

## 위계적 선형모형을 이용한 초등학교 학부모의 자녀의 학교여건 만족도 영향 분석 - 안양시 사례

김호일<sup>1</sup> · 전희주<sup>2</sup>

<sup>1</sup>안양대학교 정보통계학과

<sup>2</sup>부산외국어대학교 데이터경영학과

접수 2010년 7월 23일, 수정 2010년 9월 26일, 게재확정 2010년 10월 30일

### 요약

본 연구에서 활용하는 안양시 교육 자료는 학생과 학교간이 위계적인 수준의 형태이기 때문에 상위 수준의 특성을 반영하는 위계적 선형모형을 이용하여 안양시 초등학교 학부모의 자녀가 다니는 학교에 대한 제반 여건들에 대한 만족도에 대한 영향요인 분석을 하였다. 그 결과 부모들이 느끼는 자녀의 학교여건 만족도에 부모들의 시 교육정책에 대한 만족 정도, 학군선호의 신도시 지역을 나타내는 학교의 거주지가 그 다음 크게 영향을 미쳤다. 위계적 선형모형의 분석 결과를 종합해 보면, 사교육 없이 공교육만으로 학생들에 대한 교육이 충분히 가능하도록 하고 신도시와 같이 학교의 제반 교육환경을 개선하는 방향으로 안양시의 교육정책을 펼쳐 나간다면 학부모의 자녀가 다니는 학교에 대한 제반 여건들에 대한 만족도는 지금보다는 더욱 향상될 것으로 보인다.

주요용어: 고정효과, 분산 공분산 행렬, 위계적 선형모형, 임의효과.

### 1. 서론

사회학 뿐 아니라 특히 교육학에서 어떤 특성에 영향을 주는 영향 변수 확인과 그에 대한 변화의 연구는 전통적으로 회귀분석 모형을 기초로 적용하여 왔다. 그러나 조사 대상이 되는 각 개인은 개인마다 개인 고유의 특성을 가지고 있지만 또한 어떤 조직이나 단체에 속해 있으면서 그 조직이나 단체에 영향을 받을 수 있다. 특히 교육학에서는 학생들의 학습능력에 대한 차이가 개인적인 능력 뿐 아니라 학교의 특성화, 학교의 분위기, 학교의 시설과 같은 학교차원에서의 변수에 대해 매우 많은 영향을 받을 수 있다. 이는 학교 간의 차이가 있다는 것을 의미할 수 있다. 예를 들어 학생이나 학부모에 대한 특성은 학교, 지역사회, 학군 등은 그 조직의 특성이 다르게 나타날 수 있을 것이다. 즉 한 학교의 학생들은 다른 학교의 학생들과는 구별되는 특성을 가지고 있을 수 있는 것이다. 그래서 고정효과만의 설명변수로 구성된 회귀모형이나 분산분석모형은 그룹간의 단순한 비교 또는 자료의 위계적 구조를 고려하지 않은 상태에서 변인간의 인과관계에 초점을 두기 때문에 조직특성의 영향을 받는 변인들이 상관관계가 존재한다는 사실을 고려하지 않고 각 개인이 독립적이라고만 가정하여 모형을 구성하게 된다. 따라서 이러한 전통적인 모형은 개개인의 특성, 조직과 개인 간의 상호작용, 그리고 조직의 특성을 동시에 고려한 반응변수를 예측하는 데에는 적절한 방법이 아니게 된다.

<sup>1</sup> (430-714) 경기도 안양시 만안구 안양 5동, 안양대학교 정보통계학과, 교수.

<sup>2</sup> 교신저자: (608-738) 부산시 남구 우암동, 부산외국어대학교 데이터경영학과, 전임강사.  
Email: hjchun@pufs.ac.kr

따라서 과거에는 이런 학교 간의 차이를 무시하고 일률적으로 학교 간의 차이를 동일한 조건으로 간주하여 조사 및 분석하였다. 이는 학교 내에 있는 학생들의 개인적인 능력만 고려하고 학교나 동간의 차이는 고려되지 않고 분석되어 왔다. 이는 같은 학교 내 학생들 간의 동질성은 무시된 채 분석이 되어 왔다는 의미이다. 그러나 학교 간의 차이를 통계적으로 적절히 통제하는 모형의 필요성이 제기되어 왔다. 이러한 집단 내 유사한 개체들이 모여 있는 경우의 자료를 위계적 자료 (hierarchical data) 또는 다층 자료 (multilevel data)라고 하고 이러한 선형모형에 대한 분석을 위계적 선형모형 분석이라고 한다. 그러므로 예측을 하거나 유의한 영향변수를 찾는 데 있어서 이와 같은 위계적 관계가 존재하는 자료분석에는 개인 고유의 특성뿐만 아니라 그 상위조직의 특성에 대한 변인을 포함하여야 할 것이다. 이와 같은 위계적 선형모형 분석들은 교육학뿐만 아니라 여러 많은 사회과학 분야의 계층화되어 있는 자료를 분석하는데 많이 사용되어져 왔다. 즉 각 수준의 대해 변이의 양을 측정하고 (학교간의 분산 또는 학생간의 분산) 두 수준의 변수의 함수로 학생수준과 학교수준의 결과를 얻어내고 두 수준의 효과간의 상호작용을 평가하는 게 분석의 주 목적이 된다. 따라서 본 연구는 안양시에서 조사한 결과의 자료를 사용하여 이층 자료 경우에 위계적 선형모형을 이용하여 학부모가 자녀 학교 여건에 대한 만족도에 미치는 영향을 분석하고자 한다.

본 연구에서 적용하는 안양시 자료는 학생과 학교간이 위계적인 자료로 상위 수준의 특성을 반영하는 위계적 선형모형을 이용하여 안양시 초등학교 학부모의 자녀가 다니는 학교에 대한 제반 여건들에 대한 만족도에 대한 영향요인 분석을 하였다.

