

대학수학능력시험에서 표준점수제의 개선방안에 대한 연구*

황형태¹⁾

요약

2005학년도 대학수학능력시험부터는 수험생들마다 응시과목들이 서로 다르게 구성되어 있으므로, 과목별 난이도의 차이를 보정하기 위하여 표준점수제를 도입하여 운영하고 있다. 그러나 2005학년도 대학수학능력시험에서 본격적으로 표준점수제를 도입하여 운영해 본 결과, 과목선택의 조합에 따라서 과목별 표준점수의 차이가 지나치게 크게 나타나는 등 여러가지 문제점들이 노출되었으며, 이에 따라 표준점수제에 대한 전면적인 개선이 강력하게 요구되기도 하였다. 이 논문에서는 현행 표준점수제의 문제점들을 좀더 심도있게 분석해 보고, 그 개선방안으로 하나의 대안을 제시해 보고자 한다. 이 논문에서 개발한 가산점수제(Additive scoring system)에 의한 결정점수는 현행의 표준점수제가 갖고 있는 여러가지 문제점들을 원천적으로 해소하거나 크게 개선하며, 직관적으로 이해하기 쉬울 뿐만 아니라 계산이 간편하다는 등의 장점들도 갖추고 있다.

주요용어: 대학수학능력시험, 표준점수제, 가산점수제

1. 서론

우리나라의 대학입시제도는 그동안 지속적인 변화를 거쳐 오는 과정에서 대학입시에서의 선택과목 제도가 제7차 교육과정 이전에도 여러 차례에 걸쳐 제한적으로 실시된 바 있었다. 선택과목제가 실시되면 과목별 난이도를 균등하게 유지하는데 한계가 있기 때문에 난이도의 차이를 보정할 필요성이 생기게 되는데, 이에 따라 선택과목의 표준점수제 도입 및 운영에 대한 다양한 연구가 이루어져 왔다.

성태제(1994)와 허명희(1994)는 대학별 본고사 실시에 따른 문항분석, 표준점수제, 김사동등화 등에 관한 교육측정의 이론들을 소개하고 그 성과에 대하여 논의한 바 있으며, 김신영(1999)은 제6차 교육과정기간 동안 대학수학능력시험에서 도입했던 변환표준점수에 대하여 논의하고 변환표준점수의 영역별 영향력에 대하여 연구하였다. 한편, 선택과목 집단별로 학업능력의 차이가 존재하는 경우에 이를 반영하는 방법으로, 한국교육과정평가원(1998)은 공통과목을 이용하는 선형등화방법을 연구하였으며, 박성현, 김춘원, 박준오(2000)는 공통과목을 이용한 비모수적 등화방법을 소개하였다.

한편 제7차 교육과정에 의하여 교육받은 학생들이 응시한 2005학년도 대학수학능력시험부터는 수험생들에게 부여된 과목선택의 폭이 그 이전보다 대폭 확대되었다. 이에 따라 특히 탐구영역에서는 동일영역 내에서도 수험생들의 응시과목들이 모두 상이하게 구성되

* 이 연구는 2004학년도 단국대학교 교내연구비의 지원으로 연구되었음.

1) (140-714) 서울시 용산구 한남동 산 8번지, 단국대학교 정보컴퓨터학부, 교수

E-mail: hthwang@dankook.ac.kr

므로, 선택과목의 확대에 따르는 부작용을 완화하고 보다 합리적인 점수척도로의 전환을 위한 점수보정방법의 중요성이 한층 부각되었다.

한국교육과정평가원(2004a)은 과목별 난이도의 차이를 보정하기 위하여 2005학년도 대학수학능력시험에서 언어, 수리, 외국어(영어) 영역의 경우 각 과목별로 평균 100점, 표준편차 50점이 되도록, 탐구영역과 제2외국어/한문영역의 경우에는 각 과목별로 평균 50, 표준편차 10점이 되도록 원점수를 선형전환(linear transformation)하여 각 과목별 표준점수를 산출하였다. 김신영, 구창현, 임형, 박정(1998)은 이와 같은 선형전환 방법에 대하여 일반인들이 선형전환에 의한 표준점수의 산출과정을 이해하기 쉽고, 원점수 분포 양상을 그대로 유지하여 일관성 있게 원점수가 전환되며, 총점에 있어서의 백분위 변화도 적게 나타난다고 평가한 바 있다. 그러나 막상 2005학년도 대학수학능력시험에서 표준점수제를 전면적으로 도입하여 운영해 본 결과, 선형변환에 의한 현행 표준점수제의 여러 가지 심각한 문제점들이 드러나게 되었다.

