

중학교 1학년 통계단원에 나타난 분포개념에 관한 분석¹⁾

이 영 하* · 최 지 안**

이 연구는 우리나라 중학교 1학년(7-나 단계) 교과서 통계단원에서의 교과서 서술 방식에 대해 남주현(2007)의 세 통계적 핵심개념 가운데 분포개념의 관점에서 분석해 본 것이다. 따라서 이 단원의 학습목표의 핵심은 분포개념의 습득과 표현, 비교법 등이라고 전제한 것이다. 연구결과 다음을 알 수 있었다.

첫째, 학생들은 분포가 무엇이고, 어떤 것이 분포가 아닌지 구분할 수 있어야 한다.

둘째, 여러 형태의 분포의 표현법들은 각각 고유의 학습 이유가 제시될 필요가 있다. 그래야 학생들의 동기를 유발하고, 학생들은 적절한 상황에서 적절한 표현법을 사용할 줄 알 수 있을 것이다.

셋째, 장래 배울 확률밀도함수의 자료관련 이해를 증진시키려면 7학년의 히스토그램 수업에서 밀도히스토그램을 학습해 두는 것이 도움이 될 것이다. 적어도 상대도수를 넓이로 표현하는 경험은 필요해 보인다.

넷째, 두 도수분포의 비교, 특히 도수분포 다각형을 통한 비교의 필요성 여부가 재고되어야 한다.

다섯째, 통계문맹에 대한 7학년 단계부터의 완만한 준비와 점진적 학습이 필요하지 재고될 필요가 있다.

1. 서 론

현대 정보사회에서는 일상생활과 정치, 경제, 사업 등 각 분야에서 일어나는 각종 불확실한 현상에 대한 합리적인 판단 도구로서 통계적 방법이 놀라울 정도로 광범하게 사용되고 있어 통계적 소양이 현대 사회생활을 영위하는데 필수적으로 요구되고 있다. 오늘날 통계교육에 대한 관심의 증대는 이러한 사회적 요구와 과학기술 분야의 필요성에 기인하는 것이다.

그러나 우리나라 학교수학에서 통계는 문제 해결 도구로서 혹은 주변 세계를 이해하는 유

용한 도구로서보다도 초등 통계학에 나오는 특정한 내용으로 구성된 교재로 간주되고 있다(우정호, 2001, pp.479-480). 또한 학생들은 오래 전부터 “수학을 왜 배우느냐?”는 행동 중심적(동기적) 목표를 묻고 있다. 이것은 단순한 지식 습득 대신 “배워서 나에게 어떤 다른 생활, 어떤 다른 행동을 기대하느냐?”라는 행동중심의 목표를 묻고 있는 것으로 생각할 수 있으며 수학자가 수학을 하는 이유는 그것이 무엇이건간에 분명하리라 생각되지만 우리 학생들은 그렇지 않다고 생각되며 따라서 목표와 동기에 관한 지도는 지식 못지않게 중요하다고 할 수 있다. 또 중등학교 통계 교육에서 전 세계적인

* 이화여대, youngha@ewha.ac.kr

** 이화여대 대학원, checusi@hanmail.net

1) 본 논문은 최지안(2008)의 석사학위 논문을 요약한 것임

로 아직도 해결되지 않은 중요한 한 가지는 행동중심적 학습지도가 충분히 이루어지지 않고 있다는 것인데, 그 중요한 이유는 “자료의 정확한 수집방법과 통계적 방법의 사용 절차”에 대한 교육이 통계 단원에서 소홀히 취급되고 있다는 것과 같은 맥락을 갖고 있다고 하면서 지식 중심의 내용 구성으로 인해 평균, 막대그래프, 신뢰구간, 가설검정처리 지식만이 강조되고, 그 결과 수업의 출발점은 “수집된 자료를 분석하기 위해 이들 지식을 어떻게 적용하느냐”에서부터 항상 출발한다는 것이다. 즉, 수업의 출발은 항상 자료수집이후가 된다는 것이다(이영하, 2007, pp.2-4). 이영하, 남주현(2005)은 선행연구들(Bakker, 2004; Garfield & Ben-Zvi, 2004; Garfield, 2002)을 바탕으로 제안한 분포, 요약, 표본의 세 가지를 핵심적인 통계적 개념이라고 주장하였다. 현재 우리나라에서 시행되고 있는 7차 교육과정의 확률과 통계 영역에서 가르치는 내용을 살펴보면, 국민공통기본 교육과정인 1단계부터 10단계에서 모두 분포개념과 관련된 내용을 다루고 있음을 알 수 있다. 김경란(2007)은 학습자가 효과적으로 교과 내용을 학습하기 위해서는 교과 특성과 체계에 적합하도록 각 교과내용의 연계성을 고려하여 내용을 조직해야 한다(p.10)고 하였다.

본 연구에서는 남주현(2007)이 설명한 분포 개념들 중에서 7-나 통계단원과 관련된 표현과 분포, 분포와 비교, 중심경향성과 요약에서 분포개념을 살펴보고, 분포개념에 대한 연계성을 위하여 내용면에서 누락된 부분은 없는지, 적합한 내용으로 구성되어 있는지 여부를 살펴보고, 그 결과 문제점이 드러난다면 개선방안을 찾아보고자 한다. 또한 도수분포표, 히스토그램과 도수분포다각형, 상대도수와 누적도수의 그래프에서 ‘표와 그래프를 왜 배우는가?’를 제시하고 있는지, 그리하여 학생들의 지적호기심을

자극하여 태도변화를 이끌어낼 수 있는 행동중심적 학습지도가 이루어지고 있는지 살펴보고자 한다.

한편 우정호(2001)는 학생들이 일반적으로 형식적인 통계적 처리에 익숙하고 통계적 사고의 특성에 대한 의식이 결여되어 있어, 일상생활에서 광범위하게 사용되고 있는 통계치에 대한 비판적 안목이 미약하다(pp.481-492)고 하였으며 이영하(2007)는 최근 통계적 보도자료에서 정확하지 않은 자료나 잘못 사용된 통계절차나 방법 등을 수시로 접할 수 있다고 하면서, 이로 인해 학교 통계 교육의 중요성을 실감케 한다(p.4)고 하였다. 학생들이 졸업하고 사회로 진출했을 때, 실제 생활에서 접하는 통계자료나 정보에서 통계문맹의 심각함을 엿볼 수 있는 부분이라고 할 수 있다. 따라서 이런 실생활에서의 통계문맹의 퇴치나 기반을 조성하기 위해 본 연구대상인 7-나 통계단원에서 정확한 자료, 통계적 방법에 대한 내용이 구성되어 있는지 여부를 살펴보고, 그 결과 문제점이 드러난다면 개선방안을 찾아보고자 한다.

II. 이론적 배경

1. 분포개념과 그의 발달

가. 분포의 의미

분포의 사전 상의 의미는 한 변수의 변량의 패턴, 한 변수의 모든 가능한 수치들과 어떻게 각 값이 발생하였는지를 기록하는 것, 관찰했거나 이론적인 분류상의 발생의 빈도수를 보여주는 하나의 정렬이다(Oxford Dictionary, 2005). 다시 말해, 분포란 여러 가지 사회 현상, 자연 현상 속에서 발견되는 사물의 속성이 다양할 때 그 다양한 속성 각각의 빈도를 생각하는 개

념(이영하, 남주현, 2005, p.470)으로서, 분포는 통계학에서 다른 개념들과 연관된 가장 중요한 개념으로(Bethlehem & De Gooijer, 2000; Bakker, 2004), 변이성에서 패턴을 파악하도록 개념적 구조를 조직하고 총체로써 자료 집합을 볼 수 있게 한다. 다시 말해, 분포는 도수 분포로부터 확률 분포 함수에 이르기까지 많은 측면과 가능한 이해의 서로 다른 층을 가지고 있는 복잡한 개념이기 때문에 학생들이 가지고 있고 발달시킬 수 있는 분포의 서로 다른 개념들을 얻어내야만 할 필요가 있다(Bakker, 2004; 이영하, 남주현, 2005 재인용, p.460). 우리가 관찰할 수 있는 여러 가지 사회현상, 자연현상 속에서 드러나는 사물의 속성들이 다양할 때, 그 다양한 속성 각각의 빈도²⁾를 생각하는 개념이다. 여기서 속성이라는 것은 키와 같이 연속적인 양에 대한 것일 수도 있고, 과일 종류와 같은 범주적 성격³⁾을 지닌 것일 수도 있는데, 이러한 속성에 따라 대상을 분류했을 때 그 도수가 더 많은지 혹은 더 적은지를 아는 것이다. 예를 들어 사람의 키에 대한 분포가 있을 때 속성으로 ‘매우 큰’, ‘중간정도’, ‘매우 작은’등으로 구분할 수 있고, ‘키가 매우 큰 사람이나 매우 작은 사람은 적다’, ‘중간정도 키를 가진 사람은 많다’라고 표현할 수 있다. 다시 말해, 분포를 이해한다는 것은 범주의 의미와 빈도의 의미를 알고 이 두 개념을 결합시키는 것이라고 할 수 있다(김경란, 2007, p.5).

나. 분포개념의 발달

분포개념은, 자신의 선호와 관련된 판단을

보여주는 행위를 통해 그것의 초기형태는 거의 동물적 감각임을 알 수 있다. 가령 같은 조건이라면 자신의 먹이가 더 풍부한 곳에 동물이 더 자주 출현한다는 것은 동물조차도 먹이 분포에 대한 분포개념적 판단을 하고 있음을 추측케 한다.

따라서 인간에게 있어서도 이와 같은 수준의 분포개념은 매우 초기에 형성된다고 추측되며, 이 시기는 상대도수에 대한 지식도, 비례개념도 없는 상황에서 이미 시작된다고 추측할 수 있다. 그렇기 때문에 분포개념의 발달에 관한 우리의 주장은, “피아제의 확률개념 발달”에 대한 비판에서 볼 수 있는 바와 같이, 확률개념이 비례개념의 형성 이전에도 지도 가능한 직관적인 것임을 주장하는 학자들과 같은 맥락 위에 있으나, 피아제 또는 그를 비판하는 다른 학자들의 주장과 우리의 주장이 크게 다른 맥락은 그들이 수학적 확률로 확률을 이해하는데 비해 여기서의 분포개념은 통계적 확률을 확률의 의미로 생각하고 있다는 것이다.

이영하, 남주현(2005)은 통계 교육에서 핵심적인 통계적 개념을 무엇으로 볼 것인가에 대해 선행연구들이 제안한 것(예를 들어, Bakker, 2004; Garfield & Ben-Zvi, 2004; Garfield, 2002)을 토대로 하여 분포, 요약, 표본개념의 세 가지를 제시하였다. 이에 이영하, 남주현(2005)은 각 단계별 내용에 맞게 구성된 7차 교과서와 교사용 지도서를 중심으로 어떤 통계적 개념이 다루어지고 있는지를 분석하고 그 결과로 얻어진 각 단계에서 다루어진 주제들이 어떤 통계적 개념과 관련되는지 살펴보아, 7차 교육과

2) 어떤 사건에 대한 가능성의 크기 개념으로 특정한 상황에 대하여 어느 정도의 빈도를 차지하는가에 대한 판단은 초기에는 경험에 의존한다.

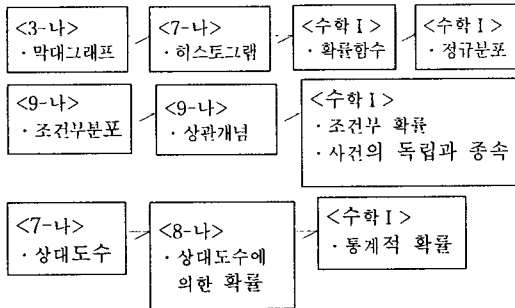
3) 범주형 자료 - 개체의 속성을 구분해 주는 자료

예) 출생지, 차의 제조회사, 평어 성적 등
수치형 자료 - 숫자로 얻어지는 자료

예) 교통사고의 횟수, 불량품의 개수, 키, 속도, 온도 등 수치형 자료에서 각각의 측정값을 셀 수 있는느냐의 여부에 따라 이산변량과 연속변량으로 구분한다.

정에서 분포 개념은 전체 교육과정에서 계속 다루어지고 있음을 주장하였다. 김경란(2007)은 앞에서 언급한 분포의 개념을 <표 II-1>와 같이 주제별로 분류, 구성하였다.

<표 II-1> 분포개념에 관한 주제별 분류(p.23)



<표 II-1>을 통해 3-나에서 7-나로 이어지는 분포개념의 주제를 살펴보면 이산변량을 다루는 막대그래프에서 연속변량을 다루는 히스토그램으로 이어지는 분포개념은 수학 I의 확률밀도함수로 연결되는 중요한 의미를 가진다는 것을 알 수 있다. 그러나 현재 교육과정에서는 7-나에서 처음 히스토그램이 다루어지고 수학 I에서 확률밀도함수가 처음 나오므로써 그 중간단계가 빠져있다. 따라서 히스토그램을 다룰 때 밀도도수 히스토그램을 넣어서 확률 구하기와 연결하는 것의 필요성이 요구된다. 즉 분포개념의 일관된 성장을 촉진할 수 있는 개선이 요구된다는 것이다(이영하, 남주현, 2005, p.469). 또한 나귀수, 황혜정, 한경혜(2001)의 연구에서는 중학교의 통계영역에 관하여 다음과 같은 개선방안을 언급하였다.

중학교에서는 자료를 도표 적으로 이해함에 있어서 외형적 그대로 보는 것에서 벗어나 그 배후의 확률변수와 분포(히스토그램)의 관점에서 볼 수 있도록 한다. …(중간생략)… 히스토그램과 누적도수와의 관계 등은 확률밀도함수, 누적분포함수로 연결되고, 정규분포를 거쳐 가설검

정 등에 이르는 중요한 시발점이기 때문에 이러한 목적의식 하에 그 흐름이 느껴질 수 있도록 관련 교육 내용 및 방법을 개발하여 반영한다(pp.253-254).

