

대입수능 선택과목 점수조정을 위한 선형계획모형 개발 및 활용*

남 보 우**

A Linear Programming Model to the Score Adjustment among the CSAT Optional Subjects

Bowoo Nam**

■ Abstract ■

This study concerns with an applicability of the management science approach to the score adjustment among the College Scholastic Aptitude Test(CSAT) optional subjects. A linear programming model is developed to minimize the sum of score distortions between optional subjects. Based on the analysis of the 377,089 CSAT(2010) applicants' performances in social science test section, this study proposes a new approach for the score equating or linking method of the educational measurement theory.

This study makes up for the weak points in the previous linear programming model. First, the model utilize the standard score which we can get. Second, the model includes a goal programming concept which minimizes the gap between the adjusting goal and the result of the adjustment. Third, the objective function of the linear programming is the weighted sum of the score distortion and the number of applicants. Fourth, the model is applied to the score adjustment problem for the whole 11 optional subjects of the social science test section.

The suggested linear programming model is a generalization of the multi-tests linking problem. So, the approach is consistent with the measurement theory for the two tests and can be applied to the optional three or more tests which do not have a common anchor test or a common anchor group. The college admission decision with CSAT score can be improved by using the suggested linear programming model.

Keywords : Test Equating, Test Linking, Linear Programming, CSAT, Admission

논문접수일 : 2010년 11월 03일 논문수정일 : 2010년 12월 20일 논문게재확정일 : 2010년 12월 28일

* 이 논문은 2010년 교육과학기술부 및 한국교육과정평가원의 지원에 의하여 연구되었고, 논문의 일부는 2010 한국 경영과학회 추계학술대회에서 발표되었음.

** 단국대학교 경상대학 경영학부

1. 서론

1.1 연구배경 및 목적

대학입학전형에서 경쟁이 공정하도록 국가는 대학수학능력시험(이하 “수능”이라 함)의 선택과목 간 점수조정을 하였지만, 최근에는 국가가 조정을 하지 않고 각 대학이 선발상황을 고려하여 가산점을 부여함으로써 선택과목 간 점수를 조정하고 있다. 외국에서도 토익과 토플, 미국 수능(SAT) 등 시험 간의 점수조정 문제에 대하여 많은 연구가 있었다.

점수조정 문제는 교육평가(educational evaluation)분야의 측정이론(measurement theory) 중 검사동등화(test equating)이론에 기초하여 연구되어 왔다. 검사동등화는 공통과목이나 공통항목을 가교검사(anchor test)로 활용하여 점수를 조정하거나, 공통응시자를 가교집단(anchor group)으로 활용하여 점수를 조정하는 것이다. 2002학년도부터 2004학년도까지 수능에서는 각 선택과목 간 공통문항이 있어서 검사동등화 방법으로 점수를 조정하였다. 2005학년도부터 수능 탐구영역의 선택과목에는 가교검사로 활용할 공통과목이나 공통문항이 없고, 모든 과목에 응시한 가교집단으로서의 공통응시자도 없어서 측정이론을 적용하지 못하였다. 따라서 점수를 조정하지 않고 수능의 각 영역별 및 과목별로 표준점수를 산출하고 있다. 즉, 언어, 외국어, 수리 영역은 영역별 평균 100점, 표준편차 20점, 탐구 영역의 각 선택과목은 과목별 평균 50점, 표준편차 10점을 갖도록 변환한 표준점수 분포에서 개인이 획득한 원점수가 어느 위치에 해당하는지 표준점수를 산출하여 제공한다. 또한 각 영역별 및 과목별 응시학생 전체에 대하여 특정 표준점수보다 낮은 표준점수를 획득한 학생의 비율로 표준점수 백분위를 계산하며, 표준점수를 9등급(stanine scale)으로 나누고 특정 표준점수가 속해 있는 등급을 표시한다.

표준점수는 과목별 시험의 난이도 차이를 반영

하여 과목 간 점수를 비교할 수 있도록 한 것이다. 그러나 과목별 응시집단의 능력 차이가 반영된 것이 아니기 때문에 서로 다른 과목의 표준점수를 직접 비교하는 것은 잘못일 수 있다. 예를 들어 윤리 과목의 표준점수 60점이 수치상으로 국사과목의 표준점수 55점보다 높다. 그러나 국사 응시자가 윤리 응시자보다 전체적으로 수학능력이 매우 높으면, 과목별로 표준화한 윤리 60점이 국사 55점보다 수학능력이 높다고 할 수 없다. 수능이 수학능력을 측정하는 것이지만 표준점수가 수학능력을 정확하게 표시하지 못하면, 표준점수로 수학능력을 평가하여 학생을 선발하는 것은 문제가 있다. 남현우[8]는 대학이 표준점수를 그대로 활용하면 엄청난 점수 왜곡 현상이 생기고 대학별로 표준점수를 다시 가공할 경우 그 방법이 천차만별이어서 혼란이 불가피 할 전망이라고 논하였다. 남보우[5]는 2010학년도 입학전형에서 대학의 90% 이상이 서로 다른 선택과목의 표준점수를 합한 총점을 비교하여 학생을 선발하고 있으며, 가산점을 부여하는 대학은 매우 적은 것으로 보고하고 있다. 선택과목의 표준점수를 비교하여 학생을 선발할 수밖에 없다면 실력이 정확하게 반영되도록 응시집단 간 능력 차이를 반영하여 표준점수를 조정할 필요가 있다.

현행 수능의 선택과목 각 쌍 사이에는 공통응시자가 있다. 이 경우 선택과목 각 쌍 간에는 공통응시자를 가교집단으로 활용하여 기존의 측정이론으로 점수를 조정할 수 있다. 그러나 선택과목이 3개 이상인 경우 과목 각 쌍의 조정결과 사이에 불일치가 있을 수 있어서 완벽한 조정이 불가능하다. 3개 이상의 과목 각 쌍 사이에만 공통응시자가 있는 점수조정 문제는 기존의 측정이론이 해법을 제시하지 못하였다. 본 연구는 완벽한 조정은 불가능하지만 최선으로 조정하려는 최적화의 개념을 이 문제에 적용하고자 한다. 즉, 능력 차이를 반영하지 못하여 왜곡되는 점수의 총합을 최소화하는 선형계획모형을 개발하여 제시하고자 한다. 모형의 최적해로부터 선택과목별 가산점의 크기를 정할 수 있으므로, 대학은 제시된 선형계획모형을 입학전형에

용이하게 활용할 수 있을 것으로 기대된다.

1.2 연구방법 및 범위

본 연구는 수능 탐구영역의 선택과목 점수조정을 위한 선형계획모형을 개발하고, 모형을 활용하여 수능 점수를 조정함으로써 개발된 모형이 유용하다는 것을 보이고자 한다. 우선 수능 점수조정과 관련된 기존의 이론과 최근의 연구결과를 검토하여 본 연구가 다루고자 하는 문제에 대한 현황과 주요 과제를 밝히고자 한다.

대입전형 현실에서 탐구영역 선택과목의 점수를 조정하는 것은 각 선택과목에 얼마의 가산점을 부여할 것인가를 결정하는 것이다. 즉, 각 선택과목 응시집단 능력의 평균 차이를 얼마의 가산점으로 환산할 것인가를 결정하는 것이다. 특정 선택과목 응시자에게 가산점을 부여하는 것은 측정이론의 평균방법(mean method)에 의한 점수조정 방법이다. 측정이론에는 평균방법 이외에도 점수 수준별로 가산점의 크기를 다르게 부여하는 선형방법(linear method)과 동백분위방법(equipercntile method)이 있다. 그러나 대학들은 선택과목마다 일정한 가산점을 정하는데 중점을 둔다는 측면에서 평균방법에 의한 조정이 현실적으로 필요하다. 따라서 본 연구는 선택과목별 응시집단의 평균 능력 차이를 반영하여 점수를 조정할 수 있는 선형계획모형을 개발하여 제시하고, 선형방법이나 동백분위방법을 도입하여 선형계획모형으로 조정하는 방법은 다음의 연구주제로 하고자 한다.

서로 다른 두 검사점수의 조정방법으로 측정이론에서 개발된 결과를 검토하고자 한다. 특히, 탐구영역 선택과목 각 쌍 간의 점수를 조정하는 방법을 선형계획모형 개발에 도입하기 위하여 측정이론의 단일집단설계(single group design)에 의한 두 검사점수 조정방법을 검토한다. 또한 선형계획모형의 적용결과와 비교하기 위하여 공통검사를 활용하여 2단계에 걸쳐 두 검사점수를 조정하는 연쇄방법(chained method)을 검토하고자 한다.

2010학년도 수능 사회탐구영역(이하 “사탐”이라 함) 선택과목 중 3과목을 선택하여 점수조정의 상황을 예시하고, 각 과목 쌍 간에 측정이론의 단일 집단설계에 의한 점수조정 방법을 도입하여 분석함으로써 점수조정 필요성, 점수조정 방법, 현행 방법의 한계 및 연구의 필요성을 밝힌다. 다음은 일반적으로 n개의 선택과목이 있고 각 선택과목 쌍 간에 공통응시집단이 있을 때, 각 과목 쌍 간에 단일집단설계에 의한 점수조정 방식을 도입하고 전체적으로 조정되지 못한 왜곡된 점수를 최소화하는 선형계획모형을 개발한다. 점수조정을 위한 기존의 선형계획모형[4, 6]은 평균능력 차이 조건을 제약식으로 포함하였으나, 현실 적용성을 높이기 위해서 평균능력 차이 반영모표를 도달하도록 목표계획법 개념의 제약식을 도입하고자 한다. 또한 목적함수를 단순 왜곡된 점수의 합이 아니라 응시자 수의 가중합계로 하여 선형계획모형을 개발하고자 한다. 모형을 개발한 후 예시 사례에 대하여 모형을 작성하고 분석하여 점수조정과 모형적용의 의미를 명확하게 하고자 한다.

