

# 컴퓨터 시뮬레이션에 의한 경제인구 예측 통계 모형에 관한 연구

## A Study on the Estimation of Economic Population Statistical Model by Computer Simulation.

정 관 회(Kwan-Hee Chung)<sup>1)</sup>

### 요 약

본 논문에서는 컴퓨터 시뮬레이션에 의한 인구예측을 통계모형을 써서 연구하였고 더불어 경제인구를 예측하였다. 과거의 인구를 토대로 하여 미래의 인구를 예측한다는 것은 불확실한 상황이 많이 개입되어 있기 때문에 매우 어려운 문제이다. 또한 예측이 되었다 하더라도 급변하는 세계적인 문화 및 국내의 문화적인 정서의 흐름에 따라서 많은 변화가 예상되므로 경제인구 예측을 적중하기에는 더 더욱 어려운 것이다. 인구 예측에 있어서 과거의 자료인즉, 1960년도부터 1990년도까지 센서스 인구를 이용하여 Box & Jenkins가 개발한 ARIMA 모형을 써서 미래 2021년도까지의 인구를 각각 표나 부록에 나타난 것처럼 경제인구를 예측하였다.

### Abstract

In this study, the economic population prediction by computer simulation has been studied by using statistical model. The forecast of future population based on that of the past is a very difficult problem as uncertain conditions are modeled in it. Even if a thought forecast is possible, world-wide cultures and the local culture emotion the cultures of the world and out country can not be predicted due to rapid change and the estimation of population is 'nowadays more and more' difficult to be made good guess. In the estimation of economic population, by using the census population from 1960 to 1990, and using ARIMA model developed by Box and Jenkins, the estimation has been done on the economic population until 2021 according to age as appeared table and appendix. This kind of forecast would have both good point and weak point of ARIMA model theory saying that prediction can be done only by the economic population.

논문접수 : 2003. 12. 11.

1) 정희원 : 경기대학교 정보과학부 교수

심사완료 : 2003. 12. 22.

\* 본 연구는 2003학년도 경기대학교 학술 연구비 지원에 의하여 수행 되었음

## 1. 서론

본 논문은 컴퓨터 시뮬레이션에 의한 경제인구 예측 통계 모형에 관한 연구이다. 경제인구란 곧 노동인구이며, 경제활동인구를 말한다. 경제인구를 예측하는 연구는 비교적 최근의 연구 분야이며, 인구 통계 사용법 역시 최근에 들어서 발전되었다. 그 방법으로는 컴포넌트 추정법 및 ARIMA 추정법 등이 있다. 인구(population)를 수량적으로 묘사한 것은 '인구 통계(demography)'로 여겨지는 경우가 가끔 있다. 인구는 일정한 종류의 통계로 표시되는 인간의 집합체로 간주된다. 해마다 죽는 사람이 있는가 하면, 출생하는 사람도 있다. 그 뿐만 아니라 인구 이동(migration)으로 실질적 증감이 일어날 수도 있다. 이런 자료는 일정한 종류의 사상(출생, 사망, 결혼 및 이혼)이 발생할 때마다 모두 기록 될 수 있게 고안된 일정한 양식의 등록에 의해서 작성된다.

마지막으로는 과거 인구를 가지고 미래인구를 추정(estimation)하는 것이다. 오늘날은 컴퓨터와 소프트웨어가 발달하여 이 방법으로 추정한다.

첫째, 예를 들어 출생률, 사망률, 정규적인 직업을 가진 사람들의 퍼센트 혹은 기혼자의 구성비와 같은 인구의 몇 가지 행태의 실적 수준은 어떠한 상태에 있는지? 둘째, 어떤 관찰 기간 중에 이런 실적수준을 변동하는 것인지? 셋째, 인구 내부구조의 변화 행태가 있는지? 이 질문에 대한 해답은 연령, 성, 인종, 직업 및 교육별 계층 기타 등과 같은 상이한 집단에 대한 수량을 비교함으로써 알아낼 수 있다.

### 1) 근로 활동인구

거의 모든 국가에서 국민의 취업활동의 분포에 관한 자료가 필요하다. 근로인구의 규모 및 구성에 관한 통계는 사회적 및 경제적 생활상태가 포함되어 있는 것이며 그렇기 때문에 이러한 계획은 국가의 인력자원 목록을 필요로 한다. 경제활동에 관한 좋은 통계를 얻는다는 것은 커다란 과업이다. 그것은 아마도 현대의 '센서스'중 가장 귀찮은 일면일 것이다. 이러한

이유로 센서스 자료가 더 좋다. '센서스'는 두 가지 방법으로 인력 분석에 기여한다.