또 누적분포함수에 관하여 이영하 외(2002)는 아래와 같이 서술하고 있다.

흔히 표본 누적분포함수에 따라 누적도수의 그래프를 그리면 교과서의 방법과 유사하나 연속함수 모양 대신 계단 모양의 형태가 된다. …(중간생략)… 히스토그램이 연속변량의 사용을 원칙으로 하고 연속변량의 누적분포함수는 연속함수이기 때문에 누적도수의 그래프도 연속함수의 모양을 유지하는 것이 좋다고 생각되기 때문이다. 더욱이 확률밀도함수는 누적분포함수를 미분한 결과로 얻어지는데 우리는 누적도수의 그래프와 히스토그램 사이에 상호 미분, 적분의 관계가 성립되려면 그렇게 하는 것이 좋기 때문이다(p.51).

이상의 여러 연구 결과와 주장들은 결국 통계단원에서 다루어지는 내용들은 하나의 구조 속에서 연계성과 위계성, 그리고 아동의 인지 발달 시기에 대한 고려가 필요함을 보여주고 있다. 따라서 본 연구는 이상의 개념적 틀, 특히 분포개념을 바탕으로, 7-나 통계단원에서 다루어지는 통계적 개념과 내용들이 그런 맥락에서도 충실히 구성되어 있는지 살펴보려는 것이다.

대개 12세경에 시작하여 16세 이후에 완성되는 형식적 조작의 단계는 구체적 조작의 발달에 의존하여 통합되고 확장된다. 구체적 조작기 사고가 논리적 사고인 반면에 이것은 “구체적”세계에 제한된다. 형식적 조작이 발달되고서야 비로소 추론이 “내용으로부터 자유롭게”나” 혹은 “구체적인 것으로부터 자유로와질” 수 있다. 초등학교에서도 분포에 대한 지도는 암묵적이지만 계속되어옴으로서, 어느 정도 동

물적 분포개념 수준 이상의 발달 가능성은 열어놓았다고 본다. 구체적 조작기 사고는 가역적 사고이다. 전도와 상보성은 독립적으로 사용되고, 그리고 두 가지 가역성이 형식적 사고로 협응된다(Wordworth, 1995; 성옥련, 김수정, 이지연(역), pp.233-234). 또한, Piaget(1963b)에 따르면, 지능을 가장 분명하게 정의하는 특징이 가역성이다(p.41). 만일 사고가 가역적일 수 있다면, 그것은 시작했던 곳으로 되돌아가는 추론이 가능하다는 것이다(Wordworth, 1995; 성옥련, 김수정, 이지연(역) 재인용, p.127). 우리나라 중학교 1학년은 대부분 14세이다. 물론 개인차가 있을지라도 대체로 구체적 조작기를 포함하여 형식적 조작기에 속한다고 볼 수 있다. 따라서 7-나의 통계단원의 내용 설명 수준에서 사고유형을 보면 분포개념에서 자료를 도수분포표로 나타내고, 도수분포표를 다시 자료를 나타내는 것, 히스토그램과 도수분포다각형의 변환, 즉 형식과 형식의 변환은 Piaget의 4단계의 인지발달 중에서 형식적 사고기의 가역적 사고에 해당되는데, 여러 가지 분포 표현법의 습득과정에서 가역적 사고의 폭을 넓힐 수 있도록 교과내용이 구성되어 있는 지 살펴보려는 것이다.

다. 7-나 단계에서의 분포개념

통계학은 기본적으로 자료수집, 자료 정리와 분석, 결과해석 등으로 나뉘고 자료정리와 분석은 자료의 요약, 정리에 관한 기술통계와 자료로부터 모집단의 특성을 추정 및 검정하는 통계추론(statistical inferences)으로 나뉜다(이외숙 외, 2002, p.6). 기술통계학에서 이루어지는 자료처리의 과정이 어떤 통계적 개념을 목적으로 관련되는가를 나타내면 다음과 같다(이영하, 남주현, 2005).

<표 II-2> 자료처리과정과 관련된 통계적 개념(p.459)

자료 수집	->	분류	->	도표	->	자료의 중심, 퍼짐에 관한 측도 구하기
요약		분포		요약		

이영하, 남주현(2005)이 통계적 개념으로 분류한 분포, 요약, 표본에서 7차 교육과정의 7-나에 해당되는 부분은 분포와 요약이다. 남주현(2007)은 통계적 방법론을 통계교육적으로 고찰하고 표현과 분포, 분포와 비교, 특이값, 변이성과 분포, 상관, 표집분포, 표본·표집·변이성, 분포의 산포와 신뢰도로 분류하였다. 남주현(2007)의 통계교육적 고찰에 따른 표현과 분포, 분포와 비교, 중심경향성의 요약은 7-나 통계단원의 내용에 해당되므로 분포의 표현과 분포의 비교, 중심경향성의 요약을 살펴보기로 한다.

1) 분포의 표현

현행 통계교육과정은 기술통계에 치우치거나, 수학적 확률론처럼 구성이 되었고, 막대그래프가 다루어질 때 분포를 나타내는 것과 단순히 자료 정리의 목적인 것과의 구분조차 없으며, 히스토그램은 범주의 도수세기에 치중되어 왜 히스토그램을 그리는지에 대한 설명이 부족함과 더불어 후에 확률밀도함수로 어떻게 연결이 되는지에 대한 언급이 없음을 쉽게 찾아볼 수 있다(남주현, 2007, p.2).

(도수)분포의 표현은 스스로 분포개념의 형식적 표현임과 동시에 분포를 표현하려는 목적을 지닌 도표적 요약이라고 할 수도 있다. (도수)분포 자체를 하나의 요약으로 보는 관점은 분포 추정이론에서 본격적으로 드러나지만, 석사수준이하의 통계에서는 함수공간에서의 수학적 이론의 도움을 받아야하는 분포추정이론까지는 배우지 않는 것이 보통이기 때문에 중등 통계 차원에서는 요약의 관점보다는 분포개념

의 일환으로 보는 것이 더 적절해 보인다.

또 남주현(2007)은 아래와 같이 서술하고 있다.

분포의 시각적 표현, 특히 표⁴⁾나 그래프들은 중요한 의미를 지닌다. 수치적으로 주어진 자료값들을 하나의 총체로써 인식하기 위해서, 다시 말해 분포상태에 대한 관찰을 위해서 표나 그래프와 같은 시각적 표현법을 가져오게 된다. 자료가 어떤 방식으로 주어지든 그것이 분포를 이루지만 시각적 표현을 통해 분포상태를 기술하고 그러한 표현법들이 분포상태를 적절하게 나타내었는가에 대한 평가 능력까지도 분포개념을 형성하는 데 필요로 된다고 볼 수 있다 (p.146).

통계적 목적을 가지고 통계 집단내의 각 개체로부터 수집되는 자료를 흔히 원시자료(raw data)라 한다. 중학교 교과서에서는 수치적 자료의 각각의 자료를 변량이라 하는데, 변량은 다른 구분 방법도 있으나 우리에게 특히 중요한 것은 값들이 연속적인가 아니면 이산적인가를 구분하는 것이 매우 중요하다. 왜냐하면 그것을 알아야 막대그래프와 히스토그램을 구분할 수 있고, 학교에서의 확률밀도함수로의 발전이 일관되게 전개될 수 있기 때문이다(이영하, 2002, p.359). 임재훈 외(2004)의 연구에서는 막대그래프는 주로 이산변량에 활용하고, 이와 유사하지만 히스토그램은 연속량 자료에 맞게 수확화된 그래프 방식이라고 하였다(김경란, 2007 재인용, p.24). 즉 이산변량의 경우 각 (이산적)변량별 도수를 나타내는 막대그래프는 상대도수의 막대그래프로, 이어서 (통계적)확률을 나타내는 막대그래프로 이어져, 결국 확률함수로 이어진다. 연속변량의 경우는 도수의 히스토그램이 상대도수의 히스토그램의 도움을 받아 밀도도수의 히스토그램으로 이어지는데

이것은 직사각형의 넓이가 상대도수를 나타내도록 고안된 것이다. 직사각형 전체의 넓이가 1이 되고, 이것은 확률밀도함수의 곡선 아래의 넓이가 1이 되는 것으로 발전하면서 넓이가 상대도수 또는 (통계적)확률로 이어져 확률밀도함수가 되는 것이다. 확률 분포 표현의 이와 같은 두 갈래는 상대도수 또는 (통계적) 확률을 높이로 나타내느냐, 넓이로 나타내느냐의 차이임으로, 도수 및 자료와 관련된 분포표현에 대한 이와 같은 구분은 조기에 이루어질 때, 장차 확률함수와 확률밀도함수의 자료 분석과의 관련성과 그 뿌리를 알게 해줄 수 있는 것이다.

히스토그램을 이용한 분포표현에서 계급의 수는 매우 중요하다. 계급의 수가 너무 많거나 너무 적으면 자료의 분포 상태를 잘 드러내기 힘들어지기 때문이다. 이영하(2007)는 히스토그램을 위한 계급의 개수에 대해 아래와 같이 설명하고 있다.

계급의 수는 히스토그램을 왜 그리느냐 하는 이유와 밀접한 관련이 있어서 이 단원의 목표에 직결된 매우 중요한 주제이다. 히스토그램을 그리는 이유는 자료의 분포상태를 알기 위함이다. 그리고 계급의 수는 자료의 분포 상태를 잘 나타내는 정도로 정하면 좋다. 계급의 수가 너무 적거나 너무 많으면 자료의 분포상태를 잘 나타내지 못한다. 그러나 5이상 15개 이하는 답이 되지 못한다. 왜냐하면 몇 개라는 답을 주지는 않기 때문이다(p.9).

김원경 외(2003, p.51)에서 계급의 개수(크기)를 정하는 방법을 살펴보면 아래와 같다.

계급의 개수를 결정하는 방법에는 울(Yule, G. U)의 방법과 스티지(Sturges, H. A)의 방법이 있다.

4) 물론 표보다는 그래프가 위치나 형태 등과 관련된 분포의 특성을 더 용이하게 파악할 수 있게 하지만, 표가 빈도를 쉽게 나타낸 준다는 의미에서 표도 분포의 시각적 표현으로 포함시킨다.

- ① 율의 방법은 경험에 의해 다소 주관적인 방법으로 계급의 수를 자료의 양에 따라 5~15의 범위 안에서 적당하게 결정하는 것이다.
- ② 스테지는 표본의 크기에 의하여 계급의 수를 결정하며 그 공식은 다음과 같다.

$$k = 1 + 3.3[\log n]$$

(단, k 는 계급의 수, n 은 자료의 크기)

이영하 외(2002, p.55)에서 스테지의 공식을 다음과 같이 소개하고 있다.

$$k = 1 + [\log_2 n]$$

(단, k 는 계급의 수, n 은 변량의 총수)

계급의 수를 적용한 후에 실제 그린 히스토그램이 자료의 분포상태를 잘 나타내지 못한다고 생각되면 계급의 수를 한두 개 늘리거나 줄이면서 더 좋은 분포 상태를 얻기 위해 노력해야 한다. 그러므로 이 귀찮은 과정은 컴퓨터를 이용하는 것이 좋다. 그러면 더 좋은 분포상태란 어떤 것인가? 한마디로 이것은 미술에서 더 잘 그린 그림은 어떤 것인가라고 묻는 것과 흡사하다. 다만 원칙적으로 변량들이 어느 값 근방에 더 모여 있고 덜 모여 있는 지를 더 잘 나타내는 것이 더 좋은 분포 상태의 표현이라고 할 수 있다(이영하, 2007, p.9). 이런 맥락에서 본 연구는 도수분포표에서 계급의 개수(크기)를 결정하는 방법을 살펴보고, 좋은 분포상태에 대한 의미와 히스토그램을 그리는 목적이 어떻게 서술되어 있는지 살펴보려는 것이다. 황해정, 나귀수, 서동엽(2000)의 연구에서 도수분포다각형은 히스토그램을 유연화(smoothing)시킨 주제로서, 초·중등 확률과 통계과정에서 유일한 유연화 주제라고 언급하였다. 히스토그램의 계급의 간격을 계속 줄여나가면 아래 [그림 II-1]과 같은 곡선을 얻는다(조태근 외, 2003).



[그림 II-1] 히스토그램의 유연화(p.287)

이와 같이 김우철 외(2005), 이외숙 외(2002), 박정식, 윤영선(2002)에서 설명한 밀도도수분포의 극한형으로 본 확률밀도함수의 그래프를 통해 도수분포다각형이 히스토그램의 부드러운 유연화 형태임을 알 수 있다.

성태제(1995)는 히스토그램에 대해 다음과 같이 설명하고 있다.

[그림 3-6]은 올해 기초통계학을 수강한 200명 학생들의 학기말 시험점수의 분포이고 만약 작년도에 수강한 학생들의 학기말 시험점수와 비교하기 위하여 그래프를 그린다면 히스토그램으로는 여러 가지 문제점이 따른다. 각 급간에 전년도 기초통계학 수강생들의 시험점수 분포를 그린다면 덧칠에 의한 그림으로 비교가 불편하다. 히스토그램은 어떤 한 집단 혹은 수집된 어떤 양적 연속변수에 대한 도식은 충족시키나 집단비교를 위해서는 좋은 그래프라 하기 어렵다(p.54).

