

제2장

표본조사에서의 가중값 조정방안 - 추정방법 중심으로 -

임경은

제1절 서론

표본조사에서의 가중값은 설계가중값, 무응답에 대한 조정, 사후층화에 대한 조정 등의 세 가지 가중값을 통합하여 산정된다. 일반적으로 모집단의 특성값인 모수에 대한 추정 단계에서 가중값을 이용하면 모수에 대한 비편의추정량(unbiased estimator)을 얻을 수 있으며, 가중값을 이용하지 않을 경우 심각한 편의(bias)가 발생할 수 있다. 특히 통계청에서 실시하고 있는 각종 표본조사와 같이 표본의 크기가 큰 대규모 조사에서는 추정량의 편의가 문제가 되므로 추정 과정에서 반드시 가중값을 이용하여야 한다. 또한 표본조사의 형태에 따른 추정량의 계산 방식에도 차이가 있으며 이에 따른 가중값의 적용도 달라지게 된다.

표본조사에서 반복조사를 하는 경우, 지정된 표본을 고정하여 조사하는 고정표본제(fixed sampling system)와 매 조사마다 표본의 일부를 교체하는 연동표본제(rotation sampling system)로 표본조사를 구분할 수 있으며, 연동표본조사에서의 추정 방법은 기존에 이용하는 추정 방법과는 다른 방법을 이용하게 된다.

연동표본조사(rotation panel survey)란 시점의 변화에 따라 일정 비율의 부표본(sub-sample)을 표본에서 제외시키고 새로운 부표본으로 표본

을 대체시키는 표본추출방법을 의미하는 것으로, ‘연속조사에서 조사 단위의 부분적인 교체추출법(sampling on successive occasions with partial replacement of unit)’ (Yates, 1949; Patterson, 1950)과 ‘시계열을 위한 표본 조사법(sampling for a time series)’ (Hansen, Hurwits, and Madow, 1953)으로도 불린다.

연동표본추출에 관한 연구는 Jessen(1942)에 의해 처음 설계되었으며, 이후 Eckler(1955)는 표본이 높은 상관관계를 가질 때 연동표본추출의 효율이 높다는 것을 보였다. Wolter(1979)는 연동표본의 추정 방법으로 최소분산추정량(Minimum Variance Unbiased Estimator: MVUE)을 이용하였으며, Rao & Graham(1964)은 복합추정량(Composite Estimator)을 이용하였다. Bailer(1975)는 연동그룹편의(rotation group bias)가 추정량에 미치는 영향에 관하여 연구하였으며, Iachan & Jones(1987)는 연동표본조사가 표본오차와 비표본오차를 동시에 줄이는 효과가 있음을 증명하였다. 박유성·김나영(1997)은 연동표본추출방식의 일반화에 관한 연구를 수행하였으며, 김영원·이수희(2001)는 연동표본설계의 효율성에 관한 연구를 진행하였다. 최근에는 조사기간 동안 동일한 표본단위를 조사한 후 교체그룹 자체를 교체하는 다수준 연동표본설계에 대한 연구가 박유성·배경화·김기환(2001), 박유성·문원기·김기환(2002), 김기환·박유성·남궁재은(2004)에 의해 활발히 이루어지고 있다.

본 연구에서는 통계청에서 실시하고 있는 여러 통계조사 가운데 매월 실시하고 있는 경제활동인구조사(이하 경활조사)와 가계조사를 대상으로 연동표본조사에서의 추정 방법과 추정량의 정도를 높이기 위한 가중값 선택에 대하여 알아보고자 한다. 경활조사와 가계조사는 2003년 표본개편에 의한 표본설계를 따르고 있으며, 2000년 5월부터 2002년 6월까지 약 2년에 걸친 시험조사를 거친 후 2004년 5월부터 인천과 경남 지역에 대해 선행 조사를 실시하고, 2005년 1월부터 전국에 걸쳐 본격적으로 연동표본제를 실시하고 있다. 그러나 조사방법의 변화에도 불구하고 추정량은 기존의 고정표본제에서 사용된 방법을 동일하게 적용하고 있는 실정이다. 따라서 본 연구에서는 연동표본제를 실시하고 있는 조사에 대하여 연동표본조사에서의 추정기법인 AK 복합추정량을 적용하는 문제에 대하여 알아보고자 한다.

먼저 제2절에서는 경제활동인구조사와 가계조사의 개요와 해당조사에서의 표본설계 및 적용 중인 연동표본조사에 대하여 살펴보고, 제3절에서는 현행추정방법에 대하여 알아본다. 이어서 제4절에서는 기존의 연구결과를 이용하여 경제활동인구조사와 가계조사에 사용할 수 있는 AK 복합추정량(AK composite estimator)과 AK 복합추정량의 분산을 유도한다. 제5절에서는 제4절에서 구한 분산식을 경제활동인구조사에 적용하여 최적의 가중값인 A 와 K 를 정하고, 이 때의 AK 복합추정량을 구하여 기존의 추정량인 비추정량과 비교하고자 한다.

제2절 경제활동인구조사와 가계조사

통계청에서 매월 실시하고 있는 경황조사는 국민의 경제활동, 즉 국민의 취업, 실업 등과 같은 경제적 특성을 조사하여 거시경제 분석과 인력자원의 개발정책 수립에 필요한 기초자료인 노동공급, 고용구조, 가용노동시간 및 인력자원 활용 정도를 제공하고, 정부의 고용정책 입안 및 평가에 필요한 기초자료로 이용된다. 경제활동인구에 관한 조사는 내무부 통계국에서 노동력조사라는 명칭으로 1957년부터 실시되어 왔으며, 1963년 통계법에 의해 지정통계로 지정, 경제활동인구조사라는 현재의 명칭으로 개편되고 경제기획원 조사통계국(현 통계청)에서 매 분기마다 면접조사를 통해 조사하고 있다. 1969년부터 인구주택총조사 자료를 이용한 다목적표본이 설계되어 현재까지 이용되고 있으며, 이후 6차례의 표본개편을 실시하였고, 1982년부터는 분기조사에서 월별조사로 전환되었다. 1982년 개편된 ILO 권고안에 따라 조사표를 전면 개편하였으며, 1987년부터는 현재와 같이 만 15세 이상의 인구에 대해 경황조사를 실시하고 있다. 조사 대상은 매월 15일 현재 대한민국에 상주하는 만 15세 이상인 자로 하며, 현역군인 및 공익근무요원, 상비예비역, 전투경찰 및 의무경찰, 형이 확정된 교도소 수감자, 소년원 및 치료감호소 수감자, 경비교도대 등은 제외된다. 조사는 매월 15일이 포함된 1주간(일요일~토요일)을 조사대상 주간으로 하며, 그 다음 주간에 조사가 이루어진다. 조사를 위하여 조사원이 대상가구를 방문하여 면접과 동시에 조사내용

을 직접 기입하는 CAPI(Computer Aided Personal Interview)방식의 면접
타계식 조사를 실시하고 있으며, 조사항목은 가구주와의 관계, 성별 등
인적사항 6개, 활동상태, 취업여부 등 확인항목 5개, 부업여부, 취업시간
등 취업자 항목 6개, 취업가능성, 구직경로 등 실업자 항목 7개, 취업희
망여부, 취업가능성 등의 비경제활동인구 항목 4개, 전직유무 및 이직시
기, 산업, 직업 등 기타항목 7개로 이루어져 있다.

경황조사와 마찬가지로 매월 실시되고 있는 가계조사는 가구의 수입
과 지출을 조사하여, 가구의 생활 실태와 변동 사항을 명확히 파악하고
가구에 대한 가계수지 실태를 파악하여 국민의 소득과 소비 수준 변화
측정, 소비자물가지수 가중값 산정, 국민소득분배 수준 측정 및 소득분
배구조 개선, 국민소득 추계, 이주대책비 산정, 영세민 구호사업, 근로자
임금기준의 결정 등에 필요한 자료를 제공하는 목적으로 이루어진다.
도시가계조사는 1942년 일제 말기에 시작되었으나 그 기록은 남아 있지
않으며, 1951년 한국은행에서 전시 하에 국민소비수준을 측정하기 위해
조사를 실시하였다. 그 후 1954년 서울의 근로자 200가구를 유의표본으
로 추출하여 조사하였고, 1960년부터는 통계적 추출방법에 의해 선정된
가구에서 조사를 실시했다. 1963년 한국은행에서 식료품비에 한하여 가
계부 기장방식으로 조사하던 것을 1975년부터 전품목에 적용했고, 1995
년부터는 현재와 같이 조사항목을 10대 비목으로 확대했다. 2003년부터
는 대상가구를 도시가구에서 전국가구로 확대하였으며, 2006년부터는 1
인가구까지 확대 실시하고 있다. 조사 대상은 농가, 어가, 외국인 가구,
비혈연자취가구, 가계수지 파악이 곤란한 가구 등을 제외한 전국에 거
주하는 모든 가구로 하며, 전국 시군의 동부 및 읍면부의 가구를 대상으
로 조사하고 있다. 가계부에 의한 조사는 매월 1일부터 말일까지 1개월
간을 조사대상기간으로 하며, 가계부 배부는 전월 28일부터 30일까지,
가계부 회수는 익월 1일부터 3일까지 각 3일간에 걸쳐 실시하고 있다.
연간소득조사표에 의한 조사는 조사대상기간을 전년 1월 1일부터 전년
12월 31일까지 1년간으로 하며, 실제조사기간은 정기조사의 경우 1월 3
일부터 10일까지, 조사대상가구로 새롭게 편입되어 가계부를 제출하는
가구의 경우 익월 1일부터 10일까지 실시하고 있다. 가구실태에 관한 사
항은 면접타계식, 가계의 수입과 지출에 관한 사항은 가계부기장에 의
한 자계식, 연간소득에 관한 사항은 면접타계식으로 각각 조사된다. 이

와 같은 가계조사는 국민생활수준 변화의 측정 및 분석 등에 필요한 자료를 제공하고 소비자물가지수 편제에 필요한 가중치 자료를 생산하며, 각종 경제 및 사회정책 입안 기초자료 및 국민소득추계 기초자료로 활용되고 있다.

1. 다목적표본 설계

통계청에서 매일 실시하고 있는 경황조사, 가계조사, 소비자전망조사 및 집세조사 등의 경상조사에서는 조사의 인력과 예산을 최소화하기 위하여, 동일한 표본에서 여러 조사를 실시할 수 있는 다목적표본으로 표본을 설계하여 운영하고 있다. 이 때 조사별 특성의 반영이 미흡하다는 단점을 보완하기 위하여 경황조사구에서 가계조사구 추출시 적격비율이 일정수준 하의 평균 가구원수 및 자동차 비율 등 다양한 모집단 분류지표를 활용하고 있다.

다목적표본 설계에서는 모집단의 분석이 가능한 2000년 인구주택총조사의 10% 표본조사구 중 실제 조사가 곤란한 섬 조사구와 시설단위 조사구 2,691개, 특수사회시설 조사구 1,201개를 제외한 24,998개 조사구를 조사모집단(survey population)으로 이용하게 된다. 이 때 추출된 표본이 모집단의 구조와 가장 흡사한 표본을 추출하기 위하여 특성지표를 적용하게 되며, 특성지표를 선정하기 위해 실업자수와 특성항목 간의 ANOVA 분석을 실시하고, 분석 결과 상관관계가 있는 것으로 판명된 30개 항목을 특성지표로 이용하고 있다. <표 2-1>은 이용 중인 특성지표 항목이다.

또한 사회통계조사 등의 연간 및 특별조사도 다목적표본에서 함께 조사되고 있으나, 2007년 표본개편에서 이들 조사는 경상조사와는 다른 조사구에서 실시하도록 하고 있으며, 다른 조사에 비하여 응답자의 부담이 큰 가계조사의 경우 전자가계부 회수율이 안정적으로 유지(40%)되는 2010년경에 별도 조사구에서 조사를 실시하는 방안을 강구하고 있다.

현재 이용되고 있는 다목적표본에서는 모집단의 대표성을 높이기 위하여 통계 공표단위에 따라 층간의 독립적 추정이 가능하도록 7개 도시와 9개 도로 층화한 후, 각 도를 다시 동부(도시지역)와 읍면부로 나누어 전국을 25개로 층화하고 있다. <표 2-2>는 경황조사에서의 표본조사구 현황이다.

