Journal of the Korean Data & Information Science Society
2010, **21**(3), 503–511

한국데이터정보과학회지

로지스틱함수모형과 비례이동평균모형에 의한 학생 수 추계와 분석†

송필준[1] · 김종태[2]

[12]대구대학교 전산통계학과

접수 2010년 4월 14일, 수정 2010년 5월 18일, 게재확정 2010년 5월 23일

요 약

본 연구의 목적은 연령진급률 혹은 학년진급률을 추정하기 위한 방법으로 비례법을 사용한 이동평균법에 의한 알고리즘을 제시하는데 있다. 학년진급률에 따른 학생 수 추계방법으로, 이동평균법과 비례이동평균법에 의한 추정방법을 제시하고, 2027년까지의 서울시의 고3학생 수를 추정하여, 한국교육개발원의 2005년, 2006년, 2007년의 로지스틱함수 추정에 의한 고3학생 수 예측결과와 비교 분석하였다. 본 연구의 결과 출생아수의 분포와 비교하여 볼 때, 본 연구에서 제시된 비례이동평균법과 이동평균법의 예측결과가 한국교육개발원의 2005년, 2006년, 2007년의 고3학생 수의 예측결과보다 더 신뢰성이 있는 것으로 나타난다.

주요용어: 교육통계서비스, 로지스틱함수, 비례이동평균법, 서울특별시, 학생 수 추정.

1. 서론

인구수와 학생 수에 대한 추계 혹은 예측에 대한 연구는 오래전부터 매우 중요한 논제로 다루어져 왔다. 고등교육 (현재의 대학) 학생 수 추계를 주제로 Royal Society (1985)는 Moor (1983)와 여러 교육기관들의 연구들에 대하여 Preece 등의 학생 수 추계에 대한 의견 및 토론들을 개제했다. Meade (1988)은 인구수를 추정하기 위한 방법으로 수정된 로지스틱모형을 제시했고, Raeside (1988)는 인구수의 경향을 찾기 위한 모형으로 로지스틱 모형을 이용하여 하였다.

교육과학기술부 산하 한국교육개발원 (2005, 2006, 2007)의 교육통계센터 교육통계서비스시스템은 지방행정자치별 초 ·중 ·고등학교 학생 수에 대하여 2020년, 2021년, 2022년까지 각각 학생 수의 예측결과를 제공하고 있다. 한국교육개발원의 학생 수 예측모형은 선형함수를 모델로 하고, 선형함수의 모수를 로지스틱성장곡선모형, 혹은 로지스틱지수평활 모형에 의하여 추정하였다.

본 연구는 m 이동평균들을 이용한 비례에 의한 초기 추정치를 구하고, m 이동평균 추정치와 초기 추정값들을 사용한 $m \times n$ 비례이동평균법 (proportional moving average)을 이용하는 학생 수를 예측하는 알고리즘을 제시하는데 목적이 있다. 또한 김종태 (2009a)의 학년진급률에 따른 학생 수 예측방법에서 제시한 $m \times n$ 이동평균법에 의한 고3학생 수 추정방법을 소개하고, $m \times n$ 이동평균법과 제시된 비례이동평균의 예측결과와 한국교육개발원의 로지스틱함수를 이용하여 예측한 2005년, 2006년, 2007년의 고3학생 수 예측결과와 비교 분석하였다.

---

† 이 논문은 2008년도 대구대학교 학술연구비 지원에 의한 연구임.
[1] (712-714) 경상북도 경산시 진량면 내리동15, 대구대학교 전산통계학과, 교수.
[2] 교신저자: (712-714) 경상북도 경산시 진량면 내리동15, 대구대학교 전산통계학과, 교수.
E-mail: jtkim@daegu.ac.kr

본 연구의 결과들을 출생아수의 분포와 비교하여 볼 때, 본 연구에서 제시된 $m \times n$ 비례이동평균법과 $m \times n$ 이동평균법의 예측결과가 한국교육개발원의 2005년, 2006년, 2007년의 고3학생 수의 예측결과보다 더 신뢰성이 있는 것으로 나타난다.