참고적으로 박찬선 등 (2007)은 학업성취도가 공간적 영향과 지리적 특징에 의해 영향을 미친다고 연구하였다. 위 논문에 의하면 자체 연구용 평가지를 만들어 영어, 수학원점 수를 표준점수화하여 이 둘을 합산한 후 상위 30%를 상위권 하위 30%를 하위권, 나머지를 중위권으로 하였을 때 안양시 전체 31개 동 중 상위권 분포비율이 가장 높은 동은 동안구의 범계동으로 52.4%이며 다음으로 귀인동 (40.0%), 평안동 (37.9%)이다. 이 3개동은 모두 안양시 신시가지라고 할 수 있는 평촌지역에 해당한다. 평촌신시가지 지역은 대부분 고층의 대규모아파트가 밀집되어 있으며 따라서 학원 또한 많이 밀집되어 있다. 따라서 사회·경제 변수가 되는 부모의 직업, 주거형태 등 다양한 변수를 통해 학업성취도가 영향을 준다는 결론을 내렸다.

본 논문은 2장에서 위계적 선형모형에 대한 이론적인 배경을 간단히 정리하고, 3장에서는 안양시 자료를 가지고 부모의 자녀 학교여건에 대한 만족도에 영향을 주는 요인들을 위계적 선형모형을 통해 분석한다. 그리고 4장에서는 위계적인 선형모형을 통해 분석한 결과와 학부모의 자녀 학교 제반 여건들에 대한 만족도를 향상시키기 위한 방안을 제안하는 맺음말로 정리를 한다.

## 2. 위계적 선형모형의 이론적 배경

위계적 선형모형 (hierarchical linear model)는 용어는 교육 사회학 분야에서 많이 사용되고 있으며 다층 선형모형 (multilevel linear model)으로도 불리며 (Goldstein, 1995; Raudenbush와 Bryk, 2002), 생명과학에서는 혼합효과모형 (mixed effects model)으로 (Singer, 1998), 통계학 문헌에서는 임의계수 모형 (random coefficient model) 또는 경험적 베이지모형 (Wolfinger, 1996)의 영역에 해당된다고 할 것이다. 국내에서는 주로 교육학 분야에서 위계적 선형모형이 많이 연구되고 있다 (강상진, 1998; 김경성, 2007; 서민원, 2003; 김경화와 김영화, 2006; 유정진, 2006; 백순근과 신호정, 2008). 그 이외 위계적 모형은 여러 개체에 대해 반복 측정된 경우에도 가능하다 (김호일, 2005).

위계적 선형모형은 특히 자료의 특성에 기초하여 명명된 용어로 측정된 자료가 모두 지분화 (nested) 되어 있거나 위계적 (hierarchical)인 특징을 갖는다. 먼저 다층 자료의 통계적 모형을 이해하기 위해 본 논문에서는 설명변수로서 하나 이상의 연속형 변수인 공분산변수가 가능한 계수들로부터 임의로 선택된

임의계수 (random coefficient)로 가정하여 설명할 것이다. 임의계수는 대상 또는 집단내의 회귀계수가 모집단 회귀모형으로부터 임의편차 (random deviation)이 있는 것으로 볼 수 있다. 위계적 회귀모형은 각 대상에 대해 임의절편 (random intercept)와 임의기울기 (random slope)를 포함한다. 본 논문에서는 단순 2층 모형의 예를 가지고 위계적 선형 모형에 대한 이론적 설명을 하고자 한다. 먼저 제 1수준인 학생에 대한 모형은 다음과 같은 회귀모형으로 표현한다.

$$y_{ij} = \beta_{0j} + \beta_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij}, \quad j = 1, 2, \dots, t, \quad i = 1, 2, \dots, n_j \quad (2.1)$$

여기서  $y_{ij}$ 는  $j$ 번째 조직에 속해 있는  $i$ 개인의 반응변수를 의미하고  $\beta_{0j}$ 와  $\beta_{1j}$ 는 각각 회귀모형의 상수와 기울기로 제 2수준 모형의 학교에 따라 다르게 나타남으로 임의값 (random value)으로 확률변수가 된다. 그래서  $\beta_{0j}$ 와  $\beta_{1j}$ 는 전체평균과 임의오차 (random error)의 표현하는 제 2수준 모형은 다음과 같다.

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \end{aligned} \quad (2.2)$$

그래서  $\beta_{0j}$ 와  $\beta_{1j}$ 의 분포는 다음과 같이 가정한다.

$$\begin{pmatrix} \beta_{0j} \\ \beta_{1j} \end{pmatrix} \sim iid, \quad N \left[ \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \end{pmatrix}, \Psi \right], \quad e_{ij} \sim iidN(0, \sigma^2), \quad \Psi = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}.$$

(2.2)의 제 2수준 모형을 (2.1)의 제 1수준 모형에 대입을 하면 혼합모형은 다음과 같이 표현된다.

$$\begin{aligned} y_{ij} &= \gamma_{00} + \gamma_{10}(x_{ij} - \bar{x}_{.j}) + u_{0j} + u_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij} \\ \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} &\sim iidN \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Psi \right], \quad e_{ij} \sim iidN(0, \sigma^2). \end{aligned} \quad (2.3)$$

모형 (2.3)에서  $\gamma_{00} + \gamma_{10}(x_{ij} - \bar{x}_{.j})$ 은 고정효과 (fixed effect) 부분이고  $u_{0j} + u_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij}$ 은 임의효과 (random effect) 부분이 된다. 그래서 위계적 선형모형인 (2.3)은 고정효과 계수와 임의효과 계수가 혼합된 혼합모형으로 표현할 수 있는 것이다. 결국 모형 (2.3)은

$$y_{ij} = \gamma_{00} + \gamma_{10}(x_{ij} - \bar{x}_{.j}) + e^*_{ij}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, n_i \quad (2.4)$$

여기 식 (2.4)에서  $e^*_{ij} = u_{0j} + u_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij}$ 는 오차항이 된다.