〈표 2-1〉 특성지표 항목

대분류	특성항목
가구 (8개)	농어가 가구수(비농어가) 주택유형별 가구수(아파트, 단독, 다가구 및 연립) 가구구분별 가구수(혈연가구) 주택점유형태별 가구수(자가) 100가구당 승용차 보유대수 100가구당 컴퓨터 보유대수
인구 (9개)	성별 인구(남, 여) 연령별 15세 이상 인구(15~19세, 20~29세, 30~59세, 60세 이상) 교육정도별 15세 이상 인구(중졸 이하, 고졸, 대졸 이상)
경활상태 (13개)	경활상태별 인구(경활, 취업자, 실업자, 비경) 산업별 취업자(농림어업, 광공업, 서비스업) 직업별 취업자(전문기술, 사무, 서비스판매) 종사상 지위별 취업자(근로자, 자영자, 사업주)

〈표 2-2〉 경찰조사 표본조사구 현황

시도	모집단 조사구수	조사구수	동부	읍면부
전국	24,998	1,629	1,242	387
서울	5,506	207	207	-
부산	1,954	117	117	-
대구	1,332	90	90	-
인천	1,282	108	108	-
광주	707	81	81	-
대전	714	81	81	-
울산	525	63	63	-
경기	4,577	198	135	63
강원	894	90	45	45
충북	806	81	45	36
충남	1,043	90	45	45
전북	1,101	81	45	36
전남	1,092	90	45	45
경북	1,565	108	54	54
경남	1,630	99	54	45
제주	270	45	27	18

이와 같이 전국을 25개 층으로 층화한 후 2002년 경찰조사 결과 중 실업자수의 상대표준오차를 이용하여 각 지역별 표본규모를 결정하게 된다. 즉, 25개 지역별로 작성된 1차 추출단위 조사구명부에서 크기의 측도에 비례하는 확률(probability proportional to size sampling)을 가진 계통추출법(systematic sampling)을 적용하여 1,629개의 표본조사구를 추출하게 된다. 추출된 표본조사구는 같은 지역 내에서 동일한 추출률을 갖는 자체가중표본(self-weight sampling)으로 약 33,000가구를 표본으로 추출하게 된다.

가계조사에서는 1998년 표본개편 때에 697개 조사구, 5,087가구가 조사되었던 것이 2003년에 999개 조사구, 7,300가구로 개편되었다. 또한 2006년부터는 1인가구(1,300가구)가 포함되었으며, 영업상 경비와 분리가 불가능한 음식점, 여관 및 하숙집을 경영하는 가구는 제외되었다. 표본의 1단 추출단위는 경찰조사에서와 같은 표본조사구이며, 2단 추출단위는 조사의 효율성을 높이기 위하여 5가구 단위로 조사구역을 구획하여 조사구내 가구를 선정하였다. <표 2-3>은 가계조사 표본조사구 현황이다.

다목적표본 설계로부터 추출된 표본은 표본으로 한번 선정되면 차기 표본개편 시까지 5년 동안 계속해서 조사를 실시하는 고정표본관리방법을 통해 관리되어 왔으나, 2005년부터는 표본의 전면개편 대신 매월 일정규모의 표본가구를 교체하는 연동표본 관리방법을 도입하였다. 그 결과 매월 약 900가구를 교체하고 매 4개월마다 181개 조사구의 조사구역을 교체하였으며, 재개발 및 적격가구 부족 등으로 인하여 2003년 이후 부적격 조사구 171개를 교체하였다. 또한 조사모집단의 규모 증가와 기본구조변화를 표본에 반영시키기 위하여, 조사구역 내에 새로운 가구가 전입하였을 때나 조사지역 내에 주택이 신축되어 가구가 전입하였을 경우에는 전입 이후부터 조사대상가구로 하여 조사를 실시하였고, 전출가구에 대해서는 조사를 중지하였다. 신축 아파트의 경우에도 새로운 조사구로 추가하였으며, 매년 주기적으로 준공 검사가 완료된 아파트 자료를 수집하여 부적격 조사구 중 39개의 새로운 조사구를 신축아파트로 대체하였다.

〈표 2-3〉 가계조사 표본조사구 현황

시도	모집단 조사구수	조사구수	동부	읍면부
전국	24,998	999	765	234
서울	5,506	126	126	-
부산	1,954	72	72	-
대구	1,332	54	54	-
인천	1,282	54	54	-
광주	707	54	54	-
대전	714	54	54	-
울산	525	36	36	-
경기	4,577	108	72	36
강원	894	63	36	27
충북	806	54	27	27
충남	1,043	63	36	27
전북	1,101	54	36	18
전남	1,092	54	36	18
경북	1,565	63	36	27
경남	1,630	63	36	27
제주	270	27	18	9

2. 연동표본제

경찰조사와 가계조사에서는 2005년 1월부터 전국에 걸쳐 본격적인 연동표본조사를 실행하고 있다. 연동표본조사란 시점의 변화에 따라 일정 비율의 부표본(sub-sample)을 표본에서 제외시키고 새로운 부표본으로 표본을 대체하는 표본 추출 방법을 의미하는 것으로, 반복조사시 표본의 일부는 계속 조사하고 나머지 부분은 교체하는 조사 방법이다. 이와 같은 연동표본조사는 현재 여러 나라에서 시행 중이며, 시행하는 나라에 따라 그 형태 또한 다양하다. <표 2-4>는 연동표본조사를 시행하고 있는 대표적인 외국사례이다.

현재 경찰조사에서 이용하고 있는 연동표본조사는 한국형 연동표본모형으로, 표본 교체 후 3년 동안 조사에 참여하다가 3년 후 완전히 표본에서 빠져나가는 형태로 약 900가구를 매월 교체하여 36개월 동안 모두 교체하는 형식(교체시 1/36 교체 시행)을 따르고 있다. 경찰조사에서

시행하고 있는 연동표본조사의 중복표본비율은 전월대비 35/36이며, 전년동월대비 24/36이다.

〈표 2-4〉 외국의 연동표본조사

국가	조사주기	연동표본내용
오스트레일리아	매월	1/8 교체, 8개월 조사 후 완전 교체
캐나다	매월	1/6 교체, 6개월 조사 후 완전 교체
독일	매년(4월)	1/4 교체, 4년 후 완전 교체
일본	매월	4-8-4 시스템, 4개월 조사, 8개월 중지, 4개월 재조사
영국	분기	1/5 교체, 5분기 후 완전 교체
미국	매월	4-8-4 시스템, 4개월 조사, 8개월 중지, 4개월 재조사
이탈리아	분기	2-2-2 시스템, 6분기 후 완전 교체
노르웨이	분기	1/8 교체, 8분기 조사 후 완전 교체
폴란드	분기	2-2-2 시스템, 6분기 후 완전 교체
오스트리아	분기	1/8 교체, 8분기(2년) 조사 후 완전 교체
그리스	분기	1/6 교체, 6분기 조사 후 완전 교체
필리핀	분기	분기별 25% 교체
포르투갈	분기	1/6 교체, 6분기 조사 후 완전 교체
에스토니아	분기	2-2-2 시스템, 6분기 후 완전 교체

주: 이 외에 연동표본조사를 시행하고 있는 국가로는 벨기에, 체코, 홍콩, 아일랜드, 뉴질랜드, 스웨덴, 핀란드, 헝가리 등이 있음.

가계조사의 경우, 경찰조사와 마찬가지로 표본 교체 후 3년 동안 조사에 참여하다가 3년 후 완전히 표본에서 빠져나가는 형태로 조사구내 조사구역 4개가 36개월 동안 모두 교체되는 형태를 따르고 있으며, 경찰조사 교체 3개월째부터 실시하게 된다. 교체주기는 4개월 중 2개월 교체(3~4월, 7~8월, 11~12월 교체)를 따르고 있으며, 교체시 1/18 교체를 시행하고 있다. 가계조사에서 시행하고 있는 연동표본조사의 중복표본비율은 전년동월대비 24/36이며, 전월대비로는 교체가 이루어지는 달의 경우 34/36, 교체가 이루어지지 않는 달의 경우 1이다.

이와 같은 연동표본조사의 도입으로 고정표본조사에 비해 응답자의 부담을 줄이고, 시계열 단절을 방지하여 안정된 추정량을 확보할 수 있게 되었으며, 모집단의 변화 반영이 시간이 경과됨에 따라 감소되는 표본 노후화 현상을 방지할 수 있게 되었다. 연동표본조사의 실시 결과 경활조사의 응답률에는 교체에 따른 차이가 거의 발생하지 않았으며, 가계조사의 경우에는 처음 3개월간 차이가 현저하게 나타났으나 그 이후에는 차이가 점차 줄어드는 것을 알 수 있다.

〈표 2-5〉 경제활동인구조사 응답률(2005년)

응답횟수	1	2	3	4	5	6
응답률	98.2	98.4	98.4	98.2	98.2	98.2

〈표 2-6〉 가계조사 가계부 회수율(2005년)

응답횟수	1	2	3	4	5	6
회수율	69.4	73.6	75.6	76.6	77.7	78.7
회수율*	44.2	54.4	62.5	69.4	73.4	77.9

주: 회수율*은 2002년 표본개편 당시 결과로 응답횟수 1~3은 병행조사 기간임.

3. 가중값(승수)

경활조사에서는 설계가중값(design weight)과 함께 성별, 시도별, 동부·읍면부별, 연령별 추계인구 등의 모집단 보조정보를 활용하여 가중값을 적용한 비추정방법(ratio estimation method)을 사후층화 조정(post-stratified weight)으로 적용하고 있다. 설계가중값은 표본설계로부터 직접 구해지며, 전국에서 표본조사구가 표본으로 추출될 확률의 역수와 각 표본조사구내 가구조사 완료율의 역수를 곱하여 다음과 같이 작성한다.

$$w_{hij}^0 = \frac{S_h}{n_h S_{hi}} \times \frac{M_{hi}}{m_{hi}} = \frac{S_h}{20 \cdot n_h}$$

여기서 $h = 1, 2, \dots, 25$ (지역), $i = 1, 2, \dots, n_h$ (표본조사구), $j = 1, 2, \dots, m_{hi}$ (표본가구), S_h 는 설계 당시 가구수, M_{hi} 는 h 지역 i 번째 표본조사구내 전체 가구수, n_h 는 h 지역내 표본조사구수, m_{hi} 는 h 지역 i 번째 표본조사구내 표본가구수(20가구)를 나타낸다. 이와 같이 작성된 설계가중값은 h 지역에서 모든 가구가 동일한 가중값을 갖게 되어 지역 층 내에서 자체 가중(self weighting)을 갖게 된다.

사후층화 조정은 다음과 같이 적용되며, 추계인구 이용시 군인 및 재소자 등 현실적으로 가구에서 조사가 불가능한 인구는 제외하여 사용한다.

$$\text{가중값} = \frac{\text{15세 이상 모집단 특성별 추계인구수}}{\text{표본에서 조사된 15세 이상 모집단 특성별 인구수}} = \frac{X_{h,sa}}{\hat{X}_{h,sa}}$$

여기서 sa 는 성별 및 연령별 등 모집단 특성을 나타낸다. 따라서 두 가중값을 모두 고려한 최종가중값은 다음과 같다.

$$w_{hijk} = w_{hij}^0 \times \frac{X_{h,sa}}{\hat{X}_{h,sa}}$$

여기서 $k = 1, 2, \dots, l_{hij}$ (표본가구원)이다.

가계조사에서의 가중값은 설계가중값, 무응답 조정 가중값(non-response adjusted weight), 그리고 사후층화 조정으로 구분되며, 표본가구에 대한 설계가중값은 전국에서 표본조사구가 표본으로 추출될 확률의 역수와 표본조사구 내의 가구조사 완료율의 역수를 곱하여 총수의 추정을 위한 확대 가중값을 산출하여 구할 수 있다. 설계가중값은 다음과 같이 나타낸다.

$$w_{hij}^0 = \frac{S_h}{n_h S_{hi}} \times \frac{M_{hi}}{m_{hi}} = \frac{S_h}{10 \cdot n_h}$$

여기서 $h = 1, 2, \dots, 25$ (지역), $i = 1, 2, \dots, n_h$ (표본조사구), $j = 1, 2, \dots, m_{hi}$

(표본가구), S 는 설계 당시 가구수, M_{hi} 는 h 지역 i 번째 표본조사구내 전체 가구수, n_h 는 h 지역내 표본조사구수, m_{hi} 는 h 지역 i 번째 표본조사구내 표본가구수(10가구)를 나타낸다.

