연구논문

우리나라 공식인구의 신뢰성 및 문제점에 대한 고찰*

Unreliability of Official Population in Korea

박유성** · 김기환*** · 김성용****
Yousung Park · Keewhan Kim · Seongyong Kim

한 국가의 공식인구에는 기준인구(base population), 조사간인구(intercensal population estimates), 조사후인구(postcensal population estimates), 그리고 추계인구(population projections) 등이 있다. 조사간인구, 조사후인구, 그리고 추계인구는 기준인구를 기반으로 하며, 대부분의 나라에서 센서스를 이용하여 기준인구를 생성하고 있다. 본 논문에서는 우리나라의 기준인구 작성 시 센서스 자료를 얼마나, 어떻게 사용하는지 살펴보고, 다른 여러 나라의 공식인구 작성방법에 대한 장·단점을 비교, 분석하고자 한다. 이를 통해 우리나라 공식인구의 문제점을 적시하고 그 해결책을 제시한다.

주제어 : 인구추계, 센서스, 기준인구, 조사간인구, 조사후인구

Base population, intercensal and postcensal population estimates, and population projections have been regularly published as official populations. Base population is usually made out based on census conducted every 5 or 10 years, and is most important from which intercensal, postcensal estimates, and population projections are derived. We investigate how base population in Korea is made out and then compare it with those of other countries. We also present problems arising from making out base population in Korea.

Key words : population projections, base population, postcensal estimates, intercensal estimates

* 이 논문은 2007년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(KRF-2007-411-J03301.)

** 고려대학교 통계학과 교수

*** 고려대학교 정보통계학과 교수

**** 교신저자(corresponding author): 고려대학교 통계학과 박사과정 김성용.
E-mail: yaba96@korea.ac.kr

Ⅰ. 인구란?

한 국가의 인구는, 국가마다 약간의 차이는 있으나, 일반적으로 해당 시점에 그 국가에 일정한 자격을 가지고 사는 사람의 집합을 의미하며, 이는 그 국가의 국적을 보유하고 있는 사람의 숫자와 학생비자 또는 노동비자 등을 보유하고 있는 적법이주자(legal immigrants)의 숫자에서 국외로 이주한 사람의 숫자를 뺀 사람 수로 계산된다(Preston et al. 2001). 국가 내의 지역별 인구는 해당지역의 거주자의 개념으로 계산되며 국제 인구이동뿐만 아니라, 국내이동도 고려하여 인구를 산출한다(Smith et al. 2001).

인구의 산출은 인구추정(population estimation)과 인구추계(population projection)로 구분된다. 인구추정은 현재(또는 최근의) 인구의 근사적인 숫자이며 인구추계는 인구의 변동요인(출생, 사망, 이동)에 가정을 부여한 후 이 가정하에서 기대되는 인구수를 말한다(Bryan 2004).

인구는 국가 및 지역개발 수립 및 평가를 위한 기초자료, 저출산, 고령화 대책, 지역별 교육 수급, 주택 수요, 환경, 식량, 교통, 사회복지 등의 정책수립 자료이며 출생률, 사망률, 취학률, 실업률, 기대여명 등 각종 사회·경제지표 및 국가 재정, 지역별 계획분배의 핵심자료이다. 또한 각종 표본조사에 있어 표본가중치 계산과 표본틀을 제공하고 기업의 사업계획, 시장조사, 투자계획 등의 기초자료로 사용된다. 그러므로 인구는 시의성(timeliness), 정밀성(accuracy), 그리고 신뢰성(reliability)이 반드시 확보되어야 한다.

Ⅱ. 공식인구

한 국가의 공식인구는 기준인구(base population), 조사간인구(intercensal population estimates), 조사후인구(postcensal population estimates), 그리고 추계인구(population projections)로 구분할 수 있다. 기준인구는 센서스를 실시한 후 센서스에서 발생한 오류를 보완한 인구를 말하고, 조사간인구는 기준인구 사이의 연도별(예를 들어, 우리나라의 경우 2001~2004) 인구의 추정치이고, 조사후인구는 최근 기준인구를 기반으로 다음 센서스까지(예를 들어, 우리나라의 경우 2006~2009)의 추정인구(preliminary estimates)이며, 끝으로 추계인구(population projections)는 최근 기준인구 또는 최근의 조사후인구를 기반으로 여러 가지 가정에 따라 예측된 미래 인구를 말한다. 그러므로 공식인구 가운데서 가장 중요

한 것은 기준인구이다.[1)]

1. 기준인구의 작성

기준인구의 작성 시기는 센서스를 기반으로 하기 때문에 센서스 시기와 일치한다. 각 나라의 기준인구의 작성방법을 살펴보면 다음과 같다.

캐나다의 경우, 센서스인구를 집계한 후 거주자가 있는 거주지를 거주자가 없는 것으로 잘못 분류(misclassified dwells)한 경우와 무응답(non-response)을 대체(imputation)기법에 의해 지역별/성별/연령별로 조정한 후(내적일치성(internal consistency)을 위해 상향식(bottom-up) 으로 조정), 이를 조정된 센서스인구(adjusted census)로 발표한다. 두 번 이상 센서스에 포함되거나 센서스에서 빠진 경우, 센서스 전에 사망 또는 이민한 경우, 그리고 일부자료의 누락(missing)이나 주소불명 등으로 인한 순누락률(under-coverage)을 성별, 연령별로 하향식(top-down)으로 추정한다. 조정된 센서스인구를 순누락률로 재조정하여 이를 기준인구(base population)로 한다. 연령별 분포를 위해 평활(smoothing) 기법을 사용하고, 지역별 순누락률을 구하기 위해 경험적 베이지안 회귀(empirical Bayesian regression)를 사용한다(Statistics Canada 2007).

미국의 경우, 센서스 후 센서스 결과에 대한 지방정부의 이의제기 프로그램인 Count Question Resolution(CQR)에 의해 센서스 결과를 조정 발표하고 이를 기준인구로 하고 있다(Census Bureau 2001; Census Bureau 2009a). 특히 순누락률을 구해 조정한 센서스 결과가 조정하지 않은 센서스 결과와 큰 차이가 없다고 보고하고 있다(Census Bureau 2004).

일본의 경우, 센서스인구를 기준인구로 사용하며, 센서스 일자와 기준인구의 일자가 10월 1일로 동일하다(Japan Statistics Bureau).

한국의 경우, 2000년 기준인구로부터 사망, 출산, 국제이동에 관한 동태자료를 인구방정식(즉, 현재인구 = 전년도 인구 + 출생 - 사망 + 인구유입 - 인구유출)에 의해 2005년 총 기준인구를 구하고 전국단위의 성별/연령별 기준인구를 작성한다. 이를 이용하여 지역별/성별/연령별 기준인구는 2005년 센서스 결과에 따라 비례 배분한다(통계청 2006; 통계교

1) 통계청에서는 "기준인구가 공식인구가 아니다."라고 밝히고 있다. 그러나 통계청 KOSIS에서 제공되는 추계인구에서 2000년 및 2005년 추계인구 등이 기준인구이며, 기준인구는 추계인구에 포함되므로 "기준인구가 공식인구가 아니다"라는 통계청의 주장은 설득력이 없다고 판단된다. 통계청의 장래인구추계(2006)에 따르면 통계청에서는 2005년 기준인구를 설정한 후, 이를 이용하여 2001년부터 2004년의 추계인구를 소급 확정한다. 그리고 이에 기반하여 2006년부터 2050년까지의 인구를 생성한 후, 이를 장래추계인구로 발표하고 있다.

육원 2008). 그러나 전국단위 성별, 연령별 기준인구의 작성방법은 명확하게 기술되어 있지 않다.