이것은  $E(y_{ij}) = \gamma_{00} + \gamma_{10}(x_{ij} - \bar{x}_{.j}), \text{Var}(y_{ij}) = [1, x_{ij}] \Psi \begin{bmatrix} 1 \\ x_{ij} \end{bmatrix} + \sigma^2$ 가 된다.

### 3. 사례분석: 안양시의 경우

#### 3.1. 자료설명

본 자료는 안양시에서 조사한 자료로 안양지역 교육에 대한 학부모 및 학생 의견조사의 결과를 이용하였다. 표본조사는 2008년 4월 한 달 동안 실시된 것으로 조사대상은 안양시소재 초중고에 재학 중인 학부모 및 학생을 대상으로 하였으나 본 분석에는 학부모를 대상으로 한 설문지 4, 302개중 초등학교에 해당하는 2701개의 설문지를 대상으로 분석하였다. 안양시의 실제 권역은 만안구, 동안구로 나뉘어져 있으나 본 조사에서는 지역에 따라 교육수준의 차이가 존재할 것이라는 가정 하에 만안구와 구동안 그리고

평촌으로 나누어 조사를 하였다. 모든 동이 조사되어 총 31개가 조사되었으며 학교는 초등학교 40곳을 조사한 결과이다.

본 연구에서는 안양시 초등학교 학부모가 생각하는 학생의 학교여건 만족도를 묻는 각 4점 척도의 통학여건, 학교 시설, 급식, 학생들 간의 분위기, 상급학교 진학 또는 학군 등에 대한 만족도를 설정하였다. 설명변수의 회귀계수 변수로 사용되는 안양시의 교육정책에 관한 만족도 설문은 각 1점에서 4점 척도로 구성된, 어학관련 사업 관련 4문항, 시설개선 관련 7문항, 기타사업에 대한 4문항 구성된 총 15개 설문으로 구성되었다. 본 연구에서는 15개 문항을 산술평균하여 사용하였다. 설명변수로는 학교가 위치하는 만안구, 구동안구, 평촌을 나타내는 지역구분(구), 부의 연령, 모의 연령, 부의 학력, 모의 학력, 월 소득수준, 자녀성별, 맞벌이 여부를 나타내는 경제활동 대상, 월 소득 중 사교육비 비중 그리고 자녀의 학급 내 성적이 범주형 변수로 초등학교 학부모의 안양시의 교육정책에 대한 만족도가 연속변수로 고려되었다(표 3.1).

표 3.1 변수설명

		변수	
수준	변수	변수설명	
종속변수	부母的 교육만족도 (각 만족도 척도는 4)	(통학 여건만족도+학교 시설만족도+급식만족도+학생들 간의 분위기만족도+상급학교 진학 또는 학군만족도)/5	
	교육정책의만족도 (각 만족도 척도는 4)	[부母的 어학관련 만족도 (4문항) + 부母的 시설개선 만족도 (7문항)+부모 기타사업 만족도 (4문항)]/15	
독립변수	수준1	지역구분(구)	1=만안구 2=구동안구 3=평촌
		부의 나이	1=40세 이하 2=40세 49세 3=50세 59세 4=60세 이상
		모의 나이	1=40세 이하 2=40세 49세 3=50세 59세 4=60세 이상
		부의 교육정도	1=중학교 졸 이하 2=고등학교 졸 3=전문대 졸 4=대학교졸 5=대학원 졸 이상
		모의 교육정도	1=중학교 졸 이하 2=고등학교 졸 3=전문대 졸 4=대학교졸 5=대학원 졸 이상
	수준2	월 소득수준	1=1,000만원 이하 2=2,000만원 이하 3=3,000만원 이하 4=4,000만원 이하 5=5,000만원 이하 6=6,000만원 이하 7=7,000만원 이하 8=7,000만원 이상
		학생 성별	1=남자 2=여자
		맞벌이여부	1=부만 활동 2=모만 활동 3=부모모두 활동 4=활동하지 않음
		사교육비중	1=9% 이하 2=19% 이하 3=29% 이하 4=39% 이하 5=49% 이하 6=50% 이상 7=없음
		자녀학급성적	1=상위 10%이상 2=상위 11 30% 3=상위 31 60% 4=상위 61 80% 5=하위 20% 이내

반응변수인 안양시 초등학교 학부모가 생각하는 학생의 학교여건 만족도는 평균이 2.8982이고 표준편차는 0.3738이고 설명변수인 초등학교 학부모의 안양시의 교육정책에 대한 만족도의 평균과 표준편차는 2.3804와 0.4815이며 두 변수의 단순상관관계는 0.3256 ( $p$ -값=0.0001)이다. 이는 안양시 초등학교 학부모는 안양시의 교육정책에 대해 만족감이 높을수록 자녀 학교 여건에 대한 만족도 또한 높아 서로 양의 상관관계가 있음을 알 수 있다.

여기서 안양시 초등학교 학부모의 자녀 학교여건에 대한 만족도보다도 안양시의 교육정책에 대한 만족도가 더 낮아 안양시의 초등학교 학부모는 안양시의 교육정책에 대해 자녀의 학교 제반 여건보다도 만족하지 못함을 나타내고 있다.

표 3.2은 각 범주형 설명변수에 대한 기초 통계 특성과 각 설명변수의 범주별에 대한 부모가 생각하는 자녀의 학교여건 만족도와 인 안양시의 교육정책에 관한 만족도 변수에 평균과 표준편차에 대한 자료이다.