가계조사는 조사의 특성상 매일 매일의 수입과 지출에 관한 품목과 금액을 가계부에 기입함으로써 타 조사에 비해 응답자의 부담이 가중되어 평균 무응답률이 15~20% 정도로 매우 높은 수준이다. 또한 가계조사에서는 가계부 기장 방식의 조사 방법을 택하고 있으므로 항목무응답이 아닌 단위무응답이 발생하게 되는데, 2006년 6월까지의 발생된 무응답을 처리하기 위하여 층내 조사 가구를 이용한 hot-deck imputation을 이용하였다. 일반적으로 무응답 처리 방법은 단위무응답인 경우와 항목무응답인 경우가 서로 다른데, 단위무응답의 경우에는 가중값 조정에 의해, 항목무응답의 경우에는 대체 방법에 의해 조정된다. 즉, 가계조사에서는 단위무응답임에도 불구하고 보조변수를 이용, 무응답 대체군을 형성하여 무응답가구를 대체한 것이다. 이에 2006년 7월부터 가중값조정법(weighting adjustment)이 적용되고 있다. 즉, 무응답을 조정하기 위하여 전국이나 지역별로 단순히 조정하는 것은 모집단이 지역별로 무응답가구와 특성항목이 동질적이지 않아 의미가 없기 때문에 이에 맞는 무응답 층을 조정하여 이용하게 된다. 이 때 추정값의 편향을 줄이고 효율을 높이기 위하여 무응답 층 내에서는 응답률이 비슷하고 표본의 크기가 50 이상이 되도록 하였으며, 특성항목이 무응답 층 내에서 동질적으로 구성되도록 하였다. 무응답 조정 가중값은 다음과 같다.

$$w_{hij}^1 = w_{hij}^0 \times \frac{n_g}{n_{g,r}}$$

여기서 n_g 는 g 번째 무응답층내 표본가구이며, $n_{g,r}$ 은 g 번째 무응답층내 응답가구이다.

가계조사에서의 소득 및 지출 항목은 주택유형에 따라 다른 특성을 보이므로 응답가구의 가중값을 주택유형에 맞게 조정하여 이용한다. 이 때 1인가구는 주택 여부에 상관없이 시도별로 설계가중값을 적용하나 2인 이상 가구의 경우에는 시도별, 동부 및 읍면부별, 주택유형별 추출률의 역수를 적용하게 된다. 즉, 단독주택, 아파트 및 기타주택 등 3개로

구분된 주택유형을 사후층화 조정하여 다음의 최종가중값을 얻게 된다.

$$w_{hij} = w_{hij}^1 \times \frac{N_{h,p}}{\hat{N}_{h,p}}$$

여기서 N 은 모집단 추정값, \hat{N} 은 표본에서 조사된 가구의 가중값 합, p 는 주택유형을 나타내는 첨자이다. 이와 같은 방식으로 1인가구 16개, 2인 이상 가구 74개로 총 90개의 층이 형성되며, 각 층에 대한 90개의 최종가중값이 <표 2-7>과 같이 적용된다.

<표 2-7> 가계조사 표본설계 가중값

시도	1인가구	2인 이상 가구					
		동 부			읍 면 부		
		단독 ¹⁾	아파트	기타주택	단독	아파트	기타주택
서울	2,730	2,736 ²⁾	1,971	2,094	-	-	-
부산	1,214	1,606	1,401	1,014	-	-	-
대구	1,542	1,288	1,269	1,902	-	-	-
인천	2,207	2,339	1,270	914	-	-	-
광주	792	834	756	495	-	-	-
대전	756	781	665	2,116	-	-	-
울산	1,037	784	600	1,476	-	-	-
경기	2,394	3,595	2,612	3,107	1,717	1,475	1,299
강원	732	886	739	744	839	715	331
충북	1,079	1,002	939	1,502	683	614	876
충남	1,280	459	505	504	1,483	889	1,655
전북	936	1,207	989	877	1,203	480	1,310
전남	1,269	635	711	649	3,399	601	1,238
경북	1,377	1,088	1,310	767	2,140	1,137	1,418
경남	1,381	1,474	1,209	3,162	1,655	771	1,624
제주	534	929	555	332	419	-	200

주: 1) 단독주택은 다가구주택을 포함.

$$2) \text{서울의 단독주택 가중값} = \frac{\text{서울 모집단 단독주택수}}{\text{서울 표본 단독주택수}}$$

모집단의 특성을 추정할 때 최종가중값은 가구수, 주택수, 가구원수 변화 등과 같은 모집단의 변동을 감안하여야 한다. 만약 어느 층에서 가구수가 크게 증가되어 표본가구수가 적게 되면 추가로 표본을 추출해야 한다. 이 때 모집단이 추출률의 역수만큼 변동이 있는 경우에 변동에 비례하여 표본을 증감해야 한다. 또한 모집단의 변화를 조사과정 중 표본에 반영해 주지 못한다면, 조사가 이루어진 이후 사후적으로 변화한 모집단의 특성을 파악하여 이들을 조사결과에 반영해 주어야 한다. 그렇지 못한 경우 모수에 대한 추정량은 큰 편의를 가지게 된다. 따라서 조사구역 내에 새로운 가구가 전입하였을 때나 신축된 주택에 전입한 가구는 전입 이후부터 조사대상가구로 포함하여 조사하고, 전출가구에 대해서는 조사를 중지하게 된다. 그러나 가계조사는 5년마다 실시하는 인구주택총조사의 10% 표본조사구를 표본틀로 이용하므로 모집단의 변동을 시차 없이 표본틀에 반영하는 것은 사실상 어려운 실정이다. 인구주택총조사 후 작성되는 장래가구추계는 시도별 가구추계가 없어 이용이 불가능하며, 유형별 주택수에 대한 연도별 자료도 따로 조사되거나 추계되고 있지 않아 이용할 수 없는 상황이다. 그러나 2007년 11월부터 시도별 가구추계가 발표될 예정이어서 이후로는 변동가중치를 적용, 보다 편의가 작은 추정값을 얻을 수 있을 것으로 기대된다.

제3절 현행 추정 방법

1. 경찰조사에서의 추정

월단위로 조사되는 경찰조사는 매월 공표되며, 분기별 및 연간 월평균 추정값을 산출하고 있다. 각 추정값은 비추정방법에 의한 가중값을 이용하여 산출하며, 각 시도에서 어떤 특성 Y 를 갖는 성별 및 연령별, 시도별 월평균 추정값은 다음과 같이 계산된다.

$$\begin{aligned}\hat{Y}_h &= \sum_i \sum_j \sum_k w_{hijk} \times y_{hijk} \\ &= \sum_i \sum_j \sum_k \frac{S_h}{20 \cdot n_h} \cdot \frac{X_{h,sa}}{\hat{X}_{h,sa}} \times y_{hijk}\end{aligned}$$

여기서

$h =$ 지역($h=1,2,\dots,25$), $i =$ 표본조사구($i=1,2,\dots,n_h$),

$j =$ 표본가구($j=1,2,\dots,m_{hi}$), $k =$ 표본가구원,

$S_h = h$ 지역내 가구수, $n_h = h$ 지역내 표본조사구수

$m_{hi} = h$ 지역, i 번째 조사구내 표본가구수(20가구)

$X_{h,sa} = h$ 지역내 성별 및 연령별 모집단 추정값

$\hat{X}_{h,sa} = h$ 지역내 성별 및 연령별 표본에서 조사된 응답자 총가중값

$y_{hijk} = h$ 지역, i 번째 표본조사구내, j 번째 가구내, k 번째 관심변수의 관측값

이다.

이를 이용하여 월평균 추정값을 나타내면 다음과 같다.

$$\hat{Y} = \sum_h \hat{Y}_h$$

각 성별 및 연령별에 따른 각 시도별 월평균 추정값의 분산 추정값과 월평균 추정값의 분산 추정값은 각각 다음과 같이 계산된다.

$$Var(\hat{Y}_h) = \frac{(1-f_h)n_h}{2(n_h-1)} \sum_i Z_{hi}^2$$

여기서 $f_h = \frac{n_h}{10N} \simeq 0$, $Z_{hi} = dY_{hi} - R_h \cdot dX_{hi}$, $dY_{hi} = Y_{hi} - Y_{hi+1}$,

$dX_{hi} = X_{hi} - X_{hi+1}$, $R_h = \sum_i \frac{Y_{hi}}{X_{hi}}$, $Y_{hi} = \sum_j \sum_k w_{hijk} \cdot y_{hijk}$,

$X_{hi} = \sum_j \sum_k w_{hijk}$ 이고,

$$Var(\hat{Y}) = \sum_h Var(\hat{Y}_h)$$

이다.

2. 가계조사에서의 추정

월단위로 조사되는 가계조사는 매 분기 및 연간 주기로 공표되며, 분기별 공표를 위하여 3개월간의 월별 자료를 집계하여 월평균 추정값을 산출하여 이용한다. 각 추정값은 각 시도의 추정값을 동부 및 읍·면부에 따른 가중값을 고려하여 먼저 산출한다. 이 때 계산되는 추정값은 설계가중값만을 고려한 가중평균이며, 추정값의 분산은 2006년 7월 이후 잭나이프 방법을 이용하고 있다. 시도별 가중평균 및 월평균 추정값은 다음과 같이 계산된다.

$$\bar{Y}_h = \frac{\sum_i \sum_j w_{hij} \cdot y_{hij}}{\sum_i \sum_j w_{hij}}$$

여기서 $y_{hij} = h$ 지역, i 번째 표본조사구내, j 번째 관심변수의 관측값을 나타내며, $w_{hij} = \frac{S_h}{n_h \cdot 10} \times \frac{n_g}{n_{g,r}} \times \frac{N_{h,p}}{\hat{N}_{h,p}}$ 이고,

$$\bar{Y} = \frac{\sum_h \sum_i \sum_j w_{hij} \cdot y_{hij}}{\sum_h \sum_i \sum_j w_{hij}}$$

이다.

또한 월평균 추정값의 분산은 잭나이프 방법에 의한 분산 추정식을 이용하여 다음과 같이 계산된다.

$$\begin{aligned} Var_J(\bar{Y}) &= \sum_{h=1}^H \left(\frac{w_{h..}}{w_{...}} \right)^2 Var_J(\bar{Y}_h) \\ &= \sum_{h=1}^H \left(\frac{w_{h..}}{w_{...}} \right)^2 \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\bar{Y}_{(hi)} - \bar{Y}_h)^2 \end{aligned}$$

여기서 n_h 는 층 h 에서의 표본조사구이고, $\bar{Y}_{(hi)}$ 는 층(h 시도) 내의 i 번째 조사구를 제외한 관심 변수에 대한 가중평균추정값이다.

잭나이프 가중치의 변화는 다음과 같다.

$$w_{j(hi)} = \begin{cases} w_j, & j \notin h \\ \frac{n_h}{n_h - 1} w_j, & j \in h, j \notin i \\ 0, & j \in (hi) \end{cases}$$

현재 미국이나 캐나다의 가계조사 등에서 이용되고 있는 잭나이프 방법은 계산이 용이하고 추출방법이나 무응답 여부에 관계없이 일반적인 추정 방식의 적용이 가능할 뿐만 아니라 분산의 실제값에 가장 가까운 추정값을 제공하는 것으로 알려져 있으므로, 실제 문제에서의 분산 추정은 잭나이프 방법을 이용하는 것이 바람직하다고 할 수 있다. 분산 추정 방식으로 잭나이프 방법을 적용할 경우, 연동표본조사를 실시하고 있는 경찰조사나 가계조사뿐 아니라 통계청에서 시행되고 있는 대부분의 통계조사에서의 분산 추정을 대체할 수 있다. 따라서 여러 통계조사에 단일 방식을 이용한 분산 추정이 가능해지므로 활용도가 높을 뿐 아니라 시간과 인력 등을 절약하는 데 큰 도움이 될 것으로 예상된다. 잭나이프 방법을 이용한 분산 추정 방식은 2006년 7월 가계조사 추정 방법에 도입된 바 있으며, 다른 통계조사에서도 확대, 활용되어야 할 것이다.

