

시계열 분석을 통한 시도별 고등학교 학생 수 예측

An Analysis of the Estimated Number of High School Students between 2016 and 2020 by Time Series Analysis

임성범*, 박선행**
단국대학교 행정학과*, 동국대학교(서울) 교육학과**

Seong-Bum Lim(iamsb711@naver.com)*, Sun-Hyung Park(sunhpark@dongguk.edu)**

요약

현재는 저출산·고령화 사회현상에 따라 한국의 교육환경이 급변하는 시점이다. 특히, 고등교육 시장에서 상당한 변화가 예상되는데 통계청의 보고에 따르면 2010년에서 2020까지 10년 동안 고등학생 수는 196만 명에서 127만명으로 35% 감소하는 것으로 나타났다. 학생 수의 변화는 교원 수급문제와 직·간접적으로 연계되며, 이는 적정한 교육재정 배정 및 예산 확보와 함께 다시 학급 당 학생 수, 교사 1인당 주당 수업시수 등 교육의 양적·질적 부문에 매우 중요하게 영향을 미치게 된다. 이러한 상황에서 교육시스템 운용에 중요한 영향요인이며 결정요인으로 기능할 것으로 예상되는 재학 학생 수에 대한 정확한 예측모형의 제시는 향후 국가의 교육행정체제와 인사관리 방향성을 결정하는데 있어서 그 가치가 매우 높다고 사료된다. 이 연구에서는 시계열 분석기법의 특성 상 축적된 과거 데이터가 충분한 11개 시도의 고등학생 수 사례를 대상으로 한다. 주어진 실측 데이터에 대한 모형적합도 검정과 오차추정을 위해 다양한 시계열 예측모형과 오차 추정방식을 적용 하였다.

■ 중심어 : | 학생 수 예측 | 시계열 분석 | 교육시스템 | 모형적합도 | 오차추정 |

Abstract

Since the number of student is regarded as the fundamental basis to calculate the future allocation of employed teachers, it needs to be systematically estimated based on statistical data. In order to achieve this purpose, the number of high school students is projected following the assumption that the teacher-student ratio of Korea should be adjusted to the level of OECD to improve the quality of education. Hence, this paper introduced the projection methods by time series model. To predict the number of high school students and error estimation, various models were adopted.

■ keyword : | The Estimated Number of High School Students | Time Series Analysis | Model Validity | Error Estimation |

I. 연구의 목적 및 필요성

한국사회는 저출산 고령화 시대의 도래로 인하여 사회전반에 있어 급격한 구조적 변화에 직면하고 있다.

생산가능 인구 부족, 인건비 상승으로 인한 기업경쟁력 약화, 노인인구 부양에 따른 복지비 지출 확대 등은 이러한 환경 변화가 초래하는 대표적인 사회 문제현상으

접수일자 : 2016년 10월 14일
수정일자 : 2016년 10월 31일

심사완료일 : 2016년 10월 31일
교신저자 : 박선행, e-mail : sunhpark@dongguk.edu

로 간주되고 있다. 특히, 저출산과 인구절벽 등으로 인해 2018년 고령사회, 2026년 초고령 사회로의 진입이 예상되는 상황에서 학령인구의 지속적 감소는 교육체제 운영과 관련 정책 실행 전반에도 부정적인 영향을 미치고 있다. 통계청의 국가통계포털에 따르면, 전국기준으로 고등학교급(만16~만18세)의 학령인구는 2015년 1,846(천명)에서 2023년 1,349(천명)으로 줄어드는 것으로 나타나고 있으며, 고정시점에서 2015년 학생 수의 73% 수준까지 감소하는 것으로 예측되고 있다. 공교육의 기본 운영구조는 궁극적으로 교육수요자인 학생 숫자에 의해서 결정된다. 학령인구의 급감은 결국에는 취학 학생 수의 감소를 초래하게 되어 학교급 별 재학생 수와 지역의 학교 수 및 학생밀도 등에 영향을 미치게 되며, 장기적으로 교육재정 투자 규모범위 결정에도 부정적 변수로서 작동한다[1][2]. 또한, 이러한 상황은 대학 입학정원 관리체제를 포함한 고등교육 시장 전반에 큰 변동을 초래한다. 학령인구의 지속적 감소는 비단 지역의 교육 불균형 뿐 아니라 대학입학 자원의 감소로 다시 이어지며, 이는 대학교육체제의 재정 건전성을 위협하는 위험 요인으로 귀결된다[3].

한편, 학생 수의 감소와 이에 따른 지역 간 학생밀도 불균형은 교육의 질적인 부분을 총괄적으로 책임지는 교원의 안정적 수급과 정원 산출, 그리고 합리적 배정 방안 모색을 어렵게 만든다. 일선 교육현장에서 학생의 학업성취도 제고와 교육 효과성 함양에 가장 강력한 영향을 미치는 교원의 불안정한 수급 배정은 창의경제 교육 패러다임에 대한 적극적 대응역량을 약화시켜 국가발전과 사회통합에 필요한 창의적 인재양성을 저해하게 된다. 따라서 교원정원 관리 시스템의 안정적 운용, 지역 간 교원배정 및 학교체제의 효율적 관리, 교육재원의 합리적 총당과 배정 등 교육기본 여건 예측과 관련 정책구안 및 실행을 결정하는 기초요인인 학생 숫자에 대한 체계적이고 실증적인 추세파악이 필요하다. 특히, 저출산 시대에 따른 학령인구 급감이 진행되고 있는 상황에서 지역별 고등학교 학생 수에 대해서 명확한 추세분석이 선행될 때에 한해서 인구추세와 지역여건이 반영된 합리적인 교원정원관리 및 교육재정 배분계획 수립이 가능할 수 있다. 이러한 문제의식에 근거하

여 본 연구는 조건이 동일하다는 가정 하에 미래는 과거의 함수라는 사실을 적극 고려하여 과거 일정기간 동안의 년도 별 데이터에 대한 체계적 분석을 통해 과거의 고등학교 학생 수 증감 패턴을 파악하고, 이 패턴을 미래에 투영함으로써 미래 시점의 고등학교 학생 수를 예측하여 향후 지역의 균형 있는 교육여건을 수립함과 동시에 교육행정 인사시스템을 안정적으로 운용하는데 필요한 기초자료 및 연구방향을 제시하는데 목적을 두고자 한다.

이를 위해 본 연구는 교육통계 자료를 참고로 하여 서울, 부산, 경기, 강원, 충북, 충남, 전북, 전남, 경북, 경남, 제주 11개 시도의 고등학교 재학생 수 시계열 분석에 초점을 두었다. 시계열 데이터는 1965년부터 2015년까지 51년간 년도 별 데이터를 사용하였고, 예측 정확성을 확인하기 위하여 사후 평가를 실시하였다. 또한 정확하고 설득력 있는 예측을 위해 2016년부터 2020년까지 5년간 각 지역의 고등학생 수를 예측하였다. 시계열 분석기법으로는 모형에 대한 모수추정, 적합도 검정과 미래 예측 값 제시를 위해 다양한 시계열 모형(ARIMA 모형, 지수평활법 등)을 적용하였다. 특히, 여기서는 미래 일정 시점의 학생 수 예측을 위해 단일변량에 대한 시계열 분석기법을 적용하는데 이것은 시계열 분석기법이 지닌 분석적 특성과 학생 수 데이터가 지닌 계열 상관의 특성, 그리고 본 연구의 주제와 관련되어 확보 할 수 있는 데이터의 한계 등을 그 이유로 볼 수 있다.

