대한수학교육학회지 수학교육학연구 제 20 권 제 3 호
Journal of Educational Research in Mathematics
Vol. 20, No. 3, 339-356. Aug 2010

초등수학영재들의 통계적 사고 특성 사례 분석: 변이성에 대한 인식을 중심으로

이 형 숙* · 이 경 화** · 김 지 원***

통계교육에서는 자료에 대한 경험을 바탕으로 한 사고의 발달이 중요하다. 특히, 자료 생성 시 발생하는 변이성에 대한 이해는 통계적 사고의 핵심이므로 변이를 고려한 학습기회를 제공할 필요가 있다. 국내외 관련 연구자들은 수학적으로 우수한 능력을 지닌 학생들이라 할지라도 통계적 사고 수준은 매우 낮으므로 적극적인 교육을 통해 이를 극복해야 한다고 주장하고 있다. 본 논문은 15명의 우리나라 수학영재아들이 자료를 통한 통계의 주요 개념들을 이해하는 다양한 방식들을 살펴보면서 그 중 모둠 활동에 참여한 세 명의 학생들이 자료와 그래프를 생성하는 과정에서 보여주는 서로 다른 통계적 사고 과정을 좀 더 세밀히 비교분석하는 것을 목표로 한다. 연구 결과, 수학적으로 매우 우수한 성취를 보이는 학생들임에도 불구하고, 선행연구에서 제시한 일반 초등학생들의 변이성에 대한 이해 양상과 별다른 차이를 보이지 않았다. 이로부터 우리나라의 초등학교 통계교육이 변이성 인식에 도움을 주지 못하고 있다는 시사점을 얻었다.

Ⅰ. 들어가는 말

통계에 관한 우리나라 초등학교 수학과 개정 교육과정에 따르면 1학년에서는 간단한 자료를 한 가지 기준에 따라 분류하고, 2~6학년에 걸쳐 표와 다양한 종류의 그래프를 학습하도록 하면서 5학년에서는 평균을 도입하고 줄기와 잎 그림 그래프를 특징적으로 다룬다. 이를 통해 통계분야는 목적에 맞게 자료를 수집하고 정리하여 적절한 그래프로 나타내고, 자료의 특성을 설명하거나 해석할 수 있도록 한다. 그리고 확률분야는 6학년에서 간단한 경우의 수 사이의 비율을 구하여 확률값을 계산하는 고전적 관점으로 간단히 취급하고 있다 (교육과학기술부, 2007). 현재 우리나라 6학년 교과서에서 비록 동전을 던져보는 활동을 소개하고는 있지만 동전을 던져서 얻은 결과가 변이성[1]을 가진 것으로 얼마든지 다양한 형태를 취할 수 있음을 본격적으로 다루지는 않는다. 이는 그 이전에 다루었던 자료의 정리와 표현에 관련된 통계적 지식과 별로 연결되지 않는 방식으로 확률 개념을 도입하는 것이며, 다시 중학교 통계의 내용과 연결되는 고리 역할도 하지 못한다. 결국 우리나라의 초등학교와 중학교에서 다루고 있는 통계와 확률은 적절한 연계성을 보이지 못하고 있음을 알 수 있다. 특히 통계적 사고를 필요로 하는 활동에서도 실제 상황

* Eastern Washington University, hslee@ewu.edu
** 서울대학교, khmath@snu.ac.kr
*** 아주대부설 과학영재교육원 강사, babybear@paran.com

1) 본 논문에서 변이성(variability)은 자료의 속성을 말하며, 변이(variation)는 자료의 속성을 정량화하는 사고 활동을 의미한다.

과 자료를 바탕으로 개인의 경험에 따른 빈도적 관점을 적절히 다루기보다, 형식적이고 피상적인 수준에서 도입한 후 주로 고전적 관점에 따른 확률 계산 문제만 반복하여 연습하도록 하고 있어, 이런 방식의 지도관점은 개선되어야 한다는 점이 수차례 지적되어 왔다 (신보미, 2007; 우정호, 2000; 이경화, 1996; 이영하, 권세림, 2009).

통계와 확률이 부드럽게 연결되도록 다루려면 빈도적 관점에서의 확률 개념에 내포된 일련의 과정, 즉, (1) 확률 실험을 하거나 관찰에 의해 자료를 생성하고, (2) 그 자료에 기초하여 확률을 구한 후, (3) 실험 상황 또는 관찰 맥락에 되돌려 구한 값의 의미를 파악하는 것이 필요하다. 이는 자료에 기초한 확률을 통해 앞으로의 실험 결과 또는 관찰 결과를 예측하는 활동과 관련되며, 결국 자료의 변이성을 파악하여 의사결정의 도구로 활용하는 통계적 사고 활동과 관련된다.

이 논문에서는 초등수학 영재아들 중 이와 같은 통계와 확률 사이의 연결성을 파악하는 경우와 파악하지 못하는 경우를 구분해보고, 각각의 경우에 어떤 내면적 수학 활동 특성과 언어적 표현 특성을 보이는지, 이것이 이들의 사고 특성에 대해 시사하는 바가 무엇인지 살펴보는 데 목표를 둔다. 특히, 자료와 그래프를 생성하도록 하였을 때, 생성한 자료에 대한 상반된 시각이 그래프 생성 활동에 어떤 방식으로 영향을 미치는지 살펴봄으로써, 기존에 개별적으로 형성된 학생들의 통계적 사고 특성이 새로운 문제 상황에서의 논의를 통해 어떻게 발현되고 조정되어가는 지에 초점을 둔다.

II. 문헌분석

확률과 통계의 연결은 학생들, 특히 어린 연령에 있는 학생들을 교육하는데 있어 중요하다 (Shaughnessy, 2003). 확률은 어떤 일이 일어날 '기회' 또는 앞으로 일어날 일의 '예측'이라는 일상적인 용어를 통해 학생들에게 직관을 통한 접근을 허용하는 반면, 통계는 자료의 정리와 표현이라는 주제 하에 형식화된 통계기법과 절차를 강조하는 형태로 교육되고 있다 (Pfannkuch & Wild, 2004). 전통적인 통계교육에서는 평균과 같은 중심경향치에 관한 탐구에 초점을 맞추었으며 자료가 내포하고 있는 변이성이나 분포에 관한 탐구는 상대적으로 간과되어왔다 (Jacobs, 1997; Shaughnessy, 2006). 이는 전통적인 통계교육에서 이론적인 확률 관점을 통계적 지식을 습득하기 위한 필수지식으로 간주한 것과 무관하지 않다. 이론적 확률, 곧 고전적 확률 관점은 자료에 근거하지 않기 때문에 변이성에 대한 고려를 배제한다. 중심경향을 표현하기 위한 척도들 또한 하나의 값으로 표현되기 때문에 자료에 기반하여 구한 값임에도 불구하고 변이성과 거리가 먼 것처럼 생각하기 쉽다. 이로 인해 중심 경향치에 대한 이해 역시 어려워졌다 (Rubin, Brice & Tenny, 1991; Shaugnessy, 2006). 특히, Mokros & Russell (1995)은 학생들이 평균을 각각의 자료에 기초하여 얻은 값이 아니라 별도의 값인 것처럼 생각함으로써 그 대표성을 파악하지 못하였음을 확인하였다.

1990년대 초부터 시작된 통계교육의 개혁은 전통적인 통계교육의 내용을 바꾸어야 한다는 것에 초점을 맞추고 있다. Cobb (1992)과 Moore (1997)는 이론적 관점, 즉 연역적인 확률적 사고로부터 출발하는 통계교육방식에서 탈피하여, 정량화 활동을 통해 수집된 자료에 근거하여 문제 상황의 해결을 추구하는 자료 분석 활동 중심의 통계교육을 주장하였다. 이는 통계

기법의 학습과 적용을 강조하는 전통적인 관점에서 벗어나, 자료에 기초한 사고를 통한 개념 형성과 자료의 이해를 바탕으로 한 자료표현과 분석이라는 관점으로의 전환을 의미한다 (Ben Zvi & Garfield, 2004). 이러한 통계교육의 관점의 변화에 따라 우리나라의 초등학교 통계교육의 목표 또한 사실의 단순한 정리와 표현 수단 획득이 아니라, 실험을 통해 생성된 자료나 혹은 실생활 문제 상황에서 획득한 자료의 일부를 체계적이고 효율적으로 정리하고 표현하여 추론하고 의사소통하는 것으로 변화되었다 (교육과학기술부, 2007).