2절에서는 학년별 학생 수 예측모형을 소개하고 한국교육개발원의 로지스틱함수 추정 방법을 소개한다. 3절에서는 $m \times n$ 비례이동평균법과 $m \times n$ 이동평균법의 알고리즘을 제시하였다. 4절에서는 3절에서 제시된 고3학생 수의 예측결과들과 한국교육개발원의 2005년, 2006년, 2007년의 고3학생 수의 예측결과를 가지고 2027년까지의 서울특별시의 고3학생 수 예측결과를 비교 분석하였다

2. 학년별 학생 수 추정 모형

인구통계의 모형을 이해하는 데는 구자홍 (2002)과 한국교육개발원의 교육통계서비스 시스템에서 잘 나타나 있다. X_t를 t년의 학생 수라고 할 때, 초 ·중 ·고등학교 학생 수 혹은 인구수 예측을 위한 일반적인 모형을 다음과 같이 설정된다.

$$X_t = \beta X_{t-1} + \phi \mathrm{O_t} + \theta \mathrm{I_t} + \epsilon_t \tag{2.1}$$

X_{t-1}은 X_t의 시차변수로 $(t-1)$년도의 인구수로서 이미 주어지거나 예측된 값이 된다.

O_t는 $(t-1)$년도에서 t 년도 사이에 사망, 질병, 품행, 유학/이민, 타시도의 전출 등의 이유로 학교를 떠난 인구수 (학생 수)이다.

I_t는 $(t-1)$년도에서 t 년 사이에 다른 시도에서 전입과 복학 및 재입학 등으로 유입된 인구수 (학생 수)이다.

O_t와 I_t는 $(t-1)$년도에서 년도 사이 1년 동안의 인구수 (학생 수) 변동을 나타내는 확률변수이다.

$X_t + k$ $(k \geq 0)$을 추정하기 위해서는 O_{t+k}와 I_{t+k}의 추정값이 필요하다. 그러나 실제로 O_{t+k}와 I_{t+k}의 신뢰할만한 통계를 구하기가 현실적으로 어렵다. 그러므로 식 (2.1)에서 O_t와 I_t는 개별적으로 추정되지 않고 통합적으로 추정될 수 있도록 다음 식 (2.2)의 식을 이용하자.

$$\beta_{t+k} \equiv \frac{X_{t+k}}{X_{t+k-1}} = \beta + \frac{1}{X_{t+k-1}} \left(\phi \mathrm{O_{t+k}} + \theta \mathrm{I}_{t+k} \right), \tag{2.2}$$

이때 식 (2.2)의 β 는 시간에 따라 일정한 값을 가지는 모수를 나타내고, 식 (2.2)를 식 (2.1)에 이용하면, 다음과 같이 모형이 단순해진다.

$$X_{t+k} = \beta_{t+k} X_{t+k-1} + \epsilon_{t+k}, \quad k = 0, 1, 2, 3, \cdots . \tag{2.3}$$

여기서 β_{t+k}는 시간 t에서 k (≥ 0)에 따라 변화하는 진급률로서 확률변수가 된다. 그러므로 학생 수 혹은 인구수의 예측은 시간 t에서 k (≥ 0)에 따른 진급률, β_{t+k}를 어떻게 추정하는가에 달려있다.

한국교육개발원 (2005, 2006, 2007) 교육통계시스템에서는 β_t를 추정하기 위하여 β_{t+k}가 경험적인 근거로 인하여 S형 성장곡선모형을 이룰 것이라는 가정 하에서, 로지스틱성장곡선 모형, 혹은 로지스틱 지수평활모형을 사용했다.

3. 비례식을 이용한 이중이동평균법

우리나라 초 ·중 ·고등학교의 학년별 학생 수의 예측을 위하여, 기존의 인구통계 데이터 및 학생 수 통계를 먼저 살펴보아야 한다.

우리나라 교육제도에서 0세 인구가 고등학교 3학년이 되까지는 18년이 걸린다. 그러므로 2009년의 0세 인구는 18년 후인 2027년에는 고등학교 3학년이 된다. 그러므로 기존의 확보된 데이터를 이용하여 학생 수를 예측할 경우에는 2010년 현재로서는 2027년까지 밖에 예측할 수가 없다.