II. 선행연구 검토

인력수요 예측은 노동력의 수요자와 공급자 간 비대칭적 고용정보에 따른 인력수급 불균형 해소와 과잉교육방지 및 직무불일치 예방을 통해 국가 인적자원의 효율적 양성과 배분에 기여할 수 있다는 측면에서 산업영역에서 중요하게 다루어지고 있다. 사회 각 분야의 인력수요 예측을 위한 대표적인 연구 결과물로는 간호사 인력의 수요예측, 봉제 산업에서의 인력양성, 실버산업 관련 인력 수요예측, 안경사 인력수급전망 등을 들 수

있다[4-7]. 인력 수요 예측을 위한 시계열 자료의 분석 방법은 기준에 따라 다양한 하위요소로 구분되는 경향이 있다. 연구자의 시각에 따라 질적·양적 예측방법으로 구분될 수도 있으며 다른 한편으로는, 시계열 자료 예측시 설명변수의 고려 여부에 따라 ARIMA 모형과 지수평활(exponential smoothing) 모형으로 대변되는 경향외삽법(trend extrapolation method), 구조적 방법(structural method), 코호트 요인법(cohort component method) 세 가지 방법으로 분류되기도 한다[2][8][9]. 이러한 기법 중에서 시계열 자료 분석방법으로 지수평활 모형과 ARIMA 모형이 주로 활용되고 있으며, 교원의 수급예측과 학생 수 예측에는 ARIMA 모형이 대표적인 통계분석 방법으로 채택되고 있다. 또한, 특정 연구들에서는 ARIMA 분석을 통해 10년(2000-2009년)간의 초·중등 교원 수요와 공급을 분석함으로써 이후 유사 연구 수행을 위한 시계열분석 방법의 시초를 제시한 바 있다[10-13]. 학생 수 추계는 국책기관을 중심으로 활발하게 수행된 바 있다. 특히, 한국교육개발원은 국책 교육기관으로서 장기 교원수급연구(1985), 한국 교육비전 2020연구(1997), 유·초·중등 학생 수에 대한 미래예측 연구(1998), 2000년대의 기초 지자체의 초·중·고등학교 학생 수 추정연구 등을 실행한 바 있다. 또한, 이 분야와 관련되어 최근에 개인연구들이 수행되고 있다. 예를 들어 학교급 별 학생 수 예측치를 산출한 선행연구 결과와 교육통계연보상의 학생 수 실적치를 비교하여 제시한 연구가 있다. 여기서는 특히 코호트 요인이 추계된 방법과 학생 수 지표의 평균고정방식, 로지스틱 선형평활 방식 등의 다양한 추계방식을 적용하여 학교급 별 학생 수 예측치를 산출한 연구결과들을 요약 제시하고 있다[2]. 다른 연구에서는 경향외삽법의 한 종류인 비선형 회귀모형을 이용하여 전국 학년별 장래 학생 수 추계를 하였다. 예측방법으로서 비선형 회귀모형인 로그모형과 거듭제곱 모형을 사용하였다. 여기서는 예측결과와 신뢰성을 검증하기 위해 제시된 두 가지 방식의 R-square의 신뢰구간을 서로 비교하는 방식을 적용하였다[14]. 그리고 이동평균법과 비레이동평균법을 이용한 서울시의 고3학생 수를 추정한 연구가 있다. 이들의 보고에 의하면 이전의 연구들에서 로지스틱 함수 추정

을 통한 예측결과보다 더 신뢰성이 있는 것으로 나타났다[15]. 또 다른 연구에서는 학생 수 추계에 있어서 통계청 자료와 주민등록 자료를 통해 장래 학생인구 추계치를 산출하고, 이 값에 일정기간 동안의 재학생 수를 학령인구로 나눈 값을 개입시켜 전국 학생 수를 제시한 바 있다[16].

이처럼 학생 수 추계에 있어서는 다양한 방식과 변수들이 고려될 수 있는데 방법론의 다양성만큼 그에 상응하는 한계점이 존재한다. 이러한 이유로 각각의 연구들이 적용한 방법에 따른 실측값과 추계 값 간에는 일정 수준의 오차가 존재하는 성향이 있다. 학생 수 추계 관련 선행연구를 총괄적으로 분석한 연구에 따르면 상향식 학생 수 지표 평균 고정법이 상대적으로 제일 낮은 수준의 오차비율을 나타내고 있으며, 진급률의 로지스틱 선형평활 방법이 그 다음으로 낮은 수준의 오차비율을 보이고 있다고 언급한 바 있다[2][17][18].

III. 시계열 모형

1. 예측모형 생성 및 예측 값 산출

먼저, 고등학생 수의 예측을 위해 11개 시도의 자료를 활용하여 1965년부터 2015년까지 51년 간 자료의 일반적 현황 즉, 기간 동안의 평균, 전년대비 증감률, 추세 존재의 통계적 유무검정을 실시하고, 시계열 분석을 위한 기술통계치(평균, 표준편차, 최대/최소값)를 분석한다. 다음으로 Durbin-Watson 검증과 자기상관 검증(Auto-correlation test)을 통해 시계열 내의 연관관계를 검토하고, 모형오차 검증과 예측모형으로 부터 주어진 데이터에 적합도가 가장 높은 모델을 선정한다. 마지막으로 고려된 기간(2016-2020)에 대한 예측 플롯을 작성하고 평균증감률 계산과 모델에 대한 평가를 최종 실시한다.

이 연구에서는 과거의 시계열 자료에서 나타난 학생 수 변화에 대한 추세를 바탕으로 미래 예측치를 산출한다. 그리고 이러한 시계열 예측에 있어서 설명변수의 개입 없이 단일변량 시계열 예측을 적용한다. 이는 학생 수 추계의 경우 교육통계연보상의 학생 수 외에도

인구 추계치, 주민등록 인구통계 등 다양한 자료의 보정이 필요하다. 그러나 선행연구들에서 알 수 있듯이 예측오차가 필연적으로 존재하는 상황에서 최근 실적치의 평균값을 코호트 요인으로 적용하는 것은 높은 분석적 효율성을 기대할 수 없다[19]. 이뿐만 아니라 데이터가 지닌 과거의 추세(특성)를 통해 미래를 투영하는데 있어서 자료 간에 서로 독립적이지 않고 전기 또는 전년 동기간 의존관계에 있기에 자료 간의 상호독립성을 가정하는 분석기법의 적용은 자체적으로 내재적 한계성을 가진다[20].

따라서, 본 논문은 11개 지역의 고등학생 수 예측에 있어서 일반적 시계열 예측방법의 하위 카테고리로 고려될 수 있는 질적·양적 예측방법과 경향외삽법, 구조적 방법, 코호트-요인법 중에서 특히, 경향외삽법(trend extrapolation method)을 적용하고자 한다.

이 방식은 ARIMA 모형, 지수평활법, 선형, 기하, 로지스틱, 다항곡선 모양 등 다양하게 구분할 수 있지만 여기서는 ARIMA 모형과 SMA(Single Moving Average), SES(Single Exponential Smoothing), DMA(Double Moving Average), DES(Double Exponential Smoothing), DTS(Damped Trend Smoothing-non seasonal)를 동시에 적용하여 모형의 적합도와 오차분석을 통해 최적의 모형을 선정한다.

학생 수를 추계하는 연구유형 중 대표적인 것을 살펴보면 비선형 회귀모형을 위해 추정 수식에 적용되는 초기 값을 추정하고 이를 통해 다시 모수를 추정하거나 코호트로 각 학교 급에서의 인원을 합하여 시계열 예측 모형(exponential smoothing with damping trend)을 적용한 후 다시 시뮬레이션을 하는 연구들이 있다[14][19]. 이외에도 비례법을 이용한 비례이동평균법(proportion moving average method)을 활용한 연구도 있다[15]. 특히, 비례이동평균법을 활용한 연구에서는 미지의 진급률 계산에서 이동평균법의 적용이 시간이 지남에 따라 많은 오차가 수반됨을 지적하고, 비례이동평균법의 필요성을 역설하고 있다. 본 연구에서는 이러한 연구들과 달리 고등학교 재학 학생 수라는 실측 데이터를 기반으로 과거의 추세가 미래에 유지될 것이라는 가정을 가지고 접근한다. 이 경우에는 서로 간 장단점이

존재한다. 모수를 추정하는데 있어서 실측데이터를 기반으로 모형에 대한 검정을 한다는 장점과 더불어 코호트를 통해 모수를 추정한다는 부정적 측면이 공존한다. 본 연구에서는 장기적 예측보다는 단기의 추세를 예측하고 다양한 시계열 모형의 상호간 비교를 통해 모수추정과 모형오차를 파악하는 데 초점을 둔다. 또한, 앞에서 언급한 것처럼 코호트 요인의 적용에서 평균치를 활용하는 것보다 실측치를 통해 추세를 연장하는 것이 보다 효과적일 수 있다[19]. 따라서 실측데이터에 대한 모형의 적합도를 기반으로 시계열 예측분석을 시행하고자 한다.