2. 조사후인구(Postcensal Estimates)

가장 최근의 기준인구로부터 출생, 사망, 인구이동에 관한 동태자료를 인구방정식에 의해 구한 인구를 조사후인구라고 한다. 일반적으로 조사후인구는 전국, 지역별, 소지역별로 독립적으로 인구를 구한 후, 소지역 성별/연령별 인구의 합이 지역별/성별/연령별 인구와 일치하도록 조정하고 이를 다시 전국 성별/연령별 인구와 일치하도록 조정(raking)한다(Census Bureau 2008; Census Bureau 2009a; Census Bureau 2009b).

조사후인구를 구하기 위해 이용되는 출생과 사망 자료는 행정자료를 이용하여 산출하기 때문에 지연신고, 국외출생, 그리고 사망원인의 불규명 등에 의해 오차가 발생한다. 이러한 문제 때문에 일반적으로 예비추정치를 구하고, 2~3년 후에 다시 최종추정치를 구한다. 인구이동은 국제이동과 국내이동으로 구분된다. 국제이동은 이민(immigrant, emigrant), 돌아온 이주자(returning emigrants), 순단기출국자(net temporary emigration), 그리고 순단기체류자(net non-permanent residents)로 정의되며 미국과 캐나다는 이민국자료와 표본조사를 통해 추정한다. 국내이동은 지역간 이동인구로서 child tax benefit 자료와 income tax return 자료 등을 이용하여 추정한다(Census Bureau 2009a; Statistics Canada 2007). 이러한 동태자료의 한계 때문에 조사후인구는 신뢰도가 높지 않은 예비 공식인구로 해석하여, 센서스가 실시되면 센서스 결과를 이용하여 조사간인구로 대체하게 된다(Census Bureau 2005; Statistics Canada 2007).

한국의 경우, 조사후인구는 사용하지 않고(동향인구라는 개념으로 사용하고 있으나 이는 단지 참고자료이다) 가정에 의해 만들어진 장래추계인구(projection population)를 사용하고 있다.

3. 조사간인구(Intercensal Estimates)

조사간인구는 2개의 센서스 사이에 있는 연도별 인구추정으로 기존의 조사후인구를 대체한 인구추정치이다. 최근의 기준인구가 작성되면 이 기준인구는 기존의 조사후인구와 차이가 발생하게 된다. 이를 울타리오차(the error of closure)라고 하며

울타리오차 = 조사후인구 − 기준인구

로 정의된다 (Preston et al. 2001; Siegel & Swanson 2004). 일반적으로 조사간인구는 울타리오차를 최근의 기준인구에서 멀수록 선형적으로 감소시켜 각 시점에 배분 후, 조사후인구에 더하여 구한다.

캐나다의 경우, 매 5년마다 센서스가 실시되고 가장 최근의 센서스는 2006년에 시행되었으므로 2006년 기준인구를 작성하고, 2001년의 기준인구로부터 이미 계산된 2006년 조사후인구를 이용하여 울타리오차(= 2006년 조사후인구 − 2006년 기준인구)를 계산한다. 2005년 조사간인구(= 2005년 조사후인구 − 울타리오차 × 0.8), 2004년 조사간인구(= 2004년 조사후인구 − 울타리오차 × 0.6), ..., 2002년 조사간인구(= 2002년 조사후인구 − 울타리오차 × 0.2)를 구한다. 즉,

$$INTR_{2006-t} = POST_{2006-t} - \left(1 - \frac{5-t}{5}\right) \times error\ of\ closure$$

가 된다(Statistics Canada 2007). 여기에서 $INTR_t$는 시점 t에서의 조사간인구이고 $POST_t$는 시점 t에서의 조사후인구를 나타낸다.

미국의 경우도 각 시점에서 현재의 기준인구에 가까울수록 울타리 오차를 많이 보정한 조사간인구를 구하고 있다. 1990년과 2000년 사이의 조사간인구를 구하기 위해서 Census Bureau는 다음의 식을 이용하였다(Census Bureau 2005).

$$INTR_t = POST_t \left(\frac{BasePopulation_{3653}}{POST_{3653}} \right)^{\frac{t}{3653}}$$

여기서 $POST_{3653}$은 2000년 4월 1일의 조사후인구이며, $BasePopulation_{3653}$은 2000년 4월 1일의 기준인구이며, 3653은 10년간의 날짜 수이다. 울타리오차는 전국, 지역별 총인구의 울타리오차와 전국, 지역별 성별/연령별 울타리오차를 구해 지역별, 연령별, 성별 총합(marginal total)이 일치하도록 조정(즉, 내적 일치성)한 후 위의 공식에 의해 조사간인구를 구한다.

한국의 경우 기준인구가 결정되면 인구동태자료를 역산하여 조사간인구를 구한다(통계

청 2006; 통계교육원 2008). 예를 들어 2005년 기준인구가 결정되면 출생, 사망, 그리고 인구이동을 역으로 계산하여(즉, 인구방정식을 역으로 적용하여) 2000년까지 계산한 후, 이 값을 2000년 기준인구로부터 빼주어 울타리 오차를 구한다. 그러므로 한국의 2005년 울타리 오차는 미국과 캐나다와 달리 2000년 기준인구를 대비하여 구하고 있다. 그리고 울타리 오차의 배분방식은 명시되어 있지 않지만 선형적 감소형태가 아닌 균등배분(즉, 0.2 × 울타리 오차)을 하는 것으로 알려져 있다. 그러므로 한국의 경우, 울타리 오차를 5년 전인 2000년 결과에 따라 구하게 되어 오차 보정의 의미가 퇴색되었다고 볼 수 있다. 지역별, 연령별, 성별 울타리 오차의 계산방식과 배분방식은 동일하지만 울타리 오차의 내적일치성은 어떻게 해결하는지에 대해서는 알려진 바가 없으며 지역별 기준인구는 센서스 결과에 따라 비례 추정하므로 내적 일치성 만족은 거의 불가능한 것으로 판단된다.

Ⅲ. 우리나라 공식인구의 문제점

1. 총인구의 결정

조사간인구와 조사후인구의 근간은 기준인구이므로 공식인구 중 기준인구를 중심으로 우리나라 공식인구의 문제점을 살펴보고자 한다. 2.1절에서 언급하였듯이 우리나라 2005년 총 기준인구(total base population)는 2000년 기준인구로부터 구한 총 조사후인구이고 전국 단위 성별, 연령별 기준인구의 작성방법은 알 수 없으며 이 성별, 연령별 기준인구는 다시 지역별로 2005년 센서스 결과에 따라 비율 배분하여 지역별 기준인구를 결정한다.

우리나라 공식인구의 가장 큰 문제는 총인구가 센서스와는 무관하게 결정된다는 것이다.[2] 센서스에서 발생하는 순누락률이 보완되면 총인구는 센서스로부터 얻을 수 있는 가장 정확한 통계이다(Bryan 2004). 센서스를 실시하는 미국 및 캐나다, 일본의 경우 총인구는 센서스 결과를 그대로 이용하거나, 순누락을 보정한 센서스의 총인구를 이용한다. 그러나 우리나라의 경우 센서스를 실시함에도 불구하고, 유독 조사후인구로 총인구를 구하여 이용하는 타당한 이유가 있을까?

---

2) 통계청에서는 “우리나라 총인구의 결정시 센서스가 직접 이용되지 않을 뿐이며 센서스를 이용하지 않는다고 말할 수 없다”고 밝히고 있다. 그러나 통계청(2006)과 통계교육원(2008)에 따르면 통계청에서는 센서스와 무관한 인구방정식에 의해 총인구를 결정하였다.