본 연구의 목적은 유용한 점수조정 방법을 제시하는 것이므로 사탐 11개 선택과목의 표준점수를 조정하고 분석하여 제시한 선형계획모형의 유용성을 보이고자 한다. 현행 수능은 원점수를 제공하지 않으므로 표준점수와 백분위를 조정대상으로 하여야 한다. 다만, 과학탐구영역이나 직업탐구영역의 선택과목 표준점수 조정 및 백분위의 조정도 사탐의 표준점수 조정과 동일한 방법을 사용할 수 있으므로 점수조정과 분석을 생략하고자 한다. 2010학년도 사탐 선택과목의 표준점수를 조정하기 전에 상관계수 분석과 평균점수 차이 분석을 통하여 조정의 타당성을 검토한다. 또한 언어, 수리, 외국어 영역의 종합평균 표준점수를 공통검사(anchor test)로 활용하여 측정이론의 2단계 연쇄방법으로 조정된 결과와 본 연구에서 제시한 선형계획모형으로 조정된 결과를 비교하고자 한다. 선형계획모형의 최적해는 엑셀의 해 구하기 기능(solver.exe)을 활용하여 구하고자 한다[13].

본 연구는 선택과목별 전체 응시집단을 대상으로 표준점수를 조정한다. 다만, 특정대학의 입학전형에는 일정한 수학능력 범위의 지원자가 경쟁한다는 현실을 고려하여 언어, 수리, 외국어 종합평균 수능등급이 2.00등급부터 3.99등급 집단의 사탐 선택과목 표준점수를 조정하여 추가로 제시하고자 한다. 종합등급이 일정범위에 속한 집단의 선택과목 점수를 조정하는 것을 제시함으로써 특정대학이 본 연구에서 제시한 모형을 학생선발에 활용할 때 참고할 수 있도록 한 것이다.

2. 기존연구의 고찰

2.1 검사결과의 비교방법

대학수학능력을 평가하는 수능은 검사에 속한다. 검사란 인간에 내재된 속성을 간접적으로 측정하기 위한 도구이다. 동일한 속성을 측정하기 위하여 두 가지 이상의 검사가 이루어지는 경우가 있다. 이때 서로 다른 검사결과를 비교하기 위한 동등화(equating), 연계화(linking), 척도화(scaling)에 대하여 많은 연구가 있었다[2, 9, 11, 16].

동등화란 동형검사의 점수를 상호 교환적으로 사용할 수 있도록 조정하는 통계적 과정이다. 평균동등화(mean equating)는 두 검사점수의 분포 형태가 동일하고 단지 평균이 다르다고 가정하고 평균을 조정하는 것이고, 선형동등화(linear equating)는 평균과 표준편차가 다르다고 가정하고 이를 조정하는 것이며, 동백분위동등화(equipercentile equating)는 두 검사의 분포특성에 관계없이 백분위가 같은 점수를 동일한 점수로 변환하여 누적분포를 일치시키도록 조정하는 것이다. 검사 총점을 분석하지 않고 각 문항의 고유한 문항특성을 분석하는 문항반응이론(item response theory)도 동등화를 위하여 활용된다.

검사 동등화를 위해서 반드시 만족해야 하는 조건은 동일한 내용과 동일한 통계적 특성을 가져야 한다는 것이다[16]. 현행 수능의 서로 다른 선택과

목 시험이 동일한 내용인 “대학수학능력” 또는 동일한 “탐구능력”을 검사하고, 각 선택과목의 점수가 동일한 통계적 특성을 가져야 동등화를 할 수 있다. 이외에도 동등화를 위하여 필요한 특성들로 공정성(equity), 대칭성(symmetry), 집단불변성(group invariance) 등이 있다. 일반적으로 정교하게 고안된 가교검사(anchor test)가 없는 경우 동등화 조건을 만족하기 어려운 것으로 알려져 있다.

검사유형간 비교 가능성을 찾는 경우를 검사동등화라고 한다면 서로 다른 검사 간 비교 가능성을 찾는 경우는 연계화 또는 척도화라고 한다. 현행 수능은 동등화에 필요한 요건을 갖춘 설계가 이루어지지 않았기 때문에 연계화가 필요한 상황이고, 이들은 개념상 다르기는 하지만 통계적인 절차는 다르지 않다[8]. 검사점수 간에는 어느 것이나 연계가 가능하지만 추론, 구성, 모집단, 측정특성과 조건 측면에서 의미가 있어야 한다. 임의의 두 변수 간에 회귀식을 구할 수 있지만 의미가 있어야 하는 것과 같다. 검사점수를 상호 비교 가능한 점수 체제로 변환시키는 연계화(linking) 방법은 일반적으로 동등화(equating), 추정(calibration), 예측(projection), 조절(statistical moderation, social moderation)의 방법이 있다[11, 14, 16]. 척도화(scaling)는 검사결과에 숫자나 순서를 부여하는 과정이며, 그 목적은 검사결과를 해석하는데 도움이 되도록 하는 것이다. 수능에서 표준점수를 산출하는 것은 이미 척도화 방법을 사용한 것이다.

수능의 경우는 검사점수의 사용 목적상으로는 검사 점수의 맞비교를 위한 동등화라고 할 수 있으나, 동등화를 사용하는 경우와는 전제가 다르다. 다만, 동등화, 연계화, 척도화는 조건에 따라 유사한 통계적 방법을 적용한다. 따라서 수능 점수조정에는 연계화 또는 척도화라는 용어를 상황에 따라 사용할 수도 있으나, 점수조정이라는 용어로 사용하는 것이 적절하다고 본다.

2.2 수능 선택과목 점수조정 연구

우리나라에서 검사점수의 조정 또는 동등화에

대한 논의와 연구는 1990년대에 이르러 활발해지기 시작했고, 1999학년도부터 수능에 표준점수를 도입하면서 선택과목 점수조정에 관한 연구들이 상당한 진전을 보였다[9]. 2004학년도까지 선택과목의 점수는 동일한 영역의 공통과목 또는 문항을 가교점사로 사용하여 조정하였다.

2005학년도부터 수능 선택과목에는 공통문항이 없고 응시자가 모든 과목을 자유롭게 선택하기 때문에 점수조정을 위한 동등화의 조건이 충족되기 어려웠다. 따라서 2005학년도 수능점수 보고체제를 검토하기 위하여 다양한 논의와 연구가 진행되었다[3]. 결국 국가의 수능점수 보고체제는 탐구영역 간은 물론 선택과목 간에도 점수를 조정하지 않고 각각 독립적으로 표준화하도록 하였다. 그 이유는 첫째, 선택과목은 독립적인 특성을 가지며, 둘째, 각 선택과목은 각 영역의 공통문항이 없어서 직접 비교하기 어려운 검사체제를 구성하고 있고, 셋째, 각 영역 점수의 상호상관을 과거 자료로 구해보면 동등화시키기에는 상관계수가 너무 낮으며, 넷째, 조정에 대하여는 각 대학이 선택할 문제이며, 조정이 필요한 경우 선발상황을 고려하여 대학 나름대로 조정하는 것이 바람직하다는 것이었다[5].

그러나 수능 선택과목 별 점수 표준화는 응시집단의 능력 차이를 반영하지 못하여 점수가 왜곡되므로 점수를 조정하여 공정성을 제고하려는 연구가 진행되었다. 전통적 측정이론으로 공통과목을 가교점사로 활용하여 합집단을 구성하고 선형방법 및 동백분위방법으로 점수를 조정할 수 있는 가능성에 대한 연구[7, 8]와 문항반응이론을 활용한 점수조정 가능성에 대한 연구[9]가 있었다. 전통적 측정이론에 의한 점수조정 방법을 적용하는 경우 공통과목 선정에 어려움이 있고, 문항반응이론을 적용하는 경우 점수의 역전현상 등이 발생하여 국가 수능점수 보고체제를 보완하거나 대학들이 대입전형에서 제시된 방법을 활용하는 데는 한계가 있었다.

한편 입학전형의 실용성 측면에 중점을 둔 연구로 각 과목별 원점수 평균을 활용하여 보정상수를

도입하고 가산점수를 간단하게 계산하는 가산점수 방법이 제시되었으나[15], 2005학년도 이후의 수능에서는 원점수를 보고하지 않았으며, 보정상수 도입 등 현실적 적용에 어려움이 있었다. 선형계획모형으로 각 선택과목 간 능력 차이를 반영하여 점수를 조정하는 방법에 대한 연구도 있었다. 남보우[4]는 2005학년도 수능의 모의자료를 활용하여 표준점수를 산출하는 선형계획모형을 제시하였다. 선형계획모형의 목적함수는 총 후회를 최소화하는 것으로 하고, 능력이 큰 응시집단의 표준점수 평균이 능력이 작은 응시집단의 평균보다 작아서는 안 된다는 조건을 제약식으로 포함하였다. 이 연구도 원점수를 활용하여 표준점수의 기준점을 산출하는 것이므로 원점수를 제공받지 못하는 대학입학전형에는 활용할 수 없었다. 이후 수능 선택과목 점수 조정에 관한 연구는 활발하게 진행되지 못하였다.

