록기반 센서스를 실시 준비중

- 이탈리아는 2011년 이전에는 조사를 통해 센서스를 실시하였으나, 2011년 이후에 지속적인 자료 업데이터 시스템으로 부터 매년 센서스 결과를 활용할 수 있게 되었음
- 현재 빅 파일럿 조사를 실시하고 있고, 150개 지방자치단체와 16만개 가구 대상으로, 통계적인 모형의 수용력 예측, 다른 원천으로부터 통합된 자료에 IT를 접목, 모바일, 태블릿 기타 등등과 같은 웹적용을 통해 데이터 동기화 테스트 실시 중

○ 세션시사점

- 다양한 국가에서 등록자료, 행정자료, 조사자료 등을 연계하여 향후 센서스에 도입하려는 시도가 계속되고 있으며, 특히, 스페인과 아부다비는 2021년과 2020년의 센서스를 위해 등록자료의 보완, 빅 데이터와 지리적인 정보 등을 활용하여 등록 센서스를 위한 준비를 지속적 실시
- 우리나라는 2015년 등록센서스를 실시하여 기존의 조사를 활용한 센서스보다 효율성과 활용성에서 성공적인 결과를 이끌어 내어 다른 국가(현재 등록센서스를 준비 중인 국가)의 모범사례가 될 수 있도록 하여야 할 것

10. (IPS 025) 거버넌스와 혁신/통계 품질 관리 시스템 : 구축, 이행, 모니터링

○ 세션요지

- 영국에서의 품질 프레임은 이론과 과정, 응답 부담, 비용적 효율성, 관련성, 정확성, 시의성, 일관성, 접근성 및 명확성으로 정의하며, 유엔에서는 공식통계는 실질적인 유용성을 검정하는 것이고 OECD에서는 품질은 이용적합성이라 정의
- Portugal은 2001년부터 2015년까지 품질관리 방법의 변천에 대하여 개괄적으로 설명하였으며, 2006년부터 2008년까지 1차 라운드의 상세한 리뷰를 하였으며, 이를 통해 2009년에 규칙 223호를 만들었고 이는 세가지로 구성되어 있으며, 첫째로 ESS(European Statistical System)가 어떻게 해야 하는지에 대한 기본 원칙과 역할을 제공, 두번째는 누가 무엇을 할지, 결정은 누가 할지, 어떻게 할지를 명확화, 셋째는 통계작성에 대한 것을 규정으로 정의
- Columbia는 품질보증체계의 주요요소로는 실천 코드, 통계적 과정, 통계품질 요소 등으로 구성되어 있으며, 이를 위해 촉진, 평가, 모니터링 순으로 관리하고, 품질인증의 기준을 3가지 유형으로 관리
 - . Type A 인증 : 점수 94~100(3년)
 - . Type B 인증 : 점수 87~93(2년)

. Type C 인증 : 점수 80~86(1년)

. 인증하지 않음 : 80점미만

- 한국의 국가통계 품질관리체계에 대하여 소개, 품질관리 방법 및 최근에 개선한 정기통계품질진단 방법을 소개하고, 통계법에 3가지 유형의 품질진단을 규정하고 있으며, 이는 정기, 수시, 자체통계품질임을 설명

○ 세션시사점

- 유럽통계청은 이미 ESS(Eurostat Statistical System)을 만들어 품질 보증 및 인증 제도를 도입하고 또한, 이를 향상시키기 위해 다양한 노력중이고, 일정기간의 진단 및 인증이 끝나고 나면 지난 결과에 대한 리뷰를 통해 개선방안 및 원칙 등을 수정 및 추가하여 통계의 품질을 향상을 위해 지속적인 노력 실시
- 현재 우리나라의 경우, 정기진단은 2006~2010년까지 1기 진단, 2011~2014년까지 2기 진단을 완료하였으나, 세부적인 검토가 부족한 것으로 생각되며, 향후 3기 진단 2015~2019년 완료 후에는 종합적인 개선방안 및 원칙 등을 재정립할 수 있는 분석이 필요할 것으로 생각
- 또한 3기 진단이 끝나기 전에 품질인증을 위한 기초적인 연구와 방법 마련을 위해 개선방안을 마련하여야 할 것이며, 기본적으로 유럽통계청의 ESS 체계와 인증, 감사 등에 대한 방법을 사전 연구 필요

11. (STS 034) 공식통계에서 새로이 부상하는 주제 및 기술

○ 세션요지

- Eurostat은 공식통계 작성의 당면과제로 모바일을 활용한 통계의 생산 및 서비스, 이용자 참여의 소셜 미디어 활용 등에 대해 ESS-2020 VISION에 담아 추진 중임을 설명 * ESS: European Statistical System
- 즉 기술적인 분석과 시각화한 자료를 모바일을 통해 이용자와 인터랙티브한 통계 서비스에 많은 혁신과 투자 예정하고, 각국의 통계청은 통계기구의 데이터 이용능력 제고, 글로벌 통계기구와의 파트너십 이용, 고위급 리더십이 강화가 필요함을 강조
- UN은 모바일 데이터, 공간 지리정보, 위성 이미지 등을 결합한 공식통계 및 지표의 활용에 대한 계획 및 사례 설명하고, UN은 Post-2015 아젠다로 17개 Goals, 169개 Targets을 설정하여 300여개의 Indicators로 모니터링 할 예정

- UN의 Global Pulse는 뉴욕, 자카르타, 우간다 Kampala에 네트워크를 설치하여 실시간의 조기경보(질병, 기후변화 등) 및 피드백에 활용 중
- 또한 UN은 2014년 글로벌 워킹그룹을 결성하여 공식통계에 빅데이터 활용전략을 마련 빅데이터의 Methodologies, 법적 제도기반, 개인정보 보호, 데이터 안전, 비용 등(8개 T/F 구성)의 이슈 해결에 실질적인 활용과 지원 추진 중
- 현재까지 모바일 데이터 및 공간지리정보는 결합을 통해 관광통계, 질병발생 분포, 인구이동의 역동성(Population dynamics), 재배 농작물 및 생산예측 정보, 지질특성 및 이용, 지역고용 및 사업체 현황 등에서 활용되고 있음
- 이집트 통계청 및 관계기관은 소비자물가지수(CPI) 자료를 이용한 가구소비 및 지출조사의 정확성 제고, 2016년 Census에 모바일 조사 준비, 통계정보 및 프로그램을 이용한 데이터마이닝의 중요성 등에 대해 설명

○ 세션시사점

- 이집트 통계청이 마련한 세션으로 이집트의 설명자료는 새로운 기법의 적용사항은 없으나 초청논문인 Eurostat, UNSD의 발표내용은 국제기구의 동향을 알 수 있는 좋은 기회
- 한국의 통계청도 국내 통계기구들의 통계 및 데이터 활용현황과 개선사항에 대해 스스로 평가하여 최신 동향을 학습할 수 있는 특별세션 구성 필요

12. (STS 068) 공식통계에서 행정자료의 역할

○ 세션요지

- UN은 행정자료 활용의 당면과제, 행정자료와 SBR (statistical business registers), 행정자료와 무역통계, SBR과 무역통계의 연계에 대해 설명
- 행정자료는 통계생산의 저비용, 응답부담 경감 등 많은 이점이 있으나 법제도적, 데이터 표준화, 단위.변수 상이 및 시의성 문제 등이 존재
- SBR 자료는 기업 등록/면허, 조세자료: VAT, 근로소득세, 기업/협회, 상공회의소 등록, 사회보장 등록, 고용 등록, 행정기관별 회계자료, 비영리기구 등록자료, 협회등록, 농업사업체 등록, 상수도/전력 등록, 학교, 병원 등록부, 중앙은행의 금융분야별 자료, 외국인 직접투자 감독자료, 주주명부, 보

조금 정보 등으로 구성

- 상품거래 무역통계(IMTS; International Merchandise Trade Statistics) 생산에 활용되는 행정자료(세관신고서, 화물송장/ 송장기록, 국제거래 신고시스템, 부가세)를 예시하고 전력자료를 결합한다면 좋은 통계를 서비스 할 수 있음을 설명
- 남아공이 조세행정자료 활용 개선 효과에 대해 설명하고, 조세자료 이용의 현대화는 근로자를 고용할 경우 고용주의 신고의무를 강화하여 데이터 품질, 시의성, 완성도를 크게 개선시키는 효과 발생, 특히 납세자의 경제활동 실태를 예전보다 정확하게 보여줌
- 뉴질랜드 조세자료는 공식통계의 작성에 광범위 하게 활용되고 기회요인은 뉴질랜드 통계청 및 다른 기관에 공급, 조세자료 통합데이터시스템을 통해 자료 접근 가능, 비밀 보호 파일 생산, 국제시스템으로 집계된 통계데이터는 정부데이터 포털로 바로 공개, 관련 연구 및 분석 프로젝트에 지원이고, 주요과제로는 용어정의/메타데이터 관리, 데이터 품질제고, 고객중심 데이터 생산, 통계적/분석적 기법 개발 등을 설명

○ 세션시사점

- UNSD의 코스타리카 기업등록부와 무역 DB 연계, 남아공.뉴질랜드에서의 조세자료를 활용한 경제통계 작성 등 행정자료(조세자료 등) 활용사례가 증가. 확대됨에 따라 우리나라의 경우에도 통계의 완성도를 높이는 다양한 활동 필요

13. (IPS 101) 통계 및 공간지리정보의 통합을 향하여

○ 세션요지

- 중국은 총리가 위원장으로 각 부처 및 행정통계국이 참여하여 100억 위안이 소요된, 제1차 국가지리센서스를 2013년부터 2015년까지 실시하여, 토양 특성, 자원, 환경, 지질안전, 토지인프라, 공공서비스 등을 결합한 공간정보를 개발함. 중국은 지리센서스 결과를 통해 지리정보와 통계정보를 잘 활용한다면 세상을 스마트하게 관리가 가능하다고 설명하고, 모든 센서스는 지역기반으로 실시하여야 하고, 세계는 통일된 스마트한 지도가 필요함을 역설하고, 2020년에 전세계적인 세계지리총조사를 할 필요가 있음을 주장
- 멕시코 INERGI는 지리통계정보의 생산을 위해 1983년에 설립하였으나 미국, 중국 등의 위성보다 소형의 위성에 의존하여 정확한 정보의 생산에 한계가 있다고 토로

○ 세션시사점

- 참석자들은 대부분 지리정보 및 인구센서스 등의 정보를 결합하여 다양한 공간 정보의 생산에 동의하고, 공간정보와 행정정보를 활용한 센서스 등의 각종 통계의 개발에 적극적인 바, 우리나라에서도 현황 및 발전방안 대한 검토 필요

14. (IPS 116) 통계역량구축: 최근 경험, 수요, 도전

○ 세션요지

- 벨기에 루벤카톨릭대는 ISI와 관련기구는 전략적 계획 과정에 참여하고 있고 ISI는 통계역량구축의 센터 역할 수행을 설명
- 오슬로대는 수리통계 모형개발, 개발도상국 박사학위 과정, 박사학위 과정의 장거리 멘토링 통계역량 구축사례 설명하고, 이태리의 경우 가족중심 교육, 낮은 생활수준 등의 환경에서 멘토링을 통해 국제적인 EU펀딩 국가로 발전하였음을 설명
- 버지니아대는 개발도상국에 대한 통계협력역량구축 추진사항 설명, 특히 LISA 2020은 비통계인과 협력하여 개발도상국의 통계인 및 데이터 사이언티스트를 대상으로 교육 및 훈련을 실시하는 통계분석제휴실험실 프로그램임을 설명하고, 교육 후 자기 나라에 귀국, 고국의 대학에 통계분석제휴실험실을 만들게 하고 연구원, 공무원, 중소기업, NGO 들에게 통계를 통한 의사결정능력을 갖출 수 있도록 하여, 2020년까지 전세계적으로 20개의 실험실 네트워크를 구축 실시, 버지니아 공대의 LISA에는 LISA 2020 Fellow, LISA2020 자문위원회, 2020 멘토링 네트워크 등 운영 설명
- ISI는 최근 한국 대전 등에서 High level Workshop을 통한 통계역량구축회의를 개최한적이 있고 이를 통해 공식통계의 미래 리더를 발굴하고 혁신적이고 강한 커뮤니티를 구축하여 POST-2015 개발수요 발굴에 협력중임을 설명

○ 세션시사점

- ISI, 벨기에루벤즈대, 오슬로대 등에서 개발도상국 통계역량강화사업의 추진사례를 소개한 바, 우리나라도 유관연구기관 등과 협력하여 통계와 데이터 전문가를 양성하고 연구원, 공무원, NGO 등을 대상으로 통계활용능력을 강화하는 프로그램 개발 검토 필요

15. (IPS 050) 방법론과 기술이 변화하는 시대의 센서스

○ 세션요지

- UNSD에서 조사한 2010년 라운드 센서스 조사결과 214국이 참가하고 이중 2011년에 60개국이 참가하였음. 금번 조사 다양한 대체자료 활용, 행정자료 및 연간조사 자료 등을 활용하여 효율성을 증진시키고, 면접방식에서 전화조사, 인터넷, 등록조사 등 다양한 조사방식이 증가되고 있는 것을 특징
- 네덜란드 센서스는 1829년에서 1889년까지 내무부에서 실시하다 1899년 이후로 통계청에서 주관하고 있음. 비용절감, 사생활 보호, 무응답 증가 등 환경변화에 직면하여 1981년 이후 Virtual 센서스로 전환하였음. 현재 2020년 센서스 준비를 위해 교육정도, 소프트웨어 개발, 지리정보 개선 등 다양한 프로젝트를 진행중에 있으며 가장 중요한 사항은 제도기반 구축 및 믿을 수 있는 등록자료 구축 및 유지를 강조
- 호주통계청은 1911년 이후 센서스를 실시해 오고 있고, 전통적인 센서스는 적정성, 시의성, 예산제약, 기술발전 등 변화요인으로 인해 조사방식 변화 필요성을 인지함. 바람직한 현대센서스는 센서스의 본질을 지키며 연간가구조사, 행정자료, 대체자료 원 발굴 등을 통해 해결(Solution)중심의 센서스로 변화되어야 함을 강조

