韓國數學教育學會誌 시리즈 A <數學敎育>
2007. 11. 제 46권, 제 4호, 407-421

J. Korea Soc. Math. Ed. Ser. A: The Mathematical Education
November. 2007, Vol. 46, No. 4, 407-421

대학수학능력시험 수리 영역 문항 난이도 예측을 위한 회귀모형 추정

이 상 하 (한국교육과정평가원)
이 봉 주 (한국교육과정평가원)
손 홍 찬** (한국교육과정평가원)

I. 들어가며

2008학년도부터 대학수학능력시험(이하 대수능)에 적용되는 9등급 점수 체제의 성공적인 정착을 위해서는 대수능 과목의 점수 분포가 적절하게 이루어질 수 있도록 다양한 난이도의 문항들로 검사를 구성할 필요가 있다. 사전에 문항의 난이도를 보다 정확하게 예측할 수 있어 그 정도를 원하는 만큼 조절할 수 있다면 대수능 과목의 점수분포는 보다 적절하게 나타날 것으로 기대할 수 있다. 그런데 현행 대수능 출제 체제에서는 예비검사를 실시하여 문항의 난이도를 추정하고 있지 않기 때문에, 고등학교 교사 및 대학 교수로 구성된 출제위원과 검토위원의 경험과 주관적인 판단에 전적으로 의존하고 있다. 즉, 대수능 출제에서 문항을 제작하여 하나의 검사지를 구성할 때, 교과 전문가들이 검사의 내용 관련 타당도를 검토하고 문항의 난이도를 예상하게 된다. 그러나 문항 난이도는 응시자 집단의 특성에 따라 달라질 수 있고, 현행 대수능의 난이도 추정치는 출제위원과 검토위원의 구성에 따라서도 달라질 수 있다. 따라서 해마다 응시자 집단과 출제위원과 검토위원이 바뀌는 현실에서는 추정한 난이도의 정확성과 항상성에는 한계가 있을 수밖에 없다. 또한 언어 영역과 외국어 영역을 제외한 모든 과목의 문항 수가 20~30개로 많지 않고 응시자 수가 수천 명에 불과한 과목도 있다. 이러한 조건들은 출제위원의 경험에 의한 주관적인 난이도의 추정으로는 모든 과목에서 일정한 비율로 응시자들의 등급을 나누어야 하는 대수능의 현실적인 목적을 달성하는 데 한계가 있음을 의미한다. 따라서 응시자 집단의 특성이나 출제위원의 인적 구성과는 상관없이 문항 자체 또는 검사지 자체로부터 난이도에 영향을 미치는 변인을 규명하여 그 영향의 정도를 파악하는 것이 필요하다.

미국의 대학입학시험인 SAT와 ACT의 경우, 예비검사를 통하여 문항 난이도를 비롯한 문항의 심리측정학적 정보를 추정하고, 그 정보를 이용하여 검사의 난이도 차이에 따른 학생들의 점수 차이를 보정하는 검사 동등화(test equating)까지 이루어지고 있다. 따라서 우리나라의 대수능처럼 문항 난이도를 내용 전문가의 주관적인 판단에만 의존하기 때문에 발생하는 문제로 고민하는 일은 없다. 또한 미국의 각 주에서 대규모로 시행하고 있는 각종 성취도 평가의 경우에도 예비검사가 이루어진 문항들을 사용하기 때문에 내용 전문가가 난이도를 가능한 정확하게 예측하기 위해 고민하는 일은 없어 보인다. Luppescu(1996)는 독해 시험 문항의 난이도에 영향을 주는 문항 변인들을 도출하여 난이도를 예측할 수 있는 모형을 추정하고, 이를 활용하여 예비검사 없는 검사 동등화 방안을 연구하였으나, 예비검사 없는 검사 동등화 방법의 필요성이 높지 않아 유의미한 후속연구로 이어지지는 않았다.

한편, 한국교육과정평가원은 수학 시험의 문항 난이도를 가능한 정확하게 예측하는 데 도움이 되는 국내 선행연구가 거의 없는 상황에서, 2003년도에 2002학년도와 2003학년도 대수능 수리 영역 문항의 특성에 근거하여

---

* 2007년 10월 투고, 2007년 11월 심사 완료
* ZDM분류 : D64
* MSC2000분류 : 97D60
* 주제어 : 대학수학능력시험, 대학입학시험, 수리 영역, 문항 난이도, 회귀방정식
** 교신저자: 손 홍 찬, 서울시 종로구 삼청동, 한국교육과정평가원 대수능연구관리처 출제연구부, 전화: 02-3704-3698, Fax: 02-3704-3740, E-mail: hcson@kice.re.kr

2004학년도 9월 모의평가 문항의 난이도를 예측할 수 있는 회귀방정식 추정 연구를 자체적으로 수행하였다(이종승·김성훈·김재철·송현정·박문환·장경숙·서재영, 2003). 이 연구에서 문항 난이도를 유의미하게 설명할 수 있는 독립변인은 문항의 내용 영역 및 행동 영역, 답지의 형식, 제재의 생소성, 문제의 복잡도, 계산의 복잡도로 나타났고, 이 6개의 독립변인들이 문항 난이도 분산의 55% 정도를 설명할 수 있는 것으로 나타났다. 추정된 회귀방정식으로 2004학년도 9월 모의평가의 문항 난이도를 예측한 결과, 예측 난이도와 실제 난이도 사이의 상관관계가 0.791로 나타났다. 이것은 2002학년도, 2003학년도 대수능 수리 영역 문항에 근거하여 추정된 회귀방정식 모형이 2004학년도 9월 모의평가 수리 영역 문항 난이도 분산의 63%를 설명할 수 있었다는 의미이다.

그러나 이 선행 연구에서 문항 난이도를 설명하는 문항 특성 변인들은 학습 이론에 근거한 것이 아니라 대수능 출제에서 사용되는 이원분류표와 교과 전문가들의 경험과 주관적인 판단에 의해 도출되었다. 또한 문항 난이도를 설명하는 문항 특성 변인들 중 범주 변인(categorical variable)에 해당하는 문항의 내용 영역 및 행동 영역, 답지의 형식은 각 변인의 하위 범주별로 대수능 기출문항들의 평균 정답률을 계산하고 그 크기에 따라 순위를 정하여 연속 변인(continuous variable)으로 변환되었다. 이 연속 변인은 각 범주 변인의 하위 범주들 간에 평균 정답률 차이가 일정하지 않기 때문에, 각 하위 범주의 평균 정답률 순위를 그 범주에 속하는 문항의 난이도 수준으로 정의한 순위 척도(ordinal scale)라고 할 수 있다. 선행연구에서는 또한 이 변인들과 함께 다른 문항 특성 변인들도 각각의 준거에 따라 연속 변인으로 정의하였고, 이 연속 변인들도 각 변인의 수준 간 난이도 차이를 일정한 것으로 가정하고 있다. 그러나 각 문항 특성 변인의 수준 간 문항 난이도 차이에 일정한 규칙성이 없음에도 일정한 것으로 가정하고 추정된 회귀방정식은 문항 난이도를 설명하고 예측하는 능력이 떨어진다고 볼 수 있다.

또한 이 선행 연구 결과를 현행 대수능 수리 영역의 문항 난이도 추정에 적용하기에는 다음과 같은 세 가지 제한점을 가지고 있다. 첫째, 제7차 교육과정에 따른 2005학년도 이후 수리 영역의 시험 범위와 그 이전 대수능 수리 영역의 시험 범위가 다르다. 2005학년도 이후 대수능 수리 영역의 시험 범위는 수학Ⅰ, 수학Ⅱ, 심화선택과목인 반면에, 2002·2003·2004학년도 대수능 수리 영역은 공통수학, 수학Ⅰ, 수학Ⅱ 외에도 고등학교 내용과 수준에서 다룰 수 있는 통합교과적 소재를 포함하고 있다. 둘째, 제7차 교육과정에 따른 2005학년도 이후 대수능 체제에서는 그 이전의 대수능과 달리 인문·자연·예체능계의 구분이 없어졌다는 것이다. 즉, 현행 수리 영역의 두 유형인 '가'형 또는 '나'형에 응시하는 집단은 그 이전에 자연계 또는 인문계에 응시했던 집단과 차이가 있다. 셋째, 이 선행 연구에서는 문항 난이도에 미치는 변인 중의 하나인 내용 영역의 분류가 세분화되어 있지 않았다. 즉, 대수능 수리 영역의 평가목표 이원분류표의 내용 영역 분류인 시험 과목의 대단원명을 기준으로 분류하는 대신에, 수학적 기초, 대수, 해석, 기하, 확률과 통계의 5개 영역으로 분류하여 난이도 모형을 추정하고 있다. 이러한 제한점을 고려하면 현행 대수능 수리 영역 출제 체제에 적합한 난이도 추정 모형을 재검토할 필요가 있다.