우리가 다루는 자료는 필연적으로 변이성을 내포하고 있으므로 통계교육에서는 변이성을 인식하도록 하는 것이 중요하다 (Moore, 1997; Reading & Shaughnessy, 2004; Shaughnessy, 2006; Shaughnessy & Pfannkuch, 2002; Wild & Pfannkuch, 1999). 문제는 복합적이고 다차원적인 변이성을 학생들이 다룰 수 있는 형태로, 그리고 적절한 활동으로 어떻게 구현하는가이다. 통계교육 연구자들은 이에 대한 다양한 제안을 제시하고 있다. 예를 들어, Reading & Shaughnessy (2004)는 다음 네 단계에 따라 변이성을 인식하고 통제하는 가운데 통계적 개념 학습을 유도할 수 있다고 보았다: (1) 변이성 인식, (2) 예측이나 분석결과의 설명을 위한 변이성의 정량화와 모델화, (3) 정량화된 변이성 즉 변이의 설명, 그리고 (4) 변이와 관련된 탐구방법 개발. 변이성을 인식하는 가운데 문제해결을 위해 수집한 자료를 어떻게 정리하고, 표현하고 해석하겠는가라는 물음을 지속적으로 제기하고 해결하는 것이 필요하다. 자료를 그래프로 나타내고, 그래프에 포함된 자료를 통해 추론하는 것은 변이성을 인식하여 적절하게 다루는 학습 기회를 제공한다 (Shaughnessy & Pfannkuch, 2002). 그럼에도 불구하고 우리나라의 통계교육은 그래프 그리기 활동을 통계적 사고발달을 위한 자료탐구나 사고 방법을 표현하고 개발하기 위한 기회로 이용하기 보다는, 다양한 형태의 그래프를 각각의 개별적인 정형화된 틀로만 다루고 있다 (이경화, 지은정, 2008). 본 연구는 현행 교과서에서 간과하고 있는 그래프의 이해와 해석을 통한 변이성 인식 과정의 세부적인 특징을 살펴보고자 한다.

반복실험을 통해 앞으로 일어날 일을 예측하는 활동은 통계와 확률을 연결시키는 좋은 기회를 제공한다 (Canada, 2006, 2008; Reading & Shaughnessy, 2000, 2004; Shaughnessy & Ciancetta, 2002; Torok & Watson, 2000; Zawojewski & Shaughnessy, 2000). 반복실험 결과를 해석하는 과정에서 변이성을 인식하고 이를 적절히 예측에 반영한다면 이미 학습한 이론적인 관점으로서의 확률 개념이 자료에 기반한 빈도적 관점으로서의 확률 개념으로 전환될 것이다. 이 과정에서 통계와 확률이 개념적으로 연결될 수 있다. 반복실험 결과는 서로 다른 세 수준의 변이성을 토대로 해석될 수 있다: (1) 개별 결과와 그 빈도수에 초점을 맞춘 덧셈적인 사고, (2) 중심 경향치를 고려한 비례적인 사고, 그리고 (3) 비례적인 사고와 더불어 분포에 주목하는 사고 (Shaughnessy, Ciancetta, & Canada, 2004). 그 동안 학교수학에서는 변이를 고려하면서 예측하는 활동을 적절히 제공하지 못하였다. 다시 말하여, 확률이나 평균을 구하여 미래에 대해 예측할 수 있다고 보았지만, 주어진 자료가 가진 본질적인 측면이라고 볼 수 있는 변이성을 고려하지는 않았다. 그러나 1990년 발표된 Moore의 논문을 계기로 변이성에 기초한 예측이라는 개념에 주목하게 되었다 (Watson, Callingham, & Kelly, 2007). 통계적 사고의 본질이라고 할 수 있는 변이성의 인식과 이에 기초한 예측은 낱낱의 자료값 또는 특정

한 하나의 자료값이 아니라 구간에서의 자료값에 주목하여 이루어지며, 그 구간을 구체화하는 방법을 발전시키면서 통계적 사고가 발달한다고 할 수 있다. 특히, 중심 경향치와 변이성을 직관적으로 관련지어 이해하면서 자료를 다루도록 하고, 이 두 개념 요소 사이의 통합에 의해 자료 집단으로서의 분포 개념을 형성하도록 하는 것이 중요하다(Shaughnessy, 2006).

III. 연구방법

1. 연구 대상

본 실험에 참여한 학생들은 우리나라 영재교육진흥법과 그 시행령에 따라 정부의 재원으로 운영되는 대학부설 과학영재교육원의 수학분야 영재로 선발된 4~5학년의 학년통합(기초)반에 속한 15명이다. 이들은 전년도 가을에 자신이 속한 학교장의 추천(1차)을 받고 학년별 학력고사(2차)를 거쳐 이들을 직접 교육할 분야의 지도교수와 지도교사로 구성된 선정추천 심사위원회로부터 선발하고자 하는 인원의 5배수에 선정된 후 학년 통합의 사고력 검사(3차)와 구술고사(4차)를 거쳐 최종 선발되었다. 이들은 1년간 120시간의 교육을 이수하면 이듬해 심화반으로 편성된다. 이들은 3월부터 6월까지 매월 6시간 (2시간 사전 원격교육, 4시간 주말 집중교육)의 수업을 받았으며 연구 당시 3박 4일간의 여름방학 집중교육 중 2일차 수업에 참가중이었다. 학생들은 본 연구에 참가하기 이전에 과학영재교육원으로부터 확률과 통계에 관한 사전 수업을 받은 경험이 전혀 없었지만 수년간 영재아들을 지도한 경험이 있는 이들의 담임강사는 이 학생들의 사전 학습 정도나 사고력 수준이 또래 학령보다 1~2년 정도 앞선다고 평가하였다.

본 논문은 수업을 진행한 연구팀의 일원(제1저자)이 두드러진 반응을 보인 것으로 판단하여 수업 중에 임의로 선정한 그룹에 속하는 세 명의 학생에 초점을 맞추었다. 그들에 대한 개인적 특성과 담임강사의 평가는 아래 <표 III-1>과 같다. 학급내 수준의 점수는 지도교사가 학생활동 평가표에 A, B, C, D, E로 평가한 평점을 5, 4, 3, 2, 1점으로 환산한 점수이며 평가1은 수업당시(기초반)의 담임강사가 학년말에 종합적으로 서술한 내용이며, 평가2는 실험이후 현재(심화반)의 담임강사가 학생에 대해

<표 III-1> 연구의 대상자

학생 ID	학년	성별	학급내 수준	담임강사의 평가
학생1	5	남	상 (4.67)	평가1. 주어진 과제가 명확하게 이해 되었을 때 문제 해결력이 뛰어남. 주어진 과제를 추론하여 논리적으로 사고하는 능력이 우수함. 친구들과 어울리기를 좋아하고 모둠 활동에서 자신의 의견을 자신 있게 표현함. 학기 초의 모습보다 학기 후에 더욱 발전한 모습을 볼 수 있었음. 평가2. 머리 회전력이 빠르고 도형과 수리에 모두 강함.
학생2	4	남	중 (4.3)	평가1. 4학년이어서 수학의 내용에 대한 이해는 깊지 않으나 수업에 적극적으로 참여하고 새로운 사실을 알아내기 위해 열심히 노력함. 교사의 발문에 대해서는 자신감이 부족하지만 자신이 알아낸 사실에 대해서는 자신 있게 발표함. 문제 해결을 위한 다양한 경험이 더욱 요구됨. 평가2. 조용하고 차분하며 사고에 깊이가 있음.
학생3	5	남	중 (4.4)	평가1. 다른 친구들에 비해 속진의 정도가 높지 않지만 문제 해결을 위한 좋은 아이디어를 자주 제시함. 정형화된 문제에 대해서는 일반적인 행동을 하지만 비정형화된 문제에 대해서는 뛰어난 능력을 발휘하는 경우가 있음. 자신이 흥미를 가진 부분에 대해서는 높은 성취 욕구를 나타내지만 그렇지 못한 부분에서는 자신감이 떨어짐. 모둠 활동 속에서 친구들과 의사소통 과정에서 문제가 생길 시 감정적인 반응을 나타내는 경우가 있음. 평가2. 자기표현력은 강하나 산만하고 의욕이 떨어짐

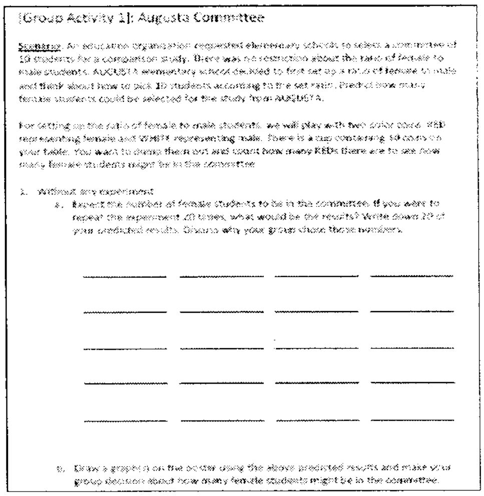
간략히 평가한 내용이다.