○ 세션시사점

- 2020 라운드 센서스는 기존 조사방식에서 벗어나 등록센서스 기반으로 전환하는 국가가 많아 질 것으로 보이는 바, 한국의 2015 센서스를 많은 국가들이 예의주시
- 2015 등록 센서스를 성공적으로 마무리 하고, 이를 세계적으로 잘 전파하여 한국 통계청의 위상도 제고하고 통계기준을 선도하는 위치확보가 필요해 보임

16. (IPS 147) 인터넷 조사-과제 및 기회

○ 세션요지

- 인터넷조사의 개념, 현황, 방법, 미래에 대한 일반적인 내용을 설명하면서, 1970년 Internet 출현, 2004년 Facebook 대중화, 2007년 I-Phone 보편화로 인터넷 조사 환경 기반은 성숙되고, 인터넷을 통한 조사의 경우 응답자선정의 경우 비확률 및 확률방식으로 선정될 수 있음을 설명
- 확률방식 경우 online intercept survey, 목록기반표집, 확률기반패널 등이 있으나,

포괄범위 오차, 무응답 오차, 측정오차 등은 해결해야 할 과제

- 향후 인터넷조사는 정보화 시대에 맞추어 저렴하고 신속한 장점과 아울러 신뢰성을 높이기 위한 좋은 이론 및 추정모델의 개발 등 도전과제를 해결하여야 함

○ 세션시사점

- 인터넷조사의 경우 공식통계 작성의 보조수단으로써 이미 널리 활용되고 있는 바 모바일환경에 친숙한 방식으로 더욱 개선해 나갈 필요성이 있음
- 이용자 만족도 조사 등 기업의 마케팅 용도로 편의성, 비용절감 등 장점이 많아 그 활용도는 높아질 예정임 감안하여 공식통계에도 활용할 수 있는 모델구축, 조정 기법 개발 등 품질제고를 위한 노력을 지속적으로 실시

17. (PS 105) 표본조사: 과거, 현재와 미래 방향

○ 세션요지

- 표본조사의 이론 및 방법의 과거, 현재, 미래방향에 대해 두 발표자가 차례로 발표를 실시하고, Kier(1817), Laplace(1814)의 무작위 추출에 이어 1925년 ISI 위원회의 이론 및 실제 보고서, Bowley(1926)의 층화무작위 추출, Neymen (1934) 최적할당추출, 미국센서스국(1940-1960)의 층화 2단계추출, Hansen and Hurwitz(1946)의 비응답처리를 위한 2단계 표본 등 표본조사의 획기적인 발전을 도모한 이론을 간략히 소개함
- Wayne A. Fuller는 계산과 분석에 대한 개념설명을 시작으로 결측자료 처리방법인 대체근접치, Raking, 회귀분석, 불완전유사표본, Little Models등을 소개하고, 아울러, Hot Deck imputation(1947), Multiple Imputation(1976,1987), 소지역추정기법 등 현재까지 영향 주는 표본자료 보정기법 중요성 설명
- 표본조사가 직면 혹은 직면할 도전과제인 시의성 측면에서 자료수집방식 Web 활용 증대, 지리정보소프트 웨어 출현 및 활용, 보조자료인 행정자료의 증가와 추정의 어려움, 로봇을 활용한 Data 수집 및 Monster 자료 출현 등을 지목하고, 모바일 자료 수집, e-census 출현 등 기존의 현장조사방식에서 탈피한 새로운 자료수집방식의 출현에 어떻게 대응할 지에 대한 청객들의 추가적인 질의가 있었고, 신 조사방식에 함께 대응하고 고민할 시기임을 강조

○ 세션시사점

- 통계학에서 역사적으로 발전해 온 가구 및 사업체 단위 조사의 표본기법 정리, 무응답처리, 임putation 등 표본이론 대한 종합점검은 매우 유의
- 급격히 변화되고 있는 조사환경에 맞추어 통계인들이 현명하게 대처하는 방안을 공동으로 강구할 필요성이 대두되고 있음을 공감

IV 맺음말

1. 국제협력교류 무대로 적극 활용

- ISI는 2년 주기 통계올림픽으로 국제협력의 장(국제기구, 각국 통계청, 학계, 민간기관, 협회 등)으로 활용 가능
- 국제기구 인사로는 유럽통계처장, UNECE 통계국장, ILO.UN. WB 등 고위직 및 중견간부 등이 참석
- 독일, 네덜란드, 베트남, 아부다비 등 양자협력국가 및 ISI와 산하 협회 회장단과 민간기관(Google 등)에서도 참석

2. 통계인들의 국제무대 활동 역량 증강 필요

- 한국 통계청의 업무현황을 국제적으로 알리고, 세계무대에서 활동할 수 있는 직원들의 국제역량 강화 필요성 대두
- 통계청은 4편의 논문(IPS 발표 1, CP 3편)을 발표하였으나, 세션을 구성하고 토론이 가능한 인적 역량강화가 필요

3. Big Data

- 빅데이터⁴⁾를 활용한 주요사례는 에스토니아 통계청과 구글임
- 에스토니아 통계청은 모바일 로밍자료*를 활용하여 국제수지 및

4) Big Data는 통계목적이 아닌 전자매체로 생성된 비표본화, 대량 등의 성질을 지닌 행정 및 민간자료를 의미

관광통계 편제의 기초자료로 활용하고 있음

* 방문객수 및 체류일수 추계에 Sim-Cards와 연계된 로밍 활동수와 첫 로밍일과 마지막로밍일을 활용

- Google의 경우 북미지역을 통해 **Consumer Surveys, Correlate, 시계열 분석기능을 탑재**하여 Big Data 활용을 실시하고 있음

○ 공식통계에서 **Big Data 활용 범위**를 어떻게 설정하는지가 관건

- 행정자료에 국한할지 민간자료로 확장할지에 따라 접근방식 상이
- 민간자료로 확장 시 **입수가능성, 파트너십 구축, 비밀보호 등 다양한 선결과제 해결**이 필요

○ Big Data 활용의 실용성 판단 확보 이전까지는 대체보다는 **보완수준**에서 접근하는 것이 합리적

4. 센서스

○ 센서스 방식은 **면접조사, 혼합조사(면접+ 행정자료원 등 대체), 등록 센서스(100% 행정자료 활용)** 3가지 방식으로 이행됨

○ **스페인**의 경우 2011년 10% Sample을 활용한 혼합조사방식을 성공적으로 수행하고, **2021년은 완전 등록 센서스로 이행 준비** 중

- 대규모(약 4,800만명) 국가의 첫 조사 없는 센서스 사례라 의미가 큼

○ **브라질**의 경우 2억 인구를 가진 대국으로 **2010 센서스를 Digital 방식 채택**(150,000개 스마트폰 활용) 하여 정확성 및 효율성을 제고

- 주거 정확성과 관련해 Digital 방식이 0.5%의 오차가 있는 반면 전통센서스는 5% 가량의 오차가 있음을 추정

- **조사환경 변화에 적극 대처하는 사례로 활용** 가능성이 있음

□ 세션개요

- 일시/장소 : 2015. 07. 27. 14:00 ~ 15:40 / Room 210
- 세션구성
 - Organizer/Chair : Renato Assuncao(Universidade Federal de Minas Gerais, Belo Horizonte, Brazil)
 - Presenter ① : Sudipto Banerjee(University of California, Los Angeles, USA)
"On Bayesian Modelings and Gaussian Process Models for High-Dimensional Geostatistical Datasets"
 - Presenter ② : Erica Rodrigues(Universidade Federal de Ouro Preto, Ouro Preto, Brazil)
"Exploring Multiple Evidences to Infer Users Location in Twitter"
 - Presenter ③ : Brian Reich로 예정되었으나 부친상으로 불참함
"Policy optimization for dynamic spatiotemporal systems"
 - Discussant : Gilberto Camara(Big data working group, Census Bureau, USA)
- 참석인원 : 약 190명

□ 세션요지

- 계층적 공간과정모형(hierarchical spatial process models)은 통계 기법들(statistical tools)을 효율적으로 이용하여 공간변수의 복잡성을 잘 설명하고 있으나, 이 모형을 추정하기 위해서는 복잡한 매트릭스 분해(matrix decomposition)가 필수불가결했음. 이에, 실질적으로 계산을 용이하게 하는 “잘 정의된 공간확률과정”을 고안하기 위해 ① 저차원 하위공간(lower-dimension subspace)에서 적용할 수 있는 Low rank process와 ② 정확한 베이지언 추론(Bayesian inference)를 하기 위하여, 다양하고 유연한 계층모형 구조 속에 결합되어 있는 사전인자(dimension-reducing prior)로써 활용될 수 있는 최근접 가우스 과정(Nearest-Neighbor Gaussian Process)을 소개하고, 미국 산림통계 데이터베이스의 자료를 이용하여 생물체량(biomass)의 공간분포에 대한 추론 사례 설명함
- 온라인 사회관계망(SNS)은 지진이나 전염병과 같은 실시간 사건을 모니터링하기 위한 중요한 정보원천(resources of information)이며 특히 이들의 행동 특성은 공간위치 파악에 중요한 정보를 제공해 줄 수 있으므로, 트위터 사용자의 공간위치를 추정할 수 있는 방법에 대해 논의함. 기존의 방법과는 달리, 지리정보 파악을 위해 두 가지 정보원천(친구관계망과 이용자가 작성한 글)을 사용하는데, 친구관계를 나타내주는 그래프 모양으로 구성된 트위터 글과 지리적 표시를 결합

하여 모형을 만든 확률적 접근법(①네트워크는 마코브 임의확률모형(Markov random field probability model)을, ②사후확률분포 추정은 마코브 체인 몬테카를로 시뮬레이션 기법(Markov chain Monte Carlo)을 사용)을 제안하였으며, GPS에 의한 위치정보를 이용하여 트위터 이용자들의 대용량 자료(dataset) 속에서 모형의 정확성을 시험하였음

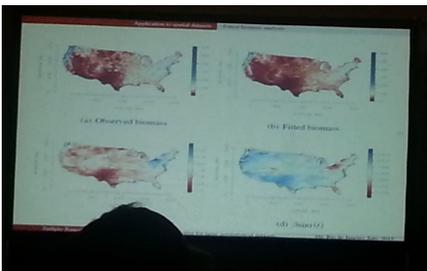
- (Discussant 논평) 빅데이터 분석의 경우 불확실성을 줄이고 정확성을 추구하는 것이 중요하므로, 두 논문 모두 측정오차에 대한 문제를 간과해서는 안 된다는 점을 지적하고 싶음. 전자는 Low rank process의 편의(bias) 문제를 논의하지 않았는데 이에 대한 해결이 필요하며, 후자는 측정오차 문제와 함께 표본이 큰 대도시에서만 사전인자(prior)를 사용해야 하나, 소도시에서도 사전인자를 사용한 것은 적절하지 않음.

□ 시사점

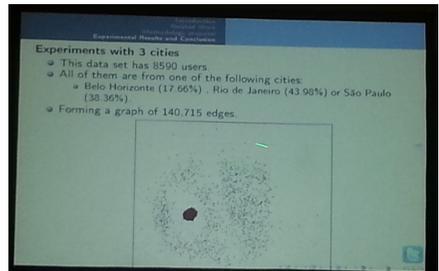
- 공간확률과정의 두 접근법 모두 많은 공간위치들 속에서 선형을 이루는 부동산수준 연산(floating point operations)을 가진 마코브 체인 몬테카를로 알고리즘(Markov chain Monte Carlo algorithms)에 수렴함
- 트위터 글과 지리적 표시를 결합하여 모형을 만든 확률적 접근법은 정보의 두 가지 유형 중 하나를 사용하는 two baseline algorithms(Naive Bayes Model and Integrated Data Aggregation)에 비교되는데, 결과적 측면에서 이 방법이 baseline methods보다는 확연히 더 우수함

□ 첨부자료

미국 산림 DB를 이용한 공간분포 추정



트위터 사용자 위치를 이용한 추정



□ 세션개요

- 일시/장소 : 2015. 07. 27. 16:00 ~ 17:40 / 203A Room
- 세션구성
 - Organizer/Chair : Yichao Wu / Dungan Liu
 - Presenter ① : Ming_Yen Cheng(National Taiwan University, Taipei, Taiwan)
"Forward variable selection for sparse ultra-high dimensional varying coefficient models"
 - Presenter ② : Ping Ma(University of Georgia, Georgia, USA)
"Leveraging in big data regression"
 - No Discussant
- 참석인원 : 약 20명

□ 세션요지

- 계수 모형(Coefficient models)을 다양화하는 것은 학술편야의 많은 영역에서 적용을 할 수 있으며, 공변량 동적 충격(dynamic impacts of covariates)를 모형화하는데 매우 유용함. 빅데이터 시대의 도래로 수많은 데이터가 범람하는 가운데 관련 변수를 선택하는 것은 쉽지 않은데, 최근 희소성 가정(sparsity assumption)하에 수행된 몇몇 연구가 있지만 아직 일부 한계점이 있음.
본 논문에서는 선행변수 선택절차(forward variable selection procedure)를 소개하는데, 이는 희소성 가정하에서 제곱 합 기준(sum of squares criterion)에 따라 중요 변수를 연속해서 선택하고 EBIC⁵⁾-또는 BIC⁶⁾-기반 정지 규칙(EBIC- or BIC-based stopping rule)을 적용하는 것으로, 실행하기 쉽고 계산도 빠르며 이론 및 수리적 관점에서도 적합한 특성을 가짐. 일부 유연한 규칙 패턴 조건하에서, EBIC나 BIC 중 하나를 정지 기준(stopping criterion)으로 사용하면 확실히 선택일관성을 갖춘 결과가 도출되는데, 시뮬레이션과 실증분석(에스트로겐 수용체를 인덱스 변수로 사용하여 암 크기 추정)을 통해 효율성과 유용성을 입증함
- 빅데이터에서 유용한 정보와 지식을 추출하는 것은 과학부문 뿐 아니라 전체 사회 측면에서도 부담이 되는 과제이며, 이를 위해서는 빅데이터 분석에 있어 효율적인 계산방법 및 통계적 접근방법 측면에서 괄목할 만한 성과가 필요함