여기에 더해서 ≪인구통계 기초와 응용≫(통계교육원 2008)에 따르면, 센서스 직후에 실시하는 사후표본조사(post enumeration survey) 결과 0.9%(425천 명)정도의 순누락률이 발생하였으나 총 기준인구를 조사후인구(48,181천 명)로 한 결과 센서스 총인구(47,279천 명)와의 차이가 1.9%(902천 명)로 증가하였다고 밝히고 있다. 이 1%(약 477천 명)의 차이를 어떻게, 어떤 근거로 채웠는지도 논의되어야 한다.[3)]

그러나 이 순누락률마저 문제가 있는 것으로 드러났다. 국가통계 포털인 KOSIS에 제공된 센서스 총인구(47,041천 명)를 이용하여 순누락률을 재산출해 보면 센서스의 순누락률은 2.42%(1,138천 명)로 나타났고, 이 중 남자의 순누락률은 3.17%(743천 명) 그리고 여자는 1.68%(395천 명)로 계산되었다. 이러한 순누락률의 차이는 어디에서 발생하는 것이며 성별로 순누락률이 차이가 생기는 근거는 무엇인지에 대하여서 진지하게 논의되어야 할 것이다.

이러한 모든 논의가 전제된 후에, 현재의 총인구를 구하는 방법에 대하여 그 신뢰도를 점검하여야만 한다. 그러나 이러한 논의를 차치하더라도, 2005년의 공식 총인구인 2005년 총 조사후인구가 신뢰성이 있기 위해서는 동태자료(즉, 출생, 사망, 국제 인구이동)의 정밀성이 전제되어야 한다. 따라서 다음 절에서는 동태자료의 신뢰성에 대해 파악해 보도록 한다.

2. 출생, 사망, 인구이동에 대한 동태자료는 믿을만한가?

통계청에서는 출생기준 익년도 4월까지의 지연신고 접수 후 출생자 수를 작성하고 있다(통계청 2009a). 〈표 1〉은 행정안전부가 제공한 2000년생과 2005년생의 연도별 지연신고건수이다. 2000년생의 2001년 5월부터의 지연신고를 날짜수에 비례하여 추정하면 25,766명이며, 2005년생의 경우 15,026명으로 추정된다. 이는 통계청의 출생자 수가 과소 집계되고 있음을 보이고 있다.[4)]

---

3) 통계청에서는 "2005년 인구총조사 당시 외국인에 대한 누락이 많아 법무부에 의뢰하여 파악한 외국인 수를 이용하였기 때문에 1%의 차이가 발생한다"라고 밝히고 있다. 그러나 외국인 인구를 성별/연령별/지역별로 어떻게 추정하는지에 대해 충분한 논의가 있어야 한다. 이와 더불어 외국인의 포함범위 역시 정의되어야 한다. 법무부의 외국인 수는 불법체류자를 포함한 수이며, 미국, 캐나다, 일본은 공식 인구에서 불법체류자를 포함하지 않는다.

4) 통계청에서는 2000년생의 경우 5,706명, 2005년생의 경우 2,791명의 지연신고가 있다고 밝히고 있다.

〈표 1〉 2000년 및 2005년 연도별 출생자 지연 신고

(단위: 명)

연 도	2000년생	연 도	2000년생	2005년생
2001	32,641	2006	348	20,139
2002	1,368	2007	218	819
2003	788	2008	83	412
2004	604	2009	57	277
2005	390			

〈표 2〉 2000년~2005년생의 출생자 수 비교

(단위: 명)

구분	2000년생	2001년생	2002년생	2003년생	2004년생	2005년생
주민등록 최대	639,800	559,670	494,660	492,914	472,935	434,320
출생자 수	634,501	554,895	492,111	490,543	472,761	435,031

〈표 3〉 국제이동 통계

(단위: 명)

연령	2000	2001	2002	2003	2004	2005	2006	2007	2008
0세	-3,593	-3,629	-3,186	-3,366	-3,387	-3,562	-3,545	-2,979	-3,250
1세	-1,051	-1,386	-1,014	-1,159	-1,086	-1,588	-1,485	-1,475	-1,024
2세	-884	-1,132	-699	-840	-815	-837	-1,015	-781	-372
3세	-878	-1,163	-720	-867	-956	-792	-1,211	-866	-590
4세	-1,039	-1,259	-921	-976	-938	-1,156	-1,379	-1,149	-837
5세	-1,097	-1,567	-985	-1,275	-1,249	-1,509	-1,970	-1,307	-1,115
6세	-956	-1,359	-1,165	-1,140	-1,055	-1,080	-1,782	-1,267	-822
7세	-1,294	-1,685	-1,247	-1,430	-1,369	-1,471	-2,061	-1,641	-1,278
8세	-1,408	-1,841	-1,539	-1,724	-1,649	-1,595	-2,420	-1,815	-1,259
9세	-1,718	-2,312	-1,906	-2,076	-1,981	-2,291	-3,311	-2,349	-1,730
10세	-1,647	-2,248	-2,212	-2,355	-2,504	-3,015	-4,259	-2,945	-2,369

〈표 2〉는 2008년까지의 주민등록인구를 분석한 것으로 2000년생은 8세까지의 최대인구, 2001년생은 7세, 2002년생은 6세, 2003년생은 5세, 2004년생은 4세, 2005년생은 3세까지의 최대주민등록인구와 통계청 출생자 수를 비교한 것이다. 표에서 볼 수 있듯이 관찰치가 가장 작은 2005년생을 제외하고 2000년생~2004년생 모두 출생자 수보다는 주민등록 최대인구가 큰 것으로 나타났다. 이는 인구의 국내유입이 많을 경우에 일어날 수 있다. 그러나 〈표 3〉의 국제이동 통계에서 볼 수 있듯이 2000년부터 2008년까지 0세~10세의 국제이동은 모두 음의 값을 가지고 있어 국외유출이 더 많다는 것을 의미하므로 0세~8세 사이의 주민등록 최대인구가 출생자 수보다 크게 나오는 것은 불가능하다. 이는 출생자 수에 대한 통계가 틀렸거나 국제이동 통계가 틀렸다는 것을 의미한다. 또한 〈표 3〉에서 0세의 국외유출이 다른 연령에 비해 매우 크게 나타나는 것에 대해 정밀점검이 필요하다고 할 수 있다.

〈표 4〉는 0세의 추계인구와 7월 1일 기준으로 산출한 월별 출생자 수를 비교한 표이다. 추계인구가 7월 1일 기준으로 작성되기 때문에 기준일자의 차에 의한 오류를 없애기 위하여 출생자 수를 7월 1일 기준으로 산출하였다. 예를 들어 〈표 4〉에서 첫 번째 열의 01/7/1의 출생자 수는 2000년 7월 1일부터 2001년 6월 30일 사이에 태어난 신생아 수를 의미한다. 사망이나 국외이동을 하지 않으면 이들은 모두 2001년 7월 1일 0세로 집계된다. 그러므로 이들의 국제이동이 양이 아닌 한 (즉 국내유입이 많지 않은 한) 0세 인구가 출생자 수보다 클 수가 없다. 〈표 3〉을 보면 0세의 국외유출이 2000년부터 2008년까지 더 많은 것으로 나타나, 〈표 4〉의 02/7/1, 03/7/1, 05/7/1와 같이 추계 0세 인구가 출생자 수보다 많은 것은 출생통계나 국제이동 통계의 오류가 아니고는 설명할 수 없다.

다음은 사망자 통계의 신뢰도를 살펴보자. 〈표 5〉는 2000년~2008년까지의 행정안전부 사망통계와 통계청의 사망자 동태자료를 비교하고 있다.

행안부의 사망자 통계가 통계청의 사망자 동태자료보다 항상 크게 나타나고 있다. 통계청의 사망자료가 익년도 4월까지의 접수를 반영하여 작성된다는 점을 감안할 때 통계청의 사망자 동태자료 역시 그 신뢰성이 의심스러울 수밖에 없다고 할 수 있다.

