

---

# 2010 Australian Statistical Conference

## 참가 결과보고서

---

2010. 12.

## < 목 차 >

I. 회의참가(출장) 개요 .....	1
II. ASC2010 소개 .....	2
III. 표본설계(Sampling Design) .....	7
□ 회의주요안건 .....	7
□ 개별 발표 내용 .....	8
1. 호주 소득 및 노동패널조사에서 CAPI 도입의 영향 .....	8
2. 호주 원주민조사 표본관리 및 조정 .....	12
3. 실제세계에서 샘플링 .....	14
IV. 결론 .....	18
V. Post-Conference Workshop .....	20
□ 개요 .....	20
□ R의 주요특징 및 시사점 .....	21
< 부 록 >	
1. 2010 Austrian Statistical Conference 참가 결과 보고서 제1권	
2. 2010 Austrian Statistical Conference 참가 결과 보고서 제2권	
3. 2010 Austrian Statistical Conference CD	

## I 회의참가(출장) 개요

### □ 목적

- 새로운 통계분석기법 도입 및 분석능력 향상을 위한 새로운 분석 프로그램 활용능력 습득
- 선진국의 통계분야에 대한 역할, 발전 방안 등을 파악하여 향후 우리청의 다양한 분야의 통계에 대한 역할 정립을 위한 정보수집
- 통계청의 경제통계, 사회통계, 환경통계, 조사방법론 업무담당자들의 훈련참가를 통하여 각 분야의 이론 및 실무능력을 향상시키고, 세계 각국의 통계전문가들과의 인적네트워크 형성을 통해 지속적으로 개선 개발을 도모

### □ 기간

- 2010.12.4.(토) - 13.(월) (10일)
- 호주 프리멘틀

### □ 출장자 : 총4명

- 정구현(5급, 사회통계국 복지통계과)
- 강동환(6급, 경제통계국 경제기획통계과)
- 오정화(6급, 통계개발원 연구기획실)
- 안다영(7급, 통계개발원 조사연구실)

## II ASC2010 소개

- 표제회의(ASC2010)는 20번째 회의이며, 특히 Western Australia에서 개최되는 첫 번째 회의임
    - “Statistics in the West : Understanding Our World” 라는 주제로 각 분야의 전문가 및 저명한 국내외 인사들이 참가
  
  - 회의목적
    - 통계관련 새로운 연구결과 및 진행사항 등의 정보를 참가자에게 제공
    - 통계관련 다양한 분야의 연구자간 상호 교류 협력
    - 세계수준의 통계학자를 활용한 발표 및 토론을 통한 전문 지식을 공유
  
  - 금년회의에는 특히 10명의 저명한 기초연설자\*의 발표 및 주제별 5~6명 토론자가 참가하여 열띤 토론 진행
- \* Professor Barry Marshall(The University of Western Australia, winner of the 2005 Nobel Prize)
- Dr Alan M Zaslavsky(Harvard Medical School, USA)
  - Professor Denise Lievesley(King's College London, UK)
  - Professor Adrian Baddeley(The University of Western Australia)
  - Professor Tadeusz Bednarski(Wroclaw University, Poland)
  - Professor Noel Cressie(The Ohio State University, USA)
  - Professor Persi Diaconis(Stanford University, USA)
  - Professor Jerry Friedman(Stanford University, USA)
  - Dr Gordon Smyth(WEHI, Melbourne)
  - Professor Chris Wild(University of Auckland, NZ)
- 
- 40개 세션(통계분야)별 4~6개의 연구결과 발표 및 질의답변 진행 (총 220여개 논문 발표)