5) Extended Bayesian information criteria(EBIC)

6) Bayesian information criteria(BIC)

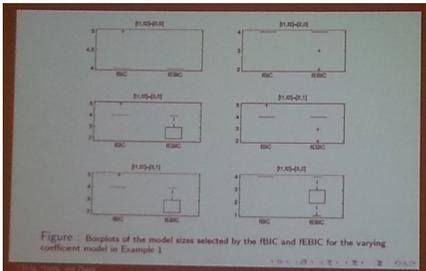
이러한 엄청난 과제를 해결하는데 있어 중요한 기여를 하는 레버리징 알고리즘(leveraging algorithms)을 소개하고, 컴퓨터로 조정되는 알고리즘들을 사용하여 아주 작은 대표적인 하위 데이터셋(sub-dataset)를 샘플링함으로써, 작은 하위 데이터 셋에서 방대한 데이터셋 관련 정보를 효율적으로 추출할 수 있으며, 이러한 알고리즘(leveraging algorithms)들은 빅데이터로 확장가능함. 이러한 노력들은 빅데이터 분석으로의 전반적인 접근을 가능하게 해 주는데 특히 직접 슈퍼 컴퓨터를 사용할 수 없는 사람들에게 유용함

□ 시사점

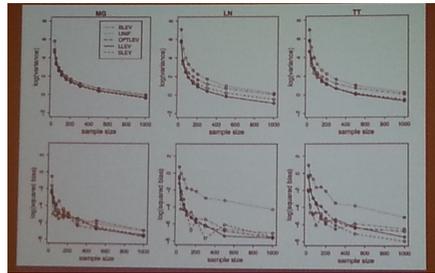
- 선행변수 선택절차는 기존 방법들과는 달리 선택일관성을 검증하기 위한 추가 확인단계 (extra screening step)가 필요 없으며, 비규칙적인 패턴 조건일지라도 다소 덜 제한적인 구조 (less restrictive setup)라면 효과적인 확인 절차(effective screening procedure)로써 유용함
- 레버리징 알고리즘(leveraging algorithms)은 효율적인 데이터 마이닝을 가능케 하므로, 일반 사용자가 태블릿컴퓨터(tablet computers)를 이용하여 빅데이터 분석을 할 수 있는 등 빅데이터 분석 대중화에 기여할 것임

□ 첨부자료

BIC & EBIC에 의해 선택된 모형크기 비교



Leveraging algorithms



□ 세션개요

- 일시/장소 : 2015. 07. 29. 10:30 ~ 12:30 / 201AB Room
- 세션구성
 - Organizer/Chair : Siu-Ming Tam / Ray Chambers
 - Presenter ① : Paul Schubert(Australian Bureau of Statistics, Australia)
"Big data Inferences and Modelling for Official Statistics, with Application to Satellite Imagery Analysis"
 - Presenter ② : Giulio Barcaroli(Istat, Rome, Italy)
"Use of Big data in official statistics"
 - Presenter ③ : Ronald Prevost(U.S. Census Bureau, USA)
"REVEAL: A Paradigm for Official Statistics"
 - Discussant : EUROSTAT staff
- 참석인원 : 약 120 명

□ 세션요지

- 주로 빅데이터는 포괄범위 편익(coverage biases), 자기선택 편익(self-selection biases), 대표도 편익(representation biases), 응답편익(response biases) 또는 측정오차(measurement errors) 문제에 봉착해 있는데, 이러한 비표본 오차 문제를 해결하기 위해 빅데이터 추론을 위한 베이저언 구조(Bayesian framework)-빅데이터 측정에 적용된 ①개념 변형(conceptualized transformation), ②샘플링 및 ③검증과정(censoring process)-에 대해 논의하고자 함. 적합한 추론은 세 개의 구조 모두 모형화해야 하며, 이 모든 것을 모형화하는 것이 가능하다면 매우 복잡하고 까다로울 수 있지만, 어느 정도의 샘플링 및 검증 무시성 조건(sampling and censoring ignorability conditions)이 충족된다면 빅데이터 측정에 대한 추론이 가능함. 베이저언 구조를 위성이미지 자료 분석에 적용한 사례로 농작지와 농작물 생산량을 예측을 설명하기 위해, 농작물 유형에 따른 이항 로지스틱 회귀분석 모형(binomial logistic regression model), 농작물 생산량을 위한 로그 선형 모형(log linear model) 그리고 두 모형의 회귀 계수(regression coefficients)를 위한 자기회귀모형(autoregressive model)을 이용하였음
- 고품질 및 다양성을 갖춘 데이터가 요구되는 현대사회에서 예산부담 및 응답부담 관련 비용은

조사 범위를 제한하므로, 공식통계에서 빅데이터 이용과 새로운 데이터 원천을 활용하는 것은 정보와 지식의 충분한 공급원이 될 수 있음. 데이터 수집과 추정의 전통적 방법을 대체하고, 조사 통계와 인터넷 데이터를 모두 이용하여 추정치를 작성하는 통합 접근법으로의 결함 목표를 가지고, 이탈리아 통계청은 웹자료수집(web scraping)과 문자 마이닝 기법(text mining techniques)을 이용한 통계작성을 위해 “기업 정보통신기술(Information and Communication Technology in enterprises)조사”를 실시함. 그리고 빅데이터 잠재력을 평가하기 위한 통합연구프로젝트에서, 이탈리아 통계청은 노동시장 지표(Labor Market Indicator)와 구글 트렌드 지표(Google Trends Indicators)를 추정하여 작성했으며, 현재 진행 중인 실험은 “사람과 지역(Persons & Places)”으로, 핸드폰 자료(mobile phone tracking data) 사용을 근거로 하여 지방자치단체 공간 단위로, 업무용과 학습용 일일 이동성(daily mobility)의 기원/출발 매트릭스(origin/destination matrix) 작성을 목표로 함. 개개인의 존재와 이동성의 대용물(proxy)로서 빅데이터를 사용하여 도출된 결과는 공식 자료와 함께 수집된 유사통계와 비교해 보았을 때, 고무적이고 매우 조짐이 좋음(promising)

- 빅데이터는 공식통계작성기관에 있어 도전과 위협이 되는 특별한 과제(challenges)임. 통계 이용자들은 그 어느 때보다도 보다 신속한 통계, 보다 다양한 통계를 요구하고 있는 실정이며, 빅데이터가 그 해답이 될 수도 있지만, 정확성 문제의 소지가 있음. 이러한 맥락에서 공식통계작성기관이 정책관련 정보를 가장 잘 제공할 수 있는 방법은 무엇인가? 조사나 총조사(census)와 같은 전통적인 접근법은 항상 시의적절하게 통계를 생산하는 것은 아니지만, 품질 측면에서 지속적인 단기예측을 위한 벤치마킹의 기초를 제시하므로, 공식통계를 위한 패러다임으로써 “REVEAL”을 제안하며, 그 과정은 다음과 같음. 인터넷과 행정자료와 같이 잘 알려진 자료원천으로부터 정보를 재수집(Reassemble)하는 것, 사전 조사나 총조사 같은 기준점(benchmark)으로부터 추정(Estimate)하는 것, file operation을 가지고 추정치를 확인(Validate)하는 것, 측정 차이를 설명하고 평가(Evaluate)하는 것, 오차 원인(sources of error)을 분석(Analyze)하고, 향후 이러한 자료 원천(data sources)을 사용할지 여부를 결정하는 것, 예측 방법론을 개선하기 위해 품질측정과 추론(inference)을 재 연계하는 것. 기존 조사들을 검토하기 위해 추정치를 사용하기보다, 조사는 추정치의 적합성을 확인할 수 있는 것이어야 함.
- (질의응답) 이탈리아 통계청 실험은 매우 흥미로운데 정확성 문제는 없는지? 편의(biases) 문제는 항상 염두에 두고 있으며, “기업ICT조사” 결과를 검토하면서 편의를 줄일 수 있는 방법을 모색하고 있음
- (Discussant) 기존 통계는 공식통계 작성기관에서 생산하였기에 자료 접근이 용이했고, 잘 정돈된 데이터였으나, 빅데이터의 경우 ①편의 문제 ②reliability and changes(핸드폰사용 및 사용자 수가 계속 변화하므로 자료원천에 대한 문제), 정확성 및 적시성 문제 그리고 ③빅데이터의 강점인 지리공간적 연결을 잘 활용하여, 기존의 자료원천 및 정보와의 연계 문제를 잘 해결해야 함

□ 시사점

- 공식통계로서의 빅데이터의 기능성과 함께, 일반적인 논의사항(품질관리, 정확성 및 적시성, 자료 연계, 프로파일링 방법 및 시각화 도구) 뿐 아니라, 최근 대두되고 있는 프라이버시 문제(법적 구조, 윤리적 가이드라인 및 사생활 보호를 위한 technology solution 등) 그리고 IT 이슈(cloud computing 및 machine learning)는 시급히 선결해야할 과제임
- 공식통계작성기관에서 빅데이터를 통계작성 방법으로 사용하기 위해서는 기존 조사의 장점들을 적극 활용하여 새로운 통계작성 패러다임으로의 변화를 도모해야 함

□ 첨부자료

호주 통계청 발표자



EUROSTAT 토론자



참고 4

(IPSO26) How can National Statistics Offices(NSOs) take advantage of new data sources and collection modes?

□ 세션개요

- 일시/장소 : 2015. 07. 29. 14:00 ~ 15:40 / 101 Room
- 세션구성
 - Organizer/Chair : Nancy Gordon(불참) / Nancy Mcbeth
 - Presenter ① : Tuulikiki Sillajoe(Statistics Estonia, Estonia), Jaanus Kroon(Bank of Estonia)
"Use of mobile devices and other advances in data collection"
 - Presenter ② : Ramachandran Ramasamy(National ICT Association of Malaysia, Malaysia)
"Meeting users' needs for indicators on the knowledge-based economy(KBE), while reducing collections on the agro-industrial economy"
 - Presenter ③ : Duncan Young(Australian Bureau of Statistics, Australia)
"Using New Data Sources to Modernize Official Statistics"
 - Presenter ④ : Andreas Mayer(Australian Bureau of Statistics, Australia)
"Big Data and Semantic Technology : A Future for Data Integration, Exploration and Visualization "
 - Discussant : Vince Garvin(Statistics New Zealand, New Zealand)
- 참석인원 : 약 200 명

□ 세션요지

- 정보통신기술(ICT)의 발전과 함께 새로운 자료 원천(data sources)이 많은데, 이러한 자료 원천은 새로운 통계 작성과 기존 통계의 품질향상 측면에서 공식통계 작성에 사용될 수 있음. 새로운 자료 원천을 사용하고자 하는 에스토니아 통계청의 포부(aspirations)를 설명하고, 자료 접근 및 자료처리를 위한 전제조건들과 법적 측면(legal aspect)을 제시한 후, 관광통계 작성을 위해 위치동자료(mobile positioning data) 사용에 중점을 두었던 국제 컨소시엄 결과를 간략히 소개함. 2008년, 에스토니아 은행은 위치동자료(mobile positioning data) 연구 경험이 있는 연구원들과 공동으로 새로운 자료수집 방법론을 개발하기 시작했고, 그 결과 개발된 방법론은 에스토니아 국경을 건너 해외로 여행하는 사람들과 에스토니아로 입국하는 사람들을 대상으로 작성한 통계에서 신뢰성을 입증받았으며, 2012년 이후 에스토니아 은행은 공식적으로 국경 여행 통계(cross-border travel statistics)를 작성하여 공표하고 있음.