2010년에 국가 수능점수를 연구자에 한하여 공개할 수 있도록 대법원 판결(공 2010상, 663) 및 국가정책이 정해졌다. 이에 따라 남보우[5]는 실제 수능점수를 활용하여 선형계획모형으로 표준점수를 조정한 결과와 교육측정이론의 다양한 조정방법으로 표준점수를 조정한 결과를 비교함으로써 선형계획모형이 선택과목 점수조정에 유용하다는 것을 보였다. 그러나 이 연구는 2002년에 제시된 선형계획모형[4]을 활용하여 실제 표준점수 자료를 분석한 것이었다. 남보우[6]는 2010년 연구[5]를 기반으로 각 선택과목의 표준점수를 활용하여 응시집단 능력 차이를 측정하고 이를 반영하여 표준점수와 백분위를 조정하는 선형계획모형을 제시하였다. 제시된 선형계획모형은 측정이론에 기반으로 둔 조정모형이고, 2010학년도 수능 사탐 377,089명의 실제 표준점수와 백분위 자료를 활용하였으며, 선택과목 간 능력차이 이상의 과도한 조정이 이루어지지 않도록 제약조건을 포함시켰고, 왜곡된 점수를 최소화하는 것을 목적으로 하였다. 그러나 이러한 제약은 실제적으로 점수조정의 폭을 협소하게 함으로써 왜곡의 정도를 최소화하는데 한계가 있었으며, 선택과목의 응시자 비중을 포함하지 않아서

소수가 응시한 선택과목도 동일한 비중으로 점수 조정에 영향을 미치게 되었다.

본 연구는 기존연구의 한계를 극복하기 위하여 다음의 네 가지 측면에서 선형계획모형을 보완하고자 하였다. 첫째, 현행 수능점수 보고체제에서 제공하는 각 영역별 및 각 선택과목별 표준점수, 백분위, 등급을 활용한다. 둘째, 남보우[6]의 연구에서 제시한 선형계획모형에서 능력 차이의 범위의 제약을 없애고, 능력 차이를 최소한으로 줄이도록 조정하는 목표계획법의 개념을 도입한다. 셋째, 목적함수를 왜곡된 점수의 가중합계로 하여 선택과목별 응시자 수의 비중이 포함되도록 한다. 넷째, 사탐 내의 11개 모든 선택과목의 표준점수를 조정대상으로 한다.

3. 점수조정을 위한 선형계획모형 개발

3.1 모형 개발을 위한 일반사항

3.1.1 측정이론에서 평균점수 차이 조정 방법
측정이론에서 평균차이를 조정하는 전통적인 방법은 평균방법(mean method)이다. 이는 평균이 같도록 조정하는 것이다. 두 검사의 분포가 같고 평균만 다를 때 두 검사의 평균차이로 모든 점수를 조정하는 모형을 활용할 수 있다. 평균차이로 조정하는 방법은 활용이 간편하므로 많은 대학들이 가산점수제 방식으로 채택할 수 있다. 검사 X의 평균을 $\mu(X)$, 검사 Y의 평균을 $\mu(Y)$ 라고 하면 검사 X의 특정점수 x 는 검사 Y의 점수척도로 표시한 y 값으로 식 (1)에 의하여 조정된다. 즉, 단일 집단이 검사 X와 검사 Y를 받았다면 X의 모든 점수에 Y의 평균에서 X의 평균을 뺀 차이만큼 가산하여 조정 후 평균이 같아지도록 조정하는 것이다.

$$y = x - \mu(X) + \mu(Y) \quad (1)$$

검사 X를 받은 집단을 1, 검사 Y를 받은 집단을

2라 하고 두 집단이 모두 공통검사 V를 받았다고 하자. 이때는 검사 X의 점수 x 를 검사 Y의 척도로 직접 조정할 수 없다. 검사 X의 점수 x 를 공통검사 V의 점수척도 v 로 조정하고, 공통검사 V의 점수 v 를 검사 Y의 점수척도 y 로 2단계에 걸쳐 조정하는 연쇄방법(chained method)를 활용할 수 있다. 즉, 공통검사 V를 매개로 하여 검사 X의 점수를 검사 Y의 척도로 조정할 수 있다. 검사 X의 평균을 $\mu_1(X)$, 검사 Y의 평균을 $\mu_2(Y)$, 집단 1의 공통검사 V의 평균을 $\mu_1(V)$, 집단 2의 공통검사 V의 평균을 $\mu_2(V)$ 라 하고, 식 (1)을 두 번 적용하면 다음 식이 된다.

$$v = x - \mu_1(X) + \mu_1(V) \quad (2)$$

$$y = v - \mu_2(V) + \mu_2(Y) \quad (3)$$

식 (2)를 식 (3)에 대입하면 식 (4)가 된다.

$$y = x + [\mu_2(Y) - \mu_1(X)] - [\mu_2(V) - \mu_1(V)] \quad (4)$$

식 (4)에 의하여 검사 X의 특정점수 x 는 검사 Y의 점수척도로 표시한 y 값으로 조정된다. 즉, 두 집단이 서로 다른 검사를 받은 점수를 조정하려면 두 검사의 평균차이를 더하고, 공통검사의 두 집단 평균차이를 빼서 조정하면 된다. 검사 X와 공통검사 V를 동시에 응시한 집단을 기준으로 검사 X의 점수를 공통검사 V의 점수척도로 조정하는 것은 단일집단설계에 의한 조정이다. 즉, 단일집단의 상이한 두 검사 결과를 조정하는 것이다. 검사 X의 점수 x 가 공통검사의 점수척도 v 로 조정되고, v 를 검사 Y의 점수척도 y 로 조정하는 연쇄방법을 적용할 때 y 는 검사 X의 점수 x 를 조정한 점수이다. 두 검사의 서로 다른 척도를 공통검사를 통하여 하나의 검사 척도로 조정하였기 때문에 비교 가능하게 한 것이다.

세 개 이상의 검사를 측정이론의 평균방법으로 조정하려면 모든 검사에 공통적으로 적용되는 공통검사가 있거나 공통응시집단이 있어야 한다. 그러나 두 검사 각 쌍 사이에만 공통과목이나 공통응

시집단이 있는 경우, 점수조정이 불가능하다. 예를 들어 검사 X의 평균점수가 검사 Y보다 높아야 하고, 검사 Y의 평균점수가 검사 Z보다 높아야 하며, 검사 Z의 평균점수는 검사 X보다 높아야 한다면 불일치가 발생하기 때문이다. 이러한 경우 조정되지 못한 왜곡된 점수를 최소화하도록 조정하는 것이 최선이다. 따라서 점수조정을 위하여 선형계획 모형 등의 최적화 방법을 활용해야 한다.

3.1.2 선택과목 점수조정 상황예시

3과목 이상의 점수조정 상황을 나타내기 위해서 2010학년도 수능에서 종합 평균등급 2.00~3.99 등급자 중 국사과목, 윤리과목, 경제과목의 응시자 수, 평균점수 및 각 과목 쌍에 공통응시자 수 및 공통응시자의 각 과목 평균점수를 <그림 1>에 예시적으로 나타냈다. 국사과목에 응시한 응시자는 18,854명이고, 국사과목의 표준점수 평균은 54.05점이다. 또한 국사과목 응시자 중 윤리과목에도 응시한 공통응시자는 7,948명이며 이들의 국사과목의 평균점수는 53.79점이고, 이들의 윤리과목 평균점수는 61.18점이다. 국사과목 응시자 중 경제과목에 응시한 공통응시자는 4,023명이고 이들의 국사과목 평균점수는 53.50점이며, 경제과목 평균점수는 58.02점이다. 마

찬가지로 윤리과목과 경제과목에 대한 응시집단별 응시자 수 및 평균점수를 <그림 1>에 나타냈다.

윤리과목 평균점수와 같아지도록 국사 점수를 조정하려면 식 (1)에 의하여 국사의 모든 점수에 $58.66 - 54.05 = 4.61$ 점을 가산하면 된다. 마찬가지로 경제의 모든 점수에 4.01점을 가산하면 윤리의 평균점수와 같게 된다. 그러나 이렇게 과목 평균점수 차이로 조정하려면 두 과목에 응시한 응시자가 같아야 한다. 즉, 동일한 집단이 서로 다른 두 검사를 받은 경우 두 검사의 평균을 동일하도록 조정하면 검사결과를 상호 비교할 수 있다. 그러나 각 과목의 응시자가 동일하지 않기 때문에 과목 쌍 간의 공통응시자가 응시한 두 과목의 평균점수 차이로 조정해야 한다.

윤리와 국사 공통응시자의 윤리 평균점수와 국사 평균점수의 차이가 $61.18 - 53.79 = 7.39$ 점이므로, 국사 점수에 7.39점을 가산하면 공통응시자의 평균점수가 같아진다. 이 경우 국사에 응시한 다른 응시자의 점수도 7.39점 가산되어야 국사과목 내에서 점수체계가 유지되며 공정하다. 그러면 국사의 평균점수는 윤리의 평균점수보다 $54.05 + 7.39 - 58.66 = 2.78$ 점이 높게 된다. 이는 국사 전체응시자의 평균점수가 윤리 전체응시자의 평균점수보다 2.78점이



<그림 1> 과목과 응시집단 간의 표준점수 평균 예시

크도록 조정하여야 하는 것을 의미하고, 국사 전체 응시자의 평균 수학능력 또는 사회탐구능력이 윤리과목의 능력보다 2.78점 높다고 볼 수 있다.

