

위계적 선형모형을 이용한 강의평가 결정요인 분석[†]

조장식¹

경성대학교 정보통계학과

접수 2013년 8월 27일, 수정 2013년 9월 29일, 게재확정 2013년 10월 10일

요약

강의평가 결과에 영향을 미치는 특성변수로는 교과목 수준의 다양한 강좌특성 변수들과 수강생 수준의 다양한 인적특성 변수들이 있다. 특정 수강생은 다수의 교과목을 이수하기 때문에 다수의 교과목들은 동일한 수강생 안에 속하게 됨으로써 공유되는 특성이 있게 된다. 즉 강의평가 결과는 교과목 수준의 강좌특성 (1-수준)과 수강생 수준의 인적특성 (2-수준)에 의해 영향을 받는 다층구조 (multi-level)를 가지게 되며, 위계적 자료 특성을 가지는 복수의 분석단위의 구조가 된다. 따라서 전통적인 회귀분석에서와 같이 개별 교과목들이 독립이라는 가정을 할 수 없게 된다. 본 논문에서는 강의평가 결과에 영향을 미치는 다층구조의 특성을 가진 변수들의 영향력을 보다 타당하게 분석하기 위한 방법으로 위계선형모형 (HLM; hierarchical linear model)을 이용하였다. 분석결과는 다음과 같다. 먼저 교과목 수준의 특성변수들 중에 강좌규모, 개설학년, 담당교수의 전임여부, 해당 교과목의 총 평균 평점, 원어강좌 여부가 통계적으로 유의하게 강의평가 결과에 영향을 미친 것으로 나타났다. 또한 수강생 수준의 인적특성 변수들 중에는 성별, 학과계열, 대입당시 전형방법, 평균평점 등이 유의하게 강의평가 결과에 영향을 미친 것으로 나타났다.

주요용어: 강의평가, 다중대응분석, 데이터마이닝, 상호작용효과, 위계선형모형, 의사결정 나무분석.

1. 서론

교육의 질 관리 및 개선의 노력을 위해 1990년대부터 도입한 강의평가 제도는 현재 대부분의 대학에서 실시하고 있다. 특히 대학교육의 질적 수준에 대한 반성과 제고를 위한 노력과 훌륭한 인적자원 개발에 대한 관심으로 인해 대학교육의 질에 대한 관심이 점점 높아지고 있다 (Newman 등, 2004). 대학교육의 질적 평가는 해당 교과목에 대한 수강생들의 강의평가 결과가 대표적인 지표이다. 수강생들의 강의평가 결과는 해당 교과목의 강좌특성 변수들과 수강생의 인적특성 변수들에 의해서 영향을 받는 것으로 선행연구들은 언급하고 있다 (Baek과 Shin, 2008; Son과 Kim, 2007 등). 한편 특정 수강생은 다수의 교과목을 수강하므로, 다수의 교과목들은 동일한 수강생 안에 속하게 됨으로써 공유되는 특성이 있게 된다. 따라서 전통적인 회귀분석에서와 같이 개별 교과목들이 독립이라는 가정을 할 수 없게 된다. 즉 강의평가 결과는 교과목 수준의 강좌특성 (1-수준)과 수강생 수준의 인적특성 (2-수준)에 의해 영향을 받는 다층구조 (multi-level)로서, 위계적 자료 특성을 가지는 복수의 분석단위의 구조가 된다. 따라서 강의평가 결과에 영향을 미치는 변수들의 영향력을 보다 타당하게 분석하기 위한 방법으로 분석단위의 문제를 해결해야 한다는 지적이 제기되어 왔다 (Ethington, 1997; Nasser와 Hagtvet, 2006). 그러나 회귀분석과 같은 전통적인 선형모형들은 복수의 분석단위를 가지는 위계적 자료의 분석에서 분산을 하

[†] 이 논문은 2013학년도 경성대학교 학술연구비지원에 의해 연구되었음.

¹ (608-736) 부산광역시 남구 수영로 309 번지, 경성대학교 정보통계학과, 교수.
E-mail: jscho@ks.ac.kr

위 수준과 상위 수준의 분산으로 분해하지 못하며, 하위 수준의 개체특성 효과가 상위 수준의 집단에 따라 변하는 구조적 관계를 규명하지 못하는 등의 방법론적으로 한계를 갖게 된다 (Kang, 1998).

한편 다층구조의 속성을 지니는 위계적 자료를 전통적인 회귀모형으로 분석하는 경우, 집단수준의 자료를 인위적으로 개인수준에 분산시키게 되어 분산이 과대 추정됨으로써 가설검정의 오류가 커지게 되거나 개인수준의 자료를 평균하여 집단수준의 자료로 활용하게 되어 집합화의 오류 (aggregation bias)를 범하게 된다. 따라서 다층구조의 위계적 구조를 갖는 자료를 전통적인 회귀분석 방법으로 분석한다면 집단의 분산이 주는 오차를 고려하지 않는 한계가 있다.

한편 위계선형모형 (HLM; hierarchical linear model)은 여러 학문분야에서 널리 활용되는 통계분석 기법으로서 자료의 위계적 특성을 적절하게 분석할 수 있는 기법이다. Raudenbush와 Bryk (2002)은 위계적 선형모형이 횡단적 다층 자료구조의 통합모형, 다변량 모형, 잠재변수 모형, 베이지안 추론모형 등에 폭넓게 적용될 수 있음을 제시하고 있다. Jeon과 Kang (2005)은 다층자료의 구조적 특성에 따른 위계적 선형모형의 모수추정을 비교한 바 있으며, Im (2002)은 다층모형을 이용한 발달연구에 대한 횡단적 접근법에 대한 모형을 설정하고 분석하는 방안을 제시하였다.

강의평가에 대한 선행연구로 Han (2001)은 교수, 학생, 수업관련 요인을 중심으로 강의평가 관련요인을 분석한 바 있다. 그리고 Han 등 (2005)은 한국대학의 강의평가실태에 대해서 분석하였으며, Cho 등 (2009)과 Cho (2010)는 교과목 특성 변수에 기초해서 의사결정나무모형과 로지스틱 회귀모형을 이용하여 강의평가점수의 균등화 방법에 대해서 연구한 바가 있다. 그리고 Cho (2012)는 영어강좌의 효과성에 대해서 강의평가 결과를 이용하여 분석한 바 있다. 한편 위계적 선형모형을 이용한 강의평가에 관련된 연구로 Ryu와 Lee (2003)는 강의평가에 영향을 미치는 학생관련 수준의 인적특성만을 고려해서 분석하였다. 그러나 교과목 수준의 강좌특성 관련 요인을 제외했다는 한계가 있다. 그리고 Son과 Kim (2007)은 강의평가에는 교과목 수준의 강좌특성뿐 만 아니라 수강생 수준의 인적특성도 영향을 미치므로 그 결과를 활용하기 위해서는 신중을 기해야 하며, 이와 관련된 체계적이고 경험적인 연구들을 수행해야 함을 제안하였다. Baek과 Shin (2008)은 수강생 수준의 인적특성과 교과목 수준의 강좌특성을 모두 고려한 위계선형모형을 이용해서 강의평가를 체계적으로 분석하였지만, 고정효과 (fixed effect)만을 모형에 포함시켰다는 한계를 갖고 있다.