- 공식통계시스템하에서 지식기반경제(Knowledge Based Economy)지표를 전반적으로 분석할 필요가 있음을 설명함. 공식통계이용자의 수요는 최근 몇 년간 급격히 변화하고 있으며, 정보통신기술(ICT) 혁신은 우리 삶 전반에 걸쳐 상당한 영향을 주었는데, 이러한 변화는 공공정책 입안을 어렵게 할 뿐만 아니라 통계 개념 및 정의 그리고 자료 적합성에 의문을 제기함. 지금의 공식통계시스템은 농업 및 산업현장 그리고 정보화 시대 성장을 설명하는데 있어 타당성 및 포괄성 부족에 근거를 두고 있음. 인터넷 등 광범위한 커뮤니케이션은 지식기반경제의 규모를 계속 증대시키고 있으며, 시·공간적 제약없이 디지털 프로세스와 콘텐츠를 강화하고 있으므로, 지식기반경제를 측정할 수 있는 새로운 통계조사들을 수용하는 것을 포함하여, 現 공식통계시스템을 면밀히 검토하고, 재정비해야함을 말레이시아 지식기반경제 통계 경험을 통해 설명함
- 새로운 자료원천(data sources)의 대두는 공식통계를 위한 많은 전략적·방법론적 논의를 제기했음. 지식, 전문성 그리고 기법과 방법론을 공유할 때, 효율적인 협업(collaboration)은 공식통계산업이 빠르고 효과적으로 현대화하는데 있어 매우 중요함. 통계생산 및 서비스를 위한 고위급 그룹⁷⁾(HLG)는 통계생산과정의 효율성과 사용자 니즈를 충족시키는 결과를 생산하는 능력을 향상하는 것을 목표로 하고 있으며, 그 핵심 사업⁸⁾ 중 하나가 공식통계에서의 빅데이터 프로젝트임. 이에 공식통계 현대화를 위한 HLG 활동이 새로운 자료원천 사용의 효과적인 협업을 어떻게 지지하는지를 호주 통계청 사례를 통해 설명함. “Modernization Committee on Products and Sources”는 빅데이터 인벤토리, 모바일 장치, 행정자료, 연계된 공개 자료(Linked Open Data), 혼합모드(Mixed Mode), 커뮤니케이션의 성과를 이루었고, 새로운 자료 원천과 관련된 공식통계 현대화 노력에 대한 향후 과제로써 품질(quality), 개인정보 보호(privacy) 그리고 동반자관계(partnership)를 언급함.
- 이용가능한 데이터의 증가와 이용자들의 요구 증가로, 국가통계작성기관들은 관련성과 적시성을 갖춘 통계를 생산해야 하는 난관에 봉착하였고, 통계가치를 극대화하기 위해 비전통적인 자료원천과 조사 자료원천을 통합함으로써 그리고 고급분석을 할 수 있는 새로운 기술(technology)을 활용함으로써, 빅데이터 특히 행정자료의 장점을 이용하기 위하여 업무 관행을 변경할 필요가 있음. 호주통계청에서 의미기술(semantic technology)을 사용하여 개발하고 있는 시험용 GLIDE(prototype Graphically Linked Information Discovery Environment)가 바로 그 사례이며, 이 환경은 국세청 자료와 세부적인 미시분석이 가능한 ABS 사업 등록 자료를 통합한 시험용 고용주-피고용주 연계 DB(Linked Employer-Employee Database)를 포함하고 있음. 그러나 데이터 구조가 보다 복잡해지고 고차원화(multi-dimensional) 됨에 따라, 의미기술은 데이터 셋 상에서 유연한 데이터 구조, 재사용 가능한 분류 및 기준, 네트워크 분석, 기계적 추론 및 추론(machine reasoning and inference)을 가능하게 함. 어떻게 시험용 GLIDE가 전통적인 데이터 탐구(data exploration)와

7) High-Level Group for the Modernization of Statistical Production and Services

8) Generic Statistical Business Process Model (GSBPM), a standard Generic Statistical Information Model (GSIM), a Common Statistical Production Architecture (CSPA) and project on Big Data in Official Statistics.

시각화를 보다 더 효율적으로 만들어 주는지, 그리고 시험용 GLIDE가 어떻게 새로운 네트워크 분석을 하는지를 사례를 통해 설명함

- (Discussant) 경직적(rigid) 통계작성 프로세스와 방법, 전통적인 자료수집 방법의 비용증가 그리고 디지털 시대 도래로 인한 데이터 범람 등 통계의 현대화는 당면과제임. 이에 새로운 자료원천을 활용 방법으로 ①공식통계작성기관들이 협력하여 함께 문제를 해결하는 것, ②새로운 자료원천을 사용했던 우수사례를 공유하는 것 그리고 ③계속해서 변화를 반영하고자 하는 지속적인 노력임을 각각의 발표자료에 대해 논평하면서 언급함
- (질의응답) 에스토니아는 데이터 접근성과 confidentiality에 중점을 두고 있으며, 말레이시아는 빅데이터 분석을 위해서는 통계지식, 비즈니스지식 그리고 technology지식이 필요함을 강조.
 Q : 빅데이터는 개념 정의가 불명확한데 호주에는 빅데이터 인벤토리가 있음. 빅데이터 인벤토리의 데이터 수집 방법과 데이터 접근성 문제는 어떻게 해결하냐?(by IAOS회장 가와사키 시게루)
 A : 빅데이터 정의는 어렵고 다양한데, 이것은 통계인들이 해결해야 할 문제이며 그 정의에 따라 수집 방법이 달라짐. 접근성 문제는 국가마다 그 상황과 구조가 상이하므로 내부적(domestic)으로 해결해야 할 문제이며, 데이터 사용은 데이터에 접근해 있는 동안 해결할 수 있어야 함(by Duncan Young)

□ 시사점

- 급변하는 시대적 변화를 반영하고 이용자들의 니즈를 충족시키기 위해, 새로운 자료원천을 공식통계 작성에 사용할 수 있는 방법을 모색할 필요가 있음
- 새로운 자료원천을 통계생산 및 서비스의 현대화에 활용하기 위해 공식통계작성기관들은 서로 공조하여야 하며, 기존 통계와 새로운 자료와의 통합을 통해 보다 부가가치가 높은 통계를 생산해야 할 시점임

□ 첨부자료

세션 구성원들



IAOS 젊은 통계인상 수상자



□ **세션개요**

- 일시/장소 : 2015. 07. 31. 14:00 ~ 15:40 / 202B Room
- 세션구성
 - Organizer/Chair : Chellam Panyandy / Oliver Chinganya
 - Presenter ① : Lesile TANG(Census and Statistics Department, Hong Kong)
"Big Data Applications for Official Statistics: Methodologies and Challenges"
 - Presenter ② : Ronald Jansen(United Nations Statistics Division, USA)
"Challenges of using Big Data for Official Statistics"
 - Presenter ③ : Olivier Dupriez(World Bank, USA)
"Big Data and the Modernization of Statistical Systems in Developing Countries"
 - Discussant : Fred Gorden(Haishan Fu 대행)
- 참석인원 : 약 100 명

□ **세션요지**

- 현대사회가 점점 더 복잡다단해지면서 보다 더 높은 수준의 관련성과 시의성을 갖춘 새로운 통계에 대한 수요가 증가하고 있음. 전통적인 조사 방법론은 목적을 가지고 이미 만들어진 자료를 수집하는 것이므로 목적부합성은 높을지라도, 정부와 응답자의 부담은 상당함. 동시에 공식통계 작성을 위한 빅데이터 적용이 그 어느 때보다도 강조되고 있음. 다양한 빅데이터 원천들 중에서도 행정자료는 공식통계생산에 가장 적합하고 관련성이 높은 것으로 간주되는데, 이는 행정자료가 조사자료보다 비용대비 효율적이고, 적시성을 갖춘 자료원천이기 때문임. 이러한 이유로 홍콩통계국(Census and Statistics Department of Hong Kong)은 정책입안을 뒷받침하기 위해 지역 의료서비스 케이스 스터디를 통해 행정자료의 적극적 활용의 실현가능성(feasibility)을 포괄범위의 완전성 및 신뢰성 측면에서 검토 하였음. 이에, 새로운 통계 생산에 있어 행정자료 활용을 개괄적으로 살펴보고, 분석에 적용한 방법론을 설명한 후, 그 과정에서 직면한 장단점 그리고 행정자료 이용의 어려움(데이터 소유자로 하여금 데이터 제공에 대한 확신부여 방법 및 데이터 프라이버시)을 극복할 해결책(법령 수정, 정부 주도의 데이터 표준화 등)들을 논의함.
- 신속성, 다양성 그리고 대용량의 특성을 가진 자료 원천인 빅데이터의 이용가능성은 정보 기술의 급속한 발전과 전자기기의 대중적인 보급과 함께 계속해서 증가추세이며, 이에

빅데이터에서 의미있는 정보를 찾기 위해 새로운 툴(tools)과 방법들이 개발되고 있음. 공식통계를 위한 빅데이터 UN 작업반 그룹(UN Working Group on Big Data for Official Statistics)은 빅데이터에 대한 범세계적 업무 프로그램(global work program)을 제공하고, 빅데이터 원천의 실질적 사용을 도모하며, 공식통계를 위한 인간의 빅데이터 사용에 공적 신뢰를 구축해야 할 의무가 있기 때문에, 현재 이러한 작업⁹⁾이 진행 중 임. 빅데이터는 GPS 위치 추적 장치, 위성 이미지, 전자상거래, 핸드폰, 소셜 미디어, 스마트 계량기(smart meters)와 같은 광범위한 자료원천에서 비롯되는데, 이는 우리 일상생활의 거의 모든 분야에서 대량으로 발생하고 있으므로, 빅데이터는 공식통계 및 포스트 2015 개발 의제(post-2015 development agenda)를 모니터링하기 위한 의미있는 지표 전달에 사용될 상당한 잠재력을 가진 자료로 간주됨. 그러나 빅데이터 장점을 인식하기 전에 선결해야 할 몇몇 과제(several challenges)들을 체계적으로 검토하고자 함. UN 빅데이터 전략 및 프로젝트 목록 조사(United Nations Survey on Big Data Strategy and Project Inventory)와 관련 문헌을 토대로, 국가통계시스템이 공식통계에 빅데이터 사용시 직면하게 되는 **주요 난제**(①데이터 접근성, ②프라이버시와 신뢰, ③빅데이터 품질관리 제도, ④빅데이터 프로젝트 이행에 있어서의 IT 요건 그리고 ⑤비용을 강조하면서, 이러한 난관을 해결할 수 있는 방법으로서 ①민·관합작(public-private partnerships)과 ②데이터 접근을 위한 원칙 개발과 법규와 제도를 통한 프라이버시 보호 및 신뢰유지, ③비통계작성 목적으로 작성된 빅데이터의 품질을 위해 새로운 방법론과 품질프레임워크(ex. UNECE의 sandbox)개발, ④ 기존 데이터와 빅데이터 통합을 위한 IT(cloud computing 및 big data analytics(Hadoop, R, SAS)) 그리고 ④통계작성기관들 간의 비용 공동부담과 정부지원을 언급함.

- 개발도상국에서의 사회경제 데이터에 대한 수요는 끊임없이 증가하고 있으며, 국가기관이 생산한 데이터의 이용가능성과 공공의 접근성(public access)은 점점 개선되고 있음. 그러나 데이터갭, 관련성, 적시성 및 정확성은 많은 개도국에 부담이 되기 때문에 “Data Revolution”이 요구되는 상황이며, 이에 공식통계작성자들은 빅데이터 사용에 많은 관심을 가지고 기대를 하고 있음.

아직 시험단계(pilot stage)에 있긴 하지만 일부 선진국 통계작성기관에서는 빅데이터 사용에 대한 그들의 경험¹⁰⁾을 공유하고 있으며, 이를 통해 전통적인 자료 수집 방법을 보완하고, 새로운 지표 작성에 있어서의 빅데이터 잠재력을 확인하였으며, 상당한 진입비용이 소외됨을 인지함. 개도국에서 공식통계를 작성하는 통계인들의 제약과 필요성을 중심으로 선진국의 경험을 통해 얻은 교훈을 설명함. 개도국은 혁신보다는 통계시스템 현대화가 요구되며, machine learning algorithms에 의존하는 빅데이터 적용에 있어 데이터셋은 중요한 input이기 때문에, 품질개선, 적시성, 정확성 및 전통적인 자료 수집(조사나 센서스)에 있어서의 공공의

9) World POP, United Nations University에서의 Bigdata technology 교육, Global Land Project인 FutureEarth, 구글 등과의 업무협약 등

10) 성공사례 및 실패사례, 위험이나 비용측면에서의 정보, 기술적 제약조건 그리고 법적 및 윤리적 장에 대한 정보 등

이용가능성(public availability)에 우선순위를 두어야 함.

- (Discussant) 빅데이터는 진입장벽이 높아 당면 과제(민간데이터로의 접근성, 일치성, 품질 및 공식통계와의 관련성, 프라이버시, 대용량 데이터 셋 저장 및 이용 능력, 직원 교육 등)이 많음. 공식통계작성기관은 데이터 공개와 이용가능성을 더욱더 강화해야 하며, 기존 데이터가 빅데이터와 호환할 수 있도록 변형해야 함. 그리고 모든 것을 지리좌표와 연계하고, 디지털 영역과의 결합을 통해 새로운 데이터의 유용성을 증가시켜야 함.
- (질의응답) 모니터링을 위한 빅데이터 적용 방법, 빅데이터 처리 및 quality 그리고 데이터 공개(openness)에 대한 질의에 대한 답변으로, 빅데이터를 위한 UN 작업반 그룹은 빅데이터와 지표들과의 관계를 모니터링하고 있으며, 현재 가계조사와 빅데이터의 교차적용 방법을 고려중임. 네덜란드 통계청의 경우 twitter를 통해 consumer sentiment data를 이용하여 consumer confidence를 모니터링 하고 있으며, 이와 같이 트위터나 페이스북은 모니터링을 위한 빅데이터 적용에 좋은 수단임(by Ronald Jansen from UNSD).
데이터 개방성은 공식 통계작성기관이 해야 하는 임무이며, 이는 엄청난 돈이 들기 때문에 쉽게 바뀌기는 쉽지 않음. 세계은행의 경우 영국 DP에서 자금을 지원받아 microdata set을 만들었으며, 이용자들이 접근하기 쉬운 dataset과 documentation을 만드는 것이 중요함(by Olivier Dupriez from World Bank). 말레이시아의 경우 통계 공개를 법령으로 명문화하여 데이터를 잘 활용할 수 있도록 함(Lesile TANG from Hong Kong)

□ 시사점

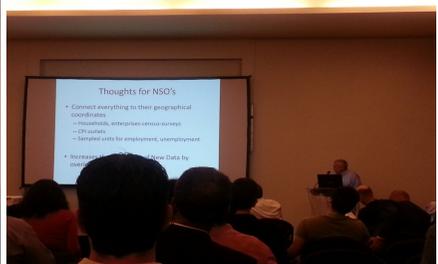
- 非통계작성 목적의 빅데이터를 공식통계 작성에 이용하고, 빅데이터의 유용성을 제고하기 위해서는 선결해야 할 많은 당면과제들이 있으며, 이러한 문제를 해결하기 위한 노력이 이미 국제기구를 중심으로 이루어지고 있기 때문에 각국의 공식통계작성기관들의 적극적 참여가 요구됨