본 연구의 자료 분석은 SAS 9.1를 활용하였고 위계적 선형모형 적용을 위해 SAS PROC MIXED 프

표 3.2 기초 통계 분석

변수	빈도	%	부모의 학교만족도		교육정책 만족도		표준편차
			평균	표준편차	평균	표준편차	
지역구분 (구)	만안구	961	35.6	2.83	0.38	2.39	0.49
	구동안구	815	30.1	2.85	0.37	2.39	0.5
	평촌	925	34.3	3	0.35	2.36	0.44
부의 나이	40세 이하	693	25.7	2.87	0.36	2.38	0.46
	40세 49세	1887	69.9	2.9	0.38	2.37	0.48
	50세 59세	92	3.4	2.93	0.43	2.49	0.49
	60세 이상	29	1	2.96	0.35	2.48	0.6
모의 나이	40세 이하	1415	52.4	2.88	0.37	2.35	0.48
	40세 49세	1225	45.4	2.91	0.37	2.4	0.47
	50세 59세	30	1.1	2.89	0.38	2.44	0.42
	60세 이상	31	1.1	2.99	0.31	2.61	0.49
부의 교육정도	중학교 졸업 이하	65	2.4	2.95	0.3	2.49	0.58
	고등학교 졸업	714	16.4	2.91	0.36	2.5	0.51
	전문대 졸업	472	17.5	2.88	0.38	2.41	0.48
	대학교 졸업	1163	43.1	2.88	0.37	2.3	0.47
모의 교육정도	대학원 졸업 이상	287	10.6	2.93	0.39	2.35	0.35
	중학교 졸업 이하	68	2.5	2.99	0.29	2.53	0.51
	고등학교 졸업	1087	40.2	2.88	0.37	2.44	0.5
	전문대 졸업	561	20.8	2.88	0.38	2.37	0.46
월 소득수준	대학교 졸업	888	32.9	2.9	0.36	2.31	0.46
	대학원 졸업 이상	97	3.6	2.92	0.4	2.23	0.39
	1, 000만원 이하	60	2.2	3	0.34	2.7	0.67
	2, 000만원 이하	236	8.7	2.93	0.37	2.62	0.53
학생 성별	3, 000만원 이하	516	19.1	2.88	0.37	2.36	0.51
	4, 000만원 이하	663	24.6	2.86	0.36	2.29	0.43
	5, 000만원 이하	501	18.5	2.89	0.35	2.38	0.46
	6, 000만원 이하	324	12	2.91	0.36	2.38	0.4
	7, 000만원 이하	167	6.2	2.93	0.41	2.35	0.35
	7, 000만원 이상	234	8.7	2.92	0.4	2.35	0.53
맞벌이여부	남자	1413	52.3	2.9	0.37	2.36	0.42
	여자	1288	47.7	2.89	0.37	2.39	0.49
사교육비중	부만활동	1385	51.3	2.87	0.37	2.39	0.47
	모만 활동	53	1.9	2.89	0.36	2.63	0.59
	부모모두 활동	1234	45.7	2.91	0.36	2.34	0.47
	활동하지 않음	29	1.1	3.17	0.39	2.91	0.71
자녀학급성적	9% 이하	469	17.3	2.94	0.46	2.45	0.46
	19% 이하	1158	42.9	2.9	0.45	2.38	0.45
	29% 이하	571	21.1	2.84	0.48	2.3	0.48
	39% 이하	316	11.7	2.89	0.5	2.37	0.5
	49% 이하	94	3.5	2.95	0.47	2.46	0.47
	50% 이상	45	1.7	2.94	0.66	2.38	0.66
전체	없음	48	1.8	2.96	0.63	2.79	0.63
	상위 10% 이상	512	18.9	2.85	0.35	2.31	0.43
	상위 11 30%	1096	40.6	2.89	0.37	2.36	0.46
	상위 31 60%	883	32.7	2.9	0.37	2.39	0.49
	상위 61 80%	159	5.9	2.99	0.39	2.6	0.61
하위 20% 이내	51	1.9	3.02	0.38	2.46	0.46	
전체		2701	100	2.90	0.37	2.38	0.48

로시저를 사용하여 분석을 하였다.

### 3.2. 모형설정

#### 3.2.1. 초기 모형 설정

먼저 가장 기본적인 모형을 생각할 수 있을 것이다. 학생수준 모형 ( $y_{ij} = \beta_{0j} + e_{ij}$ )과 학교수준 모형 ( $\beta_{0j} = \gamma_{00} + u_{0j}$ )의 혼합으로  $y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$ 를 고려할 수 있다. 이 경우에 학교수준의 분산과 잔차에 대해 유의하며 학교여건의 만족도의 추정치는 총 평균과 동일한 2.8912점을 나타내었고 집단내 상관 ( $\rho$ )은 11.11%로 학교여건 만족도에서 학교 간 변동이 11.11%이며 학교 내 변동은 88.89%임을 알 수 있다.

표 3.3 기본모형의 결과

고정효과	추정치	표준오차	t-값	Pr >  t
학교여건의 만족도	2.8912	0.02102	137.55	<.0001
임의효과	분산 추정치	표준오차	t-값	Pr >  t
학교수준 ( $\tau_{00}$ )	0.0155	0.00399	3.89	<.0001
잔차	0.1243	0.00364	34.11	<.0001
집단내 상관 ( $\rho$ )	0.1111			

이제 좀 더 발전된 모형으로, 먼저 하위수준인 부모들의 자녀의 학교여건 수준 만족도에 대한 개별수준 모형으로 (2.1)과 같이 생각할 수 있을 것이다.

부모가 생각하는 자녀의 학교여건 수준 만족도는 안양시 교육정책에 대한 만족도에 의한 회귀모형의 두 모수인 절편항  $\beta_{0j}$ 과 기울기  $\beta_{1j}$ 에 의해 특성화 된다. 부모 개별수준에 대한 자녀의 학교여건 수준 만족도에 대한 예측값은 학교수준 평균에 중점을 두고 있기 때문에 절편항인  $\beta_{0j}$ 은 동수준의 평균값이 될 것이다. 여기서  $e_{ij}$ 는 독립적인 평균이 0이고 분산이  $\sigma^2$ 인 정규분포를 따른다고 가정한다.