〈표 4〉 출생자 수와 추계 0세 인구의 비교

(단위: 명)

날 짜	01/7/1	02/7/1	03/7/1	04/7/1	05/7/1
연간 출생자수	602,530	519,829	483,634	484,057	449,623
0세 추계인구	596,105	533,117	494,291	480,092	453,778

〈표 5〉 행안부 및 통계청 사망통계 비교

(단위: 명)

연 도	사망자료	
	통계청	행안부
2000	246,163	249,197
2001	241,521	249,302
2002	245,317	248,711
2003	244,506	247,904
2004	244,217	247,102
2005	243,883	267,485
2006	242,266	240,092
2007	244,874	245,015
2008	246,113	247,389

그러나 지금까지 살펴본 동태자료(출생, 사망, 국제이동)의 부정확성은 우리나라뿐만 아니라 미국, 캐나다, 일본을 포함한 거의 모든 나라에서 발생하는 문제들이다. 이러한 동태자료의 한계가 동태자료로만 작성된 조사후인구를 센서스 후에는 조사간인구로 교체하는 주된 이유이며, 기준인구로 센서스 인구를 이용하는 이유이다. 이러한 관점에서 볼 때 동태자료로만 계산된 총 조사후인구인 2005년 총 기준인구는 이론적인 토대가 상실된 추정인구라고 말할 수 있다.

3. 성별, 연령별 기준인구의 특징은 무엇인가?

3.1절에서 밝힌 바와 같이 총인구는 센서스 총인구를 사용하지 않고 있다. 3.3절에서는 기준인구의 성별, 연령별 인구구조가 센서스의 성별, 연령별 인구구조를 따르고 있는지를 파악하기 위해 인구구조의 시계열적 변화와 2000년 기준인구와 2005년 기준인구의 인구피라미드를 살펴보도록 한다.

〈표 6〉은 주민등록 인구로, 2000년 0세~5세의 인구가 2005년에는 각각 5세~10세가 되므로 이 인구의 변화를 보여 주고 있다. 남자의 경우 2000년 0세~2세 인구가 5년 후인 2005년까지 증가하고 있고, 여성의 경우 2000년 0세~3세의 인구가 2005년까지 증가하는 것으로 나타났다. 이러한 증가현상은 〈표 7〉, 〈표 9〉에서 볼 수 있듯이 우리나라의 센서스, 일본의 센서스, 미국의 센서스 모두 다소의 차이가 있을 뿐 동일하게 관찰되고 있다. 특히, 2000년

0세~2세까지는 남녀 모두 2005년까지 인구가 증가하고 있는 것을 알 수 있으며 2000년 기준으로 다른 연령대에서도 무시해도 좋을 만큼의 미세한 인구감소가 관찰되고 있다. 이러한 현상은 지연신고나 국제인구 이동의 증가에 기인한 것으로 파악된다.

그러나 〈표 10〉의 기준인구(또는 추계인구)는 2000년 여자 0세를 제외하고 모두 2000년에서 2005년까지 인구가 감소한 것으로 나타나, 우리나라의 주민인구와 센서스뿐만 아니라 다른 나라의 센서스 결과와 다른 모습을 보여 주고 있다. 이는 2005년 기준인구 작성 시 0세~4세의 인구보정이 높은 것과 무관하지 않으며, 유독 기준인구만 0세~5세의 인구구조 시계열이 다른 인구구조 시계열과 다른지 의문이 있을 수밖에 없다.

〈표 6〉 주민등록 인구(12월 31일 기준)

(단위: 명)

남 자				여 자			
2000년		2005년		2000년		2005년	
연령	인구	연령	인구	연령	인구	연령	인구
0	318,375	5	334,651	0	289,685	5	304,827
1	323,232	6	324,142	1	295,212	6	296,393
2	335,558	7	335,899	2	304,886	7	305,565
3	350,029	8	349,730	3	323,560	8	323,805
4	362,323	9	362,063	4	325,148	9	325,104
5	377,334	10	376,513	5	333,801	10	333,308

〈표 7〉 한국: 센서스(11월 1일 기준)

(단위: 명)

남 자				여 자			
2000년		2005년		2000년		2005년	
연령	인구	연령	인구	연령	인구	연령	인구
0	315,765	5	322,240	0	283,308	5	294,871
1	307,738	6	310,808	1	281,693	6	285,496
2	327,973	7	328,836	2	299,046	7	301,198
3	337,686	8	338,321	3	310,009	8	313,432
4	352,004	9	354,023	4	315,036	9	319,662
5	364,263	10	364,196	5	321,957	10	323,941

〈표 8〉 **일본: 센서스**(10**월** 1**일 기준**)

(단위: 명)

남 자				여 자			
2000년		2005년		2000년		2005년	
연령	인구	연령	인구	연령	인구	연령	인구
0	600,466	5	607,278	0	571,186	5	575,699
1	596,910	6	604,292	1	569,250	6	575,444
2	610,326	7	610,904	2	581,831	7	582,445
3	607,950	8	608,086	3	581,353	8	580,785
4	606,869	9	605,943	4	577,957	9	577,619
5	615,685	10	616,199	5	588,448	10	588,325

〈표 9〉 **미국: 센서스** (4**월** 1**일 기준**)

(단위: 명)

남 자				여 자			
1990년		2000년		1990년		2000년	
연령	인구	연령	인구	연령	인구	연령	인구
0	2,019,081	10	2,191,244	0	1,928,232	10	2,082,812
1	1,928,747	11	2,108,157	1	1,840,807	11	2,006,936
2	1,894,973	12	2,087,228	2	1,807,706	12	1,988,614
3	1,863,343	13	2,054,008	3	1,778,306	13	1,956,842
4	1,896,723	14	2,079,560	4	1,807,163	14	1,972,671
5	1,860,600	15	2,065,127	5	1,774,863	15	1,954,277

〈표 10〉 **한국: 추계인구**(7**월** 1**일 기준**)

(단위: 명)

남 자				여 자			
2000년		2005년		2000년		2005년	
연령	인구	연령	인구	연령	인구	연령	인구
0	325,856	5	325,188	0	296,242	5	296,691
1	330,941	6	325,419	1	295,159	6	294,667
2	335,555	7	334,268	2	308,212	7	306,044
3	355,177	8	347,904	3	318,379	8	317,365
4	368,619	9	362,740	4	325,643	9	323,257
5	378,213	10	372,208	5	331,183	10	327,141

〈표 11〉 2005년 기준, 현재, 주민, 센서스인구 비교

(단위: 명)

	남 자		여 자		전 체	
	오차합	오차절대합	오차합	오차절대합	오차합	오차절대합
기준-현재	-33,536	236,152	48,153	191,382	14,616	427,534
기준-주민	-253,091	514,114	-361,735	583,135	-614,827	1,097,249
기준-센서스	739,195	780,635	390,879	399,036	1,130,075	1,179,671

4. 성별, 연령별 기준인구는 센서스 인구를 사용하고 있는가?

우리나라의 2005년 성별, 연령별 기준인구의 작성은 자세한 수식이나 절차가 없이 통계교육원 발간 ≪인구통계 기초와 응용≫ p.419에 "주민등록인구 및 2005년 기준 현재인구(인구변동요인에 의해 작성)의 성별 연령별 인구와 비교 분석하고, 2005년 인구주택총조사 사후조사 결과에 의한 총 조사인구의 순누락률 0.9% 수준을 반영하여 … 연령별 집중현상(윤달영향 등)을 제거하기 위해 3개 연령층 이동평균을 적용한다"로 되어 있으며, 2006년 발간된 통계청의 ≪장래인구추계≫에도 이와 동일하게 성별, 연령별 기준인구 작성방법을 설명하고 있다.