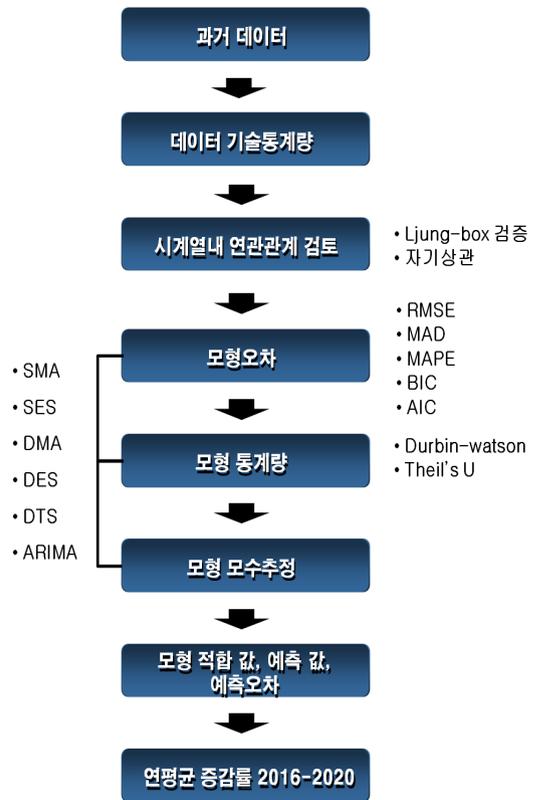


그림 1. 시계열 모형 생성

2. 예측모형

본 연구에서는 계절성을 고려하지 않은 상태에서 추세가 있는 경우와 없는 경우로 구분하여 데이터를 분석

한다. 추세가 없는 데이터의 경우 단순이동평균(Single Moving Average)과 이중 이동평균(Double Moving Average)을 적용하며, 추세가 있는 데이터의 경우 단순 지수평활(Single Exponential Smoothing)과 이중 지수평활(Double Exponential Smoothing)을 적용한다 [20][21].

먼저, (2) 단순이동평균(Single Moving Average)은 계절성과 추세가 모두 없는 데이터에 적용이 가능하며, 과거 실제 데이터의 평균값을 이용해 프로젝트 향후 실제 값 예측에 적용된다. (3) 이중이동평균(Double Moving Average)은 원래 데이터 셋의 이동평균으로 구성된 데이터 셋에 대한 이동평균을 계산함으로써 과거 데이터를 평활 시키는 방법이다. 즉, 이 방식은 앞의 단순이동평균(SMA)에 대한 이동평균을 다시 계산하는 것으로 이중 이동평균을 계산하면 데이터의 추세를 파악할 수 있다[21]. 다음으로 (4) 단순지수평활(Single Exponential Smoothing)은 데이터에 대한 명확한 추세나 계절성이 없는 경우 사용할 수 있는 두 번째 시계열 방법이다. 이 방법은 과거 데이터에 가중치를 부여하며, 이때 가중치는 과거로 갈수록 지수적으로 작아진다. 추가되는 가중치로 이동평균 모델 또는 백분율 변화모델(percentage change model)의 한계를 대부분 극복할 수 있다. 여기서의 가중치를 alpha scale(α -scale)이라고 한다[21]. (5) 이중 지수평활(Double Exponential Smoothing)은 추세는 있지만 계절성이 없는 데이터에 적용 가능한 방식이다. 이중 지수평활은 원래 데이터에 단순지수평활을 적용하여 나온 값에서 다시 한 번 지수평활을 적용하는 방식이다. 가중모수 α 는 단순 지수평활의 2차 시행(DES)에서 적용한다. 예를 들어 추세가 있는 선형모형 $y_t = \beta_0 + \beta_1 t + \epsilon_t$ 에 대하여 β_0 와 β_1 의 수정된 추정 값을 결정하기 위해 사용되는 방법을 이중지수평활이라 할 때 시점 t 에서 β_1 에 대한 수정 추정 값을 $\hat{\beta}_1(t)$ 라 하며, 다음과 같다(1)[20].

$$\hat{\beta}_1(t) = \frac{\alpha}{1-\alpha} (S_t - S_t^{(2)}) \quad (1)$$

또한 여기서는 일반적인 시계열 예측모형 외에도 감

쇠추세-비계절(DTS)방식을 적용한다(6). 이 방식은 지수평활방식에 비해 보다 정확한 예측 값을 제공할 수 있다는 측면에서 매우 인기 있는 방식이다. ARIMA 모델과의 비교에서도 모수를 줄임으로써 완전 통제가 가능하고, 보다 정확한 예측 값을 제공할 수 있다. ARIMA 모델의 경우 모형 작성에 있어서 박스-젠킨스(Box-Jenkins method)을 따른다(7). 크게 3단계의 과정을 거치며, 1단계에서는 ARIMA(p,d,q)모형의 동정(同定)을 실시하며, 2단계에서는 최우법(the method of maximum likelihood estimation)등으로 모수 a_i, b_j 를 추정한다. 3단계에서는 잔차에 대한 임의성(randomness) 여부를 검토하여 random일 경우 최종모형으로 결정하게 된다[22][23]. 특히, ARIMA모형의 경우 사례 수(N)에 대한 충분한 확보가 어려울 경우 실측치를 통한 추세연장에서 데이터의 변동성에 따라 오차범위가 같이 증가하는 경향이 있다[2].

그리고 주어진 데이터(1965-2015)에 대한 연평균 증감률 x (%)과 이 데이터를 통해 예측된 향후 5년간의 예측 데이터에 대해서는 평균 증감률을 계산한다(8). 특히, 일정기간동안의 데이터 변동에 대한 증감률 계산에 있어서 산술평균이 아닌 기하평균을 적용하는 이유는 예측에 있어서 과대 또는 과소 값이 산출되는 것을 방지하기 위해서 기하평균(geometric mean)을 적용한다. 또한, 부산, 강원을 제외하고 51개 기간 동안 행정구역의 변경 등으로 인해 데이터의 추세에 상당한 변동이 발생하는 지역이 대부분이었기에 이들 데이터에 대해 보간법을 실시하였다. 보간법은 크게 Cubic spline interpolation과 Neighbor interpolation 두 가지 방법 중 전체 데이터 세트의 상태를 평가할 때 주어진 점들을 매끄럽게 연결시킬 수 있는 Cubic spline interpolation으로 결정 하였다. 즉, 서로 떨어져 있는 두 점 사이를 연결해야 할 때 연결하는 선을 3차 다항식($a_0 + a_1x + a_2x^2 + a_3x^3$)으로 만들어서 적용하였다[24].

