韓國數學教育學會誌 시리즈 A <數學教育>
http://dx.doi.org/10.7468/mathedu.2020.59.1.47
2020. 2. 제 59권, 제 1호, 47-62.

J. Korean Soc. Math. Ed. Ser. A: The Mathematical Education
February. 2020, Vol. 59, No. 1, 47-62.

이산확률분포에 대한 예비수학교사의 이해 분석

이봉주(경북대학교, 교수) · 윤용식(제주대학교, 교수) · 임해미(공주대학교, 교수)†

A study on the understanding of mathematics preservice teachers for discrete probability distribution

Lee, Bongju(Kyungpook National University, leebj@knu.ac.kr)
Yun, Yong Sik(Jeju National University, yunys@jejunu.ac.kr)
Rim, Haemee(Kongju National University, rimhaemee@kongju.ac.kr)†
†Corresponding Author

초록

본 연구에서는 이산확률분포 파악에 필요한 지식을 표본공간의 각 원소에 정의된 확률, 이산확률변수의 정의, 이산확률변수에 정의되는 확률, 그리고 이들 사이의 관계에 대한 지식으로 정의하고, 예비수학교사가 해당 지식을 어느 정도 이해하고 있는지에 대하여 살펴보았다. 이를 위해 검사 도구를 개발하고 사범대학생 47명을 대상으로 조사하였다.

Abstract

Understanding the concept of probability distribution becomes more important. We considered probabilities defined in the sample space, the definition of discrete random variables, the probability of defined discrete probability distribution, and the relationship between them as knowledge of discrete probability distribution, and investigated the understanding degree of the mathematics preservice teachers. The results are as follows. Firstly, about 70% of preservice teachers who participated in this study expressed discrete probability distribution graphs in ordered pairs or continuous distribution. Secondly, with regard to the two factors for obtaining discrete probability distributions: probability for each element in the sample space and the concept of random variables that convert each element in the sample space into a real value, only 13% of the preservice teachers understood and addressed both factors. Thirdly, 39% of the preservice teachers correctly responded to whether different probability distributions can be defined for one sample space. Fourthly, when the probability of each fundamental event was determined to obtain the probability distribution of the discrete random variables defined in the undefined sample space, approximately 70% habitually calculated by the uniform probability. Finally, about 20% of preservice teachers understood the meaning and relationship of binomial distribution, discrete random variables, and sample space. In relation, clear definitions and full explanations of concept need to be provided from textbooks and a program to improve the understanding of preservice teachers need to be developed.

* 주요어 : 이산확률분포, 예비수학교사, 교사 지식
* **Key words** : discrete probability distribution, mathematics preservice teacher, teacher knowledge
* **Address**: Department of Mathematics Education, Kongju National University, Gongju, Korea
* **ZDM Classification** : D40
* **2000 Mathematics Subject Classification** : 97D40
* **Received**: January 14, 2020 **Revised**: January 30, 2020 **Accepted**: February 14, 2020

Ⅰ. 서론

분포라는 말의 뜻을 모르는 학생은 거의 없을 것이다. 그러나 확률분포에 대하여 설명하도록 하면 많은 학생들이 선뜻 대답하지 못한다. 이는 우선 확률에서 사용하는 용어의 일상성에서 비롯된다. ‘분포’에 대한 국어사전 상의 의미는 ‘일정한 범위에 흩어져 퍼져 있음’ 등으로(Standard Grand Korean Dictionary, 2019), ‘우주에 분포된 물질’, ‘우리 지역의 야생동물 분포’와 같이 분포라는 말은 일상적으로 사용된다. 따라서 확률과 통계 영역에서의 분포의 개념을 이해하기 위해 해석학이나 대수학의 용어 예를 들어 도함수의 개념 등을 이해하는 것보다는 상대적으로 적은 노력을 기울일 가능성이 있다. 무엇보다 학생들이 분포 개념을 설명하기 어려워하는 이유는 도수분포부터 확률분포 함수에 이르기까지 분포 개념이 확률과 통계 영역 내에서도 여러 개념과 연관되어 있기 때문이다. 관련하여 Bakker(2004)는 ‘분포’란 여러 측면과 이해의 다양한 층이 있는 복잡한 개념이기 때문에 그 개념을 명확히 파악해야 한다고 언급한 바 있다.

확률에서 ‘분포’는 변이성이 내재되어 있는 현상을 바라보는 하나의 도구로, 현상에서 일어날 수 있는 여러 값들이 흩어져 있는 상태를 나타낸다(Bakker, 2004). 분포는 변이성의 패턴을 이해하고 찾아내기 위해 표본을 추출하여 자료를 얻고 분석할 때, 이 패턴을 설명하고 예측할 수 있는 도구가 된다. 특히 확률분포는 확률과 통계학을 연결 짓는 개념이다(Cho, 2007). 확률변수는 표본공간에서 실수로 가는 함수이며, 확률변수의 ‘확률분포’는 실수치 함수인 확률변수의 확률이 ‘실수 전체에서 어떻게 분포하고 있는가?’를 의미한다. 이때 ‘어떻게 분포하고 있는가?’에 중요한 역할을 하는 것이 표본공간에 정의된 확률, 확률변수의 정의, 확률변수에 정의되는 확률, 그리고 이 세 요소 사이의 관계이다.

확률분포와 관련된 기초적인 개념에 대한 정의가 공고하게 수립되지 않으면 학생들은 통계학 이론이 전개될수록 학습에 어려움을 겪게 된다(Cho, 2011; Choi, Yun, & Hwang, 2014). 또한 확률분포는 빅데이터 분석, 컴퓨터 그래픽 등 다양한 영역에서 활용의 범위가 확대되고 있기 때문에(Goodfellow et al., 2014), 제4차 산업혁명 시대에 다양한 직업군에서 종사하게 될 학생들을 고려하여 수학 수업에서 해당 개념을 충실히 다룰 필요가 있다.

한편 이를 위해서는 확률분포에 대한 수학교사의 지식이 선행되어야 한다. 교사의 지식은 교수 학습 전반과 학생의 이해에 직접적인 영향을 주는데, 이와 관련하여 Ball, Thames, & Phelps(2008)는 수학을 가르치는 데 필요한 교사의 지식(Mathematical Knowledge for Teaching)을 교과 내용 지식(Subject Matter Knowledge)과 교수학적 내용 지식(Pedagogical Content Knowledge)으로 구분하여 강조한 바 있다. 교사가 갖추어야 할 교과 내용 지식에는 가르칠 수학 내용에 대한 일반적인 이해와 더불어 수학적 아이디어를 정확하게 표현하고 수학적 규칙과 절차에 대해 개념적 설명을 하며, 수학적 개념의 구조를 이해하고 학생들이 갖는 오류의 원인을 분석할 수 있는 지식 등이 포함된다(Han, 2016).

이상의 논의를 종합하면 학교수학에서 확률분포의 개념이 충실히 다루어지려면 수학교사 양성 단계에서부터 해당 지식에 대한 교사의 이해를 점검할 필요가 있다. 학교수학에서는 이산확률분포와 연속확률분포를 다루는데, 본 연구에서는 확률분포의 도입이 이루어지는 이산확률분포에 중점을 두어 살펴보고자 한다. 이때 이산확률분포를 파악하는 데 필요한 지식을 표본공간의 각 원소에 정의된 확률, 이산확률변수의 정의, 이산확률변수에 정의되는 확률, 그리고 이들 사이의 관계에 대한 지식으로 정의하고, 예비수학교사의 이산확률분포에 대한 이해 정도를 파악하고자 한다. 본 연구의 결과는 예비수학교사의 이산확률분포에 대한 이해를 도울 수 있는 교육 프로그램을 개발하는 기초 자료로 활용될 수 있을 것으로 기대된다.