따라서 본 연구에서는 강의평가 결과에 영향을 미치는 영향력을 분석하기 위해 교과목 수준의 강좌특성과 수강생 수준의 인적특성을 모두 고려하여 위계적 선형모형으로 분석하고자 한다. 이를 위해 다중대응분석 (multiple corresponding analysis)을 이용하여 강의평가결과와 교과목 수준의 강좌특성 및 수강생 수준의 인적특성 변수들 간의 상호 관련성을 분석한다. 또한 의사결정나무분석 (decision tree analysis)을 이용하여 강의평가 결과에 영향을 주는 강좌특성 변수들과 인적특성 변수들의 상호작용 효과를 분석하고자 한다. 여기서 동일한 수강생들이 다수의 교과목을 수강한다는 판단에 근거하여 교과목 수준의 강좌특성 변수를 1-수준 (level 1)으로, 수강생 수준의 인적특성 변수를 2-수준 (level 2)로 설정하였다. 또한 위계적 선형모형을 이용하여 4가지의 분석모형을 설정하여 2가지 수준의 특성변수들이 강의평가 결과에 미치는 영향을 분석하기 위해서 고정효과와 확률효과를 분석한다.

본 논문의 구성은 다음과 같다. 제 2절에서는 자료소개와 기술통계를 하고, 제 3절에서는 다중대응 분석을 통해서 강의평가 결과와 교과목 수준의 강좌특성 및 수강생 수준의 인적특성들 간의 상호관련성을 각각 분석한다. 그리고 의사결정나무분석을 이용하여 강좌특성 및 인적특성 변수들의 상호작용효과를 각각 분석한다. 제 4절에서는 위계적 선형모형을 활용하여 4가지 분석모형을 설정하고 교과목 수준의 강좌특성과 수강생 수준의 인적특성들이 강의평가 결과에 미치는 영향력을 분석한다. 그리고 마지막 제 5절에서는 결론을 제시한다.

2. 자료 및 변수설명

본 논문에서 사용된 자료는 모 대학의 2011년도 2월 졸업생들의 3학년 재학기간까지 이수한 교과목들의 강의평가 결과와 함께 교과목 수준의 강좌특성과 수강생 수준의 인적특성이 포함된 자료이다. 여기서 수강생들의 강의평가 결과는 평가대상이 되는 교과목 수준의 강좌특성뿐만 아니라, 평가주체인 수강생들의 인적특성에 영향을 받는다고 선행연구들을 밝히고 있다. 동일한 수강생 속에는 다수의 교과목들이 포함되므로 다층구조 (multi-level)의 속성을 지니게 되는 위계적 자료구조를 갖게 된다. 본 논문에서는 교과목 수준의 강좌특성을 1-수준 (level-1), 수강생 수준의 인적특성을 2-수준 (level-2)으로 설정하였다.

아래 Table 2.1은 교과목 수준의 강좌특성 (1-수준)에 대한 변수설명과 기술통계를 제시한 결과이다. 여기서 범주형 변수들에 대한 가변수 (dummy variables)는 위계적 선형모형에서 독립변수로 사용되며 각각의 가변수에 대한 기준범주 (reference category)는 가변수의 첨자에 0으로 표기했다.

Table 2.1 Descriptive statistics : level-1 (subject characteristics)

variables	categories	notations	sample size	E_{score}^5	
				mean	$s.d.^4$
class size	$size \leq 20$	$size_0$ ($r.c.^2$)	22,027	3.944	0.933
	$20 < size \leq 40$	$size_1$	43,807	3.902	0.921
	$40 < size \leq 60$	$size_2$	18,429	3.885	0.943
	$size > 60$	$size_3$	2,064	3.897	0.978
$cGPA^1$	$cGPA \leq 2.5$	$cGPA_0$ ($r.c.^2$)	8,985	3.657	0.971
	$2.5 < cGPA \leq 3.5$	$cGPA_1$	59,015	3.885	0.930
	$cGPA > 3.5$	$cGPA_2$	18,327	4.110	0.872
grade of establishment subject	liberal art	$grade_0$ ($r.c.^2$)	13,622	3.873	0.962
	first	$grade_1$	40,768	3.811	0.915
	second	$grade_2$	28,990	4.028	0.919
	third	$grade_3$	2,560	4.259	0.887
	fourth	$grade_4$	387	4.286	0.873
native english class	no	$native_0$ ($r.c.^2$)	85,147	3.908	0.931
	yes	$native_1$	1,180	4.014	0.880
$tenure^3$	no	$tenure_0$ ($r.c.^2$)	58,689	3.866	0.944
	yes	$tenure_1$	27,638	4.000	0.896
total			86,327	3.909	0.931

$cGPA^1$: mean grade point average of the class. $r.c.^2$ reference category in HLM. $tenure^3$: professor position. $s.d.^4$: standard deviation. E_{score}^5 : course evaluation score.

위의 결과에서 강좌규모 (size)가 커질수록 강의평가 점수가 낮아지는 경향이 있으며, 해당 강좌의 총 평점 ($cGPA$)이 높아질수록 강의평가점수는 높아짐을 알 수 있다. 또한 교과목의 개설학년 (grade)이 높아질수록 강의평가점수가 높아지며, 원어강좌 ($native_1$)의 경우가 일반강좌 ($native_0$)에 비해서 강의평가점수가 높고, 전임교원 ($tenure_1$)이 비전임교원 ($tenure_0$)보다 강의평가점수가 높음을 알 수 있다.

아래 Table 2.2는 수강생 수준의 인적특성 (2-수준)에 대한 변수설명과 기술통계를 제시한 결과이다.