이 연구에서는 2005학년도 대학수학능력시험의 결과를 바탕으로 현행 표준점수제가 갖고 있는 문제점들을 심층적으로 분석하고, 이에 대한 대안으로 새로운 보정방법을 개발하고자 한다. 여기에서 개발하는 가산점수제는 과목별 난이도 차이와 수험생 집단 간의 평균 학력 차이를 동시에 고려하여 과목별 가산점을 결정하는 방법이다. 이 방법은 현행 표준점수제가 갖고 있는 문제점들을 원천적으로 해소하거나 크게 완화시키며, 원리적으로 매우 간단하여 일반인들이 직관적으로 받아들이기에 용이하고, 계산이 간편하다는 특징도 갖고 있다.

다음의 2절에서는 2005학년도 대학수학능력시험에서 드러난 표준점수제의 문제점들에 대하여 심층적으로 고찰하고, 3절에서는 새로운 개선방안으로 가산점수제를 개발하여 2005학년도 대학수학능력시험의 결과에 적용해 보며, 종합적인 결론은 4절에서 논의한다.

2. 현행 표준점수제의 타당성 및 문제점 분석

이 논문에서는 인문계고등학교 학생들이 주로 응시하는 평가영역 중에서 선택과목만으로 이루어진 사회탐구영역, 과학탐구영역, 제2외국어/한문영역의 표준점수제에 대하여 중점을 두고 고찰해보도록 한다.

사회탐구영역과 과학탐구영역은 각각 11과목과 8과목으로 구성되어 있으며, 수험생들은 이 2개의 평가영역 중 하나의 영역을 택한 후, 택한 영역에서 4개 이내의 과목을 선택하여 응시할 수 있고, 각 과목의 원점수 배점은 50점이다. 또한 제2외국어/한문영역은 8개의 과목으로 구성되어 있는데, 수험생들은 이 중 한 과목을 선택하여 응시할 수 있으며, 각 과목의 원점수 배점은 역시 50점으로 주어진다.

이 과목들에 대하여 학생들이 취득한 원점수는 성적표에 기재되지 않으며, 각 과목별로 평균 50점, 표준편차 10점이 되도록 선형변환하여 산출한 표준점수와, 표준점수를 바탕으로 산출한 과목별 백분위, 백분위를 바탕으로 산출한 과목별 등급 등이 성적표에 기재되도록 되어 있다.

2.1. 표준점수제의 타당성 검토

여기에서는 현행의 표준점수제가 도입되어 원활하게 기능하기 위한 두 가지 전제조건들을 제시하여 검토해 보고자 한다.

첫 번째 전제조건으로, 허명회(1994)에 따르면 선형변환을 이용한 표준점수제는 기본적으로 정규성의 가정이 깔린 모수적 방법이다. 즉, 서로 다른 과목에서의 성적들을 표준점수화하여 그 값의 크기로 우열을 판단하는 경우에는 각 과목에서의 성적분포가 정규분포를 따른다는 전제조건 하에서 유효하다는 것이다. 각 과목에서의 성적분포가 정규분포를 따른다면 이들을 표준화했을 때의 분포가 모두 표준정규분포로 동일하므로, 과목이 다르더라도 동일한 표준점수는 동일한 백분위를 의미한다는 점에서 표준점수제가 통계적 의미의 근거를 갖게 되는 것이다. 따라서 성적의 분포가 정규분포가 아닌 경우에 표준점수제의 통계적 근거는 매우 제한적이라고 볼 수 있을 것이다. 그런데 지난 2005학년도 대학수학능력시험의 결과에서 나타나듯이 탐구영역의 많은 과목에서 만점자의 비율이 10%를 상회하는 등, 특히 고득점자들의 성적이 정규분포와 동떨어지게 분포하고 있는 것으로 판단된다.

두 번째 전제조건으로는, 현행 제도에서는 모든 탐구영역 과목의 평균을 50점, 표준편차를 10점으로 동일하게 고정시키는 방법으로 표준점수를 산출하였는데, 이는 각 과목을 응시한 수험생 집단의 평균학력이 균등하다는 전제 하에서만 유효한 것이다. 즉, 과목 간 평균점수의 차이는 시험문제의 난이도 차이와 과목별 수험생 집단 간 평균학력 차이의 두 가지 요인에 의하여 나타나는 것으로 볼 수 있는데, 후자의 요인을 완전히 무시하고 과목별 표준점수의 평균을 똑같이 50점이 되도록 선형변환 함으로써, 과목별 난이도 차이에 대한 과보정(over-correction)의 위험성이 크게 나타나게 되는 것이다.

허명회(1994)는 수차례 모의고사 자료분석을 통하여 선택과목별로 수험생들 간에 체계적인 학력차이가 존재하고 있음을 지적한 바 있다. 실제로 과목별 수험생 집단의 평균학력이 균등하다는 전제는 매우 위험하며 현실적으로도 적절하지 않다. 학력수준이 높은 수험생 집단의 과목선택의 패턴과 상대적으로 학력수준이 낮은 수험생 집단의 과목선택의 패턴이 같다고 볼 근거가 없으며, 재수생과 재학생, 또는 출신고교별로 과목선택의 패턴이 다르고, 서울대학교의 인문계열에서 국사를 필수과목으로 지정한 바와 같이 특정대학에서 특정과목을 요구하는 경우도 있어서, 각 과목별 수험생 집단의 평균학력이 모두 동일하다는 전제는 매우 무리한 것으로 볼 수 있다. 이런 관점에서 현행의 표준점수제는 상당히 큰 과보정의 위험성을 본질적으로 내포하고 있으며, 결과적으로 과목선택에 따라 표준점수의 극단적인 차이를 초래할 우려가 매우 큰 것으로 판단된다.