Ⅱ. 이론적 배경

1. 확률공간과 확률분포

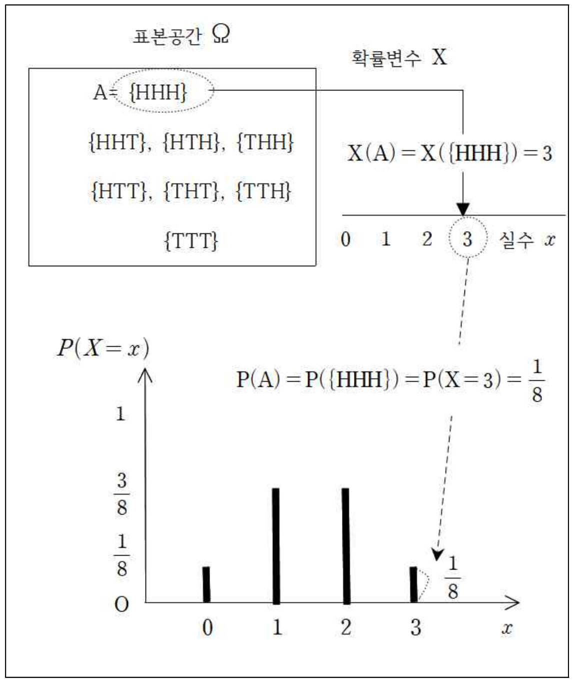
확률론에서 정의되는 확률분포는 실해석학의 측도론에 바탕을 둔다. 측도론에서 가측공간(measurable space)이란 공집합이 아닌 집합 X와 X 상의 $\sigma-$집합체($\sigma-field$) M, M의 원소를 측정하는 역할을 하는 M에서의 측도 μ로 이루어져 있으며, (X, M, μ)로 나타낸다. 어떤 시행에서 일어날 수 있는 모든 결과의 집합을 X대신 Ω로 나타내며, Ω는 표본공간(sample space)이라 부른다. F가 Ω에서의 $\sigma-field$이고, P가 F 위의 확률측도일 때, (Ω, F, P)

를 확률공간(probability space)이라 한다(Kim et al, 2006; Yun, 2005).

실수 전체의 집합 R에 대하여 $X: \Omega \rightarrow R$가 함수일 때 각 $\lambda \in R$에 대하여 $\{\omega \in \Omega \mid X(\omega) \leq \lambda\}$가 사건이면, X를 확률공간 (Ω, F, P) 상의 확률변수(random variable)라고 부른다. 확률변수는 표본공간과 실수 공간을 연결시켜 준다. 확률분포는 표본공간의 측도 P와 대응되는 실수 공간에서의 측도 P_X를 의미한다. P_X는 PX^{-1}로 정의하기 때문에 표본공간의 원소 ω에 대한 확률 $P(\omega)$와 ω에 대응되는 실수의 원소 $X(\omega)$의 확률 $P_X(X(\omega))$는 같다. 즉 어떤 사건에 대한 확률을 구할 때에는 표본공간의 원소 대신 그 원소에 대응되는 실수를 이용할 수 있다. 확률변수는 확률공간 (Ω, F, P)에서 표본공간 Ω가 이산인지 연속인 값인지에 따라 이산확률변수와 연속확률변수로 구분된다.

확률분포는 실수 공간의 측도 P_X이므로 만약 우리가 어떤 집합의 분포를 알면 그 집합의 원소에 해당하는 확률을 구할 수 있다. 이때 확률을 가중치라고 생각하면, 분포는 표본공간의 각 값 또는 구간에 가중치를 부여한 것으로 분포는 확률이 흩어져 있는 상태를 나타낸다. 다시 말해 확률변수 X에 따라 확률이 어떻게 흩어져 있는지를 합이 1인 양수로 나타낸 것을 X의 확률분포라고 한다(Kim et al, 2006; Lee, 2016).

이상에서 살펴본 바와 같이 확률분포의 개념은 표본공간, 시행과 사건, 확률변수에 대한 이해를 토대로 한다. 표본공간, 확률변수, 확률분포의 관계를 간단한 예를 통해 살펴보면 다음과 같다. 동전을 세 번 던지는 확률실험에서 표본공간은 집합 Ω={HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}이 된다. 이때 '앞면이 나온 횟수'를 확률변수 X라 하면, 이는 앞면이 세 번 나오는 사건 A, 앞면이 두 번 나오는 사건 B, 앞면이 한 번 나오는 사건 C, 앞면이 나오지 않는 사건 D의 네 가지 경우로 구분되며, 확률변수 X는 실수 $0, 1, 2, 3$을 취할 수 있고 각 확률변수가 취하는 확률은 $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$, $\frac{1}{8}$이다. 즉 확률변수는 $0, 1, 2, 3$에서 $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$, $\frac{1}{8}$의 확률로 분포한다([Fig. 1] 참조).



[Fig. 1] Relationship between sample space, probability variables, and probability distributions

종합하면, 표본공간은 확률을 내포하고 있으며, 확률변수는 표본공간과 실수 공간을 연결하는 함수이고, 확률분포는 확률변수가 취하는 실수(치역의 값)들이 어느 정도의 가중치를 가지고 있느냐(분포하느냐)를 수치(확률값)로 나타내는 확률이다. Steigerward(2013)는 확률변수의 원어인 'random variable'은 'random'도 'variable'도 아니며 잘못 붙여진 이름으로, 확률변수는 수학적으로 함수로 다루어져야 함을 강조한 바 있다. 확률변수의 확률분포를 파악하려면 확률변수가 취하는 모든 값에 대하여 확률변수가 그 값을 취할 확률을 구해야 한다. 이를 위해서는 표본공간을 이루는 모든 근원사건들의 확률을 이용한다. 확률변수가 갖는 각 값에 대응하는 확률은 이산확률변수의 경우, 그 값을 갖는 근원사건의 확률을 합하여 구할 수 있다(Kim, 2016).

한편 학교수학에서는 표본공간, 시행이 정의된 상태에서 확률변수, 확률분포를 구하는 순서로 내용이 전개되는데, 이러한 전개 방식은 학생들에게 확률과 통계가 갖는 변이성에 대한 사고 기회를 제공하지 못할 수 있다. 학생

들은 표본공간에 내포된 확률의 의미, 표본공간과 확률변수가 동일하지만 확률이 다를 때의 확률분포, 표본공간과 확률이 동일하지만 확률변수가 다를 때의 확률분포는 어떻게 될 것인지 등을 유추하는 경험을 통해 확률분포의 개념을 다각적이고 종합적으로 학습할 필요가 있다. 이러한 학습 경험을 통해 확률분포의 개념이 명확히 정립되었을 때 통계적 추정에 대한 이해, 나아가 통계와 확률을 연계하고 통계적 사고를 할 수 있을 것이다.

2. 학교수학에서의 확률분포

21세기를 살아갈 학생들은 수학적 이해 및 기능, 특히 통계적 이해와 기능을 토대로 데이터로부터 의미 있는 정보를 수집하고 창조하고 소통하는 능력인 데이터 소양(data literacy)을 갖추어야 한다(Lee et al., 2018b). 이는 OECD의 새로운 역량 프로젝트인 OECD Education 2030에서 제시한 주요 역량 중 하나로, 학생들이 통계 및 시각적으로 표시된 정보에 대해 비판적으로 사고하고, 데이터를 분석하여 데이터에 기반한 주장과 근거의 해석이 정확한지를 판단할 수 있어야 함을 강조한다.

수집한 데이터로부터 올바른 추론을 하려면 자료에서 나타나는 변이성을 살펴야 하는데, 이때 변이성을 살펴볼 수 있는 렌즈의 역할을 하는 '분포'에 대한 이해가 필수적이다(Wild, 2006; Lim, 2008). 분포 개념은 자료를 요약하고 모델화하는 것과 관련되며 확률과 통계를 연계하는 중요한 연결 고리가 된다. 자료의 분포를 표현하는 기본적인 방법은 표와 그래프, 분포를 대신하는 수치인 대푯값과 산포도 등으로 요약하는 것으로 초등학교와 중학교 <수학>에서 다루며, 확률분포는 <확률과 통계>에 포함되어 있다. 2015 개정 수학과 교육과정에서 '분포'와 관련된 내용 요소는 [Table 1]과 같다(Ministry of Education, 2015).