### 2) 경제활동 인구의 정의와 경제인구

어떤 사람이 경제 활동적(economically active)이고 또 어떤 사람들이 그렇지 않은가를 결정하는 방안을 설정하기란 쉽지 않다. 이는 생산적인 일(productive work)이 어떠한 활동으로 구성되는가를 판단할 수 있는 기준과 또 어느 정도 일을 해야만 어떤 사람을 활동적(active)이라고 분류하기에 적절한가를 판단할 수 있는 어떤 일관성 있는 기준을 의미하는 것이다. 대다수 사람들은 위의 두 가지 범주 중 하나에 들어가는 것이 아마 틀림없을 것이다. 15세부터 64세까지의 연령을 가진 사람들의 수는 경제활동인구의 대다수를 공급하는 집단으로 취급하는 것이 관례이며 이를 근로연령인구(population of working ages)라고 부르며, 이것은 경제 연령인구(population of economic ages)라고 지칭한다.

### 3) 연구방법

경제인구 즉 경제활동인구를 예측한다고 하는 것은 어려운 문제이다. 본 연구 모델의 이론상 기법은 ARIMA(autoregressive integrated moving average : 자기회귀 누적이동평균) 모형을 세워서 경제인구 추정(estimation)하였다.

프로그램은 SAS/ETS 패키지 프로그램 내에 있다. 특히 ARIMA모형의 기본자료는 남한의 경우 과거 31년간 인구(1960~1990년)를 관찰값으로 삼아서 미래의 31년간 경제인구 즉 경제인구(15세부터 64세까지)별로 예측하였다.

## 2. 관련연구

남북 통일을 대비한 경제인구(economic population)를 추측하기 위해서는 우선 먼저 사회지표(R. A. Bauer 교수가 처음으로 사용한 용어)의 인구(유아인구, 경제인구, 노령인구)에 대하여 인식하고 다음으로 우리 나라 전체의 과거, 현재 그리고 미래에 대한 총인구, 출생

를, 사망률, 이민율, 자연증가율 인류사회의 인구 성장률 등을 예측하여야 한다. 특히 인구통계는 정치, 경제, 사회, 그리고 문화 등의 척도에 기초가 되어 왔다. 오늘날 인구의 정태통계는 민주정치에서 국민의 심부름꾼을 선출할 기초가 되며, 또 근세의 부과, 장기 경제발전계획, 장기 교육발전정책계획, 사회 간접시설계획, 국민의 생활규제계획, 사회의 후생복지, 사회보장의 사회보험, 국민경제의 장. 단기 계획 등은 통계자료 즉, 인구 통계가 항상 우선 되어야 한다. 대한민국의 총 국토면적(남북한)은 220,951.08km<sup>2</sup>이며 그중 남한의 면적은 98,477.48km<sup>2</sup>이다. 2000년도 남한의 인구는 44,256,967명으로 조사되었으며 경제인구는 32,730,624명으로 예측된다.

1) 선진국들의 총인구 및 성장률

선진국들 중에 사천만 인구가 넘는 나라는 한국, 일본, 프랑스, 영국, 독일, 이탈리아, 미국 등이 있다. 그 중에 일억이 넘는 나라는 일본과 미국이며 자체인구가 일억은 넘어야 성장한다는 것을 알 수 있다 인구 성장률은 1.1이 넘지 않는 한 현존인구를 유지하기 어렵다. 결론적으로 인구가 있어야 경제인구가 많고, 경제인구가 많아야 경제활동인구가 많고, 그렇게 함으로서 GNP가 올라갈 수 있다.

표 (2-1) 선진국들의 총인구 성장률, GNP, 경제활동인구, 실업률 근로시간

항목 국명	총인구 (성장률)	GNP (\$)	경제활동 인구	실업 률 (%)	주 근로 시간 수
한국	44,606(0.9)	10,548	20,797	2.4	48.7
일본	123,611(0.3)	34,630	66,140	2.9	37.7
중국	1,160,045(0.3)	530	X	2.8	X
프랑스	56,634(0.5)	23,470	25,870	12.5	38.6
영국	56,352(0.4)	18,410	X	X	X
독일	61,077(0.7)	25,580	32,580	9.2	38.0
이탈리아	59,104(-0.2)	19,270	22,680	11.3	X
스웨덴	8,587(0.6)	23,630	4,266	8.0	37.8
미국	248,710(1.1)	25,860	131,056	6.1	42.0
캐나다	27,297(2.4)	19,570	14,832	10.4	38.6
호주	16,851(1.1)	17,980	8,776	9.8	38.7