또한 김우철 외(2006)는 상대도수의 분포다각형에 대해서 “두 개 이상의 자료집합의 분포를 같은 그림 위에 놓고 비교할 수 있어 편리하다(p.33).”고 설명하고 있다. 분포의 표현 관점에서 유연화는 자체로서 중요한 통계적 의미를 지니고 있음은 분명하지만 초중등 전 과정을 통해 연결성이 없는 고립된 주제일 뿐 아니라, 통계 이론 측면에서도 특수한 벡터공간에서 전개되는 유행이론(spline theory)으로 연결되는, 통계학 석사 과정까지도 등장하지 않는 유연화 주제를 중등학교에서 소개할 필요가 있는지는 의심스러우며, 따라서 도수분포다각형의

작성목적은 분포표현의 유연화 문제보다는 위 김우철의 주장과 같이 분포의 비교 측면에서 다루어지는 것이 더 바람직하다고 생각된다. 이와 관련하여 본 연구에서는 도수분포다각형을 왜 그리는 지에 대해 어떻게 서술하였는지 에 대해 살펴보려는 것이다.

2) 분포의 비교

남주현(2007)은 아래와 같이 서술하고 있다.

두 분포의 비교라는 맥락에서 아동들의 추론 능력을 조사한 Watson & Moritz(1999b)의 연구에서 3학년에서는 주로 개별적인 값들에서의 절대도수에 의한 크기 비교가 이루어졌고 5-7학년에서는 크기가 다른 두 집단의 분포를 비교하는 맥락에서 비례적 추론을 이용하기 시작하였다. …(중간생략)… Watson & Moritz(1999b), Shaughnessy et al.(2004), Bakker & Gravemeijer (2004)의 연구에 의하면, 6, 7학년의 경우에는 각 분포를 평균으로 요약하여 비교하는 반응을 보이기 시작하였으며, …(중간생략)… Shaughnessy et al.(2004)의 연구에 의하면 아동들이 두 분포를 비교할 때 대부분의 6-12학년 학생들은 두 분포의 비교를 할 때 산포도나 평균과 같은 통계적 요약을 주로 적용시킨다는 것을 발견할 수 있었다(pp.111-115).

이 설명은 분포의 비교 중에서 조건부분포의 비교에 관한 선행연구들에 대한 것이다. 가령 두 가지 조건(예; 사건 A와 A')에 따라 다른 관찰(예; B 와 B')의 조건부 분포에 변화가 있다면 이들 사이에 서로 상관이 있어 보인다는 직관을 통해 상관개념으로 이어지는 분포개념을 설명하고 있는 것이다. 이 설명은 결국 7-나 단계의 분포에 관한 지도에서 (조건부)분포의 비교는 이후의 상관, 종속사건과 관련되는 중요한 징검다리로서의 역할로 이해되고 강조되어야 한다는 것을 알 수 있다. 아울러 분포의 비교는 그 스스로도 중요한 가치를 보여 줄 수

있는데, 간혹 한 개의 분포에서 분포 상태의 파악이라는 분포 표현보다는 둘 이상의 분포를 비교하는 목적의 분포표현이 실생활에서 더 자주 등장한다는 점 때문이다.

분포의 표현에서 높이 또는 넓이를 이용한 두 갈래의 표현은 엄밀한 수학자의 시각에서는 다소 문제가 있는 표현 양식이다. 특히 확률밀도함수는 같은 분포에 대해 유일(unique)하지 않으며(측도론적으로 almost everywhere 유일하지만) 또 수학적으로만 생각하면 확률함수로도, 확률밀도함수로도 표현할 수 없는 분포(singular distribution)도 존재하기 때문이다. 적어도 확률함수와 확률밀도함수의 두 갈래 표현을 하나로 통합할 수 있는 표현 양식이 누적분포함수이다. 이런 맥락에서 누적도수의 그래프는 누적상대도수의 그래프, 누적 확률분포로 이어지는 교육과정 상의 의미를 생각해 보게 한다.

이외숙 외(2002)는 이산변량과 연속변량의 누적분포함수에 대해 이산확률변수의 분포함수 그래프는 아래 [그림 II-2]의 $F(x) = P[X \leq x]$ 에서 $P(X=x) > 0$ 인 x 들에서 점프(jump)가 일어나는 계단함수(step function)임을 설명하고 있다.

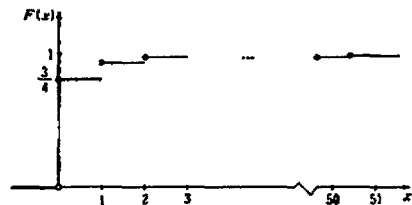


그림 4.6 분포함수 (4.2)의 그래프

[그림 II-2] 이산변량의 누적분포함수(p.87)

이산확률변수의 분포함수는 $P(X=x_i) > 0$ 인 x_i 에서 점프가 일어나는 계단함수였던 것에 반해 연속확률변수의 분포함수 $F(x) = P(X \leq x)$ 는 모든 x 에 대해 $P(X=x)=0$ 이므로 연속함수이다.

분포의 비교라는 관점, 특히 분포가 이산변량

의 것이나, 아니냐에 구분 없이 비교하는 관점에서는 두 갈래 분포 표현의 통합은 필수적이라고 할 수 밖에 없다. 확률함수와 확률밀도함수를 비교하는 것은 두 표현의 높이가 의미하는 것이 다르므로 의미가 없어지기 때문이다.

한편 7-나 단계에서는 사실상 연속변량(히스토그램)만을 취급함으로써 이런 통합은 중요하지 않다고 볼 수도 있다. 그렇다 해도 꺾은 선 그래프 모양의 누적(상대)도수의 그래프는 비교의 목적에서 좋은 시각적 효과(두 그래프를 함께 그려 비교할 수 있다는 점)를 갖는다. 그리고 이 점은 도수분포다각형에서와 같다. 그러나 분포 상태의 비교는 도수분포다각형에서 더 확연히 드러나며, 누적도수의 그래프에서는 관련 전문가가 아니면 비교의 결과를 확률밀도함수의 형태(히스토그램 또는 도수분포 다각형)로 바꾸어 이해해야 하므로 다소간의 사고 과정을 필요로 한다.

여기서 7-나 단계의 내용 분석과 관련된 몇 가지 중요한 분석 관점을 생각할 수 있다. 즉 상대도수와 누적도수의 그래프는 어떤 지도 목표의 관점에서 어떤 방법으로 서술되고 있으며, 도수분포와 누적도수분포의 표현간의 번역 문제는 어떻게 다루어지고 있는가에 관한 것이다.

3) 중심경향성과 요약

남주현(2007)은 아래와 같이 서술하고 있다.

Mokros & Russell(1995)는 통계학의 목적들 중 하나는 “크고, 다루기 힘들고, 무질서한 정보들의 모음을 요약된 표현으로 줄이는 것”이라고 지적한다. 이러한 목적을 반영하는 것 중 하나가 자료의 중심 측도적 특성을 요약하는 활동이라고 볼 수 있다. …(중간생략)… 특히 평균, 중앙값, 최빈값의 세 가지 측도들은 다양한 맥락에서 자료 집합을 위한 대푯값을 찾는 것으로 많은 나라의 교육과정에서 다루어져 왔다(p.119).

중심위치의 측도는 주어진 자료가 어떤 값을

중심으로 분포되어 있는 가를 알려주며, 중심위치의 측도로는 평균, 중앙값, 최빈값 등이 있으며 이 중 평균은 산술평균을 의미하며 중심위치의 측도로서 가장 많이 사용되고 양적 자료에만 사용된다(김우철 외, 2006, p.48). 그래프만을 사용하여 통계조사 결과를 비교할 수 없을 때, 자료를 두어 가지의 수치로 대표할 수 있는 측도를 정의하여 사용하게 되는데, 가장 많이 사용되는 두 가지의 기술측도는 자료의 중심에 대한 측도와 자료의 퍼짐에 대한 측도이다(이외숙 외, 2002, p.21). 이방수 외(2002)는 자료의 특성을 수치화하는 방법으로 분포의 중심위치를 나타내는 측도(평균, 중앙값, 최빈값), 평균치에서 흩어짐(산포)을 나타내는 측도, 분포의 형태를 나타내는 측도(왜도, 첨도), 상대적 위치의 측도(특정값이 주어진 자료가 어떤 위치에 있는가)를 이용한다(pp.246-247)고 하였다. 7-나 통계단원에서 소개되는 평균의 이와 같은 분포요약의 의미 관련성과 교육과정 상의 의미를 생각해 볼 필요가 있다. 즉 중심측도로서 평균이 소개되고 있는지, 장차 기댓값과의 연결을 도모하고 있는지 등을 살펴보는 것이다. 유한모집단에서의 평균은 무한 모집단에서의 기댓값이다. 현재 소개되고 있는 도수분포표에서의 평균은 상대도수를 이용하여 고쳐 적으면 계급값과 상대도수의 곱들의 합이다. 상대도수를 (통계적) 확률로 바꾸어 적으면 곧 기댓값 공식이 된다. 이것을 밀도도수의 히스토그램에 적용하면 확률밀도곡선에서의 기댓값인 적분 표현에 대한 구분구적법 표현과 흡사해지는 것을 곧 알 수 있다. 교육과정에서의 이런 흐름에 대한 7-나 단계에서의 준비는 어떠한 지를 살피는 것이 기댓값과의 연결성을 보겠다는 의미이다.

라. 통계문맹

통계학은 자료로부터 정보를 얻는 자료과학

이며(우정호, 2000, p.4), Moore(1990)가 강조했던 것처럼 통계교육에 있어서 자료가 중요하다(남주현, 2007 재인용, p.4). 자료는 흔히 특별한 계획을 갖고 있는 사람에 의해 수집되고 제시된다. 자료처리에서 자료 이면에 있는 자료 수집자의 신념과 태도는 자료를 조직하고 분석하는 방법과 마찬가지로 중요하다. 편향, 특정한 자료를 감추려는 기도, 자료로 오도하려는 기도, 한 관점에서부터만 자료를 제시하는 기도, 통계치의 모든 그러한 오용과 남용은 학생들을 위한 자료처리 경험의 중요한 부분이 되어야 한다(우정호, 2000, p.19). 요즘은 언론이나, 정부 기관, 각종 단체에서는 매일 통계자료에 근거한 주장들을 쏟아내고 있는데, 심각한 것은 정확하지 않은 자료(예: 가구별 물 사용량=가구별 수도요금/노후 주택일 수록 이 차이가 큼), 잘못 사용된 통계 절차나 방법 등(용어 오류(취업률, 물가), 변인통제 오류(confounding effect))이 붓물을 이루고 있음을 생각할 때 학교 통계교육의 중요성을 새삼 실감케 한다(이영하, 2007, p.4). 또한 ‘정보(information)와 자료(data)’는 오늘날 우리의 사회에 상식화된 용어이다. 통계의 가치는 의미있는 정보를 추출하는 데에 있다. 따라서 자료와 정보의 개념적 차이를 잘 이해하고, 궁극적으로 의미있는 정보를 추출하는 활동이 학교의 통계 학습의 지향점이 되어야 하며, 기계가 대신할 수 있는 계산 과정에 머물러서는 안된다(이광복 외, 2004, p.254).

하지만 학생들은 일반적으로 형식적인 통계적 처리에 익숙하고 통계적 사고의 특성에 대한 의식이 결여되어 있다. 그 결과 일상생활에서 광범하게 사용되고 있는 통계치가 부정확하고 오용되는 경우가 있음에도 불구하고 학생들은 통계치에 대한 비판적 안목이 미약하다. 그

러므로 통계는 대중교육으로서의 수학교육의 중심적인 내용의 하나가 되어야 하며 통계적 사고⁵⁾는 학교수학에서 보다 더 강조되어야 한다. 현대 사회의 모든 시민은 통계의 사회적 역할과 가치를 인식하고, 그 힘과 한계를 이해해야 하며, 양적인 정보를 이용하여 생산적인 의사결정을 할 수 있도록 통계를 배워야 한다(우정호, 2000, pp.10-22). 중등 수학에서의 통계 단원의 가장 중요한 목표는 귀납적 사고 방법을 익혀 이를 실생활의 다양한 상황에 맞게 활용할 수 있도록 하는데 있다. …(중간생략)… 새로운 사실의 발견, 창조의 출발점은 자료이기 때문에 통계교육에서 가장 중요하게 생각할 것은 자료를 소중히 여기는 태도일 것이다. 이 과정에서 자료를 소중히 여기는 태도, 깨끗하고 신뢰할 만한 자료 확보의 중요성, 통계분석과 수집과정의 관계, 통계적 속성의 비 결정론적 이해 등을 교과과정에 소개되어야 한다(이영하, 2002, pp.353-357). 우리나라 중학교 영역에 해당하는 미국의 K9-K12학년에서는 NCTM의 기준 2⁶⁾인 ‘자료를 분석하기 위한 적절한 통계적 방법을 선택하고 사용할 수 있다’라는 목표에서 올바른 통계적 방법의 중요성을 인식시키고 있다고 본다. 학생들이 학교를 졸업하고 실생활에서 여러 정보(예를 들어, 투표율, 범죄율, 인구 증가율, 질병의 확산속도, 산업생산량, 교육 성취도, 고용추세 등)를 잘 이해하고 또는 선택을 해야 하는 상황에 닥쳤을 때 통계적 방법을 적용하기 위해서는 자료처리만이 아니라 자료의 수집이나 절차의 중요성에 대해서도 인식을 하고 있어야만 결과에 대한 올바른 이해, 선택을 할 수 있을 것이다. 즉 민주시민으로써 통계문맹⁷⁾ 관련능력을 가지고 생활할 수 있을 것이다. 본 연구에서는 7-나 통계

5) 측정된 데이터를 가지고 미지의 불확실성을 가진 현상을 객관적으로 탐구하고자 하는 귀납적인 사고(김웅환, 이석훈, 2007, p.24)

6) 류희찬 외(2007). 학교수학을 위한 원리와 기준, p.463. 서울:경문사.