윤리와 경제를 동시에 응시한 공통응시자를 기준으로 경제의 점수를 윤리 점수와 비교 가능하도록 조정하려면 경제의 모든 점수에 3.90점을 가산해야 하고, 이 때 경제의 평균점수는 윤리의 평균점수보다 0.11점 낮게 된다. 이렇게 조정하면 국사의 평균점수는 윤리보다 2.78점 높고, 윤리의 평균점수는 경제보다 0.11점 높으므로, 국사의 평균점수는 경제 평균점수보다 2.89점 높아야 한다. 그러나 경제와 국사의 공통응시자를 기준으로 직접 조정하면 국사에 $58.02 - 53.50 = 4.52$ 점을 가산하고, 국사의 평균점수는 경제의 평균점수보다 3.92점 높게 된다. 따라서 국사와 경제의 평균점수 차이는 직접 조정할 경우 3.92점이 되지만, 윤리과목을 통하여 조정할 경우는 2.89점이 되어 일치하지 않는다.

국사의 평균점수를 경제의 평균점수보다 3.92점 높게 조정하면 국사와 경제 공통응시자의 두 과목 평균이 같게 되어 점수왜곡이 없다. 또한 윤리를 경제보다 0.11점 높게 조정하면 윤리와 경제 사이에 점수 왜곡이 없다. 이 경우 국사의 평균점수가 윤리의 평균점수보다 3.81점이 높게 되고, 국사와 윤리를 직접 조정할 때의 2.78점보다 1.03점의 왜곡이 발생한다. 국사와 윤리의 공통응시자는 7,948명이므로 왜곡된 점수의 총합은 8,186점이 된다. 또한 국사의 평균점수를 경제의 평균점수보다 2.89점만 높게 조정하면 공통응시자의 국사 평균점수는 1.03점 만큼 낮게 조정되어 점수가 왜곡된다. 이 경우 국사와 경제의 공통응시자는 4,023명이므로 왜곡된 점수의 총합은 4,144점이 된다. 따라서 두 가지 방법 중에서 왜곡된 점수의 총합이 적은 조정방법이 바람직하다고 할 수 있다. 이러한 문제에는 어떠한 조정방법을 적용한다고 하더라도 완전하게 조정하여 왜곡된 점수를 모두 없앨 수는 없다.

점수 조정 전에 국사 및 윤리를 동시에 응시한 7,948명의 국사 평균점수가 윤리 평균점수보다 7.39점이 낮으므로 왜곡된 점수의 합은 58,736점이고,

왜곡된 점수의 총합은 107,851점이다. 조정을 통하여 왜곡된 점수는 4,144점으로 감소하였으므로 점수왜곡 현상이 96.2%완화되었다. 응시인원을 고려하지 않은 왜곡된 평균점수의 합은 조정 전 15.81점에서 1.03점으로 93.5%가 완화되었다.

기존의 측정이론에서 이러한 문제는 완전한 조정 불가능하므로 더 이상 연구를 진전하지 않았다. 그러나 본 연구는 완전하게 조정이 되지 않는 경우에도 최적의 방법으로 조정하는 접근방법을 도입하여 선형계획모형을 개발하고자 한다.

3.1.3 변수정의

I, J : 과목명

i, j : 과목명 또는 과목 I, 과목 J의 응시집단

ij, ji : 과목 I와 과목 J를 응시한 공통응시집단

n_{ij} : 과목 I와 과목 J를 응시한 공통응시자 수

x_i : 과목 I의 임의의 점수

y_j : 과목 J의 점수 x_i 가 과목 J의 점수척도로 조정된 값

$\mu(I)$: 집단 I(과목 I)의 평균점수

k_i : 집단 i(과목 I)의 조정된 평균점수

$\mu_{ij}(I)$: 공통집단 ij의 과목 I의 평균점수

r_{ij} : 과목 I와 과목 J의 평균점수 차이의 조정 목표치(집단 i, j의 능력 차이)

u_{ij} : 과목 I와 과목 J의 조정 목표치 r_{ij} 에 부족한 조정점수(왜곡된 평균점수)

s_{ij} : 과목 I와 과목 J의 조정 목표치 r_{ij} 를 초과한 조정점수(왜곡된 평균점수)

P : r_{ij} 가 비음이면서 i와 j가 같지 않은 i, j의 순서쌍 집합

K : 선택과목 전체 응시자의 점수평균

L : 선택과목 수

3.2 평균점수 차이 조정을 위한 선형계획모형

3.2.1 모형의 개발

평균방법에 의하면 두 과목을 응시한 공통집단의 두 과목 평균이 같도록 조정해야 한다. 공통응

시집단을 기준으로 과목 I의 점수를 과목 J의 척도로 조정한다면, 식 (1)을 적용하여 공통응시집단의 과목 I의 임의의 점수 x_i 는 과목 J의 점수척도 y_j 로 다음과 같이 조정된다.

$$y_j = x_i + [\mu_{ij}(J) - \mu_{ji}(I)] \quad (5)$$

과목 I의 모든 점수를 식 (5)로 조정하면 과목 I의 조정된 점수평균은 $\mu_i(I) + [\mu_{ij}(J) - \mu_{ji}(I)]$ 가 된다. 과목 J 전체의 평균점수 $\mu_j(J)$ 에서 과목 I 전체의 조정된 평균점수를 뺀 차이 r_{ji} 는 다음의 식 (6)과 같이 된다.

$$r_{ji} = [\mu_j(J) - \mu_i(I)] - [\mu_{ij}(J) - \mu_{ji}(I)] \quad (6)$$

r_{ji} 는 공통응시집단의 두 과목 평균이 동일하도록 식 (5)에 의하여 모든 점수를 조정할 후 과목 J의 평균점수와 과목 I의 조정된 평균점수의 차이를 나타낸 것이다. 이는 과목 J에 응시한 전체집단과 과목 I에 응시한 전체집단의 능력 차이를 나타낸다고 할 수 있다. 식 (4)는 공통검사를 활용한 점수조정연쇄방법을 나타냈고, 식 (6)은 공통응시자를 활용하여 점수조정의 폭을 계산한 것이다.

과목 I의 조정된 평균점수를 k_i 라 하고 과목 J의 조정된 평균점수를 k_j 라 할 때 다음 식 (7)을 만족하면, 공통응시집단의 두 과목 평균점수가 동일하게 되어 과목별 응시집단 간 능력 차이가 완전하게 반영되고 점수 왜곡이 없게 된다.

$$k_i - k_j = r_{ji} \quad (7)$$

식 (7)을 만족하도록 k_i 를 구한 후, 식 (8)과 같이 과목 I의 모든 점수 x_i 에 $[k_i - \mu_i(I)]$ 를 가산하면 과목 I 전체의 조정된 평균점수는 k_i 가 된다.

$$y_j = x_i + [k_i - \mu_i(I)] \quad (8)$$

식 (8)은 조정된 평균점수에서 조정 전 평균점수의 차이를 모든 점수에 가산하여 조정하는 방법이

고, 식 (5)는 공통응시집단의 두 과목 평균점수 차이를 모든 점수에 가산하여 조정하는 방식이므로 다르다. 그러나 식 (6)과 식 (7)로부터 다음의 식 (9)가 성립한다.

$$k_i - \mu_i(I) = [k_j - \mu_j(J)] + [\mu_{ij}(J) - \mu_{ji}(I)] \quad (9)$$

식 (9)를 식 (8)에 대입하면, 식 (10)이 된다.

$$y_j = x_i + [k_j - \mu_j(J)] + [\mu_{ij}(J) - \mu_{ji}(I)] \quad (10)$$

과목 J의 점수척도로 과목 I의 점수를 조정한다면, 과목 J의 평균점수는 변하지 않는다. 이러한 경우 식 (10)에서 $k_j = \mu_j(J)$ 로 놓으면, 식 (10)은 식 (5)와 일치한다. 과목 J의 점수에 모두 $[k_j - \mu_j(J)]$ 를 가산하여 조정할 경우, 과목 I의 점수도 과목 J의 조정 점수만큼 더 조정하여야 공통응시집단의 평균점수가 같아진다. 예를 들어 <그림 1>에서 윤리과목의 모든 점수에 2점의 가산점을 주어 평균점수가 2점 높아졌다면, 조정 전의 점수왜곡 7.39점에 국사과목 점수에 가산하여야 공통응시집단의 평균점수가 같게 된다. 식 (10)은 곧 식 (8)이므로 k_i 를 구한 후 식 (8)을 적용하여 점수를 조정하면 공통응시집단의 평균점수가 같아져서 각 과목의 전체 응시집단의 능력 차이가 반영된다.

과목 I의 점수를 과목 J의 점수척도로 조정하기 위하여 과목 I에 부여해야 하는 가산점의 크기는 식 (5)에 의하여 공통응시집단의 과목 간 평균점수 차이 $[\mu_{ij}(J) - \mu_{ji}(I)]$ 이다. 각 과목의 가산점 크기를 변수로 하여 과목 간 가산점 차이가 공통응시집단의 평균점수 차이가 되도록 식 (7)과 같은 모양으로 모형을 구성할 수도 있다. 그러나 식 (6)과 식 (7)에 의하여 각 과목의 평균점수 수준을 결정하고, 식 (8)을 활용하여 가산점을 계산하려는 이유는 세 가지이다. 첫째, 각 선택과목별 응시집단의 상대적인 능력을 측정할 수 있고, 둘째, 조정 전과 조정 후의 점수 왜곡정도를 용이하게 계산할 수 있으며, 셋째,

조정된 평균점수를 기준으로 가산점을 계산하는 것이 편리하며 해석하기가 용이하기 때문이다.

n 개의 선택과목이 있을 때 각 과목 쌍 간의 평균점수의 차이는 nC_2 개가 있으므로, <그림 1>의 예시에서 설명한 바와 같이 일반적으로 식 (7)을 모두 만족시키지는 못한다. 식 (7)을 모두 만족하는 해가 없을 때 측정이론에서는 조정이 불가능한 문제로 제외하였으나, 최적화의 개념을 적용하여 가능한 가깝게 조정하도록 할 필요가 있다.