□ 회의기간 중 세션별 발표 논문 수

세션구분	발표 논문(수)	비고
<b>1. 단독세션</b>	<b>10</b>	
1.1. KEYNOTE ADDRESS	8	
1.2. AMSI LECTURER	1	
1.3. E K FROEMAN LECTURER	1	
<b>2. 공동세션(통계분야별)</b>	<b>220</b>	
2.1. Time Series(1)	5	
2.2. Time Series(2)	6	
2.3. Statistical Inference(1)	5	
2.4. Statistical Inference(2)	5	
2.5. Official Statistics	5	
2.6. CSIRO Biosciences	4	
2.7. Yong Statisticians(Where statistics can you take)	6	
2.8. Prediction	6	
2.9. Social Science	5	
2.10. Bayesian Statistics(Finance)	6	
2.11. Bayesian Statistics(Environment)	6	
2.12. Bayesian Statistics(Medicine)	6	
2.13. Mining & Multivariate Statistics	6	
2.14. Data Linkage Session	6	
2.15. Wildlife Analyses	6	
2.16. BCA Session(Current topics in Biostatistics)	4	
2.17. Survey Analysis	6	
2.18. Australian Pharmaceutical Biostatistics Group	5	
2.19. Stochastic Models	6	
2.20. Stochastic Models	5	

2.21. Stochastic Modelling	5	
2.22. Multivariate Environmental Statistics	6	
2.23. Environmental Theory	5	
2.24. Sample Design	5	
2.25. Spatial Statistics	6	
2.26. Spatial Statistics Theory	6	
2.27. Spatial/Environmental Statistics	6	
2.28. Imputation and Anslysis	6	
2.29. Inference	4	
2.30. Biostatistics(3)	5	
2.31. Biostatistics Analysis	6	
2.32. Biostatistics/Bioinformatics	6	
2.33. Biostatistics(1)	5	
2.34. Biostatistics(Design)(2)	6	
2.35. Applied Biostatistics	6	
2.36. Transport and Experimental Design	5	
2.37. Agriculture	6	
2.38. Finance	6	
2.39. Climate	6	
2.40. Environmental & Agricultural Statistics	5	

## □ 주요내용(접근방법 및 시사점)

### 1. Trustworthy statistics

(Professor Denise Lievesley, Chair, European Statistical Advisory Committee)

- (주제) 입증된 사실에 기반한 정책수립을 위한 공식통계의 중요성, 공식통계의 정합성에 대한 위협요인에 대한 고찰, 신뢰할 수 있는 통계작성을 위한 전문 통계인들의 의무와 역할
  - 신뢰구축의 핵심요소로 9가지 사항을 제시
    - 통계청의 자치권, 통계법, (독립적) 통계위원회, 행동강령 개발, 법위반에 대한 집행, 정치적 중립성, 이용자 의견 반영, 작성과정의 외부 감독, 국회 보고
- (접근방법) 다소 이상적인 입장에서 종종 사용되는 특정개인이나 그룹의 검증되지 않은 여론에 의한 의사결정에 대한 문제점 지적 및 정책순환\* 과정별 사실(통계) 기반 의사결정의 당위성 역설
  - \*아젠다 형성, 쟁점을 명확화, 구체적 대안설정, 실행방법 명확화, 과정별 실적이나 효과에 대한 모니터링 과정
  - 통계기반 정책수립의 기본조건은 통계의 신뢰성이며, 이는 통계의 품질, 통계제도의 품질, 통계인의 품질에 의존
- (시사점) “통계에 대한 신뢰는 장기간에 걸쳐 천천히 쌓여가지만, 신뢰의 붕괴는 순간적” 이라는 생각을 갖고 통계업무 실행

" Trust comes on foot, But leaves on horse back"

- Johan Thorbecke, Dutch statesman

## 2. The impact of introducing computer assisted personal interviewing to the HILDA\* survey

(Nicole Watson, University of Melbourne)

\* Household, Income and labour Dynamics in Australia

- (주제) HILDA 조사관련 기존의 종이조사표 활용방식에서 컴퓨터 활용의 대중화, 사생활보호에 대한 관심 증가를 반영하여 CAPI 방식으로 전환함에 따른 통계품질, 응답률, 시계열 단절 측면에서 효과분석
  - CAPI 도입에 따른 항목무응답, 주요변수들에 대한 연속성 관련 검토
  - CAPI 도입에 따른 전반적인 응답부담에 대한 효과 검토
  - 종속 면접방식의 효과 검토
- (접근방법) 성별, 나이, 주택소유여부, 가구규모, 혼인상태, 영어구사능력, 고용상태를 반영한 응답대상 선정 및 Split-Sample 검정 실시
- (연구결과) 응답률, 응답자 반응, 면접자 반응에는 차이가 유의하지 않았으나, 응답시간 감소로 조사량 증가, 서술형응답은 증가하였으며, 누락항목 및 임금 잘못 항목 감소