단원을 분석함에 있어서 통계 문맹 관련 모든 사항을 살필 수도, 그럴 필요도 없으므로, 다만 정확한 자료의 중요성 즉 자료를 목적에 맞게 조사하였는지, 자료의 합목적성과 자료를 통계적 방법을 이용하여 분석하는데 있어서 목적에 맞게 분석하는 즉 합목적적 분석을 하는 것을 얼마나 강조하고 있는지 살펴보려는 것이다.

III. 연구방법 및 분석(내용)

1. 분석대상 및 분석방법

본 연구에서 분석하고자 하는 중학교 7-나 교과서는 모두 16종⁸⁾으로 분석대상교과서를 다음 내용의 서술방식과 관점에 따라 하나씩 읽어보아 유사한 것끼리 분류하여 각 내용의 서술방식과 관점을 상호·비교하였다. 상호·비교하는 주제별 관심내용은 다음과 같다.

<표 III-1> 분석틀

분석내용	
1. 도수분포표	① 분포의 의미 ② 분포표 작성목적
2. 분포의 표현	① 막대그래프와 히스토그램 ② 도수분포다각형 ③ 평균
3. 분포의 비교	① 상대도수와 그래프 ② 누적도수와 그래프
4. 통계적 방법	① 자료수집의 목적과 정확성 ② 통계적 방법의 타당성

위의 4개의 분석내용에서 도수분포표, 분포

의 표현은 7차 교육과정 7-나 통계단원에서 명시한 주제에서 추출하였다. 비교하는 주제별 관심내용에 대해 살펴보면, 첫 번째 도수분포표에서 분포의 의미란 단순히 도수분포표가 어떻게 구성되고, 어떻게 만드는 지에 관한 서술에 그치는 것이 아니라 '분포'라는 개념을 형성시키기 위한 과정, 즉 분포란 무엇이고 그 개념의 한 외연으로서 도수분포표를 이해시킬 수 있는 준비를 하고 있는 지 그 여부를 살펴보려는 것이다. 또한 도수분포표 작성의 목적이 어떻게 기술되어있는지 보려는 것이다. 앞서 분포의 의미가 이해되었다고 할 때 이 의미와 관련된 도수분포표 작성의 과정 못지않게, 어떤 자료에서 많은 자료 값들의 분포 상태를 잘 나타낼 어떤 도구를 생각할 수 있는지, 그 맥락에서 왜 도수분포표를 작성하며, 그것이 왜 필요하고 중요한지 등에 관한 서술방식을 살펴보려는 것이다.

두 번째 앞에서 언급한 것처럼 분포 개념의 표현들 중의 하나인 히스토그램과 도수분포다각형을 살펴보고, 히스토그램과 도수분포다각형을 단순히 그리는 방법만을 습득하는 것이 아니라 분포를 표현하는 그래프로서의 서술여부와 서술 방법을 살펴보려는 것이다. 또한 초등학교에서 배운 이산변량의 막대그래프와 이 단원에서의 연속변량의 히스토그램과의 연결이 어떻게 구성되고 있는 지 살펴보고자 한다. 분포의 표현에서 분포를 한 숫자로 나타내 주는 평균에 대하여 도수분포표에서 계급값과 도수의 곱을 전체의 도수로 나누는 계산공식만을 학습하는 것이 아니라 분포의 위치를 파악하는 하나의 값으로

7) Wallman(1993)의 statistical literacy란 통계적 사고가 공적, 사적 부분과 개인적, 전문적 영역에서 내리는 결정에 공헌하는 바를 이해하고, 일상생활에서 흔히 접할 수 있는 통계적 결과를 이해하며 비평적으로 평가할 수 있는 능력(D. Ben-Zvi and J. Garfield, p.48. 2004). 통계문맹은 statistical literacy의 번역이 곤란하여 statistical illiteracy로 고쳐서 번역한 것임

8) M1-M16의 분류는 최지안(2008, p.31)의 학위논문에서의 분류코드와 같음. 여기서는 익명성을 위해 세부내용은 밝히지 않기로 함.

나타내는 평균을 살펴보려는 것이다.

세 번째 한 집단, 혹은 도수의 합이 같은 두 집단에서의 자료의 분포 상태를 파악한 후에는 도수의 합이 다른 두 집단의 자료의 분포 상태를 비교하기 위해서 전체에서의 비율, 즉 상대도수의 개념을 이용하여 두 자료의 분포의 비교가 이루어져야 한다. 따라서 ‘분포의 비교’관점에서 상대도수의 그래프를 살펴보려는 것이다. 더불어 누적도수의 그래프도 어떤 변량이 전체에서 차지하는 위치, 혹은 분포의 형태를 파악하고 두 개 이상의 분포의 형태를 상호 비교할 수 있는 서술내용과 그 방법 여부를 살펴 보려는 것이다.

네 번째 앞에서 자료의 분포 표현, 분포의 비교를 하기 위해서 정확한 자료의 수집이 필수이다. 학생들이 앞으로 실생활에서 접할 자료에서는 정확한 자료의 중요성을 아무리 강조하여도 지나치지 않을 것이다. 또한 정확한 자료의 중요성과 함께 자료들을 목적에 맞게 분석하는 통계적 방법의 타당성도 중요하다. 따라서 목적에 맞는 결과해석을 위한 통계적 분석방법의 타당성을 살펴봄으로써, 실생활에서 일어나는 오해와 편견을 막고 통계문맹을 벗어나게 하는 하나의 준비과정으로서의 서술의 충실성을 살펴보려는 것이다.

2. 분석내용과 결과

가. 도수분포표

1) 분포의 의미

도수분포표의 정의에서 분포의 의미에 대한 서술여부를 확인해 본 결과 16종의 교과서는 분포의 의미를 모두 서술하였고, 분포의 의미 서술방법을 정리하면 “주어진 전체 자료를 몇

개의 계급으로 나누고, 각 계급에 속하는 도수들을 구한 후, 이것을 표로 나타낸 것”이다. 도수분포표에서 도수분포표의 구성요소를 소개하기 위해 계급, 도수 등을 설명하고 있지만 16종의 교과서 모두가 이것이 분포라는 것의 외연중의 하나임을 분명히 드러내고 있지 않다. 분포의 의미는 ‘범주와 빈도의 순서쌍들의 집합’으로 서술될 수 있지만 7-나 단계에서는 다소 무리한(이해하기 힘든) 표현이며, 다만 7-나 단계에서는 분포와 분포가 아닌 것의 차이만 구분가능하면 충분하다고 본다. 그러나 분석대상이 되는 16종의 교과서 모두 이것은 서술되어 있지 않다. 초등학교에서 막대그래프 관련 도수분포표를 만들어 보지만 이것은 이산변량이므로 ‘분포’의 개념형성을 위한 ‘형식화’ 즉 ‘구체적 자료’와 ‘형식적 분포’사이의 가역적 사고를 위한 수업은 이루어지지 않았다고 추측되며 형식화된 분포개념은 7-나 단계에서 가르쳐야 할 통계단원의 핵심과제라 할 수 있다. 그러나 7-나 단계 교과서 16종에서 역시 이점을 분명히 드러낸 것은 없다. 최소한 ‘분포’인 것과 아닌 것의 차이를 서술해 주어야 이 점이 분명히 드러난 것으로 생각된다. 또한 도수분포표의 정의에서 ‘주어진 전체 자료’로 분류한다 하면 그 뜻이 모호하여 분포가 아닌 경우⁹⁾에도 그럴 수 있는 지 분명치 않으므로 ‘주어진 전체 자료’보다는 ‘측정값’을 범주별로 분류한다는 것이 더 명확한 표현일 것이다. 따라서 분포의 의미를 도수분포표의 정의에서만 그칠 것이 아니라 좀 더 명확하게 명시해야 비로소 분포인 것과 아닌 것을 구분할 수 있고, 이후 도수대신 확률로 표현되는 확률분포를 도수분포와 연결시켜 이해할 수 있을 것이라 생각한다. 또한 7-나에서의 도수분포표는 분포개념의 외연적 표현들(표, 막대그래프, 히스토그램, 도

9) 예) 날짜별 평균기온을 적은 표

수분포다각형 등)중의 하나일 뿐이다. 다음으로 도수분포표 작성목적에 관한 서술여부와 서술 방법을 살펴보았다.

2) 분포표 작성목적

도수분포표 작성목적의 서술방법을 살펴보면 교과서는 크게 2가지 유형-자료의 분포상태 파악을 서술한 교과서는 13종, 집단 전체에서의 위치, 집단 전체의 경향 파악으로 서술한 교과서는 3종-으로 나눌 수 있다. 도수분포표 작성목적이 명확하게 서술된 교과서¹⁰⁾도 있지만 대부분의 교과서는 간접적으로 서술되어 있으므로 도수분포표 작성목적은 좀 더 명확하게 제시할 필요가 있다. 도수분포표의 작성목적으로 서술한 '집단 전체의 위치나 그 집단 전체의 경향'에서 차지하는 위치에 대한 설명이 모호하며 집단 전체의 경향, 집단 전체 기록의 성격, 기록의 경향이 무슨 경향, 어떤 성격을 말하는 것인지 또한 명확하지 않다. 이것은 분포표 작성목적은 '분포상태의 표현'으로 한 것 보다 모호하고 포괄적인 것으로 볼 수 있다. 또한 앞에서 살펴본 것처럼 '분포상태의 표현'으로 했더라도 '분포상태'가 무엇인지 설명이 없으면 이것 역시 모호한 서술이라 할 수 있을 것이다. 엄밀히 말하면, '집단 전체에서 차지하는 위치나 집단 전체의 경향'을 알아보는 것은 도수분포표 작성목적이라기보다는 도수분포표를 작성한 후 분포표의 활용목적에 가깝다고 볼 수 있을 것이다.

16종 모든 교과서는 자료에서 제시한 문제를 통해 "...자료를 목적에 맞게 정리할 필요가 있다."고 한다. 이 때 그 목적이란 것이 자료의 범주의 빈도(도수), 자료의 분포를 구하는 것이고 서술방법의 차이는 있지만 결국 자료의 분

포상태를 파악하는 것이 도수분포표 작성목적임을 알 수 있다. 도수분포표를 만드는 방법이 중요한 것이 아니라 자료를 정리함으로써 '자료의 분포상태'를 파악하는 것, 즉 도수분포표 작성의 목적이 중요한 것이다. 그러나 모든 교과서에서는 교과서의 문제를 통해 자료의 '분포'를 묻지만 '분포'의 의미를 명확하게 설명하지 않고 도수분포표의 정의에서 '분포'의 의미를 암묵적으로 포함하고 있다. 또한 '분포'와 '분포상태'의 의미를 명확히 구분하고 도수분포표 작성목적인 자료의 분포상태를 알기 위해 '분포상태'의 의미도 명확하게 제시되어야 할 것이다. 따라서 도수분포표에서 자료의 분포상태의 의미의 서술여부에 따라 교과서를 살펴보면, 자료의 분포상태 의미를 명확하게 서술한 교과서 1종, M2는 자료의 분포상태 의미를 분명히 드러내어 서술하였다.

도수분포표를 만들면 자료의 분포 상태, 즉 어떤 값 근방에 변량이 밀집되어 있고, 어떤 값 근방에 변량의 개수가 적은지를 쉽게 알 수 있다(p.7).

(1)분포의 의미에서 언급한 것처럼 16종의 모든 교과서는 도수분포표 정의에서 분포의 의미는 서술된 것으로 보았다. 3종의 교과서(M7, M10, M13)에서는 도수분포표의 정의에서 '분포상태'를 기술하였으나 분포상태의 의미를 명확히 제시하지 않았다.

자료의 분포상태의 의미를 명확히 서술한 교과서 M2를 제외한 모든 교과서에서 도수분포표 작성목적, 도수분포표의 정의에서 '분포상태'를 직접적, 혹은 간접적으로 서술하였지만 분포상태의 의미는 명확히 보이지 않고 있다. 결국, 우리나라 교과서는 '자료의 분포상태를

10) ...의 분포 상태를 알아보기 위하여(M8, M10)

알기위해서' 도수분포표를 만드는 것이라고 하면서 분포 혹은 분포상태라는 용어는 제시하고 있지만, 결국 '분포상태의 의미'를 명확하게 제시하고 있지 않다. 교육부(1999)는 도수분포와 그래프 단원에 대하여 "통계 자료는 적절한 계급구간을 사용하여 자료들을 정리하면, 그 자료의 특징을 어느 정도 파악할 수 있다(p.64)."고 하였다. 이 때 자료의 특징은 관점에 따라 가운데(평균, 최빈수 등), 산포도(표준편차 등), 분포상태(표, 그래프), 최대값, 최소값, 첨도, 왜도 등 다양하다.

하지만 본 논문의 연구대상인 중학교 7-나에서 '도수분포표의 정의와 만들기'에서 특징을 살펴보는 것이므로 특징이 도수분포표의 '분포상태'에 관한 것으로 설명한다. 현행 모든 교과서에서는 계급의 개수(크기)를 어떻게 정하느냐에 따라 분포를 파악하기 쉽거나 어려운 도수분포표가 될 수 있음을 보여주고 있다. 김원경외(2003)는 다음과 같이 서술하고 있다.

도수분포표의 작성에 있어 계급의 수는 계급의 폭과 분포의 모양을 결정하므로 매우 중요하다 (p.51).

따라서 16종 교과서를 살펴본 결과, 모두 계급의 개수(크기)를 어떻게 결정하는지에 대하여 서술하였고 그 서술방법을 살펴보면, 계급의 수를 명확히 몇 개로 구하는 방법 제시한 교과서는 1종, M2(p.8)는 계급의 개수(크기)를 정할 때, 계급의 수를 명확히 몇 개로 구하는 방법을 아래와 같이 설명하고 있다.