2.2. 2005학년도 대학수학능력시험에서 나타난 표준점수제의 실제 문제점 분석

앞에서 살펴본 대로 표준점수제가 원활하게 기능하기 위한 두 가지 전제조건들이 충족되지 못하는 상황에서 2005학년도 대학수학능력시험이 본격적인 표준점수체제로 실시되었다. 그 결과 과연 어떤 문제점들이 초래되었는지 구체적인 항목별로 살펴보기로 한다.

(1) 앞서 예상했던 대로 과보정의 문제점이 매우 심각했던 것으로 밝혀졌다. 원점수 만점인 경우에, 사회탐구영역에서는 윤리과목이나 한국지리과목의 표준점수가 61점에 그쳤었던데 반해 사회문화과목에서는 68점으로 한 과목에 최대 7점의 차이를 보였으며, 과학

탐구영역에서는 최대 6점의 차이를 보였고, 제2외국어/한문영역에서는 한 과목의 선택에서 무려 36점의 차이가 벌어지기도 했던 것이다. 예를 들어, 사회탐구영역에서 4과목을 선택하여 모두 원점수 만점을 받았다고 하더라도, 과목선택의 조합에 따라 최대 20점의 차이까지 나타나게 되므로, 최상위권 대학에는 지원기회조차 상실할 수 있을 정도로 치명적인 것으로 나타났다. 실제의 수능과 동일한 방식으로 한국교육과정평가원이 전국적으로 실시했었던 2004년 9월의 모의고사에서는, 원점수 만점인 경우에 정치과목의 표준점수는 74점, 세계사과목의 표준점수는 61점으로 한 과목에서만 최대 13점의 차이까지 벌어진 바 있었다. 이런 이유 때문에 대학과 수험생들로 하여금 표준점수의 안정성을 신뢰하지 못하게 하는 결과를 초래하게 되었으며, 많은 대학에서 표준점수를 직접적인 전형자료로 사용하지 않는 현상까지 나타나고 있는 것이다.

(2) 교육적인 고려라는 이유로 표준점수는 반올림하여 정수화된 점수를 부여하고 있는데(한국교육과정평가원, 2004a), 이러한 반올림 제도가 표준점수제의 특성과 결부되어 또 다른 왜곡현상을 초래했던 것으로 나타났다. 동일과목에서 원점수의 일정한 차이가 점수대에 따라서 표준점수에서는 상이한 차이로 나타나는 일이 전반적으로 발생한 것이다. 원점수에서 1점의 차이가 표준점수에서는 경우에 따라 0점, 1점, 2점의 차이로 나타나게 되었는데, 예를 들어, 어떤 과목에서 원점수 45점, 46점, 47점을 취득한 3명의 학생들이 표준점수에서는 각각 58점, 58점, 59점을 받는다는가, 반대로 57점, 58점, 60점을 받는 경우들이 발생하였다. 특히 전자의 경우에는, 원점수 46점을 받은 학생들이 원점수 45점을 받은 학생들과 같은 표준점수를 받게 되므로, 표준점수를 바탕으로 산출한 백분위에서도 큰 불이익을 보게 되어 이 문제가 법정으로 비화된 바도 있다. 성실한 노력에 의하여 취득한 점수가 운에 의하여 희석된다면, 이것이야말로 반드시 시정해야할 비교육적인 현상이 아닐 수 없을 것이다.

(3) 현행 표준점수제에서는 이론적으로 점수의 상한과 하한이 존재하지 않는다. 음의 점수도 가능하고, 탐구영역과목의 경우 100점이나 200점을 넘는 점수도 가능하다. 물론 그런 경우에 강제적으로 0점이나 100점의 점수를 부여하고는 있지만 자연스러운 방법은 아니다.

(4) 이번에 지적하려고 네 번째 문제점은 표준점수제의 과목별 난이도 보정이라는 본질적 기능에 관한 것이다. 앞에서 논의했던 바와 같이 표준화된 점수가 비교척도로서 기능하기 위한 중요전제 중의 하나는 원점수들의 분포가 정규분포이거나 적어도 정규분포에 가까워야 한다는 것이다. 그러나 그 전제조건이 충분히 만족되지 못하는 상황에서 원점수를 표준편차의 단위로 전환하는 표준화 작업이 갖는 통계적 의미는 제한적일 수밖에 없는데, 표준화의 과정에 의하여 어떤 문제점이 초래되고 있는지 살펴보려고 한다. 이를 위해서는 표준점수와 원점수 사이의 연계성을 살펴보는 일이 필요한데, 원점수나 원점수에 관계된 일체의 통계자료는 공개되지 않고 있는 실정이다. 그러나 다행스럽게도 한국교육과정평가원(1994b)이 자체홈페이지를 통해서 공개했던 <영역/선택과목별 표준점수 도수분포표>를 이용하여 각 과목별로 원점수의 평균 및 표준편차들을 산출할 수 있었다. 먼저 그 방법을 간단하게 소개한다.