특히, 본 연구에서는 지역단위 데이터를 통한 상향식 접근을 지향하는데 관련 선행연구는 학생 수 추계에 있어서 이러한 필요성에 대해 언급하고 있다[2]. 예를 들어 시도단위의 취학을 단위를 전국단위로 단순화할 경우 취학을 추이를 예측하는 값이 과소 또는 과대 추정

될 수 있으며, 학생 수 추계의 경우 교육통계연보상의 학생 수 외에도 인구추계, 주민등록 인구통계 등 다양한 자료를 통한 코호트요인의 보정이 필요하기에 ARIMA 보다는 지수평활이 보다 효과적이라 주장하고 있다. 본 연구에서는 이러한 부분을 종합적으로 고려하여 다양한 시계열 예측모형을 통한 데이터의 적합도 산출과 예측 값을 제시하고자 한다.

$$SMA_n = \frac{\sum_{i=1}^n Y_i}{n} \quad (2)$$

$$DMA = 2MA_{1,t} - MA_{2,t} + \frac{2}{m-1}[MA_{1,t} - MA_{2,t}] \quad (3)$$

$$ESF_t = aY_{t-1} + (1-a)ESF_{t-1} \quad (4)$$

$$DES_t = \beta(SES_t - SES_{t-1}) + (1-\beta)DES_{t-1} \quad (5)$$

$$SES_t = \alpha Y_t + (1-\alpha)(SES_{t-1} + DES_{t-1})$$

$$DTS(\hat{X}_{t(m)}) = S_t + \sum_{i=1}^m \varnothing^m T_i, \quad (6)$$

여기서, S_t 는 t 기 점의 수준(level),

T_i 는 t 기의 추세

$$ARIMA(Z_t) = (\alpha_1 Z_{t-1} + \dots + \alpha_p Z_{t-p}) + \epsilon_t - \beta_1 \epsilon_{t-1} - \dots - \beta_q \epsilon_{t-q} \quad (7)$$

$$x(\%) = 100 \left(\frac{a_n}{a_0} \right)^{\frac{1}{n-1}} - 100 \quad (8)$$

여기서, $a_n = a_0 \left(1 + \frac{x}{100} \right) n$

a_0 는 당해 년도의 값

3. 오차추정

오차의 추정은 최적 시계열 예측모형을 선정하는데 있어서 매우 중요한 기준이다. 여기서는 대표적인 오차 추정 지표들(RMSE: root mean squared error, MSE: mean squared error, MAD: mean absolute deviation, MAPE: mean absolute percent error)과 모형적합도 값으로 고려되는 Theil's U 값을 평가기준으로 적용하였다.

먼저, MSE는 오차 값이 서로 상쇄되지 않도록 오차

를 제공하여 구하는 절대 값 오차의 척도이다. 이 방식을 적용할 경우 오차를 제공하기에 오차 값의 크기에 따라 각 시점에서 상대적으로 그 영향력이 다소 과장되게 나타날 수 있으며, 이러한 특성은 오차의 크기와 효과에 따라 모형 선정에서 중요하게 기능할 수 있다(9). RMSE는 MSE의 제곱근으로 예측오차의 평균값으로 정의할 수 있다. 이 방식은 특히 예측오차와 관련된 비용이 예측오차의 절대 값 크기와 비례할 때 사용하기에 적합한 방식이다(10). MAD는 실제 데이터와 예측 데이터 값 간의 거리의 절대 값으로 계산되는 오차 통계량이다[21]. 이 방식은 예측오차의 손실비용이 예측오차의 절대크기와 비례할 때 사용하기 적합한 방식이다(11). MAPE는 과거 실제 데이터에 대한 비율로 표시된 오차의 평균으로 계산되는 상대적 오차 통계량으로 예측오차와 관련된 비용이 오차의 크기보다 실제 데이터에 대한 비율에 더 상관있을 때 사용한다[21]. 특히, MAPE는 다른 오차추정 방식에 비해 모델 간 비교가 용이하고, 신뢰도가 높기 때문에 자주 이용된다(12). MAPE는 다음과 같이 해석이 가능하다: $0\% \leq MAPE \leq 10\%$ 는 '매우 정확한 예측', $10\% \leq MAPE \leq 20\%$ 는 '비교적 정확한 예측', $20\% \leq MAPE \leq 50\%$ 는 '비교적 합리적 예측', 그리고 $MAPE \geq 50\%$ 는 '부정확한 예측'[25]. Theil's U 값은 통계량 값이 1.0 보다 작을 때 사용된 예측방법을 통해 추정된 값이 어렵짐작한 값보다 통계적으로 우수하다는 뜻이다(13)[26]. ARIMA 모형의 순서를 지정하는데 있어서 도움이 되는 통계량으로 AIC(Akaike's Information Criterion)과 BIC(Bayesian Information Criterion)이 있다. AIC는 해당모형이 관측된 계열에 얼마나 적합한지를 고려하고자 할 때 유용하게 사용되며, 기본적으로 모수의 수와 모델 효율성과의 관계를 통해 산출된다[20]. 같은 데이터에 대해 AIC가 큰 만큼 좋지 못한 모형이 된다(14). AIC가 지닌 문제점으로 지적된 자기회귀 차수의 과대추정 문제를 해결하기 위해 BIC가 제시되었으며, 해석은 역시 가장 적은 BIC 값을 지닌 모형을 최선으로 본다(15).

$$MSE = \sum_{i=1}^n \frac{(Error_i)^2}{n} = RMSE^2 \quad (9)$$

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(Error_i)^2}{n}} = \sqrt{MSE} \quad (10)$$

$$MAD = \sum_{i=1}^n \frac{|Error_i|}{n} \quad (11)$$

$$MAPE = \sum_{i=1}^n \frac{\left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|}{n} * 100 \quad (12)$$

$$Theil's U = \sqrt{\frac{\sum_{i=1}^n \left[\frac{\hat{Y}_t - Y_t}{Y_{t-1}} \right]^2}{\sum_{i=1}^n \left[\frac{Y_t - Y_{t-1}}{Y_{t-1}} \right]^2}} \quad (13)$$