이때면, 변량이 50개인 자료에서는
 $2^1=4, 2^2=8, 2^3=16, 2^4=32, 2^5=64$
 에서 64가 처음으로 변량의 개수 50보다 크므로 계급은 6개로 정한다.

그 외 교과서 15종은 계급의 개수(크기)가 다

른 2개의 도수분포표를 보여주면서 "계급의 크기를 너무 크게 하거나 작게 하면 자료 전체의 분포상태를 알아보는 데 불편하다. 일반적으로 계급의 개수는 자료의 양에 따라 달라지지만, 5~15개 정도가 되도록 하는 것이 좋다."고 설명하고 있다. 통계영역은 중등 수학의 다른 영역과 달리 통계적 사고인 귀납적 사고능력을 기르는 것이다. 따라서 계급의 개수가 작거나 많으면 분포를 파악하기 어려우므로 "계급의 개수는 5~15개 정도로 한다."라고 정리하기 전에 컴퓨터, 계산기를 사용하여 계급을 나누는 활동을 통하여 계급이 결정될 때 나타나는 자료의 분포를 보고 계급의 개수를 결정하는 과정을 학생 스스로 경험할 수 있도록 하는 것이 학생들의 통계적 사고, 즉 귀납적 사고능력을 길러주는데 도움이 될 것이다. 물론, 컴퓨터를 사용한다 해도 처음 컴퓨터를 이용하여 계급을 정할 때, 계급의 수를 몇 개로 명확히 정할 필요가 있을 때는 사실상 Sturges방법이라고 할 수 있는 $2^{k/5}n$ 인 k 로 정하는 방법을 사용하는 것이 더 효율적이라 볼 수 있을 것이다. 또한 자료의 분포상태를 잘 나타나기 위해서 계급의 개수(크기)를 정하는 것이므로 여기서 '좋은 분포상태'가 무엇인지에 대한 설명이 필요하다. '좋은 분포상태'란 분포상태가 잘 드러난 것, 즉 변량 수가 많은 부분과 적은 부분을 간단하게 말할 수 있는 것이다. 가령 히스토그램이 높이가 일정치 않게, 들쭉날쭉한 모양보다는 가급적 한, 두 곳에서 높고 다른 곳은 낮은 높이를 갖는 것이 더 좋은 표현이다. 그러나 이것은 자료마다 특성이 있어 어떤 규칙을 말하기 어려우며, "아름다운 그림"과 같이 수학적 서술이 어려운 측면이 있다. 따라서 중학교 1학년에게 너무 어렵고 복잡한 자료를 피하는 것이 좋다고 한다면 히스토그램의 높은 부분은 한 개이고 다른 부분은 점진적으로 낮아지는 형태를 갖는 자료만 다루는 것도 고려해 볼 필요가

있다. 실제 통계학에서 가장 많은 자료는 단일모드(unimodal)분포이기 때문이다.

나. 분포표현에 관한 특징적 기술

1) 막대그래프와 히스토그램

분포의 표현에 대한 서술여부를 살펴본 결과 16종의 교과서는 직접적으로 ‘분포의 표현’이라고 하지 않았지만 ‘히스토그램 그리는 방법을 제시한 후 히스토그램을 정의’하는 방법으로 분포의 표현을 간접적으로 서술하고 있다. 도수분포표에 비해 히스토그램이 도수가 가장 큰 계급, 가장 작은 계급 등 각 계급에 속하는 자료의 수가 많고, 적음을 한눈에 알 수 있고 찾기 쉽고 편한 것으로 히스토그램이 도수분포표보다 자료의 전체적인 분포상태를 더 쉽게 표현한다는 것을 알 수 있다. 또한 히스토그램에 대해 도수를 직사각형 그림으로 나타낸 ‘한 눈으로 볼 수 있는 그림(M7)’이라고 표현했다. 이것은 단순히 히스토그램을 그리는 방법, 순서보다는 히스토그램을 통해 자료의 분포를 나타내고, 분포상태를 파악하는 것이 중요한 것이다. 한편 히스토그램은 연속변량을 다루는 그래프이기 때문에 이산변량을 다루는 막대그래프와 차이가 있다. 따라서 막대그래프와 히스토그램의 차이를 설명하고 있는 교과서를 살펴보았다. 분석 결과 막대그래프와 히스토그램의 차이를 서술한 교과서 4종에서 교과서 M6(p.15)은 다음과 같이 이산변량의 막대그래프와 연속변량의 히스토그램의 차이점을 서술하고 있다.

히스토그램과 막대그래프의 차이점

히스토그램은 각 직사각형의 가로가 계급의 크기로 정해져 있지만, 막대그래프는 막대의 폭에는 의미가 없다.
따라서, 막대의 폭은 그리기 쉽게 적당히 잡는다.

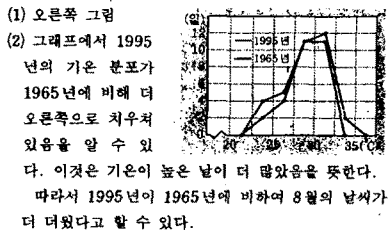
표보다 그래프가 더 자료의 분포상태를 알아보기 쉽다는 것을 보이기 위해 3-나에서 배운 막대그래프를 상기시킬 때, 이산변량을 나타내는 막대그래프와 연속변량을 나타내는 히스토그램의 차이점을 이해할 수 있도록 교과서가 구성되어야 할 것이다. 하지만 7차 교육과정에는 이산변량과 연속변량의 구분에 대한 명백한 안내가 없다. 따라서 교과서에 앞서 교육과정이 개선될 필요가 있다고 할 수 있다.

2) 도수분포다각형

분포의 표현에 대한 서술여부를 살펴본 결과 16종의 교과서는 히스토그램과 함께 도수분포표에서 언급한 것처럼 분포의 외연적 표현의 방법들 중의 하나로 분포의 표현을 ‘도수분포다각형 그리는 방법 제시 후 도수분포다각형 정의’로 간접적으로 서술하고 있다. 도수분포다각형은 히스토그램과 함께 ‘한 눈에 들어올 수 있는 또 하나의 그림(M7)’이라는 의미에서 히스토그램과 같이 자료의 분포상태를 표현한 것이다.

교과서 M1에서는 “(3) 등교 시간이 10번째로 오래 걸리는 학생이 속하는 계급을 말하여라 (p.16).”의 질문은 도수의 위치를 묻는 문제이므로 누적도수의 그래프에서 더 적절한 질문일 것이다. (1)막대그래프와 히스토그램에서 언급한 것처럼 교육부(1999, p.66)에서도 명시하고 있지만 교육과정에 반드시 따라야 하는 것은 아니므로 위의 질문은 누적도수의 그래프에서 더 적절한 질문이라고 할 수 있지만 도수를 나타낸 히스토그램, 도수분포다각형에서도 가능한 질문이다. 그러나 히스토그램을 통해 알 수 있는 것을 다시 도수분포다각형에서 묻는다는 것은 도수분포다각형을 학습해야 하는 동기를 저하시킬 우려가 있다. 또 후에 누적도수의 학습동기도 저하될 우려가 있다. 결국 도수분포다각형 학습의 특징적 목표가 서술되지 않았다

고 할 수 있다. 그래프 형태의 특징은 서술되고 있지만 그런 형태가 필요한 특징적 목표는 제시되지 않은 것이다. 따라서 도수분포다각형의 특징에 대한 서술여부를 살펴본 결과 도수분포다각형의 특징에 대해 서술한 교과서 3종이 “2개 이상의 자료의 분포를 비교할 때, 도수분포다각형을 한 모눈종이에 그래프로 나타내며 비교”로 [그림 III-1]처럼 도수의 합이 같은 두 집단에서 도수분포표보다는 도수분포다각형을 이용하는 것이 두 자료의 분포상태를 비교, 파악하기가 더 쉽다는 것을 보여주고 있다.



[그림 III-1] 도수의 합이 같은 두 집단의 도수분포다각형(M2, p.15)

도수분포다각형은 ‘히스토그램을 유연화시킨 개념¹¹⁾’으로서 두 개 이상의 자료의 분포 상태를 비교하는 목적으로 유용하다. 자료의 분포 상태를 쉽게 파악하기 위해서는 도수분포표보다는 히스토그램, 도수분포다각형을 이용하는 것이 도움이 된다는 것을 보여주고 있다. 도수의 합이 같은 두 집단의 자료의 분포상태를 비교하는 데 있어서 도수분포표보다 도수분포다각형이 더 효과적인 것을 알 수가 있다. 한 자료의 분포를 파악할 때는 한 모눈종이 위에서 히스토그램 위에 도수분포다각형을 그려봄으로써 히스토그램처럼 도수분포다각형도 자료의 분포상태를 파악하기 쉽다는 것을 보여주지만

두 집단의 자료의 분포상태를 비교할 때 히스토그램을 왜 사용하지 않는지에 대해 기술하는 교과서는 16종 모두 없다. 이것은 학생들에게 좀 더 친절하게 도수분포다각형의 필요성을 서술함에 있어, 다른 원인, 가령 교과서 쪽 수의 제한 등이 원인 일 수 있다고 생각된다. 또한 김원경 외(2003)는 아래와 같이 설명하고 있다.

히스토그램이나 도수분포다각형은 도수분포표를 일목요연하게 나타낸 그래프이다. 따라서 두 개의 그래프 중에서 하나만 그려도 자료의 분포상태를 충분히 파악할 수 있다. 그러나 히스토그램을 그리고, 또 도수분포다각형을 그리는 이유는 앞으로 키, 몸무게 등과 같이 연속적인 값을 취하는 변수에 대한 확률밀도함수를 도입하기 위한 것이다(p.55).

그러나 이와 같은 도입을 위한 어떤 준비나 이후 고등학교 교육과정에서 도수분포다각형에 관한 언급을 찾을 수 없다는 점에서 이 부분에 관한 연구나 개선이 필요해 보인다.

3) 평균

16종의 교과서에서 분포의 위치에 대한 서술 여부를 살펴보았다. 분석결과 분포의 위치를 서술한 1종, 교과서 M2는 “가로축에 축구 대표 선수들의 평균 키의 위치를 나타내어라(p.13).”로 질문을 하고 있다. 앞에서 언급한 자료의 분포의 표현인 히스토그램의 가로축에 평균을 표시함으로써 평균이 자료(축구 대표 선수의 키)의 분포에서 어느 위치에 있는지를 알게 하여 평균이 중심경향값임을 느끼게 하는 과정임을 분명히 알 수 있게 해 준다. 도수분포표보다 자료의 분포상태를 좀 더 쉽게 파악, 비교하는 방법으로 히스토그램과 도수분포다각형을 서술하였고, 이

11) 황혜정, 나귀수, 서동엽(2000). 수학과 교육목표및 내용 체계화 연구, p114. 한국교육과정평가원연구보고 PRC2000-3

때 도수분포표, 히스토그램, 도수분포다각형을 통해 자료의 분포상태, 분포의 위치를 파악할 수 있다. 하지만 표나 그래프로 나타낸 분포를 한마디로 간단하게 설명하기란 쉽지 않다. 따라서 분포상태를 모두 다 설명할 수는 없어도 최소한 자료의 값들이 어느 정도의 값인지 간단히 설명할 수 있는 대푯값이 필요하고 그 대푯값들¹²⁾ 중에서 가장 많이 사용되는 것이 평균이다.

한편 평균은 변량을 크기순으로 배열했을 때, 자료의 중간 근방을 나타내므로 자료의 분포상태를 알아볼 때 대푯값, 즉 분포의 위치를 한 개의 숫자로 나타낼 때 평균을 이야기한다. 따라서 평균은 자료의 분포상태와 분포위치를 간단히 나타내는 하나의 숫자라고 할 수 있다. 자료의 분포상태를 나타내는 도수분포표에서 평균은 자료의 분포상태를 모두 설명할 수 없지만 한마디로 요약 표현하는 방법들 중에서 하나의 방법이라는 사실을 나타내야 할 것이다.

그러나 자료에 따라 다른 변량과 구별되는 특이값(지나치게 크거나 작은 값)이 존재할 경우에는 평균이 자료의 중앙에서 멀리 벗어날 수도 있다. 이런 경우에는 중앙값이 더 적절한데, 평균은 분포의 가운데에 위치하지 않지만 중앙값은 이런 경우에도 여전히 분포의 중앙부근에 위치하기 때문이다. 따라서 분포의 중앙을 나타내기 위해 구한 평균이 정말 중앙을 잘 나타내고 있는지 확인할 필요가 있으며 그와 같은 확인과정 자체에 대한 학습과 대개의 경우 평균이 분포의 중앙을 나타냄을 지도할 필요가 있다.