Table 2.2 Descriptive statistics : level-2 (student characteristics)

variables	categories	notations	sample size	E_{score}	
				mean	s.d.
sex	female	sex_0 (r.c.)	1,645	3.832	0.492
	male	sex_1	1,265	4.035	0.514
changing major	no	$Cmajor_0$ (r.c.)	2,859	3.920	0.513
	yes	$Cmajor_1$	51	3.909	0.435
department category	art	$Cdpt_0$ (r.c.)	526	3.927	0.537
	natural science	$Cdpt_1$	250	3.976	0.518
	humanities	$Cdpt_2$	1,523	3.907	0.510
	engineering	$Cdpt_3$	611	3.922	0.489
admission time	nonscheduled	$sched_0$ (r.c.)	1,573	3.919	0.531
	scheduled	$sched_1$	1,337	3.921	0.488
admission method	others	$type_0$ (r.c.)	789	3.895	0.535
	SAT^1	$type_1$	369	3.956	0.472
	academic	$type_2$	727	3.903	0.502
	regular	$type_3$	944	3.913	0.499
double major	no	$Dmajor_0$ (r.c.)	2,661	3.918	0.513
	yes	$Dmajor_1$	249	3.937	0.491
$sGPA^2$	$sGPA \leq 2.5$	$sGPA_0$ (r.c.)	15,181	3.596	1.020
	$2.5 < sGPA \leq 3.5$	$sGPA_1$	23,115	3.808	0.929
	$sGPA > 3.5$	$sGPA_2$	48,031	4.057	0.868
total			2910	3.920	0.512

SAT^1 : scholastic aptitude test. $sGPA^2$: students's grade point average for each subject.

위의 결과에서 성별 (sex)은 여자 (sex_0)의 3.832에 비해서 남자 (sex_1)가 4.035로 강의 평가점수가 높으며, 전과 (changing major)를 하지 않은 학생 ($Cmajor_0$)이 3.920으로 전과를 한 학생 ($Cmajor_1$)의 강의 평가점수(3.909)가 낮게 나타났다. 계열 (department category)별로는 자연계열 ($Cdpt_1$)이 가장 높게 나타났으며, 그 다음으로 예술계열 ($Cdpt_0$), 공과계열 ($Cdpt_0$)의 순으로 나타났다. 모집시기 (admission time)별로는 정시모집 ($sched_1$)이 수시모집 ($sched_0$)에 비해서 높게 나타났고, 전형방법 (admission method)의 경우는 수능전형 ($type_1$)이 가장 높게 나타났으며, 그 다음으로 일반전형 ($type_3$), 인문계고전형 ($type_2$) 등의 순으로 나타났다. 복수전공 (double major)에 따라서는 복수전공을 한 학생 ($Dmajor_1$)이 복수전공을 하지 않은 학생 ($Dmajor_0$)에 비해서 강의 평가점수가 높게 나타났으며, 해당 교과목성적 ($sGPA$)이 높을수록 강의 평가 점수가 높게 나타났다.

3. 수준별 특성변수들의 관련성 분석

이 절에서는 1-수준 (교과목수준)과 2-수준 (개인수준)의 특성변수들 각각이 강의 평가결과 (E_{score})와 상호 관련성을 알아보고, 또한 강의 평가결과에 미치는 각 수준별 특성변수들의 상호작용효과를 알아보고자 한다.

먼저 수준별 특성변수들이 강의 평가결과와 상호관련성을 알아보기 위해 각 특성변수들의 범주 수준별로 다중대응분석 (multiple correspondence analysis)을 실시하였다. 다중대응분석은 개체 (케이스)와 범주에 계량적 수치를 부여함으로써 범주형 데이터를 수량화하는 분석기법으로서, 내적 일관성의 원리로부터 범주의 수량화를 실시하는 분석기법이다.

여기서 편의상 강의 평가점수를 2.5점 미만 (E_{score_0}), 2.5~3.5점미만 (E_{score_1}), 3.5점이상 (E_{score_2})로 범주화 하였다. 아래 Figure 3.1은 1-수준과 2-수준 특성변수들에 대해 각각 다중대응분석을 실시한 결과이다.

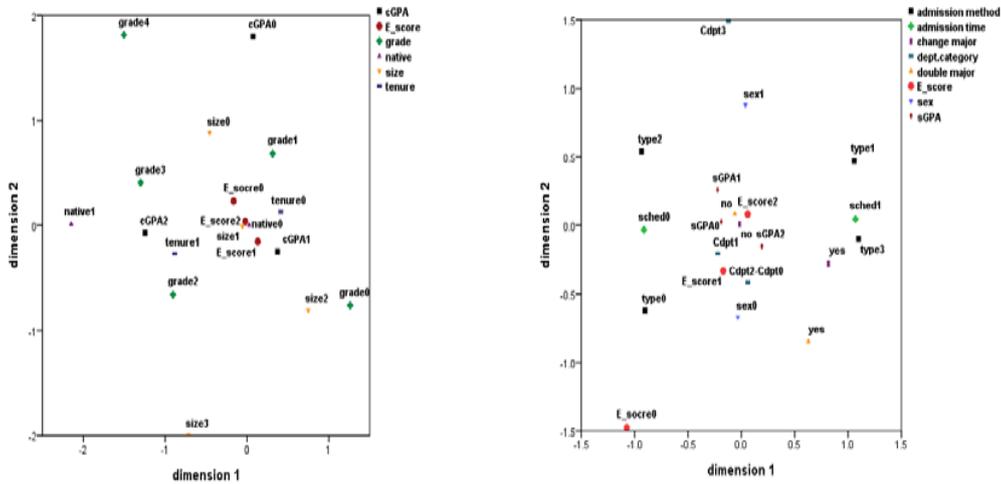


Figure 3.1 Multiple corresponding analysis (left : level-1, right : level-2)

위의 Figure 3.1의 왼쪽 그림은 1-수준 특성변수들에 대한 다중대응분석의 결과로서, 강의평가점수가 높은 그룹 (E_{score}_2)의 수강생들은 담당강좌 평균평점이 높은 그룹 ($cGPA_2$)과 전임교원 ($tenure_1$), 개설학년이 높은 그룹 ($grade_2$), 그리고 원어강좌 ($native_1$)와 상대적으로 가까운 거리에 위치하고 있어서 이들과 상호 관련성이 높은 것으로 나타났다. 반면 강의평가점수가 낮은 그룹 (E_{score}_0)의 수강생들은 개설학년이 낮고 ($grade_1$) 비전임교원 ($native_0$)이며, 담당강좌 평균평점이 낮은 그룹 ($cGPA_0$)와 상대적으로 가까운 거리에 위치하고 있어서 이들과 상호 관련성이 높은 것으로 나타났다.