모든 i, j 에 대하여 $r_{ij} = -r_{ji}$ 이므로 식 (7)에서 비음인 r_{ij} 를 선택할 수 있으며, 음의 r_{ij} 에 대한 식은 중복이므로 고려하지 않아도 된다. 여기서 집합 P 를 r_{ij} 가 비음이면서 i 와 j 가 같지 않은 다음과 같은 i, j 의 순서쌍 집합이라고 하자.

$$P = \{(i, j) | r_{ij} \geq 0, i \neq j\} \quad (11)$$

r_{ij} 가 비음인 경우 조정 후 과목 I와 과목 J의 평균점수의 차이($k_i - k_j$)가 능력차이 r_{ij} 만큼 되도록 하는 것이 목표이다. 평균점수 차이($k_i - k_j$)가 목표 r_{ij} 에 도달하지 못한 부족한 양을 u_{ij} 라 하고, 목표를 초과한 양을 s_{ij} 라 하면 식 (7)에 비음인 u_{ij} 와 s_{ij} 를 포함하여 식 (12)로 나타낼 수 있다. 목표에 부족하거나 초과하여 조정되지 못한 왜곡된 점수를 최소화하는 경우 u_{ij} 와 s_{ij} 가 모두 양수가 되지는 않는다. u_{ij} 와 s_{ij} 가 모두 0이면 능력 차이가 모두 반영되도록 조정된 것이며, u_{ij} 가 양이면 목표에 부족하게 조정된 것이고, s_{ij} 가 양이면 목표를 초과하여 조정된 것이다. 만일 u_{ij} 가 r_{ij} 이면 과목 I와 과목 J의 평균점수가 같도록 조정된 것이다. u_{ij} 가 r_{ij} 보다 크면 능력이 큰 집단의 평균점수가 오히려 낮게 역방향으로 조정된 것이다. 따라서 s_{ij} 를 모두 0으로 하고, u_{ij} 를 r_{ij} 보다 작거나 같게 제한하면 초과하거나 역방향으로 조정되지 않는다. 이러한 제약은 점수조정에 대한 사회적 합의에 따라 추가할 수 있다.

$$k_i - k_j + u_{ij} - s_{ij} = r_{ij} \quad (12)$$

식 (12)에서 조정 전의 수능점수의 평균값으로 $k_i = \mu_i(I)$, $k_j = \mu_j(J)$ 를 놓으면($u_{ij} - s_{ij}$)는 조정 전의 공통응시집단의 두 과목 평균점수의 차이인 [$\mu_j(J) - \mu_i(I)$]가 되어 왜곡된 점수를 나타낸다. u_{ij} 와 s_{ij} 가 모두 양수 일 수 없으므로($u_{ij} + s_{ij}$)에 공통응시자 수를 곱하여 조정 전 왜곡된 점수의 합을 구할 수 있다.

식 (7)은 두 집단 간의 평균차이만 정하였으므로, (k_i, k_j)값을 확정할 수 없기 때문에 어느 하나의 값을 정하고 조정하면 된다. 예를 들어 과목 I의 점수를 기준으로 다른 선택과목들의 점수를 조정한다면 $k_i = \mu_i(I)$ 로 놓으면 된다. 조정 후 과목 전체의 점수평균이 K 가 되도록 조정하려면 과목 I의 응시자 수는 n_{ii} , 과목 J의 응시자 수는 n_{jj} 이므로 식 (13)을 만족하면 된다. 식 (13)으로 표현하면 조정된 전체 점수의 평균수준을 정할 수 있어 바람직하고, 평균점수 수준인 K 값을 조정하여 조정의 기준이 되는 과목의 평균점수 수준도 정할 수 있어서 편리하다.

$$(n_{ii}k_i + n_{jj}k_j) / (n_{ii} + n_{jj}) = K \quad (13)$$

식 (12)에서 과목 I와 과목 J의 조정목표에 비하여 부족하거나 초과하여 조정된 경우, 완전하게 조정되지 못하였으므로 u_{ij} 또는 s_{ij} 만큼 평균점수가 왜곡된 것이다. 왜곡된 점수의 합은 과목 I와 과목 J의 공통응시자 수를 u_{ij} 및 s_{ij} 에 곱하면 된다. 따라서 모든 과목의 각 쌍 사이에 왜곡된 점수의 총합은 식 (14)와 같고 이를 최소화할 필요가 있으므로 목적함수가 된다.

$$\sum_{(i,j) \in P} n_{ij}(u_{ij} + s_{ij}) \quad (14)$$

일반성을 유지하면서 과목 I와 과목 J에 대하여 논의하였으므로, 과목이 다수인 경우 조정하지 못한 왜곡된 점수의 총합을 최소화하는 문제는 다음의 선형계획모형으로 정식화된다.

[모형 1] Minimize $\sum_{(i,j) \in P} n_{ij}(u_{ij} + s_{ij})$
 Subject to
 $k_i - k_j + u_{ij} - s_{ij} = r_{ij}$ for all $(i, j) \in P$
 $\sum_{i=1}^L n_{ii}k_i / \sum_{i=1}^L n_{ii} = K$
 $0 \leq u_{ij}, 0 \leq s_{ij}$ for all $(i, j) \in P$
 $0 \leq k_i$ for all i

전술한 바와 같이 [모형 1]은 남보우[6]가 제시한 선형계획모형을 보완하여 능력 차이에 대한 제약을 하는 대신 능력 차이를 최소한으로 줄이도록 조정하는 목표계획법의 개념을 도입하였다. 또한 왜곡된 점수의 가중합계를 목적함수로 설정하여 선택과목별 응시자 수의 비중이 포함되도록 하였다. [모형 1]에서 $s_{ij} = 0$ 으로 하고, $u_{ij} \leq r_{ij}$ 로 제한하며, 공통응시자 수를 곱하지 않은 목적함수로 구성한 것이 기존의 모형[6]이다.

[모형 1]은 3개 이상의 과목에서도 각 과목 쌍에 동시에 응시한 공통응시집단의 과목 간 평균점수의 차이가 최소화되도록 조정하는 것이다. [모형 1]은 공통과목이 있는 경우뿐만 아니라 모두 선택과목으로만 이루어진 경우에도 조정할 수 있다. 모두 선택과목으로 이루어진 3과목 이상의 점수조정은 측정이론의 조정방법으로 해결할 수 없었으나, 본 연구에서 [모형 1]을 제시함으로써 이 문제에 대한 새로운 조정방법을 도출하였다. 또한 기존의 선형계획모형에 목표계획 개념을 도입함으로써 모형이 유연성을 갖도록 하였다.

3.2.2 예시 사례에 대한 선형계획모형의 구성
 <그림 1>에 나타난 상황을 예시 사례로 하여 선형계획모형을 적용해 봄으로써 모형 구성에 대한 과정을 명확하게 하고자 한다. 국사과목을 1, 윤리과목을 2, 경제과목을 3이라고 하고 식 (6)을 적용하면, $r_{12} = 2.78, r_{13} = 3.92, r_{23} = 0.11$ 이다. 전체응시자 수에 대한 국사과목 응시자 수의 비율은 0.22이고, 윤리과목은 0.48, 경제과목은 0.30이다. 전체응시자의 평균점수를 60점으로 조정하려면 [모형 1]은 다음과 같이 구성된다.

[모형 2]
 Minimize
 $7948(u_{12} + s_{12}) + 4023(u_{13} + s_{13}) + 7931(u_{23} + s_{23})$
 Subject to
 $k_1 - k_2 + u_{12} - s_{12} = 2.78$
 $k_1 - k_3 + u_{13} - s_{13} = 3.92$
 $k_2 - k_3 + u_{23} - s_{23} = 0.11$
 $0.22k_1 + 0.48k_2 + 0.30k_3 = 60$
 $0 \leq k_1, 0 \leq k_2, 0 \leq k_3$
 $0 \leq u_{12}, 0 \leq u_{13}, 0 \leq u_{23}$
 $0 \leq s_{12}, 0 \leq s_{13}, 0 \leq s_{23}$

조정 전에 각 과목의 평균점수로 $k_1 = 54.05, k_2 = 58.66, k_3 = 54.65$ 로 놓으면, $(u_{12} - s_{12}) = 7.39$ 가 되고, $(u_{13} - s_{13}) = 4.52, (u_{23} - s_{23}) = -3.90$ 이 되어 왜곡점수의 총합은 107,851점이다. 엑셀의 최적화 도구인 해 찾기(Excel Solver) 기능을 이용하여 최적해를 구하면 $k_1 = 62.20, k_2 = 59.42, k_3 = 59.31, u_{13} = 1.03$ 이 되고 나머지 변수는 모두 0이다. 이는 식 (8)에 의하면 국사에 8.15점의 가산점, 윤리에 0.76점의 가산점, 경제에 4.66점의 가산점을 주어야 하는 것이다. 최적해의 목적함수 값은 조정 후의 왜곡점수의 총합 4,144점이 되어 조정 전의 107,851점보다 96.2%가 해소되었다. 이는 <그림 1>의 예시 사례에서 논의한 바와 같다.