공개된 <영역/선택과목별 표준점수 도수분포표>에서 각 과목별로 만점과 0점의 표준

점수를 알 수 있었는데, 원점수 만점이 50점인 어떤 특정과목에 대하여 원점수 50점에 해당하는 표준점수와 원점수 0점에 해당하는 표준점수를 각각 S_1 과 S_2 라고 하자. 또한 그 과목에서 원점수 평균과 표준편차를 각각 μ 와 σ 라고 하자. 표준점수는 평균이 50점, 표준편차가 10점이 되도록 선형전환한 값이므로, μ 와 σ 에 대한 다음의 연립방정식을 얻을 수 있다.

$$\frac{50 - \mu}{\sigma} \times 10 + 50 = S_1, \quad \frac{0 - \mu}{\sigma} \times 10 + 50 = S_2 \quad (2.1)$$

식(2.1)로부터 μ 와 σ 를 구하면 다음과 같다.

$$\mu = \frac{50 - S_2}{S_1 - S_2} \times 50, \quad \sigma = \frac{500}{S_1 - S_2} \quad (2.2)$$

다음의 표는 2005학년도 대학수학능력시험의 사회탐구영역, 과학탐구영역, 제2외국어/한문영역의 과목들에 대하여, 식(2.2)로부터 구한 각 과목들의 원점수 평균들과 각 과목 만점의 표준점수를 도표화한 것이다(한국교육과정평가원에서 공개한 표준점수 자체가 반올림하여 정수화한 값들이기 때문에, 이 표에서 나타난 과목별 원점수의 평균들은 모두 근사값들이며, 그 오차는 최대 ± 0.7 점 이내에 있는 것으로 계산되었다).

표 2.1: 2005 수능 과목별 원점수 평균과 만점의 표준점수

(i) 사회탐구영역

과목명	한지	윤리	세지	정치	국사	근현	법사	경지	경제	사문	세계
원점평균	39.8	38.5	36.0	35.9	35.0	33.7	33.7	33.7	33.3	33.0	30.6
표점만점	61	61	62	63	62	65	66	67	65	68	64

(ii) 과학탐구영역

과목명	지구과학I	생물 I	화학 I	생물II	물리 I	물리II	지구과학II	화학II
원점평균	36.2	35.1	34.7	32.2	31.7	31.6	28.8	27.4
표점만점	63	64	65	66	65	64	67	69

(iii) 제2외국어/한문영역

과목명	일어	프랑스어	독일어	스페인어	중국어	러시아어	한문	아랍어
원점평균	31.1	30.7	30.0	29.5	29.3	28.3	27.6	17.1
표점만점	64	67	66	66	67	63	72	100

이 표에서 동일영역내의 과목순서는 원점수의 평균이 높은 순으로 정렬한 것이다. 일반적으로 원점수의 평균이 낮을수록 그 과목의 평균난이도는 높은 것으로 평가할 수 있기 때문에, 오른쪽으로 갈수록 난이도가 높은 과목으로 간주할 수 있을 것이다.

이제 표2.1의 사회탐구영역의 과목들에 대해서 고득점자의 경우인 원점수 만점의 표준점수들을 살펴보도록 하자. 대체로 오른쪽으로 갈수록, 즉, 난이도가 높을수록 원점수 만점의 표준점수도 높아지는 경향이 있음을 알 수 있다. 그러나 세계사 과목의 경우 원점수 평균이 가장 낮아서 난이도가 가장 높은 과목으로 인정할 수 있음에도 불구하고, 만점자에 대하여 불과 64점이라는 낮은 표준점수가 부여됐음을 알 수 있다. 그보다 난이도가 낮은 근현대사, 법과사회, 경제지리, 경제, 사회문화 과목들에 비하여 표준점수가 최소 1점에서 최대 4점까지 불리한 점수인 것이다. 이는 표준점수제가 보여주는 또 다른 왜곡현상으로 지적할 수 있다. 이런 현상은 표에서 알 수 있는 바와 같이 과학탐구영역의 물리II 과목이나 제2외국어/한문영역에서의 러시아어 과목에서도 찾아볼 수 있어서, 표준점수제의 일반적인 왜곡현상을 보여주는 예라고 볼 수 있을 것이다. 이와는 별도의 문제지만, 아랍어 과목에서는 무려 100점의 표준점수까지 나타나고 있다는 사실도 표준점수의 불안정성을 보여주는 또 다른 예가 된다.