$$AIC = T \ln \hat{\sigma}_{MV}^2 + 2K \quad (14)$$

$$BIC = T \ln \hat{\sigma}_{MV}^2 + K \ln T \quad (15)$$

여기서, T 는 샘플 사이즈
 $\hat{\sigma}_{MV}^2$ 는 분산의 최우추정치
 K 는 모수

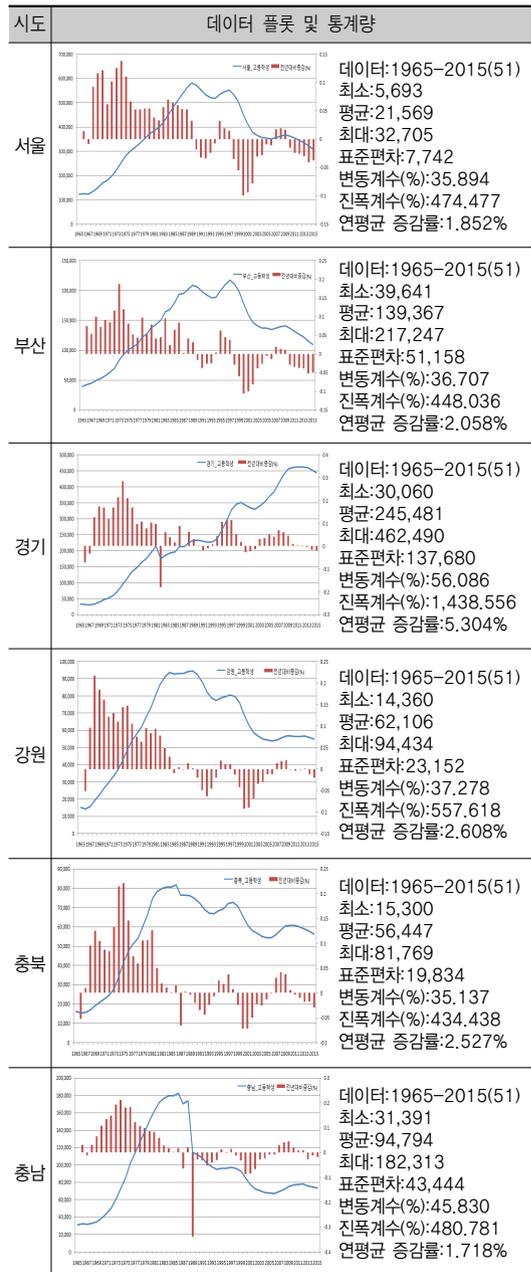
IV. 분석결과

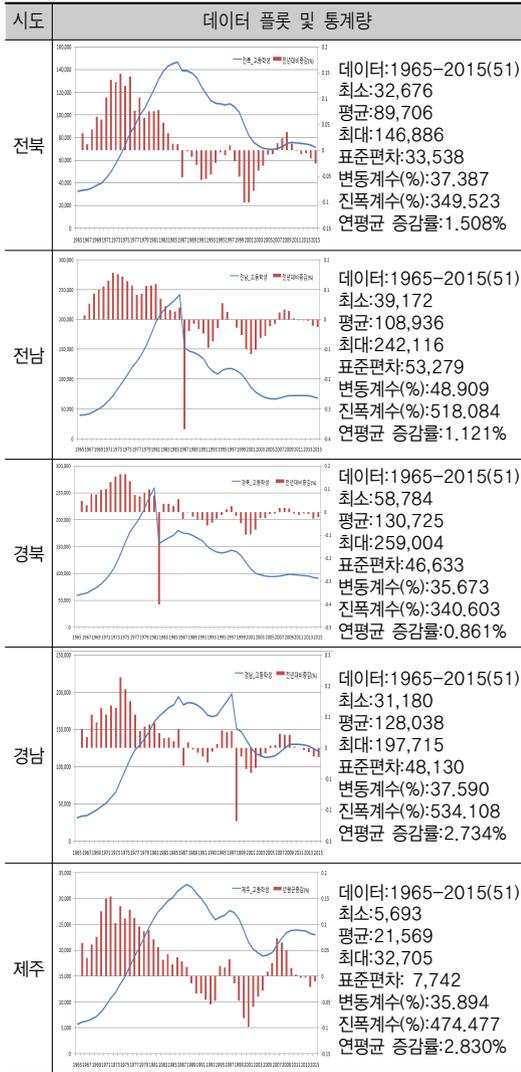
1. 11개 시도의 고등학생 수 예측

1.1 과거 데이터의 기술통계량

1965년부터 2015년까지 고등학생 수를 토대로 2016년부터 2020년까지의 각 시도별 학생 수를 예측하였다. 여기서의 학생 수는 미래의 학생 수 변동에 대한 특정한 가정을 전제로 학교에 재적 또는 등록된 학생 수를 향후 일정 기간에 걸쳐 산출하는 것을 의미한다. 제시된 11개 시도가 상이한 데이터 계열의 평균과 표준편차를 지니고 있기에 이를 동일한 기준에서 데이터의 변동을 파악하기 위해 변동계수(coefficient of variance)를 적용하였다. 변동계수가 가장 큰 지역은 경기지역으로 56.086이고, 가장 작은 지역은 충북으로 35.137로 나타났다. 또한, 동일 기간 동안 최솟값 대비 최댓값과의 상대적 비율을 통한 진폭의 정도를 파악하기 위해 진폭계수를 계산하였다. 역시 경기 지역이 가장 큰 진폭의 정

도를 나타내었으며, 가장 작은 지역은 경북으로 나타났다. 기하평균을 적용하여 51년 동안의 연평균 증감률을 계산하였으며, 그 결과 경기지역을 제외한 대부분의 지역이 2% 수준의 증가를 나타내고 있었다.





주: 변동계수=(sd/mean)*100, 진폭계수=(max-min)/min*100

그림 2. 각 시도별 학생 수 추세

1.2 모형 적합도 검증

각 시도의 데이터의 추세를 가장 잘 반영할 수 있는 시계열 예측 모형을 도출하였다. 이에 대한 도시 간 비교에서 중요한 오차통계량으로 고려되는 MAPE는 모든 모형에서 10% 이하로 나타나 매우 정확한 예측으로 나타났으며, 그 중 서울(ARIMA 모형)의 경우 MAPE가 가장 낮은 1.36%였으며, 가장 높은 지역은 경북(ARIMA 모형)으로 4.50%로 나타났다. 다음으로 중요

한 예측오차의 평가기준으로 Theil이 제안한 U 값은 1 이하일 때 일반적으로 예측된 값의 정확도가 높다고 본다[26]. 이에 대해 특정 연구자들은 U 값의 한계범위가 다소 모호할 수 있음을 지적하고 보다 엄격하게 0.3과 0.4 사이에 놓일 때 예측결과의 우수함을 경험적으로 고려할 수 있다고 지적하였다[20]. 본 연구에서 상이한 시계열 예측모형을 적용한 결과 모든 모형에서 U 값이 1 이하로 나타났으며, 가장 높은 지역은 경북(ARIMA)으로 0.7547, 다음으로 충남(DTS)이 0.6212로 나타났다. 이 외에도 경남(DTS)DL 0.6% 수준, 부산(DTS)·경기(ARIMA)가 0.5% 수준이었으며, 그 외 지역들은 0.4% 이하로 매우 안정적인 상태를 보이고 있다.

그리고 시계열자료에서 종속구조를 검정하기 위하여 즉, 여기서는 오차항의 독립성-종속성 여부를 검정하기 위한 방법으로 Durbin-Watson 검정을 실시하였다. 이 값에 대한 검정은 DW 값이 0에 근접하면 양의 자기상관, 2에 근접하면 자기상관이 없으며, 4에 근접하면 음의 자기상관이 존재한다고 본다[21]. 분석결과 모든 시계열 데이터가 2에 근접하여 자기상관이 없는 것으로 나타났다.

최초 연구에서는 모형의 적합도를 파악하기 위해 6개의 시계열 예측모형을 적용하기를 제안하였으나 분석결과 SMA, SES, DMA의 경우 실제 값과 예측을 위한 모형과의 적합도가 현격하게 낮아 이를 제외한 3개 모형 DES, DTS, ARIMA만을 [표 1]에서 적합도 검증 결과로 제시한다. 또한, 각 예측모형 적용을 통한 예측 값 산출에 있어서 모수추정과 모형 간 비교를 위해 모수 값을 [0, 1]의 범위에서 임의 지정하는 경우가 있다(경북). 이 경우는 예측 값의 변동이 매우 미약하거나 변동이 발생하지 않는 경우들이 있었다. 이때는 오차추정 및 적합도 검증에서는 최적이지만 차선모형을 최적모형으로 고려하는 것이 보다 효과적인 것이다.

[표 1]에서는 실측데이터에 대한 최적모형의 모수추정과 오차추정 및 모형적합을 크게 3가지 예측모형(ARIMA, DTS, DES)을 중심으로 제시하고 있다. 오차추정 및 모형적합에서 2015년 예측값과 실측치의 상대적 비교를 통한 상대오차 계산결과 전체 11개 시도 중 강원(1.19%), 충북(1.45%), 전북(1.13%), 경남(1.15%),