2005년 성별, 연령별 기준인구가 센서스 자료를 어떻게 이용하였는지를 파악하기 위해서는 세세한 작성법을 알아야 비교가 가능하지만, 유일한 작성법인 위의 문장을 바탕으로 우리나라의 2005년 성별, 연령별 기준인구 작성방법을 파악해 보도록 한다. 본 논문에서는 동태자료를 이용하여 2005년 11월 1일 기준의 조사후인구(통계청 표현으로는 현재인구)를 작성하여 3개 연령 이동평균을 적용하여 이를 현재인구라고 정의한다. 또한 12월 1일 기준 주민등록인구를 3개 연령 이동평균으로 취한 후 이를 주민등록인구라고 정의하고 11월 1일 기준으로 작성된 센서스인구를 3개 연령 이동평균을 취한 후 이를 센서스인구라고 정의하였다. 사후표본조사(post enumeration survey) 후 순누락률이 조정된 센서스인구를 사용해야 하지만 1세 단위의 연령별 순누락률 자료가 없어 조정되지 않은 센서스인구를 사용하기로 한다. 이 3개의 인구를 기준인구(즉, 추계인구)와 비교하기 위해 기준인구 역시 11월 1일 기준으로 재 산출하였다.

〈표 11〉은 2005년 성별, 연령별 기준인구 구조가 3개의 인구 중 어떤 인구구조에 가장 가까운가를 측정하기 위해 다음의 3개의 오차, 즉, 기준인구－현재인구, 기준인구－주민등록인구, 기준인구－센서스인구를 성별, 연령별로 구해 오차합과 오차절대값의 합을 정리한

것이다. 이 표에 따르면 남·여 모두 현재인구가 기준인구에 가장 가까운 인구구조를 가지고 있는 것으로 나타났다. 특히, 현재인구와 기준인구의 총합의 오차는 불과 14,616명으로 현재인구와 기준인구의 총인구 추정치는 거의 일치하고 있으며, 남자의 경우 오차의 합이 −33,536명, 여자의 경우 48,153명으로 성별로도 기준인구구조와 매우 유사하다는 것을 알 수 있다. 그러나 기준인구의 구조는 주민등록인구구조와 센서스인구구조와는 상당한 거리가 있는 것으로 나타났다. 오차절대값의 측면에서도 현재인구가 가장 우수한 것으로 나타났으며 주민인구와 센서스인구는 현재인구에 비해 약 3배의 오차절대합을 갖는다.

이 결과는 우리나라의 2005년 기준인구(추계인구)는 2000년의 성별, 연령별 기준인구를 이용하여 산출한 조사후인구일 가능성이 매우 높다고 할 수 있다. 그러므로 통계청의 2005년 기준인구는 센서스 성별, 연령별 인구를 참고는 했을지 모르지만 기준인구 작성 시 센서스인구를 썼다는 증거는 없다고 해도 무리가 없는 것으로 보인다.[5)]

다음으로는 2005년 11월 1일 기준의 기준인구, 주민등록인구, 센서스인구를 또 다른 주요 통계인 학령인구(교과부 제공)와 비교함으로써 어느 인구가 현실적으로 정확한가를 간접적으로 점검할 수 있다. 이들을 비교할 수 있도록 초등학생 등록수(7세:1학년, 8세:2학년, … 12세:6학년)를 기준인구, 주민등록인구, 센서스인구와의 차이를 비교하여 〈표 12〉에 정리하였다.

〈표 12〉 2005년 초등학생 등록수와 기준, 센서스, 주민등록인구 비교

(단위: 명)

연 령	초등학생 등록수 - 기준인구		초등학생 등록수 - 주민등록인구		초등학생 등록수 - 센서스 인구	
	남 자	여 자	남 자	여 자	남 자	여 자
7	-3,409	-3,682	-8,268	-7,320	-1,205	-2,953
8	-1,601	2,742	-8,343	-7,808	3,066	2,565
9	-6,157	-4,697	-10,967	-8,919	-2,927	-3,477
10	-3,673	-600	-11,844	-8,611	473	756
11	-7,163	-6,606	-14,773	-12,283	1,936	-941
12	-4,411	-4,868	-9,793	-7,614	363	-4,868
합 계	-26,414	-17,711	-63,988	-52,555	1,706	-8,918

5) 2010년 4월 9일에 개최된 고려대학교 경제연구소 컨퍼런스인 “우리나라 공식인구, 이상하다”에서 통계청 관계자들은 통계청의 기준인구는 1) 주민등록인구 및 법무부 외국인 인구, 2) 센서스 내국인(사후조사 반영) 및 법무부 외국인 인구, 3) 동향인구(조사후인구) 가운데 하나를 성별, 연령별로 각각 선택하여 작성한다고 밝혔다.

이에 따르면 남자 8세, 여자 10세를 제외한 모든 연령에 걸쳐 남녀에 관계없이 초등학생 등록수와 가장 작은 오차를 보이는 것은 센서스인구이다. 특히, 학령과 학생의 실제연령이 일치하지 않으므로 초등학교 학생의 전 연령대(7세~12세)의 오차합을 구해 본 결과, 놀랍게도 센서스인구가 초등학교 학령인구와 거의 일치하고 있음을 볼 수 있다. 그러나 기준인구와 주민등록인구는 초등학교 학령인구와는 어느 정도 거리가 있는 것으로 나타났다. 이는 센서스인구가 문제가 있다는 통계청의 생각(증거를 제시한 적이 없기 때문에 이렇게 표현함)은 설득력이 떨어진다는 것을 반증하고 있다.

위에서 밝힌 바와 같이, 우리나라의 성별, 연령별 기준인구는 조사후인구를 이용하는 것으로 파악된다. 그러나 이로 인해 센서스와는 전혀 다른 인구구조를 가지게 되는데, 그 대표적인 예가 남녀 성비의 변화이다. 〈표 13〉은 2005년 기준인구와 11월 1일 기준의 기준인구, 2005년의 센서스인구 및 순누락률을 적용하여 보정한 센서스인구의 남녀 성비이다.

〈표 13〉에 따르면 센서스인구의 경우 순누락 보정여부에 상관없이 남자가 여자보다 더 많음에 반해, 기준인구는 여자가 남자보다 더 많음을 알 수 있다. 남자 인구와 여자 인구의 차이는 결혼적령기인 20세에서 39세의 인구에서 더 크게 나고 있는데, 이는 〈표 14〉와 같다.

〈표 14〉에 따르면 결혼적령기인 20~39세의 경우 기준인구는 센서스인구에 비해 여자 1000명당 남자가 약 24명 많은 것으로 파악된다. 이뿐만 아니라, 연령별 인구 역시 기준인구와 센서스인구의 구조는 상이한 것으로 나타난다.

〈표 13〉 남녀 성비

(단위: 명)

전체연령	전체	남	여	남녀 성비
기준	48,138,077	24,190,906	23,947,171	1.0102
기준(11)	48,180,190	24,209,169	23,971,021	1.0099
센서스	47,041,434	23,465,650	23,575,784	0.9953
순누락 보정센서스	47,464,807	23,665,108	23,802,112	0.9942

〈표 14〉 결혼적령기의 남녀 성비

(단위: 명)

20~39세	전체	남	여	남녀 성비
기준	16,121,080	8,284,910	7,836,170	1.0573
기준(11)	16,054,551	8,254,263	7,800,289	1.0582
총조사	15,543,037	7,899,815	7,643,222	1.0336
순누락 보정센서스	15,605,972	7,925,852	7,680,121	1.032

〈표 15〉 11월 1일 기준인구 – 순누락 보정 총조사 인구

(단위: 명)

연 령	전 체	남	여
0~4	64,324	41,166	23,183
5~9	-3,989	5,760	-9,733
10~14	3,571	12,218	-8,789
15~19	12,189	5,734	6,582
20~24	47,043	17,630	33,045
25~29	155,682	109,203	46,186
30~34	121,437	107,674	13,557
35~39	121,505	93,904	27,379
40~44	64,184	60,777	3,478
45~49	60,101	51,115	9,213
50~54	56,890	33,899	22,713
55~59	-9,808	-291	-9,623
60~64	14,747	8,100	6,713
65~69	4,269	4,847	-545
70~74	1,126	1,892	-665
75~79	3,138	2,127	1,033
80~84	4,065	1,846	2,191
85+	12,311	3,103	9,166
계	715,383	544,061	168,909

〈표 15〉는 2005년의 5세 연령별 11월 1일 기준의 기준인구와 순누락을 보정한 센서스인구의 차이이다. 여기서 알 수 있듯이, 기준인구는 0~4세 유아 및 25~54세의 노동인구에서 총조사인구보다 훨씬 더 많으며, 65세 이상 인구에서는 유사하다. 이러한 차이는 성별, 연령별 기준인구 작성 시 순누락 보정 센서스인구를 이용했다고 말하기도 어렵다는 것을 나타낸다.