따라서 이 연구에서는 현행 대수능 출제 체제에서 보다 정확하고 안정적으로 수리 영역 문항의 난이도를 예측하기 위하여, 학습 이론과 선행 연구에 근거하여 문항 난이도를 유의미하게 설명할 수 있는 변인들을 수정·보완하고, 각 변인의 수준 간 문항 난이도 차이가 일정하도록 각 변인을 척도화하여 대수능 수리 영역 문항의 난이도를 보다 잘 예측할 수 있는 회귀방정식 모형을 탐색하고자 한다.

Ⅱ. 수리 영역 문항 난이도 예측 변인

문항 난이도[1]는 '전체 피험자 중에서 정답을 선택한 피험자의 비'로 정의되며 전체 피험자들이 문제를 해결할 때 느끼는 어려움의 정도를 나타내는 척도이다. 이 절에서는 문항 난이도에 영향을 미칠 수 있는 문항 특성

1) 응시자 수를 N이라 하고 정답자 수를 R이라고 하면, 문항 난이도는 $\frac{R}{N}$로 정의되어 0~1 사이의 값이 되고, 문항 정답률은 $\frac{R}{N}\times 100$ (%)로 정의되어 0~100%의 값을 가진다.

을 학습 모형이나 문제해결 모형을 통해 고찰한 다음, 문항 난이도 추정에 관한 선행 연구를 토대로 문항 난이도를 잘 설명할 수 있을 것으로 예상되는 요인들을 찾고자 한다.

문항의 난이도에 영향을 주는 문항 특성은 주어진 문항을 해결하는 과정을 설명할 수 있는 구체적인 학습 모형 또는 문제해결 모형으로부터 찾아볼 수 있다. 이러한 모형은 문항 난이도에 영향을 주는 문항 특성을 규명하는 데 중요한 방향을 제시할 수 있기 때문이다. 이 연구에서는 일반적인 학습 과정 및 문제해결 과정을 설명할 수 있는 자기-조절적 학습(self-regulated learning) 모형과 정보처리이론(information processing theory) 모형, 그리고 구체적인 문제해결 과정을 설명하는 Polya의 문제해결(problem solving) 모형에 근거하여 피험자들이 문제해결을 어렵게 느끼게 하는 요소들을 찾는다. 이론적으로 이들 모형에서 문제해결의 어려움에 영향을 주는 요인들 중에서 주로 문제의 속성과 관련된 것이 바로 문항 난이도에 영향을 주는 문항 특성이라고 할 수 있다.

먼저, 자기-조절적 학습 모형에 따르면 주어진 문제를 잘 해결하기 위해서는 우선 문제를 해결하려는 피험자의 동기(motivation)가 선행되어야 하고, 문제해결에 필요한 관련 지식 및 전략과 관련된 인지 능력, 문제해결에 필요한 모든 요소들을 적절하게 통합하고 조절할 수 있는 초인지(meta-cognition)가 필요하다(Davidson & Sternberg, 1998).

인지적 능력 측면에서는 피험자의 서술적 지식(declarative knowledge), 절차적 지식(procedural knowledge), 전략(strategy)이 문제해결에 중요한 역할을 한다. 초인지 측면에서는 어떤 지식과 전략을 언제 사용해야 할지 결정하는 초인지적 지식인 조건 지식(conditional knowledge)이 중요하다(Silver, 1982; Pressley, 1995). 피험자들은 수준 높은 내용 지식, 복잡한 절차 지식, 상황-구체적인 전략을 필요로 하는 문항에서 어려움을 느낄 가능성이 크다. 특히 문제해결에 필요한 지식과 전략의 수준이 높지 않다고 해도 어떤 지식과 전략을 사용해야 하는지에 대해 수준 높은 조건 지식을 필요로 하는 문항은 피험자로 하여금 문제를 어렵게 느끼게 할 것이다.

그러나 피험자의 초인지는 문제해결에 큰 영향을 주는 부분이지만 문항의 특성보다는 피험자 개인의 특성과 관련되어 있다. 즉, 문제해결에 필요한 피험자의 동기와 초인지는 문항 자체의 고유한 특성이라기보다 학습자의 특성에 보다 많이 관련되어 있기 때문에 문항 난이도에 영향을 주는 문항 특성에 관한 논의에서 제외한다. 따라서 자기-조절적 학습 모형으로부터는 문제해결에 필요한 관련 지식 및 전략과 관련된 인지 능력에 관련된 요인이 문항 난이도에 영향을 주는 문항 특성이 될 수 있음을 알 수 있다.

다음으로, 정보처리이론의 관점에서 보면 전문가의 마인드맵에는 많은 관련 개념들이 연결되어 있으며, 개념과 개념들 사이의 연결이 복잡하게 연결되어 있다. 이것은 그 개념이 주어졌을 때, 전문가의 작동기억(working memory)에 많은 양의 관련 정보들이 올라올 수 있음을 의미한다. 따라서 문제를 해결할 때, 전문가의 경우에는 문제해결에 필요한 모든 정보를 가지고 있을 가능성이 높기 때문에 비전문가에 비해서 상대적으로 문제해결의 가능성이 높다. 그리고 연결이 강하다는 것은 한 개념에 자극을 주면 나머지 개념들이 자동적으로 자극을 받을 수 있다는 것을 의미하는데, 그 거리가 멀어질수록 자극을 전달하기가 어려워진다. 전문가와 비전문가의 마인드맵의 차이로부터 실제로 문항의 난이도에 영향을 주는 변인들을 추론하는 것이 가능하다. 문제를 해결하는 데 필요한 개념의 개수가 많으면 많을수록 문항이라는 자극을 통해 작동기억에 올려야하는 개념의 개수가 증가하기 때문에 문제해결의 어려움에 영향을 줄 수 있다.

또한 이러한 개념들을 조합해서 효율적으로 문제를 해결할 수 있는 전략이 필요하다. 그 전략의 난이도가 문제해결의 어려움에 영향을 미치는 변인으로 고려될 수 있다. 한편, 정보처리이론의 관점에서도 문항 자극을 통해 올라온 개념 그리고 전략 중에서 어느 것을 사용하는 것이 필요한지를 결정하는 초인지적 조건 지식은 문항 난이도에 영향을 주지만, 앞서 언급한대로 개인적 특성과 관련이 깊어 문항 특성 논의에서는 제외한다.

마지막으로, 문제해결 모형에 따르면 문제해결의 첫 단계는 주어진 문제에 대한 정신적 표상(mental representation)을 작동기억에 구성하는 것이다(Davidson et al., 1994). 문제표상(problem representation)은 문제해

결과 관련된 서술적 지식, 절차적 지식, 전략, 조건 지식을 장기기억(long-term memory)에서 추출하여 형성하기 때문에, 문제의 해결책을 찾는 데 필요한 정보 또는 문제해결의 방향까지도 포함한다고 할 수 있다(Gredler, 1997). 문제해결의 첫 단계는 문항 관련 정보들을 장기기억에서 추출하여 문제표상을 작동기억에 구성하는 것으로, 작동기억의 용량을 많이 필요로 하는 문항 특성들은 문항의 난이도에 영향을 미치는 잠재적인 변인들이 될 수 있다. 즉, 작동기억에서 문제표상을 구성할 때, 익숙하지 않은 수학적 기호가 포함되거나, 일상생활에서 잘 접할 수 없는 상황들이 포함되거나, 문항의 형식이 새롭거나, 문제를 이해하는 데 필요한 개념의 개수, 문제를 해결하는 데 필요한 개념의 개수 등이 작동기억의 용량에 영향을 주는 변인이 될 수 있다. 이와 같은 변인들은 같은 문제라고 하더라도 개인에 따라 문제 표상을 구성하는 데 필요한 작동기억의 용량이 얼마든지 다를 수 있기 때문에 유의미한 설명 변인이 될 수 있다고 단정할 수 는 없다.

그러나 학습자의 능력과 상관없이, 문제해결에 필요한 개념의 수가 한 개인 것보다 2개인 문제의 표상을 올바르게 구성하는 데 좀 더 많은 작동기억의 용량이 필요하다는 것을 가정하는 것에 무리가 따르지는 않는다. 좀 더 많은 작동기억을 필요로 한다는 것은 상대적으로 좀 더 어려운 문항일 가능성이 높다는 것을 의미하기 때문이다. 이와 같이 문제표상 구성에 필요한 작동기억의 용량과 문항 난이도의 관계처럼, 문항의 난이도에 영향을 줄 수 있는 문항 특성들은 문제를 해결하는 데 필요한 개념의 개수, 문항의 생소성과 관련 있는 척도들인 것으로 추론할 수 있다.