2015 개정 수학과 교육과정에 따른 고등학교 확률과 통계 교과서 9종을 분석한 결과, 확률분포 단원은 확률에서의 시행 → 표본공간 → 확률변수 정의 → 확률분포 정의의 순으로 다루고 있었다. 이때 확률분포의 개념을 별도로 제시하기보다 이산확률변수에 대한 확률질량함수의 대응 관계로 이산확률분포를 설명하는 경우가 많았다. 교육과정 문서의 학습요소에 '표본공간'은 포함되지 않기 때문에, 교과서에서는 표본공간을 어떤 시행에서 일어날 수 있는 모든 결과의 집합으로 구체적인 사례를 제시하는 정도로만 다루고 있었다. 2015 개정 교육과정에 따른 수학 교과서의 확률분포의 정의는 [Table 2]와 같이 대체로 세 유형으로 구분되었다. 9종의 <확률과 통계> 교과서는 편의상 A~I로 구분하여 제시하였다.[1)]

[Table 1] Contents related to 'distribution' in the curriculum

School Level	Content Elements	Learning Elements
Elementary school	• Data processing • Average	picture graph, bar graph, line graph, average, band graph, pie chart, chance
Middle school	• Organizing and interpreting data • Probability and basic properties of probability • Representative value and degree of scattering • Correlation	variation, stem-and-leaf plot, class, class interval, frequency, frequency distribution table, histogram, frequency distribution polygon, relative frequency, case, probability, median, mode, representative values, degree of scattering, deviation, variance, standard deviation, scatter plot, correlation
High school	• Probability distribution • Statistical estimation	random variable, discrete random variable, probability distribution, continuous random variable, expected value, binomial distribution, the law of large number, normal distribution, population, sample, complete enumeration survey, sample survey, random sampling, population mean, population variance, population standard deviation, sample mean, sample variance, sample standard deviation, estimation, reliability, confidence interval

1) 9종의 교과서는 참고문헌에 제시함(Bae et al., 2018; Hong et al., 2018; Hwang et al., 2018; Kwon et al., 2018; Kim et al., 2018;, Koh et al., 2018; Lee et al, 2018a; Park et al., 2018; Ryu et al., 2018).

[Table 2] Definition of probability distribution in textbooks

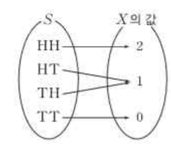
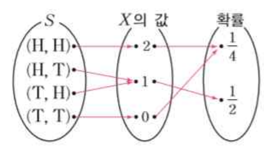
Text books	Definition of probability distribution
A B C D E H I	When all values of the discrete random variable X have are $x_1, x_2, x_3, \cdots, x_n$ and the probability of X having these values is $p_1, p_2, p_3, \cdots, p_n$ respectively, the corresponding relationship of $x_1, x_2, x_3, \cdots, x_n$ and $p_1, p_2, p_3, \cdots, p_n$ is called probability distribution of the discrete random variable X, and the function $P(X=x_i)=p_i \quad (i=1,\ 2,\ 3, \cdots,\ n)$ representing this correspondence is called probability mass function of the discrete probability variable X.
F	*No separate definition of probability distribution. Indicates the properties of the probability mass function $P(X=x_i)=p_i \ (i=1,\ 2,\ 3,\cdots,\ n)$, which indicates the probability distribution of X when the value taken by the discrete random variable X is $x_1, x_2, x_3, \cdots, x_n$.
G	The corresponding function $P(X=x_i)=p_i$ $(i=1,\ 2,\ 3,\cdots,\ n)$ of probability $p_1, p_2, p_3, \cdots, p_n$ of having this value for all values $x_1, x_2, x_3, \cdots, x_n$ of the discrete random variable X is called the probability mass function of the discrete random variable X. The corresponding relationship of the probability mass function is called the probability distribution of the discrete random variable X.

한편 9종의 확률과 통계 교과서 중 8종에서는 표본공간, 확률변수가 갖는 값, 확률 사이의 대응 관계를 [Table 3]과 같이 대응도로 나타내면서 대응 관계를 시각적으로 제시했는데, 이는 앞서 제시한 [Fig. 1]을 단순화한 것이라 볼 수 있다. 8종의 교과서 중 4종은 표본공간과 실수 공간을 연결하는 함수로서의 확률변수를 나타내는 경우만 대응도로 나타냈으며, 4종은 확률변수가 갖는 값과 각 값이 가질 확률의 관계도 대응도로 나타냈다.

확률변수는 표본공간과 실수 공간을 연결하는 함수이지만, 확률변수의 각 값에 확률을 대응시키는 것은 함수가 아니다. 따라서 표본공간과 실수 공간을 연결하는 함수로서의 확률변수를 표현하기에는 대응도가 적합하지만, 확률변수에서 각각에 대응되는 확률을 대응도로 나타내는 것에 대해서는 논의가 필요하다. 이산확률변수에서 각 점에서의 확률의 크기를 나타낼 때 이를 대응도로 시각화하는 것은 수학적인 함수가 아님에도 불구하고 수학적인 함수로 인식하는 오개념을 유발할 수 있기 때문이다.

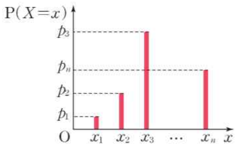
즉, 확률질량함수를 확률변수의 값의 모임에서 [0, 1]로 대응시키는 '함수'라고 잘못 인식할 수 있다. 특히 4종의 교과서 C, D, H, I에서 표현하고 있는 대응도는 합성함수를 설명할 때 주로 사용하기 때문에 표본공간에서 실수의 집합으로의 함수인 확률변수와, 확률변수의 값의 집합에서 확률의 집합으로의 대응을 합성하는 오개념을 생성할 여지가 있다. 여기서는 확률변수의 공역과 치역이 같지 않아 합성함수와 유사한 그림을 사용하는 것에 주의해야 한다. 또한 확률변수의 개념에 대한 명확한 이해 없이 확률변수가 갖는 값에 대응되는 확률을 대응도로 개념 이미지를 구성한 학생은 이산확률분포의 그래프를 수학적 함수의 그래프와 같이 순서쌍(점)으로 나타내는 오류를 범할 수 있으며, 나아가 연속확률변수에 대한 확률밀도함수에서의 확률을 이해하는 데 어려움을 겪을 수 있다.

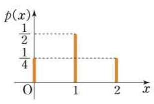
[Table 3] Diagram of random variable and probability distribution in textbooks

Text books	Random variable	Probability distribution
A, B, E, G		No Picture
C, D, H, I		
F	No picture	

한편 확률과 통계 9종의 교과서에 제시된 이산확률변수에 대한 확률분포표와 그래프는 [Table 4]와 같다. 9종의 교과서 중 6종의 교과서에서 이산확률변수의 각 값에서의 확률의 크기를 선막대로 나타냈으며, 2종의 교과서에서는 그래프를 제시하지 않았다. 확률과 통계 수업에서는 이산확률분포의 그래프를 선막대로 나타낸다는 것을 명확히 하고, 선막대의 길이가 확률변수 각 값에서의 확률의 크기를 나타낸다는 것을 지도해야 한다.

[Table 4] Table and graph of probability distribution in textbooks

<table>
<tr><th>Text books</th><th>Table and graph of probability distribution</th></tr>
<tr><td>A, B, C, E, G, H</td><td>
<table>
<tr><td>X</td><td>x_1</td><td>x_2</td><td>x_3</td><td>⋯</td><td>x_n</td><td>합계</td></tr>
<tr><td>$\mathrm{P}(X=x_i)$</td><td>p_1</td><td>p_2</td><td>p_3</td><td>⋯</td><td>p_n</td><td>1</td></tr>
</table>


</td></tr>
<tr><td>D</td><td>
<table>
<tr><td>X</td><td>x_1</td><td>x_2</td><td>x_3</td><td>⋯</td><td>x_n</td><td>합계</td></tr>
<tr><td>$\mathrm{P}(X=x)$</td><td>p_1</td><td>p_2</td><td>p_3</td><td>⋯</td><td>p_n</td><td>1</td></tr>
</table>
No graph
</td></tr>
<tr><td>F</td><td>No table and graph</td></tr>
<tr><td>I</td><td>
Represent as a specific example
<table>
<tr><td>X</td><td>0</td><td>1</td><td>2</td><td>합계</td></tr>
<tr><td>$\mathrm{P}(X=x)$</td><td>$\frac{1}{4}$</td><td>$\frac{1}{2}$</td><td>$\frac{1}{4}$</td><td>1</td></tr>
</table>


</td></tr>
</table>

확률과 통계 영역은 연역 과학인 수학과는 다른 방법의 접근이 요구되며 그 개념과 방법에 대한 적절한 사용과 해석을 수반하는 의미 있는 문제 상황을 통해 학습되어야 한다(Woo, 2017). 그러나 확률과 통계 교과서의 확률분포 단원은 대체로 확률변수와 확률분포 → 이산확률변수의 기댓값과 표준편차 → 이항분포와 그 성질 → 정규분포와 그 성질 → 통계적 추정의 순으로 구성되고, 새로운 개념을 제시하고 예제를 통해 확인하는 연역적 방식으로 서술되었으며, 학생들이 스스로 확률을 정의하여 이산확률분포를 구성하는 과제를 제시하는 경우는 드물었다.

예비수학교사는 스스로 확률을 정의하여 이산확률분포를 구성하는 경험을 할 필요가 있다. 즉 표본공간과 확률변수가 동일하지만 확률이 다를 때의 확률분포를 구하거나, 표본공간과 확률이 동일하지만 확률변수가 다를 때의 확률분포를 구성하는 경험은 매우 중요하다고 볼 수 있다. 이러한 활동을 통해 예비수학교사는 확률분포에 대한 개념을 명확히 할 수 있게 된다.