기존의 인구 통계로, 0세-6세 인구는 주민등록인구통계를 사용하였다. 시도별 연령별 주민등록인구통계는 통계청 (1998-2009)에서 구할 수 있다. 그러나 1998년 이전의 시도별 연령별 주민등록인구통계 데이터를 구하기란 무척 어렵다. 기존의 0세-6세 인구를 구하는 다른 방법은 통계청 (2006)의 장래인구추계통계에서 찾는 방법이다. 그러나 장래인구추계는 추계데이터이므로 데이터의 신뢰성이 많이 떨어진다.

주민등록인구통계의 약점은 0세 인구가 성장하여 6세 혹은 7세가 될 때까지 각 세별 인구가 꾸준히 증가한다는 것이다. 이것에 대하여 김종태 (2009b)는 역추정 방법을 사용하여 0세-6세 인구를 다시 추정해 놓았다.

초 ·중 ·고등학교의 학생 수에 대한 통계는 한국교육개발원 (1982-2009)의 '교육통계연감'에서 발표한 1982년에서 2009년까지의 학생 수 자료를 이용하였다.

$X_{t,d}$를 t연도의 d $(d = 0, 2, \cdots, 18)$세 인구수, 혹은 d학년의 학생 수로 다음과 같이 정의하자.

$$d = \begin{cases} 0, 1, \cdots, 6, & \text{각각 0세, 1세,} \cdots, \text{6세 인구수,} \\ 7, 8, \cdots, 12, & \text{각각 초등1, 2,} \cdots, \text{6학년 학생수,} \\ 13, 14, 15, & \text{각각 중학 1, 2, 3학년 학생수,} \\ 16, 17, 18, & \text{각각 고등1, 2, 3학년 학생수.} \end{cases} \tag{3.1}$$

2009년 0세는 1년 후인 2010년에는 1세, 2011년에는 2세, $\cdots$, 18년 후인 2027년에는 고3학생이 된다. 1982년의 고3학생은 실제로 1964년에 0세로 출발하여 18년 후에 1982년에 고3이 된 것이다.

이런 규칙에 따라서 t를 기준연도로 두면, $t = 1982 + (18 - d) - m, \cdots, 2009$와 학년 (연령)인 $d = 0, 1, \cdots, 18$에 대하여, 학년 혹은 연령진급률은 다음과 같이 정의 된다.

$$\beta_{t,d} = \frac{X_{t,d}}{X_{(t-1),(d-1)}}. \tag{3.2}$$

여기서 d $(d = 0, 1, \cdots, 18)$는 식 (3.1)에 정의 된 것이다. 그러므로 초、중、고등학교 학생 수 혹은 인구수 예측을 위한 일반 모형은 다음과 같이 설정된다.

$$X_{t,d} = \beta_{t,d} X_{(t-1),(d-1)} + \epsilon_{t,d}. \tag{3.3}$$

여기서 $\beta_{t,d}$는 기준연도 t에서 연령 d (≥ 0)에 따라 변화하는 진급률로서 확률변수가 된다. 식 (3.3)의 학년 (연령) 진급률 $\beta_{t,d}$의 추정치 $\widehat{\beta}_{t,d}$을 구하기 위한 알고리즘은 다음과 같다.

절차1. 기존의 데이터들인 주민등록인구통계와 한국교육개발원에 초 ·중 ·고등학교의 학생 수 $X_{t,d}$에 대한 식 (3.2)의 $\beta_{t,d}$의 연령진급률 혹은 학년진급률을 계산 한다.