지금까지 살펴본 바와 같이 경찰조사와 가계조사에서는 연동표본계를 활용하여 조사를 시행하고 있으나, 추정 방법은 여전히 고정표본계에서의 추정 방법(비추정량)을 이용하고 있는 실정이다. 이럴 경우 추정값의 편의 문제가 발생할 수 있으므로 표본의 특성에 맞는 추정 방법을 도입할 필요가 있다. 다음 절에서는 연동표본조사에서 계산되는 AK 복합추정량에 대하여 알아보려고 한다.

제4절 AK 복합추정량

1. AK 복합추정량

AK 복합추정량(AK composite estimator)이란 연속조사시 $t-1$ 달까

지의 알려진 정보를 t 달의 모수 추정에 이용하는 방법으로, 일반식은 다음과 같다.

$$y_t^{AK} = \frac{1}{l} \left\{ (1-K+A) \sum_{j=1}^b X_{t,j} + \left(1-K - \frac{b}{l-b} A \right) \sum_{j=b+1}^l X_{t,j} \right\} + K(y_{t-1}^{AK} + \Delta_{t,t-1})$$

여기서 y_t^{AK} : AK 복합추정량

$X_{t,i}$: 표본에 i 번째로 포함된 부표본에 기초한 응답자의 가중값
 $\sum_{j=1}^b X_{t,j}$, $j=1, \dots, b$ 는 t 시점에 표본에 처음 포함되는 부표본의 단순추정량, $X_{t,j}$, $j=b+1, \dots, l$ 은 $t-1$ 시점과 t 시점 동시에 표본에 포함되는 부표본의 단순추정량

$\Delta_{t,t-1}$: $t-1$ 달과 t 달의 월별 변화량의 추정량으로, t 달과 $t-1$ 달에 공통으로 포함된 부분 표본에 기초한 통계량

A, K : 가중값, $0 \leq A \leq 1$, $0 \leq K < 1$ 이다.

위의 식에서 A 와 K 의 값을 제한하는 경우, AK 복합추정량의 특수한 형태인 단순 K 복합추정량(y_t^K), current 복합추정량(y_t^c), 단순추정량(y_t) 등을 생각할 수 있다. 먼저 $A=0$ 인 경우를 단순 K 복합추정량이라고 하고 $y_t^K = (1-K)X_t + K(y_{t-1}^K + \Delta_{t,t-1})$ 로 나타내며, $A=0$, $K=0.5$ 인 경우는 current 복합추정량이라고 하고 $y_t^c = 0.5X_t + 0.5(y_{t-1}^c + \Delta_{t,t-1})$ 로 나타낸다. 또한 $A=0$, $K=0$ 인 경우에는 $y_t = \frac{1}{l} \sum_{j=1}^l X_{t,j}$ 인 단순추정량이 된다. 여기서 단순추정량이란 한 시점에 속한 모든 부표본의 추정값들의 단순평균을 의미하며, 부표본의 추정값은 조사의 목적에 따라 단순평균이나 비추정값이 될 수 있다.

경황조사의 경우, 매 시점 36개의 부표본이 존재하며, 매월 하나씩의 표본교체가 일어나게 된다. 즉, 한번 표본에 포함된 부표본은 36개월 동안 표본에 포함되어 있다가 표본에서 완전히 빠져나가게 된다. 이 때 매달 표본에 처음 포함되는 부표본은 기존에 포함되어 있는 35개의 표본들과는 다른 방향의 편향을 갖게 된다(Bailer, 1975). 따라서 AK 복합추정량은 처음 표본에 포함되는 표본에 다른 표본들보다 큰 반대부호의

가중값을 주게 된다. 이와 같은 특성을 일반식에 적용시키면, 경찰조사에서의 AK 복합추정량을 다음과 같이 나타낼 수 있다.

$$\begin{aligned} y_t^{AK} &= \frac{1}{36} \left\{ (1-K+A)X_{t,1} + \left(1-K-\frac{1}{35}A\right) \sum_{j=2}^{36} X_{t,j} \right\} \\ &\quad + K(y_{t-1}^{AK} + \Delta_{t,t-1}) \\ &= \frac{1}{36} \left\{ (1-K+A)X_{t,1} + \left(1+\frac{1}{35}K-\frac{1}{35}A\right) \sum_{j=2}^{36} X_{t,j} \right\} \\ &\quad - \frac{1}{35}K \sum_{j=2}^{36} X_{t-1,j-1} + Ky_{t-1}^{AK} \end{aligned}$$

여기서 $\Delta_{t,t-1} = \frac{1}{35} \sum_{j=2}^{36} (X_{t,j} - X_{t-1,j-1})$ 이다.

가계조사의 경우에는 경찰조사와 마찬가지로 매 시점에서 36개의 부표본이 존재하나, 표본교체가 일어나는 달(3, 4, 7, 8, 11, 12월)에만 2개의 새로운 부표본이 표본에 포함되게 된다. 이 때 표본에 처음으로 포함되는 첫 번째와 두 번째 부표본이 나머지 34개의 중복된 부표본들과는 다른 방향의 편향을 갖게 된다. 이와 같은 특성을 지닌 가계조사에서의 AK 복합추정량은 표본교체가 일어나는 달과 일어나지 않는 달의 경우로 나누어 각각을 일반식에 적용하여 다음과 같이 나타낼 수 있다.

☒ 표본교체가 일어나는 경우(3, 4, 7, 8, 11, 12월인 경우)

$$\begin{aligned} y_t^{AK} &= \frac{1}{36} \left\{ (1-K+A) \sum_{j=1}^2 X_{t,j} + \left(1-K-\frac{2}{34}A\right) \sum_{j=3}^{36} X_{t,j} \right\} \\ &\quad + K(y_{t-1}^{AK} + \Delta_{t,t-1}) \\ &= \frac{1}{36} \left\{ (1-K+A) \sum_{j=1}^2 X_{t,j} + \left(1+\frac{2}{34}K-\frac{2}{34}A\right) \sum_{j=3}^{36} X_{t,j} \right\} \\ &\quad - \frac{1}{34}K \sum_{j=3}^{36} X_{t-1,j-2} + Ky_{t-1}^{AK} \end{aligned}$$

여기서 $\Delta_{t,t-1} = \frac{1}{34} \sum_{j=3}^{36} (X_{t,j} - X_{t-1,j-2})$ 이다.

☑ 표본교체가 일어나지 않는 경우(1, 2, 5, 6, 9, 10월인 경우)

$$\begin{aligned} y_t^{AK} &= \frac{1}{36} \left\{ (1-K) \sum_{j=1}^{36} X_{t,j} \right\} + K(y_{t-1}^{AK} + \Delta_{t,t-1}) \\ &= \frac{1}{36} \sum_{j=1}^{36} X_{t,j} - \frac{1}{36} K \sum_{j=1}^{36} X_{t-1,j} + K y_{t-1}^{AK} \end{aligned}$$

여기서 $\Delta_{t,t-1} = \frac{1}{36} \sum_{j=1}^{36} (X_{t,j} - X_{t-1,j})$ 이다.

이와 같은 AK 복합추정량은 t 시점 이전의 정보를 t 시점의 모수를 추정하는 데 이용하므로 단순추정량을 사용할 때에 비하여 작은 분산을 가질 것으로 예상된다. 그러나 각 교체 시점에서 표본에 포함되어 있는 부표본에 따라 각기 다른 기대값을 가지게 되므로 연동그룹편의(rotation group bias)가 발생(Bailer, 1975)하는 등 복합추정량 자체에서 발생하는 여러 편익에 의해 그 효율성이 상쇄될 수 있다. 즉, 연동표본에서의 각 부표본 중 표본에 처음으로 들어오는 부표본의 편익이 다른 부표본에 비하여 크기 때문에(Hanson, 1978) AK 복합추정량은 처음 표본으로 들어오는 부표본과 나머지 부표본들에 서로 다른 가중값을 주어 추정의 정도를 높이고자 하였다.

2. AK 복합추정량의 분산

y_t^{AK} 의 분산은 여러 개의 부표본 추정량들의 분산과 공분산을 포함하게 되므로(Kumar & Lee, 1983), y_t^{AK} 의 분산을 구하기 위해서는 다음과 같은 분산-공분산 구조에 대한 가정이 필요하다.

가정 1) $Var(y_{t,i}) = \sigma^2$ for all t, i

여기서 t 는 해당월, i 는 연동그룹을 나타낸다.

가정 2) $Cov(y_{t,i}, y_{s,j}) = \rho_{|t-s|} \sigma^2$, $\rho_0 = 1$

가정 3) $Cov(y_{t,i}, y_{t,j}) = 0$, $i \neq j$

위 가정은 중복된 연동그룹에서의 공분산은 정상성을 만족하며, 주어진 시점에서 다른 연동그룹간의 상관관계는 없다는 것을 의미한다.

일정 시점의 수준을 추정하기 위한 AK 복합추정량은 다음과 같다.

$$y_t^{AK} = y_t + K(y_{t-1}^{AK} + \Delta_{t,t-1})$$

여기서 $y_t = \frac{1}{l} \left\{ (1-K+A) \sum_{j=1}^b X_{t,j} + \left(1-K - \frac{b}{l-b} A\right) \sum_{j=b+1}^l X_{t,j} \right\}$ 이므로, 위 가정이 성립할 경우 AK 복합추정량에서의 수준의 분산은 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \text{Var}(y_t^{AK}) = & [\text{Var}(y_t) + K^2 \text{Var}(\Delta_{t,t-1}) + 2K \text{Cov}(y_t, \Delta_{t,t-1}) \\ & + 2 \sum_{g=1}^{36} K^g \{ \text{Cov}(y_t, y_{t-g}) \\ & + K \text{Cov}(y_t, \Delta_{t-g,t-g-1}) \\ & + K \text{Cov}(\Delta_{t,t-1}, y_{t-g}) \\ & + K^2 \text{Cov}(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) \}] / (1-K^2) \end{aligned}$$

여기서 경찰조사의 경우 분산의 세부 항목 값은 다음과 같다.

$$\text{Var}(y_t) = \frac{1}{1260} [35(1-K)^2 + A^2] \sigma^2$$

$$\text{Var}(\Delta_{t,t-1}) = \frac{2}{35} (1-\rho_1) \sigma^2$$

$$\text{Cov}(y_t, \Delta_{t,t-1}) = \frac{1}{1260} [35(1-K) - A] (1-\rho_1) \sigma^2$$

$$\begin{aligned} \text{Cov}(y_t, y_{t-g}) = & \frac{1}{1587600} [1225(1-K)^2(36-g) \\ & + 70A(1-K)(g-18) - gA^2] \rho_g \sigma^2 \\ & g = 1, 2, \dots, 35 \end{aligned}$$

$$\begin{aligned} \text{Cov}(y_t, \Delta_{t-g,t-g-1}) = & \frac{1}{44100} [35(35-g)(1-K) \\ & - (35-g)A] (\rho_g - \rho_{g+1}) \sigma^2 \\ & g = 1, 2, \dots, 34 \end{aligned}$$

$$\begin{aligned} \text{Cov}(\Delta_{t,t-1}, y_{t-g}) = & \frac{1}{44100} [35(1-K)(36-g) + gA] \\ & (\rho_g - \rho_{g-1}) \sigma^2 \\ & g = 1, 2, \dots, 35 \end{aligned}$$

$$Cov(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) = \frac{1}{1225} (35-g) (2\rho_g - \rho_{g-1} - \rho_{g+1}) \sigma^2$$

$$g = 1, 2, \dots, 34$$

가계조사의 경우 매달 하나의 부표본이 표본에 포함되는 경황조사와는 다르게 표본교체가 일어나는 달에는 두 개의 부표본이 포함되며, 교체가 일어나지 않는 달에는 전 달과 같은 표본이 유지되므로 교체 여부에 따른 분산의 차이가 발생하게 된다. 따라서 각각의 경우로 나누어 분산내 항목들을 계산하여야 한다. 조건에 따른 각 항목은 다음과 같다.