이러한 학습 이론을 고찰한 결과, 문항 자체로부터 내용 지식, 절차적 지식, 문제해결에 필요한 개념의 수, 익숙하지 않은 기호나 상황이 주는 문항의 생소성 등이 문항의 난이도에 영향을 미치는 요인이 될 수 있음을 추정할 수 있다. 또한 문항의 난이도는 피험자의 동기나 전략 또는 초인지와 같이 문항 자체보다는 피험자의 특성에 종속되는 요인들 때문에 문항 난이도에 영향을 미치는 문항 자체의 요인만으로 난이도를 추정하는 것에는 어느 정도 한계가 있을 수 있음을 알 수 있다.

한편, 이종승 등(2003)의 '문항 난이도 추정 모형 개발 연구'와 박문환(2004)의 '대학수학능력시험 난이도 변인 탐색' 연구에서는 수리 영역 문항의 난이도에 영향을 미치는 요인으로 문항의 내용 영역 및 행동 영역, 문항 유형, 제재의 생소성, 계산의 복잡도, 문제의 복잡도, 보조적 자료, 풀이 방법의 다양성 등을 제시하고 있다. 따라서 이러한 선행 연구와 학습 이론의 공통 요인을 토대로 하여 대수능 수리 영역 문항 난이도에 영향을 미치는 문항 특성 요인으로 문항의 내용 영역 및 행동 영역, 문항 유형, 문항의 생소성, 계산량, 개념의 수를 도출하였다.

이와 같은 문항의 특성 외에도 문항 난이도를 예측할 수 있는 외적인 변인으로 출제위원의 예상정답률과 검사지에 배치된 문항의 위치를 고려할 수 있다. 현행 대수능 출제 체제에서 출제위원의 예상정답률은 문항 제작이 완료되는 시점에서 산출되고 이것은 검토위원의 예상정답률과 함께 난이도 조정에 유용하게 사용되고 있기 때문이다.

출제위원의 예상정답률은 위에서 열거한 여러 가지 요인들을 직관적으로 종합하여 산출되는 것으로 볼 수 있고, 실제정답률에 가까울수록 회귀방정식에서 다른 요인들이 난이도에 미치는 영향은 떨어진다고 볼 수 있다. 그러나 예상정답률이 실제정답률과 차이가 클 때는 다른 요인을 고려하여 난이도를 예측할 필요성이 증대된다.

또한 문항의 위치는 수리 영역 검사지의 문항의 배열의 관습에서 연유한 것으로, 문항 배열은 일반적으로 대단원의 순서와 난이도를 고려하며, 특히 단답형 문항의 위치가 고정된다는 특징이 있다. 문항 배열에서의 특정 위치는 앞에서 주어진 요인들과 대략적으로 관련이 있을 수 있지만, 시간제한으로 뒤에 위치한 문항일수록 문항 자체가 지니는 특성으로부터 추정한 것보다 정답률이 낮아질 수 있기 때문에 고려할 필요성이 있다. 문항의 위치를 고려한 회귀방정식에서 위치 변인이 의미 없게 나타날 수도 있지만, 그렇지 않은 경우에는 검사지를 구성한 다음 배열한 위치에 따라 문항 자체의 특성에서 비롯된 요인에 더하여 위치 요인을 보정하여 난이도를 계산할 필요가 있다.

지금까지의 논의의 결과로 본 연구에서는 대수능 수리 영역의 문항 난이도를 예측할 수 있는 요인으로 문항의 내용 영역 및 행동 영역, 문항의 유형, 문항의 생소

성, 계산량, 개념의 수, 출제위원의 예상정답률, 문항의 위치를 선정하였다.

Ⅲ. 연구 방법

1. 분석 대상

본 연구의 목적은 현행 대수능 수리 영역 출제 체제에 적합한 난이도 예측 모형을 추정하는 것으로, 2005~2007학년도 대수능 수리 영역 문항을 대상으로 '가'형 문항을 분석하였다. 특히 선택과목 중에서는 응시자 비율이 96% 이상을 차지하는 미분과 적분 문항을 대상으로 하였다. 이는 수리 영역 '가'형 응시자 집단의 점수 분포가 정규분포에 가까운 특징이 있는 반면에, '나'형 응시자 집단의 점수 분포가 대개 낮은 점수대에 학생이 몰려 있는 편포의 형태를 띠는 특징이 있기 때문이다. 따라서 이 연구에서는 학년도별로 대수능 수리 영역 '가'형의 공통문항 25개와 미분과 적분 문항 5개가 분석 대상이 되었다. 근래 2006학년도와 2007학년도 대수능에서 '가'형의 응시자 비율은 각각 26.4%와 23.4%였다.

분석 대상 문항의 예상정답률, 실제정답률, 내용 영역, 행동 영역, 문항 유형에 관한 정보는 2005~2007학년도 대수능 수리 영역의 문항카드 및 시험 결과 분석 보고서에서 수집하였다(이명준·이봉주·이대현, 2005; 이양락·이봉주·조윤동·변희현·손홍찬·민경석·시기자, 2006; 이양락·이봉주·손홍찬, 2007). 그러나 문항의 생소성, 계산량, 개념의 수는 본 연구의 저자들이 각각의 문항을 심층 분석하여 도출하였다.

2. 문항 특성 변인의 수준

내용 영역은 각 시험 범위 내 과목의 대단원에 해당하는 것으로 '수학 I'에서 행렬, 지수와 로그, 지수함수와 로그함수, 수열, 수열의 극한, 순열과 조합, 확률, 통계, 그리고 '수학 II'에서 방정식과 부등식, 함수의 극한과 연속성, 다항함수의 미분법, 다항함수의 적분법, 이차곡선, 공간도형과 공간좌표, 벡터, 그리고 '미분과 적분'에서 삼각함수, 함수의 극한, 미분법, 적분법의 19개 영역이다. 각 문항의 내용 영역 분류는 대수능 출제매뉴얼의 평가목표 이원분류표에 근거하여 이루어졌다(박선화·박문환·이봉주, 2004).

행동 영역은 대수능 수리 영역 평가목표 이원분류표에 나타나 있는 '계산', '이해', '발견적 추론', '연역적 추론', '수학 내적 문제해결', '수학 외적 문제해결'의 6개 영역에서 '이해'를 '기초 이해'와 '심화 이해'로 나누어 7개 영역으로 나누었다. '이해' 영역을 '기초 이해'와 '심화 이해'로 세분한 것은 그 영역의 문항 정답률이 다소 광범위하게 분포하는 경향이 있기 때문이다. 이종승 등(2003), 박문환(2004)의 연구에서도 '기초적 이해'와 '심화된 이해'로 세분하여 행동 영역을 총 7개 항목으로 구분하여 정답률의 차이를 분석한 바 있다. 각 문항의 행동 영역 분류는 평가목표 이원분류표에 근거하여 이루어졌다. 단, '기초 이해' 영역으로 분류한 문항은 평가목표 이원분류표에서 '이해' 영역으로 분류된 문항 중에서 주어진 문제와 관련된 수학적 개념을 파악하거나 주어진 문제 상황을 수학적으로 표현할 수 있는지를 묻는 문항이고, 평가목표 이원분류표의 '이해' 영역에 해당하는 문항 중에서 '기초 이해' 영역으로 분류되지 않는 문항은 '심화 이해' 영역으로 분류하였다.

문항의 유형은 최선답형, 합답형, 단답형의 3가지 유형으로 나누었다. 문항에 대해 수험자가 제한된 답지에서 선택하여 표기하는지 또는 구성하여 작성하는지에 따른 분류에 의거하면 선택형인 5지선다형과 서답형인 단답형의 두 가지로 볼 수 있지만, 5지선다형을 답지 구성 방식에 따라 최선답형과 합답형 두 가지로 나눌 필요가 있다. 흔히 3개의 명제가 주어지고 그 진위 여부를 묻는 합답형 문항의 정답률은 최선답형 문항에 비해 정답률이 낮아서(이양락 등, 2006), 합답형 문항 수의 변화는 전체 검사지의 난이도에 영향을 미치게 된다.

계산량은 문항을 해결하는 데 필요한 계산의 길이를 뜻하는 것으로 크게 4가지로 나누었다. 계산의 길이는 그 길이가 길수록 많은 시간이 소요되고 실수의 가능성도 높기 때문에 정답률에 영향을 미치는 요인이 된다. 계산량에서는 계산이 복잡하거나 어려운 정도는 위에서 열거한 다른 요인들에 의해 영향을 받을 수 있기 때문에 고려하지 않았고, 계산의 길이만 고려하였다. 수험자들이 출제 의도를 반영하는 풀이 방법으로 대수능 시험지의 여백에 풀었을 때를 기준으로, 계산량을 1~3줄의 길이,

4~6줄의 길이, 7~10줄의 길이, 11줄 이상으로 하여 구분하였다.