다음은 2-수준 (수강생 특성) 변수들에 대해 다중대응분석을 실시한 Figure 3.1의 오른쪽 그림을 살펴보면, 강의평가 점수가 높은 그룹 (E_{score}_2)은 복수전공을 하지 않고 ($Dmajor_0$) 기타전형을 제외한 전형 ($type_1$, $type_2$), 정시모집 ($sched_1$), 교과목 평균평점이 높은 그룹 ($sGPA_2$), 남자 (sex_1) 등이 상대적으로 가까운 거리에 위치하고 있어서 이들 간의 상호 관련성이 높은 것으로 나타났다. 한편, 강의평가 점수가 낮은 그룹 (E_{score}_0)는 기타전형 ($type_0$), 여자 (sex_0), 예체능계열 ($Cdpt_0$) 등과 상대적으로 가까운 거리에 위치하고 있어서 이들 간의 상호 관련성이 높은 것으로 나타났다.

다음으로는 1-수준 특성변수와 2-수준 특성변수들이 강의평가 결과에 미치는 효과를 각각 분석하기 위해 비모수적인 방법인 의사결정나무 분석을 이용하고자 한다. 이를 통해서 강의평가 결과에 영향을 미치는 각 수준별 특성변수들에 대한 고차의 상호작용효과를 분석하고자 한다. 의사결정나무분석을 수행하기 위해 지니지수 (Gini index)를 분리기준으로 사용하였으며, 이지분리를 수행하는 CART (classification and regression trees; Breiman 등, 1984) 알고리즘을 사용하였다. 의사결정나무 분석과 관련된 선행연구로는 Cho와 Park (2012)과 Jung과 Min (2013) 등이 있다. 강좌특성 (인적특성) 변수들에 대한 정지규칙으로는 최대나무깊이 (maximum tree depth)는 3 (3)으로 설정하였으며, 최소 케이스 수 (minimum number of cases)에서 부모마디 (parent node)는 4000 (400), 자식마디 (child node)는 400 (40)으로 설정하였으며, 가지치기 (pruning)를 병행하였다.

아래 Figure 3.2와 Figure 3.3은 수준-1과 수준-2의 특성변수들에 대한 의사결정나무분석의 결과를 각각 제시한 것이다.

먼저, Figure 3.2의 결과에서 강의평가 점수에 가장 많은 영향을 미치는 1-수준 (교과목수준) 특성변수는 개설학년 (grade)이고 그 다음으로 담당강좌 평균평점 (cGPA), 강좌규모 (size) 등의 순으로 나타났다. 특히 개설학년이 3학년 이상이고 교과목 평균평점이 3.5 이상인 그룹이 4.303으로 강의평가점수

가 가장 높게 나타났다. 한편 개설학년이 1학년과 교양강좌이고 교과목 평균평점이 2.5 이하이면서 강 좌규모가 20명 이상의 그룹이 3.605로 강의평가점수가 가장 낮게 나타났다.

다음으로 Figure 3.3의 결과에서 강의평가 점수에 가장 많은 영향을 미치는 2-수준 (수강생수준) 특 성변수는 성별이고 그 다음으로는 전형방법, 계열의 순으로 나타났다. 특히 성별이 남자이고, 계열이 예 체능과 자연계열이면서 기타전형으로 입학한 수강생들의 강의평가 점수가 4.301로 가장 높게 나타났다. 반면 성별이 여자이고 전형방법이 일반, 기타 및 인문계고 전형으로 입학한 공학계열 수강생들의 강의평 가점수는 3.748로 가장 낮게 나타났다.

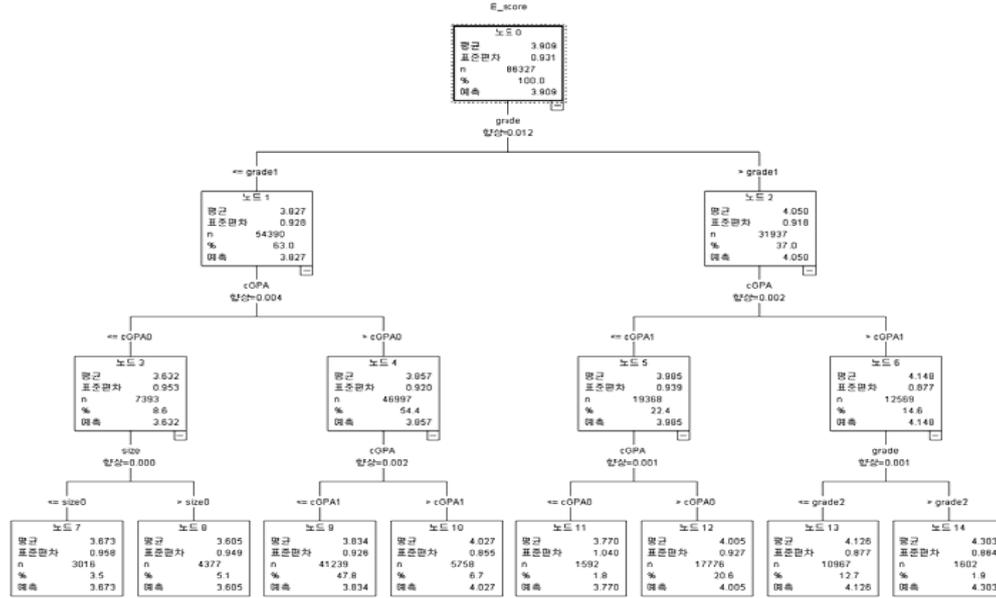


Figure 3.2 Decision tree analysis for level-1 (subject characteristics)

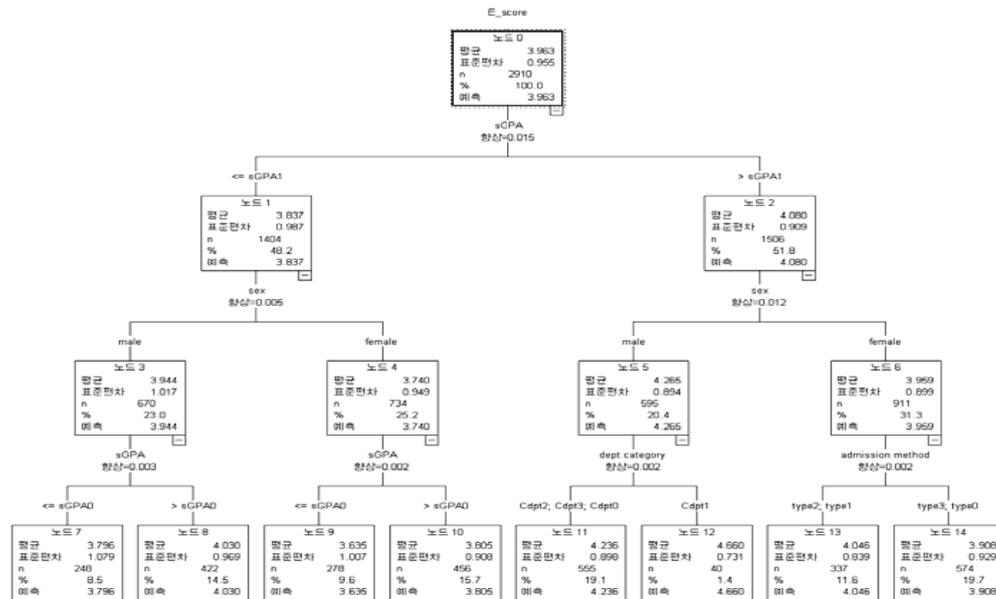


Figure 3.3 Decision tree analysis for level-2 (student characteristic)