3. 가산점수제 개발 및 적용 예

표준점수제를 도입하는 중요한 두 가지 목적은 첫째, 과목별 난이도의 보정이고, 둘째, 과목별 평균과 표준편차를 일정하게 함으로써 기대할 수 있는 과목별 영향력의 균등화라고 설명할 수 있다. 그러나 이 가운데 표준점수제를 도입하는 주된 목적은 과목별 난이도의 보정이며, 과목별 영향력의 균등화는 부차적인 기능이라고 볼 수 있다. 그 이유로, 만일 선택과목 없이 공통과목만 있는 경우라면 굳이 표준점수제를 도입하지 않고, 과목별 영향력이 다르더라도 원점수를 그대로 사용하는 경우가 일반적이라는 사실에서 찾아볼 수 있다. 그러나 과목별 영향력의 균등화라는 부차적 기능까지 얻기 위하여 표준편차까지 감안한 표준점수제를 도입했을 때는, 앞 절에서 살펴본 바와 같은 여러 가지 심각한 문제점들이 필연적으로 나타나게 된다.

일반적으로 과목별 난이도를 직접적이고도 단조적으로 나타낼 수 있는 주된 지표는 과목별 평균점수라고 간주할 수 있을 것이다. 표준편차의 값은 난이도가 아주 높을 때나 아주 낮은 상반된 경우에 공통적으로 작아지는 경향이 있기 때문에, 난이도를 단조적으로 나타낼 수 있는 지표는 아니다. 따라서 여기에서는 단순하게 과목별 평균점수들만을 이용하여 과목별 평균난이도의 차이를 보정하는 새로운 보정방법을 개발하고자 한다. 앞으로 이 방법을 가산점수제(Additive scoring system)라고 하자.

가산점수제는 과목별 평균점수를 이용하여 과목별 평균난이도를 산정함으로써 과목별 가산점을 결정하고, 원점수에 과목별 가산점을 합산하여 최종적으로 과목별 결정점수를 구하는 방법이다. 이 방법은 과목별 영향력의 균등화라는 표준점수제의 부차적인 기능 대신에, 표준점수제가 전혀 반영하지 못한 과목별 수험생집단 간의 평균적인 학력 차이를 반영하는 기능을 갖게 된다.

3.1. 가산점수제에 의한 결정점수의 산출방법

원점수 배점이 50점으로 주어지는 탐구영역, 또는 제2외국어/한문 영역의 어떤 특정과

목을 생각해 보자. 이 과목에 응시한 학생들의 원점수평균을 m 이라고 하고, 과목 배점 50에서 m 을 뺀 값 C 를 이 과목의 여평균(complement of mean score)이라고 하자. 즉,

$$C = 50 - m. \tag{3.1}$$

여평균 C 는 그 과목의 난이도가 높을수록, 또는 그 과목을 응시한 학생들의 평균학력이 낮을수록 그 값이 커지게 된다. 즉, C 의 값을 결정하는 요인은 크게 나누어, (i) 과목의 난이도와, (ii) 수험생 집단의 평균학력의 두 가지로 요약될 수 있다.

이제 어떤 상수 w (단, $0 \leq w \leq 1$)에 대하여, 여평균에서 과목의 난이도에 따르는 기여분을 wC , 수험생 집단의 평균학력에 따르는 기여분을 $(1-w)C$ 라고 하자. 여기에서 w 는 보정상수, wC 는 과목별 가산점이라고 부르도록 한다. 결론적으로, 가산점수제에서 어떤 과목의 결정점수는 과목별로 일정한 가산점 wC 를 수험생의 해당과목 원점수에 합산함으로써 결정된다. 즉,

$$\text{결정점수} = \text{원점수} + \text{과목별 가산점}(wC) \tag{3.2}$$

과목별 가산점 wC 의 값은 그 과목의 난이도가 높기 때문에 응시수험생들의 여평균이 높아지게 된(즉, 과목평균이 떨어지게 된) 점수에 해당하므로, wC 의 값을 각 수험생들의 과목원점수에 가산함으로써 과목별 난이도 차이를 보상하자는 것이다. 반대로, $(1-w)C$ 는 그 과목 응시수험생들의 평균학력이 낮기 때문에 여평균이 높아지게 된(즉, 과목평균이 떨어지게 된) 점수에 해당하므로, 이는 학력이 자연스럽게 점수에 반영된 결과로 보아야 하며 보상해서는 안된다. 따라서 이렇게 산출된 최종적인 결정점수는 과목별 난이도 차이를 보정하여 상쇄하고, 응시수험생들의 평균학력차이는 그대로 반영하는 결과를 얻게 된다.

보정상수 w 의 값이 커질수록 과목별 난이도에 대한 보정력이 커짐을 의미하며, 다음과 같은 두 가지 성질은 쉽게 알 수 있다.

(1) w 의 값이 0인 경우, 모든 과목의 가산점은 0점이 되므로 결정점수는 원점수와 같게 된다.