☒ $t = 3, 4, 7, 8, 11, 12$ (교체가 있는 달)

$$Var(y_t) = \frac{1}{1224} [34(1-K)^2 + 2A^2] \sigma^2$$

$$Var(\Delta_{t,t-1}) = \frac{2}{34} (1-\rho_1) \sigma^2$$

$$Cov(y_t, \Delta_{t,t-1}) = \frac{1}{1224} [34(1-K) - 2A] (1-\rho_1) \sigma^2$$

☒ $t = 1, 2, 5, 6, 9, 10$ (교체가 없는 달)

$$Var(y_t) = \frac{1}{36} (1-K)^2 \sigma^2$$

$$Var(\Delta_{t,t-1}) = \frac{2}{36} (1-\rho_1) \sigma^2$$

$$Cov(y_t, \Delta_{t,t-1}) = \frac{1}{36} (1-K)(1-\rho_1) \sigma^2$$

☒ $t = 3, 7, 11$ (교체 첫 번째 달)

$$Cov(y_t, y_{t-g}) = \begin{cases} \frac{1}{36^2} \left[\begin{array}{l} (1-K)^2 (36-2a) \\ - \frac{(72-8a)}{34} A(1-K) - \frac{8a}{34^2} \end{array} \right] \rho_g \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 3, 4, 7, 8, \dots, 32, 35 \\ \frac{1}{36^2} \left[\begin{array}{l} (1-K)^2 (38-4a) \\ - \frac{(76-8a)}{34} A(1-K) \end{array} \right] \rho_g \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \quad g = 1, 2, 5, 6, \dots, 33, 34 \end{cases}$$

$$Cov(\Delta_{t,t-1}, y_{t-g}) = \begin{cases} \frac{(38-4a)}{1224} (1-K)(\rho_g - \rho_{g-1})\sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \quad g = 1, 2, 5, 6, \dots, 33, 34 \\ \frac{1}{1224} \left[(36-2a)(1-K) + \frac{4a}{34}A \right] (\rho_g - \rho_{g-1})\sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 3, 4, 7, 8, \dots, 32, 35 \end{cases}$$

$$Cov(y_t, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{(38-4a)}{36^2} \left[1 - K - \frac{2}{34}A \right] (\rho_g - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \quad g = 1, 2, 5, 6, \dots, 33, 34 \\ \frac{(34-2a)}{1224} \left[1 - K - \frac{2}{34}A \right] (\rho_g - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 3, 4, 7, 8, \dots, 32, 35 \end{cases}$$

$$Cov(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{38-4a}{1224} (2\rho_g - \rho_{g-1} - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 1, 2, 5, 6, \dots, 33, 34 \\ \frac{32-2a}{34^2} (2\rho_g - \rho_{g-1} - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 3, 4, 7, 8, \dots, 32, 35 \end{cases}$$

☒ $t = 4, 8, 12$ (교체 두 번째 달)

$$Cov(y_t, y_{t-g}) = \begin{cases} \frac{1}{36^2} \left[(1-K)^2(36-2a) - \frac{(72-8a)}{34} \right. \\ \quad \left. A(1-K) - \frac{8a}{34^2} \right] \rho_g \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 1, 4, 5, 8, 9, \dots, 32, 33 \\ \frac{1}{36^2} \left[(1-K)^2(38-4a) - \frac{(76-8a)}{34} A(1-K) \right] \rho_g \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \quad g = 2, 3, 6, 7, \dots, 34, 35 \end{cases}$$

$$Cov(\Delta_{t,t-1}, y_{t-g}) = \begin{cases} \frac{1}{1224} \left[(36-2a)(1-K) + \frac{4a}{34}A \right] (\rho_g - \rho_{g-1}) \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 1, 4, 5, 8, 9, \dots, 32, 33 \\ \frac{(36-4a)}{1224} (1-K) (\rho_g - \rho_{g-1}) \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \quad g = 2, 3, 6, 7, \dots, 34, 35 \end{cases}$$

$$Cov(y_t, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{(34-2a)}{1224} \left[1 - K - \frac{2}{34}A \right] (\rho_g - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 1, 4, 5, 8, 9, \dots, 32, 33 \\ \frac{(36-4a)}{36^2} \left[1 - K - \frac{2}{34}A \right] (\rho_g - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 2, 3, 6, 7, \dots, 34, 35 \end{cases}$$

$$Cov(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{34-2a}{34^2} (2\rho_g - \rho_{g-1} - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ \quad g = 1, 4, 5, 8, 9, \dots, 32, 33 \\ \frac{36-4a}{1224} (2\rho_g - \rho_{g-1} - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 2, 3, 6, 7, \dots, 34, 35 \end{cases}$$

☐ $t = 1, 5, 9$ (교체 없는 첫 번째 달)

$$Cov(y_t, y_{t-g}) = \begin{cases} \frac{1}{36^2} [(1-K)^2 (36-2a) + \frac{4a}{34}A(1-K)] \rho_g \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 2, 5, 6, 9, 10, \dots, 33, 34 \\ \frac{1}{36^2} [(1-K)^2 (40-4a)] \rho_g \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \quad g = 1, 3, 4, 7, 8, \dots, 32, 35 \end{cases}$$

$$Cov(\Delta_{t,t-1}, y_{t-g}) = \begin{cases} \frac{(40-4a)}{36^2} (1-K)(\rho_g - \rho_{g-1})\sigma^2 \\ \text{if } a = 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 1, 3, 4, 7, 8, \dots, 32, 35 \\ \frac{1}{36^2} \left[(36-2a)(1-K) + \frac{4a}{34}A \right] (\rho_g - \rho_{g-1})\sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \quad g = 2, 5, 6, 9, 10, \dots, 33, 34 \end{cases}$$

$$Cov(y_t, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{(36-2a)}{1224} (1-K)(\rho_g - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ \quad g = 2, 5, 6, 9, 10, \dots, 33, 34 \\ \frac{(36-4a)}{36^2} (1-K)(\rho_g - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 1, 3, 4, 7, 8, \dots, 32, 35 \end{cases}$$

$$Cov(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{36-2a}{1224} (2\rho_g - \rho_{g-1} - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ \quad g = 1, 2, 5, 6, 9, 10, \dots, 33, 34 \\ \frac{36-4a}{36^2} (2\rho_g - \rho_{g-1} - \rho_{g+1})\sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 3, 4, 7, 8, \dots, 32, 35 \end{cases}$$

☒ $t = 2, 6, 10$ (교체 없는 두 번째 달)

$$Cov(y_t, y_{t-g}) = \begin{cases} \frac{1}{36^2} \left[(1-K)^2 (36-2a) + \frac{4a}{34}A(1-K) \right] \rho_g \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ \quad g = 3, 6, 7, 10, 11, \dots, 34, 35 \\ \frac{1}{36^2} [(1-K)^2 (40-4a)] \rho_g \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ \quad g = 1, 2, 4, 5, 8, 9, \dots, 32, 33 \end{cases}$$

$$Cov(\Delta_{t,t-1}, y_{t-g}) = \begin{cases} \frac{(40-4a)}{36^2} (1-K)(\rho_g - \rho_{g-1}) \sigma^2 \\ \text{if } a = 1, 1, 2, 2, 3, 3, \dots, 9, 9, \\ g = 1, 2, 4, 5, 8, 9, \dots, 32, 33 \\ \frac{1}{36^2} \left[(36-2a)(1-K) + \frac{4a}{34}A \right] (\rho_g - \rho_{g-1}) \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ g = 3, 6, 7, 10, 11, \dots, 34, 35 \end{cases}$$

$$Cov(y_t, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{(36-2a)}{1224} (1-K)(\rho_g - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ g = 2, 3, 6, 7, \dots, 34, 35 \\ \frac{(40-4a)}{36^2} (1-K)(\rho_g - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 2, 2, 3, 3, \dots, 9, 9, \\ g = 1, 4, 5, 8, 9, \dots, 32, 33 \end{cases}$$

$$Cov(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) = \begin{cases} \frac{40-4a}{36^2} (2\rho_g - \rho_{g-1} - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 2, 2, 3, 3, \dots, 17, \\ g = 1, 4, 5, 8, 9, \dots, 32, 33 \\ \frac{36-2a}{1224} (2\rho_g - \rho_{g-1} - \rho_{g+1}) \sigma^2 \\ \text{if } a = 1, 2, 3, \dots, 17, \\ g = 2, 3, 6, 7, \dots, 34, 35 \end{cases}$$

분산과 관련된 자세한 유도 과정은 <부록>에 수록하였다.

제5절 사례연구

본 절에서는 2004년 12월부터 2006년 12월까지의 경찰조사자료를 이용하여 비추정방법에 의한 추정량과 AK 복합추정방법에 의한 추정량의 결과를 비교해 보고자 한다.

먼저 AK 복합추정량 분산식을 적용하여 분산을 구하고, 이를 비추정량을 사용하였을 때와 비교하여 상대효율(Relative Efficiency)을 비교하였다. 상대효율은 다음과 같다.

$$\text{상대효율} = \left\{ \frac{\text{Var}(y_{ratio})}{\text{Var}(y_t^{AK})} \right\} \times 100$$

AK 복합추정량의 분산은 과거 자료와의 상관관계를 이용하여 계산되므로 상대효율은 2006년 12월 자료를 기준으로 구하였으며, 비추정방법과 AK 복합추정방법을 이용하여 각각의 경우에 따른 추정값과 구성비도 구하였다.

AK 복합추정량의 경우 해당 자료에 대하여 최소의 AK 복합추정량 분산을 갖는 가중값 A 와 K 를 구한 후 계산하였다. 각 가중값은 A 를 고정시킨 상태에서 K 를 0.01에서 0.09까지 증가시켜 분산을 최소화시키는 값을 선택하였으며, 선택된 K 값에 대하여 A 를 0.01에서 0.09까지 증가시켜 분산을 최소화하는 값을 선택하였다. 이와 같은 방법을 통해 AK 복합추정량의 최적 가중값은 $A = 0.5$, $K = 0.9$ 로 나타났다. 이하 AK 복합추정량 I이라 한다. 또한 비추정방법을 이용하여 얻은 15세 이상 인구수와 근접한 추정값을 얻기 위하여, A 를 고정시킨 상태에서 K 를 0.01에서 0.09까지 증가시켜 비추정값과 가장 근접한 추정값을 갖는 값을 선택하고, 선택된 K 값에 대하여 A 를 0.01에서 0.09까지 증가시켜 비추정값과 가장 근접한 추정값을 갖는 값을 선택하였다. 이와 같이 새로운 가중값을 고려한 AK 복합추정량에서 선택된 최적 가중값은 $A = 0.17$, $K = 0.21$ 이다. 이하 AK 복합추정량 II라 한다.

다음은 위의 두 가지 AK 복합추정량에서의 비추정량에 대한 상대효율이다.

〈표 2-8〉 AK 복합추정량의 상대효율

종류	A	K	상대효율
AK 복합추정량 I	0.5	0.9	294.587
AK 복합추정량 II	0.17	0.21	209.985

<표 2-8>에서 알 수 있듯이 AK 복합추정량 I 과 AK 복합추정량 II 모두가 비추정량을 사용하는 것보다 효율이 높다는 것을 알 수 있으며, AK 복합추정량 I의 경우가 AK 복합추정량 II의 경우보다 높은 상대 효율을 가지고 있음을 알 수 있다. 즉, 표본들 간의 상관계수가 매우 큰 경우(0.9 이상), 분산을 최소화하는 가중값을 선택하여 계산한 AK 복합추정량 I 이 추정을 하는 데 있어 가장 효율적인 방법임을 보여준다.

비추정방법에 의한 추정량은 기존에 사용하던 추정방법을 그대로 이용하여 구한 것으로, 결과는 <표 2-9>와 같다. AK 복합추정량과의 비교를 위하여 2005년 1월부터 2006년 12월까지의 추정값을 구하였으며, 추정 결과 구성비를 함께 나타내었다. <표 2-10>은 AK 복합추정량 I 을 이용하여 구한 AK 복합추정값과 구성비이다. 비추정량과 AK 복합추정량 I 을 이용하여 구한 추정값과 구성비를 살펴보면, 추정값의 경우 AK 복합추정량에 의한 15세 이상 인구수의 추정값이 연동이 지속됨에 따라 비추정값에 비해 점차 작게 추정됨을 알 수 있다. 구성비의 경우에는 연동이 시작된 후 1년 정도는 비추정값의 구성비와 AK 복합추정값의 구성비에 큰 차이가 발생하지 않으나, 연동이 지속됨에 따라 AK 복합추정방법에서의 15세 이상 인구수 대비 경활인구수가 비추정방법에 비해 비교적 큰 값을 추정하고, 15세 이상 인구수 대비 비경활인구수의 경우에는 비추정방법에 비해 비교적 작은 값을 추정하고 있음을 알 수 있다. 즉, 연동이 지속될수록 비추정방법에 의한 15세 이상 인구수 대비 비경활인구수가 AK 복합추정방법을 이용할 때보다 점차 큰 비중을 차지하게 됨을 알 수 있다.