4. 위계적 선형모형의 분석결과

4.1. 연구모형

수강생의 강의평가 결과에 대한 교과목 수준의 강좌특성 변수들과 수강생 수준의 인적특성 변수들을 구분해서 각 수준에서 유의하게 영향을 미치는 변수들을 파악하고자 한다. 이를 위해서 Table 2.1과 Table 2.2에 있는 교과목 수준의 강좌특성 변수와 수강생수준의 인적특성 변수들을 사용하였다. 그리고 모든 범주형 변수들은 위계적 선형모형에 독립변수로 투입하기 위해 아래 Table 4.1과 같이 가변수 (dummy variable)로 변환하였다. 여기서 *r.c.*는 기준범주 (reference category)를 의미한다.

Table 4.1 Dummy variables for level-1 and level-2 variables

level-1 variables			level-2 variables		
independent variables	categories	dummy variable	independent variables	categories	dummy variable
class size	size ≤ 20	<i>r.c.</i>	sex	female	<i>r.c.</i>
	20 < size ≤ 40	<i>size_{d1}</i>		male	<i>sex_{d1}</i>
	40 < size ≤ 60	<i>size_{d2}</i>	changing major	no	<i>r.c.</i>
	size > 60	<i>size_{d3}</i>		yes	<i>Cmajor_{d1}</i>
<i>cGPA</i>	<i>cGPA</i> ≤ 2.5	<i>r.c.</i>	department category	art	<i>r.c.</i>
	2.5 < <i>cGPA</i> ≤ 3.5	<i>cGPA_{d1}</i>		natural science	<i>Cdpt_{d1}</i>
	<i>cGPA</i> > 3.5	<i>cGPA_{d2}</i>		humanities	<i>Cdpt_{d2}</i>
grade	liberal art	<i>r.c.</i>	admission time	engineering	<i>Cdpt_{d3}</i>
	first	<i>grade_{d1}</i>		nonscheduled	<i>r.c.</i>
	second	<i>grade_{d2}</i>	scheduled	<i>sched_{d1}</i>	
	third	<i>grade_{d3}</i>	others	<i>r.c.</i>	
native english class	fourth	<i>grade_{d4}</i>	admission method	SAT	<i>type_{d1}</i>
	no	<i>r.c.</i>		academic	<i>type_{d2}</i>
	yes	<i>native_{d1}</i>		regular	<i>type_{d3}</i>
tenure	no	<i>r.c.</i>	double major	no	<i>r.c.</i>
	yes	<i>tenure_{d1}</i>		yes	<i>Dmajor_{d1}</i>
<i>sGPA</i>				<i>sGPA</i> ≤ 2.5	<i>r.c.</i>
				2.5 < <i>sGPA</i> ≤ 3.5	<i>sGPA_{d1}</i>
				<i>sGPA</i> > 3.5	<i>sGPA_{d2}</i>

본 논문에서 고려한 모형으로는 (1) 기본모형 (Model I : null model), (2) level-1의 교과목 수준 변수만 포함한 모형 (Model II : random coefficient model), (3) level-2의 수강생 수준 변수만 포함한 모형 (Model III : mean as outcomes model), (4) level-1과 level-2의 교과목 수준 변수와 수강생 수준 변수를 모두 포함한 모형 (Model IV : intercepts and slopes as outcomes model)을 설정하고 분석하였다. 각 분석 모형에 대한 수리적 표현은 다음과 같다.

먼저 Model I (null model)은 강의평가 점수에 미치는 영향을 분석함에 있어서 강좌특성변수와 인적 특성변수들을 모두 포함하지 않은 기본모형 (일원분산모형)을 통하여 수강생 간 강의평가점수의 평균차이 존재 여부와 그 비율을 분석하고자 한다. 즉 level-1 (교과목특성)과 level-2 (수강생특성)에 대한 기

본모형은 다음과 같다.

① level-1 (교과목 수준) : $y_{ij} = \beta_{0j} + e_{ij}$, $e_{ij} \sim N(0, \sigma^2)$.

② level-2 (수강생 수준) : $\beta_{0j} = \gamma_{00} + u_{0j}$, $u_{0j} \sim N(0, \tau_{00}^2)$, $i = 1, \dots, n_i, j = 1, \dots, m$.

여기서 y_{ij} 는 j 번째 수강생에 대한 i 번째 교과목의 강의평가 점수를 나타내며, β_{0j} 는 j 번째 수강생의 강의평가 평균점수를 나타낸다. 그리고 e_{ij} 는 i 번째 교과목의 강의평가 점수가 j 번째 수강생의 강의평가 평균점수로부터 편차를 나타내는 것으로, 이 값의 분산인 σ^2 이 교과목 수준의 변량이 된다. 그리고 γ_{00} 는 표본 전체 평균을 나타내며, u_{0j} 는 j 번째 수강생의 효과, 즉 j 번째 수강생 간의 차이를 나타낸다. 또한 τ_{00}^2 는 수강생 수준에서의 변량을 의미한다.

두 번째 Model II (random coefficient model)는 기본모형에 level-1 (교과목 수준)의 특성변수들만을 추가한 것으로서 분석모형은 다음과 같다.

① level-1 (교과목 수준)

$$y_{ij} = \beta_{0j} + \beta_{1j} \times size_{d1} + \beta_{2j} \times size_{d2} + \beta_{3j} \times size_{d3} + \beta_{4j} \times grade_{d1} + \beta_{5j} \times grade_{d2} \\ + \beta_{6j} \times grade_{d3} + \beta_{7j} \times grade_{d4} + \beta_{8j} \times tenure_{d1} + \beta_{9j} \times cGPA_{d1} + \beta_{10j} \times cGPA_{d2} \\ + \beta_{11j} \times native_{d1} + e_{ij}, e_{ij} \sim N(0, \sigma^2).$$

② level-2 (수강생 수준)

$$\beta_{0j} = \gamma_{00} + u_{0j}, u_{0j} \sim N(0, \tau_{00}^2).$$

$$\beta_{ij} = \gamma_{i0} + u_{ij}, u_{ij} \sim N(0, \tau_{ii}^2), i = 1, 2, \dots, n_i, j = 1, \dots, m.$$

세 번째 Model III (mean as outcomes model)은 기본모형에 level-2 (수강생 수준)의 특성변수들만을 추가한 것으로서 분석모형은 다음과 같다.