선진국들 중에 GNP가 표(2-1)에서 보듯이 이만 달러가 넘는 나라는 일본, 프랑스, 독일, 스웨덴 그리고 미국이 있으며, 일본의 경우는 삼만 달러가 훨씬 넘는다 만 달러 범주에 있는 나라는 영국, 이탈리아, 캐나다 그리고 호주이다. 우리 나라는 요즈음에 일만 달러를 넘고 있으며, 중국은 아직도 일 천 달러에 달하고 있다. 경제활동인구는 곧 경제인구인데 표(2-1)에서 보듯이 이천 만 명 이상인 나라는 한국, 일본, 프랑스, 독일(구 서독), 이탈리아, 미국 등의 나라이고, 영국의 경우는 자료가 입력되지 않았다. 이중 일본은 66,140천명이며, 미국은 131,056천명이다. 호주나 캐나다는 각각 8,776천명과 4,266천명이다.

2) 컴포넌트에 의한 인구예측모형

인구예측방식은 인구정태 통계(census data)와 인구 동태 통계(출산율, 사망률)등을 토대로 하여 인구증감요인에 의한 예측 즉, 조성방법(component method)에 의한 인구통계가 정부관청을 비롯한 일반연구원에서 많이 이용된다. 따라서 본 항에서는 1990년도에 통계청이 발표한 컴포넌트 방법에 의해서 2021년까지 PROJ-5 P패키지에 의해서 장래인구를 예측한 것을 표(2-2)에 수록하였다. 따라서,

$$\text{인구의 총수} = \text{성} \cdot \text{년령별 기준인구} + \text{출생수} + (t+1)\text{년도 0세 인구수} + \text{유입인구수} - \text{유출인구수} - \text{사망자수}$$

라는 것이 성립한다. 표 (2-2)는 교육인구, 경제인구, 노령인구수를 구분하기 위하여 각 세별로 표시하지 않았다. 1990년도가 기준인구이며, 그 이후로 예측인구이다. 표 2-2에서 나타나 듯이 우리 나라의 장래 경제인구는 감소추세를 보이므로 경제성장의 어두운 면을 보이고, 노령인구수는 증가하므로 정부의 예산이 증가할 것이므로 여기에 관한 정책을 세워야 될 것이다. 따라서 노동인구 즉, 경제활동인구를 증가시키려면 하루빨리 남북통일이 성취되어야 한다는 절박한 상황이 예측된다.

표 2-2 컴포넌트 방법에 의한 장래인구 예측

년형 년도	0~14 세	15세~ 17세	18세~ 59세	60세~ 65세	66세~	합 계
1990	11,07 7,251	2,626, 717	25,865 ,363	1,350, 857	1,949,095	42,869 ,283
1995	10,39 9,800	2,401, 460	28,012 ,640	1,730, 394	2,306,507	44,850 ,801
2000	9,916, 572	2,120, 610	29,768 ,268	2,131, 242	2,852,682	46,789 ,374
2005	9,841, 461	1,948, 195	30,774 ,508	2,268, 044	3,601,894	48,434 ,102
2010	9,510, 144	1,974, 079	31,372 ,908	2,502, 984	4,323,318	49,683 ,433
2015	8,790, 139	1,993, 556	31,554 ,608	3,022, 727	4,985,354	50,346 ,444
2021	7,988, 695	1,840, 536	30,503 ,327	4,262, 272	5,990,844	50,578 ,196

인구예측 시뮬레이션을 하기 위한 수리식은 과거 인구 총수의 자료를 이용하여 다음 식들에 의하여 추세를 도출한 후 장래인구를 추정하는 방법이다. 추세선 도출방법직선식은 식(2. 1)이고, 곡선식은 (2, 2), (2, 3), (2, 4)등이 있다.

$$p_t = p_0(1 + r + t) \quad (2. 1)$$

$$p_t = p_0(1 + r)^t \quad (2. 2)$$

$$p_t = p_0 \times e^{rt} \quad (2. 3)$$

$$p_t = k / (1 + \exp(a^* \beta^t)) \quad (2. 4)$$

식 (2, 1)은 직선방정식이고, 식 (2, 2)는 복리방정식이며, 식 (2, 3)은 대수 방정식이고 그리고 식(2, 4)는 로지스틱 방정식이다. 이 수리방정식들의 기초자료는 년도별, 시도별인구와 행정구역별 자료는 제공되나 각 세별(연령별) 인구자료 작성이 어렵다는 단점이 있다.