절차2. 학년 (연령)인 $d = 0, 1, \cdots, 18$에 대하여, t연도의 진급률을 포함하는 $(t-1)$, $(t-2), \cdots, (t-m-1)$의 최근 m개의 진급률에 대한 평균 $\widehat{\beta}^{m}_{t,d}$을 t연도의 진급률의 대표값로 둔다. $t = 1982 + (18 -$

$d) - m, \cdots, 2009$에 대하여,

$$\widehat{\beta}^{m}_{t,\ d} = \frac{1}{m}\sum_{k=1}^{m}\beta_{(t-k+1),d}, d = 0, 1, \cdots, 18. \tag{3.4}$$

2010년의 1세 인구, 2010년-2011년까지의 2세 인구, 2010년-2012년까지의 3세 인구, $\cdots\ \cdots$, 2010년에서 2027년까지의 고3학생 수들을 예측하기 위해서 연령 혹은 학년 진급률을 식 (3.4)와 같은 방법으로 김종태 (2009a)은 다음과 같이 제시했다. $d = 0, 1, \cdots, 18$에 대하여,

$$\widehat{\beta}^{m\times n}_{t,\ d} = \frac{1}{n}\sum_{k=1}^{n}\widehat{\beta}^{m}_{(t-k),d}, t = 2010, 2011, \cdots, 2027 - (18 - d). \tag{3.5}$$

절차 2에서 구한 미지의 진급률 $\beta_{t,d}$의 추정치 $\hat{\beta}^{m\times n}_{t,d}$을 이용하여, $d = 0, 1, \cdots, 18$와 $t = 2010, 2011, \cdots, 2027 - (18 - d)$에 대하여, $m \times n$ 이동평균법 (moving average method) 을 이용한 학생 수 추정은 다음과 같다.

$$\mathbf{m}\times\mathbf{n}\text{이동평균법} : \mathbf{mnMAE} \equiv \widehat{X}^{m\times n}_{t,\ d} = \hat{\beta}^{m\times n}_{t,\ d} \cdot \widehat{X}^{m\times n}_{(t-1),(d-1)}. \tag{3.6}$$

여기서 $\widehat{X}^{m\times n}_{t,\ d} = \hat{\beta}^{m\times n}_{t,\ d} \cdot \widehat{X}^{m\times n}_{(t-1),(d-1)}$은 이미 주어진 값이거나 혹은 직전에 예측된 값이다.

미지의 진급률을 계산하는 다른 방법은 당해 연도의 진급률이 직전 연도의 진급률과 같은 비례식으로 작동할 것이라고 가정하고 구하는 방법이다. 그러나 이 비례법을 이용한 방법은 시간이 지날수록 일정한 비율을 가지고, 많은 오차를 수반한다. 이를 보완하기 위한 방법으로 마코브체인이나 시계열모형, 로지스틱모형 등이 있다. 그러나 시계열분석 모형들을 적합 시켜 보았지만 좋은 결과를 얻지 못했다 (김종태, 2009a). 다음의 절차 3에서는 비례법의 단점을 보완하기 위해서, 식 (3.4)에서 구한 진급률을 비례법에 적용시켜서 $d = 0, 1, \cdots, 18$와 $t = 2010, 2011, \cdots, 2027 - (18 - d)$에 대한 진급률의 초기값들을 계산한다. 절차 4에서는 절차 3에서 구한 초기값 진급률들을 다시 이동평균법을 이용하여, 미지의 진급률들을 다시 추정한다. 그리고 다시 절차 3에서 구한 초기값 진급률 계산하여 절차 4의 과정을 반복한다.

절차3. 각 학년 (연령)의 미지의 진급률 $\beta_{t,d}$의 추정을 위해 식 (3.4)에서 구한 진급률을 이용하여, 다음과 같은 비례식을 계산하여. 진급률 $\beta_{t,d}$의 초기값 추정치 $\widetilde{\widetilde{\beta}}^{m}_{t,d}$을 구한다. $d = 0, 1, \cdots, 18, t = 2010, 2011, \cdots, 2027 - (18 - d)$에 대하여,

$$\hat{\beta}^{m}_{(t-1),(d-1)} : \hat{\beta}^{m}_{(t-1),d} = \hat{\beta}^{m}_{t,(d-1)} : \widetilde{\widetilde{\beta}}^{m}_{t,d}, \tag{3.7}$$
$$\widetilde{\widetilde{\beta}}^{m}_{t,d} = \hat{\beta}^{m}_{(t-1),d} \times \frac{\hat{\beta}^{m}_{t,(d-1)}}{\hat{\beta}^{m}_{(t-1),(d-1)}}$$