① level-1 (교과목 수준)

$$y_{ij} = \beta_{0j} + e_{ij}, e_{ij} \sim N(0, \sigma^2).$$

② level-2 (수강생 수준)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times sex_{d1} + \gamma_{02} \times Cmajor_{d1} + \gamma_{03} \times Dmajor_{d1} + \gamma_{04} \times Cdpt_{d1} + \gamma_{05} \times Cdpt_{d2} \\ + \gamma_{06} \times Cdpt_{d3} + \gamma_{07} \times sched_{d1} + \gamma_{08} \times type_{d1} + \gamma_{09} \times type_{d2} + \gamma_{010} \times type_{d3} \\ + \gamma_{011} \times sGPA_{d1} + \gamma_{012} \times sGPA_{d2} + u_{0j}, u_{0j} \sim N(0, \tau_{00}^2), i = 1, \dots, n_i, j = 1, \dots, m.$$

마지막으로 Model IV (intercepts and slopes as outcomes model)는 기본모형에 level-1 (교과목 수준)과 level-2 (수강생 수준)의 특성변수들을 모두 추가한 모형으로서 분석모형은 다음과 같다.

① level-1 (교과목 수준)

$$y_{ij} = \beta_{0j} + \beta_{1j} \times size_{d1} + \beta_{2j} \times size_{d2} + \beta_{3j} \times size_{d3} + \beta_{4j} \times grade_{d1} + \beta_{5j} \times grade_{d2} \\ + \beta_{6j} \times grade_{d3} + \beta_{7j} \times grade_{d4} + \beta_{8j} \times tenure_{d1} + \beta_{9j} \times cGPA_{d1} + \beta_{10j} \times cGPA_{d2} \\ + \beta_{11j} \times native_{d1} + e_{ij}, e_{ij} \sim N(0, \sigma^2).$$

② level-2 (수강생 수준)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times sex_{d1} + \gamma_{02} \times Cmajor_{d1} + \gamma_{03} \times Dmajor_{d1} + \gamma_{04} \times Cdpt_{d1} + \gamma_{05} \times Cdpt_{d2} \\ + \gamma_{06} \times Cdpt_{d3} + \gamma_{07} \times sched_{d1} + \gamma_{08} \times type_{d1} + \gamma_{09} \times type_{d2} + \gamma_{010} \times type_{d3} \\ + \gamma_{011} \times sGPA_{d1} + \gamma_{012} \times sGPA_{d2} + u_{0j}, u_{0j} \sim N(0, \tau_{00}^2).$$

$$\beta_{ij} = \gamma_{i0} + u_{ij}, u_{ij} \sim N(0, \tau_{ii}^2), i = 1, \dots, n_i, j = 1, \dots, m.$$

4.2. 분석결과 및 해석

본 논문의 관심은 교과목 수준 강좌특성 변수와 수강생수준의 인적특성 변수들이 강의평가 결과에 대해 어느 정도 영향을 미치는지를 알아보고자 한다. 4.1절에서 제안한 4가지의 분석모형에 대한 위계적 선형모형의 결과는 Table 4.2에 제시하였다.

먼저 기본모형인 Model I (null model)의 결과에서는 집단 내 상관 (ICC; intra class correlation)이 약 27%로 나타났다. 이는 수강생의 인적특성에 따른 차이가 설명할 수 있는 강의평가 결과의 최대 분산이 27%이며, 나머지 약 73% 정도는 교과목 특성들에 의한 수강생 내의 차이라는 것을 의미한다.

다음으로 기본모형에 교과목수준의 강좌특성 변수들만 추가한 Model II (random coefficient model)의 분석결과는 다음과 같다. 먼저 Model I에 비해서 추가한 독립변수들이 갖는 설명력은 1-수준과 2-수준에서 각각 13%와 8%로 증가한 것으로 나타났다. 또한 강좌규모는 기본범주 (reference category)인 20명 미만에 비해서 20명 이상 40명 미만 ($size_{d1}$)이 통계적으로 유의하게 강의평가점수가 낮음을 알 수 있다. 그리고 개설학년은 교양과목에 비해서 개설학년이 올라갈수록 통계적으로 유의하게 강의평가점수가 향상됨을 알 수 있다. 또한 비전임교원에 비해서 전임교원 ($tenure_{d1}$)이, 해당교과목의 총평균평점이 높을수록, 그리고 일반강의에 비해서 원어강의 ($native_{d1}$)의 강의평가점수가 통계적으로 유의하게 높음을 알 수 있다. 다음으로 랜덤효과 (random effect)의 결과를 보면, $grade_{d4}$, $cGPA_{d2}$, $native_{d1}$ 를 제외한 모든 변수에 대해서 랜덤효과가 통계적으로 유의함을 알 수 있다. 즉 이들 변수들이 강의평가 결과에 미치는 효과는 수강생에 따라 유의하게 차이가 있음을 의미한다.

기본모형에 수강생수준의 인적특성 변수들만 투입한 Model III (mean as outcomes model)의 결과에서는 Model I에 비해서 추가한 독립변수들이 갖는 설명력은 1-수준과 2-수준에서 각각 0%와 8%로 증가한 것으로 나타났다. 또한 남자 (sex_{d1})가 여자에 비해서 통계적으로 유의하게 강의평가점수가 높으며, 인문계열 ($Cdpt_{d2}$)과 공학계열 ($Cdpt_{d3}$)이 예체능계열에 비해서 강의평가점수가 유의하게 낮게 나타났다. 그리고 기타전형에 비해서 수능전형 ($type_{d1}$)의 강의평가점수는 유의하게 높고, 인문계고전형 ($type_{d2}$)은 유의하게 낮게 나타났다. 그리고 수강생의 평균평점이 3.5이상 (GPA_{d2})인 경우는 평균평점이 2.5미만인 경우보다 통계적으로 유의하게 강의평가점수가 높게 나타났다.