### Ⅲ

## 표본설계(Sampling Design)

### □ 회의 주요 안건

- 본 섹션은 호주의 대표적인 표본조사를 중심으로 표본설계 및 관리에 대한 실제 조사사례를 가지고 발표함

- 호주의 대표적 패널조사인 HILDA\*는 최근 종이 조사표에서 CAPI로 전환함에 따라 발생하는 영향, 장단점, 시사점을 소개함

\* Household Income and Labor Dynamics in Australia

- 호주는 원주민(aborigine)에 대해 정책적, 문화적으로 각별한 배려를 하고 있으며, 원주민들의 실태 및 삶(복지상태) 파악을 위해 여러 표본조사를 실시

그러나 그들이 지역적 산재가 넓게 퍼져 있어 조사마다 상이한 결과가 추정되어 정확한 결과를 보정하기 위한 별도의 보정(adjustment) 작업 필요성 및 방법 제시

- 호주의 산업분류는 1993년 17개(ANZSIC93) 대분류로 구분하고 있으나, 2006년 신분류 체계로 전환되어 19개(ANZSIC06) 대분류를 사용

신분류의 연착륙 및 과거 시계열연장을 위해 2009년 조사는 두 가지 분류를 동시에 사용하여 조정하는 방법을 제공

## □ 개별 발표 내용

### 1. 호주 소득 및 노동패널조사에서 CAPI 도입의 영향(The impact of introducing CAPI to the HILDA Survey)

□ 발표자 : Nicol Watson([n.watson@unimelb.edu.au](mailto:n.watson@unimelb.edu.au))

#### □ 주요내용

- 호주 소득 및 노동패널조사(HILDA) 소개
  - 호주의 가족, 소득, 노동에 관한 호주의 대표적인 패널조사
  - 연간조사로 지속적인 패널조사로 가구 및 15세이상 가구원\*을 조사
    - \*가구원 질문서 및 자계식(self completion) 질문서 작성
- 자료수집 방법 전환 영향
  - CAPI는 다양한 장점제공
    - 조사과정 단순화, 응답자 일치성 현장검증, 조사와 동시 자료입력, 자료 질적 제고, 파라데이터\* 수집가능, 기존자료 활용 등
      - \*조사시간, 방문회수(call back) 등
  - 시계열 단절 위험 상존
- CAPI 도입배경
  - 컴퓨터 사용 일반화, 사생활 보호의식 강화
- 자료수집 방법 전환을 위한 시험조사(split-sample test) : 2007년
  - 특정년도(2007) 종단패널을 위해 질문지 설계 및 절차 테스트
  - 가구 보다는 응답개인에 초점을 두어 비교

	종이	CAPI
가구 조사표	366	343
가구원 조사표	702	671

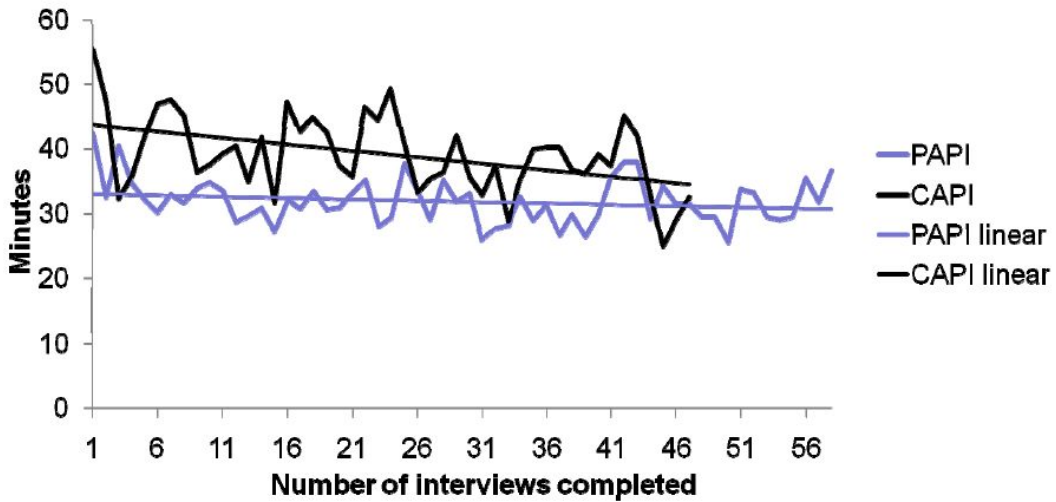
○ 응답자 특성(2007)