□ 첨부자료

Ronald Jansen from UNSD



Fred Gorden filling in for Haishan Fu



□ 세션개요

- 일시/장소 : 7/28/2015, 09:00~10:10 Room 101
- 세션구성
 - 좌장(성명, 소속 등) : Jean-Louis Bodin
 - 발표자(성명, 소속, 논문제목 등)
 - . Privacy, Confidentiality, and Big Data(Stephen Fienberg, Heinz College)
 - 토론자(성명, 소속 등)
 - . Jean-Louis Bodin
- 참석인원 : 150명

□ 세션요지

- (발표내용)
 1. 센서스 자료(census data)의 변천 소개 : 1950년~2010년까지의 센서스 자료의 관리 방안 설명 및 2010년에는 결과를 아이폰에 적합하도록 제공
 2. 빅데이터(Big Data) : 다양한 자료의 결합(센서(sensors)), 비디오, 운전자 인구통계(demographic of drivers). 기타 등등 과 구글(Google), 페이스북(Facebook), 국가안정보장국(NSA:National Security Agency) 등의 자료
 - > 전통적인 조사나 센서스를 대체할 수 있음
 3. 개인정보보호 방안 : 개인 식별정보 제거, 센서스 자료의 지리적 병합(geographic aggregation)방법으로 해결, 표본(Sample)자료 제공(공적으로 마이크로데이터를 사용하는 경우에 한하여)
 - > 데이터 생성 프로세스 도입 : 잡음삽입(noise addition), 데이터 교환(data swapping), 사후-임의방법(post-randomized method), 병합(agggregation), 기타 등등
 4. 개인정보(Confidentiality), 노출(Disclosure), 위해(Harm)에 대하여 [공식통계에 대하여 위해(Harm)는 노출(Disclosure)의 "인지(Perceptions)"와 관련]이 있다고 하며, 데이터의 활용을 통해 위험을 줄일 수 있음
 5. 개인정보의 개념 : 유용성(Usability), 투명성(Tranparency), 이중성(Duality)
 - 유용성 : 데이터를 시스템 제약으로부터 자유롭게 활용할 수 있도록 확장
 - 투명성 : 노출제한(Disclosure limitation)으로 부터 편의(bias)와 변동(variability) 결과에서 직접 또는 암시적인 정보를 제공하는 방법론을 확장

- 이중성 : 노출제한(Disclosure limitation)과 거대한 양의 자료를 분석에 활용할 수 있도록 확장

6. 표준적인 정보보호와 개인보호방법

- 사례를 제거하는 방법, 병합하는 방법, 노이즈를 추가하는 방법, 데이터 스와핑 등 ($X \rightarrow Z = AXB + C$) : Duncan and Pearson(1991)

7. 차별적인 개인정보(Differential Privacy) 방법 : 입실론(ϵ)-Differential Privacy, 베이지안(Bayesian)방법 및 복합적인 대체(multiple imputation)를 활용
 사례) 현재는 베이지안 방법을 이용하여 표본의 자료를 기반으로 모수의 특성을 파악하는 방식으로 변천

모수 \rightarrow 표본 \rightarrow (표본|모수) \rightarrow (표본, 모수) \rightarrow (모수|표본)

○ (토론내용/질의응답)

실제자료(real data)에 대한 개인정보 노출, 정보보호를 위해 어떻게 해야 하는가?

- 실질적인 연습(잡음추가(noise addition), 데이터 교환(Data swapping), 기타 등등을 통해 적합한 방법을 지속적으로 적용하여 해결하는 방법이 있음

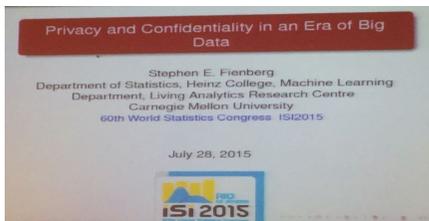
□ 시사점

○ 통계 이용자들은 지속적으로 마이크로 데이터에 대한 제공을 요청하고 있으나 통계작성기관에서는 정보보호와 개인정보 노출 등의 문제로 제공에 대하여 부정적인 의견이 많이 있음

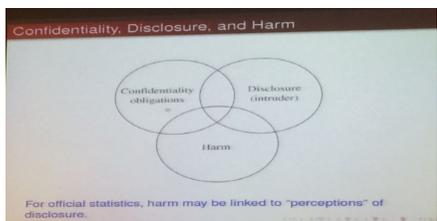
- 통계청도 개인정보 보호 및 정보노출을 줄이기 위한 다양한 방법(노이즈 추가, 데이터 결합 등을 통해 마이크로데이터 제공을 위한 연구를 지속적으로 실시하여야 할 것으로 사료됨
- 또한, 빅데이터(Big data)시대를 준비하기 위해 자료의 정보보호와 개인정보 노출에 대한 선진국의 방법 연구 필요

□ 첨부자료

발표자료



개인정보, 노출, 위해의 관계



참고 7

(IPS016) Measuring Progress, Sustainable Development and Wellbeing in a Post-2015 environment

□ 세션개요

- 일시/장소 : 7/27/2015, 10:30~12:30 Room 101
- 세션구성
 - 좌장(성명, 소속 등) : Ronald Jensen
 - 발표자(성명, 소속, 논문제목 등)
 - . Measuring Progress, sustainable Development and Well-being in a Post-2015 environment(Conal Smith, OECD)
 - . Measuring Australia's progress in a Post-2015 environment(Peter Harper, ABS)
 - . Measuring National Well-being in the UK: Four years of progress(Glenn Everett, Abigail Self, UK)
 - . Measuring Sustainable Development, Review and Reporting on a Post-2015 Agenda(Lidia Bratanova, UNECE)
- 토론자(성명, 소속 등)
 - . Ronald Jensen
- 참석인원 : 100명

□ 세션요지

- (발표내용)

영국 : 국가적인 웰빙을 어떻게 정의할 것인가, 어떻게 측정할 것인가, 궁극적인 목적은 무엇인가 등을 고민하고 있으며, (6개월에 한번씩 자료를 업데이트 하고 있으며, "Life in UK" 라는 책자도 발간

국가적 웰빙의 정의를 개인, 지역사회, 국가로써 '어떻게 우리가 하고 있는가'와 어떻게 미래에도 지속시킬 것인가로 정의하고 있으며, '어떻게 우리가 하고 있는가'에 대한 전체 그림을 위해 GDP와 같은 경제, 사회, 환경정보의 넓은 범위가 필요하다. 지속적인 발전 목표를 위해서 정부통계서비스(the Government Statistical Service : GSS)에 대한 리포팅과 모니터링의 지속적인 관리가 필요할 것이다.

호주 : 2002년부터 2013년까지 호주의 진보 측정(measure of Australia's progress : MAP)내용을 발간, 웰빙의 영역을 사회, 경제 환경, 정부, 4가지로 구성하고 있음
MAP의 참고를 위한 주요 수단으로는 온라인 블로그, 비디오, ABS 페이스북 등을

활용하고 추가적으로는 컨퍼런스, 전통적인 미디어, 전문가 그룹, 워크숍, 공식회의 등을 활용하고 있음

UNECE : 58개국의 회원국을 가지고 있으며, 8개국에 대하여 사전조사(pilot test)를 실시하였으며, 2016년에는 지속가능한 개발 목표(SDG : sustainable development goal)의 측정을 다시 시작하려고 하고 있다.

지속가능한 발전을 측정하기 위한 주요 메시지 4가지는 소리 개념적 프레임워크(Sound conceptual framework)와 통계인과 기타 다른 이해관계자들과의 대화가 필요하며, 통계청의 역할은 자료 수집보다는 국가통계 시스템의 관리일 것이며, 포스트 2015 아젠더는 통계적인 발전과 역량구축을 더 높이는 것이며, 새로운 데이터 소스 발전을 위해 데이터 혁명을 뒷받침할 방법과 도구들이다. 결론으로는 모니터링과 리포팅을 위한 관련 메카니즘을 정립하는 것이 필요하다.

OECD : 삶의 웰빙 프레임은 GDP와 같은 경제 보다는 사람, 입력과 출력보다는 결과(Outcomes), 평균과 불평등, 주관적과 객관적인 면, 오늘과 내일 등과 같이 크게 4가지로 나타내고 있다.

국제적인 웰빙 지표들 중 사회적 연결(Social connection)과 주관적인 웰빙(subjective wellbeing) 등을 나타내는 데 한계가 있다고 설명하고 있으며, 기존 정책의 한계를 개선(Marginal improvement to existing policy)한 국가는 오스트리아, 호주, 이탈리아, 뉴질랜드, 영국이며, 주요한 정책과 관련된 웰빙(Well-being a significant policy concern)을 적용한 나라는 중국, 이스라엘, 한국, 멕시코 이다.

- (토론내용/질의응답) 국가적인 웰빙 영역의 구성과 국제적인 측면에서의 모니터링 방안 필요
국가적으로는 크게 3~4개의 영역을 구성하며 지표의 수에 대하여서는 다양하게 결정하고 있는 상황이며, 국제적인 관리를 위해서는 각 국가들의 공통지표에 대하여 모니터링 및 공표를 하고 있으며, 지속가능한 개발 목표를 발전시켜 나가는 것이 필요함

□ 시사점

- 여러 국가나 국제기구에서 웰빙에 대하여 정의하고 있으며, 지속가능한 개발 목표를 향해 다양한 노력을 하고 있고, 영역과 지표의 기준을 다르게 설정하고 있다. 한국도 국가차원에서의 웰빙 정의 및 기준을 명확하게 정의하고 다양한 분야로의 지속적인 개발이 필요하다고 생각됨

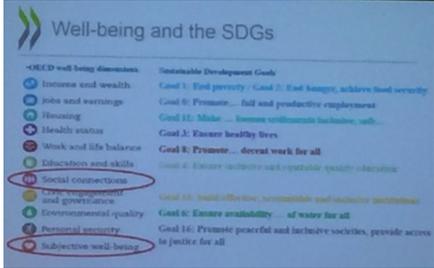
궁극적으로 통일된 기준으로 지수를 나타내는 것은 한계가 있으며 사회적 연결, 주관적인 웰빙 등에 대해서는 국제기구나 선진국과의 공조를 통해 다양한 시도가 필요할 것으로 생각됨

□ 첨부자료

세션 참석자



OECD 발표자료



□ **세션개요**

- 일시/장소 : 7/27/2015, 14:00~15:30 Room 203B
- 세션구성
 - 좌장(성명, 소속 등) : Jean-Michel Durr
 - 발표자(성명, 소속, 논문제목 등)
 - . Improving the Production Process of the Irish Retail Sales Index(Satisfying Customs' Needs Using Lean Six Sigma Thinking)(Steve Macfeely, Patrick Foley, University college Cork, Ireland)
 - . Quality Management System annd Further Challenges – Innovation Approach in Re-designing of the Statistical Business process(Blagica Novkovska, Slovenia)
 - . Improvements on the statistical quality practices at the Brazilian Institute of Geography and Statistics(Zelia Bianchini, IBGE)
 - . Improving the management of statistical processes via the use of Quality Gates(Paul Schubert, Siu-Ming Tam, ABS)
 - 토론자(성명, 소속 등)
 - . Andrea Diniz da Silva
- 참석인원 : 100명

□ **세션요지**

- (발표내용)

아일랜드 : Lean* Six-sigma를 활용하여 효율적이며 시의성(Timeless) 개선을 위해 증거 위주의 결정을 강조하고 분산 및 시간낭비를 줄이고 예측을 높이는 방안을 제시

 - * Lean ? 도요타의 생산시스템에서 시작, 제조업 뿐만아니라 서비스업을 포함한 다양한 분야에 적용, 예를 들면, 과다생산, 필요 없는 공정단계 진행, 불필요한 이동, 대기, 불량등이며, 노력, 시간, 공간을 줄이고 더 많은 것을 얻는 방법

진행과정에 대한 흐름도를 활용하여 불필요한 데이터의 이동, 대기시간, 무가치의 행동 등을 줄여서 효율적이며 시의성 있는 통계 생산에 도움이 되도록 하며, 이를 통해 인력 요구사항(staff requirement)을 50%까지 줄일 수 있게 되었다.

슬로베니아 : 품질관리를 위해 통계청은 2007년부터 지속적인 개선을 실시하였다.

품질관리를 위해 탑다운(Top down)방식과 버텀업(Bottom up) 방식을 병행하여 실시, 탑다운 방식에는 공통적인 평가 프레임(Common Assessment Framework:CAF)와 EU의 프랙티스 코드(Code of Practice)와 상세한 리뷰들로 구성되어 있으며, 버텀업 방식은 조사품질 리포팅(Survey quality reporting)과 DESAP(자체통계품질진단)으로 구성되어 있다. 또한, SBPM을 도입하여 통계작성체계를 일원화 및 통계작성전반을 관리하기 위해 eStat-system(온라인 자료수집체계를 구축) 구축하였다.

브라질 : 2013년에 2011년 아메리카 컨퍼런스(2011 statistical Conference of the Americas)에서 제시한 가이드라인, 원칙, 지표 및 우수사례 등을 기반으로 Cop를 구성하였다. 17개의 키 원칙(Key principles)과 80개의 지표(indicators)로 코드를 구성하였으며, 사회통계 위원회(Social statistics committee)를 2007년에 설립하였으며, 연합정부의 사회통계 생산자들로 구성되어 있다. 경제통계 위원회(Economic Statistics Committee) 구성을 위해 노력중이며, 통계조사의 응답부담, 통계생산의 효율성 제고 등을 위해 행정자료의 통계적 사용 확대를 위한 노력을 하고 있다. 통계 생산자료에 대한 메타데이터의 공통형식 등을 구성하고 있으며, IBGE는 다양한 전문성과 기술 등의 확대를 위해 훈련 계획 등을 수립하고 있다, 정확성, 비교성, 접근성 등의 향상을 위해 지속적인 노력을 하고 있으며, 향후에는 품질보증제도를 구축 예정

호주 : 통계작성과정에 대한 체크포인트와 결정 포인트를 마련하여 프로세스의 품질 또는 명확한 평가를 위해 활용하고 있다. 그런 의미에서 품질 게이트(Quality Gates)체계를 구축하여 신호등(블루, 옐로우, 레드)과 같은 개념으로 관리하고 있으며, 지속적으로 개선하고 있음. 품질 게이트의 이점으로는 품질을 유지하고 위험을 최소화 할 수 있으며, 효율성을 높일 수 있으며, 성공적인 수행 및 지식공유에 도움이 되며, 책임성을 명확하게 할당할 수 있는 것 등이 있다.