다. 분포비교에 관한 특징적 기술

1) 상대도수와 그래프

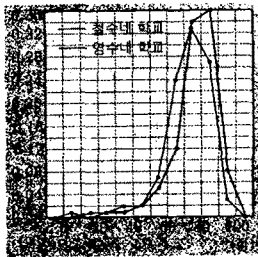
16종의 교과서를 분석한 결과 분포의 비교에

대해 모두 서술하였다. 상대도수의 필요성을 서술한 교과서 9종에서 상대도수의 학습의 특징적 목표를 분포의 비교로 분류한 교과서 1종은 단순히 비율, 도수분포표에서 상대도수를 구하는 계산방법만을 제시하는 것이 아니라 도수의 비교만으로 판단할 수 없을 때, 두 도수의 비율, 상대도수를 비교하고 있다. 즉 도수의 합이 다른 두 자료의 분포를 비교하는 방법, 상대도수 학습의 특징적 목표를 경우의 수로 설명하고 있다. 단순히 비율의 비교를 하는 데 편리하다로 분류한 8종의 교과서는, 같은 자료에서 두 계급의 비교를 상대도수로 설명하고 있는데 이것은 도수로 비교하여도 쉽게 파악할 수 있으므로 두 계급의 상대도수로 비교하면 쉽게 알 수 있다는 것은 도수의 합이 다른 두 자료의 분포를 비교하는 방법에 대한 상대도수 지도를 하기 전의 준비로 해석된다. 또한 ‘자료를 비교하기에 편리하다’고 언급하였지만 ‘30회 이상 40회 미만’과 같이 한 계급의 상대도수를 구하는 것과 상대도수의 합이 1이 됨을 설명한 것에 그치고 두 계급의 상대도수의 비교에 대해서도 설명하지 않았다. 결국 8종의 교과서는 ‘자료를 비교하기에 편리하다’는 설명만 덧붙여 상대도수 학습의 특징적 목표를 서술하기보다는 상대도수의 계산방법만 제시한 것이라고 볼 수 있다. 상대도수의 그래프 지도의 특징적 목표의 서술여부를 살펴보면, 16종 교과서는 모두 심화과정에서 상대도수의 그래프 지도의 특징적 목표를 서술하고 있다. 두 분포의 비교는 조건부분포의 비교로 이어져 상관개념의 형성에 직접적 영향을 주는 주제이므로 이에 관한 지도는 학습내용 구성에서 매우 중요하다고 할 수 있을 것이다.

심화과정에 앞서 나머지 16종의 교과서 내용을 살펴보면, 모든 교과서는 ‘분포상태를 알아

12) 중앙값, 최빈값

보기 쉽다', '비율을 비교하면 차이점을 쉽게 알아볼 수 있다'처럼 '분포상태'를 알려주는 방법으로 도수분포표, 히스토그램, 도수분포다각형에 이어 상대도수의 히스토그램과 분포다각형으로 방법을 제시하고 있다. 또한 도수로 비교해도 분포상태를 파악할 수 있는 것을 상대도수의 히스토그램과 도수분포다각형을 설명하는 것은 앞서 학습한 그래프(히스토그램과 도수분포다각형)와 그래프의 학습목표가 반복되는 상황을 경험하는 학생들에게 학습동기를 저하시킬 우려가 생길 수 있다고 본다. 따라서 도수분포표, 히스토그램, 도수분포다각형과 차별된 상대도수 학습목표의 특징적 목표를 명확히 제시할 필요가 있다고 본다. 심화과정에서 16종의 모든 교과서는 도수의 합이 다른 두 자료를 비교할 경우 도수의 합이 다른 경우에는 도수만으로 자료의 분포를 비교하는 것은 의미가 없고, 각 계급의 상대도수를 이용하여 분포상태를 비교하는 방법을 제시하고 있다. 즉 상대도수의 그래프 지도의 특징적 목표, 도수의 합이 다른 두 자료의 분포의 비교를 서술하였다. 이 때 한 모눈종이에 상대도수의 분포다각형모양의 그래프를 제시하고 설명한 교과서는 12종¹³⁾에서 6종의 교과서는 그래프를 제시하고 있다.



[그림 III-2] 도수의 합이 다른 두 집단의 상대도수의 분포다각형(M2, p.21)

도수의 합이 다른 두 집단의 분포상태를 비

교하기 위해서는 상대도수를 구하여 서로 비교하고, 비교를 좀 더 합리적으로 하기 위해서는 상대도수를 구하여 상대도수의 분포표를 만들어 비교하고, 두 집단의 상대도수의 분포표를 한 모눈종이에 그래프로 나타내면 두 집단의 분포 상태를 보다 쉽게 비교할 수 있다.

7차 교육과정에는 상대도수의 그래프 지도의 특징적 목표, 도수의 합이 다른 두 자료의 분포의 비교가 심화과정 내용으로 한정되어 있다. 도수의 합이 다른 자료의 분포 비교를 하는 상대도수의 학습목적이 심화과정에만 제한되는 결과를 초래한 것으로 보인다.

16종 교과서에서 모두 밀도도수라는 용어는 사용하지 않았지만 문제로 다룬 예(M2), [그림 III-3]의 문제가 있었다.

● 오른쪽 그래프는 앞 쪽 예제 1의 원수대 한 학생들의 수학 성적을 나타낸 그래프에서, 각 계급의 상대도수를 계급의 크기로 나눈 값을 세로축으로 하여 그린 히스토그램과 분포다각형이다. 다음 질문에 답하여라.

- (1) 위의 히스토그램에서 각 직사각형의 넓이는 무엇을 나타내는가?
- (2) 위의 히스토그램에서 직사각형 전체의 넓이는 얼마인가?
- (3) 위의 분포다각형의 넓이는 얼마인가?

[그림 III-3] 밀도도수 히스토그램(p.65)

위의 문제(밀도도수 히스토그램)에서 한 개의 직사각형의 넓이(계급의 크기×밀도도수)는 상대도수를 나타낸다. 따라서 밀도도수의 히스토그램에서 직사각형 전체의 넓이는 1이 된다. 이 때 계급의 수가 한없이 늘면 계급의 크기는 대부분 0으로 수렴하게 되고, 직사각형 모양은 매우 가는 띠 모양이 되는데 그렇게 되면 결국 밀도도수의 히스토그램은 직사각형들의 계단식 모양에서 부드러운 곡선모양으로 변하게 된다.

13) 문제로만 제시하고 그래프를 보여주지 않은 교과서 6종

이 곡선 아래의 넓이는 여전히 1이며 이 곡선을 나타내는 함수가 확률밀도함수이다. 이와 같은 설명이 먼저 이루어진다면 확률밀도곡선 아래의 넓이가 상대도수와 유사한 통계적 확률이 됨을 이해하는 것이 쉽게 이루어질 것이다.

2) 누적도수와 그래프

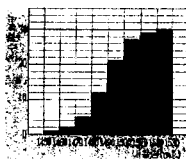
16종의 교과서를 분석한 결과 분포의 비교에 대해 모두 간접적으로 서술하였다. 누적도수의 필요성을 명확히 서술한 교과서 13종에서 누적도수 학습의 특징적 목표에 대한 서술방법을 살펴보면, 10종의 교과서는 처음 계급부터 어떤 계급(연속된 계급)까지의 도수의 합으로 누적도수 학습의 특징적 목표를 서술하였다. 이것은 ‘처음 계급부터 어떤 계급까지의 도수의 합’은 누적도수의 정의이므로 정의를 학습한다는 것을 학습의 목표로 하는 것은 다소 무리가 있어 보인다.

자료 전체 중 어떤 대상이 차지하는 위치로 설명한 3종의 교과서는 교육부(1999, p.66)에서 언급한 것처럼 누적도수 학습의 특징적 목표인 ‘자료 전체 중 어떤 대상이 차지하는 위치’를 서술하고 있다.

누적도수의 그래프 지도의 특징적 목표에 대한 서술여부를 살펴보았다. 분석결과 누적도수의 그래프 지도의 특징적 목표를 명확하게 서술한 교과서는 10종 모두 ‘자료 전체 중 어떤 대상이 차지하는 위치’로 누적도수의 그래프 지도의 특징적 목표를 서술하고 있다.

16종의 교과서에서 교과서 1종, M5에서만 누적도수의 히스토그램을 설명하였다.

오른쪽 그래프는 (표-13)의 누적도수의 분포표를 히스토그램으로 나타낸 것이다. 도수분포표를 히스토그램으로 나타내면 분포의 상태를 쉽게 알아볼 수 있었던 것과 마찬가지로, 누적도수 그림으로 나타내면 누적도수를 쉽게 알 수 있다.

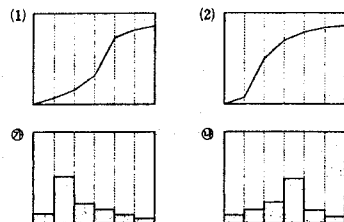


[그림 III-4] 누적도수의 히스토그램(M5, p.25)

이 단원이 연속변량의 분포표현, 비교 등에 관한 것이므로 표본누적분포함수 역시 연속함수 형태가 더 바람직하며 이 경우 계단함수([그림 III-4]의 히스토그램의 윗부분만 그려 나간 것)는 이산변량에 더 적합하다.

또한, [그림 III-4] 누적도수의 히스토그램의 마지막 계급 ‘1520~1530(mL)’까지 칠해진 부분은 히스토그램과 달리 누적분포의 그래프이므로 ‘1530(mL)’이후 모두 칠해져야 한다. 또 연속변량의 확률밀도함수가 누적분포함수를 미분하여 얻어진 것을 고려할 때, 구간별 선형누적분포함수의 그래프를 미분하면 계급별로 상수함수가 되고 이것은 앞에서 배운 히스토그램(직사각형의 윗부분만 그린)이 된다. 교과서 M2(pp.25-26)에서만 “누적도수가 앞 계급에 비하여 가장 많이 증가한 계급은 도수가 가장 큰 계급임을 알 수 있다.”고 하며 [그림 III-5]를 통해 그 가파른 정도에 따라 어떤 계급에서 가파르게 올라가면 그 계급에서 누적도수가 많이 증가한 것을 파악하는 형태의 문제를 제시하였다. 누적도수의 그래프는 히스토그램에서 그 계급의 직사각형의 높이가 높을수록 도수가 많기 때문이다.

다음은 두 개의 자료에서 얻은 누적도수의 그래프와 히스토그램이다. 같은 자료를 나타내는 것끼리 짝지어라.

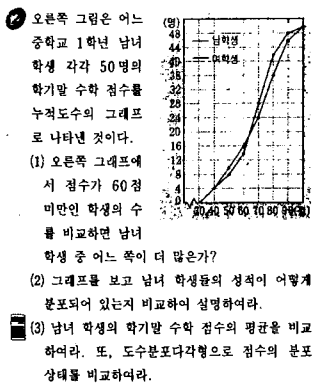


[그림 III-5] 히스토그램과 누적도수의 그래프(M2, p.28)

[그림 III-5]의 문제는 누적도수의 그래프와 히스토그램의 변환을 통해(7-나에서 미분을 학습하지 않으므로 앞에서 언급한 것처럼 가파른

정도에 따라 어떤 계급에서 가파르게 올라가면 그 계급에서 누적도수가 많이 증가한 것을 파악하는 것으로) 누적도수의 그래프와 히스토그램 사이, 즉 ‘형식적 분포’와 ‘형식적 분포’사이의 가역적 사고를 위한 것으로 보인다. 하지만 나머지 15종의 교과서에서는 이런 형식적 조작의 가역적 사고에 대한 서술이 없다. 또한 교과서 M2를 제외한 모든 교과서에서는 누적도수의 분포다각형에서 형식적 조작의 가역적 사고에 대한 서술이 없다. 학생들의 형식적 조작의 가역적 사고를 돕기 위해 관련 내용이 추가되는 것이 도움이 될 것이라 본다.

누적도수의 분포다각형 역시 자료의 분포를 비교하기 위한 것이므로 두 자료의 분포 비교에 대한 설명이 제시되어야 할 것이다. 따라서 두 자료의 분포 비교를 하는 누적도수의 그래프 서술여부를 살펴보았다. 분석결과 두 자료의 분포 비교에 대해 누적도수의 그래프를 제시한 교과서는 3종에서 2종은 두 자료의 분포 상태를 비교하기 위하여 도수의 합이 같은 두 누적도수의 그래프를 제시하였다.



[그림 III-6] 도수의 합이 같은 두 집단의 누적도수의 분포다각형(M2, p.30)

[그림 III-6]에서 보여준 누적도수의 그래프는 통계학, SPSS에서 P-P plot과 분포적으로 유사한 것이다. 원래 P-P plot은 두개의 누적도수의

그래프에서 한 쪽 그래프의 증가분을 x 축에, 다른 그래프의 증가분을 y 축에 나타내어 $y=x$ 의 그래프와 비교하는 것이다. 그러나 7-나 단계에서 이것을 소개하기 어려우므로 각 누적도수의 그래프를 포개어 그려서 비교하는 방법을 소개하는 것이라고 생각된다. 7-나 단계에서는 P-P plot을 소개하지는 않지만 그에 대한 준비를 하는 것으로 보아진다. 나머지 1종의 교과서는 도수의 합이 다른 두 집단의 누적도수분포표를 제시하고 그래프에 대한 문제를 제시하고 있다. 하지만 도수의 합이 다른 두 집단이므로 누적도수의 그래프가 아니라 누적상대도수의 그래프를 그리는 것이 더 적합하다고 볼 수 있다. 누적상대도수는 누적확률과 거의 같은 것으로 생각할 수 있기 때문이다. 현재까지 분포의 표현과 관련된 주제를 살펴보면 도수로부터 상대도수, 누적도수로 논의를 진행해왔으나 이 과정에서 상대도수의 역할이 모호하다. 연속변량에서는 상대도수보다는 밀도도수가 더 중요한데, 밀도도수에 관한 언급은 없다. 한편 상대도수 이후에 누적상대도수가 설명된다면 밀도도수와는 별개로 상대도수의 역할이 드러난다. 그런데 누적상대도수도 언급하지 않는다. 결국 상대도수의 역할은 도수의 합이 다른 분포의 비교라는 실용적 목적 외 이후의 학습주제와는 연결되지 않은 상태로 남게된다. 7차 교육과정에는 누적상대도수의 내용에 대한 것이 언급되지 않고 있다.