4. 사회탐구영역 선택과목 점수조정

[모형 1]을 활용하여 수능 선택과목의 점수를 조정하고 그 결과가 타당한지 분석하고자 한다. 이를 위하여 각 탐구영역의 선택과목 표준점수 및 백분위 점수를 모두 분석할 수도 있으나, 제시된 모형의 유용성을 검토하는 것이 목적이므로 탐구영역 중 하나인 사탐 11개 전체의 선택과목별 모든 응시자의 표준점수를 조정하고자 한다. 또한 특정 대학의 입학전형 경쟁자는 수학능력이 일정범위에 있는 경우가 보통이므로 언어, 수리, 외국어 영역 수능 종합평균 등급이 2.00~3.99인 대상자의 점수를

조정하여 대학이 조정방법을 활용하는데 참고하게 하고자 한다. 다른 탐구영역의 선택과목 또는 백분위의 점수조정은 동일한 방법을 활용하면 되기 때문에 본 연구에서는 생략하고자 한다.

우선 사람 선택과목의 상관계수 분석을 통하여 점수조정이 타당한가를 검토하고, 선택과목간의 평균차이가 존재하는지 분석한다. 평균차이가 존재하면 측정이론의 연쇄방법을 적용하여 평균차이만큼 점수를 조정할 결과와 선행계획모형을 적용하여 조정할 결과를 비교하고자 한다.

4.1 상관계수 및 평균차이 분석

두 검사가 동일한 특성을 측정해야 두 검사의 점수를 비교 가능하도록 조정하는 것이 타당하다[16]. 사람의 선택과목 간 점수를 조정하려면 각 선택과목들이 “대학수학능력” 또는 “사회탐구능력”이라고 하는 동일한 특성을 측정해야 한다[7]. 각 선택과목이 동일한 특성을 측정하는지 구인타당도를 검토하기 위해서 상관계수법을 활용할 수 있으며, 각 선택과목 사이의 상관계수가 높으면 구인타당도가 있다고 판단한다. 따라서 선택과목의 각 쌍 및 공통

과목과의 상관관계를 분석하여 점수조정의 타당성을 검토할 필요가 있다.

사람의 11개 선택과목 표준점수 간 피어슨 상관계수를 통해 상관관계 정도를 확인해 본 결과 <표 1>와 같이 유의수준 0.01에서 모든 상관계수가 유의하게 나타났다. 또한 언어, 수리, 외국어 영역(이하 “언수외”라 한다)을 공통과목으로 선정하는 것이 바람직하므로[5], 언수외와 각 선택과목 표준점수 사이의 상관계수를 검토한 결과 유의수준 0.01에서 유의하게 나타났다. 일반적으로 상관계수가 0.4~0.6이면 상관이 있고, 0.6~0.8이면 상관이 높은 것으로 판단한다[12]. 이러한 기준을 <표 1>에 적용하면 선택과목 표준점수 사이에는 상관이 있고, 선택과목과 언수외 표준점수 사이에는 상관이 높으므로, 선택과목 간 점수조정을 하는 것은 타당하다고 볼 수 있다.

2010학년도 사람을 응시한 총 377,089명의 각 선택과목 응시자 수와 각 선택과목 쌍에 응시한 공통응시자 수를 <표 2>에 나타냈다. 수능의 특정 응시자는 사람의 선택과목을 4과목까지 선택할 수 있으므로 과목 쌍 간의 공통응시자가 있게 된다. 표에서 윤리과목 응시자는 1,888백 명이고, 국사와

<표 1> 선택과목 표준점수 쌍 간의 상관계수

선택과목	윤리	국사	한국지리	세계지리	경제지리	근현대사	세계사	법과사회	정치	경제	사회문화	언수외(공통)
윤리	1.0											
국사	.68**	1.0										
한국지리	.54**	.64**	1.0									
세계지리	.59**	.64**	.64**	1.0								
경제지리	.45**	.49**	.56**	.55**	1.0							
한국근현대사	.61**	.74**	.54**	.54**	.40**	1.0						
세계사	.58**	.74**	.49**	.59**	.41**	.67**	1.0					
법과사회	.60**	.68**	.52**	.48**	.47**	.56**	.54**	1.0				
정치	.64**	.59**	.52**	.52**	.47**	.56**	.52**	.66**	1.0			
경제	.57**	.61**	.54**	.47**	.46**	.49**	.52**	.59**	.56**	1.0		
사회문화	.64**	.61**	.55**	.48**	.46**	.56**	.46**	.57**	.61**	.56**	1.0	
언수외(공통과목)	.74**	.68**	.68**	.67**	.62**	.68**	.65**	.70**	.73**	.67**	.74**	1.0

주) **: 상관계수가 0.01수준에서 유의함.

〈표 2〉 두 선택과목에 동시에 응시한 응시자 수(백 명)

선택과목	윤리	국사	한국 지리	세계 지리	경제 지리	근현대사	세계사	법과 사회	정치	경제	사회 문화	연수 외
윤리	1,888											
국사	300	696										
한국지리	1,143	325	2,481									
세계지리	123	53	354	491								
경제지리	125	36	499	181	613							
근현대사	1,146	458	1,477	194	209	2,333						
세계사	144	114	148	59	27	246	387					
법과사회	262	79	305	38	48	305	41	636				
정치	561	153	681	89	119	703	83	253	1,278			
경제	287	130	469	72	161	393	55	105	234	848		
사회문화	1,428	389	1,854	276	400	1,694	209	421	871	588	2,802	
연수 외	1,888	695	2,481	491	613	2,333	387	636	1,278	848	2,802	3,771

〈표 3〉 응시집단별 선택과목 표준점수 평균

과목 응시집단	윤리	국사	한국 지리	세계 지리	경제 지리	근현대사	세계사	법과 사회	정치	경제	사회 문화	연수 외
윤리	50.00	49.20**	48.72**	47.72**	47.59**	49.44**	49.29*	48.78**	49.66**	49.22**	49.58**	49.71**
국사	53.35**	50.01	53.48**	51.40**	53.43**	54.17**	53.48**	54.52**	54.57**	56.90**	52.42**	52.79**
한국지리	49.68**	49.85**	50.02	50.48**	50.28**	49.47**	48.72**	49.58**	49.35**	49.78**	50.00**	50.38**
세계지리	48.57**	48.67**	51.58**	50.01	50.73**	49.61**	49.26**	47.39**	49.20**	48.78**	48.47**	49.83**
경제지리	49.21**	49.34**	53.22**	52.57**	50.01	50.24**	48.92**	49.49**	50.01**	50.54**	50.60**	51.68**
근현대사	50.43**	51.43**	49.85**	50.18**	49.92**	50.00	51.35	50.25**	50.40**	50.18**	50.09**	50.56**
세계사	49.14*	50.64**	48.18**	48.44**	47.89**	51.41	49.98	47.96**	49.13**	48.80**	47.48**	48.75**
법과사회	48.00**	49.12**	48.17**	46.10**	47.56**	48.64**	47.52**	49.95	49.65**	48.79**	47.95**	48.82**
정치	50.76**	50.57**	49.78**	50.17**	50.40**	50.74**	50.46**	51.53**	49.99	49.96**	50.41**	50.98**
경제	51.26**	52.79**	52.17**	51.01**	51.79**	51.68**	50.82**	52.16**	52.43**	50.00	51.90**	52.81**
사회문화	49.66**	49.02**	49.54**	49.27**	49.67**	49.37**	48.64**	49.14**	49.46**	49.50**	49.92	50.19**
연수 외	50.00**	50.01**	50.02**	50.01**	50.01**	50.00**	49.98**	49.95**	49.99**	50.00**	49.92**	50.18

주) * : p < 0.05, ** : p < 0.01.

윤리의 공통응시자는 300백 명을 나타낸 것이다. 과목 쌍 간의 공통응시자는 표의 대각선을 중심으로 대칭이므로 아래 삼각부분에만 나타냈다.

〈표 3〉에 응시집단별 선택과목과 연수의 표준점수 평균을 제시하였다. 윤리과목 응시집단이 윤리과목에서 획득한 표준점수 평균은 50.00점이고, 윤

리과목 응시자 중 국사과목 시험의 표준점수 평균은 49.20점이며, 한국지리 시험의 표준점수 평균은 48.72점이고, 연수외의 표준점수 평균은 49.71점을 나타낸 것이다. 또한 국사과목과 윤리과목 공통응시자의 국사과목 평균점수는 49.20이고, 윤리과목 평균점수는 53.35점을 나타낸 것이다. 이때 연수 외

의 표준점수는 각 선택과목의 표준점수 척도와 동일하도록 1/2을 곱한 후 평균점수를 구하였다.

과목 각 쌍의 공통응시집단이 획득한 각 과목의 평균점수 사이에 차이가 있는지 대응 T-검정을 할 수 있다. 엑셀의 't-검정 : 쌍체비교' 기능을 활용하여 국사와 윤리과목 공통응시집단 30,048명 각 수험생의 국사점수에서 윤리점수의 차이를 구하고, 차이의 평균이 0인지 검정하면, t 통계량은 -82.46이 된다. 유의수준 0.01에서 자유도 30,047의 양측 임계치는 2.576이므로 공통응시집단의 국사점수와 윤리점수의 평균이 같다는 가설을 기각하고 매우 낮은 P값을 가진다. 따라서 국사와 윤리과목 공통응시집단의 국사점수와 윤리점수의 평균값에 차이가 있다. <표 4>에서 ** 부분은 0.01유의수준에서 평균에 차이가 없다는 가설을 기각하는 과목 쌍이며, * 부분은 0.05유의수준에서 기각하는 과목 쌍이

다. 세계사 과목과 근현대사 과목은 평균점수에 차이가 없다는 것을 기각하지 못하며, 윤리과목과 세계사 과목은 0.01유의수준에서는 기각하지 못하지만, 0.05유의수준에서는 기각한다. 나머지 53개의 쌍에는 평균점수 사이에 차이가 없다고 볼 수 없으므로 전체적으로 조정이 필요하다.