개념의 수는 문항을 해결하기 위해 필요한 개념들의 수로 주로 교과서의 소단원의 명칭에 해당하는 개념을 단위로 하였다. 교과서마다 단원의 구성은 약간씩 다르지만 최소 구성단위인 소단원의 경우 거의 일치함을 확인할 수 있다. 개념 수의 척도는 1개, 2개, 3개, 4개 이상의 네 가지로 나누었다. 문제를 푸는데 필요한 개념의 수를 명확히 한다는 것은 어려운 일이지만 그 중요성을 고려하였다. '수학 I' 이전에 배운 내용일지라도 문항의 해결에 중요한 역할을 하는 것은 개념의 수에 포함시켰다. 어느 한 개념을 알아야 다른 개념을 알 수 있는 경우에는 한 개로 세지 않고 2개로 셌으며, 경우에 따라서는 소단원명과 일치하지는 않지만 문제 풀이에 중요한 역할을 하는 요소도 경우에 따라 개념의 수에 포함시켰다.

문항의 생소성은 문항의 형태, 아이디어의 참신성, 교과서나 참고서에 자주 나타나는 정도 등을 고려하여 매우 익숙한 편, 익숙한 편, 생소한 편, 매우 생소한 편의 4가지로 나누었다.

출제위원의 예상정답률은 모든 출제위원들이 산출한 문항별 예상정답률의 평균을 사용하였고, 문항의 위치는 문항 번호를 사용하였다.

3. 문항 특성 변인의 척도화

대수능 기출 문항들의 난이도에 근거하여 문항 특성 변인들의 수준 간 난이도 차이가 일정하도록 척도화하였다. 이러한 척도화는 각 문항 특성 변인과 문항 난이도 간의 관계를 일차식으로 단순화할 뿐만 아니라 문항 난이도와의 상관관계를 향상시켜 회귀방정식의 모형의 설명력을 높일 수 있는 장점이 있다. 특히 문항 특성 변인 중에서 범주 변인인 내용 영역, 행동 영역, 문항 유형도 연속 변인으로 변환하면서 수준 간 난이도 차이가 일정하도록 척도화 하였다.

문항의 내용 영역은 대단원 수준에서 문항들을 분류하고 각 단원별로 문항의 평균 정답률을 계산하였다. 평균 정답률을 1의 자리에서 반올림하고 0.1을 곱하여 각 대단원의 난이도 수준을 나타내는 척도로 정의하였다. 이와 같은 방법으로 문항의 행동 영역 변인도 척도화하였다. 그러나 문항유형 변인의 경우에는 수준 간 평균 정답률 차이가 일정하였기 때문에 수준별 평균 정답률의 내림차순 순위를 그대로 사용하였다. 이러한 척도화 과정은 범주 변인을 연속 변인으로 변환하는 것일 뿐만 아니라 연속 변인의 수준 간 난이도 차이를 일정하게 만들어 주는 효과가 있다. 특히 범주 변인을 연속 변인으로 변환하는 것은 회귀분석에서 내용 영역 범주의 개수와 관계없이 내용 영역 변인의 자유도(degree of freedom)를 1로 고정시켜 독립변인의 유의미성을 검정하는 데 유리하다.

한편 개념의 수, 계산량, 문항의 생소성 변인들은 연속 변인으로 볼 수 있으나, 각 변인의 수준 간 평균 정답률 차이가 일정하지 않았다. 따라서 이들 변인에 대해서도 문항의 내용 영역 변인을 척도화한 방법과 마찬가지로 각 변인의 수준과 관련된 문항들의 평균 정답률에 근거하여 각 수준을 척도화 하였다. 각 문항 특성 변인의 수준별 척도는 <표 Ⅲ-1>, <표 Ⅲ-2>, <표 Ⅲ-3>, <표 Ⅲ-4>와 다음과 같다.

<표 Ⅲ-1> 내용 영역 수준별 척도

수준	행렬, 지수와 로그	확률, 함수의 극한과 연속성	수열, 통계, 방정식과 부등식, 다항함수의 미분, 다항함수의 적분, 삼각함수, 함수의 극한	지수로그와 함수, 수열의 극한, 순열과 조합, 이차곡선	공간도형과 공간좌표, 벡터, 적분법	미분법
척도	8	7	6	5	4	2

<표 Ⅲ-2> 행동 영역 수준별 척도

수준	계산	기초적 이해, 연역적 추론, 수학 외적 문제해결	발견적 추론	심화된 이해, 수학 내적 문제해결
척도	8	6	5	4

<표 Ⅲ-3> 문항 유형 수준별 척도

수준	최선답형	합답형	단답형
척도	1	2	3

<표 III-4> 개념의 수, 계산량, 문항의 생소성, 수준별 척도

개념의 수		계산량		문항의 생소성	
수준	척도	수준	척도	수준	척도
1개	8	1~3줄	9	매우 익숙하다	8
2개	6	4~7줄	6	익숙한 편이다	6
3개	4	7~10줄	5	생소한 편이다	5
4개 이상	3	11줄 이상	4	매우 생소하다	4

4. 자료 분석 방법 및 절차

가. 2005, 2006, 2007학년도 대수능 수리 영역 문항 난이도를 유의미하게 설명할 수 있는 회귀방정식을 추정하였다.

나. 추정된 회귀방정식의 유용성과 실용성을 비교하기 위하여 각각의 회귀방정식을 이용하여 2007학년도 대수능 9월 모의평가 수리 영역 문항의 난이도를 예측하였다.

다. 각각의 회귀모형을 이용한 문항 난이도 예측치가 실제 난이도를 어느 정도 설명할 수 있는지를 알아보기 위하여 회귀분석을 하였다.

Ⅳ. 연구 결과

1. 회귀모형 추정

가. 회귀모형 1: 예상정답률

현행 대수능 출제 체제에서는 문항 제작자가 대수능 기출 문항들의 정답률 자료를 참고하여 문항의 정답률을 예상하고 있다. 이러한 예상정답률은 대수능 기출 문항들의 정답률에 근거를 두고 있지만, 문항 제작자들의 경험에 의한 주관적인 판단에도 크게 의존하고 있다. 따라서 대수능 문항 제작자의 구성이 달라짐에 따라 문항의 예상정답률과 실제정답률에 차이가 있을 수 있기 때문에, 학생들의 바람직한 점수 분포를 담보하는 데 어려움이 있다. 그러나 예상정답률 산출은 현행 대수능 출제 체제에서 사전에 문항의 정답률을 예측하는 유일한 방법이었다. <표 Ⅳ-1>은 출제 과정에서 문항 제작자들이 주관적으로 판단한 문항의 예상정답률이 어느 정도 문항 난이도를 설명할 수 있는지를 알아보기 위한 회귀분석 결과이다. 문항 제작자들이 예상한 정답률이 문항의 난이도 분산의 78%를 설명할 수 있었으며, 예상정답률 x_0와 실제정답률[2] y의 관계를 나타내는 회귀방정식은

$$y = -10.06 + 1.22x_0 \quad \cdots \text{(회귀모형 1)}$$

인 것으로 나타났다. 이 식에 따르면, 예상정답률이 50%인 문항들의 실제정답률은 51%(문항 난이도는 0.51)가 된다. 이는 예상정답률이 50%인 문항들의 평균 실제정답률은 대략 51%라는 의미이다. 또한 예상정답률이 50%인 문항들에 대한 실제정답률의 90% 신뢰구간은 34%와 68% 사이인 것으로 나타났다.

<표 Ⅳ-1> 예상정답률에 따른 문항 난이도 추정

Source	DF	Sum of Squares	Mean Square	F	p
Model	1	33546.64	33546.64	318.31	<.0001
Error	88	9274.29	105.39		
Total	89	42820.93			

나. 회귀모형 2: 문항 특성 변인

<표 Ⅳ-2>는 이 연구에서 정의된 문항 특성 변인들이 문항 난이도를 어느 정도 설명할 수 있는지를 알아보기 위한 회귀분석의 결과이다. 유의수준 5% 수준에서 문항의 생소성을 제외한 5개의 문항 특성 변인들이 유의미하게 문항의 난이도를 설명하는 것으로 나타났으며, 이 5개의 변인들이 문항 난이도 분산의 70%를 설명할 수 있는 것으로 나타났다. 이 문항 특성변인들과 실제정답률 y의 관계를 나타내는 회귀방정식은

$$\begin{aligned} y = &-16.75 + 4.56x_1 + 4.50x_2 - 6.13x_3 \\ &+ 2.91x_4 + 2.64x_5 \quad \cdots \text{(회귀모형 2)} \end{aligned}$$

인 것으로 나타났다. 예를 들어, [함수의 극한과 연속성($x_1 = 7$), 발견적 추론($x_2 = 5$), 합답형($x_3 = 2$), 개념수

2) 종속변인인 문항 난이도는 1보다 작은 수이고, 독립변인인 예상정답률은 0~100 사이의 수이기 때문에, 회귀방정식의 계수들의 크기가 너무 작아지는 경향이 있었다. 따라서 종속변인을 문항 난이도 대신에 실제정답률을 사용하였다. 정답률을 100으로 나누면 문항 난이도가 된다.