교과서 M2에서 언급한 것처럼 ‘어느 계급에서의 누적도수의 증가와 감소’를 알아본다는 것은 비교되는 자료와 비교하는 자료가 있어야 한다. 즉 도수분포표에서 분포개념의 외연의 한 방법인 그래프, 누적도수의 분포다각형에서 분포의 비교를 간접적으로 서술한 것이다. 또한 ‘표보다는 그래프가 더 편리하다’는 것은 역시 도수분포표보다 히스토그램, 도수분포다각

형이 더 편리하다고 설명한 것처럼 누적도수에
 서도 누적도수의 분포표보다는 누적도수의 히
 스토그램, 도수분포다각형이 분포 상태를 파악
 하기 쉽다는 것을 의미하고 있다. 따라서 교과
 서는 누적도수의 그래프를 그리는 방법 뿐 아
 니라 누적도수의 그래프를 보고 한 집단에서의
 자료의 위치 파악 및 두 자료의 분포의 비교와
 정확한 이해를 파악할 수 있도록 구성되어야
 할 것으로 보인다.

라. 이산변량의 그래프

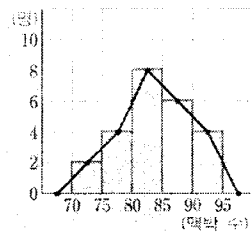
이번 절에서는 도수분포표, 히스토그램과 도
 수분포, 상대도수와 누적도수의 그래프에서 제
 시되는 자료를 살펴보았다. 7-나 단계 통계단원
 에서 소개하는 모든 내용은 초등학습 내용의
 복습이 아닌 한 모두 연속변량만 다루는 것이
 학생들의 혼란을 막는데 도움이 될 것이라고
 본다. 그리고 고등학교 이후에는 필요에 따라
 예외임을 분명히 하고 예외도 가르치는 방식이
 더 적절할 것이다. 첫 번째 통계단원에서 도수
 분포표는 히스토그램을 그리기 위해 만들기 때
 문에, 평균도 그런 경우(계급이 있어 계급값을
 사용해야 하는 경우)만 다루기 때문에 도수분
 포표를 만들 때 제시되는 자료¹⁴⁾를 살펴보면
 다음과 같다.

1 다음은 어느 중학교 여학생 30명의
 읽음일지기 측정 결과를 나타낸
 것이다. 5회부터 시작하여 계급의
 크기를 5로 하는 도수분포표를 만들
 고, 각 계급의 계급값을 구하여라.

읽음일지기(회)	도수(명)	계급값
5~		

[그림 III-7] 이산변량의 도수분포표(M3, p.11)

두 번째 히스토그램은 확률밀도함수의 도입
 을 위한 것이므로, 그래프에서 나타나는 자료
 의 형태가 연속변량인지를 살펴보면, 히스토그
 램과 도수분포다각형에서 교과서 3종은 이산변
 량을 제시하였다.



[그림 III-8] 이산변량의 히스토그램과
 도수분포다각형(M5, p.17)

이것은 히스토그램과 분포다각형이 확률밀도
 함수로의 개념으로 발전시키기 위해서는 히스
 토그램은 연속변량으로 나타내어져야 한다.

세 번째 평균을 계산할 때 제시되는 도수분
 포표의 자료에서 교과서 5종은 이산변량을 제
 시하였다.

아래 표는 어느 학교 학생 40명이 졸업기를 한 기록을 나타낸 것
 이다. 처음 계급을 0회 이상 40회 미만으로 하는 도수분포표를
 만들고, 물음에 답하여라.

<표-5> 졸업기

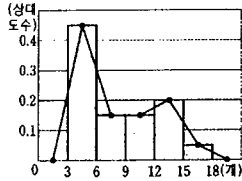
계급	도수	상대도수	누적도수	누적상대도수
1	125	9	238	17
2	15	10	148	18
3	225	11	256	19
4	133	12	57	20
5	248	13	172	21
6	42	14	303	22
7	163	15	92	23
8	315	16	191	24
			288	32
			213	40
			210	
			169	34
			128	35
			205	36
			23	37
			170	38
			296	39
			213	40

[그림 III-9] 평균계산에서 이산변량(M4, p.14)

네 번째 6종의 교과서는 상대도수의 그래프
 에서 제시되는 자료를 제시하고, 이산변량으로
 상대도수의 히스토그램과 도수분포다각형을 그

14) 이산변량 제시한 교과서 6종

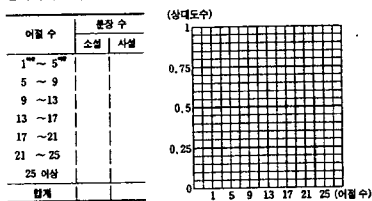
리는 방법을 설명하고 있다.



[그림 III-10] 이산변량의 상대도수의 그래프(M4, p.21)

살펴본 결과, 교과서 M4의 3종은 이산변량으로 상대도수의 그래프에서 그리는 방법을 설명하거나 문제로 그래프를 그리는 활동을 제시하고 있다. 이것은 상대도수의 그래프(히스토그램과 분포다각형)는 확률밀도함수로의 개념으로 발전시키기 위해서는 상대도수는 연속변량으로 나타내어져야 한다. 심화과정에서 16종의 모든 교과서에서 도수의 합이 다른 두 자료를 비교하였고, 이 중 4종의 교과서는 이산변량을 제시하였다.

2 두 자료의 분포를 비교하기 위하여 다음 상대도수의 분포표를 완성하고, 소실과 사실에 해당하는 상대도수의 분포를 색깔을 달리하여 도수분포다각형 모양의 그래프로 나타내어라.

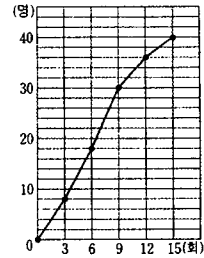


[그림 III-11] 두 자료 분포비교의 상대도수의 분포표와 분포다각형(M7, p.21)

초등학교 때 학습한 이산변량과 범주형 자료를 연속변량을 나타내는 상대도수의 그래프(히스토그램과 분포다각형)에서 이용하는 것은 적절하다고 하기 어렵다.

다섯 번째 연속변량으로 그려지는 누적도수의 그래프는 계급의 간격이 없이 그려져 있어야 한다. 누적도수의 그래프에서 교과서 4종은

이산변량을 제시하고 누적도수의 그래프를 설명하고 있다.



[그림 III-12] 이산변량의 누적도수의 그래프(M4, p.25)

살펴본 결과, 교과서 M1의 3종은 이산변량으로 누적도수의 그래프에서 그리는 방법을 설명하거나 문제로 그래프를 그리는 활동을 제시하고 있다. 앞에서 살펴본 결과 연속변량이 아닌 이산변량으로 표와 그래프를 제시하고 있다. 7차 교육과정에는 이산변량과 연속변량의 구분에 대한 명백한 안내가 없다. 따라서 교과서에 앞서 교육과정이 개선될 필요가 있다고 할 수 있다.

마. 통계적 방법

1) 자료수집의 목적

자료수집 목적을 명확히 서술한 교과서는 2종, 간접적으로 서술한 교과서는 3종, 서술하지 않은 교과서는 11종이다. 우선, 자료수집 목적을 명확히 서술한 교과서 2종에서 교과서 M1(p.10)은 “횡단보도를 건널 때 시간이 부족해서 도로 한 가운데에서 기다리거나 뛰는 경우가 가끔 있다. 지체네 반에서는 학교 앞 횡단보도의 녹색 신호 시간이 적당한지 알아보기로 하였다. 무엇을 조사해야 할까?”로 문제를 제시하고 있다. 이어 ‘보행자 30명을 대상으로 횡단보도를 통과하는 데 걸리는 시간’을

조사한 자료를 제시하고 “20초 이상 걸린 사람은 모두 몇 명인가?”라는 자료의 분포를 묻는 질문과 함께 “녹색 신호 시간은 몇 초로 하는 것이 적절하겠는가?”를 묻고 있다. 즉 횡단보도의 적당한 녹색 신호 시간을 알기 위해서 보행자 30명을 대상으로 횡단보도를 통과하는 데 걸리는 시간을 조사하고 있다. 교과서 M2(p.6)는 우리 반 학생들의 운동복을 단체로 구입하기 위해서 “(2) 우리 반 전체의 경우 운동복은 치수별로 각각 몇 벌씩 필요한지 조사하여라.”고 질문하고 있다. 치수 80에서 110에 해당되는, 즉 각 치수에 해당되는 운동복을 몇 벌이 필요한지 알기 위해서(예를 들면, 80에 해당되는 140-145(cm)인 학생이 몇 명인지 알아야 하므로) 혹은 구입하기 위해서는 학생들의 키를 조사할 필요가 있는 것이다. 자료를 수집하는 목적은 자료를 수집하여 규칙을 발견하고 앞으로의 일을 예측, 계획하고 어떤 일을 결정함에 있어 올바른 판단을 하기 위해서이다. 교과서 M1은 보행자 30명의 횡단보도를 통과한 시간을 조사하여 녹색 신호 시간의 적절한 시간을 예측하는 것이 목적이고, 교과서 M2는 학생들의 키를 조사하여 운동복을 치수별로 몇 벌씩 주문할 것인가에 대한 올바른 판단을 하기 위해서이다. 간접적으로 자료수집의 목적을 서술한 교과서 3종에서 교과서 M14는 자료수집 목적을 아래와 같이 서술하고 있다.

학급 친구들과 야구공 던지기를 하여 누가 가장 멀리 던질 수 있는지 알아보기로 하자. (단, 한 사람이 두 번씩 던져서, 그 중 더 멀리 던진 것을 그 사람의 기록으로 인정하기로 한다.)

▶ 이 때, 자기보다 더 멀리 던진 사람이 몇 명이 있는지 쉽게 알아보려면, 주어진 자료를 어떻게 정리하면 될까?(p.9)

학급 친구들 중 누가 가장 멀리 야구공을 던질 수 있는지 알기 위해서 야구공의 던진 기록이 필요한 것이다. 또한 여기서는 다른 교과서에서 언급하지 않은 측정방법 중 어떻게(단, 한 사람이 두 번씩 던져서, 그 중 더 멀리 던진 것을 그 사람의 기록으로 인정하기로 한다.) 자료를 얻었는지를 부분적이지만 서술하였다.

자료수집의 목적을 직접적으로 서술한 교과서 M1과 M2에 비해, 교과서 M14는 ‘누가 가장 멀리 야구공을 던지는가를 알기 위해’라고 수집목적을 서술하였지만 단순히 누가 가장 멀리 던지는가에 끝나는 것이 아니라 이 자료를 통해 예를 들면, 야구선수의 외야수를 누구로 결정한 것이라는 어떤 일을 결정함에 있어 올바르게 판단할 수 있어야 할 것이다. 즉 구체적이고 실제적인 자료수집의 목적이 되어야 할 것이다.

앞에서 언급한 것처럼 자료 수집을 통하여 예측, 계획, 올바른 판단을 하기 위해서는 정확한 자료가 수집되어야 하는 것이다. 따라서 정확한 자료의 중요성과 정확한 자료를 얻기 위한 방법에 대한 서술여부와 서술방법을 살펴보았다.

2) 정확한 자료의 중요성과 정확한 자료를 얻기 위한 방법

16종의 교과서를 분석한 결과 모든 교과서는 정확한 자료의 중요성에 대하여 서술하지 않았다. 도수분포표의 정의와 만들기의 도입에서 제시된 자료¹⁵⁾를 살펴보면 모든 교과서에서는

15) 직접 측정하거나 수집한 자료 : 학생들의 키, 등교(통학)시간, 여름방학봉사 활동시간 등
보관된 기록자료를 조사한 자료 : 대관령 평균기온, 우리나라에서 대도시를 제외한 지방에 소재한 30개 도시의 인구밀도, 월드컵본선에 3회 이상 진출한 국가와 그 횟수 등

직접 측정하거나 수집한 것으로 제시된 자료가 목적에 맞도록 설계, 실험됐는지 서술되어 있지 않다. 자료를 수집할 때는 '정확한 자료의 수집'이 중요하다. 수행과제 문제¹⁶⁾처럼 비만도나 그 외 사생활을 침해할 수 있는 질문 혹은 마약취급 등 불법행위 등의 자료를 조사할 때 거짓응답이나 부정확한 자료가 수집될 경우도 있다. 이럴 때는 신분을 노출하지 않는 약속하에 자료를 수집해야 할 것이다. 따라서 자료를 직접, 측정하거나 수집할 때 정확한 자료를 얻기 위해서 그 방법이나 절차에 대해서도 교과서에 서술하여야 할 것이다.

반면 보관된 기록자료를 조사한 자료는 직접, 측정하거나 수집한 것으로 제시된 자료보다는 목적에 맞도록 설계, 실험됐다고 할 수 있지만 이와 다른 측면에서 즉 기록조사자의 잘못이 옮겨올 가능성(목적과 차이가 있는 경우)이 있을 수 있다. 보관된 기록자료를 조사한 자료의 경우는 자료 자체에 대한 신뢰도가 높기 때문에 오히려 부정확한 자료인지 더 알기 어려울 수도 있다. 모든 교과서에서 정확한 자료의 중요성을 서술하지 않은 이유가 '자료'라는 것은 항상 옳바를 것이라는 암묵적 전제를 반영하는 것이라고 볼 수 있다.

하지만 실생활에서 제시되는 자료는 정확한 자료가 아닌 경우도 존재하기 때문에, 자료 수집을 통하여 예측, 계획, 올바른 판단을 하기 위해서는 자료수집 목적에 대해 타당한 '정확한' 자료¹⁷⁾의 중요성을 아무리 강조하여도 지나치

지 않을 것이다. 또한 정확한 자료라도 사사로운 이득을 위해 자료에 맞지 않는 통계적 방법을 이용하여 잘못된 결과, 자료를 제시하는 경우도 있다. 따라서 통계적 방법의 타당성의 서술여부와 서술방법을 살펴보았다.

3) 통계적 방법의 타당성

16종의 교과서를 분석한 결과 모든 교과서는 통계적 방법의 타당성에 대하여 올바르게 서술하였다. 사용된 통계적 방법이 올바르게 아냐는 것은 도수분포표, 히스토그램과 도수분포다각형, 상대도수의 분포표, 누적도수의 분포표 등에서 통계적 분석의 목적에 어긋난 통계적 도구를 활용하는 경우를 뜻한다¹⁸⁾. 교과서 16종은 도수분포표, 분포의 표현, 분포의 비교에서 살펴본 것처럼 통계적 분석의 목적에 맞게 자료를 활용하고 있다.