〈표 2-9〉 비추정방법을 이용한 추정값 및 구성비

(단위: 천 명, %)

년/월		2005.1월	2월	3월	4월	5월	6월
추정값	경활	23,058	23,074	23,531	23,839	24,051	24,123
	취업자	22,078	22,085	22,576	22,934	23,199	23,246
	실업자	980	989	956	904	852	878
	비경활	14,865	14,937	14,565	14,346	14,223	14,245
	15세 이상 인구수 ¹⁾	37,923	38,011	38,096	38,185	38,274	38,368
구성비	경활	60.8	60.7	61.8	62.4	62.8	62.9
	취업자	95.7	95.7	95.9	96.2	96.5	96.4
	실업자	4.3	4.3	4.1	3.8	3.5	3.6
	비경활	39.2	39.3	38.2	37.6	37.2	37.1
	년/월	7월	8월	9월	10월	11월	12월
추정값	경활	24,072	23,689	23,918	24,056	23,976	23,526
	취업자	23,184	22,847	23,048	23,186	23,191	22,699
	실업자	888	843	870	870	785	827
	비경활	14,331	14,739	14,530	14,412	14,511	14,977
	15세 이상 인구수 ¹⁾	38,403	38,428	38,448	38,468	38,487	38,503
구성비	경활	62.7	61.6	62.2	62.5	62.3	61.1
	취업자	96.3	96.4	96.4	96.4	96.7	96.5
	실업자	3.7	3.6	3.6	3.6	3.3	3.5
	비경활	37.3	38.4	37.8	37.5	37.7	38.9
	년/월	2006.1월	2월	3월	4월	5월	6월
추정값	경활	23,340	23,365	23,769	24,088	24,267	24,320
	취업자	22,471	22,412	22,848	23,242	23,484	23,501
	실업자	869	953	921	846	783	819
	비경활	15,206	15,227	14,871	14,597	14,463	14,455
	15세 이상 인구수 ¹⁾	38,546	38,592	38,640	38,685	38,730	38,775
구성비	경활	60.6	60.5	61.5	62.3	62.7	62.7
	취업자	96.3	95.9	96.1	96.5	96.8	96.6
	실업자	3.7	4.1	3.9	3.5	3.2	3.4
	비경활	39.4	39.5	38.5	37.7	37.3	37.3
	년/월	7월	8월	9월	10월	11월	12월
추정값	경활	24,270	23,972	24,096	24,253	24,225	23,773
	취업자	23,447	23,164	23,330	23,463	23,458	22,989
	실업자	823	807	766	789	767	784
	비경활	14,544	14,864	14,758	14,621	14,667	15,132
	15세 이상 인구수 ¹⁾	38,814	38,836	38,854	38,874	38,892	38,905
구성비	경활	62.5	61.7	62.0	62.4	62.3	61.1
	취업자	96.6	96.6	96.8	96.7	96.8	96.7
	실업자	3.4	3.4	3.2	3.3	3.2	3.3
	비경활	37.5	38.3	38.0	37.6	37.7	38.9

주. 1) 경제활동인구수 + 비경제활동인구수

〈표 2-10〉 AK 복합추정방법을 이용한 추정값 및 구성비(A=0.5, K=0.9)

(단위: 천 명, %)

년/월		2005.1월	2월	3월	4월	5월	6월
추정값	경찰	23,061	23,106	23,588	23,920	24,128	24,225
	취업자	22,074	22,118	22,635	23,027	23,290	23,361
	실업자	987	988	953	893	838	865
	비경찰	14,855	14,916	14,553	14,322	14,197	14,220
	15세 이상 인구수	37,916	38,022	38,141	38,242	38,325	38,445
구성비	경찰	60.8	60.8	61.8	62.5	63.0	63.0
	취업자	95.7	95.7	96.0	96.3	96.5	96.4
	실업자	4.3	4.3	4.0	3.7	3.5	3.6
	비경찰	39.2	39.2	38.2	37.5	37.0	37.0
년/월		7월	8월	9월	10월	11월	12월
추정값	경찰	24,175	23,813	24,012	24,284	24,272	23,814
	취업자	23,300	22,978	23,153	23,429	23,512	23,016
	실업자	875	835	859	855	760	798
	비경찰	14,327	14,798	14,601	14,361	14,339	14,755
	15세 이상 인구수	38,502	38,611	38,613	38,645	38,611	38,569
구성비	경찰	62.8	61.7	62.2	62.8	62.9	61.7
	취업자	96.4	96.5	96.4	96.5	96.9	96.6
	실업자	3.6	3.5	3.6	3.5	3.1	3.4
	비경찰	37.2	38.3	37.8	37.2	37.1	38.3
년/월		2006.1월	2월	3월	4월	5월	6월
추정값	경찰	23,728	23,793	24,237	24,951	24,910	25,009
	취업자	22,887	22,870	23,342	23,827	24,149	24,213
	실업자	841	923	895	824	761	796
	비경찰	14,903	14,875	14,492	14,173	14,018	13,977
	15세 이상 인구수	38,631	38,668	38,729	39,124	38,928	38,986
구성비	경찰	61.4	61.5	62.6	63.8	64.0	64.1
	취업자	96.5	96.1	96.3	95.5	96.9	96.8
	실업자	3.5	3.9	3.7	3.3	3.1	3.2
	비경찰	38.6	38.5	37.4	36.2	36.0	35.9
년/월		7월	8월	9월	10월	11월	12월
추정값	경찰	25,120	24,862	25,044	25,285	25,440	25,183
	취업자	24,323	24,080	24,304	24,526	24,694	24,427
	실업자	798	782	741	759	746	756
	비경찰	14,043	14,318	14,256	14,120	14,135	14,609
	15세 이상 인구수	39,163	39,180	39,300	39,405	39,575	39,792
구성비	경찰	64.1	63.5	63.7	64.2	64.3	63.3
	취업자	96.8	96.9	97.0	97.0	97.1	97.0
	실업자	3.2	3.1	3.0	3.0	2.9	3.0
	비경찰	35.9	36.5	36.3	35.8	35.7	36.7

〈표 2-11〉 AK 복합추정방법을 이용한 추정값 및 구성비($A = 0.17, K = 0.21$)
(단위: 천 명, %)

년/월		2005.1월	2월	3월	4월	5월	6월
추정값	경찰	23,061	23,090	23,548	23,856	24,060	24,143
	취업자	22,078	22,103	22,594	22,956	23,210	23,265
	실업자	982	987	954	900	850	878
	비경찰	14,862	14,936	14,565	14,340	14,219	14,245
	15세 이상 인구수	37,923	38,026	38,113	38,196	38,279	38,388
구성비	경찰	60.8	60.7	61.8	62.5	62.9	62.9
	취업자	95.7	95.7	95.9	96.2	96.5	96.4
	실업자	4.3	4.3	4.1	3.8	3.5	3.6
	비경찰	39.2	39.3	38.2	37.5	37.1	37.1
년/월		7월	8월	9월	10월	11월	12월
추정값	경찰	24,083	23,700	23,922	24,093	24,009	23,538
	취업자	23,196	22,857	23,053	23,224	23,228	22,714
	실업자	887	843	869	868	781	824
	비경찰	14,335	14,755	14,531	14,384	14,475	14,954
	15세 이상 인구수	38,418	38,455	38,453	38,477	38,484	38,492
구성비	경찰	62.7	61.6	62.2	62.6	62.4	61.2
	취업자	96.3	96.4	96.4	96.4	96.7	96.5
	실업자	3.7	3.6	3.6	3.6	3.3	3.5
	비경찰	37.3	38.4	37.8	37.4	37.6	38.8
년/월		2006.1월	2월	3월	4월	5월	6월
추정값	경찰	23,376	23,397	23,803	24,131	24,310	24,361
	취업자	22,507	22,445	22,882	23,286	23,529	23,544
	실업자	868	952	921	845	781	818
	비경찰	15,171	15,195	14,844	14,568	14,441	14,431
	15세 이상 인구수	38,547	38,592	38,647	38,699	38,751	38,792
구성비	경찰	60.6	60.6	61.6	62.4	62.7	62.8
	취업자	96.3	95.9	96.1	96.5	96.8	96.6
	실업자	3.7	4.1	3.9	3.5	3.2	3.4
	비경찰	39.4	39.4	38.4	37.6	37.3	37.2
년/월		7월	8월	9월	10월	11월	12월
추정값	경찰	24,338	24,021	24,145	24,312	24,311	23,871
	취업자	23,517	23,215	23,379	23,525	23,542	23,088
	실업자	821	806	766	788	769	783
	비경찰	14,514	14,832	14,744	14,601	14,639	15,117
	15세 이상 인구수	38,852	38,853	38,889	38,913	38,950	38,988
구성비	경찰	62.6	61.8	62.1	62.5	62.4	61.2
	취업자	96.6	96.6	96.8	96.8	96.8	96.7
	실업자	3.4	3.4	3.2	3.2	3.2	3.3
	비경찰	37.4	38.2	37.9	37.5	37.6	38.8

<표 2-12> 비추정방법과 AK 복합추정방법을 이용한 추정값의 차이

(단위: 천 명)

비추정 vs AK I	년/월	2005.1	2	3	4	5	6
	15세 이상 인구수 차이	7	-11	-45	-57	-51	-77
비추정 vs AK II	년/월	2006.1	2	3	4	5	6
	15세 이상 인구수 차이	-85	-76	-89	-139	-198	-211
비추정 vs AK I	년/월	2005.1	2	3	4	5	6
	15세 이상 인구수 차이	0	-15	-17	-11	-5	-20
비추정 vs AK II	년/월	2006.1	2	3	4	5	6
	15세 이상 인구수 차이	-1	0	-7	-14	-21	-17
비추정 vs AK I	년/월	7	8	9	10	11	12
	15세 이상 인구수 차이	-99	-183	-165	-177	-124	-66
비추정 vs AK II	년/월	7	8	9	10	11	12
	15세 이상 인구수 차이	-349	-344	-446	-531	-683	-887
비추정 vs AK I	년/월	7	8	9	10	11	12
	15세 이상 인구수 차이	-15	-27	-5	-9	3	11
비추정 vs AK II	년/월	7	8	9	10	11	12
	15세 이상 인구수 차이	-38	-17	-35	-39	-58	-83

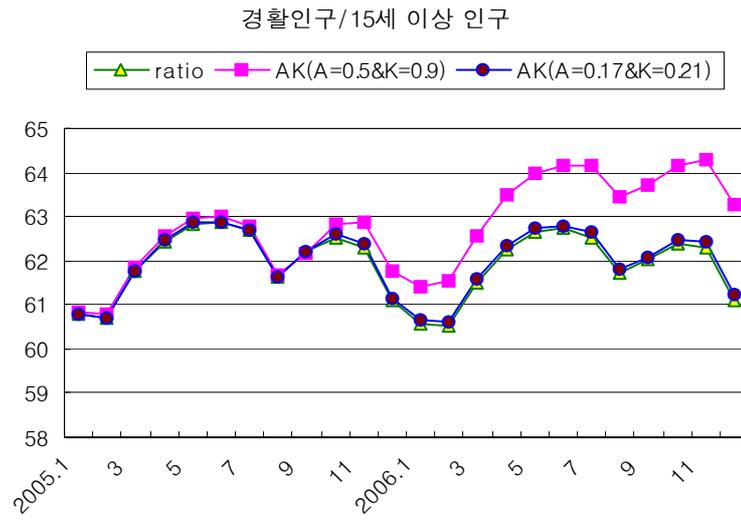
<표 2-11>은 AK 복합추정량 II를 이용하여 구한 추정값과 구성비이다. 여기서의 결과를 <표 2-9>의 비추정값 및 구성비와 비교하면, 추정값의 경우 AK 복합추정량 I을 이용하는 경우와 마찬가지로 연동이 지속됨에 따라 15세 이상 인구수에 대한 추정값에 점차 큰 차이가 발생함을 알 수 있다. 그러나 AK 복합추정량 I에 비하여 비추정량에서의 15

세 이상 인구수와의 차이가 비교적 작게 나타나며, 구성비에 있어서는 두 추정량이 매우 유사한 형태를 나타냄을 알 수 있다. 이와 같은 현상은 해당월의 비추정값에 큰 가중을 주고 전월확정값 및 전월과 금월 간의 변동분 추정값에 가중을 작게 하며, 해당월에 새로 진입하는 진입표본의 영향력에도 작은 가중을 줌으로써 나타나는 것으로 볼 수 있다. 즉, AK 복합추정량에 어떤 가중값을 주느냐에 따라 계산되는 추정값에는 큰 차이가 발생하게 된다는 것을 알 수 있다. <표 2-12>는 세 가지 방법에 의해 얻은 15세 이상 인구수 추정값의 차이를 나타낸 표이다.