두 모수인 절편항  $\beta_{0j}$ 과 기울기  $\beta_{1j}$ 은 학교수준에 따라 다를 것으로 보았기에 본 연구에서는 상위수준인 학교수준은 부모교육수준, 가정소득수준, 사교육비정도, 부모연령 등 변수에 의해 설명되지 않는 부분이 있음을 고려하여 보고 상위수준인 학교수준 모형을 다음과 같이 설정하였다.

$$\begin{aligned}
 \beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{지역구분})_j + \gamma_{02}(\text{부연령})_j + \gamma_{03}(\text{모연령})_j + \gamma_{04}(\text{부학력})_j \\
 &\quad + \gamma_{05}(\text{모학력})_j + \gamma_{06}(\text{월소득수준})_j + \gamma_{07}(\text{자녀성별})_j \\
 &\quad + \gamma_{08}(\text{맞벌이여부})_j + \gamma_{09}(\text{사교육비비중})_j + \gamma_{010}(\text{자녀학급성적})_j + u_{0j}, \\
 \beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{지역구분})_j + \gamma_{12}(\text{부연령})_j + \gamma_{13}(\text{모연령})_j + \gamma_{14}(\text{부학력})_j \\
 &\quad + \gamma_{15}(\text{모학력})_j + \gamma_{16}(\text{월소득수준})_j + \gamma_{17}(\text{자녀성별}) \\
 &\quad + \gamma_{18}(\text{맞벌이여부})_j + \gamma_{19}(\text{사교육비비중})_j + \gamma_{110}(\text{자녀학급성적})_j + u_{1j},
 \end{aligned} \tag{3.1}$$

여기서  $u_{0j}$ 와  $u_{1j}$ 는 평균이고 분산공분산 행렬이  $\Psi$ 인 다변량 정규분포를 따른다.  $\Psi$ 의 값들은 학교수준의 설명들을 통제된 후에  $\beta_{0j}$ 과  $\beta_{1j}$ 에 대한 오차가 될 것이다. 결국 식 (3.2)를 식 (3.1)에 대입을 하면 학생수준과 학교수준을 결합한 모형은 다음과 같이 될 것이다.

표 3.4 초기 위계적 선형 모형 결과

고정효과	분자 자유도	분모 자유도	F-값	Pr > F	
교육정책의 만족도	1	39	55.74	<.0001	
지역구분 (구)	2	829	10.06	<.0001	
부의 나이	3	829	0.41	0.7455	
모의 나이	3	829	0.28	0.8402	
부의 교육정도	4	829	1.55	0.1854	
모의 교육정도	4	829	0.56	0.6944	
가족 수입	7	829	0.33	0.9409	
학생 성별	6	829	2.03	0.0590	
맞벌이여부	1	829	0.02	0.8949	
사교육비중	3	829	2.61	0.0501	
학생성적	4	829	1.31	0.2649	
임의효과	분산	공분산 추정치	표준오차	Z-값	Pr > F
학교평균 ( $\tau_{00}$ )	0.1107	0.05956	1.86	0.0316	
기울기 ( $\tau_{11}$ )	0.01689	0.02334	-1.75	0.0805	
$\tau_{01}$	-0.04080	0.00946	1.79	0.0370	
잔차	0.1157	0.00568	20.38	<.0001	
적합 통계량					
-2 REML	835.5				
AIC	843.5				
BIC	850.2				

$$\begin{aligned}
y_{ij} = & \gamma_{00} + \gamma_{01}(\text{지역구분})_j + \gamma_{02}(\text{부연령})_j + \gamma_{03}(\text{모연령})_j + \gamma_{04}(\text{부학력})_j \\
& + \gamma_{05}(\text{모학력})_j + \gamma_{06}(\text{월소득수준})_j + \gamma_{07}(\text{자녀성별})_j \\
& + \gamma_{08}(\text{맞벌이여부})_j + \gamma_{09}(\text{사교육비비중})_j + \gamma_{010}(\text{자녀학급성적})_j \\
& + \gamma_{11}(\text{부연령})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{12}(\text{모연령})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{13}(\text{부학력})_j(x_{ij} - \bar{x}_{.j}) \quad (3.2) \\
& + \gamma_{14}(\text{모학력})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{15}(\text{월소득수준})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{16}(\text{자녀성별})_j(x_{ij} - \bar{x}_{.j}) \\
& + \gamma_{17}(\text{맞벌이여부})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{18}(\text{사교육비비중})_j(x_{ij} - \bar{x}_{.j}) \\
& + \gamma_{19}(\text{자녀성적})_j(x_{ij} - \bar{x}_{.j}) + u_{0j} + u_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij}
\end{aligned}$$

여기서  $u_{0j} + u_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij}$ 는 임의효과부분이고 나머지 부분은 고정효과 부분이 될 것이다. 모든 변수를 고려한 초기 위계적 선형모형에 대한 결과는 표 3.4에 제시 되어 있다. 위계적 선형모형에 대한 중요한 관심은 초등학교 학부모의 안양시 교육정책 만족도 변수와 다른 설명변수에 의해 설명되어지지 않은, 즉 학교수준 회귀모형의 절편항과 회귀계수의 분산과 공분산이다. 최대우도추정량에 의해

$$\widehat{Var} \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} = \begin{pmatrix} \widehat{\tau}_{00} & \widehat{\tau}_{01} \\ \widehat{\tau}_{10} & \widehat{\tau}_{11} \end{pmatrix} = \begin{pmatrix} 0.1107 & -0.0408 \\ -0.0408 & 0.01689 \end{pmatrix}$$

이 된다. 또한  $\hat{\rho} = \widehat{\tau}_{01} / \sqrt{\widehat{\tau}_{00}\widehat{\tau}_{11}} = -0.944$ 로 제 2 수준인 학교에 대한 상수항의 오차와 회귀계수에 대한 오차 사이에는 매우 강한 음의 상관관계를 보이고 있다. 이는 학교에 따라 자녀 학교 여건에 대한 부모의 평균 만족도가 낮을수록 안양시 교육정책에 대한 만족도가 클수록 부모의 자녀 학교 여건에 대한 만족도 효과가 더욱 더 커짐을 의미한다 할 것이다.