예를 들어, 경시대회 1차 성적결과 중 본선에 진출한 학생은 상위 20명이라 한다. 이 때 누적도수의 그래프가 좋은 것인가? 누적상대도수의 그래프가 더 좋은 것인가? 이 경우 구하고자 하는 것이 본선에 진출한 학생 수는 경시대회에 참가한 학생수가 100명, 1000명이던지 상관없이 항상 '상위 20명'이기 때문에 이 경우 누적도수의 그래프를 이용하는 것이 적합하다. 반대로 학생 수가 다른 두 학교 학생들의 성적 분포 비교에서 어느 학교 학생들의 성적이 더 좋은가를 비교할 때 도수분포다각형이 더 좋은지, 상대도수의 분포다각형이 더 좋은가? 이 경

16) 세계보건기구(WHO)의 '표준 비만도 산출 공식'을 이용하여 우리 반 친구들의 비만도를 조사하는 문제 (M16, p.30)

17) 예를 들면, 우리나라사람들의 물 사용량은 세계 전체와 비교할 때 많은가? 라는 목적으로 물 사용량을 조사할 때 각 세대별 물 사용량을 조사(실제 측정 자료는 수도요금으로 산출)할 경우 수도관시설의 낙후 등으로 물이 새서 요금이 많이 나오면 이것은 물 사용량과 무관하다고 볼 수 있다. 정확한 물 사용량을 조사해야 할 것이다.

18) 예를 들면, 상대도수를 바탕으로 그린 상대도수의 분포다각형에 근거한 비교를 해야 하는데 도수를 바탕으로 한 것을 쓰거나, 반대로 도수이어야 하는데 상대도수로 비교한 경우, 또는 분포자체를 비교해야 하는데 분포의 평균만으로 비교하는 경우 등이라고 할 수 있다.

우는 학생 수가 다르므로 상대도수를 사용하여야 한다. 따라서 이 경우는 상대도수의 분포다각형을 이용하는 것이 적합하다.

모든 교과서에서 분석하고자 하는 목적에 맞게 통계적 방법을 사용하였지만 위의 2가지의 예처럼 목적에 맞는 통계적 방법을 제시하여 어떤 통계적 방법이 타당성 있는지 판단하는 내용도 도움이 될 것이다. 또한 실생활에서 제시된 자료들 중에서는 앞에서 언급한 '정확한 자료의 수집'도 중요하지만 '정확한 자료'라 하더라도 그 자료를 통계적 방법의 타당성에 어긋난 방법으로 산출한 결과를 제시할 수도 있다. 통계적 소양을 가지고 올바른 판단할 수 있게 혹은 경험을 할 수 있도록 교과서와 함께 교육과정 역시 개선의 여지가 있다고 생각된다.

IV. 결론 및 제언

교과서 분석결과 분포개념과 함께 내용연계성, 행동중심적(동기적) 목표, 통계적 소양 측면에서 개선되어야 할 점을 제안하면 크게 다섯 가지로 요약된다.

첫째 자료를 도수분포표, 히스토그램과 도수 분포다각형, 상대도수와 누적도수의 그래프로 나타내는 것은 분포의 표현이며, 하나의 분포 내에서 범주 간 도수비교, 두 개 이상의 분포 간 절대도수 또는 상대도수에 의한 비교 등 상대도수와 누적도수는 분포의 비교에 대한 내용으로 분석된다. 또한 두 집단의 분포비교로 이용되는 상대도수뿐 아니라 평균으로도 이용될 수 있으며, 평균이 중심위치의 측도로서 5-나에서 배웠던 수학적 문제가 아니라 통계적 문제이다. 따라서 7-나 통계단원의 지도핵심은 '분포'이고 분포의 표현, 분포의 비교, 분포의 위치에 대한 내용임을 알 수 있다. 교사 입장에서

무엇을 가르치는 지, 학생입장에서 무엇을 배우는 지 처음 접하는 단원명을 통해 알 수 있으므로 7-나 통계단원의 중단원명을 '분포의 표현'과 '분포의 비교'로 수정할 것을 제안한다. 또한 중심위치의 측도의 대푯값으로 평균, 중앙값, 최빈값이 이용되고 가장 많이 이용되는 대푯값이 평균이므로 '평균'을 분포의 개념, 통계적으로 인식하기 위해 '분포의 위치'로 수정할 것을 제안한다.

둘째 표와 그래프를 작성할 때 학생들이 "왜 배우는가?"에 대한 해답이 될 수 있도록 즉 동기유발을 할 수 있는 목표를 제시할 필요가 있다. 따라서 각 그래프의 특징적 목표를 서술하도록 교과서가 구성되어야 할 것이다. 그 세부적 내용은 앞장에서 제시되었다.

셋째 7차 교육과정에서 통계단원은 분포개념이 국민공통기본과정인 1단계에서 10단계까지 모두 다루어지고, 밀도도수히스토그램은 분포개념의 발달 측면에서 고등학교서 배우게 되는 확률밀도함수로 연계되는 핵심적 개념이라고 할 수 있다. 밀도도수의 히스토그램에서 한 개의 직사각형의 넓이(계급의 크기 × 밀도도수)는 상대도수를 나타낸다. 밀도도수의 히스토그램이 계급의 간격을 점차 좁혀나가면 직사각형들의 계단식 모임에서 부드러운 곡선 모양으로 변하고 이 곡선 아래의 넓이는 여전히 1이며 이 곡선을 나타내는 함수가 확률밀도함수이다. 7-나 통계단원에서 배우는 분포는 범주와 빈도의 순서쌍의 집합으로 이것이 고등학교의 확률함수와 확률밀도함수로 관련지을 수 있게, 분포의 일반화의 준비라 할 수 있다. 따라서 연계성 확보를 위한 대안으로 밀도도수의 히스토그램을 교과내용에

추가할 것을 제안한다. 더불어 이산변량의 막대그래프와 연속변량의 히스토그램을 구분해 주면 히스토그램과 확률밀도함수의 연계성 측

면에서도 일관된 전개가 분명해 질 것이다. 아울러 본 연구에서는 논의되지 못하였으나, 이를 위해 8단계와 9단계의 확률통계 단원과의 연결성도 중요하리라고 생각된다.

넷째 앞에서 살펴본 것처럼 7차 교육과정에서 제시한 도수분포표, 히스토그램과 도수분포다각형, 도수분포표에서 평균, 상대도수와 누적도수의 그래프, 심화과정(도수의 합이 다른 두 집단의 분포를 비교하는 방법)을 16종 교과서에서 모두 서술하였다. 개정된 수학과 교육과정(교육인적자원부, 2007)에서는 7차 교육과정의 7-나 통계단원의 심화과정(도수의 합이 다른 두 집단의 분포를 비교하는 방법)이 삭제되었다. 도수의 합이 다른 두 집단의 분포를 비교하는 방법은 교과서 저자의 해석에 따라 통계단원 본문 내용에 삽입할 수도 있고 삭제할 수도 있다. 이것은 다양화와 표준 교육과정 두 측면에서 살펴볼 수 있는 데 개정될 교과서에서 도수의 합이 다른 두 집단의 분포를 비교하는 방법에 대한 삽입과 삭제의 여부는 본 연구자가 판단하기 어렵다. 따라서 교육과정에서는 도수의 합이 다른 두 집단의 분포를 비교하는 방법에 대해 명확히 제시할 필요가 있다고 본다.

다섯째 학생들이 실생활에서 접하는 자료의 정확성, 해석할 수 있는 능력 즉 통계문맹을 통계교육을 통해 극복할 수 있도록 정확한 자료의 중요성, 목적에 맞는 통계적 분석의 방법에 대한 이해를 할 수 있도록 교과서가 구성되는 것이 필요하다고 본다. 세부적인 대안과 예는 앞장에서 제시되었다.

참고문헌

강육기 외(2001). **중학교 수학 7-나**. 두산.
강행고 외(2001). **중학교 수학 7-나**. 중앙교육

진흥연구소.
김경란(2007). **분포개념 발달 관점에서 본 초·중등 확률통계 교육과정 및 교과내용의 분석**. 이화여자대학교 교육대학원 석사학위논문.
김우철 외(2005). **현대통계학(제4개정판)**. 서울:영지문화사.
김우철 외(2006). **통계학개론(제4개정판)**. 서울:영지문화사.
김응환·이석훈(2008). **통계교육(2판)**. 서울:경문사.
김원경 외(2003). **고등학교 확률과 통계 교사용 지도서**. 천재교육.
고성은 외(2001). **중학교 수학 7-나**. 블랙박스.
교육부(1999). **교육부 고시 제 1997-15 중학교 교육 과정 해설**. 대한교과서주식회사.
교육인적자원부(2007). **수학과 교육과정 [별책 8]**. 대한교과서주식회사.
금중해 외(2001). **중학교 수학 7-나**. 고려출판.
나귀수·황해정·한경혜(2001). **수학과 교육 목표 및 내용 체계화 연구(II)**. 한국교육과정평가원 연구보고 PRC 2001-9.
남주현(2007). **초·중등 통계교육을 위한 통계적 방법론에 대한 연구**. 이화여자대학교 대학원 박사학위논문.
류희찬 외(2007). **학교수학을 위한 원리와 기준**. 서울:경문사.
박규홍 외(2001). **중학교 수학 7-나**. 두레교육.
박두일 외(2002). **중학교 수학 7-나**. 교학사.
박윤범 외(2001). **중학교 수학 7-나**. 대한교과서.
박정식·윤영선(2002). **현대통계학(제4판)**. 서울:다산출판사.
배종수 외(2001). **중학교 수학 7-나**. 한성교육연구소.
신항균(2001). **중학교 수학 7-나**. 형설출판사.

- 성태제(1995). **현대기초통계학의 이해와 적용**. 서울: 양서원.
- 이광복 외(2004). **고등학교 수학 10-가 교사용 지도서**. 새한교과서.
- 이방수 외(2002). **고등학교 수학 10-가 교사용 지도서**. 천재교육.
- 이영하 외(2001). **중학교 수학 7-나**. 교문사.
- 이영하 외(2002). **중학교 수학 7-나 교사용 지도서**. 교문사.
- 이영하(2002). **확률과 통계**. 2002년 서울시 1급 정교사 연수자료.
- 이영하(2007). **확률과 통계의 목표, 내용의 계열화 및 평가**. 광주시 연수자료.
- 이영하·남주현(2005). 통계적 개념 발달에 관한 인식론적 고찰. **한국수학교육학회지** 44(3), pp.457-475.
- 이외숙 외(2002). **통계학 입문(2판)**. 서울:경문사.
- 이준열 외(2001). **중학교 수학 7-나**. 디딤돌.
- 우정호(2000). 통계교육의 개선방향 탐색, **학교수학**, 2(1), pp.1-27.
- 우정호(2001). **학교수학의 교육적 기초(증보판)**. 서울: 서울대학교출판부.
- 양승갑 외(2001). **중학교 수학 7-나**. 금성출판사.
- 조태근 외(2001). **중학교 수학 7-나**. 금성출판사.
- 조태근 외(2003). **고등학교 수학 I**. 금성출판사.
- 전평국 외(2002). **중학교 수학 7-나**. 교학연구사.
- 최용준(2002). **중학교 수학 7-나**. 천재교육.
- 최지안(2008). **중학교 1학년 통계단원에 나타난 분포 개념에 관한 분석**. 이화여자대학교 교육대학원 석사학위 논문.
- 황혜정·나귀수·서동엽(2000). **수학과 교육 목표 및 내용 체계화 연구**. 한국교육과정평가원 연구보고 PRC 2000-3.
- 황석근·이재돈(2001). **중학교 수학 7-나**. 한서출판사.
- D. Ben-Zvi and J. Garfield (2004). **The Challenge of Developing Statistical Literacy. Reasoning and Thinking**(p.48). Kluwer Academic Publishers. Printed in the Netherlands.
- Oxford Dictionary (2005).
- Wordworth, B. J. (1995). **빠아제 인지적·정의적 발달**. (성옥련, 김수정, 이지연 역). 서울: 중앙적성출판사. (영어 원작은 1989년 출판)

An Analysis of Statistics Chapter of the Grade 7's Current Textbook in View of the Distribution Concepts

Lee Young Ha (Ewha Womans University)

Choi Ji An (The Graduate School of Education Ewha Womans University)

This research is to analyze the descriptions in the statistic chapter of the grade 7's current textbooks. The analysis is based on the distribution concepts suggested by Nam(2007). Thus we assumed that the goal of this statistic chapter is to establish concepts on the distributions and to learn ways of communication and comparison through distributional presentations.

What we learned and wanted to suggest through the study is the followings.

1) Students are to learn what the distribution is and what are not.

2) Every kinds of presentational form of distributions is to given its own right to learn so that students are more encouraged to learn them and use them more adequately.

3) Density histogram is to be introduced to extend student's experiences viewing an area as a relative frequency, which is later to be progressed into a probability density.

4) Comparison of two distributions, especially through frequency polygons, is to be an hot issue among educational stakeholder whether to include or not. It is very important when stochastic correlations be learned, because it is nothing but a comparison between conditional distributions.

5) Statistical literacy is also an important issue for student's daily life. Especially the process ahead of the data collection must be introduced so that students acknowledge the importance of accurate and object-oriented data.

* **Key words** : textbooks analysis(교과서 분석), concepts on the statistics(통계적 개념), distribution(분포)

논문접수: 2008. 7. 9

논문수정: 2008. 8. 11

심사완료: 2008. 8. 25