2개($x_4=6$), 복잡한 계산($x_5=5$)]의 특성을 가지는 문항의 실제정답률은 53%(문항 난이도는 0.53)정도인 것으로 추정된다. 이것은 주어진 문항 특성을 가진 문항들의 평균 실제정답률은 대략 53%라는 의미이다. 또한 이러한 특성을 가진 문항들에 대한 실제정답률의 90% 신뢰구간은 32%와 74% 사이인 것으로 나타났다.

<표 IV-2> 문항 특성 변인에 따른 문항 난이도 추정

Source	DF	Sum of Squares	Mean Square	F	p
Model	5	29852.97	5970.59	38.67	<.0001
Error	84	12967.96	154.38		
Total	89	42820.93			

Source	DF	Type III SS	Mean Square	F	p
내용영역(x_1)	1	2199.29	2199.29	14.25	0.0003
행동영역(x_2)	1	2093.85	2093.85	13.56	0.0004
문항유형(x_3)	1	2357.64	2357.64	15.27	0.0002
개념수(x_4)	1	990.33	990.33	6.41	0.0132
계산난이도(x_5)	1	646.55	646.55	4.19	0.0438

다. 회귀모형 3: 문항 특성 변인과 문항 위치

대수능 출제 체제에서는 경험과 직관에 근거하여 상대적으로 쉬운 문항들은 전반부로 배치하고 어려운 문항들은 후반부에 배치하고 있다. 따라서 시험에서 문항의 위치를 나타내는 문항 번호는 문항 제작자의 경험과 직관에 의한 문항 난이도와 관련이 있다고 할 수 있다. 문항 제작자들도 문항의 난이도에 영향을 미치는 것으로 나타난 문항 특성뿐만 아니라 자신의 경험과 직관에 근거하여 문항들의 난이도를 예상하고 문항들을 배치하는 것이다. 따라서 문항 위치 정보는 문항 특성 정보에 의해 설명되지 않는 문항 난이도를 추가적으로 설명할 수 있는 가능성을 갖고 있다.

<표 IV-3>은 문항 특성 변인뿐만 아니라 문항 위치 정보까지 포함시켜 문항 난이도를 추정하는 회귀분석의 결과이다. 유의수준 5%수준에서 문항의 생소성을 제외한 5개의 문항 특성 변인뿐만 아니라 문항위치 변인이 유의미하게 문항의 난이도를 설명하는 것으로 나타났으며, 이들 변인들이 문항 난이도 분산의 73%를 설명하는 것으로 나타났다. 즉, 문항위치 정보는 5개의 문항 특성 변인들이 설명할 수 있는 문항 난이도 분산의 양을 3% 정도 증가시켰다. 따라서 문항제작자의 경험과 직관에 따른 문항배치 정보는 5개의 문항 특성 변인들이 설명할 수 없는 문항 난이도를 추가적으로 설명할 수 있다는 것을 알 수 있다. 이 회귀분석에서 유의미한 문항관련 변인들과 실제정답률 y의 관계를 나타내는 회귀방정식은

$$y=-0.25+3.06x_1+4.43x_2-4.06x_3+2.78x_4+2.42x_5-0.61x_7\cdots \text{ (회귀모형 3)}$$

인 것으로 나타났다. 예를 들어, [함수의 극한과 연속성($x_1=7$), 발견적 추론($x_2=5$), 합답형($x_3=2$), 개념수 2개($x_4=6$), 복잡한 계산($x_5=5$), 문항위치($x_7=20$)]의 특성을 가지는 문항의 실제정답률은 52%(문항 난이도 0.52)인 것으로 추정된다. 이것은 주어진 문항 특성을 가진 문항들의 평균 실제정답률은 대략 52%라는 의미이다. 또한 이러한 특성의 문항들에 대한 실제정답률의 90% 신뢰구간은 33%와 71% 사이인 것으로 나타났다.

<표 IV-3> 문항 특성 변인 및 문항 위치에 따른 문항 난이도 추정

Source	DF	Sum of Squares	Mean Square	F	p
Model	6	31304.24	5217.37	37.60	<.0001
Error	83	11516.70	138.76		
Total	89	42820.94			

Source	DF	Type III SS	Mean Square	F	p
내용영역(x_1)	1	852.72	852.72	6.15	0.0152
행동영역(x_2)	1	2029.12	2029.12	14.62	0.0003
문항유형(x_3)	1	873.15	873.15	6.29	0.0141
개념수(x_4)	1	905.31	905.31	6.52	0.0125
계산난이도(x_5)	1	539.57	539.57	3.89	0.0519
문항위치(x_7)	1	1451.27	1451.27	10.46	0.0018

라. 회귀모형 4: 예상정답률과 문항 특성 변인

대수능 출제 체제에서는 문항 제작자로 하여금 대수능 기출문항의 정답률을 참고하여 문항의 정답률을 예상하게 한다. 그런데 문항 제작자는 문항 난이도에 영향을 주는 문항특성을 나름대로 고려하여 정답률을 예상하기 때문에, 회귀모형 1의 예상정답률이 문항 난이도 분산의

많은 부분을 설명할 수 있었던 것으로 보인다. 그러나 예상정답률 산출은 문항 제작자들의 경험에 의해 주로 추정될 뿐, 문항 특성 변인들을 체계적으로 고려하여 문항 난이도를 추정하는 체계는 아니다. 따라서 문항제작자의 구성이 바뀌게 되면 예상정답률 추정치가 여전히 달리질 가능성이 있다.

문항 제작자가 추정한 예상정답률과 문항 위치를 포함하는 문항 특성 변인들이 각각 문항 난이도를 어느 정도 설명할 수 있는지를 알아보기 위하여 회귀분석을 하였다. 이 회귀분석에서는 예상정답률과 개념 수의 변인만이 유의미한 것으로 나타났다. <표 Ⅳ-4>는 유의수준 5%에서 유의미하지 않은 독립변인들을 제외하고 유의미한 독립변인인 예상정답률과 개념의 수로 문항 난이도를 추정하는 회귀분석 결과를 보여주고 있다.

<표 Ⅳ-4> 예상정답률과 개념의 수에 따른 문항 난이도 추정

Source	DF	Sum of Squares	Mean Square	F	p
Model	2	34134.44	17067.22	170.94	<.0001
Error	87	8686.50	99.84		
Total	89	42820.94			

Source	DF	Type Ⅲ SS	Mean Square	F	p
예상정답률(x_1)	1	19968.69	19968.69	200	<.0001
개념수(x_4)	1	587.80	587.80	5.89	0.0173

<표 Ⅳ-4>는 두 독립변인들이 문항 난이도 분산의 80%를 설명하는 것으로 나타났다. 이 회귀분석에서 유의미한 문항관련 변인들과 실제정답률 y의 관계를 나타내는 회귀방정식은

$$y = -17.35 + 1.12x_1 + 2.21x_4 \cdots \text{(회귀모형 4)}$$

인 것으로 나타났다. 예를 들어, [예상난이도($x_1 = 50\%$), 개념수 2개($x_4 = 6$)]의 특성을 가지는 문항의 실제정답률은 52%(문항 난이도는 0.52)인 것으로 추정된다. 이것은 주어진 문항 특성을 가진 문항들의 평균 실제정답률이 대략 52%라는 의미이다. 그리고 이러한 특성의 문항들에 대한 실제정답률의 90% 신뢰구간은 35%와 68% 사이인 것으로 나타났다.

요약하면, 문항의 난이도에 대한 문항 제작자의 직접적인 판단이 제외된 회귀모형 2의 경우 문항 난이도 분산의 70%를 설명할 수 있었으며, 대수능 출제에서 문항의 상대적인 난이도 차이에 대한 직접적인 판단의 척도가 될 수 있는 문항 위치가 설명 변인으로 추가된 회귀모형 3은 문항 난이도 분산의 73%를 설명하였다. 또한 문항 제작자들의 경험과 대수능 기출 문항의 난이도 자료에 근거하여 주관적으로 판단한 예상정답률로 문항 난이도를 추정하는 회귀모형 1은 실제 문항 난이도의 78%를 설명할 수 있었으며, 문항 특성 변인 중에서 개념의 수가 설명 변인으로 추가된 회귀모형 4는 실제 문항 난이도 분산의 80%를 설명할 수 있었다.