	종이	CAPI	차이
남자	48.6	45.3	-3.4
평균연령	42.4	44.7	2.3**
NSW	51.4	50.7	-0.8
자가	70.2	71.6	1.4
가구원	3.2	3.1	-0.1
결 혼 상 태			
결혼	50.3	51.4	1.1
사실혼	9.1	9.5	0.4
별거/이혼/사별	14.5	14.9	0.4
미혼	26.1	24.0	-2.1

	종이	CAPI	차이
언어 적합도			
영어만 사용	80.2	80.6	0.4
비영어 및 영어활용 좋음	17.2	17.1	-0.1
비영어/영어활 용도좋지않음	2.4	2.2	-0.2
노 동 상 태			
취업	63.5	63.6	0.1
실업	4.6	3.1	-1.4
비경	31.9	33.2	1.3

○ 도입 후 영향

- (-) : 응답률, 응답자 반응, 조사원 반응은 크게 변화 없음
- (↑) 응답시간 감소로 조사량(건수) 증가, 정부 보조금 항목에서 잘모름 증가, 중복항목 응답 늘어남, 서술형 응답 증가
- (↓) 누락항목 감소, 임금 항목 잘모름 감소



○ CAPI 조사표는 종이조사표와 달리 설계

종이조사표	CAPI 조사표																														
<p>F53b For how many weeks last financial year did you receive the [specify pension / allowance]? [FOR EACH ONE RECEIVED, WRITE IN NUMBER IN COLUMN B BELOW.]</p> <p>F53c Including only your share, how much did you receive in total income from the [specify pension / allowance] last financial year? Please include any lump sum advances you received, but do not include any bonus payments previously mentioned. [FOR EACH ONE RECEIVED, WRITE IN AMOUNT IN COLUMN C BELOW.] IF RESPONDENT DOES NOT KNOW YEARLY AMOUNT ASK: What about the average received per fortnight from the [specify pension / allowance]? Are you able to estimate what that amount was? WRITE IN AMOUNTS IN COLUMN D BELOW.</p> <table border="1"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>OR</th> <th>D</th> </tr> <tr> <th></th> <th></th> <th>No. of weeks received</th> <th>Annual amount</th> <th></th> <th>Average per fortnight</th> </tr> </thead> <tbody> <tr> <td>Age Pension (from Australian Govt) .....</td> <td>01</td> <td></td> <td>\$</td> <td></td> <td>\$</td> </tr> <tr> <td>Newstart Allowance .....</td> <td>02</td> <td></td> <td>\$</td> <td></td> <td>\$</td> </tr> <tr> <td>Mature Age Allowance .....</td> <td>03</td> <td></td> <td>\$</td> <td></td> <td>\$</td> </tr> </tbody> </table>		A	B	C	OR	D			No. of weeks received	Annual amount		Average per fortnight	Age Pension (from Australian Govt) .....	01		\$		\$	Newstart Allowance .....	02		\$		\$	Mature Age Allowance .....	03		\$		\$	<p>F53b. For how many weeks last financial year did you receive the Age Pension (from Australian Govt) ?</p> <p><input checked="" type="checkbox"/> No. of weeks received 52</p> <p><input type="checkbox"/> Refused</p> <p><input type="checkbox"/> Don't know</p> <p>F53c. Including only your share, how much did you receive in total income from the Age Pension (from Australian Govt) last financial year? Please include any lump sum advances you received, but do not include any bonus payments previously mentioned.</p> <p><input checked="" type="checkbox"/> Annual amount (whole \$) 13000</p> <p><input type="checkbox"/> Refused</p> <p><input type="checkbox"/> Don't know</p> <p>If necessary, please provide a comment for any unusual responses.</p> <p>HHID Z851 - ALFRED (88) PQ - F - Income</p>
	A	B	C	OR	D																										
		No. of weeks received	Annual amount		Average per fortnight																										
Age Pension (from Australian Govt) .....	01		\$		\$																										
Newstart Allowance .....	02		\$		\$																										
Mature Age Allowance .....	03		\$		\$																										