확률과 통계 영역에서의 교사의 지식 수준은 학생들의 이해에 직접적인 영향을 준다. 이와 관련하여 Byun(2005)은 중등수학교사의 확률과 통계에 대한 지식이 교과서에 대한 의존도가 높고 통계적 이론에 대해 낮은 지식 수준을 보인다고 하였으며, Son(2009)은 고등학교 교과서를 기준으로 수학 교사에게 필요한 확률 내용에 대한 대학 수준의 지식을 정리하였다. Kang(2014)은 학생이 어려워하는 확률변수의 개념을 지도하기 위한 교사의 수학적 지식(MKT)을 제시한 바 있다.

교사는 교과서에 제시된 표본공간, 확률변수, 확률분포에 대한 개념을 명확히 이해하여 학생들에게 전달해야 하며, 이때 고등학교 교과서에 직접적으로 제시되어 있지는 않지만 대학 수준에의 확률분포에 대한 지식을 토대로 학교수학의 내용을 전체적으로 조망할 수 있어야 한다. 이를 위해서는 예비수학교사 교육 단계에서 중등 수학과 교육과정 및 교과서와 대학 수학을 연계하여 논의하는 과정이 필요하며, 교과서도 표본공간, 확률분포에 대한 개념을 명확히 제시하고, 이를 바탕으로 이산확률변수와 분포, 연속확률변수와 분포를 각각 정의하고 이들을 연계하여 이해할 수 있도록 구성할 필요가 있다.

Ⅲ. 연구 방법

1. 연구 절차

본 연구는 이산확률분포에 대한 예비수학교사의 이해를 분석하기 위해 [Fig. 2]의 절차로 연구를 진행하였다.

Set up research topics and explore prior studies
⇩
Analyze high school probability and statistical textbooks
⇩
Develop the pilot test tool
⇩
Implement the pilot test and analyze results
⇩
Modify the pilot test tool and confirm the final test tool
⇩
Select research targets and implement the final test
⇩
Analyze test results and derive implications

[Fig. 2] Research procedure

우선 연구주제를 수립하고 관련 문헌과 선행연구를 탐색하여 확률과 통계 교과서를 분석한 결과를 토대로 예비검사 도구를 개발, 적용하였다. 이후 예비검사 결과를 분석한 결과를 토대로 예비검사 도구를 수정·보완하여 본검사 도구를 확정하였다. 이후 본검사를 실시하고 결과를 분석하였다. 본 연구의 예비검사는 2018년, 본검사는 2019년에 실시하였다.

2. 연구 대상

이 연구에서는 예비수학교사의 이산확률분포에 대한 이해를 조사하기 위하여 3개 사범대학의 수학교육과 학생을 연구 대상으로 설정하였다. 2019년 9월 첫 주에 대학별로 예비수학교사 70명을 대상으로 검사 도구를 배포하였고, 최종적으로 47부가 회수되었다. 47부는 모두 최소한 문항 이상에 응답된 것으로 나타나, 본 연구의 분석 대상을 47명으로 최종 확정하였다. 검사에 참여한 연구 대상에 대한 정보는 [Table 5]와 같다.

[Table 5] Research targets

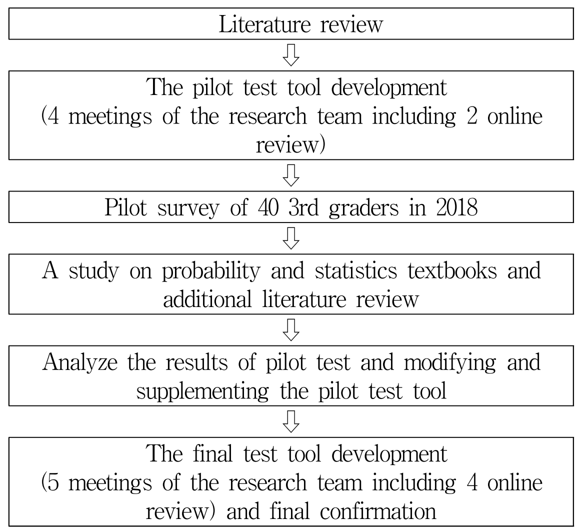
Content		Freq.	%
Grade	3	41	87.2
	4	6	12.8
2009 Revision of math curriculum	before	28	59.6
	after	19	40.4
Number of taking courses in probability and statistics (Including liberal arts courses)	0	8	17.0
	1	30	63.8
	2	9	19.2

연구 대상의 학년을 살펴보면, 3학년이 41명, 4학년이 6명이었다. 고등학교 과정에서 학습한 확률과 통계 과목 또는 내용을 살펴보면, 2009 개정 수학과 교육과정 이전의 내용으로 학습한 예비수학교사는 28명이고, 나머지 19명은 2009 개정 수학과 교육과정으로 학습하였다. 교원양성기관에서 수강한 확률과 통계 관련 강좌 수를 살펴보면, 한 강좌를 수강한 예비수학교사가 30명으로 가장 많았고 두 강좌를 수강한 경우는 전공 과목과 교양 과목을 을 각각 하나씩 수강하였다고 응답하였다. 한 과목도 수강하지 않은 학생도 8명인 것으로 나타났다. 이 연구는 예비수학교사의 이산확률분포 지식에 대한 오류의 근거를 교원양성기관의 확률과 통계 교육 내용에서 탐색하는 것이 아니기 때문에 확률과 통계 관련 수강 강좌 수에 따른 이해 정도를 구분하여 분석하지는 않았다.

3. 검사 도구

1) 검사 도구 개발 절차

검사 도구 개발은 [Fig. 3]의 절차로 진행되었다.



[Fig. 3] Development procedure for test tool

우선 선행연구를 토대로 예비검사 도구 초안을 개발하고, 이후 검토 협의회 2회, 온라인 검토 2회를 통해 2018년 8월에 예비검사 도구를 완성하였다. 예비검사 도구는 표본공간과 이산확률변수의 이해, 표본공간에 확률이 정의되지 않은 경우, 동일한 표본공간에 대해 다르게 정의된 확률변수의 확률분포, 표본공간에서 확률이 다른 경우, 종합적인 이해를 묻는 문항들로, 총 9문항으로 구성하였다. 이후 예비검사 도구를 이용하여 2018년 당시 3학년인 예비수학교사 40명을 대상으로 예비검사를 실시하였다. 예비검사 결과를 분석한 결과, '확률분포표를 구하라.'고 제시한 문항에 대하여 계산 과정에 오류를 보이는 경우가 있기는 하지만 대부분의 예비수학교사가 확률분포표를 옳게 구하는 것으로 나타났다. 그러나 이산확률분포에 대한 종합적인 이해를 묻는 다음 설문 문항에 대한 응답은 거의 무응답이나 오답인 것으로 나타났다.

> 이산확률변수 X가 이항분포를 따를 수 있게 표본공간과 표본공간 상의 확률을 정의하고 표본공간 위의 확률변수 X를 정의하시오. 표본공간, 표본공간 상의 확률 또는 확률변수를 다른 방법으로 정의할 수도 있는지 설명하시오.

예비검사 결과에 근거하여, 첫째, ‘확률분포표’를 구체적으로 명시하는 것보다 ‘이산확률분포’를 어떻게 이해하고 있는지를 조사할 필요가 있음을 확인하였다. 둘째, 이항분포를 따르는 이산확률변수에 대한 확률에 대한 문제의 난이도를 조정하여 문항을 수정할 필요가 있음을 확인하였다. 더 나아가 분석한 9종의 확률과 통계 교과서를 토대로 이산확률분포 그래프에 대한 이해 정도를 추가할 필요가 있음을 확인하였다. 이에 예비검사 도구를 수정하고 보완하여 재구성하였다. 본검사 도구를 개발하는 협의회는 총 5회에 걸쳐 이루어졌는데, 2회는 온라인 협의회로 검사 도구의 문항 수정, 추가 등을 논의하였다. 3회차에는 오프라인 협의를 통해 문항 선제, 문항 나열 순서 등에 합의하였다. 이후 2회의 온라인 협의회를 통해 본검사 도구를 확정하였다.

2) 본검사 도구 구성

본검사 도구는 [Table 6]에 제시된 5개의 문항으로 구성되었다.