절차4 절차2과 절차3에 구한 $\widetilde{\widetilde{\beta}}^{m}_{t,d}$을 가지고 n개씩 묶은 이동평균법을 이용하여 $d = 0, 1, \cdots, 18$와 $t = 2010, 2011, \cdots, 2027 - (18 - d)$에 대하여, 미지의 진급률 $\beta_{t,d}$의 추정치 $\widetilde{\widetilde{\beta}}^{m\times n}_{t,d}$ 을 추정한다.

$$\widetilde{\widetilde{\beta}}^{m\times n}_{t,d} = \frac{1}{n}\sum_{k=1}^{n}\widetilde{\widetilde{\beta}}^{m}_{(t-k+1),d}. \tag{3.8}$$

절차 3과 절차 4을 반복하여 구한, 미지의 진급률 $\beta_{t,d}$의 추정치 $\hat{\beta}_{t,d}^{m\times n}$을 이용하여, 미지의 학생 수에 대한 예측값$\tilde{\tilde{X}}_{t,\ d}^{m\times n}$, $d=0,1,\cdots,18$ 와 $t=2010,2011,\cdots,2027-(18-d)$을 추정하기 위하여 $m\times n$ 비례이동평균법 (proportion moving average method)을 이용하여 다음과 같이 구한다.

$$\mathbf{m}\times\mathbf{n}\text{비례이동평균법} : \mathbf{mnPMAE} \equiv \tilde{\tilde{X}}_{t,\ d}^{m\times n} = \tilde{\tilde{\beta}}_{t,\ d}^{m\times n} \cdot \tilde{\tilde{X}}_{(t-1),(d-1)}^{m\times n}. \tag{3.9}$$

이 때 $\tilde{\tilde{X}}_{(t-1),(d-1)}^{m\times n}$은 이미 주어진 값이거나 혹은 직전에 예측된 값이다.

식 (3.4)와 식 (3.5)에서 이동평균의 정의는 시계열에서의 이동평균의 정의와는 다소 차이가 있다. 시계열에서의 이동평균은 데이터들의 중간값들을 추정하는 특성을 가지지만 본 연구에서의 이동평균의 개념은 가장 최근의 진급률에 대하여 가중치를 더해가는 특성을 가지고 있다.

식 (3.9)와 식 (3.6)의 $m\times n$ 비례이중이동평균 방법은 $m\times n$ 이동평균들 보다 더 예측오차가 적은 것으로 나타난다. 그리고 모의실험결과 $m\times n=4\times 4$일 때, 학생 수에 대한 예측결과가 좋은 것으로 나타난다.

4. 로지스틱모형 예측결과와의 비교

연령진급률인 β_t가 경험적인 근거로 인하여 S형 성장곡선모형을 이룰 것이라는 가정 하에서, 한국교육개발원 (2005, 2006, 2007) 교육통계시스템은 β_t를 추정하기위한 방법으로 로지스틱성장곡선 모형 혹은 로지스틱 지수평활모형을 사용했다.

그러나 한국교육개발원 예측결과에서 특이한 점이 발견된다. 2005년과 2007년도의 예측결과는 비슷한 분포모양을 가지는 대신에 2006년도의 예측결과는 2005년과 2007년의 예측결과의 분포 모양에 있어서 많은 차이를 나타낸다. 이 예측결과들은 분명히 서로 다른 방법들을 사용했을 것이다. 한국교육개발원의 결과들의 차이가 로지스틱성장곡선 모형과 로지스틱지수평활 모형의 차이에서 생긴 결과 인지는 확실하지 않다.

또 다른 특이점은 2006년의 예측결과는 2005년과 2007년의 예측결과에 비하여 과대추정 되어있고, 2005년과 2007년의 고3학생 수 예측에 있어서 2016년과 2019년 사이에 지나치게 고3학생 수를 과소 추정하고 있다는 것이다.