○ CAPI 조사표로 변경 후 영향

- 객관적 문항
  - 노동문항(19개 항목), 주택(2개)은 변화 없음
  - 임금(5개) 중 정부보조 항목, 흡연 증가
  - 채식여부 증가
- 주관적 문항
  - 고용만족도, 건강(2개 항목)은 감소
  - 가족 및 아동(4개), 독서 및 계산 능력(4개) 변화없음

- 2009년(wave 9) 전면적 CAPI 도입 및 영향
  - CAPI는 컴퓨터에 펜을 사용하여 활용 쉬움 , 비용절감 기대
  - 최고의 응답율 기록(96.3%), 장기간 인터뷰 가능
  - 금액 변수에 대해 잘모름/거절 증가
  - 산업/직업에 대한 자세한 내용 조사 증가
  - 기존(보조)자료 사용에 따른 일치성 증가

□ 결론

- 2007년 시험조사는 CAPI 도입에 대한 긍정적 효과 제공
  - 데이터 질적 개선( 현장에서 입력 및 내검, 개방형 항목 응답증가, 응답내용 질적제고\*  
\*사회적 기대 항목 편향감소
  - 조사건수 확대는 우려
  
- 2009년 조사결과는 성공적이거나 금액항목에 대해 항목불응 증가 우려

---

호주의 CAPI 시스템

---



## 2. 호주 원주민조사 표본관리 및 조정(Sampling monitoring and adjustments for indigenous surveys)

□ 발표자 : Tamie Anacotta([tamie.anakotta@abs.gov.au](mailto:tamie.anakotta@abs.gov.au))

### □ 주요내용

#### ○ 발표내용 요약

- 호주 원주민은 호주 사회, 문화에 있어 독특하고 중요한 특징을 지님
- 호주 통계청은 원주민에 대해 정확한 자료수집을 위해 노력 경주
- 다만 원주민을 확인하고 찾기 위해 비용 대비 효율적인 조사도모
- 높은 이동을 및 응답의 협조가 조사결과에 큰 영향 미침
- 불확실성은 표본규모 큰 변동, 비용의 비효율, 부정확한 추정값 발생 가능
- 호주 통계청은 조사원으로부터 표본 모니터링에 초점 집중
- 조사기간 중 주간마다 응답율 및 표본탈락율(sample loss rate) 고려하여 표본을 증가하는 표본설계를 적용

#### ○ 원주민 조사 결과

State	Fully Responding% Target = 85%	Partly Responding%	Incomplete%	Non Response%	Refusal%	Sample Loss%
AUST	85.49	1.76	0.96	9.23	2.56	13.86
ACT	86.5	3.27	0.61	6.75	2.86	9.44
NSW	82.18	1.93	0.45	12.17	3.28	14.17
NT	83.71	1.81	0.45	8.82	5.2	12.13
QLD	88.98	0.84	1.85	5.59	2.74	12.57
SA	87.71	1.14	1.2	8.3	1.65	15.25
TAS	93.38	0.41	0.14	4.14	1.93	14.5
VIC	80.98	2.97	1.55	12.46	2.04	14.45
WA	87.63	1.53	0.21	8.27	2.36	14.04

○ 조정내용 및 영향

사건	영향	조정
기대보다 높은 응답율	비용 증가	표본크기 감소
기대보다 낮은 응답율	조사목적 미달	표본크기증가(편향발생가능)

○ 호주 원주민 조사

- 매 3년마다 두 조사\*를 번갈아 가면서 원주민 조사 실시

\* 원주민 건강조사(National Aboriginal and Torres Strait Islander Health Survey),  
원주민 사회조사(National Aboriginal and Torres Strait Islander Social Survey)

○ 호주 원주민(토레스해협 거주자 포함) 인구

	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	전체
원주민	133,184	29,143	121,766	24,483	54,883	16,410	51,055	3,737	434,661
인구구성비	2.1%	0.6%	3.3%	1.7%	3.0%	3.5%	27.8%	1.2%	2.3%