2. 학습 과제

본 실험을 위한 수업은 해당 과학영재교육기관에서 실시하는 방학 중 집중교육의 한 가지 특색 사업으로 영어로 진행하는 수학 수업의 일환으로 이루어졌기에 과제도 영문으로 제시되었다. “What’s next? Expecting the unexpected!”[2]라는 제목으로 수업을 하는 동안 [그림 Ⅲ-1]의 활동지를 제공하였다. 수업의 목표는 자료에 대한 변이성 인식을 통한 통계적 사고의 발달이었다. 구체적인 학습활동으로는 변이를 고려할 수 있는 문제 상황으로 (1) 동전을 이용한 확률 실험과 생산된 자료값에 근거한 확률의 의미 생각해보기, (2) 반복시행으로부터 생산된 평균의 의미 생각해보기, (3) 반복 실험의 결과값들을 예측해보고 그 예측값들을 생산하게 된 이유 생각해보기, 그리고 (4) 예측을 위한 그래프 생성하기 등의 활동을 제시하였다.

3. 자료 수집

수업은 15명의 학생들을 무작위로 3명씩 5모둠으로 편성하여 저자 중 한 명(제1저자)이 4시간에 걸쳐 수업을 진행하였고 이 중 한 모둠에 속한 3명의 학생을 집중 관찰한 자료를 분석하였다. 수업 중 토론은 영어와 한국어가 모두 허용되었으나 영어에 능숙하지 않은 학생들로 인해 대부분의 토론은 한국어로 이루어졌다. 영어로 제시된 문제의 이해에 어려움이 있는 학생들을 위해 교사가 기본적으로 해석을 해주었고 영어와 한국어에 모두 능숙한 2명의 도우미 한인 중학생이 영어에 어려움이 있는 학생들을 개별적으로 도왔다. 수업은 2대의 비디오카메라로 녹화하였는데 한 대는 교사의 활동과 발문을 다른 한 대는 모둠 활동을 촬영하였다. 각 모둠에는 디지털 음성녹음기를 두어 한 대의 카메라가 이동하는 동안 벌어지는 모둠별 토론을 빠짐없이 녹음할 수 있었다. 학생들에게 배부하였던 교재와 수업 중 제작된 그래프들도 모두 수거하여 수업 분석에 참고하였다.


[Group Activity 1]: Augusta Committee

Scenario: An education organization requested elementary schools to select a committee of 10 students for a comparison study. There was no restriction about the ratio of female to male students. AUGUSTA elementary school decided to first set up a ratio of female to male and think about how to pick 10 students according to the set ratio. Predict how many female students could be selected for the study from AUGUSTA.

For setting up the ratio of female to male students, we will play with two color coins: RED representing female and WHITE representing male. There is a cup containing 10 coins on your table. You want to dump them out and count how many REDs there are to see how many female students might be in the committee.

1. Without any experiment
   a. Expect the number of female students to be in the committee. If you were to repeat the experiment 20 times, what would be the results? Write down 20 of your predicted results. Discuss why your group chose those numbers.

   b. Draw a graph on the poster using the above predicted results and make your group decision about how many female students might be in the committee.

[그림 Ⅲ-1] 연구의 과제(수업 활동지)

Ⅳ. 자료 분석 및 결과

1. 학습 활동 분석 관점

수업을 통해 수집한 자료를 분석하기 위해 수업 중 토론 과정에서 떠오른 주요 주제들을 바탕으로 아래의 3가지 질문을 만들고 이를 학습 활동의 분석 기준으로 설정하였다.

기준1: 반복시행에 의한 확률문맥에서 변이성에 대한 이해는 평균에 대한 이해

2) 수업계획 단계에서 도움을 준 Dan Canada에 감사의 말을 전한다.

와 어떻게 관련되는가?

기준2: 학생들이 독립시행을 균등분포로 이해하는 경향이 있는데 변이성에 대한 이해는 독립시행에 대한 이해와 어떻게 관련되는가?

기준3: 반복시행에 기초하여 예측할 때, 변이성을 고려하는 경우와 그렇지 않은 경우, 그 차이점은 무엇인가?

2. 수업 자료의 분석 결과

분석 1: 반복시행에 의한 확률문맥에서 평균과 변이성의 이해

반복시행 문맥에서 확률의 의미를 설명해보는 활동이 주어졌을 때, 자료에 기초하여 확률을 이해하는 네 가지의 다른 사고방식이 관찰되었다. 그 중 평균을 통해 확률을 이해하는 시각이 주목받았다. 우리는 학생들이 어떻게 이미 형성해 놓은 확률 개념과 그것의 통계적 측면을 관련시키는지를 관찰하였다.

교사(연구자)는 수업을 위해 준비한 문제들에 본격적으로 들어가기에 앞서 학생들에게 확률에 대해 무엇을 알고 있는가라는 물음으로 그들의 사전지식을 확인해 보았다. 학생들은 확률을 "전체 경우의 수 분에 특정 경우의 수" 또는 "어떤 일이 일어날 수 있는 가능성"이라고 정의하였다. 이에 대해 교사는 구체적으로 양면이 각각 빨간색과 하얀색으로 이루어진 동전을 보여주고 그 동전을 한 번 던졌을 때 빨간색이 나올 확률이 무엇이냐고 물었다. 모든 학생들은 주저하지 않고 곧바로 1/2 이라하였고, 교사는 그 의미를 설명하도록 하였다. 학생들이 발표한 의견은 다음의 4가지였다.

의견 1: "빨간색이 나올 경우의 수랑 하얀색이 나올 경우의 수가 같다."

의견 2: "두 번 중의 한번, 네 번 중의 두 번은 빨간색이 나온다."

의견 3: "두 번 던졌을 때 두 번 다 빨간색이 나올 수도 있지만, 그런 건 확실하지 않다. 많이 던지면 빨간색이 나오는 개수가 반에 가까워진다."

의견 4: "두 번 중에 한번이 나온다는 것은 평균인 것 같다."

학생들은 "확률이 $\frac{1}{2}$이다"라는 의미를 다음과 같은 네 가지의 서로 다른 논리로 이해하고 있었다. : 확률에 대한 원리(의견 1, Beckmann, 2005), 비율 개념(의견 2), 큰수의 법칙에 대한 직관(의견 3), 그리고 평균 개념(의견 4). 의견 4를 표현한 학생에게 교사가 '평균'이라는 말을 좀 더 자세히 설명해 달라고 했을 때 그는 다음과 같이 말했다: "어떨 때는 하얀색이 두 번 나올 수도 있고 어떨 때는 빨간색이 두 번 나올 수도 있고, 어떨 때는 하얀색 한 번 빨간색 한 번, 그래도 평균적으로 계산하면 둘 중의 하나가 나온다는 뜻입니다."

이 학생은 다른 학생들과는 달리 각각의 개별적인 시행의 결과보다는 자료 전체를 표현할 수 있는 평균이라는 용어로 확률을 설명하였다. 그럼에도 불구하고 우리는 이 학생이 평균을 어떤 특정 시행의 결과로 생각하는지 반복되는 시행에서 발생될 변이를 고려한 사고인지 좀 더 지켜볼 필요가 있었다. 다음은 학급 전체 학생들과 교사 사이의 평균에 관한 대화이다. (교사는 실제 수업에서 경어와 평어를 적절히 섞어서 사용하였지만 이하 각 에피소드에서는 학생들의 말과 혼동되지 않도록 평어로 통일하여 나타낸다.)