2. 회귀모형을 이용한 난이도 예측

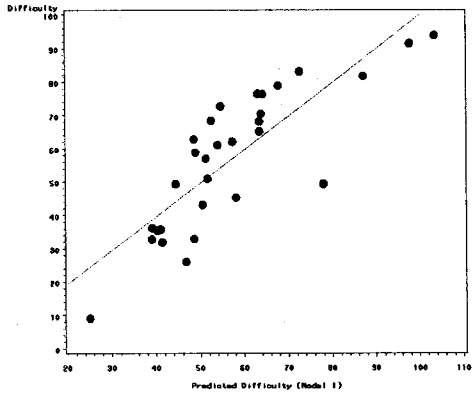
2005 · 2006 · 2007학년도 대수능 수리 영역 문항의 난이도를 설명하는 회귀모형들의 유용성을 검정하기 위하여, 추정된 각각의 회귀모형을 이용하여 2007학년도 대수능 9월 모의평가 수리 영역의 문항 난이도를 예측하였다. <표 Ⅳ-5>에 각 회귀모형에 의한 예측 난이도와 실제 난이도 간의 피어슨 상관계수와 p값이 제시되어 있다. 예상정답률을 독립변인으로 가지고 있는 회귀모형에 의해 예측된 문항 난이도와 실제 문항 난이도의 상관계수는 약 0.8 정도였으며, 문항 특성만을 독립변인으로 가지고 있는 회귀모형에 의해 예측된 문항 난이도와 실제 문항 난이도의 상관계수는 약 0.7 정도였다.

<표 Ⅳ-5> 예측 난이도와 실제 난이도 사이의 피어슨 상관계수

		예측 난이도			
		회귀모형 1	회귀모형 2	회귀모형 3	회귀모형 4
실제 난이도		0.8312	0.6984	0.6942	0.8181
	p*	<.0001	<.0001	<.0001	<.0001
예측 난이도	회귀모형 1		0.8422	0.8755	0.9889
	p		<.0001	<.0001	<.0001
	회귀모형 2			0.9788	0.8759
	p			<.0001	<.0001
	회귀모형 3				0.9101
	p				<.0001

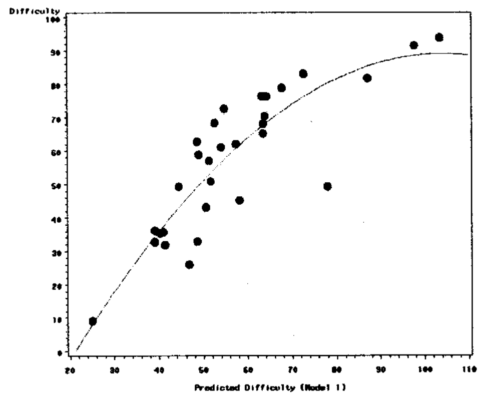
* 귀무가설(r=0) 하에 모수가 추정치보다 크게 될 확률

각각의 회귀모형에 근거하여 2007학년도 대수능 9월 모의 평가 수리 영역 문항의 난이도를 예측하고, 이 예측 문항 난이도가 실제 문항 난이도를 어느 정도 설명할 수 있는 지를 알아보기 위하여 예측 문항 난이도를 독립변인으로 하고 실제 문항 난이도를 종속변인으로 하는 회귀분석을 실시하였다.



$R^2 = 0.6908$

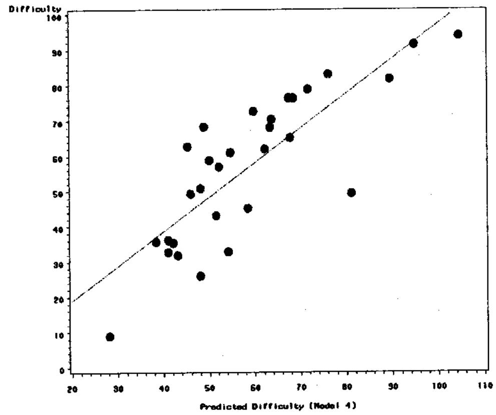
<그림 1> 회귀모형 1의 예측 난이도(직선)



$R^2 = 0.7568$

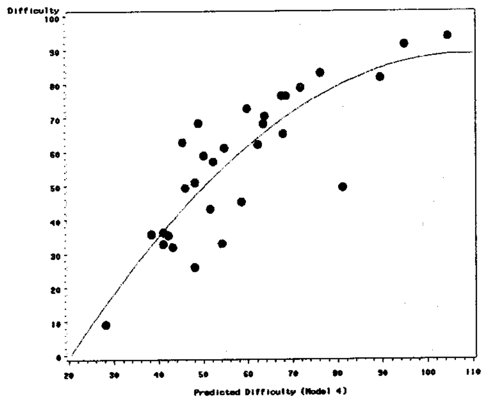
<그림 2> 회귀모형 1의 예측 난이도(이차곡선)

<그림 1>은 예상정답률을 독립변인으로 가지는 회귀모형 1을 이용한 문항 난이도 예측치로 실제 문항 난이도를 설명하기 위하여 직선 모형을 적용하였을 경우를 나타내고 있는데, 이 직선 모형은 실제 문항 난이도 분산의 69%를 설명할 수 있는 것으로 나타났다. <그림 2>는 직선 대신에 이차곡선을 적용하였을 경우를 나타내고 있으며, 이 이차곡선 모형은 실제 문항 난이도 분산의 76%를 설명할 수 있는 것으로 나타났다.



$R^2 = 0.6693$

<그림 3> 회귀모형 4의 예측 난이도(직선)

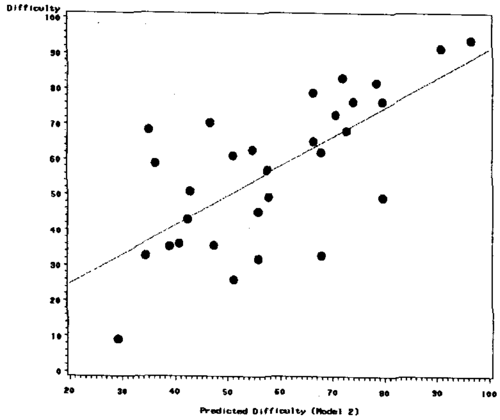


$R^2 = 0.7133$

<그림 4> 회귀모형 4의 예측 난이도(이차곡선)

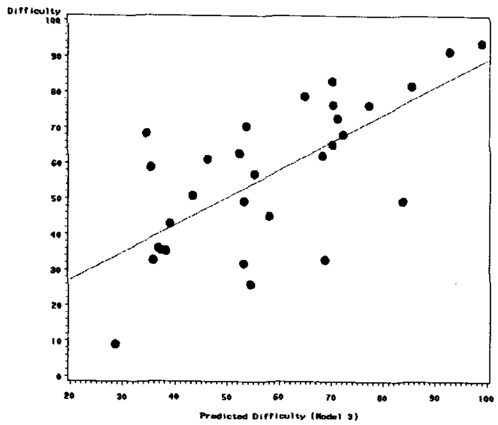
<그림 3>과 <그림 4>는 예상정답률과 문항의 개념수를 독립변인으로 가지는 회귀모형에 의해 예측된 문항 난이도와 실제 문항 난이도 사이의 관계를 나타내고 있

다. <그림 3>에서 예측 문항 난이도와 실제 문항 난이도의 관계를 설명하기 위하여 직선 모형을 적용하였는데, 실제 문항 난이도 분산의 67% 정도를 설명할 수 있는 것으로 나타났다. <그림 4>에서 직선 모형 대신에 이차곡선 모형을 적용하였는데, 실제 문항 난이도 분산의 71%를 설명할 수 있는 것으로 나타났다.



$R^2 = 0.4877$

<그림 5> 회귀모형 2의 예측 난이도



$R^2 = 0.4818$

<그림 6> 회귀모형 3의 예측 난이도

<그림 5>와 <그림 6>은 예상정답률을 독립변인으로 가지지 않는 회귀모형에 의해 예측된 문항 난이도와 실제 문항 난이도 사이의 관계를 나타내고 있다. <그림 5>는 문항 특성변인만을 독립변인으로 가지는 회귀모형 2에 의해 예측된 문항 난이도와 실제 문항 난이도 사이에 직선 모형을 적용한 경우를 보여주고 있으며, 실제 문항 난이도 분산의 약 49%를 설명하였다. <그림 6>은 문항 특성 변인과 문항 위치 변인을 독립변인으로 가지는 회귀모형에 의해 예측된 문항 난이도와 실제 문항 난이도 사이에 회귀직선을 적용한 경우를 보여주고 있으며, 실제 문항 난이도 분산의 약 48%를 설명할 수 있는 것으로 나타났다.

요약하면, 각각의 회귀모형에 의해 추정된 예측 난이도와 실제 난이도 사이의 상관계수의 크기는 0.69~0.83으로 모두 유의미하였다. 회귀모형의 의해 추정된 예측 난이도와 실제 난이도 관계를 설명하기 위하여 직선을 적용하였을 때, 예상정답률을 설명변인으로 가지는 회귀모형 1과 4의 경우 각각 실제 난이도 분산의 69%와 67%를 설명할 수 있었다. 예상정답률을 설명변인으로 가지지 않는 회귀모형 2와 3은 실제 난이도 분산의 49%와 48%를 설명할 수 있었다. 특히, 회귀모형 1과 4에 의해 예측된 난이도와 실제 난이도 사이의 관계를 설명하는 회귀모형은 직선보다 이차곡선을 적용하는 것이 바람직한 것으로 나타났는데, 이차곡선을 적용하였을 때 실제 난이도 분산의 76%와 71%를 설명할 수 있는 것으로 나타났다.