고정효과부분에 대한 결과를 보면, 부연령, 모연령, 가족소득수준, 자녀성별 변수는 명백히 부모의 자녀 학교여건 만족도에 통계적으로 유의한 영향을 주지 않고 있음을 알 수 있다.

**3.2.2. 선택된 모형**

초기 설정 모형의 상위수준인 학교수준 모형에서 명백히 유의하지 않은 부연령, 모연령, 가족소득수준, 자녀성별 변수를 제거하여 다시 모형을 설정하면 상위수준 모형은 다음과 같이 표현할 수 있다.

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{지역구분})_j + \gamma_{02}(\text{부학력})_j + \gamma_{08}(\text{맞벌이여부})_j + \gamma_{09}(\text{사교육비중})_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{지역구분})_j + \gamma_{12}(\text{부학력})_j + \gamma_{18}(\text{맞벌이여부})_j + \gamma_{19}(\text{사교육비중})_j + u_{1j} \end{aligned} \quad (3.3)$$

결국 식 (3.4)를 하위수준 (3.1)에 대입하면, 선택된 최종 위계적 선형모형은

$$\begin{aligned} y_{ij} &= \gamma_{00} + \gamma_{01}(\text{지역구분})_j + \gamma_{02}(\text{부학력})_j + \gamma_{08}(\text{맞벌이여부})_j + \gamma_{09}(\text{사교육비중})_j \\ &\quad + \gamma_{11}(\text{지역구분})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{12}(\text{부학력})_j(x_{ij} - \bar{x}_{.j}) + \gamma_{18}(\text{맞벌이여부})_j(x_{ij} - \bar{x}_{.j}) \\ &\quad + \gamma_{19}(\text{사교육비중})_j(x_{ij} - \bar{x}_{.j}) + u_{0j} + u_{1j}(x_{ij} - \bar{x}_{.j}) + e_{ij} \end{aligned}$$

가 된다.

표 3.5에는 최종 선택된 위계적 선형모형에 대한 임의효과 모수에 대한 추정치와 고정효과 부분의 모수에 대한 유의성 검정을 보여준다. 최대우도추정량에 의해 추정된 학교수준 회귀모형의 절편항과 회귀계수의 분산 공분산 행렬은

$$\widehat{Var} \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} = \begin{pmatrix} \widehat{\tau}_{00} & \widehat{\tau}_{01} \\ \widehat{\tau}_{10} & \widehat{\tau}_{11} \end{pmatrix} = \begin{pmatrix} 0.1019 & -0.0378 \\ -0.0378 & 0.0159 \end{pmatrix}$$

가 된다. 여기서도 마찬가지로  $\hat{\rho} = \widehat{\tau}_{01} / \sqrt{\widehat{\tau}_{00}\widehat{\tau}_{11}} = -0.942$ 로 제 2 수준인 학교에 대한 상수항의 오차와 회귀계수에 대한 오차 사이에는 매우 강한 음의 상관관계를 보이고 있다. 또한 본 논문에서 주장하는 학교 간 차이 등의 이질성으로 인해 위계적 선형모형의 변량효과의 타당성을 검증하기 위해 분산요소 중  $\tau_{00}$ 와  $\tau_{11}$ 의 유의성 여부를 검토하면,  $\tau_{00}$ 와  $\tau_{11}$ 의  $p$ -값이 모두 0.05 이하로 유의미하여 위계적 선형모형의 타당성을 뒷받침해주고 있다.

표 3.5 최종 선택된 위계적 선형모형

고정효과	분자 자유도	분모 자유도	F-값	Pr > F		
교육정책의 만족도	1	39	63.36	<.0001		
학교위치 (구)	2	851	10.38	<.0001		
부의 교육정도	4	851	2.57	0.0366		
맞벌이 여부	6	851	2.51	0.0576		
사교육 비중	4	851	2.41	0.0256		
임의효과	분산	공분산	추정치	표준오차	Z-값	Pr > F
학교평균 ( $\tau_{00}$ )	0.1019		0.05679	1.79	0.0364	
기울기 ( $\tau_{11}$ )	0.0159		0.02242	-1.69	0.0919	
$\tau_{01}$	-0.0378		0.009143	1.73	0.0415	
잔차	0.1147		0.005554	20.65	<.0001	
적합 통계량						
-2 RML	759.5					
AIC	767.5					
BIC	774.3					

위계적 모형의 모수 추정에는 주로 REML (Maximum Likelihood) 방법이 사용된다. ML (Maximum Likelihood) 추정의 단점은 분산 공분산의 추정이 오로지 고정효과에 의존한다는 점이다.



ML과 REML 방법은 일반적으로  $\sigma^2$ 에 대해서는 비슷한 추정결과를 보여주지만 분산-공분산 행렬  $\Psi = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}$ 를 추정하는 데에는 뚜렷한 차이를 보인다. 조직의 수  $j$ 가 매우 크다면 두 추정값은 유사한 결과를 보여주지만, 조직의 수  $j$ 가 작을 때에는 REML 방법에 의한 추정치  $\hat{\tau}_{11}$ 은 ML에 의한 방법보다 약  $(j - f)/j$ 의 배수만큼 작아지게 된다. 여기서  $f$ 는 고정효과 추정치의 수이다 (Raudenbush와 Bryk, 2002). 학교수준에서 모든 설명을 고려한 초기 모형과 최종 선택된 모형의 결과를 비교하기 위해서 두 모형의 -2 REML의 차이를 두 모형의 모수의 차이인 자유도로 나누어  $\chi^2$  검정을 통해 모형의 유의성을 비교한다.

$$\frac{(2REML_{full} - 2REML_{reduced})}{df = 22} = \frac{835.5 - 759.5}{22} = 3.45 \sim \chi^2_{df=22}$$

유의수준  $\alpha = 0.05$ 에서 축소모형은 통계적으로 초기 모형과 차이가 없다는 귀무가설을 기각하지 못하여 축소모형을 선택하게 된다. 또한 표 3.4와 표 3.5의 AIC와 BIC 비교에 의해서도 표 3.5의 모형이 선택되는 결과를 보인다.