### 3. ARIMA모형에 의한 경제 인구 예측

시간의 흐름에 따라 변화는 현상을 관찰한 결과를 나타내는 수치적 함수를 변수 또는 변량이라 한다. 이 변수가 일정기간 동안 변화할

때 동일 간격의 각 시정 (t)에서 관찰 전 값들의 계열을 시계열이라 한다. 예측 방법에는 시계열 모형(time series model)과 회귀 모형(regression model)두 가지가 있다.

첫째 시계열 모형이란 과거의 관찰 값들이나 오차 등을 이용하여 과거 시계열에 존재하는 계열의 형태를 파악하고, 이것을 이용하여 미래에 발생하는 시계열의 형태를 예측하는 모형이다. 둘째 회귀 모형은 예측 (Prediction)하고자 하는 요인(종속변수)이 하나 또는 여러 개의 독립변수와 원인과 결과(cause-effect relation)를 갖는다는 가정 하에서 자료를 이용하여 종속 변수의 미래 값을 예측하는 모형이다.

본 절에서의 연구를 시계열 분석 (time series analysis)에 경제인구 예측을 시뮬레이트(simulate)함에 있다. 예측 방법에는 질적인 예측 방법과 양적인 예측 방법이 있는데, 이 방법은 다음 세 가지 조건이 존재함을 가정하고 예측을 하는 방법이다.

- ① 과거에 대한 정보를 가지고 있다.
- ② 과거의 정보는 자료의 형태로 변량화 시킬 수 있다.
- ③ 과거 정보의 패턴이 미래에도 같은 형태로 유지한다.

이런 조건하에서 통계적예측 모형들 중에서 ARIMA(autoregressive integrated moving average)모형을 이용한다. 시계열 분석의 대표자격인 BOX & JENKINS 의 방법에서는 비정상시계열은 차분을 이용해서 ARIMA모형을 적용시킬 수 있는 정상시계열로 변환시켜 사용한다. P차 AR부분과 q차 MA부분의 정상시계열 혼합모형 ARIMA(p, q)는 다음 식 (3, 1)과 같다.

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \dots - \theta_q \alpha_{t-q} \quad (3, 1)$$

여기서, AR(p) 모형은 식(3, 2)와 같다.

$$Z_t = \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + \alpha_t \quad (3, 2)$$

여기서

$$a_t = \varphi_p(B)Z_t \quad (3, 3)$$

$$\varphi_p(B) = (1 - \varphi_1 B - \dots - \varphi_p B^p) \quad \text{이고,}$$

$a_t$ 는 평균이 0이고 분산이  $\delta_a^2$ 이며 서로간에 독립인 백색잡음(white noise)이다.

차분을 이용하면 비정상 시계열(nonstational time series)을 정상 시계열(stational time series)로 변환시켜야 할 경우가 많다. Box & Jenkins 방법에서는 비정상 시계열차분을 이용해서 ARIMA 정상시계열로 변환시켜 분석한다.

### 1) ARIMA (p, d, q) 모형

원 시계열  $Z_t$  의  $d$ 차 차분시계열을

$$W_t = \nabla^d Z_t$$

라하자. 여기서  $Z_t$ 가 ARIMA(p, d, q) 과정을 하면 이 시계열은 차수 p, d, q인 ARIMA 즉, 자기회귀루적이동 평균 모형을 갖는다고 한다. 이때 차수는 2차 이하로 하여, 보통 d=0, 1, 2차수에서 하나를 식별 선택한다.

ARIMA 모형식은

$$W_t = \nabla^d Z_t = (1 - B)^d Z_t \quad \text{일 때,}$$

$$\Phi(B)W_t = \theta(B)a_t \quad (3, 4)$$

로 표시되고, 일반식은

$$\Phi(B) \nabla^d Z_t = \theta(B)a_t \quad \text{이다.}$$

여기서는

$$\Phi(B) : 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\theta(B) : 1 - \theta_1 B - \dots - \theta_q B^q$$

$$\nabla : 1 - B$$

$$a_t = \theta^{-1}(B)\Phi(B)(1-B)^d Z_t$$

이다.  $a_t$ 는 평균이 0 이고 분산이  $\sigma_a$ 의 백색잡음(white noise)이다.

누적부분의 특정식은

$$\Phi(B) = \theta(B)(1 - B)^d \quad (3, 5)$$

로 된다.