1번 문항은 구체적인 확률분포표를 제시하고 해당 확률분포 그래프로 그리도록 한 것이다. 이산확률분포 그래프는 9종의 확률과 통계 교과서 중에서 2종의 교과서에서는 다루지 않고, 다루는 7종의 교과서도 대부분 한 번만 언급하고 제시된 문제도 확률분포표를 구하는 문제로 제한되었다. 이는 이산확률분포를 이산확률분포표로 나타내는 것이 그래프로 표현하는 것보다 더 일반적인 것으로 이해된다. 이에 예비수학교사가 다소 소홀하게 다룰 수 있는 이산확률분포에 대한 지식으로 간주하여 정의하고 최종 검사 도구에 추가하였다.

2번 문항은 이산확률분포는 표본공간의 각 원소에 확률이 정의되고, 표본공간의 원소를 실수로 변환하는 이산확률변수에 대하여 구한다는 것을 인식하고 이해하는지를 알아보기 위한 것이다. 3번 문항은 하나의 표본공간에 대하여 확률분포가 다른 이산확률변수를 정의할 수 있는지에 대한 것이다. 4번 문항은 표본공간의 각 원소에 포함된 확률을 어떻게 다루고 있는지를 알아보기 위한 것이다.

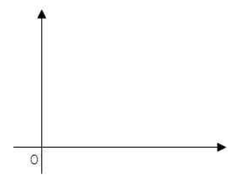
5번 문항은 구체적인 표본공간에 대하여 이산확률변수를 정의하고 그 이산확률변수가 이항분포를 따르도록 표본공간의 각 원소의 확률을 정의할 수 있는지를 알아보기 위한 것이다. 즉, 이 문항은 표본공간에서의 확률, 이산확률변수의 정의, 이산확률변수에 정의된 확률 등의 관계를 일반적인 상황에 적용할 수 있는지를 파악하기 위한 것이다.

[Table 6] The final test tool

Items

1. Look at the next discrete probability distribution table and draw a graph of the discrete probability distribution on the right coordinate plane.

X	0	1	2	합계
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1



2. Choose what you can obtain the probability distribution, and find the probability distribution.

① sample space $\Omega=\{a, b, c, d\}$
② sample space of a single dice
$\Omega=\{1, 2, 3, 4, 5, 6\}$
③ random variables X defined by the number of front faces in an experiment where two coins are thrown

3. Can you define a random variable with different probability distributions for one sample space? Explain the reason using examples.

4. Let Ω is the sample space for selecting students from 1 to 6 in a class.
$\Omega=\{\text{April}, \text{Bekky}, \text{Devy}, \text{Emma}, \text{Frank}, \text{Paul}\}$
When defining random variable $X: \Omega \rightarrow R$ on Ω as $X(\text{April})=1,\ X(\text{Bekky})=2,\ \cdots,\ X(\text{Paul})=6$,
Find the probability distribution of the discrete random variable.

5. The discrete random variable $X: \Omega \rightarrow R$ defined on sample space $\Omega=\{0, 1, 2, \cdots, n\}$ is $X(x)=x$. Define the probability of each element in the sample space so that the discrete random variable according to the binomial distribution.

4. 자료 분석

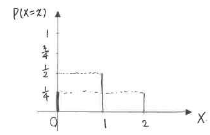
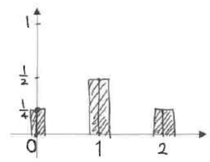
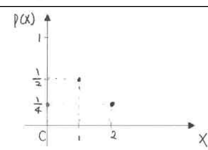
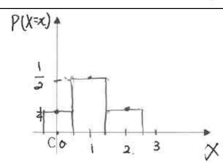
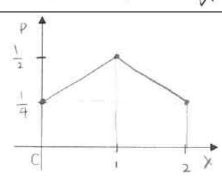
수집한 설문 응답 자료를 다음과 같이 분석하였다. 우선 예비수학교사 47명의 응답을 범주별로 구분하고, 범주별 빈도수와 백분율을 산출하였다. 이에 각 범주별로 대표적인 응답 사례를 토대로 예비수학교사의 이산확률분포에 대한 이해의 근거를 제시하였다.

Ⅳ. 결과 분석 및 논의

1. 이산확률분포 그래프

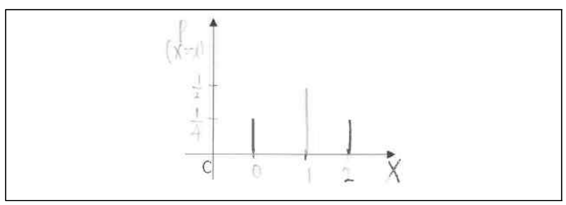
이산확률분포 그래프는 이산확률변수의 확률분포를 나타내는 방법 중 하나이다. 예비수학교사가 이산확률변수의 그래프를 어떻게 표현하는지 알아보기 위한 1번 문항의 응답 결과는 [Table 7]과 같다.

[Table 7] Understanding discrete probability distributions graph

Responses		Freq.	%
line bar graph		8	17.0
bar graph		6	12.8
ordered pair (point)		25	53.2
histo-gram shape		4	8.5
line graph shape		4	8.5
Total		47	100

선막대 또는 막대로 표현한 예비수학교사는 응답자 전체의 약 30%에 불과했는데, 약 70%의 예비수학교사는 점으로 표현하거나 히스토그램 및 꺾은선 등의 이산확률변수가 취하는 값을 연속으로 표현함으로써 이산확률분포 그래프에 대한 오류를 드러냈다.

예비수학교사의 이산확률분포 그래프에 대한 인식을 좀 더 심층적으로 살펴보면 다음과 같다. 먼저, 이산확률분포 그래프는 7종의 확률과 통계 교과서에서 선막대로 표현하고 있다. 이는 이산확률변수의 분포 상태를 쉽게 보여주기 위하여 이산확률분포를 보통 확률막대그래프(probability line graph)로 나타낸다고 설명한 확률과 통계 전문서(예: Kim, Kim, Park, Park, & Soh, 2006; Lee & Kim, 2017)의 내용과 일치한다. 선막대로 그래프를 정확하게 표현한 예비수학교사는 응답자 전체의 17%에 불과한 것으로 나타났다. 이 중에서 한 명의 예비수학교사는 이산확률변수가 갖는 이산점 0의 위치를 좌표평면에서 원점이 아닌 점을 새로 지정하여 표현하였다([Fig. 4] 참조). [Fig. 4]의 응답자와 같이 가로축(x축)에 확률변수 X, 세로축(y축)에 각 확률변수의 값의 확률 $\mathrm{P}(X=x)$을 지정하여 표현하였지만, 이산점 0의 위치를 새로 지정한 것은 좌표평면에서 원점 O의 x값이 0임을 인식하지 못하였거나 확률의 크기를 나타내는 y축에 선을 표시하는 것에 대한 거부감에서 비롯된 것으로 해석할 수 있다. 또 다른 한 명의 예비수학교사는 각 축을 표시하지 않는 것으로 나타났다. 이는 관습적으로 가로축을 x축, 세로축을 y축으로 인식하는 것처럼 보이지만, 확률분포 그래프에서는 x축이 확률변수 X를 나타내고 y축이 $X=x$에서의 확률을 나타낸다는 것을 표시하는 것의 필요성을 간과하고 있음을 드러낸다.



[Fig. 4] Location error case of discrete point zero (0)

다음으로, 이산점에서 확률의 크기를 막대그래프로 표현한 예비수학교사는 응답자 전체의 약 13%에 해당한다.

막대그래프로 나타낸 6명 중 5명이 [Fig. 4]와 같이 $x=0$의 위치를 좌표평면 위에서 원점이 아닌 새로운 점으로 지정하고 각 축도 표시하지 않았다. 이는 이산확률분포 그래프를 막대로 표현하는 것은 적절하지만, 좌표평면에서 각 축이 나타내는 값의 의미를 드러내고 확률변수 X의 값이 실수이므로 $x=0$에서의 확률을 원점에서 표시하는 것임을 예비수학교사에게 인식시켜 줄 필요가 있음을 보여준다.

이어서 이산확률변수의 각 이산점과 확률을 순서쌍(점)으로 표현한 예비수학교사는 응답자 전체의 절반이 넘는 약 53%를 차지하였다. 이는 이산확률분포를 각 이산점에 그에 해당하는 확률의 대응 관계로 설명하고, 대응 관계를 나타내는 함수를 확률질량함수로 정의함에 따른 것으로 해석된다. 즉, 수학에서 다루는 함수의 그래프로 인식하고 있음을 드러내는 것으로 볼 수 있다. 한편 이 중 2명은 이산확률변수가 갖는 이산점 0의 위치를 새로 지정했으며, 9명은 각 축을 표시하지 않았다.