[에피소드1] 데이타의 개수와 평균

(확률을 평균이라는 용어를 통해 설명하는 대답이 어떤 학생으로부터 나오고, 그에 덧붙여 '여러 번의 시행을 통해 제일 근접한 값을 평균이라 한다.'는 또 다른 의견이 나왔다.)

교 사: 그것(평균)은 변해?
학생들: 변해요.
교 사: 변해? 그런데 데이터(를) 많이 걷을수록 더 정확한 평균을 구할 수 있다고 생각해?
학생들: 네.
교 사: 데이터 만 개와 백 개를 걷는 것 중에 어떤 것으로부터 구한 평균이 좀 더 정확할까? 아니면 상관없을까?
학생들: 만 개요.
교 사: (만 개의 경우가) 좀 더 정확할거 같아?
학생들: 네.
학생A: 사실 상관은 없지만, 그것에 대한 정보를 어느 정도는 알 수 있으니까.
교 사: 더 많은 걸 평균내면 더 정확해?
학생들: 네.
학생B: 백 개(번) 던졌을 때하고 만 개(번) 던졌을 때 하고 나올 수 있는 경우의 가지 수가 다르잖아요. 그러니까 그것에 의존하면 훨씬 더 잘 할 것 같아요.
교 사: 경우의 수(시행횟수)가 많아지면 더 정확하다는 말인가?
(여기저기서 '더 자세하게', '더 세밀하게' 라고 말한다)
교 사: 자세하게가 무슨 의미지?
학생C: 경우의 수 중에 가장 잘 나올 수 있는 (잘 들리지 않는다).
학생B: 비율은 어디서나 같잖아요.
교 사: 무슨 비율?
학생B: 만약 4000개 6000개라고 하면 4대 6이니까, 그러니까 경우의 수가 많은데(로부터) 비율이 나오면 훨씬 정확하니까.

앞의 의견 4에 대해 구체적인 생각을 제시한 학생의 평균을 통한 확률의 이해는 시행결과에 대한 도수들의 합과 나누기에 의한 계산에 바탕하고 있는 듯이 보이는 반면, 위의 [에피소드 1]은 확률의 개념이 평균이라는 통계적인 개념을 통해 학생들 사이에서 어떻게 구체화되는지 보여주고 있다. 확률을 "제일 근접한 값"이라는 연역적인 관점에서, 혹은 "나올 수 있는 경우의 가지 수가 다르다"며 변이성을 고려하면서, 또는 "가장 잘 나올 수 있는"이라 말하며 직관과 변이성을 동시에 고려하면서 이해하고 있었다. 특히, "경우의 가지 수"라는 도수의 변화를 고려한 발언을 한 학생B의 비율을 통한 확률의 이해는 단순한 연역적인 비율에 기초한 확률지식이 아닌 변이성에 대한 이해에 기초한 통계적인 사고 활동으로 여겨진다.

반복시행과 평균이라는 내용을 학생들이 어떻게 확률과 연결시키는지 좀 더 자세히 살펴보기 위해 교사는 학생들과 독립시행의 의미에 대해 논의하였다. 교사는 학생들에게 동전을 한 번 던진 후 그 결과를 알 경우와 모를 경우 각각에 대해, 그 다음 시행의 결과를 예측하라고 질문을 받았을 때 같은 답을 하고 싶은지 다른 답을 하고 싶은지 물었다: "보여주건 안 보여주건, 아무리 확률이 ½이라고 해도 똑 같은 게 두 번 나올 수도 있고, 보여주나 안 보여주나 상관이 없을 것 같아요", "앞에 거랑 뒤에 거랑 상관이 없으니까…." 학생들은 동전을 던지는 시행이 반복될 경우 각 시행은 서로 독립적인 사건을 만든다는 데 의심이 없어 보였다. 이러한 독립시행에 대한 이해를 바탕으로 교사는 "그렇다면 100번 던지나 10번 던지나 다음 것을 예측하는 데는 아무 상관이 없네. 그런데 왜 100번 던지면 [10번 던지는 것보다] 더 정확한 답을 할 수 있다고 생각하지?"라고 질문했다. 이에 대해 학생들은 다음과 같이 네 가지의 다른 의견을 보여주었다:

반응 1: "아무 상관없을 것 같다."
반응 2: "확률이 어차피 1/2인데 아무리 많이 던졌다고 변하는 값은 아니다."
반응 3: "선거같은 경우에는 더 많은 사람들에게 물어보면 진짜 값에 더 가까워진다."
반응 4: "열번 던지면 빨간면이 0번부터 10까지지만 백번 던지면 빨간면이 0번부터

100, 분모가 더 커지니까 더 세밀해 질 것 같다."

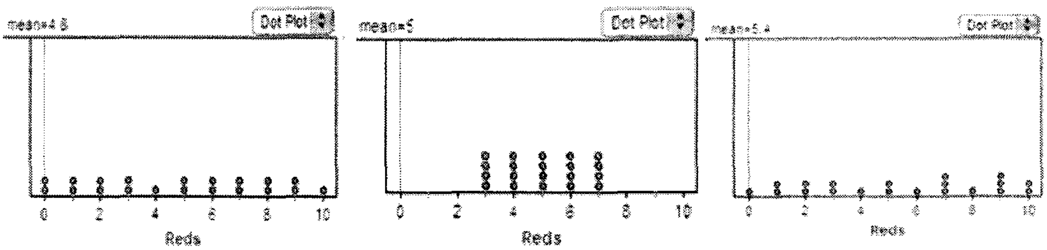
통계적인 사고를 요구하는 상황에서 학생들은 자료와 관련된 확률에 대한 이해 방식을 좀 더 분명히 드러냈다. 반응 1과 반응 2에서 학생들은 확률을 연역적으로 정해진 값으로 이해하고 있는 반면, 반응 3과 반응 4는 자료에 기초하여 확률을 이해하고 있다. 즉, 반응 1과 반응 2는 자료가 도입되는 상황에서 어떠한 변이도 고려하지 않고 있고, 반응 3과 반응 4는 수집된 자료는 전체의 부분이라는 인식과 함께 확률 $\frac{1}{2}$의 의미를 이해하고 있다. 그러나 반응 3과 반응 4는 그 사고 방법에 있어 현저한 차이를 보인다. 반응 3은 반복시행에 대한 이해를 전수집단을 찾는 과정으로 이해함으로써, 시행의 결과 값으로 얻어진 자료가 필연적으로 포함하고 있을 변이성을 절대적인 오차의 의미로 해석하고 있다. 즉, 자료의 크기가 크면 자료의 신뢰도가 더 높아진다는 것을 인식하고 있지만, 반복시행과 연결하여 설명하지 못하고 있다. 반응 4는 자료를 개별적인 값의 단순 집합이 아닌 전체로 봄으로써, 확률을 변이를 고려한 분포적인 사고에 기초한 평균에 의해 이해하고 있다.

예비문제와 질문을 통해 학생들의 확률에 대한 다양한 형태의 이해 방식을 알아본 뒤, 본격적인 수업을 통해 그러한 이해가 통계적 사고를 요구하는 구체적인 문제 상황인 자료를 생성하고 정리하는 활동에서 어떻게 표출되고 발전되는지를 알아보았다.

분석 2: 변이성 고려의 결여로 인한 독립시행의 균등분포적 표현

어떤 시행의 결과값이 그 다음 시행의 결과값에 영향을 주지 않을 때 그 시행들은 서로 독립이라고 말한다. 학생들은 결과값들 사이의 관련 없음을 모든 가능 결과값들이 같은 정도의 가능성을 가지고 나타난다는 의미로 해석하였다. 이러한 방식의 이해는 변이성에 대한 이해 결여와 깊은 관련이 있을 것으로 생각되었다. 즉 변이성에 대한 이해의 부족이 분포에 대한 편향된 시각을 심어주었을지도 모른다고 예상되었다.