5. 지역별, 성별, 연령별 기준인구와 기타 문제점

센서스는 전수조사이기 때문에 막대한 비용(2005년 1,290억원; 2010년 1,809억원)이 소

요된다. 전수조사를 하는 근본이유는 총인구구조뿐 만아니라 소지역 인구구조를 정확하게 파악해야 하기 때문이다. 통계청에서는 지역별, 성별, 연령별 추계인구를 구할 때 전국단위로 작성된 성별, 연령별 기준인구를 센서스의 지역별 비율에 따라 배분하여 사용한다(통계청 2006). 예를 들어 전국단위의 20세 남자의 인구가 결정되면, 이를 센서스의 남자 20세 인구의 시도별 비율에 따라 배분하여 시도별 20세 남자의 인구를 생성한다.

미국 및 캐나다는 전국단위의 성별, 연령별 인구와 지역단위의 성별, 연령별 인구를 독립적으로 구한 후, 이들이 서로 일치하도록 조정해 준다. 그러나 우리나라는 지역별 총인구 및 지역단위의 성별, 연령별 인구를 독립적으로 구하지 않고 있다. 더군다나 성 및 연령 집단 내에서의 지역분포를 이용함에 따라, 지역별 인구의 구조를 반영하지 못하는 한계점을 지니고 있다.

본 논문에서는 성 및 연령 집단 내의 지역분포가 아닌, 각 지역에서의 성별 비율을 이용하여 2005년 지역별 인구를 재산출하고 이를 기준인구와 비교하였다. 〈표 16〉에서 볼 수 있듯이, 2005년의 지역별 기준인구는 센서스에 비해 전 지역에서 남자는 과대추정되고, 여자는 과소추정되는 것으로 여겨진다.

여기에 연령을 더하여 지역별로 성별, 5세 연령별 비율을 이용하여 2005년 인구를 재산출하여 이를 기준인구와 비교한 결과는 〈표 17〉과 같다. 각 지역에서의 성별, 5세 연령별 오차절대합의 총합은 850,786 으로 그 차이가 매우 크며, 각 지역별 오차합의 경우도 규칙성이 없어 기준인구가 지역별 인구구조를 반영하고 있지 않음을 나타내고 있다.

〈표 16〉 기준인구 – 지역별 센서스의 성비를 이용한 추정치

(단위: 명)

지 역	남 자	여 자	지 역	남 자	여 자
서울특별시	46,897	-44,173	강원도	-1,966	-5,622
부산광역시	14,526	-19,591	충청북도	2,453	-9,920
대구광역시	20,325	1,155	충청남도	4,635	-9,374
인천광역시	10,653	-12,501	전라북도	-2,276	-12,074
광주광역시	9,725	-3,161	전라남도	7,491	-5,380
대전광역시	-2,134	-6,383	경상북도	31,365	2,100
울산광역시	2,781	-861	경상남도	14,845	-4,070
경기도	30,836	-59,676	제주도	2,563	-3,191

〈표 17〉 **지역별 성별, 5세 연령별 비율을 이용한 추정치와 기준인구 비교**

(단위: 명)

지 역	기준-추정	오차절대합	지 역	기준-추정	오차절대합
서울특별시	21,191	154,754	강원도	-6,459	27,021
부산광역시	-8,345	60,915	충청북도	-3,574	27,591
대구광역시	-6,873	42,511	충청남도	-4,749	37,427
인천광역시	1,983	45,915	전라북도	-3,783	38,178
광주광역시	-2,707	21,670	전라남도	-5,371	42,680
대전광역시	-3,770	22,908	경상북도	-3,555	50,820
울산광역시	720	17,271	경상남도	-3,693	57,550
경기도	30,377	194,534	제주도	-1,390	9,041

위 결과에서 보듯이, 우리나라의 지역별 기준인구는 각 성별 및 연령 집단 내에서의 지역비율을 이용하여 산출함으로써 지역별 인구구조를 반영하지 못하고 있다. 통계청에서는 왜 이와 같이 지역별 인구를 계산해야만 하는가?

기준인구(추계인구)작성 시, 현재인구, 주민등록인구, 센서스인구 중 하나만을 이용하였다면 본 논문에서 재산출한 것과 같이 지역별, 성별, 연령별 인구는 동시에 구해질 수 있다. 그런데 굳이 이러한 방법을 이용하는 것에 대한 질문에 통계청은 답변할 수 있는가?

기준인구의 부정확성은 조사간인구와 조사후인구에도 영향을 미칠 뿐만 아니라 미래인구 예측에도 직접적인 영향을 미치므로 기준인구의 정밀도는 반드시 확보되어야 한다. 그러나 현재의 기준인구는 그러하지 못하다는 것이 자명하다.

공식인구의 부정확성은 각종 사회·경제지표를 계산할 때 문제를 야기한다. 우리나라는 현재 출생률 계산 시 15~49세 주민등록인구로 분모를 사용한 후 추계인구로 환원하며, 사망률 계산 시 주민등록인구를 이용하여 계산하고 있으며 인구이동률은 추계인구를 사용하여 사회통계의 혼란을 야기하고 있다. 특히, 연금 등 재정계산을 위한 인구를 무엇으로 할 것인지, 실업률과 1인당 국민소득 등의 계산에서 분모인구는 무엇으로 해야 하는지 아직도 통일된 기준이 존재하지 않은 실정이다.

현재 통계청에서는 2015년부터 센서스를 행정자료를 이용한 등록 센서스로 전환하고자 하고 있다. 그러나 등록 센서스로 전환하기 이전에 다음의 문제들이 해결되어야 한다(양경진 2008; 이건 2009). 행정자료와 센서스인구 및 가구 수의 불일치, 세대와 가구 개념의 분

리, 개별 거처에 대한 식별, 행정자료에서 이용되는 항목의 개념과 센서스에서 이용되는 항목 개념의 불일치 등 등록 센서스를 실시하기 이전에 고려되어야 할 많은 문제들이 존재한다. 이러한 문제점에 대한 해결이 없이는 등록 센서스를 시행한다 하더라도, 인구추계에서의 문제는 계속 발생할 것이다.

Ⅳ. 장래인구추계

장래인구추계의 목적은 미래 인구의 구조를 파악하여 이에 따른 국가경영 발전계획 수립과 향후 인구와 관련된 각종 사회·경제지표를 작성하기 위한 기초자료 및 학술자료를 제공하기 위함이다. 미래 인구구조는 출생률, 사망률 및 국제이동에 대한 가정에서 출발한다. 이 세 가지의 인구변동요인 각각에 대해 모형을 설정하고 인구변동요인의 상위, 중위, 하위 값을 설정하여 미래 인구구조를 생성하게 된다(Siegel & Swanson 2004). 이 중 중위값은 출생률, 인구이동률, 사망률의 관찰자료에 의해 생성되므로, 현재의 인구변동요인의 추이가 변하지 않고 미래에도 지속된다는 가정에서 미래인구구조를 예측하는 값이 된다. 최근 통계청의 2050년까지의 장래인구추계는 이 중위값에서 계산된 값으로 해석할 수 있다.