마지막으로 기본모형에 교과목수준과 수강생수준의 특성 변수들을 추가한 Model IV (intercepts and slopes as outcomes model)의 결과에서는 추가한 독립변수들이 갖는 설명력은 1-수준과 2-수준에서 각각 13%와 13%로 증가한 것으로 나타났다. 먼저 교과목수준의 강좌특성 변수들이 강의평가점수에 미치는 효과를 살펴보면 더미변수 $size_{d2}$ 와 $size_{d3}$ 를 제외한 모든 더미변수들이 통계적으로 유의함을 알 수 있다. 먼저 강좌규모가 20~40명 미만 ($size_{d1}$)은 20명 미만에 비해서 통계적으로 유의하게 강의평가점수가 낮으며, 개설학년이 증가할수록 통계적으로 유의하게 강의평가점수가 높음을 알 수 있다. 전임교원 ($tenure_{d1}$)은 비전임교원에 비해서, 교과목 평균평점 ($cGPA$)이 높을수록, 그리고 원어강좌 ($native_{d1}$)가 일반강좌에 비해서 통계적으로 유의하게 강의평가점수가 높아짐을 알 수 있다.

다음으로 수강생수준의 인적특성 변수들이 강의평가에 미치는 효과를 살펴보면, 남자는 여자에 비해서 통계적으로 유의하게 높으며, 인문계열과 공학계열이 예체능계열에 비해서 유의하게 낮음을 알 수 있다. 또한 인문계고전형은 기타전형에 비해서 통계적으로 유의하게 낮게 나타났으며, 수강생의 평균평점이 3.5이상인 경우는 평균평점이 2.5미만인 경우보다 통계적으로 유의하게 높게 나타났다. 랜덤효과 (random effect)의 결과를 보면, 대부분의 변수에 대해서 랜덤효과가 통계적으로 유의함을 알 수 있다. 즉 이들 변수들이 강의평가 결과에 미치는 효과는 수강생에 따라 유의하게 차이가 있음을 의미한다.

이상의 위계적 선형모형의 결과는 3절의 다중대응분석과 의사결정나무분석의 결과들과 대체로 유사함을 알 수 있다.

Table 4.2 The results of hierarchical linear model

effects	level	variables	Model I		Model II		Model III		Model IV			
			coefficient	s.e. ¹	coefficient	s.e.	coefficient	s.e.	coefficient	s.e.		
fixedeffect	subject	<i>intercept</i>	3.92***	0.01	3.66***	0.02	3.82***	0.04	3.55***	0.04		
		<i>size_{d1}</i>			-0.02***	0.01			-0.01**	0.01		
		<i>size_{d2}</i>			0.00	0.01			0.00	0.01		
		<i>size_{d3}</i>			-0.03	0.02			-0.03	0.02		
		<i>grade_{d1}</i>			-0.07***	0.01			-0.07***	0.01		
		<i>grade_{d2}</i>			0.04***	0.01			0.04***	0.01		
		<i>grade_{d3}</i>			0.15***	0.02			0.15***	0.02		
		<i>grade_{d4}</i>			0.13***	0.04			0.12***	0.04		
		<i>tenure_{d1}</i>			0.09***	0.01			0.09***	0.01		
		<i>cGPA_{d1}</i>			0.25***	0.01			0.25***	0.01		
		<i>cGPA_{d2}</i>			0.39***	0.01			0.40***	0.01		
		<i>native_{d1}</i>			0.06***	0.02			0.06***	0.02		
		student	<i>sex_{d1}</i>					0.26***	0.02	0.30***	0.02	
	<i>Cmajor_{d1}</i>						-0.05	0.07	-0.02	0.07		
	<i>Dmajor_{d1}</i>						0.00	0.03	-0.02	0.03		
	<i>Cdpt_{d1}</i>						0.05	0.04	0.05	0.04		
	<i>Cdpt_{d2}</i>						-0.05**	0.03	-0.07***	0.03		
	<i>Cdpt_{d3}</i>						-0.12***	0.03	-0.10***	0.03		
	<i>sched_{d1}</i>						-0.15	0.09	-0.13	0.09		
	<i>type_{d1}</i>						0.13***	0.09	0.13	0.10		
	<i>type_{d2}</i>						-0.07***	0.03	-0.05**	0.03		
	<i>type_{d3}</i>						0.08	0.09	0.07	0.09		
	<i>GPA_{d1}</i>						0.04	0.04	0.03	0.04		
	<i>GPA_{d2}</i>						0.16***	0.04	0.13***	0.04		
	random effect		subject	<i>size_{d1}</i>			0.01***	0.00			0.01***	0.00
				<i>size_{d2}</i>			0.04***	0.00			0.04***	0.00
				<i>size_{d3}</i>			0.07***	0.02			0.08***	0.02
				<i>grade_{d1}</i>			0.10***	0.01			0.10***	0.01
		<i>grade_{d2}</i>				0.07***	0.01			0.07***	0.01	
<i>grade_{d3}</i>					0.03***	0.01			0.03***	0.01		
<i>grade_{d4}</i>					0.02	0.04			0.02***	0.04		
<i>tenure_{d1}</i>					0.02***	0.00			0.02***	0.00		
<i>cGPA_{d1}</i>					0.01***	0.00			0.01***	0.00		
<i>cGPA_{d2}</i>					0.01	0.00			0.00	0.00		
<i>native_{d1}</i>			0.00	0.02			0.00	0.02				
variance	σ^2		0.63***		0.55***		0.63***		0.55***			
	τ^2_{00}		0.24***		0.22***		0.22***		0.21***			
	total		0.87		0.77		0.85		0.76			
	ICC		0.27		0.28		0.26		0.28			
R^2	level-1		-		0.13		0.00		0.13			
	level-2		-		0.08		0.08		0.13			
	total		-		0.11		0.02		0.13			

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. s.e.¹: standard error.

5. 결론

본 논문은 위계선형모형을 이용하여 교과목수준의 특성과 수강생의 인적특성이 강의평가 결과에 미치는 영향력을 분석하였다. 주요 결과는 다음과 같다.

첫째, 기본모형의 분석결과 강의평가점수는 교과목 간 차이가 수강생 간의 차이보다 더 큰 것으로 나타났다 (ICC=0.27).

둘째, 교과목수준의 강좌특성 변수들만 투입한 결과 강좌규모는 20명미만에 비해서 20~40명미만이 유의하게 강의평가점수가 낮게 나타났다. 그리고 개설학년이 올라갈수록 유의하게 강의평가점수가 높게 나타났다. 또한 비전임교원에 비해서 전임교원이, 해당교과목의 총평균평점이 높을수록, 그리고 일반강의에 비해서 원어강의 강의평가점수가 유의하게 높게 나타났다.