마지막으로, 이산확률분포 그래프를 히스토그램이나 꺾은선그래프 모양으로 나타낸 예비수학교사는 각각 8.5%, 인 것으로 나타났다. 이는 이산확률변수와 연속확률변수를 혼동하고 있다고 볼 수 있다. 이 두 가지 경우에도 1명은 0의 위치를 새로 지정하였고, 4명이 각 축을 표시하지 않았다.

2. 표본공간, 이산확률변수, 이산확률분포의 관계

이산확률변수의 이산확률분포를 구할 때에는 표본공간의 각 원소에 대한 확률, 표본공간의 각 원소를 실숫값으로 변환하는 함수인 확률변수의 두 가지 요인이 중요한 역할을 한다. 이에 대한 예비수학교사의 이해를 조사하기 위한 2번, 3번, 4번 문항의 분석 결과는 다음과 같다.

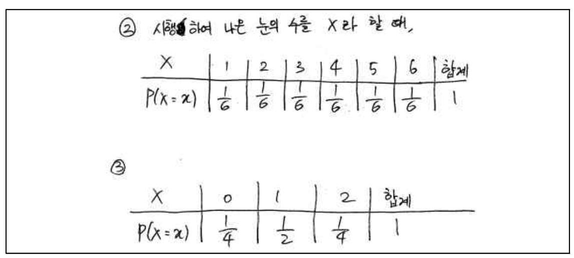
1) 이산확률분포의 결정 요인

먼저, 이산확률분포를 구할 수 있는 두 가지 요인을 모두 고려하는지에 대한 예비수학교사의 응답 결과는 [Table 8]과 같다. 표본공간을 구성하는 각 원소의 확률과 확률변수를 모두 인식하여 옳게 응답한 예비수학교사는 응답자 전체의 약 28%를 차지하고, 무응답을 포함하여 두 가지 요인을 모두 인식하지 못한 경우는 약 11%인 것으로 나타났다. 특히 확률이 정의되어야 한다는 것은 인식하지만 확률변수도 정의해야 한다는 것을 인식하지 못한 예비수학교사는 응답자 전체의 약 62%를 차지하였다. 이러한 결과는 절반 이상의 예비수학교사가 정의된 확률변수에 대하여 확률분포를 구하는 것을 간과하고 있음을 드러낸다고 할 수 있다.

[Table 8] Recognizing the two factors of discrete probability distributions

Responses	Freq.	%
understood both factors and got the right answer	13	27.7
recognized that probabilities must be defined in the sample space, but did not understand that random variable is required	29	61.7
did not recognize both factors	2	4.2
no response	3	6.4
Total	47	100

이산확률분포를 구하기 위한 두 가지 요인에 대한 예비수학교사의 응답 사례를 토대로 보다 자세하게 살펴보면 다음과 같다. 확률변수를 정의한 ③번을 선택하고 확률분포를 확률분포표 또는 확률질량함수로 옳게 구한 예비수학교사는 11명이다. 이 중에서 2명은 ②번의 경우 시행하여 나온 눈의 수를 확률변수 X를 정의한 후에 [Fig. 5]와 같이 확률분포표를 제시하였다.

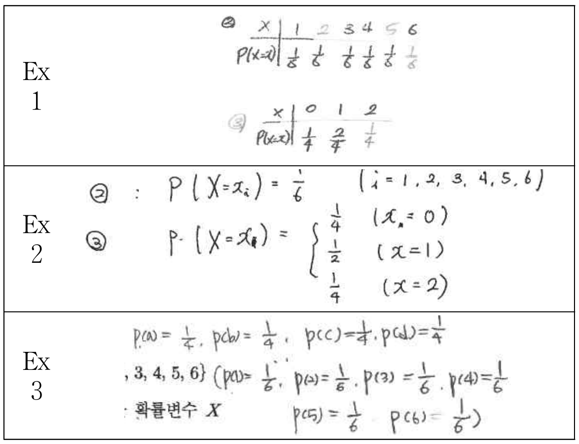


[Fig. 5] Correct responses of item 3

한편, 무응답을 제외한 오답의 사례는 [Table 8]에서 나타나듯이 크게 두 가지 유형으로 구분된다. 하나는 오류 유형은 표본공간의 각 원소에 대하여 확률이 정의되어야 한다는 것을 알지만 확률변수를 정의하지 않는 것이다. 또 다른 오류 유형은 확률의 존재를 이해하지 못하

고 습관적으로 균등 확률로 계산하는 경우이다. 2가지 유형에 대한 예비수학교사의 응답 사례는 [Fig. 6]과 같다.

[Fig. 6]의 사례 1과 사례 2는 첫 번째 오류 유형에 해당하고, 사례 2는 두 번째 오류 유형에 해당한다. 즉, 사례 1과 사례 2는 시행에서 각 근원사건에 대한 확률이 정의되어야 하는 것을 인식하였지만 확률변수를 정의하지 않는 오류를 범하였다. 한편, 사례 1은 확률분포표로 나타낸 것이고, 사례 2는 확률질량함수 식으로 표현한 것이다. 대부분의 예비수학교사는 확률분포를 확률분포표로 나타내지만 사례 2와 같이 확률질량함수 식으로 표현하는 경우도 나타났다.



[Fig. 6] Incorrect responses of item 3

2) 하나의 표본공간에 대한 다른 확률분포

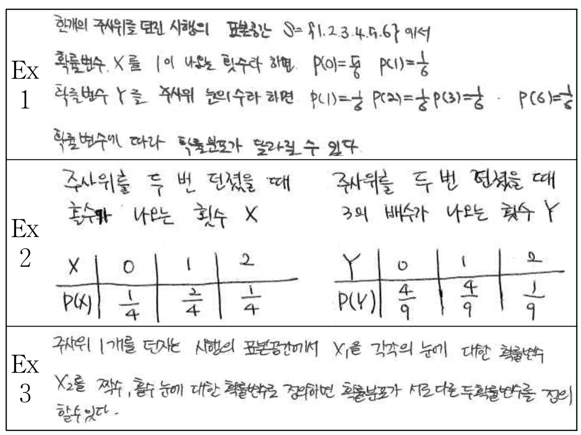
3번 문항은 확률분포가 확률변수의 정의에 따라 결정되는지에 대하여 예비수학교사가 어떻게 이해하고 있는지를 알아보기 위한 것이다. 하나의 표본공간에 대하여 확률분포가 다른 확률변수를 정의할 수 있는지에 대한 예비수학교사의 응답 결과는 [Table 9]와 같다.

동일한 표본공간에 대하여 서로 다른 확률변수를 정의함으로써 확률분포가 다르다는 것을 설명한 예비수학교사는 응답자 전체의 약 15%를 차지하였다. 이에 대비하여 약 24%의 예비수학교사는 표본공간의 원소가 같지만 각 근원사건의 확률이 다른 2개의 시행이나 실험을 예로 제시하고, 다른 확률분포를 제시하였다. 정의할 수 있다고 응답하였지만 그 과정에서 오류가 있거나 확률변수를 고려하지 않는 오류를 보인 예비수학교사도 응답자 전체의 약 37%를 차지하였다. 그리고 약 26%의 예비수학교사는 응답하지 않았다.

[Table 9] Understanding the different probability distributions for one sample space

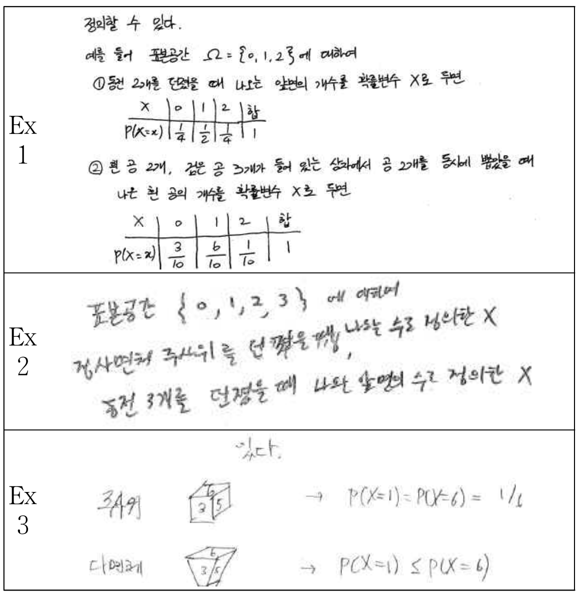
Responses	Freq.	%
defined two discrete random variables with different probability distributions for one sample space	7	14.9
two experiments with the same elements in sample space but different probability of each fundamental event	11	23.4
tried to explain by random variable but took some errors in the process	10	21.3
explained using the probability of the event	4	8.5
responded that it can be defined without an example	3	6.4
no response	12	25.5
Total	47	100

하나의 표본공간에 대하여 확률분포를 다른 확률변수를 정의할 수 있는가에 대한 이해를 예비수학교사의 응답 사례로 좀 더 자세하게 살펴보면 다음과 같다. 하나의 실험에 대하여 확률분포가 다른 확률변수를 정의한 응답 사례는 [Fig. 7]과 같다. 이러한 응답을 한 7명의 예비수학교사는 그림의 사례와 같이 모두 주사위를 던지는 시행을 예로 하여 제시하였다.