표 1. 시계열 모형 검정 및 예측

시도	최적모형상대오차 (%: 2015 기준), 년 평균 증감(%)	최적모형 모수추정			오차추정 및 모형적합				
		최적	2nd	3rd	지표	최적	2nd	3rd	
서울	· 예측: 306,305 · 실측: 308,306 · 상대오차: 0.65% · 평균증감: -3.859%	ARIMA(1,2,2) AR(1): -0.8625(0.0919) MA(1): -1.34(0.1121) MA(2): -0.6510(0.1049)	DTS alpha: 0.999 beta: 0.999 phi: 0.888	DES alpha:0.999 beta:0.999	RMSE	7,580	9,174	9,405	
					MAD	5,214	6,426	6,633	
					MAPE	1.36%	1.69%	1.75%	
					Theil' s U	0.3473	0.4479	0.4468	
					DW	1.81	1.25	1.28	
					AIC	17.99	■	■	
					BIC	18.10	■	■	
					예측 (2016 - 2020)	2016 *	297,338	297,555	296,206
						2017 *	286,618	287,999	284,107
						2018 *	275,684	279,507	272,008
2019 *	264,935	271,959	259,910						
2020 *	254,027	265,250	247,811						
부산	· 예측: 110,493 · 실측: 109,773 · 상대오차: 0.66% · 평균증감: -2.575%	DTS alpha:0.999 beta:0.999 phi:0.7874	DES alpha: 0.9682 beta : 0.9990	ARIMA(0,2,0)	RMSE	5,041	5,314	5,363	
					MAD	3,585	3,947	3,999	
					MAPE	2.54%	2.79%	2.79%	
					Theil' s U	0.5464	0.5417	0.5275	
					DW	1.89	1.97	2.11	
					AIC	■	■	17.17	
					BIC	■	■	17.17	
					예측 (2016 - 2020)	2016 *	105,200	103,880	103,966
						2017 *	101,600	98,005	98,159
						2018 *	98,764	92,129	92,352
2019 *	96,532	86,254	86,545						
2020 *	94,774	80,379	80,738						
경기	· 예측: 446,522 · 실측: 444,554 · 상대오차: 0.44% · 평균증감: -0.837%	ARIMA(1,1,0)	DTS alpha:0.999 beta:0.999 phi:0.7853	DES alpha: 0.9959 beta: 0.9990	RMSE	8,766	8,773	9,277	
					MAD	6,569	6,604	6,583	
					MAPE	2.63%	2.65%	2.63%	
					Theil' s U	0.5576	0.5666	0.5411	
					DW	1.79	1.78	1.96	
					AIC	18.20	■	■	
					BIC	18.24	■	■	
					예측 (2016 - 2020)	2016 *	438,160	438,158	436,389
						2017 *	433,139	433,134	428,222
						2018 *	429,197	429,188	420,055
2019 *	426,102	426,089	411,888						
2020 *	423,672	423,655	403,722						
강원	· 예측: 55,490 · 실측: 54,836 · 상대오차: 1.19% · 평균증감: -2.465%	ARIMA(0,2,2) MA(1): -0.6520(0.1099) MA(2): -0.6172(0.1103)	DTS alpha:0.999 beta:0.999 phi:0.9035	DES alpha:0.999 beta:0.999	RMSE	1,257	1,485	1,521	
					MAD	1,005	1,137	1,203	
					MAPE	1.57%	1.80%	1.91%	
					Theil' s U	0.2966	0.4666	0.4610	
					DW	2.32	1.25	1.30	
					AIC	14.35	■	■	
					BIC	14.43	■	■	
					예측 (2016 - 2020)	2016 *	53,457	53,809	53,699
						2017 *	52,187	52,880	52,561
						2018 *	50,917	52,041	51,423
2019 *	49,647	51,283	50,286						
2020 *	48,377	50,599	49,148						

시도	최적모형 상대오차 (%: 2015 기준), 년평균 증감(%)	최적모형 모수추정			오차추정 및 모형적합							
		최적	2nd	3rd	지표	최적	2nd	3rd				
충북	· 예측: 57,043 · 실측: 56,227 · 상대오차: 1.45% · 평균증감: -1.910%	DTS alpha: 0.9990 beta: 0.9990 phi: 0.8606	DES alpha:0.999 beta:0.999	ARIMA(0,2,0)	RMSE	1584	1641	1652				
					MAD	1266	1307	1315				
					MAPE	2.21%	2.28%	2.26%				
					Theil's U	0.4809	0.4654	0.4560				
					DW	1.46	1.56	1.55				
					AIC	■	■	14.82				
					BIC	■	■	14.82				
					예측 (2016-2020)				2016 [†]	54,769	54,532	54,530
									2017 [†]	53,513	52,837	52,833
									2018 [†]	52,433	51,141	51,136
				2019 [†]	51,503	49,445	49,439					
				2020 [†]	50,703	47,750	47,742					
충남	· 예측: 74,104 · 실측: 73,563 · 상대오차: 0.74% · 평균증감: -0.822%	DTS alpha: 0.9990 beta: 0.6728 phi: 0.8177	DES alpha:0.999 beta:0.6092	ARIMA(0,2,1)	RMSE	6,467	6,733	6,799				
					MAD	3,642	3,898	3,960				
					MAPE	3.77%	4.03%	4.04%				
					Theil's U	0.6212	0.6311	0.6303				
					DW	2.00	1.96	1.96				
					AIC	■	■	17.69				
					BIC	■	■	17.73				
					예측 (2016-2020)				2016 [†]	72,613	72,370	72,371
									2017 [†]	71,836	71,176	71,179
									2018 [†]	71,200	69,982	69,986
				2019 [†]	70,680	68,788	68,794					
				2020 [†]	70,255	67,594	67,602					
전북	· 예측: 72,349 · 실측: 71,544 · 상대오차: 1.13% · 평균증감: -2.054%	DTS alpha:0.999 beta:0.999 phi:0.999	DES alpha:0.999 beta:0.999	ARIMA(0,2,0)	RMSE	2,095	2,140	2,156				
					MAD	1,703	1,752	1,766				
					MAPE	1.87%	1.93%	1.92%				
					Theil's U	0.3813	0.3760	0.3697				
					DW	1.68	1.75	1.75				
					AIC	■	■	15.35				
					BIC	■	■	15.35				
					예측 (2016-2020)				2016 [†]	69,817	69,656	69,655
									2017 [†]	68,236	67,768	67,766
									2018 [†]	66,789	65,880	65,877
				2019 [†]	65,466	63,991	63,988					
				2020 [†]	64,254	62,103	62,099					
전남	· 예측: 67,712 · 실측: 68,407 · 상대오차: 1.02% · 평균증감: -1.972%	ARIMA(0,2,2) MA(1): 0.0603(0.1025) MA(2): 0.6998(0.1022)	DTS alpha:0.999 beta:0.999 phi: 0.6341	DES alpha: 0.999 beta:0.7655	RMSE	9,121	9,668	10,667				
					MAD	4,461	4,853	4,798				
					MAPE	3.95%	4.37%	4.31%				
					Theil's U	0.5897	0.6578	0.7144				
					DW	1.94	1.7	1.79				
					AIC	18.32	■	■				
					BIC	18.40	■	■				
					예측 (2016-2020)				2016 [†]	67,592	67,247	66,711
									2017 [†]	66,298	66,511	65,015
									2018 [†]	65,004	66,044	63,319
				2019 [†]	63,710	65,748	61,623					
				2020 [†]	62,416	65,561	59,927					

시도	최적모형 상대오차 (%: 2015 기준), 년평균 증감(%)	최적모형 모수추정			오차추정 및 모형적합			
		최적	2nd	3rd	지표	최적	2nd	3rd
경북	· 예측: 90,288 · 실측: 90,238 · 상대오차: 0.06% · 평균증감: -0.355% (DTS)	ARIMA(0,1,1)	DTS [*] alpha:0.999 beta:0.999 phi:0.5472	DES alpha:0.999 beta:0.999	RMSE	11,205	10,374	12,740
					MAD	5,995	6,615	5,175
					MAPE	4.50%	4.97%	3.88%
					Theil' s U	0.7547	0.7243	0.8399
					DW	1.65	1.91	1.95
					AIC	18.53	■	■
					BIC	18.57	■	■
					2016 ' 2020 '	90,199	89,094	88,144
					2017 ' 2020 '	90,199	88,468	86,051
					2018 ' 2020 '	90,199	88,126	83,958
2019 ' 2020 '	90,199	87,938	81,865					
2020 ' 2020 '	90,199	87,836	79,772					
경남	· 예측: 121,459 · 실측: 120,084 · 상대오차: 1.15% · 평균증감: -0.871%	DTS	ARIMA(1,2,1)	DES alpha:0.999 beta:0.7493	RMSE	6,605	6,703	7,088
					MAD	4,219	4,214	4,274
					MAPE	3.24%	3.18%	3.27
					Theil' s U	0.6071	0.5815	0.5819
					DW	2.00	1.99	1.91
					AIC	■	17.70	■
					BIC	■	17.78	■
					2016 ' 2020 '	117,571	117,824	116,492
					2017 ' 2020 '	115,894	116,543	112,900
					2018 ' 2020 '	114,775	115,896	109,307
2019 ' 2020 '	114,028	115,662	105,714					
2020 ' 2020 '	113,530	115,695	102,122					
제주	· 예측: 22,557 · 실측: 22,984 · 상대오차: 1.86% · 평균증감: -0.317%	ARIMA(1,2,0)	DTS	DES alpha:0.999 beta:0.999	RMSE	479	510	523
					MAD	370	400	407
					MAPE	1.66%	1.83%	1.86%
					Theil' s U	0.3397	0.3937	0.3842
					DW	1.94	1.13	1.18
					AIC	12.39	■	■
					BIC	12.42	■	■
					2016 ' 2020 '	22,851	22,769	22,746
					2017 ' 2020 '	22,762	22,576	22,508
					2018 ' 2020 '	22,690	22,402	22,270
2019 ' 2020 '	22,625	22,245	22,032					
2020 ' 2020 '	22,563	22,103	21,794					