우리나라 장래인구추계(통계청 2006; 통계교육원 2008)는 출산율의 추정을 위해 로그감마모형(Kaneko 2003)을 사용하였고, 국내이동은 최근 5년간 성별, 연령별 국제이동률의 평균치가 유지되는 것으로 가정하였다. 사망률은 Lee-Carter모형(Lee & Carter 1992)을 사용하였으나, 과거의 사망 연령의 부재에 의한 75세 이상 사망확률의 예측을 위해 Brass Logit모형(Brass 1971)을 사용하였다.

Park & Kim(2009)은 출산율은 현재와 동일하며 국제이동은 없다는 가정(Goerge & Perreault 2002)을 하였으며, 사망률은 GLM모형과 random walk prior분포를 이용한 Bayesian 모형을 사망원인별, 연령별 사망자료에 적용하여 장래인구를 예측하였다. 정부의 예측은 2018년에 고령사회(65이상 인구가 전체 14%)에 진입하고 2026년(20%이상)에는 초고령사회에 진입한다고 하였으나, Park & Kim은 2017년 고령사회, 2024년에 초고령사회에 진입한다고 예측하여 고령화 속도가 정부의 예측보다 빠르다고 하였다. 정부는 추계인구를 이용하여 예측을 하였고 Park & Kim은 추계인구보다 100여만 명이 적은 센서스인구를 이용하여 예측하였다. 이 결과를 정리하면 〈표 18〉과 같다.

〈표 18〉 Park & Kim(2009) 예측인구와 정부추계인구의 비교

(단위: 명)

	연 도	0~14세	15~64세	65세 이상	총 합
총조사인구(2005) 및 예측인구	2005	8,986,128	33,690,088	4,365,218	47,041,434
	2010	7,846,864	34,923,908	5,382,597	48,153,369
	2015	6,830,600	35,626,253	6,616,253	49,073,106
	2020	6,401,695	35,052,403	8,316,980	49,771,078
	2030	5,687,196	31,484,593	13,544,684	50,716,473
정부추계인구	연 도	0~14세	15~64세	65세 이상	총 합
	2005	9,151,077	34,584,511	4,444,602	48,180,190
	2010	7,906,908	35,610,778	5,356,853	48,874,539
	2015	6,732,801	36,163,474	6,380,819	49,277,094
	2020	6,118,161	35,506,403	7,701,125	49,325,689
	2030	5,525,336	31,298,528	11,810,707	48,634,571

※ 2005년 정부추계인구는 11월 1일 기준으로 환산함.

〈표 18〉에서 볼 수 있듯이 두 예측인구 모두 15~64세 인구는 2005부터 2030년까지 거의 유사한 것으로 나타났으나, 0~14세에서는 2010년까지는 Park & Kim(2009)의 예측치가 정부예측치보다 작다가 2015년부터는 커지는 현상을 보이고 있다. 그러므로 출산율의 관점에서 살펴볼 때, 정부의 경우가 고령화를 촉진하는 결과를 초래한다. 65세 이상의 인구를 살펴볼 때 2005년에는 정부의 예측치가 Park & Kim의 예측치보다 높으나, 그 이후에는 정부의 예측치가 Park & Kim 예측치보다 계속 낮음을 볼 수 있다. 이는 정부가 사망률에 대해 좀 더 보수적이며, 이로 인해 Park & Kim보다 사망자 수를 높게 추정하고 있다.

〈표 19〉 동태자료 기대수명과 통계청 추계 기대수명 비교

(단위: 세)

연 도	기대수명(남자)			기대수명(여자)		
	추계(A)	Postcensale stimates(B)	차 이 (B-A)	추계(A)	Postcensale stimates(B)	차 이 (B-A)
2006	75.3	75.7	0.4	82.1	82.4	0.3
2007	75.5	76.1	0.6	82.3	82.7	0.4

〈표 20〉 **인구변동요인별 추계치 및 실제치 비교**

(단위: 1,000명)

인 구	2005년	2006년	2007년	2008년
		인 구	인 구	인 구
2006년 추계(A)	48,138	48,297	48,456	48,607
postcensal estimates(B)	48,138	48,387	48,612	48,972
차 이 (B-A)	0	90	155	366
출 생	-	4	17	48
사 망	-	-3	-12	-24
국제이동	-	83	126	293

* : 2008년 국제이동 자료는 잠정치임

실제로 통계청(통계청 2009b)에 따르면 새롭게 관찰된 2006년과 2007년의 인구동태자료를 이용하여 기대수명을 계산했을 때, 〈표 19〉에서와 같이 통계청의 장래인구추계치보다 더 높은 값을 보이고 있다. 이는 장래추계인구에서 사망률을 과대 추정하였기 때문이다. 〈표 20〉 은 2006년부터 2008년까지의 인구변동 요인별 장래추계인구와 동태자료에 의한 조사후인구와의 차이에 대한 정부의 자료 중 일부이다.

〈표 20〉에 따르면 2006년부터 2008년까지 출생자 수는 장래추계보다 더 늘어나고 사망자 수는 장래추계보다 감소하며, 국제이동은 늘어난 것으로 나타난다. 이러한 차이를 일으킨 동태자료의 변동 때문에, 2006년에 작성한 장래추계인구의 중간점검을 2009년 5월에 시행하게 된다. 2006년부터 2008년의 인구동향을 반영한 장래추계인구를 새로 작성하여 이를 2006년 추계와 비교하였다.

그 결과 통계청에서는 "인구동향을 반영하여 작성한 인구는 추계인구의 흐름과 차이를 보이나 장래에 총인구가 크게 조정받을 것이라고 판단하기는 아직 이름"이라고 밝히고 있으나, 이는 2006년부터 2008년까지의 오직 세 시점의 동태자료만 추가하였을 뿐, 출생 및 사망에 대한 모형과 기본 가정을 그대로 설정했기 때문이다. 즉, 출생률 및 사망률 계산을 위한 모형 및 모수의 미래값에 대한 설정을 2006년 장래인구추계와 동일하게 한 상태에서, 세 시점의 자료추가는 출산율 모형인 로그감마모형의 모수추정치와 Lee-Carter의 사망률 모형의 모수추정치에서 무시해도 될 만큼의 미세한 변화만을 가지고 오게 되어, 새롭게 계산된 장래추계인구와 기존의 2006년 장래추계인구는 거의 동일할 수밖에 없기 때문이다. 그러나 〈표 20〉에서 알 수 있듯이 시간이 지남에 따라 2006년 추계치와 동태자료 관측치와

의 차이가 점점 더 커짐을 알 수 있으며, 이에 따라 인구추계에서 이용되는 모형 및 가정에 대한 재검토가 필요하다고 여겨진다.

특히, 사망률 예측의 경우 Lee-Carter 모형은 자료의 지나친 단순화로 사망률의 횡적·종적 상관관계를 반영하지 못하고 있으며(Park et al. 2006), 연령별 사망률이 시간이 지남에 따라 일정함을 가정했기 때문에(Lee & Carter 1992; Booth et al. 2001) 우리나라와 같이 연령별 사망률의 변화가 급격한 나라에서는 이용하기 적합하지 않다. 이러한 관점에서 볼 때 Park & Kim(2009)에 의한 사망율 모형이 더 현실적인 것으로 판단된다.

미국의 경우 Lee-Carter 모형을 개선한 Lee & Tuljapurkar(2000) 의 방법을 이용하고 있으며(Census Bureau 2000), 캐나다의 경우 전국 및 지역의 사망률을 추정하기 위한 Li & Lee(2005)의 방법을 이용하고 있다(Statistics Canada 2005).