4.2 측정이론의 연쇄방법과 선형계획모형 조정결과 비교

4.2.1 측정이론의 연쇄방법을 적용한 조정결과 <표 3>의 자료 중 해당과목 평균점수와 공통과목인 연수의 평균점수를 <표 4>에 보기 쉽도록 다시 정리하였다. 또한 <표 4> 측정이론의 연쇄방법을 적용한 조정결과(가산점)를 제시하였다. 세계사를 기준으로 측정이론의 연쇄방법인 식 (4)를 적용

<표 4> 전체응시자의 표준점수 조정 결과

선택과목	적용방법 및 기준	측정이론을 적용한 조정 (공통과목을 가교검사로 활용하여 연쇄방법 적용)			선형계획모형을 적용한 조정 (과목 쌍 간 능력차이 반영) (조정결과 : 가산점)		
		해당과목 평균점수	공통과목 평균점수	조정결과 (가산점)	순수 선택과목		
					[모형 1]	현실제약	
윤리		50.00	49.71	0.94	0.94	0.79	0.44
국사		50.01	52.79	4.01	4.01	4.94	3.17
한국지리		50.02	50.38	1.59	1.59	1.41	0.44
세계지리		50.01	49.83	1.05	1.05	1.21	0.44
경제지리		50.01	51.68	2.90	2.90	2.10	0.76
근현대사		50.00	50.56	1.79	1.79	1.78	0.44
세계사		49.98	48.75	0.00	0.00	0.87	0.44
법과사회		49.95	48.82	0.10	0.10	0.00	0.00
정치		49.99	50.98	2.22	2.22	0.89	0.78
경제		50.00	52.81	4.04	4.04	3.35	1.94
사회문화		49.92	50.19	1.50	1.50	0.95	0.44
왜곡점수 합계	조정 전(백점)			25,505	25,505	25,470	
	조정 후(백점)			0	0	5,838	14,350
	개선비율(%)			100	100	77.08	43.66

주) 현실적 제약은 과목 각 쌍의 왜곡정도가 현재보다 나빠지지 않으며, 왜곡정도 이상으로 과도하게 조정되지 않도록 조건을 추가한 것이다.

하기 위하여 각 과목과 세계사 과목을 비교하였고 공통과목을 가교점사로 활용하였다. 예를 들면 국사의 조정점수인 가산점은 $[49.98-50.01]-[48.75-52.79] = 4.01$ 로 계산된다. 국사 응시집단은 공통과목에서 평균 52.79점을 받았고, 세계사 응시집단은 공통과목에서 평균 48.75점을 받았으므로 국사 응시집단의 공통과목 능력이 세계사 응시집단의 공통과목 능력보다 4.04점 높다. 그러나 국사의 평균 점수가 세계사의 평균점수보다 0.03점 높으므로 가산점은 4.01점으로 계산되어야 한다. 이와 같은 방법으로 모든 선택과목의 가산점을 계산하였다. 조정을 하지 않는다면 각 과목 응시자가 가산점을 받지 못하여 왜곡된 점수의 합은 25,505백점이 된다. 여기서 왜곡된 점수는 응시집단 간 공통과목 능력과 해당과목 점수의 차이를 조정하지 않아서 발생한 것이다. 공통과목을 활용하여 연쇄방법으로 조정하면 두 선택과목은 공통과목을 가교로 하여 평균이 같도록 조정되기 때문에 조정 후에 왜곡된 점수의 합은 0이 된다.

4.2.2 선형계획모형을 적용한 조정결과

선형계획모형을 적용하여 점수를 조정한 결과를 <표 4>에 나타냈다. [모형 1]은 공통과목을 활용할 때 측정이론의 연쇄방법을 포함하는 모형이다. 즉, 각 선택과목과 공통과목 쌍 간의 능력 차이만 포함하고, 선택과목 각 쌍 간의 능력 차이를 포함하지 않도록 모형을 구성하고 최적해를 구하면 된다. 이 경우 각 선택과목과 공통과목이 하나씩 대응되므로 식 (7)이 만족되어 측정이론의 연쇄방법과 같게 조정된다. 언수의 공통과목만을 활용한 선형계획모형의 최적해는 <표 4>에 제시한 바와 같이 측정이론을 적용한 조정결과와 같다. 공통과목을 기준으로 할 때 조정 전 왜곡점수 총합은 25,505백점이 되고, 완전하게 조정되었으므로 조정 후에 점수의 왜곡은 없다. 다만, 여기서 왜곡점수는 공통과목의 점수차이가 반영되지 않아서 발생하는 왜곡점수이다.

공통과목이 없이 순수 선택과목만 있는 경우 [모형 1]을 적용하여 최적해를 구하고 가산점을 계산

하여 <표 4>에 나타냈다. 법과사회과목을 기준으로 할 때 우리는 0.79점, 국사는 4.94점을 가산하여 조정하는 것이 최적으로 나타난다. 이 때 조정 전 왜곡점수의 총합은 25,470백점에서 최적화 후 5,838백점으로 77.08%가 개선되었다. 다만 이 왜곡점수는 선택과목 각 쌍 공통응시집단의 선택과목 평균 점수 차이가 반영된 것으로 선택과목별 응시집단의 공통과목 평균점수 차이를 반영한 왜곡점수와 기준이 다르다.

점수조정에 있어서 현실적인 제약조건이 추가될 수 있다. 과목 각 쌍의 왜곡정도가 현재보다 나빠지지 않고, 왜곡정도 이상으로 과도하게 조정되지 않도록 현실적인 제약조건 하에서 점수를 조정하여 <표 4>에 나타냈다. 각 쌍의 왜곡정도가 현재보다 나빠지지 않는 조건은 [모형 1]에서 조정되지 못한 점수 u_{ij} 를 공통응시집단의 능력차이 r_{ij} 보다 적거나 같게 제한하고, 왜곡정도 이상으로 과도하게 조정되지 못하도록 초과 조정점수 s_{ij} 를 0으로 하여 최적해를 구하였다. 조정의 범위를 제한하였으므로 개선비율은 43.66%로 낮아졌고 점수조정 폭이 축소되었지만, 현실상황을 고려하여 보수적인 방향으로 점수조정을 할 때 적용 가능한 해를 도출하였다고 볼 수 있다.

4.2.3 종합등급 2.00~3.99 집단 조정결과

전술한 바와 같이 특정대학의 입학전형에서 경쟁자는 수학능력이 일정한 범위에 있다. 따라서 언어, 수리, 외국어 영역의 종합평균 등급이 2.00~3.99등급인 집단의 사람 11개 선택과목 점수를 측정이론의 연쇄방법과 선형계획모형으로 조정하여 <표 5>에 참고로 제시하였다. <표 4>에서 논의한 사항은 <표 5>에 적용된다. 2.00~3.99등급자의 경우에도 조정의 방향은 전체응시자의 경우인 <표 4>와 비슷하고 조정의 크기가 다르다. 따라서 대학이 탐구영역 선택과목 점수조정을 하려면 해당 대학 모집단위에 응시하는 집단을 대상으로 가산점의 크기를 정할 수 있다.

〈표 5〉 종합(언수외)등급 2.00~3.99 집단의 표준점수 조정 결과

적용방법 및 기준		측정이론을 적용한 조정 (공통과목을 가교검사로 활용하여 연쇄방법 적용)			선형계획모형을 적용한 조정 (과목 쌍 간 능력차이 반영) (조정결과 : 가산점)			
		해당과목 평균점수	공통과목 평균점수	조정결과 (가산점)	공통과목과 선택과목의 관계만 고려	순수 선택과목 [모형 1]	현실제약	
선택과목	윤리	58.66	60.17	0.62	0.62	0.55	0.63	
	국사	54.05	60.91	5.97	5.97	5.16	6.29	
	한국지리	57.39	60.16	1.88	1.88	1.82	1.73	
	세계지리	57.72	60.11	1.50	1.50	1.49	1.44	
	경제지리	55.83	60.19	3.47	3.47	3.38	3.13	
	근현대사	57.38	60.23	1.96	1.96	1.83	1.91	
	세계사	58.40	59.96	0.67	0.67	0.81	0.73	
	법과사회	59.21	60.10	0.00	0.00	0.00	0.00	
	정치	57.92	60.17	1.36	1.36	1.29	1.21	
	경제	54.65	60.56	5.02	5.02	4.56	4.61	
	사회문화	57.83	60.17	1.45	1.45	1.39	1.28	
	왜곡점수 합계	조정 전(백점)			6,912	6,912	9,486	
		조정 후(백점)			0	0	1,591	2,393
개선비율(%)				100	100	83.22	74.77	

4.2.4 측정이론의 연쇄방법과 선형계획모형의 조정방법 비교

각 선택과목과 공통과목의 관계만 포함하여 선형계획모형으로 점수를 조정하면 공통과목을 가교검사로 활용하여 측정이론의 연쇄방법으로 조정한 결과와 같다. 또한 선형계획모형은 공통과목이 없이 순수 선택과목 쌍 간의 공통응시집단을 활용하여 조정할 수 있다. 순수 선택과목으로 조정한 가산점과 공통과목을 활용하여 연쇄방법으로 조정한 가산점을 직접 비교하는 것은 기준이 다르기 때문에 어렵다. 다만, 사회탐구능력이 높으면 언수외의 수학능력도 높아서 가산점의 방향이 대체적으로 같게 나타난다. 여기서 논의하고자 하는 것은 측정이론의 연쇄방법을 적용하는 것과 선형계획모형을 적용하는 것의 특성을 비교하고자 한다.