학생들에게 양면이 각각 빨간색과 노란색으로 이루어진 동전 10개를 동시에 던졌을 때 나올 빨간 면의 수를 실험없이 생각만으로 20개 생성해 보라고 하였다. 그리고 그 생성된 20개의 값들을 바탕으로, 다시 10개의 동전을 동시에 던졌을 때 나올 빨간면의 개수를 예측해보고 그 예측에 도움이 되도록 생성한 값들을 그래프로 표현해보라고 했다. 각 학생들은 20개의 값들을 생성하였고 한 명을 제외한 14명 학생들의



[그림 Ⅳ-1] 골고루 퍼짐을 강조한 생성 값들의 예

자료값들의 평균은 4.5에서 5.5사이에 분포되어 있었다. 그 한 명의 학생이 만들어낸 20개 값들의 평균은 4였다. 학생들이 특정값을 평균으로 미리 정하고 그것에 맞춰 자료값들을 고안해 내었는지 아닌지는 알 수 없다. 학생들은 모둠별로 하나의 그래프를 그리도록 요청받았고, 누구의 자료를 사용하겠는지, 또 어떤 식으로 20개의 자료를 그래프로 표현하겠는지를 결정하기 위해 토론을 이어갔다.

어떻게 20개의 값들을 생성했는지에 대해 토론할 때 학생들은 각각의 값들이 골고루 퍼져있는지 아닌지에 관심을 보였다. 본 과제에 들어가기에 앞서 질문한 "동전을 10번 던졌을 때 빨간 면이 몇 번 나올 것 같으냐?"는 물음에 많은 수의 학생이 5번, 혹은 5번과 비슷한 정도로 나올 것이라고 응답했다. 소수의 학생 중에 한 명은 "이론적으로는 5번이 나오겠지만 알 수 없다."고 했고 다른 한 명은 "하얀색 면이 나올 확률과 같으므로 예측할 수 없다, 즉 0부터 10까지 골고루 가능하다"라고 답변하였다. 그럼에도 불구하고 20개의 시행 결과값들을 예측하라고 했을 때 많은 수의 학생들이 '골고루'라는 생각에 집중하였다. 이러한 경향은 학생들이 반복시행을 독립시행으로 이해하고 있으며, 독립적이라는 것을 '결과값들이 서로 관련 없음'으로 그리고 '분포에 있어 골고루의 형태로 나타나야 함'으로 이해하고 있음을 보여준다([그림 Ⅳ-1]).

독립시행과 균등분포를 관련짓는 생각은 가능한 결과값들의 중간값을 평균으로 계산되게 하여 이론적 확률과의 일치성을 확인해주지만 (이항분포의 경우), 학생들이 그 평균이 갖고 있는 자료의 대표성을 고려하지 않고 있음을 보여준다. 독립시행의 균등분포를 생각하고 있는 학생들은 그 자료들을 자신들의 목적에 맞도록 표현함에 있어 많은 갈등을 겪었다. 특히, 다음의 분석 3에서 우리는 한 모둠에서 벌어진, 자료에 대한 극히 다른 견해를 갖고 있는 세 학생들 사이의 토론을 통해 다양한 통계적 사고를 세밀히 살펴보고자 한다.

분석 3: 예측 활동에 반영된 변이성과 자료에 대한 서로 다른 인식

같은 모둠에 속한 세 명(학생1, 학생2, 학생3)이 서로 토론하는 과정을 세밀히 검토하였다. 먼저 학생2가 수업 활동지에 있는 20개의 모든 칸에 5를 써 넣었다. 즉 학생2는 10개의 동전을 던졌을 때 빨간 면이 나올 동전의 개수는 늘 5가 될 것이라고 생각한 것이다. 이런 학생2의 답변에 학생1은 불만을 토로하였다.

[에피소드2] 모든 예측값이 5가 되는 확률은 낮아

학생1: 너, 5 5 5 이거 고쳐.
학생2: 왜? 그래도 확률이 가장 높잖아.
학생1: 0부터 10까지의 숫자 아무거나 써, 아무거나 써.
학생2: 아니야, 확률이 높게 해야 되.
교 사: 왜 이렇게 써놨지? (5 5 5… 를 가리키며)
학생2: 가장 확률이 높잖아요.
교 사: 5가 확률이 높다는 거 어떻게 알아?
학생2: 그럼 아무거나 써도 되는 거예요?
교 사: 정말 그렇게 나올 거 같아?
학생1: 야, (학생2를 향해) (20개의 빈칸이 모두 5로 채워진 것을 가리키며) 한 다섯 개만 지워. (학생2가 싫다고 하자) 그럼 반만 지워.
교 사: 그럼 학생1이 한번 얘기해보자. (5 5 5… 를 가리키며) 왜 이것이 비현실적이지?
학생1: 현실적이긴 한데 경우의 수가 많으면, … (교사를 바라보며) 이거 순서 바뀌어도 되는 거예요?
학생2: 딱 정해진 것 아니에요?
학생1, 학생2, 학생3: 바꿀 수 없잖아요.

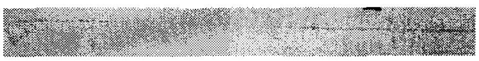
학생1: (학생2에게) 야, 이렇게 하고 싶어, 아니면 고칠거야?
학생2, 학생3: 야, 그래도 확률이 높으면 상관이 없잖아.
학생1: 야, 이렇게 하면 이거 하는 의미가 없잖아. 이걸 만들든지 말든지… 조금만 고쳐, 왜 5만 나와~~~ 계속 5만 나온다는 건 더 확률적으로 낮어.
학생2: 왜?
학생1: 야, 계속 5가 나오는 게 확률이 낮지.
학생2: 얘는 얘얘얘… (5 5 5 5 5 ….를 가리킴).
학생1: 어짜피 똑같잖아, 확률은. 확률은 계속 11분의 1이잖아

학생1은 10개의 동전을 동시에 던졌을 때 나올 수 있는 빨간 면의 수는 0에서 10까지이고 그들은 모두 같은 만큼으로 나올 기회가 있다고 생각했다. 반면, 학생2는 한 개의 동전을 던졌을 때 빨간 면이 나올 확률이 1/2이므로 10개의 동전을 던졌을 때 5개가 빨간 면이 나올 확률이 가장 높다고 생각했다. 예측은 가장 높은 확률에 의함으로 20개의 값으로 모두 5를 택한 것은 합리적인 것이라고 생각했다. 학생1은 10개의 동전을 던졌을 때 다섯 개가 빨간 면이 나온다는 사건에 대한 이견보다는 5가 연달아 나온다는 것에 불편함을 내비쳤다. 그러면서 동시에 학생들은 자신들이 생성한 값들의 순서에 주목하기 시작했다.

학생1은 또한 토론 중간에 실험의 의도를 생각하기 시작했다: "야, 이렇게 하면 이거 하는 의미가 없잖아 (5만 연달아서 써 넣으면 실험을 통해서 10개 동전을 던졌을 때 나오는 빨간 면의 개수를 예측하는 활동을 할 필요가 없잖아)."이러한 발언은 학생1이 비록 가능한 결과값들이 동일한 확률을 가질 것이라는 생각을 하고 있었지만, 시행결과는 필연적으로 변이를 동반할 것이라는 생각을 하고 있음을 보여준다. 이들은 결국 학생1이 생성한 20개의 값들을 그 모둠이 그래프로 표현할 자료로 선택했다. 하지만, 변이성을 인식하고 있는 학생1과 인식하지 못하고 있는 학생2는 자료를 그래프로 표현하는데 있어 아래의 [에피소드 3]에서와 같이 큰 의견차를 드러냈다.

[에피소드3] 도수를 나타내는 "회"와 시행의 순서를 의미하는 "회"

(학생1, 학생2, 학생3은 그래프를 그리기 위해, 우선 가로 세로 축을 그렸다. 그런데 각 축에 무엇을 쓸 것인가로 의견을 달리하고 있었다.)
학생1: 야, 이렇게 되면 여기랑 여기랑 그어야 되잖아, 그런데 너는 뭐를 [하니?]
학생2: 몇 회인지, 1회 2회 3회 4회 (가로축에 20개의 칸을 만들고 맨 오른쪽 칸에서부터 1을 쓰고 왼쪽으로 진행하며 20까지 쓴다.) ([그림 IV-2]).