- (토론내용/질의응답) GSBPM 사용을 통한 체계적인 관리 및 통계작성과정 관리방안에 대한 토론
통계작성과정별로 GSBPM을 활용하고 있으며, 또한, 워크숍이나 교육 등을 통해 작성과정 단계별 관리방안을 공유하는 것이 필요

□ 시사점

- 통계작성과정의 효율성과 시의성 확보를 위해 Lean six-sigma를 도입하고, GSBPM과 같은 체계를 구축하여 통계작성과정을 표준화하고 각 국가별로 중요시하는 품질차원을 개선하려는 노력을 하고 있음

- 또한, 기존에 국가통계의 품질관리를 위해 활용하는 품질차원들을 고도화하고 보완하는 방안을 지속적으로 노력하고 있으며, 우리나라도 시의성과 효율성 등을 확보하기 위한 방안 검토가 필요할 것으로 생각됨

□ 첨부자료

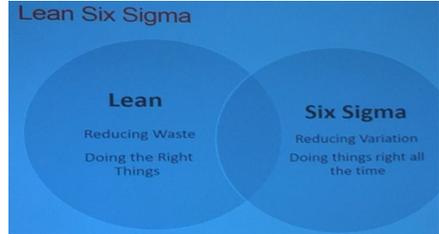
슬로베니아 발표자료(e-stat)

Part 3: eStat System

- Metadata driven system for statistical data collection
- Web-forms for "on-line" statistical data collection
- Automated solution for users
- Integration of various and isolated IT solutions
- Using internationally accepted standards

Expected benefits: Standardization of processes, Reducing burden of reporting units, Improving quality

아일랜드(Lean과 Six sigma)



□ 세션개요

- 일시/장소 : 7/31/2015, 10:30~12:30 Room 205
- 세션구성
 - 좌장(성명, 소속 등) : Edurado Barredo Capelot
 - 발표자(성명, 소속, 논문제목 등)
 - . Spain 2021. Is it possible to carry out a Census without any fieldwork operation?(Jorge L. Vega Valle, Spain NSO)
 - . Toward a register-based Census-Abu Dhabi(Aisha Turki, Badreyya Al Shehhi, Salah Qudairi, Walter Robb, Abu Dhabi)
 - . Changing Census by combining administrative source and sample surveys(Fabio Crescenzi, Italy NIS)
 - 토론자(성명, 소속 등)
 - . Pado Valente
- 참석인원 : 60명

□ 세션요지

- (발표내용)
 1. (Spain) 스페인의 경우, 2001년에는 약간의 등록자료를 활용하여 전수조사를 실시하였고, 2011년에는 등록자료를 활용하고 10%를 조사에 의존, 2021년에는 현장조사 없이 기존의 조사결과와 등록자료를 연계하여 활용할 계획임
 - 인구등록자료를 활용하는 이점은 국민의 부담을 줄이고, 주기적으로 요구되는 정보를 얻을 수 있으며, 공공예산을 최적으로 활용할 수 있으며, 산출하기 어려운 (hard-to count) 그룹을 찾기가 쉬워진다.
 - 2021년에는 인구등록자료(Population register), 인구와 관련된 다른 등록자료, 등록자료로 부터 얻어내기 어려운 자료에 대해 기존의 조사결과, 사전 센서스, 토지등록자료, 유동성에 대하여 모바일 폰으로 부터의 빅데이터(big data)등을 통합하여 실시 예정
 - 이를 위해 2014년과 2015년에는 인구등록자료와 관련자료(과거 센서스, 통계적 모델) 등을 활용하여 이주자(Migration)과 가족관계(Family relationship) 등의 자료를 구축하였다. 현재는 이용자의 요구에 부합하는 다른 형태의 결과를 제공하

기 위해 Big microdata file과 지리적인(geographical) 정보를 활용하고 있다.

- 결론적으로 현재까지 "big countries(거대 국가들)"에 대하여 등록자료를 활용한 센서스가 성공한 사례가 없다. 스페인은 2021년 조사없는 센서스를 성공적으로 실시하기 위해 최선을 다하기 위해 많은 준비를 하고 있다.

2. (Abu Dhabi) 아부다비는 매년 인구가 급속하게 증가하고 있는 상황이다. 행정자료를 활용하면 자료수집의 시간과 비용을 줄일 수 있고, 5년 주기보다 짧은 주기로 정기적으로 자료를 추출할 수 있는 이점이 있고, 수집된 자료를 활용하여 지역과 국제기구에 필요한 통계지표 등을 추출할 수 있게 된다.

- 아부다비의 등록기반에 활용되는 자료는 크게 세가지이며, 주택등록자료(Housing register), 인구등록자료(Population register), 노동력자료(Labor register)이다.
- 결론적으로 아부다비는 행정자료를 기반으로 2020년에 등록기반 센서스를 실시하기 위한 준비를 하고 있다.

3. (Italy) 센서스 근대화 1차 라운드 : 2011년 이전에는 조사를 통해 센서스를 실시하였다. 센서스 근대화 2차 라운드 : 2011년 이후에 지속적인 자료 업데이터 시스템으로 부터 매년 센서스 결과를 활용할 수 있게 되었다.

- 현재 빅 파일럿 서베이(A big pilot survey)를 실시하고 있다. 150개의 지방자치단체와 16만개의 가구를 대상으로 하고 있으며, 이를 통해 통계적인 모형의 수용력을 예측하고, 다른 원천으로부터 통합된 자료에 IT를 접목시켜보고, 모바일, 태블릿 기타 등등과 같은 웹적용을 통해 데이터의 동기화를 테스트하기 위해 실시하고 있다.
- 이러한 것을 통해 비용과 자료요구 및 책임은 변화에 따라 줄어들고, 등록시스템, 다양한 자료로 부터 자료를 추출할 수 있는 방법의 발전, 조사의 통합, 프로세스의 현대화 등은 성공의 요인이 되었다.

○ (토론내용/질의응답)

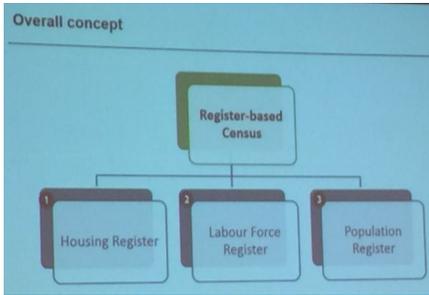
조사된 자료와 등록자료의 연계를 통해 좋은 결과를 유추해 낼 수 있으며, 특히, 스페인의 경우 2021년 센서스를 현장조사 없이 실시하기 위해 다양한 등록자료의 연계 및 통계적 모델 등을 구성하고, 또한 지리적인정보 등을 활용한 주택등록자료 보완 하려고 하고 있다. 다양한 자료의 연계 및 사전 준비 등을 통해 성공적인 사례가 될 것으로 기대된다.

□ 시사점

- 다양한 국가에서 등록자료, 행정자료, 조사자료 등을 연계하여 향후 센서스에 도입하려는 시도가 계속되고 있으며, 특히, 스페인과 아부다비는 2021년과 2020년의 센서스를 위해 등록자료의 보완, 빅 데이터와 지리적인 정보 등을 활용하여 등록 센서스를 위한 준비를 지속적으로 실시하고 있다.
우리나라는 2015년 등록센서스를 실시하여 기존의 조사를 활용한 센서스보다 효율성과 활용성에서 성공적인 결과를 이끌어 내어 다른 국가(현재 등록센서스를 준비 중인 국가)의 모범사례가 될수 있도록 하여야 할 것이다.

□ 첨부자료

아부다비의 등록 센서스 기초자료



세션참석자들



참고 10

(IPS025) Governance and innovation/Quality assurance systems for statistics : From development to implementation and monitoring

□ 세션개요

- 일시/장소 : 7/29/2015, 10:30~12:30 Room 202A
- 세션구성
 - 좌장(성명, 소속 등) : Walter Radermacher
 - 발표자(성명, 소속, 논문제목 등)
 - . Prioritising the User Perspective on Statistical Quality(William Richard Alldritt, UK)
 - . The role of peer reviews as an important instrument in the context of monitoring and enhancing quality of official statistics(Helena Cordeiro, Portugal)
 - . Strengthening Quality Assurance for National Statistical Systems(Monica Pinzon Torres, Colombia)
 - . Quality Management System of Korean Official Statistics(Hyungee Kim, South Korea)
 - . Reinforced quality assurance of Government Finance Statistics(불참)
 - 토론자(성명, 소속 등)
 - . Thomas Burg(Austria)
- 참석인원 : 50명

□ 세션요지

- (발표내용)
 1. (UK) 영국에서의 품질 프레임은 이론과 과정, 응답 부담, 비용적 효율성, 관련성, 정확성, 시의성, 일관성, 접근성 및 명확성으로 정의하고 있다. 유엔(UN Fundamental principles)에서는 공식통계는 실질적인 유용성을 검정하는 것이고 OECD(OECD quality Framework)에서는 품질은 이용적합성(fitness for use)이라 정의한다. 결론적으로 품질은 유용성과 같다고 정의하였다.(QUALITY=UTILITY) 그러나, 우리가 이야기하는 관련성(Relevance)의 개념은 유용성(Utility)과는 다르다고 설명하고 있다. 관련성은 이용자 관점이 아니라 생산자 관점에서 어떻게 하였는지를 보는 것이다.
통계청은 이용자에게 대한 증거로부터 최대의 결과를 얻을 수 있을 것이다. 이용자에게 증거를 얻기 위해서 질문지를 잘 만들어야 한다. 예를 들면, 당신은 우리사이트를 들어 본적이 있는나?는 질문이 아니라 당신은 결정을 할 때 어떤 통계로 부

터 정보를 얻는가? 라는 질문으로 물어보아야 한다.

2. (Portugal) 2001년부터 2015년까지 품질관리 방법의 변천에 대하여 개괄적으로 설명하였으며, 2006년부터 2008년까지 1차 라운드의 상세한 리뷰(Peer Review)를 하였으며, 이를 통해 2009년에 규칙 223호를 만들었다. 규칙 223호는 세가지로 구성되어 있으며, 첫째로 ESS(European Statistcial System)가 어떻게 해야 하는지에 대한 기본원칙과 역할을 제공한다. 두번째는 누가 무엇을 할지, 결정은 누가 할지, 어떻게 할지를 명확하게 한다, 셋째는 통계작성에 대한 것을 규정으로 정의한다. 2013년부터 2015년까지 2차 라운드의 상세한 리뷰는 스태프, 매니저 등에 대한 훈련이 요구되고, 근거 기반과 문서 준비 등이 필요하고, 기관간의 협조 등도 강하게 요구된다. 그러나, 좋은 결과가 나오게 하기 위해서는 대규모의 투자가 필요하고, 자체적인 규칙, 각 기관의 지식 등의 요구가 증가하게 된다. 향후에는 Cop(Code of Practice)의 개정, 품질보증체계의 개정 및 좋은 사례와 혁신적인 도구를 기초로 실질적인 혁신을 위해 노력할 것이며, 3차 라운드 상세한 리뷰는 2020년에 실시할 예정이다.

3. (Columbia) 품질보증체계의 주요요소(Key elements)로는 실천 코드(Code of practice), 통계적 과정, 통계품질 요소 등으로 구성되어 있으며, 이를 위해 촉진(promotion), 평가(Assessment), 모니터링(Monitoring) 순으로 관리하고 있다. 품질보증을 위해 독립된 전문 기구는 수집된 정보의 분석, 웹페이지 리뷰, 증거의 문서화 자료의 분석 및 리뷰, 통계적인 지속성 분석 등과 통계생산에 참여한 전문가의 의견수렴 및 이용자 인터뷰, 자료의 품질 평가 등을 종합하여 평가하고 있다.

품질인증의 기준을 3가지 유형으로 관리하고 있다

- Type A 인증 : 점수 94~100(3년)
- Type B 인증 : 점수 87~93(2년)
- Type C 인증 : 점수 80~86(1년)
- 인증하지 않음 : 80점미만

품질개선을 위한 계획에 대한 모니터링을 통해 실행계획의 준수 여부, 사전검증 및 인증된 통계 프로세스 정보를 발표하기 위한 사전 준비 등을 모니터링 하고 있다. 결론적으로 통계품질인증 모델의 개선을 위해 최고의 실질적인 개선의 촉진, 평가, 모니터링을 포함하여 실시하려고 하고 있다. 이를 통해, 통계생산과정의 약점을 파악하여 지속적인 개선을 할수 있으며, 또한, 통계생산과정 전부를 이용자에게 자신있게 제공할 수 있을 것이다. 더 나아가 이러한 작업들을 통해 표준적인 개선과 국제적으로 인정할 만한 최고의 사례 등을 만들 수 있을 것이다.

4. (Korea) 한국의 국가통계 품질관리체계에 대하여 소개, 품질관리 방법 및 최근에

개선한 정기통계품질진단 방법을 소개하였다. 한국의 통계생산체계는 분산형이며, 통계청은 일부 주요 통계를 생산하고 있다. 통계법에 3가지 유형의 품질진단을 규정하고 있으며, 이는 정기, 수시, 자체통계품질진단이다. 정기, 수시, 자체 품질진단의 주기와 필요성, 대상 등에 대하여 소개하고 품질개선이 필요한 기관에 대하여 품질컨설팅을 실시하는 내용을 소개하였다. 마지막으로 2015년부터 정기통계품질진단 체계의 개선사항을 설명하였다. 통계작성기관 담당자가 직접 통계작성과정에 대하여 작성한 통계정보보고서를 기반으로 진단에 활용하며, 진단의 효율성 및 비용을 절감하여 5년간 500종 내외의 진단이 가능하게 되었다, 또한, 통계정보보고서 작성방법 및 평가방법에 대하여 소개하고 진단결과 공표 및 생산자용 통계정보보고서와 이용자용 통계정보보고서를 제공하여 향후 통계이용자의 통계 오남용을 줄일 수 있도록 노력하고 있다.