(2) w 의 값이 1인 경우, 모든 과목의 가산점은 그 과목의 여평균이 되고 여평균 전체를 원점수에 합산하여 결정점수를 산출하므로, 결과적으로 과목별 응시수험생들의 학력차이는 무시되고 모든 과목의 결정점수들은 표준점수제에서와 같이 과목별 평균이 50점으로 일정하도록 보정된다.

또한 동일 영역내의 두 과목에 대하여 과목별 가산점을 각각 $wC_1 = w(50 - m_1)$, $wC_2 = w(50 - m_2)$ 라고 하면 다음이 성립한다.

$$\text{과목1의 가산점} - \text{과목2의 가산점} = wC_1 - wC_2 = w(m_2 - m_1) \tag{3.3}$$

즉, 과목 간 가산점의 차이는 과목 간 평균의 차이에 보정상수 w 의 값을 곱한 값과 같음을 알 수 있다. 그러므로, 예를 들어 $w = 0.3$ 인 경우의 가산점수제는 '과목 간 평균 차이의 30%를 과목 간 난이도의 차이로 인정하여 보정한 결과'라는 식의 해석이 가능하다.

3.2. 보정상수 w 값의 결정

허명희(1994)는 선택과목들에 대하여 표준점수제를 통하여 사후에 난이도를 보정해야

할 필요성 중의 하나로, 출제교수들이 경쟁적으로 자신의 과목을 쉽게 출제하려는 요인을 원천적으로 없애야 한다는 점을 지적한 바 있다. 물론 수험생들이 쉬운 과목으로만 쏠리려는 비교육적인 현상을 염두에 둔 지적이었을 것이다. 그러나 현행의 표준점수제는 2절에서 살펴본 바와 같이 과도한 난이도 보정의 위험을 안고 있으며, 이는 또 다른 왜곡을 초래할 수 있다. 그러므로 적절한 보정력의 선택은 정책적으로 매우 중요한 것이다.

사후의 난이도 보정은 '필요한 범위 내에서 최소화'하는 것이 바람직한 것으로 판단된다. 여기서 '필요한 범위'란 적어도 쉬운 과목으로의 쏠림 현상을 방지할 수 있는 정도는 되어야 할 것이다. 그런데 한국교육평가원에서는 이미 출제 단계에서부터 과목별 난이도를 균등하게 유지하려는데 상당한 노력을 기울여왔기 때문에, 수험생들이 취득한 원점수는 상당 부분을 각자 노력의 결과로 인정하여 유지하고, 필요한 최소한의 보정만을 정책적으로 가하는 것이 교육적인 측면에서도 바람직할 것으로 판단되는 것이다.

앞에서 언급한 바와 같이 가산점수제에서 보정상수 w 의 크기는 난이도 보정력에 직접적으로 연관되어 있으므로, w 값의 결정은 매우 중요하다. 기본적으로 w 의 값은 보수적인 관점에서 정책적으로 선택하는 것이 교육적인 측면에서도 바람직하고 기능적으로도 안전할 것으로 보인다. 다음의 절에서는 가산점수제의 안정적 기능을 보여주기 위해서 지난 2005학년도 대학수학능력시험의 경우에 대하여 $w = 0.3$ 으로 하여 가산점수제를 적용해본 결과를 예시해 볼 것이다. 이 경우는 앞서 살핀 대로 과목 간 평균차이의 30%를 과목 간 난이도 차이로 정책적으로 '인정'하는 것으로 이해할 수 있다. 계산과정이 대단히 간단하기 때문에 또 다른 w 의 값들에 대하여도 같은 방법으로 쉽게 적용해볼 수 있을 것이다.

다음은 이 절의 마지막으로, w 의 값을 대략적으로 어림하는데 이용될 수 있는 한 가지 수리적 방법을 소개하고자 한다.

원점수 배점이 50점인 k 개의 선택과목이 있는 어떤 탐구영역 A의 각 과목 원점수 평균들이 m_1, m_2, \dots, m_k 라고 하자. 또한 타영역의 원점수 100점인 어떤 공통과목 B가 존재해서, 탐구영역 A의 각 과목 응시생들의 공통과목 B에 대한 원점수 평균에 (배점이 50점으로 같아지도록) 50/100을 곱한 값이 각각 l_1, l_2, \dots, l_k 로 주어져 있다고 하자. 그러면 탐구영역 A의 임의의 두 과목 i, j 에 대하여 $(l_i - l_j)/(m_i - m_j)$ 는 두 과목 i, j 사이의 평균차이에서 차지하는 두 과목 응시생들의 학력차이의 기여비율로 볼 수 있을 것이다. 공통과목에서는 난이도의 차이가 없기 때문이다. 따라서, w 를 추정하는데 다음 식이 이용될 수 있다.

$$\hat{w} = 1 - \frac{2}{k(k-1)} \sum_{i < j}^k \frac{l_i - l_j}{m_i - m_j} \quad (3.4)$$

3.3. 2005학년도 대학수학능력시험 적용 예 및 결과 분석

$w = 0.3$ 으로 하여 2005학년도 대학수학능력시험의 결과에 대하여 우리가 개발한 가산점수제를 적용해 보았다. 표2.1에 주어진 과목별 원점수 평균들을 오차가 없는 것으로 가정하고 이용하였으며, 최종점수의 정수화가 한국교육과정평가원의 방침이므로 여기에서도 과목별 가산점은 반올림하여 정수화하였다. 그 결과는 다음의 표3.1과 같다.