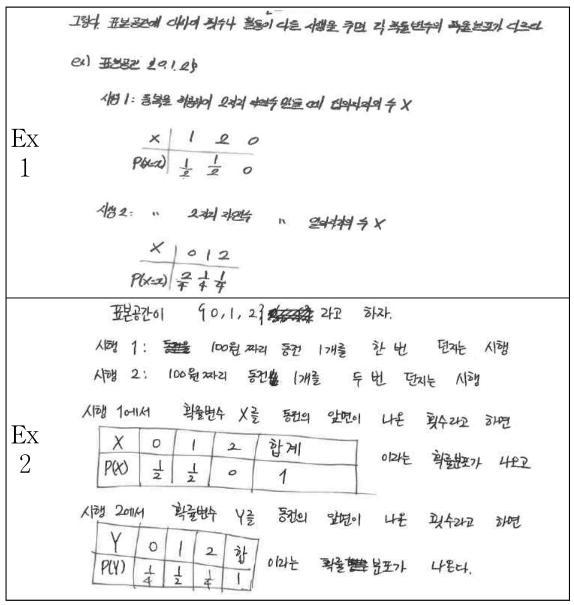
[Fig. 7] Different probability distributions for the same experiment

또한 표본공간의 원소가 같은 2개의 실험을 예로 설명한 예비수학교사의 3가지 응답 사례는 [Fig. 8]과 같다. 사례 1과 사례 3은 모두 동전 던지기, 주머니에서 공 꺼내기, 주사위 던지기 등과 같은 다른 소재를 이용한 시행을 토대로 원소가 같은 표본공간을 형성하고, 그 표본공간에 대하여 확률변수를 정의하여 다른 확률분포를 제시하였다. 사례 3은 주사위 던지기의 동일한 소재이지만 던지는 주사위의 모양이 다르게 지정한 것이다. 이와 같이 각 면의 나타날 확률이 다른 주사위를 이용한 예를 제시한 예비수학교사가 7명인 것으로 나타났다. 이와 같은 응답을 한 이유는 고등학교 확률과 통계 교과서뿐만 아니라 대학 과정의 확률과 통계에서도 표본공간의 동일성에 대하여 다루지 않기 때문인 것으로 해석된다. 이에 이러한 응답이 문항에 대한 틀린 응답이라고 할 수 없다.



[Fig. 8] Responses of different probability distribution response using examples of two trials with the same element in the sample space

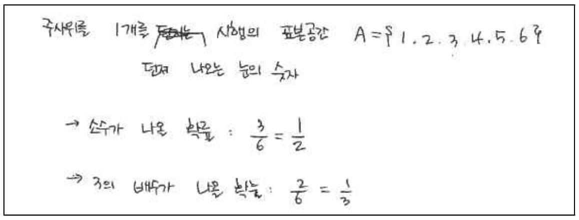
한편 하나의 표본공간에 확률변수를 다르게 정의함으로써 다른 확률분포가 될 수 있음을 설명하는 과정에서 나타난 오류에서 주의해야 할 대표적인 사례를 살펴보면 [Fig. 9]와 같다. 이 그림에 제시된 두 사례 모두에서 나타나는 오류는 크게 두 가지이다. 첫째, 표본공간을 시행의 결과의 모임으로 인식하지 않고, 확률변수가 취하는 값의 모임으로 인식하는 경향이 있음을 보여준다. 둘째, 확률의 크기가 0인 확률변수의 값도 표본공간의 원소에 포함시키는 오류이다. 표본공간은 실험에서 나타나는 모든 결과의 집합이므로 확률이 0인 원소는 표본공간의 원소가 될 수 없다는 것을 간과하였다고 볼 수 있다. 이러한 사례는 확률분포를 구하는 문제에서 표본공간과 확률변수를 이미 정의해 주기 때문에 문제의 답을 찾는 데에는 어려움이 없지만, 실제로 시행하는 실험의 결과에 대한 확률분포를 구하기 위해서는 통계 개념에 대한 명확한 이해가 뒷받침되어야 함을 드러낸다. 이외의 사례에서도 확률변수에 대한 이해가 부족하여 문제의 조건에 맞은 예를 제시하여 설명하는 데 어려움이 있음을 드러내었다.



[Fig. 9] Descriptive error case for probability distribution in one sample space

이외에도 이 문항에 응답하는 과정에서 확률변수와 확률분포에 대한 이해가 부족함을 드러내는 사례는 [Fig. 10]과 같다. 이 사례에서는 두 가지 사건, 즉 소수가 나올 사건과 3의 배수가 나올 사건의 확률을 제시하였다. 이러한 결과는 [Fig. 10]의 사례와 마찬가지로 표본공간, 확률변수, 확률분포에 대한 예비수학교사의 이해를 도와줄 필

요가 있음을 드러낸다.



[Fig. 10] Misunderstanding of probability variable and probability distribution

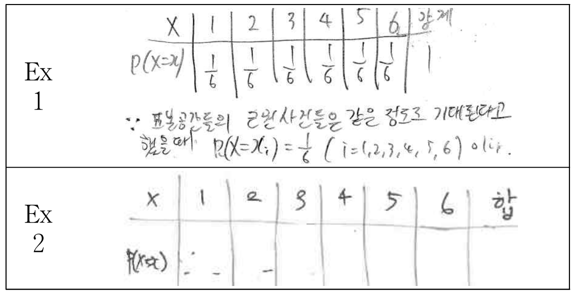
3) 근원사건의 확률과 확률분포

각 근원사건의 확률이 정의되지 않은 표본공간에 정의된 이산확률변수의 확률분포를 구하라는 문항 4에 대한 예비수학교사의 응답 결과는 [Table 10]과 같다. 이산확률분포를 구할 때 표본공간의 각 원소에 정의된 확률을 고려하지 않고 습관적으로 균등 확률로 계산하는 예비수학교사가 응답자 전체의 약 70% 이상인 것으로 나타났다. 문제에 주어진 표본공간의 원소의 개수가 6이므로 70% 이상의 예비수학교사가 각 이산점에서 확률의 크기를 $\frac{1}{6}$로 계산하여 확률분포를 제시하였다. 이는 문항 3번의 결과와 마찬가지로 이산확률분포를 구할 때 고려해야 할 요인에 대하여 강조해야 할 필요성을 시사한다.

[Table 10] Understanding the probability distributions of discrete probability variables for sample spaces with undefined probability

Responses	Freq.	%
calculated equal probabilities without considering reasonable probability	33	70.2
failed to provide a correct answer but considering that the probability is undefined	2	4.3
no response	12	25.5
Total	47	100

한편, 확률이 정의되어 있지 않은 것을 고려하였다고 추정되는 예비수학교사 2명의 응답 사례는 각각 [Fig. 11]과 같다. 사례 1은 확률의 정의가 필요하다는 것을 인식하고 표본공간의 근원사건이 같은 정도로 기대된다고 정의하여 확률분포표를 제시하였다. 사례 2는 확률을 구할 수 없어 각 이산점에서 대응되는 확률을 빈칸으로 제시한 것으로 보인다. 그 외의 무응답으로 제시한 예비수학교사의 인식은 판단하기가 어려운 측면이 있다.



[Fig. 11] Student responses that is judged to take into account probability in item 5

3. 이항분포와 표본공간의 확률

이항분포는 대표적인 이산확률분포 중의 하나이고, 고등학교 확률과 통계 교육과정에서도 이항분포를 다루고 있다. 이에 예비수학교사가 이산확률분포에 대해 종합적인 이해를 하고 있는지를 조사하기 위해 5번 문항은 이항분포에 대한 내용으로 구성하였다. 이 문항에서는 예비수학교사가 이항분포를 따르는 확률변수를 이해할 뿐만 아니라 확률변수와 표본공간의 관계도 이해하고, 확률변수의 정의역인 표본공간의 각 원소의 확률을 정의할 수 있는지를 물었는데, 5번 문항에 대한 예비수학교사의 응답 결과는 [Table 11]과 같다.