○ (토론내용/질의응답)

국가재정통계와 같은 경우에 품질인증을 어떻게 하고 있는지에 대한 질문에 통계마다 인증의 판단 기준이 다르며 통계의 특성마다 기준에 차이 발생, 그리고, 현재 제공되고 있는 품질보고서(Quality Report)가 너무 어렵게 되어 있으며, 이용자에게 활용하는 데는 한계가 있다.

또한, 품질과 유용성이 같은 의미를 가진다고 하였는데 그 의미는 품질은 이용자의 편리성과 궁극적인 목적에 부합하여야 통계로서 가치가 있는 것으로 설명하였다.

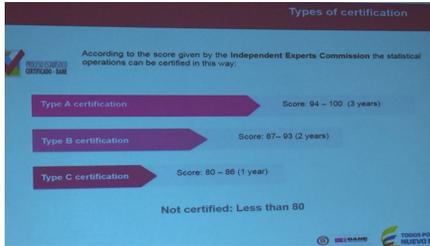
□ 시사점

○ 유럽연합(Eurosta)에서는 이미 ESS(Eurostat Statistical System)을 만들어 품질 보증 및 인증 제도를 도입하고 또한, 이를 향상시키기 위해 다양한 노력을 하고 있다. 그리고, 일정기간의 진단 및 인증이 끝나고 나면 지난 결과에 대한 리뷰를 통해 개선방안 및 원칙 등을 수정 및 추가하여 통계의 품질을 향상을 위해 지속적인 노력을 실시하고 있다.

○ 현재 우리나라의 경우, 정기진단은 2006~2010년까지 1기 진단, 2011~2014년까지 2기 진단을 완료하였으나, 세부적인 검토가 부족한 것으로 생각되며, 향후 3기 진단 2015~2019년 완료 후에는 종합적인 개선방안 및 원칙 등을 재정립할 수 있는 분석이 필요할 것으로 생각된다. 또한, 3기 진단이 끝나기 전에 품질인증을 위한 기초적인 연구와 방법 마련을 위해 개선방안을 마련하여야 할 것이다. 기본적으로 유럽통계청의 ESS 체계와 인증(Assurance), 감사(Audit) 등에 대한 방법을 사전 연구 필요할 것이다.

□ 첨부자료

콜롬비아 발표자료(평가 등급)



발표자료



□ 세션개요

- 일시/장소 : 7/28/2015, Room 204B
- 세션구성
 - 좌장(성명, 소속 등) : Abo Bakr Elgindy
 - 발표자(성명, 소속, 논문제목 등)
 - . Special topic session – ISI 2015 Emerging topics for European statistics in the informtion age (Mariana Kotzeva, Eurostat)
 - . Recent developement in the preparatory arrangements for Egypt census 2016 (Nehall ahmed Farouk, Central Agency of Public Mobilization and Statistics, Egypt)
 - . Data Mining Applications in Official Statistics (Mohamed El-Sharkawy, CAPMAS, Egypt)
 - . Importance of Harmonized Household Income and Expenditure Surveys: with an emphasis on CPI (Hala Abou-Ali, ERF data Portal, Egypt)
 - . The use of mobile phone dataand geo-spatial information (Ronald Jansen, UNSD)
 - 토론자(성명, 소속 등)
 - . Mohammed Ismil
- 참석인원 : 80명

□ 세션요지

- (발표내용)
 1. (Mariana Kotzeva) Eurostat은 공식통계 작성의 당면과제로 모바일을 활용한 통계의 생산 및 서비스, 이용자 참여의 소셜 미디어 활용 등에 대해 ESS-2020 VISION에 담아 추진 중임을 설명
 - * ESS: European Statistical System
 - 즉 기술적인 분석과 시각화한 자료를 모바일을 통해 이용자와 인터랙티브(interactive)한 통계서비스에 많은 혁신과 투자 예정
 - 이를 위해 각국의 통계청은 통계기구의 데이터 이용능력 제고, 글로벌 통계기구와의 파트너십 이용, 고위급 리더십이 강화가 필요함을 강조

2. (Ronald Jansen) 모바일 데이터, 공간 지리정보, 위성 이미지 등을 결합한 공식통계 및 지표의 활용에 대한 계획 및 사례 설명
 - UN은 Post-2015 아젠다로 17개 Goals, 169개 Targets을 설정하여 300여개의 Indicators로 모니터링 할 예정
 - UN의 Global Pulse는 뉴욕, 자카르타, Kampala에 네트워크를 설치하여 실시간의 조기경보(질병, 기후변화 등) 및 피드백에 활용 중
 - 또한 UN은 2014년 글로벌 워킹그룹을 결성하여 공식통계에 빅데이터 활용전략을 마련 빅데이터의 Methodologies, 법적 제도기반, 개인정보 보호, 데이터 안전, 비용 등(8개 T/F 구성)의 이슈 해결에 실질적인 활용과 지원 추진 중
 - 현재까지 모바일 데이터 및 공간지리정보는 결합을 통해 관광통계, 질병발생 분포, 인구이동의 역동성(Population dynamics), 재배 농작물 및 생산예측 정보, 지질특성 및 이용, 지역고용 및 사업체 현황 등에서 활용되고 있음
3. (이집트 통계청 및 관계기관) 소비자물가지수(CPI) 자료를 이용한 가구소비 및 지출조사의 정확성 제고, 2016년 Census에 모바일 조사 준비, 통계정보 및 프로그램을 이용한 데이터마이닝(DM; data mining)의 중요성 등에 대해 설명
 - (토론내용/질의응답) 데이터마이닝(DM)은 전통적인 기법인데 설명 중에 기존통계를 분석하여 오류 및 숨어있는 현상을 포착한다고 했는데 이를 통해 과거자료의 내용 보정에도 활용한다는 의미인가?
 - > DM은 과거자료의 분석을 통해 미래를 예측하거나 오류의 재발방지에 이용하는 것으로 데이터의 오류를 수정하는데 활용하지는 않고 있음

□ 시사점

- 이 세션은 이집트 통계청이 마련한 것으로 이집트의 설명자료는 새로운 기법의 적용사항은 없으나 초청논문인 Eurostat, UNSD의 발표내용은 국제기구의 트렌드를 알 수 있는 좋은 기회
- 한국의 통계청도 국내 통계기구들의 통계 및 데이터 활용현황과 개선사항에 대해서 스스로 평가하여 최신 트렌드를 학습할 수 있는 특별세션 구성 필요

□ 세션개요

- 일시/장소 : 14:00~15:40, 7/28/2015, Room 204B
- 세션구성
 - 좌장(성명, 소속 등) : Eric Nordholt
 - 발표자(성명, 소속, 논문제목 등)
 - . The use of administrative data for statistical business registers and trade statistics (Ronald Jansen, UNSD)
 - . Emerging applications of tax administrative data in South Africa (Elizabeth Gavin)
 - . Tax data in official statistics: opportunities and challenges (Michael Slyuzberg)
- 참석인원 : 100명

□ 세션요지

- (발표내용)
 1. (Ronald Jensen) 행정자료 활용의 당면과제, 행정자료와 SBR (statistical business registers), 행정자료와 무역통계, SBR과 무역통계의 연계에 대해 설명
 - 행정자료는 통계생산의 저비용, 응답부담 경감 등 많은 이점이 있으나 다음과 같은 과제(Challenges) 들이 존재
 - . 원천 행정자료의 접근 제약: 법적, 동의, 계약관계 등에 따른 제약
 - . 데이터 전송 제약: 파일 형식 상이, 메타데이터 표준화 부족 등
 - . 통계적 사용 목적의 적합성: 원자료의 단위, 변수, 분류 등이 상이하거나 시의성 부족
 - SBR은 행정자료 결합의 통계 생산에 매우 중요, 즉 잘 구축된 SBR자료는 경제통계 생산의 핵심역할
 - . SBR 자료는 다음과 같은 형식으로 존재

: 기업 등록/면허, 조세자료: VAT, 근로소득세, 기업/협회, 상공회의소 등록, 사회 보장 등록, 고용 등록, 행정기관별 회계자료, 비영리기구 등록자료, 협회등록, 농업사업체 등록, 상수도/전력 등록, 학교, 병원 등록부, 중앙은행의 금융분야 별 자료, 외국인 직접투자 감독자료, 주주명부, 보조금 정보 등

- 상품거래 무역통계(IMTS; International Merchandise Trade Statistics) 생산에 활용 되는 행정자료(세관신고서, 화물송장/ 송장기록, 국제거래 신고시스템, 부가세)를 예시하고 전력자료를 결합한다면 좋은 통계를 서비스 할 수 있음을 설명

2. (Gavin) 남아공이 조세행정자료 활용 개선 효과에 대해 설명

- 남아공의 조세자료 이용의 현대화는 근로자를 고용할 경우 고용주의 신고의무를 강화하여 데이터 품질, 시의성, 완성도를 크게 개선시키는 효과 발생, 특히 납세자의 경제활동 실태를 예전보다 정확하게 보여줌

3. (Slyuzberg) 뉴질랜드 조세자료는 공식통계의 작성에 광범위 하게 활용되고 있거나 기회와 과제에 대해 설명

- 기회

- . 뉴질랜드 통계청 및 다른 기관에 공급
- . 조세자료 통합데이터시스템을 통해 자료 접근 가능
- . 비밀보호 파일 생산
- . 국세시스템으로 집계된 통계데이터는 정부데이터 포털로 바로 공개
- . 관련 연구 및 분석 프로젝트에 지원

- 주요과제(지속적)

- . 용어정의/메타데이터 관리
- . 데이터 품질제고
- . 고객중심 데이터 생산
- . 통계적/분석적 기법 개발

○ (토론내용/질의응답)

- 유엔에서 행정자료 활용에 대해 지원한 사례는?

-> 코스타리카의 경우 행정자료를 통계청에서 활용할 수 있게끔 유엔에서 지원 하였으나, 콜롬비아는 조세자료의 이용에 제약이 많음을 설명, 행정자료 이용이 발달한 뉴질랜드의 경우는 영세사업체 통계에 과세데이터 의존하는 경향,

통계산출물의 변환의 적절성(예: 6개월 단위로 수집된 조세정보를 분기단위로 변환하는 방법 등) 등의 이슈를 제기하고 있음

□ 시사점

- 이 세션은 행정자료 활용사례를 알 수 있는 좋은 기회였고, 외국의 행정자료(조세 자료 등)의 활용사례처럼 우리나라도 이에 대한 제도적 기반확충에 이용할 필요가 있음

□ 세션개요

- 일시/장소 : 7/31/2015, Room 207
- 세션구성
 - 좌장(성명, 소속 등) : Shigeru Kawasaki
 - 발표자(성명, 소속, 논문제목 등)
 - . Praactice of Geography Census and Monitoring in China (Dr. Pengde, Deputy Administrator, Administration of Surveying, Mapping, and Geoinformation, China)
 - . Integrated Geography and Statistical Expertise (Mr. Ronaando Ocampo, Vice President of the Governing Board, INERGI, Mexico)
 - . Mr. Rochelle Morgan, Manager, Geospatial Platform Operations Branch, Staistics New Zealand
 - . Unleashing the power of data to change lives (Ms. Gemma Van Halderen, General Manager, Strategic Partnerships and Project Division, ABS)
 - 토론자(성명, 소속 등)
 - . Ms. Gemma Van Halderen, General Manager, Strategic Partnerships and Project Division, ABS
- 참석인원 : 80명

□ 세션요지

- (발표내용)
 1. (Mr. Pengde) 중국은 총리가 위원장으로 각 부처 및 행정통계국이 참여하여 제1차 국가지리센서스를 2013년부터 2015년까지 실시함
 - 조사내용 : 토양 특성, 자원, 환경, 지질안전, 토지인프라, 공공서비스, 공간개발
 - 소요예산: 100억 위안
 - 중국 지리센서스 결과 시사점
 - . 지리정보와 통계정보를 잘 활용한다면 세상은 스마트하게 관리 가능
 - . 모든 센서스는 지역기반으로 실시될 필요

- . 세계는 통일된 스마트한 지도가 필요
 - . 2020년에 세계지리총조사를 할 필요가 있음
 - . 세계지리정보는 post 2015 구현에 유용
2. (Rolando Ocampo) INERGI는 멕시코가 지리통계정보의 생산을 위해 1983년에 설립하였으나 미국, 중국 등의 위성보다 소형의 위성에 의존하여 정확한 정보의 생산에 어려움 있음
- (토론내용/질의응답)
 - (Gemma Van) 지리정보의 활용은 각 기관의 협력(행정자료+조사자료+지리정보 등)이 중요하고 국가간 자료의 비교를 위해 IAOS 등 국제기구와의 협력이 또한 중요
 - . 각국은 10년주기의 지리/인구 센서스 실시를 제안함

□ 시사점

- 참석자들은 대부분 지리정보 및 인구 센서스 등의 정보를 결합하여 다양한 공간 정보의 생산에 동의
- > 공간정보와 행정정보를 활용한 센서스 등의 각종 통계의 개발에 적극적인 바, 우리나라에서도 현황 및 발전방안 대한 검토가 필요