V. 논의

1. 문항특성 변인

문항의 난이도에 영향을 줄 수 있을 것으로 추측한 문항 특성 변인들 중에서 다른 문항 특성 변인들이 포함되었을 때 유일하게 유의미하지 않은 변인이 문항의 생소성이었다. 문항의 생소성은 피험자가 문제해결에 필요한 지식과 전략들을 불러 적절한 문제표상을 만드는 것을 어렵게 할 뿐만 아니라, 문제를 효율적으로 해결할 수 있는 상황-구체적인 전략 대신에 상황-일반적인 전략을 사용하게 할 가능성이 높다. 이러한 조건들이 결국은 피험자의 작동기억 용량을 많이 차지하게 되어 문제

해결을 더욱 어렵게 한다. 그럼에도 불구하고 문항 난이도 추정 회귀모형에서 문항의 생소성이 유의미한 변인으로 나타나지 않았던 이유는 크게 다음 세 가지로 요약할 수 있다. 첫째, 문항의 생소성은 문항 난이도를 설명하는 유의미한 변인이지만 다른 문항 특성 변인들이 문항의 생소성이 설명할 수 있는 부분을 나누어 설명한다고 볼 수 있다. 둘째, 문항의 생소성이 문항의 난이도에 유의미한 영향을 주지만 대수능 문제는 고등학교 교육과정과 연계성을 가지면서 기출 문항을 배제하고 있기 때문에 문항 생소성의 범위를 제한함으로써 대수능 수리 영역의 문항들의 생소성의 차이가 문항 난이도를 설명하는 데 유의미하지 않은 것으로 나타났을 수 있다. 셋째, 문항의 생소성은 여러 가지 차원에서 기인할 수 있으나, 각각의 차원을 하나의 독립변인으로 분리하여 정의하지 않았기 때문일 수 있다. 문항은 여러 가지 이유로 피험자에게 생소하게 느껴질 수 있으나, 본 연구에서는 문항의 형태, 아이디어의 참신성, 교과서나 참고서에 자주 나타나는 정도 등을 교과 전문가의 입장에서 개략적으로 판단하고, 이를 문항 난이도를 설명하는 회귀분석을 위하여 일차원적 척도로 변환하였었다. 그러나 문항의 생소성에 대한 이러한 척도화 방법은 다차원적인 문항의 생소성을 각 차원에 대하여 일관된 가중치를 둘 수 없게 만드는 한계를 가지고 있다.

2. 회귀모형 추정

현행 대수능 출제 체제에서는 문항 제작자와 검토자가 그들의 경험을 토대로 문항 난이도가 실린 대수능 기출 문제 자료집 등을 참고하여 문항 난이도를 예측하고 있다. 예상정답률로 문항 난이도를 설명하는 회귀모형은 문항 난이도 분산의 78%를 설명할 수 있었다. 한편, 문항 특성 변인만을 이용하여 문항 난이도를 설명하는 회귀모형은 문항 난이도 70%정도를 설명할 수 있었다. 예상정답률과 모든 문항 특성변인들로 문항 난이도를 설명하는 회귀모형을 추정하였을 때, 문항 특성 변인들 중에 유일하게 문항 풀이에 필요한 수학적 개념의 수만이 유의미하였으며 문항 난이도 분산의 80%를 설명할 수 있었다. 이 세 회귀모형을 비교하여 얻을 수 있는 시사점은 세 가지로 정리된다.

첫째, 문항 제작자들에 의해 추정된 예상정답률은 문항 특성 변인들이 설명할 수 있는 문항 난이도 분산의 대부분을 설명할 수 있었지만, 문항 특성 정보 중에서 '개념의 수'로 설명할 수 있는 문항 난이도 분산의 2%를 설명할 수는 없었다.

둘째, 개념의 수를 제외한 나머지 문항 특성 변인들로 설명되는 문항 난이도 분산의 68%가 예상정답률로 설명되는 문항 난이도 분산 78%에 포함되는 것으로 간주할 수 있다. 따라서 문항 제작자들이 여러 가지 문항 특성을 종합적으로 판단하여 예상정답률을 추정하였다고 가정한다면, 문항 제작자와의 면담이나 실제 난이도와 추정된 난이도 사이에 큰 차이가 나는 문항들을 분석함으로써 최소한 문항 난이도 분산의 10% 정도를 더 설명할 수 있는 문항 특성 변인들을 찾는 것이 가능할 것이다. 예를 들면, 유의미하지 않은 변인으로 나타난 문항의 생소성과 관련 있는 변인들이 현재 설명되지 않고 있는 문항 난이도 분산의 10% 중에서 일정 부분을 설명할 수 있을 것으로 보인다.

셋째, 현행 대수능 출제에서 수리 영역 문항들은 문항의 난이도와 문항의 내용 영역(대단원)에 따라 배열하는 경향이 있었다. 그런데 문항의 위치에 포함되어 있는 문항의 내용 영역 정보가 설명할 수 있는 문항 난이도 차이는 문항 특성 변인 중의 하나인 내용 영역 변인에 의해 충분히 설명되는 것으로 간주할 수 있다. 따라서 문항의 위치가 회귀모형에서 유의미하게 나타난 사실은 문항의 상대적인 난이도 차이에 대한 문항 제작자의 주관적 판단이 문항 특성 변인들 중에는 포함되지 않으면서 문항 난이도 분산을 추가적으로 설명할 수 있게 함을 보여준다. 즉, 문항 특성 정보와 함께 문항 난이도를 설명하는 회귀모형에서 문항 특성 정보들이 설명할 수 없는 문항 난이도 분산의 3%를 추가로 설명할 수 있었다.

이 유의미한 문항의 위치 정보가 문항 난이도 추정에 주는 시사점이 있다. 고등학교 3학년 학생들을 가르쳐 본 경험이 없거나 우리나라 전체 학생들의 수준에 대해 알지 못하는 교과 전문가들에게는 각각의 문항 난이도를 직접 추정하는 것보다 문항 특성을 종합적으로 고려하여 여러 문항들의 난이도를 상대적으로 비교하는 것이 훨씬 수월할 것이다. 따라서 문항 특성만을 고려하여 수리 영역 문항 30개의 난이도를 추정해야 하는 경우 교과 전문

가로 하여금 문항 특성뿐만 아니라 문항 난이도를 종합적으로 판단하여 문항을 나열하게 함으로써 문항 난이도 추정치의 정확도를 조금 높일 수 있을 것으로 추정된다.

넷째, 문항의 위치가 유의미하게 나타난 것은 제한된 시간 내에 모든 문제를 풀어야 하는 시험에서 각 문항에 대한 시간 배분에 실패하는 피험자들이 많이 있기 때문일 수 있다. 이 경우 시험지의 후반부에 위치하고 있는 문항일수록 문항 정답률이 떨어지는 경향이 생기게 된다. 즉, 유의미한 문항의 위치 변인은 한 문항이 시험지 내에 어느 위치에 있느냐에 따라 난이도가 달라질 수 있다는 것을 의미하고 있다. 따라서 회귀모형 추정에서 유의미하게 나타난 문항 특성 정보와 문항의 난이도 순서 정보뿐만 아니라 최종적인 문항 배치 정보까지 설명 변인으로 추가한다면 문항 난이도 추정치의 정확도를 높일 수 있을 것으로 보인다.