언수외의 공통과목을 활용하여 측정이론의 연쇄방법으로 조정한 것은 두 선택과목의 공통응시집단을 활용하여 직접 조정한 결과와 차이가 난다. <표

3>에서 국사와 세계사 두 과목을 응시한 공통응시 집단의 국사 평균점수는 50.64점이고, 세계사 평균 점수는 53.48점이다. 직접 조정하면 공통응시집단의 평균점수가 같아야 하므로 국사에 $[53.48 - 50.64] = 2.84$ 점을 가산하면 된다. <표 4>에서 공통과목을 활용하여 연쇄방법으로 조정한 국사의 가산점 4.01점은 공통응시자를 기준으로 조정한 2.84점보다 과도하게 조정된 것이다. 공통응시자를 기준으로 직접 조정하는 것은 단일집단 설계에 해당하기 때문에 공통과목을 선정하여 연쇄방법으로 조정하는 것보다 측정이론 측면에서 타당하다. 다만, 모든 과목을 응시한 공통응시자가 없으므로 기존의 측정이론을 적용하기 위해서는 공통과목을 활용할 수밖에 없다. 본 연구에서 제시하는 선형계획모형을 적용하는 방법은 과목 각 쌍 간의 공통응시자를 기준으로 직접 조정하는 것이다.

이외에도 공통과목을 활용하는 측정이론의 연쇄방법은 두 가지 한계가 있다. 첫째, 공통과목의

점수는 과학탐구영역이나 직업탐구영역 응시자의 점수가 혼재된 표준화된 점수이므로 사탐 내의 각 선택과목 간 점수를 조정하는데 타 영역의 점수들이 영향을 미쳐서 가교검사로 적합하다고 보기 어렵다. 둘째, 각 선택과목의 내용이 “수학능력”보다는 “사회탐구능력”을 측정하는 것으로 볼 수 있기 때문에 탐구능력을 기준으로 점수를 조정하는 것이 보다 합리적이다.

결국 선형계획모형을 적용하는 방법은 선택과목 각 쌍의 공통응시자를 기준으로 직접 점수를 조정하는 것이므로 측정이론 측면에서 보다 타당하며, 사회탐구능력을 기준으로 조정하는 것이다. 또한 선형계획모형은 공통과목을 활용하여 측정이론의 연쇄방법으로 조정한 결과와 같게 조정할 수 있으므로 유용하다. 다만, 선형계획모형을 활용하려면 모형을 수립하고 입력 자료를 준비해야 한다.

5. 결 론

현행 수능은 선택과목별로 표준점수를 산출하여 대학 및 수험생에게 제공하고 있다. 표준점수는 과목별 시험의 난이도 차이를 반영하여 과목 간 점수를 비교할 수 있도록 한 것이다. 그러나 과목별 응시집단의 능력 차이가 반영된 것이 아니기 때문에 수학능력 또는 탐구능력을 올바르게 나타낸다고 할 수 없다. 따라서 서로 다른 과목의 표준점수를 직접 비교하는 방식인 표준점수의 총점을 기준으로 학생을 선발하면 공정하지 않을 수 있다. 그럼에도 불구하고 2010학년도 입학전형에서 대학의 90% 이상이 서로 다른 선택과목의 표준점수를 합한 총점을 비교하여 학생을 선발하고 있으며, 가산점을 부여하는 대학은 매우 적은 것으로 보고하고 있다. 또한 가산점을 정하는 합리적이고 객관적인 이론이 없는 실정이다. 따라서 능력이 정확하게 반영되도록 응시집단 간 능력 차이를 반영하여 표준점수를 조정하는 방법을 개발할 필요가 있다.

이 연구는 수능 선택과목사이의 점수조정 문제에 대하여 경영과학 접근방법의 유용성을 보인 것

이다. 교육평가분야의 측정이론은 검사 간의 평균, 표준편차, 백분위 등을 완전하게 같도록 조정하는 방향으로 연구되었다. 그러나 일반적으로 가교검사나 가교집단이 없는 세 검사 이상의 경우에 완전하게 조정하는 것은 불가능하다. 따라서 이러한 문제에 대하여 측정이론은 더 이상 다루지 않았다. 이 연구는 완전하게 조정할 수 없는 문제에 대해서는 최선으로 조정하려는 경영과학의 접근방법을 적용하여, 선택과목 쌍 간의 공통응시집단을 기준으로 조정되지 못한 왜곡점수를 최소화하는 선형계획모형을 개발하였다. 모형의 적용가능성을 검토하기 위하여 2010학년도 수능 사탐 전체 11개 선택과목에 응시한 377,089명의 점수를 분석하고 최적해를 구하여 기존의 측정이론과 비교하였다. 분석결과 제시된 선형계획모형은 각 대학이 입학전형에서 선택과목의 가산점을 정하기 위하여 활용하는데 유용한 것으로 나타난다.

이전의 연구에서도 점수조정을 위한 선형계획모형이 제시되었지만 현실적인 적용에 어려움이 있어서 본 연구에서 몇 가지 약점을 보완하였다. 첫째, 원점수를 사용하는 대신, 수능의 실제 자료로 제공되는 표준점수를 활용하였다. 둘째, 조정목표와 조정의 차이를 최소한으로 줄이도록 목표계획법의 개념을 도입하였다. 셋째, 선형계획모형의 목적함수는 왜곡된 점수와 응시자 수의 가중합계가 되도록 하였다. 넷째, 선형계획모형을 사탐의 전체 11개 선택과목 점수조정 문제에 적용하여 현실적으로 대입전형에서 활용될 수 있음을 보였다. 따라서 수능 점수를 활용하여 대학입학에 대한 결정을 할 때, 제시된 모형을 활용하면 능력이 우수한 학생의 선발이 가능하고 입학전형에 공정성을 높일 수 있다고 본다.

이론적으로 볼 때 본 연구에서 제시한 모형은 다수 검사의 연계화에 대한 일반화이다. 그러므로 두 검사에 대하여 측정이론과 일치하며, 가교검사나 가교집단을 갖지 않는 셋 이상의 순수 선택검사의 검사점수 조정에도 적용이 가능하다. 따라서 제시된 모형은 대학이나 고등학교의 선택과목 점수

조정 및 다양한 측정 또는 평가결과의 조정에 적용될 수 있다고 본다.

이 연구는 검사 쌍 간의 공통응시자를 기준으로 조정방법을 도출하였지만, 다수 과목의 공통응시자를 기준으로 점수를 조정하는 방법 개발이 추후 연구과제가 된다. 또한 평균방법 이외에도 선형방법과 동백분위방법을 기반으로 하는 선형계획모형의 개발 및 적용가능성에 대한 연구도 필요하다.

참고 문헌

- [1] 김성훈, 타당화 개념모형을 이용한 대학수학능력시험 체제 진단, 『교육평가연구』, 제23권, 제1호(2010), pp.1-27.
- [2] 김현철, “모의실험에 의한 대학수학능력시험 선택과목 점수 동등화 방법들의 적합성 비교 (1) : 원점수와 동등화점수의 차이분석”, 『교육학연구』, 제41권, 제1호(2003), pp.363-384.
- [3] 남명호, 김신영, 남현우, 권순달, 박정, 조지민, 『2005학년도 대학수학능력시험 점수 체제 연구』, 한국교육과정평가원, 연구보고서 대수능 CAT 2002-32, 2002.
- [4] 남보우, “대입수능 동등화 표준점수 산출 보정 : 2005 사례연구”, 『한국경영과학회 2002 추계 학술대회논문집』, (2002), pp.85-89.
- [5] 남보우, 『수능 선택과목의 표준점수 심층분석 및 대입전형 활용방안』, 한국교육과정평가원, RRO, 2010.
- [6] 남보우, 수능시험 선택과목 점수조정에 대한 측정학 및 경영과학 비교, 『한국경영과학회 2010 추계 학술대회 발표논문집』, (2010), pp.164-168.
- [7] 남현우, “2005 수능의 탐구영역 내 과목 간 점수 조정에서 합집단 연계화방법의 적용가능성”, 『교육평가연구』, 제16권, 제1호(2003), pp.163-182.
- [8] 남현우, “대학수학능력시험 영역별선택과목의 연계화 가능성탐색”, 『교육평가연구』, 제17권, 제2호(2004), pp.93-105.
- [9] 박 정, “문항반응이론에 의한 2005 대학수학능력시험 영역별 선택과목 점수조정의 타당성 검증”, 『교육학연구』, 제43권, 제2호(2005), pp. 135-154.
- [10] 부재율, “검사동등화 시의 고려사항”, 『교육평가연구』, 제12권, 제1호(1999), pp.43-64.
- [11] 성태제, “다른 교과와 점수를 연결하는 통계적 조절 방법의 타당성”, 『교육평가연구』, 제13권, 제1호(2000), pp.63-99.
- [12] 윤영선, 『상관분석』, 교육과학사, 2000.
- [13] 정기호, 백천현, 『엑셀을 활용한 경영과학』, 4판, 비엔엠북스, 2009.
- [14] 최성열, 이규민, 박아청, 검사점수 연계화 방법의 상대적 적절성 비교, 『교육심리연구』, 제21권, 제1호(2007) pp.1-25.
- [15] 황형태, 대학수학능력시험에서 표준점수제의 개선방안에 대한 연구, 『응용통계연구』, 제18권, 제3호(2005), 521-532.
- [16] Kolen, M.J. and R.L Brennan, *Test Equating, Scaling, and Linking*, 2nd Ed., Springer, 2004.