제주(1.86%), 전남(1.02%) 6개 지역을 제외한 5개 시도가 예측 값과 실측 값 간에 1% 이하의 상대오차(relative error)를 나타내고 있다. 비록 상대오차가 1%를 상회하는 지역들이 다수 발생하였으나 대부분 약 1% 수준에 머무르고 있으므로 각 시도의 2016-2020년까지의 예측에 대해서 적용된 시계열 예측모형은 높은 수준의 정밀도를 보이고 있다라고 볼 수 있다. 또한, 과거 데이터와 최적 모형을 통해 예측된 기간(2016-2020) 동안의 년평균 증감률이 있어서 서울을 제외한 대부분의 지역이

약 -2% 수준이며, 서울의 경우 약 -4%로 매우 높은 수준의 고등학교 재학생수 감소가 예상된다.

V. 정책적 함의

교육에 필요한 질 높은 교원의 충원은 학생들에게 양질의 교육 서비스를 제공함은 물론 선진국 수준의 교육 여건을 확보 한다는 측면에서 의미가 있다. 그리고 이

러한 관계는 정책의 실현성 측면에서 학생 수의 증감과 중요하게 연계된다. 또한, 지역에서의 교원충원과 편중 역시 학생 수 증감 추세가 중요한 결정 요인이라고 볼 수 있다. 우리의 경우 2000년대 중반을 기점으로 저출산 현상에 대한 우려가 가속화되면서 교원수급정책이 학생 수 감소에 중요하게 영향을 받는다고 볼 수 있다.

이러한 측면에서 학생 수의 증감은 그 자체가 교육 서비스 제공의 질적 측면을 결정하는 필요조건 일뿐 아니라 국가의 핵심 성장역량인 인적자원의 양성에도 중요하게 영향을 미친다. 따라서 학령인구의 증감 추세에 대해서는 매우 주의 깊게 추적관찰 하는 동시에 이를 정책형성에 핵심변수로 개입시킴으로써 관련 정책의제 실행의 적합성과 운영방안의 실현가능성을 확보할 필요가 있다. 본 연구의 가치와 한계점을 통해 세 가지 측면에 대한 성찰적 숙고와 실천적 방안 탐색이 필요하다.

1. 데이터 기반의 최적 예측모형 산출

본 연구에서 다루는 시계열 예측방법은 주어진 데이터를 기반으로 미래 시점의 증감 추세를 예측하는 방식이다. 특히, 과거 일정기간 동안의 데이터의 평균, 전년 대비 증감, 추세 존재의 통계적 유무검정 등 일반적 현황과 기술통계치, 그리고 Ljung-Box 검정과 자기상관검정 등 시계열 내의 연관관계 검토뿐 아니라 모형오차검토를 다양한 시계열 예측모형(SMA, SES, DMA, DES, DTS, ARIMA)과 오차추정 지표들(MSE, RMSE, NAD, MAPE, Theil's U, BIC, AIC)을 통해 수행하였다.

이러한 복합적 검토는 과거의 추세를 미래에 반영하는 시계열 예측방법이 지닌 장점을 최대한 부각하고, 한계점을 보완한다는 측면에서 매우 의미가 있다고 보여진다. 즉, 주어진 데이터를 기반으로 미래추세에 대한 최적의 예측모형을 산출 할 수 있다. 본 연구에서는 특히 주요 11개 시도에 대한 고등학교 학생 수의 증감 패턴을 토대로 향후 5년간의 시계열 예측을 시행함으로써 지역에 따라 편중이 발생하는 부분과 시계열 모형의 적합도를 동시에 고려하였다. 분석결과 과거 데이터에 대한 적합도 측면에서 크게 3가지 모델이 고려되었

는데 ARIMA모델의 경우 서울, 경기, 강원, 전남, 경북, 제주가 DTS는 부산, 충북, 충남, 전북, 경남이 적합한 것으로 나타났다. 2015년 데이터를 기준으로 실측값과 예측 값 간의 상대오차는 대부분의 지역에서 1% 수준으로 나타났기에 데이터에 대한 예측모형의 적합도가 높은 수준이라고 볼 수 있다.

그러나 본 연구는 시계열 데이터의 추세를 기반으로 미래를 예측하는 방식이다. 따라서 데이터의 추세·기간·형태에 따라 모델이 적합도가 달라질 수 있다. 이러한 시계열 추세 분석 관점에서의 데이터 기반접근은 현실 대응성 차원에서 의미가 있을 수 있으나 단일변수를 기반으로 하기에 결과 자체를 장기적 예측에 적합하다고 주장하기에는 경험적 적절성 확보 차원에서 다소 무리가 있을 수 있다. 이는 지역의 고등학교 재학생이라는 변수가 지닌 특성이 여러 가지 배경요인들에 의해 결정되는 이유뿐 아니라 ARIMA 모형과 같은 특정 분석기법의 적용이 구축된 데이터의 지원이 어려운 경우 그 효과성을 반감시키기 때문이다. 물론 이러한 이유들로 인해 여기서는 16개 시도 중 분석을 위한 충분한 데이터 구축이 가능한 11개 지역만을 대상으로 하였다. 즉, 단기적 시각에서 학생 수 변동에 따라 조직운영상에 직·간접적인 영향을 받는 대학조직의 경우 예측모형에 따른 다양한 예측결과를 살펴 볼 수 있으며, 중장기적 교육정책의 패러다임과 실제 정책입안을 고려해야 하는 정부기관의 경우 예측결과는 물론 분석의 전체적 방향성과 프로세스에 보다 의미를 둘 필요가 있다고 사료된다.

2. 역동적 추세변화를 통한 체계적이고 통합적인 정책설계

최근 교육환경의 급격한 변화로 인해 지역 간, 학교 간 교육 불균형이 심각해지고 있다. 특히, 지속적인 학령인구의 감소로 인하여 일부 지역에서는 학령인구의 부족현상으로 학교 통폐합에 대한 논의가 나오고 있으며, 다른 지역에서는 과밀학급의 문제가 발생하고 있다. 이뿐 아니라 국가의 교육여건을 OECD 기준으로 충족시키기 위하여 가장 중요하게 고려되는 교원 정원을 조정하고 각 지역과 학교의 필요에 따라 적절히 배정하는

것이 중요하다.

적정 수준의 교원을 지역별로 배정하는 것은 교육시스템의 효율적 설정과 운용을 위해 중요하게 관리되어야 하는 부분이다. 따라서 학생 수의 정확한 추계는 향후 우리나라 공교육의 기본 여건을 가능하고 이를 바탕으로 향후 교육여건의 개선 및 과제 제시에 대한 방향성을 제시할 수 있는 자료로써 그 중요성이 인정된다. 특히, 대부분의 대학 입장에서 학생자원을 통한 등록금 의존이 높기에 이러한 고등학생 수의 변화와 추세를 통해 전략적으로 조직운영계획을 수립하고 관리할 필요가 있다. 즉, 대학 입학자의 지역별 추계를 파악하고 이를 통해 조직에서 어느 수준까지 가용자원과 시간, 노력 등을 집중할지에 대한 전략을 세부적으로 수립할 필요가 있는 것이다.