#### ① 모형의 식별 방법

이론적 모형식별 통계량인 ACF, PACF, IACF와 똑같은 모형을 나타내고 있는 경우는 드물게 나타난다. 따라서 주어진 시계열에 대해 여러 개의 유사한 모형이 가능할 수 있지만 적은 수의 모수를 가진 시계열 모형을 설

정하는 것이 바람직하다. 더 자세한 것은 참고 문헌[5]의 P. 90~P. 91에 나타나 있다.

모형의 식별은 실제 자료를 가지고 SAS/ETS를 이용하여 ARIMA 모형을 식별한 후에 모수를 추정하고, 추정이 부적합하면 다시 되돌아가서 모형을 식별하고 모수추정 과정을 반복한다. 이런 경우에 추정된 값이 검진을 거쳐서 실행된다. 이 과정을 그림 (3-1)에 나타나 있다.

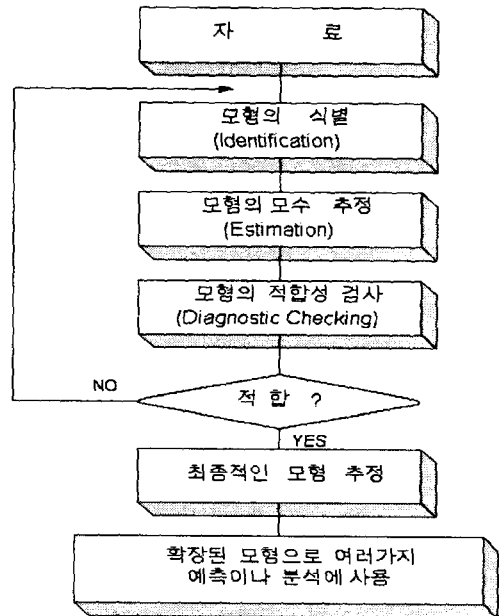


그림 (3-1) Box & Jenkins 의 ARIMA 모형 선택방법

\* 그림 자료: 자유아카데미,

SAS/ETS-시계열분석(I) P85.

#### ② ARIMA 모형의 통계량

ARIMA 모형을 식별하기 위한 통계량은 여러 가지 방법이 있다. 최후 예측 오차함수(final prediction error function:FRE)는 1970년도 Akaike에 의해서 제시되었으며, 아카이케 정보 판단 기준(Akaike's information criterion function:AIC)은 1972 연도에 Akaike에 의해서 제안되었으며, 자기회귀변동 판단 기준 함수(criterion of autoregressive transfer function:CAT)는 1974년도 Parzen에 의해서 ARIM

A 모형은 유한차수의 AR 모형으로 근사 시킬 수 있다고 가정하고 제안했고, 베이지안 정보 판단 기준(Bayesian information criterion: SBIC, BIC)은 1975년도에 Schwarz에 의해서 제안되었으며, 반복대수법칙(law of iterated logarithm)은 1979년도에 Hannon-Quinn에 의해서 판단 기준을 제시하였다. 그 각각의 함수들은 다음과 같다. 따라서 분산이 보통  $W \cdot N$ 이다\*.

$$FPE(p, q) = \hat{\sigma}_a^2(p, q) \left[ 1 + \frac{2(p+q)}{n} \right] \quad (3, 6)$$

$$AIC(p, q) = \ln \hat{\sigma}_a^2(p, q) + \frac{2(p+q)}{n} \quad (3, 7)$$

$$CAT(p) = -\left(1 + \frac{1}{n}\right), \quad p=0 \quad (3, 8)$$

$$\frac{1}{n} \sum_{j=1}^p \frac{1}{\hat{\sigma}_a^2(j)} - \frac{1}{\hat{\sigma}_a^2(p)}, \quad p=1, 2, \dots$$

$$SBC(p, q) = \ln \hat{\sigma}_a^2(p, q) + \frac{\ln(n)}{n} (p, q) \quad (3, 9)$$

$$HQ(p, q) = \ln \hat{\sigma}_a^2 + \frac{2 \ln \ln N}{N} (p+q) \quad (3, 10)$$

이상의 식들에서 통계추정량들을 결정한다.

### ③ 모형의 검진과 모형의 변형

식별하고, 추정하고, 그리고 검진을 반복하여서 「Box-Jenkins」의 시계열분석의 ARIMA모형을 선택한다. 실시 자료 즉, 60년도부터 90년도까지 우리 나라 기준 인구를 SAS/ETS를

이용하여 추정한다.