표 4.1은 서울특별시의 고3학생 수에 대한 로지스틱모형을 사용한 한국교육개발원의 예측결과와 3절에서 제시된 식 (3.6)의 $m\times n$ 이동평균법 (mnMAE)에 의한 $\hat{X}_{t,\ d}^{m\times n}$의 결과와 식 (3.9)의 $m\times n$ 비례이동평균법 (mnPMAE) $\tilde{\tilde{X}}_{t,\ d}^{m\times n}$ 의 결과를 나타낸 것이다.

김종태 (2009a)에서 예측결과들이 잘 되었는지에 대한 평가방법을 다음과 같이 제시하고 있다. "미래 예측에 있어서, 예측한 결과들이 얼마나 정확할 것인가를 아는 방법은 두 가지가 있다. 첫 번째는 이미 알고 있는 실제값들과 모의실험을 통해서 이들을 추정한 추정값들과의 오차들을 비교하는 방법이고, 두 번째는 미래의 예측값에 가장 큰 영향을 미칠 수 있는 데이터들을 예측된 시점으로 평행이동을 시킨 후에, 예측된 값들과 비교를 통해서 알 수 있다."

미래의 학생 수에 가장 큰 영향을 미칠 수 있는 이미 태어 난 서울특별시의 출생아 수를 고3이 되는 18년 후로 평행이동 시켜서 예측된 고3학생 수들과의 분포를 비교 해 봄으로서 경험적인 판단을 해 보는 것이다.

표4.2은 표 4.1의 통계를 이용하여 직전연도 대비 고3학생 수 감소율과 출생아수 감소율 나타낸 것이다. 3절에서 제시된 mnMAE와 mnPMAE의 결과는 출생아수의 감소율과 거의 비슷한 비율를 가지고 예측되고 있음을 보인다. 그러나 2017년 출생아수의 감소율은 5%에 불과하지만 한국교육개발원의 2005년과 2007년의 예측들 중 2017년의 고3학생 수 감소율을 보면 12%와 15%가 각각 감소하는 것으

표 4.1 서울특별시의 고3학생 수 예측과 18년 이동 출생아수 통계

서울	한국교육개발원 고3학생 수 추정			mnMAE	mnPMAE	18년 이동출생아수
연도	2005년도	2006년도	2007년도			
2010	121,292	123,984	124,960	121,669	121,578	182,662
2011	117,635	119,034	118,564	117,203	117,059	175,760
2012	113,467	117,669	115,385	115,325	114,905	175,433
2013	112,044	115,600	108,822	112,541	112,222	165,822
2014	106,936	110,849	103,967	107,525	107,367	151,695
2015	98,000	109,163	100,891	104,659	104,515	142,141
2016	101,167	103,811	102,781	98,684	98,469	133,761
2017	89,325	100,948	85,422	94,973	94,785	126,734
2018	86,183	101,845	83,730	95,887	95,662	131,932
2019	87,466	91,813	79,157	83,004	82,799	113,628
2020	88,623	83,425	75,888	74,030	73,615	100,919
2021		80,943	77,703	75,938	75,386	100,135
2022			80,224	73,438	72,887	98,790
2023				65,996	65,536	89,489
2024				68,279	67,892	92,885
2025				74,061	73,446	100,107
2026				70,099	69,586	94,736
2027				65,833	65,395	90,300

로 나타난다. 즉, 한국교육개발원의 2017년 학생 수의 예측은 출생아수 감소율과 비교해 볼 때, 10% 이상의 큰 차이를 보이고 있다. 또한 2019년의 학생 수 예측에 있어서 14%의 출생아수 감소율이 나타나는 반면에 한국교육개발원의 로지스틱모형 예측은 1% 증가 혹은 5%감소의 결과를 나타내고 있다.