3. 회귀모형을 이용한 난이도 예측

문항 제작자의 주관적인 판단인 예상정답률만 독립변인으로 가지는 회귀모형 1과 문항 특성 변인만을 독립변인으로 가지는 회귀모형 2가 설명하는 문항 난이도 분산은 각각 78%와 70%로 8% 정도 차이가 나지만, 이 두 모형을 이용한 2007학년도 대수능 9월 모의평가 수리 영역 예측 문항 난이도가 설명할 수 있는 실제 문항 난이도 분산은 각각 69%와 49%로 20%정도 차이가 생겼다. 난이도를 예측할 때 생기는 설명력의 차이가 회귀모형을 추정할 때 생기는 설명력 차이의 2.5배가 된 것은 회귀모형 2가 수험생의 특성을 전혀 반영할 수 없기 때문이었던 것으로 추정된다. 즉, 수험생들이 9월에 고등학교 교육과정을 아직 이수 중이라는 점에서 수능 9월 모의평가에 응시하는 전체 수험생의 특성과 11월 본수능에 응시하는 전체 수험생의 특성은 차이가 있다고 볼 수 있다. 그런데 문항 특성 변인에만 근거하여 난이도를 추정하는 회귀모형 2는 2005 · 2006 · 2007학년도 본수능 자료에 근거하여 추정되었을 뿐만 아니라, 수험생의 특성을 반영하는 독립변인이 전혀 포함되어 있지 않다. 그러나 예상정답률에 근거하여 문항 난이도를 추정하는 회귀모형 1도 역시 본수능 자료에 근거를 두고 있지만, 문항 제작자가 수능 모의평가 문항의 정답률을 예상할 때는 수험생들이 교육과정을 이수 중이라는 점을 고려한다고 볼 수 있다. 따라서 이것은 회귀모형 1의 독립변인인 예상정답률에는 수험생의 특성이 이미 어느 정도 반영되어 있다는 것을 의미한다. 그러므로 회귀모형 1이 수능 9월 모의평가의 문항 난이도를 좀 더 정확하게 추정할 수 있었던 것으로 보인다. 그러나 문항 특성 변인만을 고려하여 문항의 난이도를 추정해야 하는 경우 피험자 특성의 차이를 고려하지 않기 때문에 회귀모형을 추정하고 그 모형으로 문항 난이도를 예측하는 표본의 특성이 유사할수록 바람직하다. 구체적으로, 대수능 문항의 난이도를 예측하려면 대수능 자료에 근거하여 회귀모형을 추정하고, 대수능 9월 모의평가의 문항 난이도를 예측할 때는 대수능 9월 모의평가 자료에 근거하여 회귀모형을 추정하여 사용함으로써 보다 정확한 문항 난이도 예측치를 얻을 수 있을 것으로 보인다.

예상정답률만 독립변인으로 포함하고 있는 회귀모형 1에 의해 예측된 문항 난이도가 실제문항 난이도를 설명하는 정도는 어떤 회귀모형(일차 직선 또는 이차 곡선)을 적용하느냐에 따라 약 7%의 차이가 났다. 즉, 문항 난이도를 예측하는 회귀모형의 독립변인이 예상정답률 하나일 때, 예측 난이도와 실제 난이도 사이의 관계는 일차식보다는 이차식으로 나타났다. 이것은 예상정답률과 문항의 개념 수를 독립변인으로 가지는 회귀모형 4에서도 마찬가지로 나타났다. 즉, 이차 곡선 모형이 직선 모형보다 실제 문항 난이도 분산의 약 6%를 더 설명할 수 있는 것으로 나타났다.

VI. 나오며

2005 · 2006 · 2007학년도 대수능 수리 영역 문항 특성 변인들을 조작적으로 정의하고, 그 문항 특성 변인들이 문항 난이도와 1차 선형 관계가 되도록 척도화하여 추정한 회귀모형은 경험과 직관을 토대로 하여 문항 제작자가 판단한 예상정답률로 추정된 회귀모형이 설명하는 문항 난이도 분산의 90%를 설명할 수 있는 것으로 나타났다. 문항 제작자가 예상정답률을 판단할 때 문항 특성과 학생 특성을 모두 나름대로 고려하였다고 가정할 때, 본 연구에서 예상정답률로 설명되지만 문항 특성 변인들로 설명되지 않는 문항 난이도 분산을 추가적으로 설명할

수 있는 문항 특성들을 찾아내어 척도화한다면, 문항 특성 변인에 근거하여 추정된 회귀모형의 문항 난이도 예측치의 정확도는 더욱 높아질 수 있다고 본다. 특히, 이론적으로 문항의 생소성과 관련된 문항 특성 변인들을 찾아내고 조작적으로 정의하는 것이 필요하며, 현실적으로 문항 제작자의 예상정답률을 판단하는 과정에 대한 질적 분석을 통하여 유의미한 문항 특성 변인들을 찾는 것도 필요할 것이다. 또한 학생들의 문제해결 과정에 대한 철저한 분석과 함께 교과 내용의 위계 또는 개념들 사이의 관계를 분석하는 것이 필요하다.

추정된 회귀모형들이 문항 난이도를 어느 정도 예측할 수 있는지를 비교한 결과는 문항 제작자가 학생의 특성들을 어느 정도 고려할 수 있는 회귀모형 1 (예상정답률에 근거한 문항 난이도 추정 모형)이 가장 양호하였지만, 문항 제작자의 경험과 주관적인 판단에 전적으로 의존하기 때문에 문항 난이도 추정의 일관성이 떨어지는 것으로 나타났다. 문항 특성만을 이용한 난이도 추정 모형은 실제 난이도 분산의 49%정도를 설명할 수 있었다.

문항 특성 정보를 이용하여 문항의 난이도를 예측하고자 하는 시도는 우리나라 대수능처럼 시험의 결과가 개인에게 미치는 영향이 절대적이면서 수험생의 적절한 점수 분포가 필요한 경우에 반드시 필요하다. 예비검사를 통하여 사전에 문항의 심리측정학적 정보를 추정하여 본시험에 사용할 수 있는 문제은행 체제가 구축된다면 다른 정보를 이용하여 문항 난이도 추정치의 정확도를 높이기 위해 고민할 필요는 없을 것이다. 그러나 대수능은 사회적으로 매우 민감한 고부담 시험으로 기출 문항을 철저히 배제하고 있기 때문에 예비검사로 노출되었던 문항을 다시 사용하는 것은 현실적으로 불가능하다. 따라서 예비검사를 실시할 수 없는 대수능과 같은 고부담 시험의 난이도 및 점수분포에 일관성을 유지하기 위해서는 문항특성 변인을 이용하여 문항 난이도를 보다 정확하게 예측할 수 있는 회귀모형을 찾으려는 노력이 따라야 한다.

참 고 문 헌

박문환 (2004). 대학수학능력시험 난이도 관련 변인 탐색. 대한수학교육학회지 수학교육학연구, **14(1).** pp.71-88.

박선화 · 박문환 · 이봉주 (2004). 대학수학능력시험 출제 매뉴얼-수리 영역- . 한국교육과정평가원.

이명준 · 이봉주 · 이대현 (2005). 2005학년도 대학수학능력시험 결과 분석 연구 -수리-. 한국교육과정평가원 (대수능 CAT2005-8-2).

이양락 · 이봉주 · 손홍찬 (2007). 2007학년도 대학수학능력시험 결과 분석 연구 -수리 영역-. 한국교육과정평가원(대수능 CAT2007-8-2).

이양락 · 이봉주 · 조윤동 · 변희현 · 손홍찬 · 민경석 · 시기자 (2006). 2006학년도 대학수학능력시험 결과 분석 연구 -수리 영역-. 한국교육과정평가원(대수능 CAT200 6-8-2).

이종승 · 김성훈 · 김재철 · 송현정 · 박문환 · 장경숙 · 서재영 (2003). 문항 난이도 추정 모형 개발 연구-대학수학능력시험의 언어, 수리, 외국어(영어) 영역을 중심으로-. 한국교육과정평가원(연구보고 RRE 2003-14).

Davidson, J. E.; Deuser, R. & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 207-226). Cambridge, MA: MIT Press.

Davidson, J. E. & Sternberg, R. J. (1998). Start problem solving: How metacognition helps. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp.47-68). Manwah, NJ: Lawrence Erlbaum Associates.

Gredler, M. E. (1997). *Learning and Instruction: Theory into Practice* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Luppescu, S. (1996). *Virtual equating: an approach to reading test equating by concept matching of items.* Unpublished doctoral dissertation, The University of Chicago, 1996.

Pressley, M. & McCormick, C. B. (1995). *Cognition, Teaching, & Assessment.* New York, NY: Harper Collins.

Silver, E. A. (1982). Knowledge organization and mathematical problem solving. In F. K. Lester & Garofalo (Eds.), *Mathematical problem solving: Issues in research* (pp. 15-25). The Flanklin Institute Press.

Estimating the regression equations for predicting item difficulty of mathematics in the College Scholastic Ability Test

Sang Ha Lee
Korea Institute of Curriculum & Evaluation
E-mail : leesangha@gmail.com

Bong Ju Lee
Korea Institute of Curriculum & Evaluation
E-mail : yibongju@kice.re.kr

Hong Chan Son
Korea Institute of Curriculum & Evaluation
E-mail : hcson@kice.re.kr

The purpose of this study is to identify the item characteristics that are supposed to affect item difficulty and to estimate the regression equations for predicting item difficulty of mathematics in the College Scholastic Ability Test(CSAT). We selected six variables related to item characteristics based on learning theories: contents, cognitive domain, novelty, item type, number of concepts, and the amount of computation. With data of the CSAT mathematics test administered in 2004-2006, item difficulty was regressed on the six variables, the location of an item, and the item writer's judgment on difficulty. The novelty of an item was found to be a statistically insignificant variable in explaining item difficulty. Four regression equations with different sets of independent variables could explain 70%~80% of the item difficulty variance and were validated as predicting item difficulty of the mock CSAT in 2006.

* ZDM Classification : D64
* 2000 Mathematics Subject Classification : 97D60
* Key Words: College Scholastic Ability Test, college entrance exam, mathematics, item difficulty, regression equation