또한, 정책을 설계하는데 있어서도 이러한 데이터의 추세와 변화가 의사결정을 하는데 중요 요소로서 기능할 수 있도록 통합적인 의사결정 지원모형(Strategic Decision Making Support System)을 설계할 필요가 있다. 이러한 모형의 적용으로부터 조직은 복잡적이고 복잡한 문제의 해결에 있어서 이를 효과적으로 수행하기 위한 정형화된 틀을 고려할 수 있으며, 지식을 효과적으로 관리하고 문제해결 프로세스와 기법들을 조합하는 노력을 경주 할 수 있다.

3. 지역의 특성을 고려한 시스템 모델링 필요

연구에서 다루고 있는 11개 지역의 재학 고등학생 수의 경우 고등교육기관 또는 정부기관에서 정책적 방향성을 설정하기 위해 고려해야 할 직접적 변수이기 보다는 배경요인으로 기능하는 변수이다. 특히 우리나라의 경우 지역의 특수성과 교육환경의 높은 변동으로 인해 지역 간·학교 급간 교육 불균형이 심각한 상황이다. 따라서 각 지역에 따라 존재하는 교육환경 요인의 불균형을 완화하고, 교육의 질적 제고를 위해서는 변화하는 목표 값에 영향을 미치는 변수들을 다양하게 고려하는 동시에 이것의 비선형성과 순환성을 고려할 필요가 있다. 즉, 중단 데이터가 지니는 특성과 시스템 모형 내에서 다양한 변수 값들의 기능을 복합적으로 고려하는 시스템 모델링이 필요하다.

앞으로 교육환경에 있어서 많은 변화가 발생하고 이에 따라 구성원의 행태와 의지, 그리고 만족도와 같은 심리적 요인들이 목표 값에 상당한 영향을 미칠 가능성이 크다. 이는 행위자의 의도된 주관성이 구조를 형성하기에 선형성을 기대하기 어렵고, 부분과 전체의 구조 변화가 다시 행위자의 행태와 의지에 영향을 미치기에 상호작용성과 순환구조이기 때문이다.

본 연구에서는 비록 이러한 시스템 모델링을 위한 기준이 되는 변수 값인 고등학교 재학 학생 수만을 시계열 예측기법으로 분석하고 있지만 향후 이를 시스템 모델 내에 개입시켜 지역의 교육환경 불균형과 같은 문제를 해결하기 위한 전략을 개발하고자 할 경우 시스템 내부의 체제화 과정에서 진화적 발전과 공진화(Coevolution)에 대한 고려가 수반 되어야 할 것으로 사료된다.

참 고 문 헌

- [1] 이영, 2011-2020 중장기 교원수급계획 수립 및 교원양성기관 규모 적정화·내실화 방안 마련을 위한 정책연구, 교육부정책과제, 2011.
- [2] 엄문영, "학생수 전망과 교육재정 기획," 한국교육행정학회 소식지, 제115권, pp.10-14, 2013.
- [3] 이상준, 윤관호, 김성윤, "교육서비스 만족도 지수 모형개발 및 IPA를 통한 개선방안," 한국콘텐츠학회논문지, 제13권, 제11호, pp.510-521, 2013.
- [4] 오영호, "2025년까지의 우리나라 간호사 인력의 수요 및 공급 전망," 보건경제와 정책연구, 제16권, 제3호, pp.139-161, 2010.
- [5] 최선운, "국내 봉제산업 인력양성 실태와 전망," 한국의류학회지, 제9권, pp.46-55, 2012.
- [6] 김정순, 정인숙, 김명수, "부산지역 실버서비스 및 관련 인력수요에 대한 요구도," Journal of Korean Public Health Nursing, 제21권, 제2호, pp.182-192, 2007.
- [7] 마기중, 이학준, 이종훈, 이혜정, "안경사의 인력수급 전망," 대한시과학회지, 제3권, 제1호, pp.25-40,

2001.

[8] 김종태, “기초자치단체의 학생수 추계를 위한 알고리즘,” 한국데이터정보과학회지, 제22권, 제6호, pp.1167-1173, 2011.

[9] 조대현, 이상일, “이지역 코호트-요인법을 이용한 부산광역시 장래 인구 추계,” 대한지리학회지, 제46권, 제2호, pp.212-232, 2011.

[10] 김현철, *시계열자료의 분석과 예측. 교육·심리·사회 연구방법론 총서시리즈 <22>*, 서울: 교육과학사, 2005.

[11] 김병주, 오영수, “교원의 수요와 공급 예측치 비교,” 한국교육행정학회, 제20권, 제4호, 2002.

[12] 조동섭, *초등교원 증장기 수급 계획 및 안정적인 충원 방안 연구*, 교육인적자원부, 2003.

[13] 차양은, 서지영, 이병렬, “중등교원 수요와 공급의 미래예측,” 교육행정학 연구, 제21권, 제3호, pp.297-316, 2003.

[14] 윤용화, 김종태, “비선형 회귀모형을 이용한 학년별 학생수 추계,” Journal of the Korean Data & Information Science Society, 제23권, 제1호, pp.71-77, 2012.

[15] 송필준, 김종태, “로지스틱함수모형과 비례이동 평균모형에 의한 학생 수 추계와 분석,” Journal of the Korean Data & Information Science Society, 제21권, 제3호, pp.503-511, 2010.

[16] 이영, 한유경, 김이경, “초중등 장래 학생수 및 교원수 추정에 기반한 교원수급정책의 방향,” 교육재정경제연구, 제22권, 제3호, pp.59-79, 2013.

[17] 공은배, 김용남, 엄문영, 이선호, *지방교육재정 확충을 위한 연구*, 한국교육개발원, 2013.

[18] 강성국, *한국교육 60년 성장에 대한 교육지표 분석*, 경제·인문사회 연구원, 2005.

[19] 김기환, 이창호, 최보승, “학령인구 감소에 따른 지역별 대입 지원자 감소에 대한 예측연구,” 한국데이터정보과학회지, 제26권, 제6호, pp.1175-1188, 2015.

[20] 김연형, *시계열예측*, 서울: 형설출판사, 2002.

[21] Johnathan Mun, *시물레이션을 이용한 미래형*

리스크 분석, 경기:(주)이레테크, 2005.

[22] 노형진, *SPSS/Excel에 의한 재미있는 시계열분석*, 서울: 도서출판 호산, 2007.

[23] 서정민, “시계열 모형을 활용한 사회서비스 수요·공급모형 구축: 발달재활 서비스를 중심으로,” 한국콘텐츠학회논문지, 제15권, 제6호, pp.399-410, 2015.

[24] https://docs.oracle.com/cd/E24674_01/epm.1112/cb_predictor_user/frameset.htm?apcs09s02.html

[25] C. D. Lewis, *Industrial and business forecasting methods*, London, Butterworths, 1982.

[26] H. Theil, *Applied Economic Forecasting*, Amsterdam, North Holland, 1966.

저 자 소 개

임 성 범(Seong-Bum Lim)

정회원



- 2011년 8월 : 경희대학교 행정학과(행정학박사)
- 2014년 1월 ~ 현재 : 단국대학교 행정학과 연구전임 조교수

<관심분야> : 계량분석, 교육서비스 만족도

박 선 형(Sun-Hyung Park)

정회원



- 1998년 10월 : University of Tasmania 교육행정 전공(Ph.D)
- 2008년 3월 ~ 현재 : 동국대학교 교육학과 교수

<관심분야> : 교육행정철학, 조직이론과 지도성, 조직창의성과 혁신, 교사교육