표 3.1: 가산점수제의 2005 수능 적용 결과($w = 0.3$ 의 경우)

(i) 사회탐구영역

과목명	한지	윤리	세지	정치	국사	근현	법사	경지	경제	사문	세계
원점평균	39.8	38.5	36.0	35.9	35.0	33.7	33.7	33.7	33.3	33.0	30.6
가산점	3	3	4	4	5	5	5	5	5	5	6
결점만점	53	53	54	54	55	55	55	55	55	55	56
표점만점	61	61	62	63	62	65	66	67	65	68	64

(ii) 과학탐구영역

과목명	지구과학I	생물 I	화학 I	생물II	물리 I	물리II	지구과학II	화학II
원점평균	36.2	35.1	34.7	32.2	31.7	31.6	28.8	27.4
가산점	4	4	5	5	5	6	6	7
결점만점	54	54	55	55	55	56	56	57
표점만점	63	64	65	66	65	64	67	69

(iii) 제2외국어/한문영역

과목명	일어	프랑스어	독일어	스페인어	중국어	러시아어	한문	아랍어
원점평균	31.1	30.7	30.0	29.5	29.3	28.3	27.6	17.1
가산점	6	6	6	6	6	7	7	10
결점만점	56	56	56	56	56	57	57	60
표점만점	64	67	66	66	67	63	72	100

원점평균: 과목별 원점수 평균

결점만점: 가산점수제에 의한 과목별 만점의 결정점수

표점만점: 표준점수제에 의한 과목별 만점의 표준점수

이 표 역시 동일영역 내 과목들의 순서는 원점수 평균이 큰 순서로 정렬한 것이다. 이 표의 결과로부터 살펴볼 수 있는 가산점수제의 특징들을 몇 가지 항목으로 나누어 분석해보기로 하자.

(1) 표준점수제에 비하여 각 영역 내에서 만점의 보정점수 차이가 상당히 작아진 것으로 나타났다. 두 과목 간의 최대 차이가 사회탐구 영역의 경우 7점에서 3점으로, 과학탐구 영역의 경우 6점에서 3점으로, 제2외국어/한문영역의 경우 37점에서 4점으로 크게 좁혀져서, 과보정의 심각한 문제점이 상당히 완화되고 있음을 알 수 있다. 더 작은 w 의 값을 이용하면 과보정의 문제를 더 줄일 수 있을 것이다. 그러나 이 경우에는 과소보정으로 하여 난이도 보정이라는 본질적 기능이 손상되지 않도록 신중하게 검토해야 한다.

(2) 2.2절에서 지적했던 것처럼, 표준점수제에서는 동일과목에서 원점수의 일정한 차이가 점수대에 따라서 표준점수에서는 상이한 차이로 나타나는 일이 전반적으로 발생하게 되

는데, 가산점수제에서는 원점수의 차이가 결정점수에서도 그대로 유지되어 이런 왜곡현상이 생기지 않는다.

(3) 가산점수제에서는 당연히 음의 점수는 나올 수 없으며, 이론적인 만점이 존재한다. $w = 0.3$ 이면 어떤 경우에도 가산점이 $50 \times 0.3 = 15$ 점을 넘을 수 없으므로, 결정점수가 65점을 넘는 경우는 없다.

(4) 2.2절에서 지적했던 표준점수제의 네 번째 문제점도 가산점수제에서는 나타날 수 없음을 알 수 있다. 즉, 난이도가 높은 과목에서 결정점수가 불리해지는 왜곡현상은 나타날 수 없으며, 난이도가 높을수록, 즉, 과목평균이 낮을수록 결정점수에서 보상되는 가산점도 커지게 된다.

(5) 계산이 무척 간편하고 일반인들도 그 내용을 이해하기 쉽다. 즉, '과목별 평균 차이의 30%를 난이도의 차이로 보상하는 방법'이라는 식으로 이해할 수 있을 것이다.

3.4. 타 영역에서의 적용 방법

현행의 대학수학능력시험에서 언어영역과 외국어(영어)영역은 공통과목이므로 난이도를 보정할 필요성은 없지만, 타 영역과의 일관적인 점수체계를 유지하기 위하여 가산점수제를 적용하는 것이 바람직하다. 이 두 영역의 배점은 각각 100점이므로, 식 (3.1)에서 과목의 여평균이 $C = 100 - m$ 으로 주어지는 점만이 다르며, 가산점은 역시 wC 가 되고 결정점수의 산출과정도 동일하다. 또한 수리(나)영역의 경우도 이와 똑같은 방식으로 결정점수를 산출할 수 있다. 그러나 수리(가)영역의 경우는 약간 다르다.