출산율 예측의 경우, 통계청에서 이용하는 출산율 모형인 로그감마모형(Kaneko 2003)은 출산이 완결되지 않은 코호트에 대한 불확실성의 문제가 존재하며, 사망과 국제이동 등으로 인한 코호트 크기의 감소가 없다는 가정이 비현실적이다. 따라서 Lee(1993)의 출산율 모형 및 캐나다의 Pearson Type Ⅲ모형(Verma et al. 1996) 등 다른 여러 출산율 모형에 대한 검토 및 검증이 필요하며, 이는 우리나라 출산율 자료에 대한 모형의 적합도를 통해 판단되어야 할 것이다.

Ⅴ.결론 및 제언

우리나라 인구는 정의가 모호한 거주자 개념으로 되어 있어 공식인구 작성에 혼란을 야기하고 있다. 무엇보다도 먼저 인구의 포함범위(universe)가 명확하게 정의되어야 한다. 우리나라의 기본인구는 센서스를 바탕으로 재정립되어야 한다. 현재의 기본인구는 출처가 불분명하며 재현(reproducible)이 가능하지 않아 이론적 근거가 없으며, 무엇보다도 센서스를 이용했다는 증거를 어디에서도 찾을 수 없다. 총조사후 표본조사(post enumeration survey)를 강화하여 순누락률(net undercoverage rate)의 측정을 정밀화하거나, 미국과 같이 지방행정단위에서 센서스 결과에 이의를 제기하게 하는 방법을 도입할 필요가 있다. 최근에 가장 중요한 인구변동요인으로 떠오르는 국제이동을 측정하는 센서스 항목을 추가하고 표본조사를 고려해야 한다. 조사후인구(postcensal estimate)는 현재의 장래추계인구 대신에 사용되어야 하며 장래추계인구는 장기적인 관점에서만 사용되어야 한다. 조사후인구를 사용함으로써 울타리 오차(error of closure)를 정상적인 방법으로 측정할 수 있게 되며, 이에

따라 조사간인구(intercensal estimate)의 울타리 오차의 배분방식을 개선할 수 있게 되어 신뢰도를 높일 수 있게 된다. 지역별 기준인구의 신뢰성을 높이기 위해 전국단위 기준인구와 독립적으로 지역별, 성별, 연령별 인구를 산출하고 전국단위 인구와 내적 일치성(internal consistency)을 유지할 수 있도록 조정해야 한다. 이에 따라 전국단위뿐만 아니라 지역단위별로 출생률, 사망률, 실업률, 국민소득 등의 사회지표와 경제지표의 분모를 통일해서 사용하여야 한다. 공식인구에 대한 정의의 작성방법은 체계적이어야 하며 수식화가 필요하며 특히 전국단위 추정치와 지역단위별 추정치의 일치성 유지를 위한 조정문제를 공식화하여야 한다. 장래추계인구 모형에 사용한 출산율 모형, 사망률 모형, 그리고 국제이동 모형의 타당성이 검증되어야 한다.

현재의 통계청의 조직 및 인적자원은 이러한 문제를 해결하는 데 근본적인 문제를 가지고 있으므로 인구관련 전담 연구부서의 신설과 동시에 인구관련 학자들의 활발한 연구 참여가 선행되어야 할 것이다.

참고 문헌

이 건. 2009. “행정등록 기반 센서스의 전망과 과제-방법론적 접근.” ≪통계연구≫ 14(1): 1-29.

양경진. 2008 ≪인구센서스 방법론 연구≫. 통계개발원.

통계교육원. 2008. ≪인구통계 기초와 응용≫. 통계교육원.

통계청. 2009a. ≪인구동향조사≫. 통계청.

통계청. 2009b. ≪장래추계 중간점검≫. 통계청.

통계청. 2006. ≪장래인구추계≫. 통계청.

Booth, H., Maindonald, R., Smith, L. 2002. “Applying Lee-Carter under Conditions of Variable Mortality Decline.” *Population Studies* 56: 325-336.

Brass, W. 1971. “On the Scale of Mortality.” In W. Brass(ed.), *Biological Aspects of Demography*. London: Taylor and Francis.

Bryan, T. 2004. “Population Estimates.” In Siegel and Swanson(eds), *The Methods and Materials of Demography*. San diego: Elsevier Academic Press.

Census Buerau. 2009a. *Methodology for United State Resident Population Estimates by Age, Sex, Race and Hispanic Origin(Vintage 2009)*: April 1, 2000 to July 1, 2009. Washington, D.C: U.S. Census of Bureau.

Census Bureau. 2009b. *Methodology for the State and County Total Resident Population Estimates(Vintage 2009)*: April 1, 2000 to July 1, 2009. Washington, D.C: U.S.

Census of Bureau.

Census Bureau. 2008. *Methodology for the State and County Resident Population Estimates by Age, Sex, Race, and Hispanic Origin(Vintage 2008)*: April 1 to July 1, 2008. Washington, D.C: U.S. Census of Bureau.

Census Bureau. 2005. *Methodology*. Washington, D.C: U.S. Census of Bureau.

Census Bureau. 2004. *Coverage Measurement from the Perspective of March 2001 Accuracy and Coverage Evaluation*. Washington, D.C.: U.S. Census of Bureau

Census Bureau. 2001. *CENSUS2000 INFORMATIONAL MEMORANDUM NO. 100*. Washington, D.C: U.S. Census of Bureau.

Census Bureau. 2000. *Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100*. Washington, D.C: U.S. Census of Bureau.

Goerge, M. V. and Perreault, J. 2002. "Methods of External Migration Projections and Forecasts." In N.Kelman & H.Cruijsen(eds.), *National Population Forecasting in Industrialized Countries*. Amsterdam: Swets and Zeitlinger.

Japan Statistics Bureau. (available from http://www.stat.go.jp/english/data/jinsui/1.htm)

KaneKo, R. 2003. "Elaboration of the Coale-McNeil Nuptiality Model as the Generalized Log Gamma Distribution: A New Identity and Empirical Enhancements." *Demographic Research* 9: 223-262.

Lee, R. D. and L. Carter. 1992. "Modeling and Forecasting U. S. mortality." *Journal of American Statistical Associate* 87: 659-671.

Lee, R. D., and Tuljapurkar, S. 2000. Population Forecasting for Fiscal Planning: Issues and Innovations. In A. Auerbach & R. Lee(eds.), *Demography and Fiscal Policy*. Cambridge: Cambridge University Press.

Lee, R. D. 1993. "Modeling and Forecasting the Time Series of US Fertility: Age Distribution, Range, and Ultimate Level." *International Journal of Forecasting* 9 : 187-202.

Li, N. and R. Lee. 2005. "Coherent Mortality Forecasts for a Group of Populations." *Demography* 42: 575-594.

Park , Y. and Kim, S. 2009. "Adjustment of Sex-Cause-Age Specific Deaths and Forecast of Sex-Age Specific Population Structure." Unpublished.

Park, Y., Choi, J. W., and Lee, D. H. 2006. "A Parametric Approach for Measuring the Effect of the 10th Revision of the International Classification of Diseases." *Journal of the Royal Statistical Society Series C - Applied Statistics* 55: 677-697.

Preston, S. Heuveline and P. Guillot, M. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell Publishing.

Siegel, J. S. and Swanson D. A(eds.) 2004. *The Methods and Materials of Demography*. San Diego: Elsevier Academic Press.

Smith, S. K., Tayman, J. and Swanson, D. 2001. *State and Local Population Projection: Methodology and Analysis*. New York: Kluwer Academic/Plenum Publisher.

Statistics Canada. 2007. *Population and Family Estimation Methods at Statistics Canada*. Ottawa: Statistics Canada.

Statistics Canada. 2005. *Population Projections for Canada, Provinces and Territories* 2005-2031. Ottawa: Statistics Canada.

Verma, R.B.P., Loh, S., Dai, S.Y. and Ford, D. 1996. *Fertility Projections for Canada, Provinces and Territories*, 1993-2016. Ottawa: Statistics Canada.

[접수 2010/4/19, 1차수정 2010/5/25, 2차수정 2010/6/20, 게재확정 2010/6/21]