□ 세션개요

- 일시/장소 : 7/27/2015, 203C
- 세션구성
 - 좌장(성명, 소속 등) : Giorgio Alleva
 - 발표자(성명, 소속, 논문제목 등)
 - . A Business Architecture Model to support the Modernization Project within the Istat(Piero Dermotrio Falorsi 외 2명, Istat)
 - . Modernation strategies in the ABS(Duncan Young, ABS)
 - . Improvement of Business Architecture model at the HCSO - Experiences of HCSO(Csaba Abry, HCSO)
 - . Modernization of Official Strategies - an International Perspective(Steven Vale, UNECE)
 - 토론자(성명, 소속 등)
 - . Nadia Mignolli
- 참석인원 : 70명

□ 세션요지

- (발표내용) 공식통계의 현대화를 위한 경영기술(Business Architecture) 도입 과제
 - (ABS) 전통적인 GDP, Cesus 등의 통계수집 중심의 활동을 개혁하여 솔루션 중심의 통계서비스를 도입할 필요
 - (ISTAT) 경영기술 도입의 당면 과제 : 통계작성절차의 개선, 직원 능력 개발, 통계 거버넌스의 강화
 - (UNECE) 데이터 수집 방법, 빅데이터 활용, 비용절감 등 급속히 변화하는 통계환경 당면문제 들*의 해결을 위해 ECC VISION 2020 CSPA 서비스를 국제적 협력을 통해 강화할 필요
 - * 예시) 빅데이터의 활용의 우선적 해결 문제
 - . 비밀보호, 품질관리, 협력 등을 위한 가이드라인 개발
 - . 프로파일 기법 개발

. 모래상자(sand box) 같은 IT 및 방법론적 이슈들

* CSPA : Common Statistical Production Architecture

○ (토론내용) 공식통계의 실제와 공시공계 처리시스템과의 협력의 중요성

- EA(Enterprise Architecture)는 파산을 막기위해 사업의 비전과 성과의 수행을 사전적으로 분석함으로써 미래의 기업목표를 조정하여 미래 발전전략을 결정하는 핵심적 활동의 경영기법

- 통계조직에 있어서도 통계의 생산, 설계, 통계정보의 보존과 통계자산 등에서 성과 및 리스크를 관리하는 다음의 환경변화에 대한 모든 통계적 경영정보, 기술, 응용 활동에도 적용됨

. 고령화 및 ICT 기술환경

. 빅데이터 활용, 통계분석 및 공유 활용 등 이용자의 요구 변화

. 통계생산 비용의 증가 및 자원이용의 효율화

. 통계생산 및 정보시스템의 솔루션센터 등

○ (질의응답)

- 헝가리의 통계경영 기술의 발전전략 방향은?

=> UNECE의 포괄적 통계처리기법모델(Generic Statistical Business Process Model) 등 국제표준에 맞는 기법을 수용하고 있음

□ 시사점

○ 한국의 통계청에도 UNECE, ABS의 사례를 검토하여 새로운 환경변화에 대처할 수 있는 국제 표준의 통계경영기술 도입 및 적용 필요

□ **세션개요**

- 일시/장소 : 10:30~12:30, 7/30/2015, Room 207
- 세션구성
 - 좌장(성명, 소속 등) : Steven Heeringa(미시건대 교수)
 - 발표자(성명, 소속, 논문제목 등)
 - . Recent Experiences in SCB (Jozef Teugels, 벨기에 루벤카톨릭대 교수)
 - . Experiences in Mentoring (Arnoldo Frigessi, 오슬로대)
 - . Building Statistical Collaboration Capacity in developing countries(Eric Vance, 버지니아공대 교수)
 - . Capacity Building Experience In Southen Africa(David North, 남아공 KwaZulu대 교수)
- 참석인원 : 70명

□ **세션요지**

- (발표내용)
 1. (Jozef) ISI와 관련기구는 전략적 계획 과정 (SCB; Strategic Planning Process)에 참여하고 있고 ISI의 SCB (Statistical Capacity Building)는 센터 역할 수행
 2. (Arnold Frigessi) 수리통계 모형개발, 개발도상국 박사학위 과정, 박사학위 과정의 장거리 멘토링 (Distance Mentoring) 통계역량 구축사례 설명
 - 이태리의 경우 가족중심 교육, 낮은 생활수준 등의 환경에서 멘토링을 통해 국제적인 EU펀딩 국가로 발전
 - . 통계역량 구축은 어느 곳에서부터 착수하고, 작지만 중요한 단계에서 단기적 인센티브의 부여를 통한 장기적인 목표 달성에 중점을 두고 수행하는 것이 중요
 - . 탐다운 방식만이 아닌 초기 투자가 궁극적으로 장기적인 스폰서가 되게 하는 명확한 목표를 두고 시행
 3. (Eric Vance) 개발도상국에 대한 통계협력역량구축 (Building Statistical Collaboration Capacity in Deveveloping Countries) 추진사항 설명
 - LISA 2020 (Laboratory for Interdisciplinary Statistical Analysis 2020) 프로그램 설치
 - . LISA 2020은 비통계인과 협력하여 개발도상국의 통계인 및 데이터 사이언티스트을 대상으로 교육 및 훈련을 실시하는 통계분석제휴실험실 프로그램

- . 교육 후 자기나라에 귀국, 고국의 대학에 통계분석제휴실험실을 만들게 하고 연구원, 공무원, 중소기업, NGO 들에게 통계를 통한 의사결정능력을 갖출 수 있도록 하여, 2020년까지 전세계적으로 20개의 실험실 네트워크를 구축
 - . 버지니아 공대의 LISA에는 LISA 2020 Fellow, LISA2020 자문위원회, 2020 멘토링 네트워크 등 운영
4. (Ada Van, ISI 국장) 최근 한국 대전 등에서 High level Workshop을 통한 통계역량구축회의를 개최한적이 있음
- 워크숍은 공식통계의 미래 리더를 발굴하고 혁신적이고 강한 커뮤니티를 구축하여POST-2015 개발수요 발국에 협력중임
- (토론내용)
- 202프로그램 수행을 위해 펀딩 및 지원한 사례는?
 - > USA농무성 등으로부터 펀딩을 하여 탄자니아의 통계역량강화를 위한 통계분석제휴실험실 설립 지원 중

□ 시사점

- 이 세션은 개발도상국 통계역량강화사업의 추진사례를 공유한 것으로 국가통계역량강화사업에 참고할 수 있는 중요 사례임
- 우리나라 통계청도 비통계인과 협력하여 통계와 데이터 전문가를 양성하고 연구원, 공무원, NGO 등을 대상으로 통계활용능력을 강화하는 프로그램 개발 검토 필요

□ 세션개요

- 일시/장소 : 2015. 07. 28. 10:30 ~ 12:30 / 205
- 세션구성
 - Organizer/Chair : Arona Pistiner
 - Presenter ① : Keiko Osaki-Tomita
"Methodologies of population and housing census : revisited"
 - Presenter ② : Eric Schulte Nordholt
"What is a census : a register-based view"
 - Presenter ③ : Duncan Young
"Traditional Census-A misnomer and an oxymoron"
 - Discussant : Claude Julien
- 참석인원 : 약 40명

□ 세션요지

- UNSD에서 조사한 2010년 라운드 센서스 조사결과 214국이 참가하고 이중 2011년에 60 개국이 참가하였음. 금번 조사 다양한 대체자료 활용, 행정자료 및 연간조사 자료 등을 활용하여 효율성을 증진시키고, 면접방식에서 전화조사, 인터넷, 등록조사 등 다양한 조사방식이 증가되고 있는 것을 특징으로 함
- 네덜란드 센서스는 1829년에서 1889년까지 내무부에서 실시하다 1899년 이후로 통계청에서 주관하고 있음. 비용절감, 사생활 보호, 무응답 증가 등 환경변화에 직면하여 1981년 이후 Virtual 센서스로 전환하였음. 센서스 자료원으로는 인구등록, 일자리정보, 연금 및 보험자료, 사회보장자료, 주택등록부 등 행정자료와 노동력조사 등 연간조사 자료를 혼합해서 활용. 현재 2020년 센서스 준비를 위해 교육정도, 소프트웨어 개발, 지리정보 개선 등 다양한 프로젝트를 진행중에 있으며 가장 중요한 사항은 제도기반 구축 및 믿을 수 있는 등록자료 구축 및 유지를 강조.

- 호주통계청은 1911년 이후 센서스를 실시해 오고 있고, 전통적인 센서스는 적정성, 시의성, 예산제약, 기술발전 등 변화요인으로 인해 조사방식 변화 필요성을 인지함. 바람직한 현대센서스는 센서스의 본질을 지키며 연간기구조사, 행정자료, 대체자료원 발굴 등을 통해 해결(Solution)중심의 센서스로 변화되어야 함을 강조함
- 토론자는 2010라운드 센서스 이행상황을 점검해보고 2020라운드센서스가 나아가야 할 방향에 대한 점검의 기회로 개별 발표자료의 중요성을 부각함. 10년 주기의 센서스 자료가 정책수요, 기술발전 등 환경변화에 맞을지 고민의 필요성도 제기하면서 개별국가는 변화되는 상황에 맞추어 조사방식을 발전시켜 나갈 것을 주문.

□ 시사점

- 2020 라운드 센서스는 기존의 조사방식에서 벗어나 등록센서스 기반으로 전환하는 국가가 많아 질 것으로 보여지는 바, 한국의 2015 센서스를 많은 국가들이 예의주시하고 있음
- 2015 등록 센서스를 성공적으로 마무리 하고, 이를 세계속으로 잘 전파하여 한국통계청의 위상도 제고하고 통계기준을 선도하는 위치확보가 필요해 보임

참고 17

(IPS 147) Internet Survey—Challenges and Opportunities

□ 세션개요

- 일시/장소 : 2015. 07. 29. 09:00 ~ 10:10 / 101
- 세션구성
 - Organizer/Chair : Vijay Nair / Natalie Shlomo
 - Presenter : Mick Couper "Internet Survey—Challenges and Opportunities"

□ 세션요지

- 인터넷조사의 개념, 현황, 방법, 미래에 대한 일반적인 내용을 설명함
- 1970년 Internet이 출현하고, 2004년 facebook의 대중화에 이어 2007년 I-Phone의 보편화로 인터넷 조사환경 기반은 성숙. 인터넷을 통한 조사의 경우 응답자선정의 경우 비확률 및 확률방식으로 선정될 수 있음. 비확률방식의 경우 편의성, 저렴한 비용 등을 가지고 자발적 online Panel, 응답지향표본, Social Media 모집방식 등이 있음.
- 확률방식의 경우 online intercept survey, 목록기반표집, 확률기반패널 등이 있음. 인터넷 조사의 경우 포괄범위 오차, 무응답 오차, 측정오차 등 해결해야 할 과제와 아울러 비확률표집의 경우 통계적 조정을 요함.
- 향후 인터넷조사는 정보화 시대에 맞추어 저렴하고 신속한 장점과 아울러 신뢰성을 높이기 위한 좋은 이론 및 추정모델의 개발 등 도전과제를 해결하여야 함
- 개도국 및 지역통계의 경우 인터넷 기반이 취약하여 도입하기 어려울 수 있다는 청중의 견해에 오히려 발표자는 보조수단으로서의 인터넷 조사가 위기극복의 기회를 더 많이 가져올 수 있다는 견해를 피력

□ 시사점

- 인터넷조사의 경우 공식통계 작성의 보조수단으로써 이미 널리 활용되고 있는 바 모바일 환경에 친숙한 방식으로 더욱 개선해 나갈 필요성이 있음
- 이용자 만족도 조사 등 기업의 마케팅 용도로 편의성, 비용절감 등 장점이 많아 그 활용도는 높아질 예정임 감안하여 공식통계에도 활용할 수 있는 모델구축, 조정기법 개발 등 품질제고를 위한 노력을 지속적으로 실시

□ 세션개요

- 일시/장소 : 7. 27 / Room 101
- 세션구성
 - 좌장 : Danny Pfeffermann
 - 발표 : Jon Rao / Wayne A. Fuller
- 참석인원 : 300명

□ 세션요지

- 표본조사의 이론 및 방법의 과거, 현재, 미래방향에 대해 두 발표자가 차례로 발표를 실시함
- (발표자 : Jon Rao) Kier(1817), Laplace(1814)의 무작위 추출에 이어 1925년 ISI 위원회의 이론 및 실제 보고서, Bowley(1926)의 층화무작위 추출, Neymen(1934)의 최적할당추출, 미국센서스국(1940-1960)의 층화 2단계추출, Hansen and Hurwitz(1946)의 비응답처리를 위한 2단계 표본 등 표본조사의 획기적인 발전을 도모한 이론을 간략히 소개함
- (발표자 : Wayne A. Fuller) 계산과 분석에 대한 개념설명을 시작으로 결측자료 처리방법인 대체근접치, Raking, 회귀분석, 불완전유사표본, Little Models등을 소개함
 - 아울러, Hot Deck imputation(1947), Multiple Imputation(1976,1987), 소지역추정기법 등 현재까지 영향을 주는 표본자료 보정기법의 중요성을 언급
- 표본조사가 직면하고 있거나 직면할 도전과제로 시의성에 있어 자료수집방식 Web 활용 증대, 지리정보소프트 웨어 출현 및 활용, 보조자료인 행정자료의 증가와 추정치의 어려움, 로봇을 활용한 Data 수집 및 Monster 자료 출현 등을 지목함
- 모바일 자료 수집, e-census 출현 등 기존의 현장조사방식에서 탈피한 새로운 자료수집 방식의 출현에 어떻게 대응할 지에 대한 청객들의 추가적인 질의가 있었고, 신 조사방식에 함께 대응하고 고민할 시기임을 강조

□ 시사점

- 통계학에서 역사적으로 발전해 온 가구 및 사업체 단위 조사의 표본기법 정리, 무응답처리, 임putation 등 표본이론 대한 종합점검은 매우 유익
- 급격히 변화되고 있는 조사환경에 맞추어 통계인들이 현명하게 대처하는 방안을 공동으로 강구할 필요성이 대두되고 있음을 공감