[그림 IV-2] 학생2가 의미하는 시행순서로서의 "회"

학생1: 뭐가 몇 회?
학생2: 그러니까.
학생1: 아, 이거?
학생2: 20회를 했잖아.
(학생1, 학생2, 학생3이 각각 20개씩 생산한 값들을 모두 사용해도 되냐고 교사에게 묻는다)
학생1: (학생2, 학생3에게) 여기있는 5의 개수를 세! (세 명이 각각 생산한 자료들을 함께 보면서) 5의 개수를 살펴보라고. 1회에 여학생이 뽑히는 경우의 수가 1이잖아.

학생1은 "회"라는 용어를 시행의 각 결과값들의 도수를 표현하기 위해 사용한 반면, 학생2는 "회"를 시행의 순서를 나타내기 위해서 사용했다. 이러한 용어의 사용에 대한 혼란은 토론 내내 이어졌다. [에피소드4]는 이러한 용어

의 혼란이 학생들의 변이에 대한 인식의 차이 뿐만 아니라 자료를 보는 시각과도 관련되어 있을 것이라는 것을 보여준다.

[에피소드 4] 그래프 축들에 있는 숫자들

교 사: (학생들이 세로축에 그린 11개의 칸에 있는 0부터 10까지의 수를 가리키며) 여기에 0부터 10은 뭘 표현해?

학생들: 여자들이요.

교 사: 여자들? 그러면 여기 2에는 어떤 숫자를 대응시키고 싶어?

학생1: 어, 여기는 2명 뽑힌 게 2번이니까, 2까지.

교 사: 그럼 맥시멈 넘버는 뭘까? 여기가 (가로축에 있는 수를 가리키며) 20이면 너무 많을 것 같아? 그럼 몇이 되면 좋을 것 같아?

모두 함께: 5.

교 사: 5? 그럼 5는 뭘 뜻해?

(잠깐 침묵이 흐른다)

교 사: (가로에 있는 5와 세로에 있는 8을 번갈아 가리키며) 그럼 여기가 5다, 여기에 8이 5다, 이게 무슨 뜻이지?

학생1: 8이 5요?

교 사: 2에는 무슨 숫자 대응시키고 싶어?

학생3: 2.

교 사: 왜?

학생1: 2명 뽑힌 것이 2번.

교 사: 그럼.

학생1: 아니, 이 최대숫자 뽑힌 개수가.

(자료들을 어떻게 그래프로 표현할 것인가에 대해 계속 토론한다.)

학생3: 아니 1이 4잖아.

학생1: 1이 2랑 19일 때거든.

모둠의 학생들이 그래프를 그리기위해 택한 자료는 아래의 [그림 Ⅳ-3]과 같다. 학생3의 "1이 4잖아"라는 말은 첫 번째 시행의 결과가 4 라는 의미로 그는 가로축을 시행의 순서로, 세로축을 시행의 결과값들로 생각하며 그래프를 그리려 하고 있다는 것을 알 수 있다. 반면, 학생1의 "1이 2랑 19일 때거든"라는 말은 비록 그래프의 가로축에 1부터 20까지의 칸을 만들어 놓았지만, 학생1은 "1이 두번 나왔다"(사실은 세 번 나왔음)는 시행의 결과값인 1의 도수에 주목하고 있음을 보여준다.

[그림 Ⅳ-3] 학생3의 모둠이 사용하고 있는 자료

[에피소드 5] 자료값도 되고 시행의 순서도 되는 가로축

(학생3은 계속 꺾은선 그래프를 그려야 한다고 주장하고 있다.)

학생3: 2회[째]가 1.

학생1: 이건 꺾은선이 아니라 이상한 그래프야. … 지금 우리가 여기서 20번을 다 던진 거란 말야. 20번을 다 던져가지고 처음에는 ….

(계속 토론이 이어진다.)

학생1: 꺾은선 그래프를 생각한 이유가, 꺾은선 그래프를 그린다 쳐. 만약 0이다. 그러면, 여학생 0명이 뽑힌 그 횟수, 네가 말하는 1회 2회 3회 4회 그게 아니라, 뽑힌 개수.

학생3: 몇 번 반복했다?

학생2: 우리가 그만큼을 던지는 건데?

학생3: 개수를 알 수는 없지.

학생1: 0은 한 번 뽑혔어. 그리고 1은 세 번 뽑혔어, 그러니까 3 찍고, 그럼 되잖아. 세 번 뽑힌다고 …(잘 들리지 않는다)… 하는 건 아니잖아. 이렇게 해야 5가 몇 번 뽑혔는지 알고.

학생3: 몇 번 뽑혔는지 아는 게 아니고 지금,

몇 번 째에 뭘 선택했는지 알아야지.

학생1: 선생님, 여기에요, 5가 뽑힌다고 했으면요, 5가 몇 번 (다른 학생과 말이 엉킴).

교 사: 그건 너희들 맘대로야, 학생3 말대로 가도 되고, 학생1 말대로 가도 되고. 그러면 지금은 한사람 의견을 따르고 조금 있다가 다른 그래프를 그려.

학생3: 근데, 1회 2회 3회 합쳐가지고, 뽑힌 횟수가 아니라, 그 몇 회째에 몇 개를.

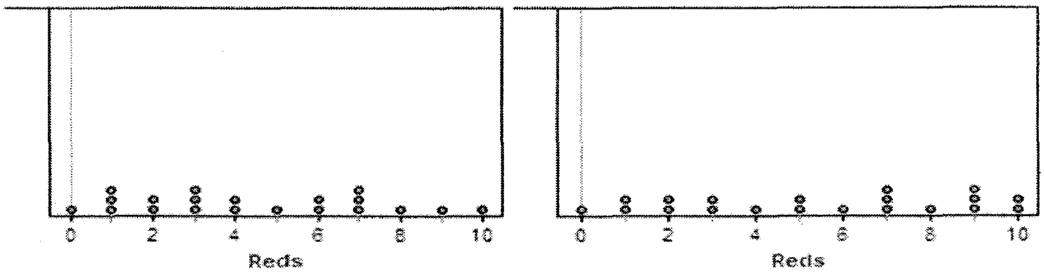
교 사: 학생3의 아이디어는 0부터 20회이니까, 학생3은 1회에 몇이 뽑혔다, 2회는 몇이 뽑혔다, 이것을 말하고 싶은 거지?

학생3: 예.

학생1: 아~~~ 1회 때는 여자 몇이 뽑혔다, 그거예요? 아~~~

학생3의 꺾은선 그래프 그리기에 대한 주장은 그의 자료에 대한 시각을 보여준다. 학생3은 "(10개의 동전 중에 빨간 면이 나타날 동전의) 개수(즉, 커미티 위원으로 뽑힐 여학생의 수)를 알 수는 없지"라고 말하며 현재 사용하고 있는 20개 값들이 시행의 순서에 따라 나타난 독립적인 결과들이라고 생각하고 있었다. 이것은 학생1의 예측을 위해서는 나타난 결과값들에 대한 도수를 알 필요가 있다는 생각과 분명한 차이를 보여준다. 왜냐하면 학생1의 생각은 20개의 자료들이 예측하고자 하는 값에 대한 변이를 품고 있을 거라는 생각을 암시하고 있는 반면, 학생3의 경우는 20개의 자료가 전체로써가 아닌 개별 값들로써 고려되고 있다는 것을 보여주며 이것은 그 학생이 자료값들에 대한 변이성을 고려하고 있지 않다는 증거가 될 수 있기 때문이다. 학생3의 "몇 번 뽑혔는지 아는 게 아니고 지금, 몇 번째에 뭘 선택했는지 알아야지"는 학생3이 왜 꺾은선 그래프에 집착하는지에 대한 또 다른 해석을 가능하게 한다. 그는 제시된 문제에 대해, 균등분포를 갖는 20개의 예측값을 생산했다([그림 Ⅳ-4], 우측그래프). 그리고 수집된 자료를 실제로 존재하는 모집단의 부분으로 여겼다(앞의 [에피소드 1]이후에 나타난 반응 3은 학생3의 의견이다). 이런 식의 자료에 대한 이해에 비추어 봤을 때, 꺾은 선 그래프에 대한 그의 의견은 수집된 자료를 예측을 위해 그 값들이 어떻게 분포되어 있는가에 관심이 갖고 바라보기 보다는 개별적인 시행결과들이 어떻게 하나하나 균등하게 되도록 균형을 잡아가는가에 집중하고 있음을 짐작케 해준다. 학생3의 개별 자료값에 초점을 맞추어 자료를 보는 시각은 그가 시행의 결과값들을 이산변수로써 생각하기 보다는 연속변수로 다루고자 했을지도 모른다는 생각을 하게 한다. 이럴 경우, 예측값을 구하기 위해 학생3이 수집된 자료를 변이를 고려하면서 이용하려 했다고 보기는 어렵다.