셋째, 수강생수준의 인적특성 변수들만 추가한 결과에서는 남자가 여자에 비해서 유의하게 강의평가점수가 높으며, 인문계열과 공학계열이 예체능계열에 비해서 유의하게 낮게 나타났다. 그리고 기타전형에 비해서 수능전형의 강의평가점수는 유의하게 높고, 인문계고전형은 유의하게 낮게 나타났다. 그리고 수강생의 평균평점이 3.5이상인 경우는 평균평점이 2.5미만인 경우보다 통계적으로 유의하게 높게 나타났다.

넷째, 교과목수준과 수강생수준의 특성변수들을 모두 추가한 결과에서는 다음과 같다. 먼저 교과목수준의 변수들이 강의평가점수에 미치는 효과를 살펴보면, 강좌규모가 20~40명 미만은 20명 미만에 비해서 유의하게 강의평가점수가 낮으며, 개설학년이 증가할수록 통계적으로 유의하게 높게 나타났다. 전임교원은 비전임교원에 비해서, 교과목 평균평점이 높을수록, 그리고 원어강좌가 일반강좌에 비해서 유의하게 높게 나타났다. 다음으로 수강생수준의 변수들이 강의평가에 미치는 효과를 살펴보면, 남자는 여자에 비해서 유의하게 높으며, 인문계열과 공학계열이 예체능계열에 비해서 유의하게 낮게 나타났다. 또한 인문계교전형은 기타전형에 비해서 통계적으로 유의하게 낮게 나타났으며, 수강생의 평균평점이 3.5이상인 경우는 평균평점이 2.5미만인 경우보다 통계적으로 유의하게 높게 나타났다. 또한 랜덤효과(random effect)의 결과를 보면, 대부분의 교과목 수준의 특성변수들에 대해서 랜덤효과가 통계적으로 유의함을 알 수 있었다.

이상의 결과는 교육의 질 관리를 위한 기초자료로 제공될 수 있을 것이다. 한편 본 논문은 3학년까지만 수강한 교과목에 대한 강의평가결과를 분석했다는 점과 4가지의 분석모형만 분석했다는 점에서 본 연구의 결과를 지나치게 일반화 시키는데는 한계가 있음을 밝혀 둔다. 보다 세밀한 분석모형의 설정과 분석은 향후의 과제로 남겨두기로 한다.

References

- Baek, S. G. and Shin, H. J. (2008). Multilevel analysis of the effects of student and course characteristics on student course evaluation - Focused on the undergraduate liberal education program. *Journal of Educational Evaluation*, **21**, 1-24.
- Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and regression trees*, Chapman and Hall Press, Washington, D.C.
- Cho, J. S., Kang, C. W., Choi, S. B. (2009). Comparison on equating methods for course evaluation. *Journal of the Korean Data Information & Science Society*, **20**, 65-75.
- Cho, J. S. (2010). A study on equating method based on regression analysis. *Journal of the Korean Data Information & Science Society*, **21**, 513-521.
- Cho, J. S. (2012). Study on the effectiveness of english-medium class. *Journal of the Korean Data & Information Science Society*, **23**, 1137-1144.
- Cho, K. H. and Park, H. C. (2012). A study on decision tree creation using marginally conditional variables. *Journal of the Korean Data Information & Science Society*, **23**, 299-307.
- Ethington, C. A. (1997). A hierarchical linear modeling approach to studying college effects. *Higher Education: Handbook of Theory and Research*, **12**, Agathon Press, NY, 165-194.
- Han, S. I. (2001). Research on student ratings of university teaching: Analysis of determinants related to professor, student, and class. *The Journal of Educational Administration*, **19**, 247-266.
- Han, S. I., Kim, H. J. and Lee, J. Y. (2005). A comprehensive study of Korean students' evaluations of university teaching. *The Journal of Educational Administration*, **23**, 379-403.
- Im, S. H. (2002). Multilevel models for cross-sectional approach to the developmental research. *Journal of Educational Evaluation*, **15**, 295-315.
- Jeon, M. J. and Kang, S. J. (2005). A comparison of multilevel models in their parameter estimation - The comparison of 2-level HLM, 3-level HLM, and CMM. *Journal of Education Evaluation*, **18**, 123-147.
- Jung, H. J. and Min, D. K. (2013). The study of foreign exchange trading revenue model using decision tree and gradient boosting. *Journal of the Korean Data Information & Science Society*, **24**, 161-170.
- Kang, S. J. (1998). Analytical comparisons between classical linear models and multilevel models as educational and social research methods. *Journal of Educational Evaluation*, **11**, 207-258.
- Nasser, F. and Hagtvet, K. A. (2006). Multilevel analysis of the effects of student and instructor/course characteristics on student ratings. *Research in Higher Education*, **47**, 559-590.
- Newman, F., Couturier, L. and Scurry, J. (2004). *The future of higher education: Rhetoric, reality, and the risks of the market*, Jossey-Bass, San Francisco.
- Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, SAGE publication, San Francisco.
- Ryu, C. H., Lee, J. H. (2003). A study on student factors associated with the student evaluation of teaching at universities. *Korean Management Review*, **33**, 89-807.
- Son, C. K. and Kim, Y. T. (2007). A study on the difference of result to variables related to students' evaluation of teaching in university. *Journal of Educational Evaluation*, **20**, 1-24.

Determinants of student course evaluation using hierarchical linear model[†]

Jang Sik Cho¹

Department of Informational Statistics, Kyungsung University

Received 27 August 2013, revised 29 September 2013, accepted 10 October 2013

Abstract

The fundamental concerns of this paper are to analyze the effects of student course evaluation using subject characteristic and student characteristic variables. We use a 2-level hierarchical linear model since the data structure of subject characteristic and student characteristic variables is multilevel. Four models we consider are as follows; (1) null model, (2) random coefficient model, (3) mean as outcomes model, (4) intercepts and slopes as outcomes model. The results of the analysis were given as follows. First, the result of null model was that subject characteristics effects on course evaluation had much larger than student characteristics. Second, the result of conditional model specifying subject and student level predictors revealed that class size, grade, tenure, mean GPA of the class, native class for level-1, and sex, department category, admission method, mean GPA of the student for level-2 had statistically significant effects on course evaluation. The explained variance was 13% in subject level, 13% in student level.

Keywords: Course evaluation, decision tree analysis, hierarchical linear model, interaction effect, multiple corresponding analysis.

[†] This research was supported by Kyungsung University Research Grants in 2013.

¹ Professor, Department of Informational Statistics, Kyungsung University, Busan, 608-736, Korea.
E-mail: jscho@ks.ac.kr