표 4.2 직전연도 대비 고3학생 수 감소율과 출생아수 감소율

서울	한국교육개발원 고3학생수 감소율			mnMAE 감소율	mnPMAE 감소율	출생아 감소율
연도	2005년도	2006년도	2007년도			
2010	0.01	-0.02	-0.02	0.00	0.00	-0.01
2011	0.03	0.04	0.05	0.04	0.04	0.04
2012	0.04	0.01	0.03	0.02	0.02	0.00
2013	0.01	0.02	0.06	0.02	0.02	0.05
2014	0.05	0.04	0.04	0.04	0.04	0.09
2015	0.08	0.02	0.03	0.03	0.03	0.06
2016	-0.03	0.05	-0.02	0.06	0.06	0.06
2017	0.12	0.03	0.17	0.04	0.04	0.05
2018	0.04	-0.01	0.02	-0.01	-0.01	-0.04
2019	-0.01	0.10	0.05	0.13	0.13	0.14
2020	-0.01	0.09	0.04	0.11	0.11	0.11
2021		0.03	-0.02	-0.03	-0.02	0.01
2022			-0.03	0.03	0.03	0.01
2023				0.10	0.10	0.09
2024				-0.03	-0.04	-0.04
2025				-0.08	-0.08	-0.08
2026				0.05	0.05	0.05
2027				0.06	0.06	0.05

그림 4.1은 표 4.1의 결과를 그림으로 나타낸 것이다. 그림 4.1에서 보듯이 3절에서 제시된 mnMAE와 mnPMAE의 고3학생 수 예측결과는 매우 동일해 보이고, 또한 출생아수의 분포와 거의 비슷한 분포 모양을 이루고 있다. 한국교육개발원 2006년도 예측결과는 mnMAE와 mnPMAE의 분포 모양과 유사하나 mnMAE와 mnPMAE의 예측보다 과대 추정되었음이 보여 진다. 2005년과 2006년의 한국교

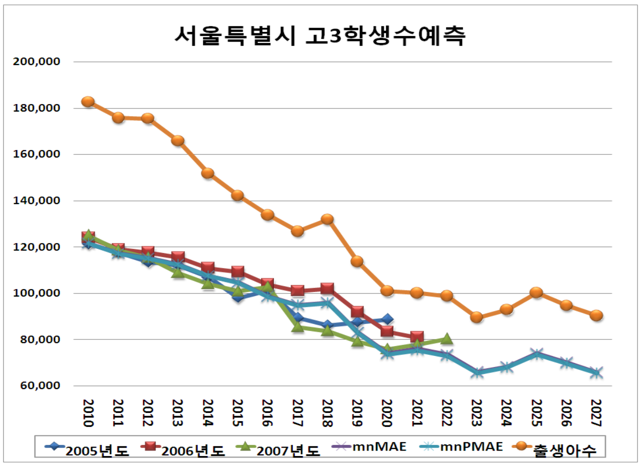

그림 4.1 서울특별시의 고3학생 수 예측과 18년 이동 출생아수 분포

육개발원의 예측결과는 과소 추정을 보이면서 출생아수의 분포와 다른 현상을 보인다.

표4.3은 표 4.1의 통계를 이용하여 2009년의 실제 고3학생 수 데이터를 기준으로 예측결과들의 고3학생 수 감소율과 2009년의 실제 출생아수를 기준으로 출생아수의 감소율 나타낸 것이다. 18년 이동 출생아수는 앞으로 10년 후인 2020년까지 약 44%가 감소하는 것으로 나타난다. 그러나 한국교육개발원 2005년, 2006년, 2007년의 고3학생 수는 각각 27%, 32%, 38%가 주는 것으로 나타났고, mnMAE와 mnPMAE의 예측은 40%가 주는 것으로 나타난다.

표 4.3 2009년 실제 데이터 대비 고3학생 수 감소율와 출생아수 감소율

서울	한국교육개발원 고3학생 수 감소율			mnMAE 감소율	mnPMAE 감소율	출생아 감소율
연도	2005년도	2006년도	2007년도			
2010	0.01	-0.02	-0.02	0.00	0.00	-0.01
2011	0.04	0.03	0.03	0.04	0.04	0.03
2012	0.07	0.04	0.06	0.06	0.06	0.03
2013	0.08	0.05	0.11	0.08	0.08	0.08
2014	0.12	0.09	0.15	0.12	0.12	0.16
2015	0.20	0.11	0.17	0.14	0.14	0.21
2016	0.17	0.15	0.16	0.19	0.19	0.26
2017	0.27	0.17	0.30	0.22	0.22	0.30
2018	0.29	0.17	0.31	0.21	0.22	0.27
2019	0.28	0.25	0.35	0.32	0.32	0.37
2020	0.27	0.32	0.38	0.39	0.40	0.44
2021		0.34	0.36	0.38	0.38	0.44
2022			0.34	0.40	0.40	0.45
2023				0.46	0.46	0.50
2024				0.44	0.44	0.48
2025				0.39	0.40	0.44
2026				0.43	0.43	0.47
2027				0.46	0.46	0.50

* 18년 이동 출생아수의 감소율은 18년 이동 2009년의 출생아수를 기준으로 함.