한편, 학생1은 학생3과 마찬가지로 반복시행의 결과값들에 대해 균등분포를 생각했음에도



[그림 Ⅳ-4] 학생1(좌)과 학생3(우)이 생성한 20개 자료값들의 분포

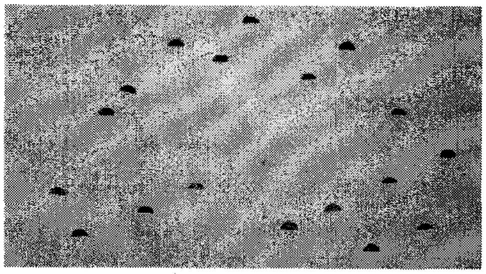
([그림 Ⅳ-4], 좌측그래프) 불구하고 자료생성에 있어서 변이를 고려했다. 그는 20개의 5를 활동지 1a에 적어 넣은 학생2에게 변화를 줄 것을 강력하게 주장했었다. 이러한 변이성에 대한 인식은 그가 빨간 면이 나오는 동전의 개수를 생각함에 있어 변화는 필연적일 것이라 여기고 있음을 뜻한다. 따라서 학생1은 학생2나 학생3이 시행순서에 주목할 때 계속 "회"를 도수로 해석하고 또 그래프를 그릴 때도 시행의 결과값들의 변화를 보기위해 그것들의 도수를 구하려고 했다. 이는 학생1이 개별 결과값에 주목하기 보다는 자료값들 전체를 보고자 했음을 뜻한다. 그러므로 시행의 결과값들은 학생1에게 있어 연속변수가 아닌 이산변수로써 의미를 갖게 된다.

각 모둠에 주어진, 예측을 위해 도움이 되는 그래프를 그리라는 과제에서 학생1과 학생2, 학생3은 많은 의견충돌이 있었음에도 불구하고 [그림 Ⅳ-5]의 그래프를 그렸다. 학생2와 학생3은 시행의 순서를 변수로 하여 각 시행의 결과값들을 그 변수의 값으로 대응시켰다. 이러한 그래프는 case value plot이 stacked dot plot으로 변환되어가는 중간단계의 사고과정을 보여준다 (Cobb, 1999; Konold & Higgins, 2003). Case value plot은 개별시행 (case)에 대한 결과값 (value)을 막대의 높이로 하여 막대그래프를 그린 것이다.

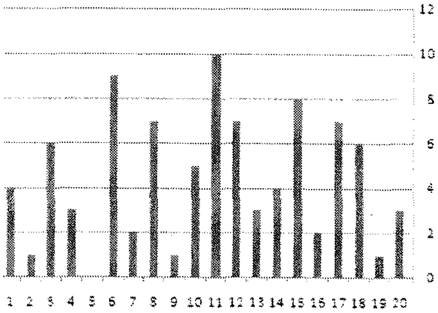
[그림 Ⅳ-5]에 사용된 자료를 사용하여 case value plot을 그리면 [그림 Ⅳ-6]이 얻어진다 ([그림 Ⅳ-5]에서 학생들은 시행순서를 오른쪽에서 왼쪽으로 잡았다). 즉, [그림 Ⅳ-5]는 case value plot에서 막대의 끝이 위치하고 있는 곳을 점들로 표시한 것이다. [그림 Ⅳ-5]에 있는 각 점들을 오른쪽 끝으로 밀어버림으로써 우리는 [그림 Ⅳ-7]을 얻을 수 있다. 이것은 시행의 결과값들의 도수들의 분포를 보여준다. 우리는 [에피소드 3]에서부터 [에피소드 5]를 통해, 학생2와 학생3이 [그림 Ⅳ-5]를 그리려고 하는 동안 학생1은 [그림 Ⅳ-7]을 마음 속에 그리고 있었음을 알 수 있다. 즉, 분석 3에 나타난 대화들은 자료를 보는 두 가지 다른 시각을 가진 학생들이 어떻게 그들의 시각을 그래프에 표현하려고 하는지를 상세하게 보여주고 있다.

한 그래프가 다른 그래프보다 자료의 정보를 표현하는데 있어 절대적으로 우위에 있다고 말할 수는 없다. 그 우위는 우리가 알고자 하는 것을 각 그래프가 얼마만큼 잘 표현하고 있는가에 따라 정해질 것이다 (Konold & Higgins, 2003). 그럼에도 불구하고 어떤 그래프는 다른 것에 비해 그 의미를 해석하기가 더 어려운 것은 사실이다. Konold & Higgins(2003)는 학생들이 stacked dot plot에 있는 점들이 개별시행의 결과값들을 나타낸다는 것을 이해하는데 어려움을 겪는다고 보고하였다.

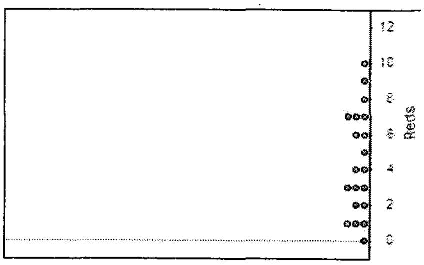
본 연구자들은 [그림 Ⅳ-6]과 [그림 Ⅳ-7]에 나타난 그래프들이 학생들의 자료에 대한 변이성의 이해와 관련이 있다고 생각한다. 학생들은 20개의 자료값들을 바탕으로 다음 번 시행의 값을 예측해보라는 활동에 참여하고 있었다. 다음 시행의 값을 예측하기 위해 도움이 되는 그래프로써 학생2와 학생3은 개별 시행의 결과값들에 초점을 맞춘 그래프를, 학생1은 각 결과값에 대한 도수를 나타낼 수 있는 그래프를 생각했다. 그러한 자료 표현에 대한 다른 접근방식은 그들의 자료를 바라보는 시각의 영향으로 여겨진다. 본 연구 진행과정에서 학생2와 학생3은 자료에 대한 변이성을 인식하지 못한 반면, 학생1은 각 자료값들이 포함하고 있을 필연적인 변화가능성을 인식하며 자료를 만들 때부터 그 변이성을 자료에 반영하려 하였다.



[그림 Ⅳ-5] 모둠의 학생들이 그린 그래프



[그림 Ⅳ-6] [그림 Ⅳ-5]를 case-value plot으로 변환하여 나타낸 모습



[그림 Ⅳ-7] [그림 Ⅳ-5]를 stacked dot plot으로 변환하여 나타낸 모습

Ⅴ. 논의 및 결론

본 논문에서는 초등학생들이 반복실험 결과를 해석하고 예측활동을 하면서 이미 학습한 확률 개념과 통계 개념을 어떤 방식으로 활용하고 발전시키는지 살펴보았다. 특히, '자료를 어떤 관점에서 파악하는가?' 그리고 '변이성을 인식하는가?' 여부에 따라 드러나는 특징을 살펴보았다. 자료를 보는 관점은 확률, 평균, 독립시행, 그리고 반복실험의 의미를 설명하거나 토론할 때, 그리고 반복실험 결과를 해석하여 예측할 때 구체적으로 드러났다. 이 관점을 해석함으로써 학생들의 변이성에 대한 인식과 통계적 사고 특성을 파악하였다. 관찰 대상 학생들 중 변이성을 인식하였던 학생 1[3]은 나머지 두 학생과 달리 분포에 대해 적절하게 파악하고 이를 반영하여 그래프를 그리려 하였다. 학생1이 의도한 그래프는 개별 자료값들을 전체 자료의 일부로 보면서 변이성을 고려하여 분포를 표현하려는 흔적이 있는 반면, 나머지 두 학생은 개별 자료값 자체에 대한 정보만을 반영한 그래프를 그리려 하였다. 두 학생이 변이성을 고려하지 못하는 것은, 이론적인 확률 관점에 고착되어 표본의 크기에 상관없이 단지 비례 관계만을 고려하는 모습에서도 확인할 수 있었다. 일부 학생들은 비례 관계만이 아니라 '평균적으로 얻을 수 있는 값으로서의 확률' 개념을 형성하기도 하였다. 또한, 독립시행의 의미를 직관적으로 파악하지 못하여 '반복실험의 모든 결과값들은 전 범위 혹은 일정 범위 안에서 균등하게 분포되어야 한다.'는 오개념을 형성하였다.