수리(가)영역의 경우에 전체 배점은 역시 100점이지만, 공통부분과 공통부분을 제외한 세 가지 선택과목으로 구성되어 있다. 100점 중 만일 공통부분의 배점이 T 이고, 공통부분의 원점수 전체평균이 m 이라면, 공통부분의 여평균은 $C = T - m$ 이고 이에 따라 공통부분의 가산점은 wC 가 된다. 다음으로, 각 선택과목별 해당 부분의 원점수 평균을 m_i , ($i = 1, 2, 3$)라고 하면, 각 선택과목의 가산점은 $w(100 - T - m_i)$ 로 주어지게 된다. 공통부분의 가산점과 해당 선택과목의 가산점을 합하여 해당 과목을 선택한 수험들의 수리(가)영역 최종 가산점을 산출하게 되는데, 이 경우 반올림은 두 가지 가산점을 합산한 이후에 할 것을 권장한다. 반올림으로 인한 오차를 조금이라도 줄여줄 수 있기 때문이다.

4. 결론

가산보정법은 난이도 보정 방법으로서 표준점수제가 안고 있는 여러 가지 심각한 문제점, 또는 부작용들을 원천적으로 해소하거나 상당히 완화할 수 있는 대안으로 개발되었다. 표준점수제와 가산점수제는 모두 난이도 보정이라는 공통의 기본적 기능을 갖추고 있지만, 표준점수제가 과목별 평가 영향력의 균등화라는 기능을 추구하는 반면에, 가산점수제는 과목별 수험생집단 간의 평균적인 학력차이를 반영하는 기능을 갖게 된다.

가산점수제에서 보정상수 w 는 과목별 난이도 차이의 보정력을 결정하는 중요한 상수이다. 이 논문에서는 가산점수제가 안정적으로 기능하기 위하여 보수적인 관점에서 정책적으로 w 의 값을 선택할 것을 권장하였다.

적용 예에서 살펴본 바와 같이, 이 논문에서 개발한 가산보정법이 대학수학능력시험이나 또는 선택과목체제를 갖고 있는 또 다른 시험들에 대하여 난이도 보정 목적으로 사용될 경우, 표준점수제에 비하여 매우 안정적으로 기능하여 수험생들의 혼란을 크게 줄여줄 수 있을 것으로 기대된다.

참고문헌

- 김신영 (1999). 대학수학능력시험의 영역별 변환표준점수 반영비율에 관한 연구, <교육과정평가연구>, 2, 289-299.
- 김신영, 구창현, 임형, 박정 (1998). 대학수학능력시험 결과보고 및 분석방안 연구, 한국교육과정평가원 연구보고서.
- 박성현, 김춘원, 박준오 (2000). 대학입시에서의 선택과목 점수 표준화에 관한 연구, <품질경영학회지>, 28, 124-132.
- 성태제 (1994). 대학별고사를 위한 문항분석, 표준점수, 검사동등화, <한국통계학회 논문집>, 1, 206-214.
- 한국교육과정평가원 (1998). 1999년도 수능표준점수제 방안, 미발표 보고서.
- 한국교육과정평가원 (2004a). 2005학년도 대학수학능력시험 Q & A 자료집, 한국교육과정평가원 홈페이지 2004.05.13일자 공개자료.
- 한국교육과정평가원 (2004b). 2005학년도 대수능 등급 표준점수 및 도수분포, 한국교육과정평가원 홈페이지 2004.12.14일자 공개자료.
- 허명희 (1994). 새 대학입시의 통계적 계획과 분석 - 문항분석과 선택과목 등화(표준점수제)를 중심으로, <한국통계학회 논문집>, 1, 215-225.

[2005년 3월 접수, 2005년 4월 채택]

A Study on an Alternative to the Standardized Scoring System in CSAT *

Hyungtae Hwang ¹⁾

ABSTRACT

In the current College Scholastic Aptitude Test (CSAT), the standardized scoring system is being adopted to adjust the differences of degrees of difficulty between the optional subjects. But it becomes clear that the system has several weak points, some of which are considered to be very serious. In this paper we propose an alternative method, so-called the additive scoring system. It determines the additional points per each subjects, according to the subject mean scores. The proposed method has been simulated using the data of 2005' CSAT, and it turns out that the additive scoring system reduces or remove the troubles caused by the standardized scoring system.

Keywords: CSAT, Standardized scoring system, Additive scoring system

* The present research was conducted by the research fund of Dankook University in 2004

1) Professor, Department of Information and Computer Science, Dankook University, san 8 Hannam-dong
Yongsangu, Seoul, 140-714, Korea
E-mail: hthwang@dankook.ac.kr