### ④ ARIMA 모형에 의한 예측

ARIMA 모형을 사용하는 예측(forecasting)은 여러 가지 방법으로 표현될 수 있다. 각 표현 방법은 계산(computing), 최신화(updating), 정도화(pricision), 장기예측형태(long-term-forecasting behavior) 하는데 목적이 있다. 실시 자료를 이용하여 1991년부터 2021년까지 18세부터 59세(이상)까지 인구를 예측한 다음 18세부터 60세 미만까지의 총인구를 모형에 의하여 예측한다.

### 2) ARIMA 모형에 의한 인구예측의 실증적 분석

본 절에서는 사례중심으로 15세부터 64세(이상)까지 과거 31년(1960~1990년)간 자료에 의해서 장래 31년(1991~2021년)간 인구예측을 5세 간격으로 ARIMA(p, d, q) 모형에 의하여 예측하였으며, 적합 모형은 부록표 ARIMA(p, d, q) 모형에 의한 인구예측에 나타나있다. 5년 간격사이의 예측은 증감에 대한 차에 대한 수치를 평균하여 균등 배분하였으며, 자세한 수치는 참고문헌[5] 및 부록 표를 인용하였다.

#### ① 자료구성 및 처리방법

기본자료는 과거 인구를 남·여 각 세별로 구분해서 구성하고 SAS/ETS의 PROCEDURE ARIMA로 처리하였다. 일반적으로 Box & Jenkins형의 모형 가정에서는 그림 (3-2)를 거쳐서 시행한다. 표 (3-1)에서와 같이 ARIMA모형을 두개를 구축하였다. 모형에 의해서 모든 인구실증적분석을 실었다.

#### ② 인구예측의 실증적 분석

여기서 남한의 경제인구, 14세(남, 여)부터 64세(남, 여)는 참고문헌[5] 및 부록 표를 인용하였으며, 예측 과정을 실었으며, 컴퓨터에서 출력된 자료이므로 그림으로 표시한 것도 있다.

표 (3-1) 5세 간격. ARIMA에 의한 적합성 모형

연령	ARIMA(p,d,q) 모형	
	남 자	여 자
15	ARIMA(2,0,4)	ARIMA(3,0,4)
18	ARIMA(3,0,2)	ARIMA(3,0,3)
23	ARIMA(1,0,4)	ARIMA(1,0,3)
28	ARIMA(3,0,3)	ARIMA(2,0,0)
33	ARIMA(1,0,3)	ARIMA(2,0,2)
38	ARIMA(1,1,3)	ARIMA(0,1,1)
43	ARIMA(0,1,2)	ARIMA(0,1,1)
48	ARIMA(1,1,0)	ARIMA(0,1,1)
53	ARIMA(3,1,0)	ARIMA(2,2,2)
58	ARIMA(2,2,4)	ARIMA(3,2,4)

③ 18세 남자 모형(I)의 식별 단계

18세 남자 모형선택 p = 3, d = 0, q = 2의 차수가 가장 접근한다고 생각된다. ARIMA(3, 0, 2) 을 찾기까지는 다음과 같이 진행되었다. AR 차수인 p = 0, 1, 2, 3 까지 변화시키고, 차분의 차수인 d = 0, 1, 2 까지 변화시켰으며, MA 차수인 q = 0, 1, 2, 3, 4 까지 변화시켰다.

모두 50회 반복한 결과, 백색잡음(White noise)을 위한 체크(Check) 자기상관에 관련된 값들(x<sup>2</sup>값, 확률값)이 같은 군을 찾았다. 18세 남자에서 같은 군의 ARIMA 모형(I)은 다음과 같이 나타난다. 이들 중에서 AIC값과 SBC 값이 낮으면서 추정을 위한 상관계수값은 높고, 편 자기 상관값의 확률값이 낮으면서, Lag의 확률값이 0.05보다 클 때 ARIMA모형으로 선택될 수 있다.

④ 18세 남자의 최종적인 ARIMA 모형

따라서 ARIMA(p, d, q) 모형의 일반식을 ARIMA(p, d, q) = φ(B)∇<sup>d</sup>Z<sub>t</sub> = θ(B)a<sub>t</sub>  
= φ(B)(1-B)<sup>d</sup>Z<sub>t</sub> = θ(B)a<sub>t</sub>이다. 다음과 같이 출력되었다.