○ 호주 통계청 모니터링

- 출현율(P<확인> \* P<응답>) 파악
- 응답율(전체 응답가구 및 부분 응답가구)
- 전체내용을 응답한 성인 및 아동 규모
- 조사비용

○ 호주통계청 모니터링 결과

- 기대보다 낮은 출현율 시현
  - 이주, 조사거부, 자료수집 방법 효과(면접조사 vs 종이신고(paper form), 확  
인 질문(screen questions) 불응

	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	전체
원주민 가구 기대수	391	511	118	243	430	229	88	81	2091
실제 원주민 조사수	322	342	80	172	346	161	46	64	1533

- 호주통계청 보완내용
  - 조사기간 동안 표본크기 증가
  - 무응답 오차를 설명하기 위한 추가적인 보조자료(benchmarks) 적용
- 호주통계청 향후계획
  - 조사기간 동안 데이터 모니터링 및 자료수집 개선
  - 확인질문 검토

### 3. 실제세계에서 샘플링(Sampling in the real world)

□ 발표자 : Martin Caruso([martin.caruso@abs.gov.au](mailto:martin.caruso@abs.gov.au))

□ 주요내용

- 발표내용 요약
  - 고용자조사는 목표모집단의 임금 및 고용과 같은 변수를 추정하기 위해 사업체 프레임으로부터 리스트 추출로 조사 실시
  - 시간에 따른 리스트 변화, 사업체 표본의 변화는 실제 세계 변화와 무관하게 조사 추정치에 영향을 미침
  - 호주 통계청은 시간에 따른 리스트 변화를 반영하여 통계목적에 적합하도록 표본 재설계 검토
  - 산업분류는 매 10-15년 마다 변경되며 이는 새로운 산업 발생, 산업 간 변경 등을 반영
  - 호주 통계청은 중소기업의 응답부담을 감소하기 위해 연동표본 사용
  - 2009년 평균임금 조사는 과소포함을 줄이기 위한 신 프레임 사용, 신 산업분류(ANZSIC06) 도입, 신 표본설계를 적용함
  - 이러한 변화를 측정하기 위해 舊 프레임을 통한 舊 표본설계, 舊 산업분류로 구 사업체를 조사
  - 과거 표본과 중복을 최대한으로 한 신 표본(parallel sample)은 신 표본설계, 2006 산업분류, 신 표본설계로 추출된 사업체로 조사



- 신·구 표본 영향을 주요 요소별 분해하고 이러한 영향이 변동과정의 요소에서 발생하는지 특정 조사 단위(사업체)에서 발생하는지 파악
- 고용자 조사 개요
  - 주간 임금은 연금의 물가연동자료 및 남녀 임금격차 비교 자료 활용
  - 지정된 주에 사업체로부터 고용 및 임금자료 수집
    - 평균 임금 = 전체 임금 / 전체 고용자
  - 이용자 관심사는 분기별 움직임에 초점 맞춤
  - 응답부담을 감소하기 위해 연동표본제(랜덤추출) 사용하여 사업체 선택
  - 평균 주간임금은 비가중(ratio weights)을 사용
- 표준산업분류 변경
  - 호주, 뉴질랜드 산업분류(ANZSIC93)는 1993년에 시작
  - 신 산업분류(ANZSIC06)는 2006년부터 도입, 정보와 통신을 분리
  - 광업, 제조업 등 산업을 총화하여 분류
  - ANZSIC93는 17개 분류에서 ANZSIC06은 19개 분류로 확대
- 표준산업분류 적용 검토
  - 변경 후 프레임으로부터 과소포함 발생, 표본규모 조정 등 표본재설계 필요
  - 전면 변경에 따른 시계열 단층 발생 우려
- 표본설계
  - 표본 재설계는 특별히 2개 분류에 대해 효율성 제고 필요
  - 기존 프레임으로 ANZSIC93과 ANZSIC06 두 개의 코드를 갖춤
  - ANZSIC06로 전면 이행을 3년간 늦춘 것은 표본설계를 위한 몇 년간의 두 개의 코드 필요
- 두 개의 평균 주간임금 표본
  - 기존 표본은 기존 프레임에 1993년 산업분류를 적용
  - 신표본은 새로운 프레임에 2006년 산업분류를 적용