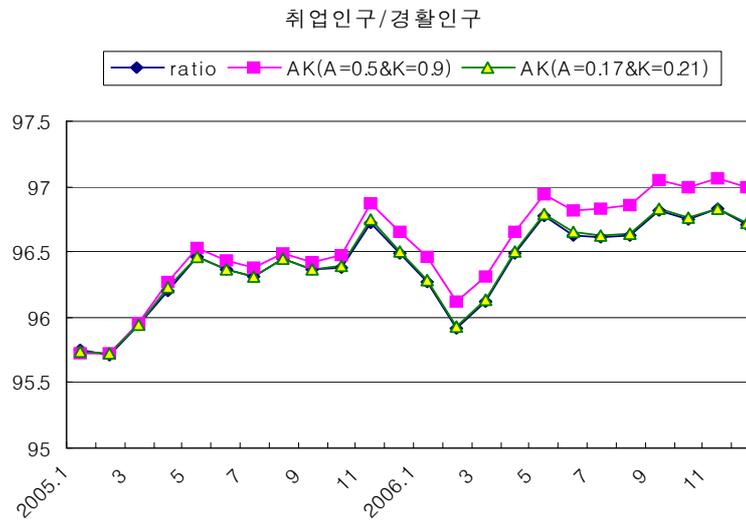
[그림 2-1]과 [그림 2-2]는 비추정방법, AK 복합추정량 I, AK 복합추정량 II를 이용하여 구한 추정값의 구성비를 그래프로 나타낸 것으로, AK 복합추정량의 가중값인 A 와 K 의 선택에 따라 구성비의 형태가 크게 달라짐을 확인할 수 있다.

AK 복합추정량(AK 복합추정량 I과 AK 복합추정량 II를 이용한 경우)의 추정값 및 구성비 형태가 연동이 지속됨에 따라 비추정량과 점차 달라지는 이유는 AK 복합추정량에 이용되는 가중값의 차이 때문이다. 당월의 비추정값과 전월확정값 및 전월과 금월 간의 변동분 추정값, 그리고 해당월에 새로 진입하는 진입표본의 영향력에 각각 얼마만큼의 가중을 주었느냐에 영향을 받게 되는 것으로, 해당월의 비추정값에 큰 가중을 줄수록 비추정값과 유사한 형태의 AK 복합추정값을 얻게 된다는 것을 알 수 있다. 이 때 AK 복합추정량의 추정값과 구성비에 영향을 주는 가중값 A 와 K 는 연속적으로 조사되는 자료들의 상관관계에 큰 영향을 받게 된다. “순환표본설계의 효율성에 관한 연구”(김영원·이수희, 2001)에 의하면, 연동표본조사법을 이용한 AK 복합추정량의 경우 월별 중복되는 표본들 간의 상관관계가 높을수록 독립표본추출법을 이용한 단순추정량 사용에 비해 상대효율이 높아진다고 하였으며, 해당월의 비추정값에 비해 전월확정값 및 전월과 금월 간의 변동분 추정값에 가중을 크게 하는 것이 보다 효율적인 추정값을 얻을 수 있는 방법이 된다는 것을 보였다.

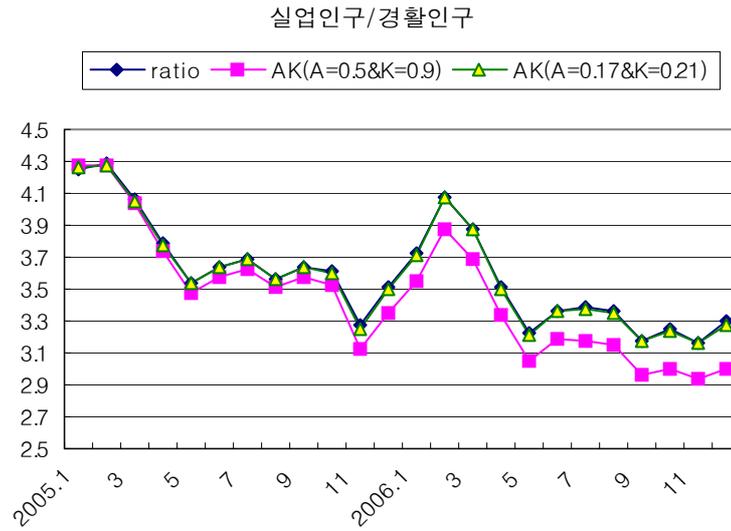
[그림 2-1] 경제활동인구비율



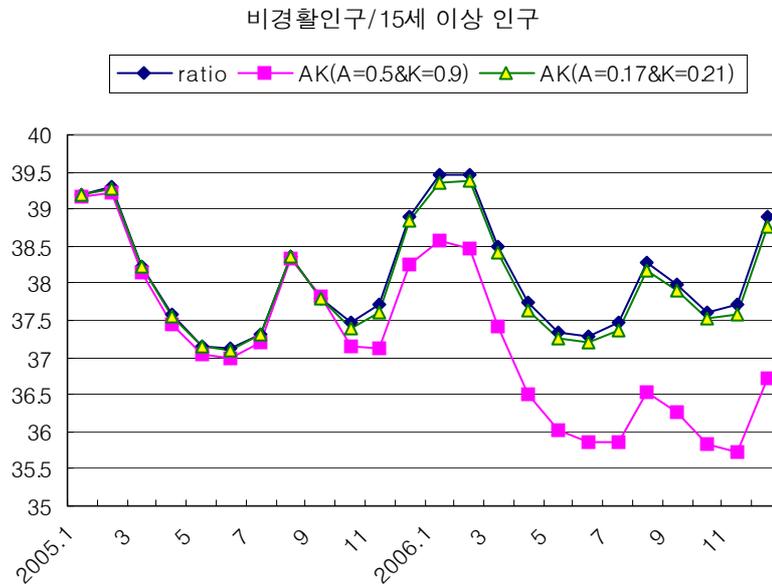
[그림 2-2] 취업률



[그림 2-3] 실업률



[그림 2-4] 비경제활동인구비율



본 연구에서 이용된 2004년 12월부터 2006년 12월까지의 자료를 살펴보면, 표본에 포함된 부표본들은 해당월과 지난 월들 사이에 높은 상관관계를 지속적으로 유지하고 있는 것을 알 수 있으며, AK 복합추정량에서의 분산을 최소화하는 가중값은 전월확정값 및 전월과 금월 간의 변동분 추정값에 가중을 크게 하는 것으로 결정되었다.

제6절 결론 및 제언

표본의 연동교체가 추정량에 영향을 미치는지 여부와 그 영향력의 형태를 알아보기 위하여 2년 동안 지속된 연동표본을 이용하여 비추정량과 AK 복합추정량을 비교하였다. 이 때 AK 복합추정량은 AK 복합추정량의 분산을 최소화하는 경우(전월확정값 및 전월과 금월 간의 변동분 추정값에 큰 가중을 주는 경우; $A = 0.5$, $K = 0.9$)와 비추정량에서의 15세 이상 인구수와 가장 유사한 추정값을 제공하는 경우(해당월의 비추정값에 가중을 크게 주는 경우; $A = 0.17$, $K = 0.21$)로 나누어 고려하였다.

AK 복합추정량의 분산을 최소화하는 가중값($A = 0.5$, $K = 0.9$)을 이용한 경우에는 연동이 지속됨에 따라 비추정량을 이용한 경우와 추정값에 있어 큰 차이를 보임을 알 수 있으며, 구성비에 있어서는 연동 시작 후 약 1년 동안은 유사한 형태를 유지하다가 연동이 지속됨에 따라 비추정값에 비하여 경활인구수는 크게, 비경활인구수는 작게 추정하는 형태를 보임을 알 수 있었다. 또한 비추정량에서의 15세 이상 인구수와 가장 유사한 추정값을 제공하는 경우($A = 0.17$, $K = 0.21$)에는 분산을 최소화하는 가중값을 이용한 경우에 비해 비추정값과 가까운 추정값을 보였으며, 구성비의 경우에는 연동 지속 여부와 관계없이 비추정값과 유사한 형태를 지속적으로 유지함을 알 수 있었다.

1/36 연동표본조사를 시행하고 있는 경찰조사에서의 부표본간 상관계수는 매우 높게 나타났으며, 이는 전월확정값이 해당월의 추정값에 큰 영향을 미치고 있음을 의미한다. 따라서 AK 복합추정량에서의 분산을 최소화하는 가중값 A 와 K 를 구하고, 이들을 이용하여 추정값을 유

도하는 것이 보다 정확한 추정값을 얻을 수 있는 방법이라고 할 수 있을 것이다.

2005년 1월부터 전국에 걸쳐 시행된 연동표본조사에서 처음 표본에 포함되었던 부표본들은 2007년 12월의 조사를 마지막으로 모두 교체된다. 따라서 3년 동안의 자료가 모두 수집된 이후에 AK 복합추정량을 구하여 기존 방법과의 효율을 비교하는 등 연동표본조사에 대한 전반적인 평가가 이루어져야 할 것이다. 또한 현재 시행 중인 1/36 연동표본조사는 한번 표본에 포함되면 3년 동안 조사가 지속되는 형태이므로, 응답자의 부담이 여전히 큰 부분으로 남아 있으며 모집단 반응률이 점차 낮아진다는 단점도 가지고 있다. 따라서 조사원의 업무량을 고려하여 조사 기간을 단축하는 방법을 고려하거나, 4-8-4 시스템과 같이 일정 기간 조사 후 조사를 중지했다가 다시 재조사하는 방법 등에 대한 연구가 지속적이고 체계적으로 이루어져야 할 것이다.

참고문헌

- 김기환 · 박유성 · 남궁재은(2004), "1/G 교체표본디자인에서의 일반화복합추정량과 평균제곱오차에 관한 연구", 「응용통계연구」, 제17권 제1호, pp.61-73.
- 김영원 · 이수희(2001), "순환표본설계의 효율성에 관한 연구", 「계간국민계정」, 제4호, pp.28-45.
- 박유성 · 김나영(1997), "표본교체방식의 비교와 일반화", 「통계분석연구」, 제2권 제1호, pp.125-142.
- 박유성 · 문원기 · 김기환(2002), "3개의 교체그룹을 갖는 2수준 교체표본설계에서의 복합추정량에 관한 연구", 「응용통계연구」, 제15권 제1호, pp.45-55.
- 박유성 · 배경화 · 김기환(2001), "P-수준교체표본에서 교체그룹내 상관관계를 고려한 일반화 복합추정량", 「응용통계연구」, 제14권 제1호, pp.81-90.
- 박홍래(1989), 「통계조사론」, 영지출판사, 서울.
- 윤연옥 · 김규영 · 이명호(2004), "통계청 가구부문 조사의 표본설계", 「조사연구」, 제5권 제1호, pp.103-130.
- 정구현(2005), "한국형 연동표본 모형에 관한 고찰", 「통계연구」, 제10권 제2호, pp.74-97.
- Bailer, B.A.(1975), "The Effect of Rotation Group Bias on Estimator from Panel Surveys", *Journal of the American Statistical Association*, 70, pp.23-29.
- Binder, D.A. and M.A. Hidiroglou(1988), "Sampling in Times", *Handbook of Statistics*, Vol.6, pp.187-211.
- Breau, P. and L. Ernst(1983), "Alternative Estimators to the Current Composite Estimator", *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp.397-402.