18년 이동 2009년의 출생아수는 2027년에 이르러 약 50%가 줄어들고, mnMAE와 mnPMAE의

고3학생 수 예측결과는 약 46%가 줄어드는 것으로 나타난다.

서울특별시의 고3학생 수가 46%까지 줄어드는 경우, 서울의 미래는 어떻게 될 것인가를 생각해 보지 않을 수 없다. 저 출산의 심각한 영향으로 인하여 우리나라 인구는 급속히 감소할 것으로 보이고, 이러한 감소율은 다른 나라에서 찾아 볼 수 없을 것이다. 그러나 수도권 집중화 현상이 계속 증가하는 한 서울특별시의 고3학생 수 감소는 수도권지역의 대학 신입생 충원에는 아무런 영향을 미치지 않을 것이다. 적어도 18년 이동 출생아수가 2027년까지 2009년 대비 50%가 줄어든다고 해도, 그러나 수도권으로 인구가 이탈하는 비수도권 지역은 심각한 지경에 이를 수도 있을 것이다.

참고문헌

구자홍 (2002). <인구통계학의 이론과 실제>, 교우사, 서울.

김종태 (2009a). 학년진급률에 따른 학생수 예측방법. <한국데이터정보과학회지>, **20**, 857-867.

김종태 (2009b). 주민등록 0세-6세 인구의 역추정과 기존 인구통계와의 출생아수 비교. <한국데이터정보과학회지>, **20**, 1145-1153.

통계청 (1998-2009). <각세별 시도별 주민등록인구통계>, 통계정보시스템, 대전.

통계청 (2006). <장래인구특별추계결과>, 통계정보시스템, 대전.

한국교육개발원 (2005). <교육통계 예측결과>, 교육통계서비스, 서울.

한국교육개발원 (2006). <교육통계 예측결과>, 교육통계서비스, 서울.

한국교육개발원 (2007). <교육통계 예측결과>, 교육통계서비스, 서울.

한국교육개발원 (1982-2009). <교육통계연보>, 교육통계서비스, 서울.

Meade, N. (1988). A method logistic model applied to human population. *Journal of Royal Statistical Society series A*, **151**, 491-498.

Moore, P. G. (1983). Higher education: The next decade. *Journal of Royal Statistical Society series A*, **146**, 213-245.

Raeside, R. (1988). The use of sigmoids in modelling and forecasting human population. *Journal of Royal Statistical Society series A*, **151**, 499-513.

Royal Society (1985). Projections of student numbers in higher education. *Journal of Royal Statistical Society series A*, **148**, 175-213.

Journal of the Korean Data &
Information Science Society
2010, **21**(3), 503–511

한국데이터정보과학회지

Projection of the student number by logistic function and proportional moving average model†

Phil Jun Song[1] · Jongtae Kim[2]

[12]Department of Computing and Statistics, Daegu University

Received 14 April 2010, revised 18 May 2010, accepted 23 May 2010

Abstract

The goal of this paper is to suggest an algorithm to get the number of student on the elementary, meddle and high-school for the forecasting of the numbers of student by the moving average method using a proportional expression. Comparing with the results of Korean education statistical system 2005, 2006, and 2007, the results of this paper are better than those of the Korean education statistical system.

Keywords: Educational statistical services, logistic function, proportional moving average, Seoul.

† This research was supported by the Daegu University Research Grant 2008.
[1] Professor, Department Computing and Statistics, Daegu University, Kyoungbook 712-714, Korea.
[2] Corresponding Author: Professor, Department Computing & Statistics, Daegu University, Kyoungbook 712-714, Korea. E-mail: jtkim@daegu.ac.kr