18세 남자 ARIMA(3, 0, 2) 모형

MODEL FOR VARIABLE MALE
ESTIMATED MEAN = 340929
AUTOREGRESSIVE FACTORS
FACTOR 1
1-1.3151B**(1)-0.1141B**(2)+0.457761B**(3)
MOVING AVERAGE FACTORS
FACOTR 1
1+0.3153378B**(1)-.684496B**(2)

⑤ 18세 남자 ARIMA(3, 0, 2)의 모형에 의한 예측 값

각 나이의 성별로 31개의 과거 자료가 되어 입력되어 처리된 관례로 일련번호가 32부터 출력되었으므로 다시 추정 연도로 바꾸어서 작성하면, 32(1991년), 33(1992년), 34(1993년), 35(1994년), ..... , 61(2020년), 62(2021년)이 된다. 특히 18세 · 59세 남녀 인구 예측을 표 (3-2)에 실었다.

10:15 MONDAY, APRIL 13, 1992  
SEX=MALE AGE=18  
P=3 : D=0 : Q=2

FORECASTS FOR VARIABLE MALE

표 (3-2) 18세 남자 ARIMA(3, 0, 2) 추정인구

OBS	FORECAST	STD ERROR	LOWER 95%	UPPER 95%
-----FORECAST BEGINS-----				
32	461770	11994.9	438260	485279
35	434739	36714.9	362780	506699
40	383359	60377.2	265022	501696
45	344256	70600.2	205883	482630
50	324326	73095.6	181061	467590
55	320403	73250.6	176835	463971
60	325511	73428.5	181594	469428
61	327012	73496.6	182961	471062
62	328567	73569.3	184374	472760

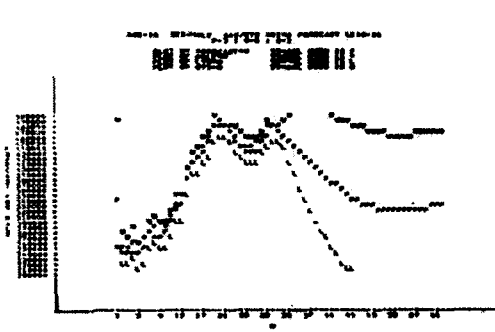


그림 (3-2) 18세 남자 인구 예측 그래프

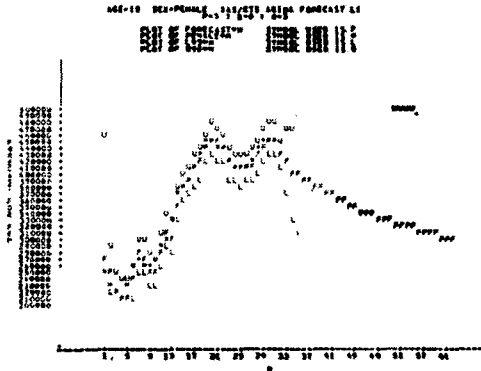


그림 (3-4) 18세 여자 인구 예측 그래프

⑥ 18세 여자의 최종적인 ARIMA모형

```

SEX = FEMALE      AGE = 19
P = 3 : D = 0 : Q = 3

MODEL FOR VARIABLE FEMALE
ESTIMATED MEAN      =      267583
AUTOREGRESSIVE FACTORS
                                FACTOR 1
1-1.45433B**(1)+1.02823B**(2)-.521849B**(3)
MOVING AVERAGE FACTORS
                                FACOTR 1
1+0.510084B**(1)+0.387881B**(2)+0.877746B**(3)
    
```

그림 (3-3) 18세 여자 ARIMA(3, 0, 3)모형

⑦ 18세 여자 ARIMA(3, 0, 3)의 모형에 의한 예측 값

각 나이의 성별(여자)의 과거 31년간 자료가 처리되었으며 부록표 ARIMA(p, d, q) 모형에 의한 인구예측에 나타나 있다. d값이 0 이므로 이 모형은 ARIMA모형으로서 점차적으로 인구가 감소 현상을 보인다. 그림 (3-4)으로 그려져 있다.