- Caldwell C.V. and P.J. Cantwell(1996), "Comparing a Fixed-Panel Sample Design with a Rotating-Panel Sample Design for Monthly Trade Surveys", Bureau of Census Annual Research Conference.
- Cantwell, P.J.(1988), "Variance Formulae for Generalized Composite Estimators under a Balanced One-Level Rotation Plan", *Bureau of the Census Statistical Research Division Research Report*.
- Cantwell, P.J.(1990), "Variance Formulae for Composite Estimators in Rotation Designs", *Survey Methodology*, 18, pp.153-183.
- Hansen, M.H., W.N. Hurwitz, and W.G. Madow(1953), *Sample Survey Methods and Theory*, Vol.2, Wiley, New York.
- Huang, E.T. and L. Ernst(1981), "Comparison of an Alternative Estimator to the Current Composite Estimator in the Current Population Survey", *Proceedings of the American Statistical Association, Section on Survey Research Methods*, pp.303-308.
- ILO(2004), Source and Methods: Labour Statistics, Vol.2, *Economically Active Population, Employment, Unemployment and Hours of Work*, ILO.
- Kumar, S. and H. Lee(1983), "Evaluation of Composite Estimation for the Canadian Labor Force Survey", *Survey Methodology*, 9, pp.178-201.
- Park, Y. and K. Kim(2000), "Generalized One-Level Rotation Designs with Finite Rotation Groups", *Journal of Korean Statistical Society*, 20, pp.29-44.
- Park, Y., K. Kim, and J. Choi(2001), "One-Level Rotation Designs Balanced on Time in Monthly Sample and in Rotation Group", *Journal of the American Statistical Association*, Vol.96, No.456, pp.1483-1496.
- Patterson, H.D.(1950), "Sampling on Successive Occasions with Partial Replacement of Units", *Journal of Royal Statistical Society, series B*, 12, pp.241-255.
- Rao, J.N.K. and J.E. Graham(1964), "Rotation Design for Sampling on Repeated Occasions", *Journal of the American Statistical Association*, 59, pp.492-509.
- United Nations(2005), *Designing Household Survey Samples: Practical Guidelines*, New York.

- Wolter, K.M.(1979), "Composite Estimation in Finite Populations",
Journal of the American Statistical Association, 74, pp.604-613.
- Wolter, K.M.(1985), *Introduction to Variance Estimation*, Springer-Verlag.
- Yates, F.(1949), *Sampling Methods for Censuses and Survey*, Charles
Griffin, London.

< 부 록 >

1. 수준추정 AK 복합추정량의 분산

36개월 연속조사(경황조사)에서의 일정시점의 수준을 추정하기 위한 AK 복합추정량은 다음과 같다.

$$y_t^{AK} = y_t + K(\Delta_{t,t-1} + y_{t-1}^{AK})$$

여기서 $y_t = \frac{1}{36} \left[(1-K+A)X_{t,1} + \left(1-K-\frac{A}{35}\right) \sum_{j=2}^{36} X_{t,j} \right]$ 이므로,

$$\begin{aligned} \text{Var}(y_t^{AK}) &= \text{Var}[y_t + K\Delta_{t,t-1} + Ky_{t-1}^{AK}] \\ &= [\text{Var}(y_t) + K^2 \text{Var}(\Delta_{t,t-1}) + K^2 \text{Var}(y_{t-1}^{AK}) \\ &\quad + 2K \text{Cov}(y_t, \Delta_{t,t-1}) + 2K \text{Cov}(y_t, y_{t-1}^{AK}) \\ &\quad + 2K^2 \text{Cov}(\Delta_{t,t-1}, y_{t-1}^{AK})] \\ &= [\text{Var}(y_t) + K^2 \text{Var}(\Delta_{t,t-1}) + 2K \text{Cov}(y_t, \Delta_{t,t-1}) \\ &\quad + 2K \text{Cov}(y_t, y_{t-1}^{AK}) \\ &\quad + 2K^2 \text{Cov}(\Delta_{t,t-1}, y_{t-1}^{AK})] / (1-K^2) \end{aligned}$$

이다.

우변에서의 y_{t-1}^{AK} 를 제거하기 위하여 해당 항을 다음과 같이 확장한다.

$$y_{t-1}^{AK} = \sum_{g=1}^{36} (K^{g-1} y_{t-g} + K^g \Delta_{t-g,t-g-1}) + K^{36} y_{t-37}^{AK}.$$

위 식을 원 식에 대입하여 정리하면,

$$\begin{aligned}
\text{Var}(y_t^{AK}) &= [\text{Var}(y_t) + K^2 \text{Var}(\Delta_{t,t-1}) + 2K \text{Cov}(y_t, \Delta_{t,t-1}) \\
&\quad + 2K \text{Cov}(y_t, \sum_{g=1}^{36} \{K^{g-1} y_{t-g} \\
&\quad + K^g \Delta_{t-g,t-g-1}\} + K^{36} y_{t-37}^{AK}) \\
&\quad + 2K^2 \text{Cov}(\Delta_{t,t-1}, \sum_{g=1}^{36} \{K^{g-1} y_{t-g} \\
&\quad + K^g \Delta_{t-g,t-g-1}\} + K^{36} y_{t-37}^{AK})] / (1 - K^2)
\end{aligned}$$

이고, 0인 항을 제거하면 다음과 같이 정리할 수 있다.

$$\begin{aligned}
\text{Var}(y_t^{AK}) &= [\text{Var}(y_t) + K^2 \text{Var}(\Delta_{t,t-1}) + 2K \text{Cov}(y_t, \Delta_{t,t-1}) \\
&\quad + 2 \sum_{g=1}^{36} K^g \{ \text{Cov}(y_t, y_{t-g}) + K \text{Cov}(y_t, \Delta_{t-g,t-g-1}) \\
&\quad + K \text{Cov}(\Delta_{t,t-1}, y_{t-g}) \\
&\quad + K^2 \text{Cov}(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) \}] / (1 - K^2)
\end{aligned}$$

경찰조사의 경우 우변 항들은 다음과 같다.

$$\begin{aligned}
\text{Var}(y_t) &= \text{Var} \left\{ \frac{1}{36} \left[(1 - K + A) X_{t,1} + (1 - K - \frac{A}{35}) \sum_{j=2}^{36} X_{t,j} \right] \right\} \\
&= \frac{\sigma^2}{1260} [35(1 - K)^2 + A^2]
\end{aligned}$$

$$\text{Var}(\Delta_{t,t-1}) = \text{Var} \left[\frac{1}{35} \sum_{j=2}^{36} (X_{t,j} - X_{t-1,j-1}) \right] = \frac{2}{35} (1 - \rho_1) \sigma^2$$

$$\text{Cov}(y_t, \Delta_{t,t-1}) = \frac{1}{1260} (1 - \rho_1) \sigma^2 [35(1 - K) - A]$$

$$\text{여기서 } y_t = \frac{1}{36} \left[(1-K+A)X_{t,1} + (1-K - \frac{A}{35}) \sum_{j=2}^{36} X_{t,j} \right],$$

$$\Delta_{t,t-1} = \frac{1}{35} \sum_{j=2}^{36} (X_{t,j} - X_{t-1,j-1}) \text{ 이다.}$$

$$\begin{aligned} \text{Cov}(y_t, y_{t-g}) &= \frac{1}{1587600} \rho_g \sigma^2 [1225(1-K)^2(36-g) \\ &\quad + 70A(1-K)(g-18) - gA^2] \end{aligned}$$

$$\text{여기서 } y_{t-g} = \frac{1}{36} \left[(1-K+A)X_{t-g,1} + (1-K - \frac{A}{35}) \sum_{j=2}^{36} X_{t-g,j} \right] \text{ 이고,}$$

$g = 1, 2, \dots, 35$ 이다.

$$\text{Cov}(\Delta_{t,t-1}, y_{t-g}) = \frac{1}{44100} (\rho_g - \rho_{g-1}) \sigma^2 [35(1-K)(36-g) + gA],$$

여기서 $g = 1, 2, \dots, 35$ 이다.

$$\begin{aligned} \text{Cov}(y_t, \Delta_{t-g,t-g-1}) &= \frac{1}{44100} (\rho_g - \rho_{g+1}) \sigma^2 [35(35-g)(1-K) \\ &\quad - (35-g)A] \end{aligned}$$

$$\text{여기서 } \Delta_{t-g,t-g-1} = \frac{1}{35} \sum_{j=2}^{36} (X_{t-g,j} - X_{t-g-1,j-1}) \text{ 이고,}$$

$g = 1, 2, \dots, 34$ 이다.

$$\text{Cov}(\Delta_{t,t-1}, \Delta_{t-g,t-g-1}) = \frac{1}{1225} \sigma^2 [(35-g)(2\rho_g - \rho_{g-1} - \rho_{g+1})]$$

여기서 $g = 1, 2, \dots, 34$ 이다.

가계조사의 경우 우변 항들은 조건에 따라 서로 다른 형태로 구해지며, g 값에 따른 차이를 고려하여 경찰조사에서의와 같이 유도가 가능하다. 자세한 내용은 생략한다.

2. 경제활동인구조사에서의 연동모형도

년	월	panel 1																																			
		조사구1				조사구2				조사구3				조사구4				조사구5				조사구6				조사구7				조사구8				조사구9			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
2004	12	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
2005	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	3		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	4			4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	5				5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	6					6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	7						7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	8							8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	9								9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	10									10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	11										11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
	12											12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
2006	1											13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	2												14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	3													15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	4														16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	5															17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	6																18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	7																	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	8																		20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	9																			21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	10																				22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	11																					23	24	25	26	27	28	29	30	31	32	33	34	35	36		
	12																						24	25	26	27	28	29	30	31	32	33	34	35	36		
2007	1																							25	26	27	28	29	30	31	32	33	34	35	36		
	2																								26	27	28	29	30	31	32	33	34	35	36		
	3																									27	28	29	30	31	32	33	34	35	36		
	4																										28	29	30	31	32	33	34	35	36		
	5																											29	30	31	32	33	34	35	36		
	6																												30	31	32	33	34	35	36		
	7																													31	32	33	34	35	36		
	8																														32	33	34	35	36		
	9																															33	34	35	36		
	10																																34	35	36		
	11																																	35	36		
	12																																		36		
2008	1																																				
	2																																				
	3																																				
	4																																				
	5																																				
	6																																				
	7																																				

(계속)

년	월	panel 2																																						
		조사구1-1			조사구2-1			조사구3-1			조사구4-1			조사구5-1			조사구6-1			조사구7-1			조사구8-1			조사구9-1														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
2004	12																																							
2005	1																																							
	2	1																																						
	3	1	2																																					
	4	1	2	3																																				
	5	1	2	3	4																																			
	6	1	2	3	4	5																																		
	7	1	2	3	4	5	6																																	
	8	1	2	3	4	5	6	7																																
	9	1	2	3	4	5	6	7	8																															
	10	1	2	3	4	5	6	7	8	9																														
	11	1	2	3	4	5	6	7	8	9	10																													
	12	1	2	3	4	5	6	7	8	9	10	11																												
2006	1	1	2	3	4	5	6	7	8	9	10	11	12																											
	2	1	2	3	4	5	6	7	8	9	10	11	12	13																										
	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14																									
	4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15																								
	5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16																							
	6	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17																						
	7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18																					
	8	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19																				
	9	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20																			
	10	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21																		
	11	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22																	
	12	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23																
2007	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24															
	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25														
	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26													
	4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27												
	5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28											
	6	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29										
	7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30									
	8	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31								
	9	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32							
	10	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33						
	11	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34					
	12	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35				
2008	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	2		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	3			3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	4				4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	5					5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	6						6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	7							7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			

3. 가계조사에서의 연동모형도

년	월	panel 1																																						
		조사구1				조사구2				조사구3				조사구4				조사구5				조사구6				조사구7				조사구8				조사구9						
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
2004	12	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
2005	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	3			3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
	4				5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	5				5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	6				5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	7					7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	8							9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	9							9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	10							9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	11										11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	12											13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
2006	1										13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36						
	2										13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36						
	3											15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36							
	4															17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	5															17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	6															17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	7																19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36						
	8																				21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	9																				21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	10																				21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
	11																					23	24	25	26	27	28	29	30	31	32	33	34	35	36					
	12																						25	26	27	28	29	30	31	32	33	34	35	36						
2007	1																						25	26	27	28	29	30	31	32	33	34	35	36						
	2																						25	26	27	28	29	30	31	32	33	34	35	36						
	3																								27	28	29	30	31	32	33	34	35	36						
	4																										29	30	31	32	33	34	35	36						
	5																											29	30	31	32	33	34	35	36					
	6																											29	30	31	32	33	34	35	36					
	7																												31	32	33	34	35	36						
	8																														33	34	35	36						
	9																															33	34	35	36					
	10																															33	34	35	36					
	11																																35	36						
	12																																							
2008	1																																							
	2																																							
	3																																							
	4																																							
	5																																							
	6																																							
	7																																							

(계속)