최종 선택된 위계적 선형모형의 모수추정치들과 학교수준의 범주형 설명변수들의 수준에 따른 최소제곱평균 (Least Square Means, LSMEANS)을 보여주고 있다. 최소제곱평균에 대한 자세한 의미는 Milliken과 Johnson (1994)를 참조하기 바란다.

표 3.6 위계적 선형모형의 모수 추정치와 최소제곱평균

고정효과	수준	추정치	표준오차	최소제곱평균	표준오차
절편	.	2.7477	0.2134	.	.
교육정책의 만족도	.	0.2569	0.0323	.	.
지역구분 (구)	만안구	-0.2129	0.0505	2.9203	0.0537
	구동안구	-0.1882	0.0523	2.9450	0.0578
	평촌	0	.	3.1332	0.0563
부의 교육정도	중학교 졸업 이하	0.2582	0.1136	3.1722	0.1049
	고등학교 졸업	0.0983	0.0416	3.0123	0.0526
	전문대 졸업	0.0257	0.0436	2.9396	0.0555
	대학교 졸업	0.0455	0.0362	2.9595	0.0506
	대학원 졸업 이상	0	.	2.9140	0.0587
맞벌이 여부	부만 활동	-0.2776	0.1337	2.9262	0.0416
	모만 활동	-0.3043	0.1557	2.8995	0.0897
	부모모두 활동	-0.2353	0.1335	2.9685	0.0406
	활동하지 않음	0	.	3.2039	0.1316
사교육 비중	9% 이하	-0.0349	0.1375	3.0699	0.0524
	19% 이하	-0.1450	0.1354	2.9597	0.0454
	29% 이하	-0.1423	0.1364	2.9625	0.0485
	39% 이하	-0.0825	0.1377	3.0223	0.0533
	49% 이하	-0.1554	0.1495	2.9494	0.0788
	50% 이상	-0.1765	0.1590	2.9282	0.0907
	없음	0	.	3.1047	3.1047

### 3.2.3. 결과 해석

표 3.4와 표 3.5의 결과를 종합해 보면, 학부모의 나이, 모의 교육수준, 가구의 소득수준, 자녀의 성별, 자녀의 학급 내의 성적 순위는 부모들이 느끼는 자녀의 학교여건 만족도에는 영향을 주지 않았다. 반면, 초등학교 학부모의 안양시 교육정책에 대한 만족도, 신도시 지역을 나타내는 지역구분 (구), 부의 교육수준, 부모의 맞벌이 여부, 사교육 비중 변수가 부모들이 느끼는 자녀의 학교여건 만족도에 유의한 변수

로 선정되었다. 표 3.5를 보면, 그 중에서 특히 안양시 교육정책에 대한 만족도가 가장 유의한 영향을 주는 변수이며 다음으로 지역구분(구), 부의 교육수준, 맞벌이 여부, 사교육 비중의 순서로 나타났다. 이는 안양시 학부모가 자녀 학교 여건에 대한 만족도에 영향을 미치는 데에 안양시의 교육정책이 가장 중요한 요소인 것이다.

표 3.6 결과를 보면, 부모들의 안양시 교육정책에 대한 만족도에 대한 변수의 추정치가 0.2569로 부모들이 안양시 교육정책에 대한 만족도를 크게 느낄수록 부모들은 자녀의 학교여건에 대해 더욱 만족함을 알 수가 있다. 학교의 거주지가 평촌에 있을 때 자녀의 학교여건 만족도는 가장 높게 나타나고 다음으로 동안구, 만안구 순으로 나타났다. 이는 신도시에 있는 학교가 제반 여러 환경이 기존 지역보다 우수하기에 학부모의 자녀 학교 여건에 대한 만족도가 높은 것은 당연한 사실일 것이다. 그래서 경제적으로 여유가 있는 부모들은 자녀들의 교육을 위해 신도시나 더 좋은 학교로 이사하는 우리의 현실을 그대로 반영하고 있는 것이다. 부모의 교육수준은 낮을수록 부모들이 느끼는 자녀의 학교여건 만족도는 더 크게 나타났고 경제활동 대상관련해서는 부모 모두가 일하는 가정의 부모가 자녀의 학교여건 만족도가 높게 나타났고 어머니 혼자 일하는 가정은 자녀의 학교여건 만족도는 가장 낮은 수치를 보였다. 마지막으로 사교육비 관련해서는 사교육비를 50% 이상 부담하는 부모가 학교여건에 대해 가장 불만족임을 알 수 있다. 반면 사교육비를 전혀 지출하지 않는 부모가 자녀 학교여건에 대해 가장 만족하고 있으며 대체로 사교육비를 많이 지출할수록 자녀 학교여건에 대한 만족도는 떨어지는 현상을 보여주고 있다. 이러한 현상은 사교육에 의존성이 높은 학부모일수록 학교의 공교육을 신뢰하지 않고 있다는 의미일 것이다. 결론적으로 사교육 없이 공교육으로 충분한 교육이 가능하도록 학교의 제반 환경을 개선하고 안양시의 교육정책에 대한 신뢰감을 주는 것이 안양시 학부모들의 자녀 학교여건에 대한 만족도를 높이는 데에 무엇보다 중요하다 할 것이다. 그래서 그러한 안양시의 공교육에 대한 교육정책에 대한 학부모들의 신뢰가 형성된다면 당연히 학부모들의 학교 제반여건들에 대한 만족도는 지금보다는 더욱 향상될 것이다.