그래프를 자료 표현의 수단으로 활용함에 있어서도, 변이성을 고려하는 관점과 그렇지 못한 관점을 확인할 수 있었다. 단지 시간(실행 순서)에 따른 결과로서의 자료값들에 주목하여 그래프를 그린 학생들은 변이성에 전혀 주목하지 못하였다. 이들은 평균적으로 얻을 수 있는 값으로서의 확률을 균등 분포 형태로 제시할 뿐, 반복실험 상황에서 부딪힐 수 있는 변이

3) 학생1은 [에피소드1]에 나오는 학생B와 동일인이며 [에피소드1]에 뒤따라 나오는 네 가지 반응 중 반응4의 의견을 내었다.

현상을 생각하지 못하였다. 학생 1을 포함한 일부 학생들은 빈도수에 대한 정보를 이용하여 그래프를 그렸는데, 이는 변이성으로 인하여 매 시행마다 달라질 수 있는 자료값을 적절히 고려한 것으로 볼 수 있다.

이 연구에 참여한 학생들은 또래의 4~5 학년 수준에서는 수학적으로 매우 우수한 성취를 보이고 실제로는 6학년이상의 사고력이 있음에도 불구하고, 선행연구(예를 들어, Reading & Shaughnessy, 2004)에서 제시한 일반 초등학생들의 변이성에 대한 이해, 곧, 반복시행에서의 변이성에 대한 이해, 표본의 크기가 자료의 신뢰도에 미치는 영향이나 표본이 가지는 변이성에 대한 이해 양상과 별다른 차이를 보이지 않았다. 이는 우리나라의 초등학교 통계교육이 변이성 인식에 도움을 주지 못하고 있음을 시사한다. 변이성에 대한 인식은 자료를 왜 수집해야 하는지, 어떤 방식으로 수집해야하는지, 수집한 자료를 어떻게 다룰 것인지, 표나 그래프로 표현된 자료를 어떻게 이해할 것인지 등 통계적 사고에 초점을 둔 교육적 조치(Ben Zvi & Garfield, 2004)에 의해서 향상될 수 있음을 시사하기도 한다. 개정 수학과 교육과정에서는 이와 같은 관점을 얼마간 반영하려고 하였으나, 이를 교과서와 수업에 반영하는 일은 교육과정에서의 변화보다 더 크고 구체적인 변화를 요구한다. 이를 위한 실행 연구가 후속되어 우리나라 통계교육이 확률교육과 적절히 연결되는 가운데 통계적 사고를 향상시키는 방향으로 개혁되기를 기대한다.

참고문헌

교육과학기술부(2007). **수학과 교육과정**. 서울: 교육과학기술부.

신보미(2007). **시뮬레이션을 활용한 확률지식의 교수학적 변환 방식**. 한국교원대학교 박사학위논문.

우정호(2000). **수학 학습 지도 원리와 방법**. 서울: 서울대학교 출판부.

이경화(1996). **확률 개념의 교수학적 변환에 관한 연구**. 서울대학교 박사학위논문.

이경화 · 지은정(2008). 그래프의 교수학적 변환 방식 비교: 우리나라 교과서와 MiC 교과서의 초등 통계 내용을 중심으로. **수학교육학연구 18**(3). 353-371.

이영하 · 권세림(2009). 정보 분석 및 활용 측면에서의 중학교 2학년 확률 단원 분석, **학교수학 11**(3). 389-413.

Ben Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-15). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Canada, D. (2006). Variability in a probability context: Developing preservice teachers' understanding. In J. Novotna, H. Moraova, M. Kratka, & N. Stehlikova (Eds.), *Proceedings of the 30th Annual Meeting of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 265-272). Prague, Czech Republic.

Canada, D. (2008). The known mix: A taste of variation. *Mathematics Teacher, 102*(4), 286 291.

Cobb, P. (1992). Report of the joint ASA/MAA committee on undergraduate statistics. In *America Statistical Association 1992 Proceedings of the Section on statistical Education* (pp. 281-283). Alexandria, VA: ASA.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning, 1*(1), 5-43.

Jacobs, V. R. (1999). How do students think about statistical sampling before instruction? *Mathematics in the Middle School, 5*(4), 240-246, 263.

Konold, C., & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *Research companion to principles and standards for school mathematics* (pp. 193 215). Reston, VA: National Council of Teachers of Mathematics.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education, 26*(1), 20-39.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistics Review, 65*, 123-137.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben& J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 17-46). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Reading, C., & Shaughnessy, J. M. (2000). Student perceptions of variation in a sampling situation. In T. Nakahar, & M. Koyama (Eds.), *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 89-96). Hiroshima, Japan.

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201-226). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Rubin, A., Bruce, B., & Tenny, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere Jones (Ed.), *Proceedings of the third international conference on teaching statistics* (pp. 314-319). Voorburg, Netherlands: International Statistical Institute.

Shaugnessy, J. M. (2003). Research on students' understanding of probability. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *Research companion to principles and standards for school mathematics* (pp. 216-226). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J. M. (2006). Research on students' understanding of some big concepts in statistics. In G.F. Burrill, & P. C. Elliott (Eds.), *Thinking and reasoning with data and chance: sixty eighth yearbook* (pp. 77-98). Reston, VA: NCTM.

Shaughnessy, J. M., & Ciancetta, M. (2002). Students' understanding of variability in a probability environment. In B. Philips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society* (CDROM), South Africa. The Netherlands: IASE.

Shaughnessy, J. M., Ciancetta, M., & Canada, D. (2004). *Students' attention to variability when comparing distributions.* Paper presented at the Research Presession of the 82nd Annual Meeting of the National Council of Teachers of Mathematics, Philadelphia, PA.

Shaughnessy, J. M., & Pfannkuch, M. (2002). How faithful is old faithful? Statistical thinking: A story of variation and prediction. *Mathematics Teacher, 95*(4), 252-259.

Torok, R., & Watson, J. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal, 12*(2), 147-169.

Watson, J., Callingham, R., & Kelly, B. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning, 9*(2), 83-130.

Wild, C., & Pfannkuch, M.(1999). Statistical thinking in empirical enquiry. *International statistical Review, 67*(3), 223-265.

Zawojewskiu, J. S., & Shaughnessy, J. M. (2000). Mean and median: Are they really so easy? *Mathematics Teaching in the Middle School, 5*(7), 436-440.

A Case Study of the Characteristics of Mathematically Gifted Elementary Students' Statistical Reasoning : Focus on the Recognition of Variability

Lee, Hyung Sook (Eastern Washington University)
Lee, Kyeong Hwa (Seoul National University)
Kim, Ji Won (Ajou Science Education Center for Gifted Students)

It is important for children to develop statistical reasoning as they think through data. In particular, it is imperative to provide children instructional situations in which they are encouraged to consider variability in data because the ability to reason about variability is fundamental to the development of statistical reasoning. Many researchers argue that even highperforming mathematics students show low levels of statistical reasoning; interventions attending to pedagogical concerns about child ren's statistical reasoning are, thus, necessary. The purpose of this study was to investigate 15 gifted elementary students' various ways of understanding important statistical concepts, with particular attention given to 3 students' reasoning about data that emerged as they engaged in the process of generating and graphing data. Analysis revealed that in recognizing variability in a context involving data, mathematically gifted students did not show any difference from previous results with general students. The authors suggest that our current statistics education may not help elementary students understand variability in their development of statistical reasoning.

* **Key word** : statistical reasoning(통계적 사고), variability(변이성), variation(변이), mathematically gifted students(수학영재)

논문접수: 2010. 7. 23
논문수정: 2010. 8. 13
심사완료: 2010. 8. 20