- 두 개 표본조사의 목적은 추정량의 변동분을 측정
- 2006년 산업분류로 과거 시계열 추정
- 2009년 5월 조사, 2009년 8월 공표시 2006년으로 변동을 점검

○ 조사결과

- 2개 주(州)에서 주별 수준이 큰 차이 발생
- 추정값의 큰 변화는 산업별 수준에서 발생

State	March 2009 A93 estimates	May 2009 A93 estimates	May 2009 A06 estimates	August 2009 A06 estimates
Average Weekly Ordinary Time Earnings Full-time Adult males				
WA	\$1,470.70	\$1,488.80	\$1,401.30	\$1,424.30
Average Weekly Earnings Full-time Adult females				
Tasmania	\$1,011.90	\$1,028.40	\$944.70	\$974.90

○ 변동분을 구성요소별로 분해

Main Components	Link to Equation
Movement/Change	$\sum_{i \in c} \frac{w_{2i}(y_{2i} - y_{1i}) - A\hat{W}E_{1c} \{w_{2i}(x_{2i} - x_{1i})\} + y_{1i}(w_{2i} - w_{1i}) - A\hat{W}E_{1c} \{x_{1i}(w_{2i} - w_{1i})\}}{\hat{X}_{2c}}$
Common Unit/Stratum Flip (comm)	$\sum_{i \in c \cap comm} \frac{w_{2i}(y_{2i} - y_{1i}) - A\hat{W}E_{1c} \{w_{2i}(x_{2i} - x_{1i})\} + y_{1i}(w_{2i} - w_{1i}) - A\hat{W}E_{1c} \{x_{1i}(w_{2i} - w_{1i})\}}{\hat{X}_{2c}}$
Framework (fw)	$\sum_{i \in c \cap fw} \frac{y_{1i}(w_{2i} - w_{1i}) - A\hat{W}E_{1c} \{x_{1i}(w_{2i} - w_{1i})\}}{\hat{X}_{2c}}$
Rotation in/Birth (rib)	$\sum_{i \in c \cap rib} \frac{w_{2i}y_{2i} - A\hat{W}E_{1c} w_{2i}x_{2i}}{\hat{X}_{2c}}$
Rotation out/Death(rod)	$\sum_{i \in c \cap rod} \frac{A\hat{W}E_{1c} w_{1i}x_{1i} - w_{1i}y_{1i}}{\hat{X}_{2c}}$

○ 조사단위별 정렬

- 위의 공식을 활용하여 사업체를 순위별로 정렬하여 특이값 발견

□ 결론

- 호주통계청은 일반화 회귀추정 및 비추정을 사용하여 보조변수를 보정
- 장기시계열을 수정하기 위함
- 표본 재설계는 추정값의 효율성을 제고

## IV 결론

### □ 우리나라 현황

- CAPI는 현재 물가조사 등에서 도입하여 조사활용
  - 다만 CAPI 도입을 자료수집 방법의 변화로만 간주하여 도입 전·후 조사에 대한 영향 평가는 부족한 실정
- 현재 통계청은 소수집단(예. 모자가구, 장애인가구, 다문화가구, 외국인 가구 등)에 대한 통계개발은 미흡한 상태
- 분류변경에 따라 조사별로 시계열 연장을 실시하고 있음
  - 다만, 소급작성시 표준화된 방법 및 차이점을 요인별로 분석하는 작업은 미흡한 실정

### □ 시사점

- CAPI는 조사표 변경을 넘어 조사원 조사용이, 응답자 부담감소 측면에서 개발 검토가 필요
- CAPI에 사용되는 노트북, PDA, 넷북을 좀 더 조사원 친화적으로 변경 필요
  - 컴퓨터 무게 및 스크린 크기 등을 고려
  - 조사표를 종이조사표와 달리 CAPI에 맞는 조사표 설계
  - 민감한 내용은 응답자가 직접 기입할 수 있도록 설계 필요
- CAPI 도입으로 응답율, 조사항목의 완전성, 항목간 일치성 등 평가 필요
  - 신기술 도입에 따른 전반적인 응답부담 영향
  - 시계열 단절 등 자료에 미치는 영향
  - 세부항목의 변화 등을 검토