⑧ 경제 인구 예측 분석

경제인구의 성장은 그 나라 경제 활동인구의 성장을 뜻하며, 이것은 곧 근로 활동인구를 나타내는 것이다. 예컨대, “노동력(labor force)”으로 정의 할 수 있다. 노동 공급을 결정하는 인구 중에서 “경제 활동적”인 부분은 대부분의 국가에 있어서 매우 비슷하다. 전일제 직업도 성인이나 중간층에 집중되어 있다. 아주 어리거나 아주 늙은 사람들은 정규적인 직업에 종사하지 않고, 다른 사람들의 부양에 의존한다. 이 양극단의 사이에 있는 사람들이 노동의 공급을 담당하는 집단인 것이다. 15세부터 64세까지의 연령을 가진 사람들의 수는 경제 활동인구의 대다수를 공급하는 집단으로 취급하는 것이 관례이며, 이를 “근로 연령인구(population of working ages)” 즉, “경제 연령인구(population of economic ages)”라고 부른다. 1991년 남한의 경제인구는 29,805,741명이었고, 1995년은 30,308,979명이었고, 2000년은 31,021,551명이었고, 2005년은 31,816,419명이었고, 2010년은 32,730,624명이었고, 2015년은 33,779,679명이었고, 2021년은 35,221,624명이다. 미래 21년간(1991년부터 2021년까지) 경제인구 증가는 5,415,883명 증가 추세를 보인다.



#### 4. 결 론

끝으로 남북 통일이 되었을 시, 경제활동인구 즉, 경제인구는 1995년도에는 45,892,550명이고, 2000년도에는 47,840,095명이며, 2005년도에는 50,012,932명이며, 2010년도에는 52,471,120명이고, 2015년도에는 54,457,717명이며, 2020년도에는 57,164,288명이다.

여기서 2020년의 남북 합계 인구 중 남한 2021년 경제인구를 합한 값이 된다. 남·북 통일이 된다는 가정 하에서의 노동력은 대단한 노력을 발휘 할 수 있다고 생각되어 질 수도 있다. 예컨대, 우리의 노동력 인구는 구라파의 영국, 불란서 등의 총 인구와 거의 비슷한 정도이다. 여기서 북한의 경제인구 예측은 보건복지포럼 97/4를 인용하였음을 밝혀둔다.

#### 참 고 문 헌

[1] 김 재연 (1987), 컴퓨터 시뮬레이션, 박영사.  
 [2] 김 은정 외 1인 (1990), 시계열분석입문, 자유아카데미.  
 [3] 박 규상 (1972), 인구문제와 인구정책, 한얼문구.  
 [4] 정 관희 (1991), 컴퓨터 시뮬레이션에 대한 인구예측통계모형에 관한 연구, 동국대학원.  
 [5] Banks Jerry (1984), Discrete-Event System Simulation, Prentice-Hall.  
 [6] Fishman S. George (1978), Principles of Discrete Event Simulation, N. Y. John wiley & Sons.  
 [7] Gorden Geoffrey (1978), System Simulation, N. J. Prentice-Hall, INC.

[8] Law M. Averill and W. David Kelton (1982), Simulation Modeling and Analysis, New York, Mcgraw-Hill, INC.

[9] Lehman S. Richard (1977), Computer Simulation and Modeling, N. J. Lawrence Erlbaum Associates.

[10] SAS/ETS User's Guide (1988), Version 6, First Edition, SAS institute INC.

[11] Wyman Paul Forrest (1970), Simulation Modeling, New York: John Wiley & Sons, Inc.

<부록 표> ARIMA(p, d, q) 모형에 의한 인구 예측

(단위 : 명)

나이	2015			2021		
	남	여	계	남	여	계
15	315,790	301,785	617,575	309,375	302,875	612,250
16	317,505	305,560	623,065	315,772	302,282	618,054
17	319,220	309,335	628,555	322,169	301,689	623,858
18	320,935	313,111	634,046	328,567	301,095	629,662
19	326,246	316,159	642,405	331,031	304,576	635,607
20	331,557	319,207	650,764	333,495	308,057	641,552
25	360,338	323,380	683,718	336,785	313,681	650,466
30	365,797	332,544	698,341	332,513	307,152	639,665
35	366,533	355,328	721,861	372,240	335,315	707,555
40	392,103	353,169	745,272	420,731	377,973	798,704
45	375,830	348,886	724,716	404,357	375,090	779,447
50	350,551	341,600	692,151	378,783	374,035	752,818
55	323,707	336,055	659,762	364,842	381,638	746,480
60	298,450	309,771	608,221	362,590	361,807	724,397
64	222,898	267,563	490,461	267,790	314,899	582,689
합	17,224,23	16,555,4	33,779,6	18,009,	17,212,4	35,221,6
계	7	42	79	211	13	24

## 정관희



1972년 동국대학교 통계학과  
학사

1975년 동국대학교 대학원 응  
용통계학과 석사

1992년 동국대학교 대학원 전  
산통계학 전공 박사

1980년 ~ 현재 : 경기대학교  
정보과학부 정교수 겸 정보 통  
신 대학원장