#### 4. 맺음말

위계적 선형모형에서의 가정은 제 1 수준에서의 회귀모형의 절편항과 기울기가 제 2수준의 단위에 대하여 임의로 변한다는 것이다. 이것은 제 2수준 내에서 제 1 수준의 측정자료가 위계적인 특성을 가지고 있기 때문이다. 특히 사회과학에서 분석되는 많은 자료들은 분석에 사용되는 개개인은 학교, 회사, 지역 등 이러한 상위조직에 직간접적으로 많은 영향을 받는다고 할 수 있다. 이러한 경우 개인의 특성과 상위 수준의 특성을 고려하여 상호작용을 다각적으로 분석한다면 복잡한 실제 연구하고자 하는 변수의 영향력을 좀 더 깊이 들여다 볼 수 있을 것이다.

본 논문에서 이용하는 안양시 자료는 학생과 학교는 위계적인 관계를 나타내고, 초등학교 학부모들이 느끼는 자녀 학교여건에 대한 만족도는 그 보다는 상위의 학교 수준에 따라 이미 영향을 받기에 본 연구에서는 상위 수준인 학교의 특성을 반영하는 위계적 선형모형을 설정하여 분석을 하였다. 그 결과 학부모의 나이, 모의 교육수준, 가구의 소득수준, 자녀의 성별, 자녀의 학급 내의 성적 순위는 부모들이 느끼는 자녀의 학교여건 만족도에는 영향을 주지 않았다. 반면, 안양시 교육정책에 대한 만족도, 신도시 지역을 나타내는 지역구분(구), 부의 교육수준, 부모의 맞벌이 상태를 나타내는 경제활동 대상, 사교육 비중은 부모들이 느끼는 자녀의 학교여건 만족도에 유의한 영향을 주는 것으로 나왔다. 부모들의 안양시 교육정책에 대한 만족도가 가장 크게 영향을 미쳤으며 신도시 지역을 나타내는 학교의 거주지가 그 다음 크게 영향을 미쳤다.

본 논문에서 위계적 선형모형의 분석 결과를 종합해 보면, 사교육 없이 공교육만으로 학생들에 대한 교육이 충분히 가능하도록 하고 신도시와 같이 학교의 제반 교육환경을 개선하는 방향으로 안양시의 교육정책을 펼쳐 나간다면 안양시 초등학교 학부모의 자녀가 다니는 학교에 대한 제반 여건들에 대한 만족

도는 지금보다는 더욱 향상될 것으로 보인다.

### 참고문헌

- 강상진 (1998). 교육 및 사회연구를 위한 연구방법으로서의 다층모형과 전통적 선형모형과의 비교분석 연구. <교육평가연구>, **11**, 207-258.
- 김경성 (2007). <다층자료분석의 이론과 실제>, 한국교육심리학회 제6차 연구방법론 워크샵.
- 김경화, 김영화 (2006). 중등학교에서 조직학습이 조직효과성에 미치는 영향에 대한 위계적 선형모형. <교육사회학연구>, **16**, 1-23.
- 김호일 (2005). 반복자료에 대한 Proc Mixed의 이용에 관한 소고. <한국자료분석학회>, **7**, 2067-2080.
- 박찬석, 주경식 (2007). 학업성취도의 공간적분포 패턴연구. <한국지리환경교육학회지>, **15**, 65-76.
- 백순근, 신호정 (2008). 위계선형모형을 활용한 대학생의 강의평가 분석. <교육평가연구>, **21**, 1-24.
- 서민원 (2003). 다층모형의 논리적 구조와 적용: 대학교육의 효과 측정과 분석. <교육평가연구>, **16**, 43-63.
- 유정진 (2006). 위계적 선형모형의 이해와 활용. <아동학회지>, **2**, 169-187.
- 조진남 (2009). 체중감량자료에 대한 적정 공분산형태모형 산출에 관한 실증연구. <한국데이터정보과학회지>, **20**, 377-385.
- 최재성 (2007) 순서형 자료에 대한 비례승산 혼합모형. <한국데이터정보과학회지>, **18**, 471-479.
- Cho, Gyo Young (2008). Modified n-level skip-Lot sampling inspection plans. *Journal of Data Information Science Society*, **19**, 811-818.
- Choi, Jaesung (2008) A marginal logit mixed-effects model for repeated binary response data. *Journal of Data Information Science Society*, **19**, 313-420.
- Goldstein, H. (1995). *Multilevel statistical models*, 2nd Edition, Edward Arnold, London.
- Milliken, G. A. and Johnson, D. E. (1992). *Analysis of messy data, Volume 1: Designed Experiments*, Chapman and Hall, New York.
- Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models*, Sage Publications Inc., Newbury Park, CA.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, **23**, 323-355.
- Wolfinger, R. D. (1996). Heterogeneous variance-covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics*, **1**, 205-230.

# Studying on parents' satisfactory factor to elementary school which their children go to. - focusing on Anyang city

Hoil Kim<sup>1</sup> · Heuiju Chun<sup>2</sup>

<sup>1</sup>Department of Information & Statistics, Anyang University

<sup>2</sup>Department of Data Management, Pusan University of Foreign Studies

Received 23 July 2010, revised 26 September 2010, accepted 30 October 2010

## Abstract

In this study, we applied a hierarchical linear model to Anyang city data because students and their schools are hierarchical data structure. As a result, main factors which affect parents' satisfaction to school which their children go to are parents' satisfaction to Anyang city's education policies and areas which their schools located at. We suggest based on the analysis by this hierarchical linear model that if Anyang city make educational policies more efficient and effective in order for students to study in public school without private education and if Anyang city improve environment related with school like those of new cities, parents' satisfaction to school which their children go to will be increased.

*Keywords:* Fixed effect, hierarchical linear model, random effect, variance-covariance matrix.

---

<sup>1</sup> Professor, Anyang University, Department of Information & Statistics.

<sup>2</sup> Corresponding author: Instructor, Pusan University of Foreign Studies, Department of Data Management. Email: hjchun@pufs.ac.kr