- 현재 증가하고 있는 다문화가구 등 소수집단에 대한 조사가 요구될 것임
  - 출현율이 적은 조사의 경우 표본의 모니터링을 통해 표본규모를 조정하는 표본설계는 획기적인 아이디어라고 판단됨
  
- 분류 개편에 따른 시계열 단절 현상을 막기 위해 신·구분류를 일정 기간 동시 조사 필요
  - 특히 신구분류를 병행 조사하여 소급자료에 대한 이론적 추론이 필요

**□ Workshop 개요**

- 주제 : R의 기초, Analysing spatial point patterns in R
- 기간 : 2010. 12. 10. ~ 11.(2일간)
- 강사 :
  - Martin Hazelton(Chair of Statistics, Massey University New Zealand) and Berwin Turlach(A/Professor of Statistics, University of WA)
  - Adrian Baddeley(CRIRO/University of Western Australia)

**□ Workshop 세부 내용****1. Introduction to R**

- 1.1. R 프로그램 설명
- 1.2. R 기본명령어(Basic Syntax)
- 1.3. Linear Modelling in R
- 1.4. R Base Graphics
- 1.5. Generalized Linear Models in R
- 1.6. 프로그램 작성 실습

**2. Analysing spatial point patterns in R**

- 2.1. Overview
- 2.2. Data Types & Data Entry
- 2.3. Intensity
- 2.4. Poisson Models
- 2.5. Interaction
- 2.6. Gibbs Models
- 2.7. Marked Point Patterns
- 2.8. Higher Dimensions and Other Spatial Data
- 2.9. 프로그램 작성 실습

## □ R의 주요 특징

- R은 R. Becker, J. Chamber, A. Wilks 등이 S 언어를 확장하여 개발한 프로그램 언어임
  - 동시에 SAS, SPSS, Minitab 등 통계패키지에서 수행할수 있는 기능을 탑재하여 사용자의 의도를 반영한 프로그램 작성에 유용함
- R은 현재 초기 프로그램 언어의 미숙으로 우리청에서 도입되지 않고 있으나, 다른 프로그램 사용에 대한 보완 기능으로 활용 가능, 통계 자료처리 및 분석 업무 활용에 많은 특징과 장점을 보유
  - 구입 및 유지비용이 전혀 들지 않는점, 타 프로그램보다 우월한 Graphic 기능, 사용자의 요구에 맞추어 새로운 함수 생성 기능 등
- 현재는 통계학 및 계량분석자들이 R 언어를 이용하여 자료 분석을 위한 새로운 프로그램을 개발하고, 이를 패키지화하여 R에 포함시키기 위한 노력이 진행 중에 있음
- R프로그램은 금융, 생명공학, 조사, 지리정보 등의 분야에서 많이 활용되고 있음

## □ 시사점

- (통계 Package 활용성) 현재 다용량 자료처리에는 SAS, SPSS 등의 프로그램의 활용이 필요하나 소규모의 표본조사, 또는 분석용 자료 처리에도 구분없이 사용료를 지불 활용하고 있음
  - 구입 및 유지보수 비용이 없고, 사용자요구에 맞는 새로운 함수 생성기능이 우월한 R을 활용함으로써 자원절약 및 통계분석 및 보고서 작성 기능 향상을 도모

- (R 프로그램 활용확대) 현재 R 프로그램은 우리청내에서 제한적으로 사용되고 있음(이용자 측면, 프로그램 기능 측면)
  
- 이번 연수를 통하여 습득한 지식의 청내 대중화를 목적으로 “R 프로그램 활용 모임” 운영 및 이를 통한 통계분석기법 향상을 기할 것임

## 부 록

1. 2010 Austrian Statistical Conference 참가 결과 보고서 제1권
  - 회의 개요 및 발표자료 모음
2. 2010 Austrian Statistical Conference 참가 결과 보고서 제2권
  - Post-Conference Workshop
3. 2010 Austrian Statistical Conference CD
  - 참가결과보고서, 발표자료 모음, R 프로그램 및 연습용데이터와 교육자료 수록함