[Table 11] Understanding of binomial distribution

Responses	Freq.	%
understood the binomial distribution and sample space and presenting correctly the probability (including calculating the success probability by $\frac{1}{2}$)	9	19.2
recognized the meaning of the binomial distribution but failed to get the probability of sample space	5	10.6
did not understand the binomial distribution	11	23.4
no response	22	46.8
Total	47	100

이항분포를 따르는 확률변수의 정의역인 표본공간의 각 원소의 확률을 옳게 정의한 예비수학교사는 약 20%인 것으로 나타났다. [Fig. 12]는 이항분포를 따르는 이산확률변수의 각 이산점에서의 확률을 옳게 구한 사례를 제시한 것이다. 3명의 예비수학교사는 사례 1 또는 사례 2와 같이 이항분포의 모수인 p에 대한 언급과 함께 이항분포에 대한 확률질량함수 식을 제시하였다. 그리고 다른 3명의 예비수학교사는 사례 4 또는 사례 5와 같이 이항확률분포를 따르는 구체적인 시행을 예로 제시하여 각 이산점에서의 확률을 제시하거나 일반적인 성공 확률 p 대신에 구체적인 확률 $\frac{1}{2}$을 이용하여 확률을 제시하였다. 나머지 3명의 예비수학교사는 사례 3과 같이 단순하게 이항분포를 나타내는 확률질량함수 식만 제시하였다. 특히 이 문항에서는 사례 5와 같이 각 이산점에서의 확률을 제시한 예비수학교사는 2명이고, 다른 7명은 모두 확률을 일반화한 식으로 표현하였다.

Ex. 1	임의의 실수 $p \in [0,1]$에 대하여 근원사건 $\{k\}$ $(k=0,1,2,\cdots,n)$의 확률을 ${}_nC_k\, p^k(1-p)^{n-k}$로 정의하면 확률변수 X는 이항분포를 따른다. ($\because$ $X(x)=x$ 이므로).
Ex. 2	$P(X=k) = {}_nC_k\, p^k q^{n-k}$ $0<p<1$, $q=1-p$, $k \in \{0,1,\cdots n\}$
Ex. 3	$p(X=x) = {}_nC_x\, p^x q^{n-x}$
Ex. 4	$P(\{k\})$ = 동전을 n번 던졌을 때, 앞면이 K번 나올 확률 ($n, k \in \mathbb{N}$, $K \le n$) $= {}_nC_K \left(\frac{1}{2}\right)^n$
Ex. 5	X: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $P(x)$: $\frac{{}_{10}C_0}{2^{10}}$, $\frac{{}_{10}C_1}{2^{10}}$, $\frac{{}_{10}C_2}{2^{10}}$, $\frac{{}_{10}C_3}{2^{10}}$, $\frac{{}_{10}C_4}{2^{10}}$, $\frac{{}_{10}C_5}{2^{10}}$, $\frac{{}_{10}C_6}{2^{10}}$, $\frac{{}_{10}C_7}{2^{10}}$, $\frac{{}_{10}C_8}{2^{10}}$, $\frac{{}_{10}C_9}{2^{10}}$, $\frac{{}_{10}C_{10}}{2^{10}}$

[Fig.12] Example of obtaining probability at each discrete point of a discrete probability variable following a binomial distribution

한편 이항분포의 의미를 일부 이해한 것으로 파악되지만 확률을 제시하지 못한 2가지 사례는 [Fig. 13]과 같다. [Fig. 13]의 사례 1은 이항분포의 확률질량함수 식을 옳게 제시하지 못한 경우로, 2명의 예비수학교사의 응답에 해당된다. 사례 2는 이산확률변수가 이항분포를 따른다는 기호를 표시하고 사례 1과 마찬가지로 확률을 구하지 못한 경우로, 3명의 응답에 해당된다.

Ex. 1	$p(X=k) = p^k(1-p)^{n-k}$, $k \in \mathbb{Z}$, $0 \le k \le n$ $\sum_{k=0}^{n} p^k(1-p)^{n-k} = 1$
Ex. 2	$X \sim B(n,p)$. $p(X=x) = \frac{1}{n+1}$

[Fig. 13] Student responses in which binomial distributions are partially recognized, but the probabilities are not answered

V. 결론 및 제언

확률분포는 확률과 통계학을 연결 짓는 중요한 개념으로, 명확한 개념 형성이 요구된다. 특히, 이산확률분포의 개념을 이해하기 위해서는 표본공간에 정의된 확률, 이산확률변수의 정의, 이산확률변수에 정의되는 확률, 그리고 이들 사이의 관계에 대한 이해가 필수적이다. 이에 본 연구에서는 이를 이산확률분포에 대한 지식으로 정의하고, 예비수학교사의 이해 정도를 조사하고 분석하였다. 본 연구의 결론과 그에 따른 시사점은 다음과 같다.

첫째, 주어진 확률분포표를 그래프로 표현하는 문항에 대하여 (선)막대로 옳게 그린 예비수학교사는 전체 응답자의 30%에 그쳤다. 이에 반해, 이산확률변수의 각 이산점과 확률을 순서쌍(점)으로 표현한 예비수학교사는 응답자 전체의 절반이 넘는 53%를 차지하였다. 이들은 이산확률분포의 그래프를 수학에서 다루는 일반적인 함수의 그래프로 인식하고 있다고 볼 수 있다. 또한 나머지 17%의 예비수학교사는 이산확률분포의 그래프를 히스토그램이나 꺾은선그래프 모양으로 나타냈는데, 이 경우는 이산확률변수와 연속확률변수를 혼동하고 있다고 볼 수 있다. 이상의 결과는 이산확률변수의 각 점에서 갖는 확률의

의미, 즉 확률질량함수를 정의역과 공역 사이의 대응 관계를 나타내는 수학적 함수와 명확하게 구분할 수 있는 설명이 제시될 필요가 있음을 시사한다. 더불어 이산확률변수와 연속확률변수의 확률분포를 구분하는 학습의 기회도 충분히 제공되어야 할 것이다.

둘째, 표본공간, 이산확률변수, 이산확률분포의 관계에 대한 예비수학교사의 이해를 조사한 결과, 이산확률변수의 이산확률분포를 구하는 두 가지 요인 즉, 표본공간의 각 원소에 대한 확률과 표본공간의 각 원소를 실숫값으로 변환하는 확률변수의 개념에 대해 예비수학교사의 13%만이 두 가지 요인을 모두 이해하고 옳게 해결한 것으로 나타났다. 하나의 표본공간에 대해 다른 확률분포를 정의할 수 있는지에 대해서는 39%의 예비수학교사만 옳게 응답하였다. 또한 각 근원사건의 확률이 정의되지 않은 표본공간에 정의된 이산확률변수의 확률분포를 구하도록 했을 때 약 70%가 습관적으로 균등 확률로 계산을 하였다. 이러한 3가지 결과와 예비검사 결과를 연계해 볼 때, 예비수학교사들이 대체로 확률 계산을 토대로 이산확률분포표를 구하는 절차적 지식을 갖추고 있지만 이산확률분포에 대한 개념을 정립하지 못한 것으로 해석할 수 있다. 이산확률분포에 대한 지식을 이해하는 것은 이후 연속확률분포 개념에도 영향을 줄 수 있기 때문에, <확률과 통계> 교과서와 교사용 지도서뿐만 아니라 대학 과정의 확률과 통계 관련 강좌에서 표본공간, 이산확률변수, 이산확률분포의 관계에 대한 명확한 설명과 다양한 예제를 제공할 필요가 있다.

셋째, 대표적인 이산확률분포인 이항분포를 따르는 확률변수에 대하여 그 치역인 표본공간을 이해하고 표본공간의 확률을 정의할 수 있는지와 이항분포, 확률변수, 표본공간의 의미와 관계를 이해한 예비수학교사는 약 20%에 불과한 것으로 조사되었다. 이러한 결과는 두 번째 연구 결과와 마찬가지로 많은 예비수학교사가 이산확률분포에 대한 지식을 갖추고 있지 못함을 드러낸 것이라고 볼 수 있다. 이산확률분포에 대한 예비수학교사의 지식은 향후 교사가 되었을 때 학생의 이해에 직접적인 영향을 줄 수 있기 때문에, 예비교사 양성과정의 확률과 통계 강좌에서 이산확률분포에 대한 종합적인 이해를 가질 수 있도록 하는 프로그램을 개발할 필요가 있다. 프로그램에서는 표본공간, 표본공간의 확률, 표본공간을 실숫값으로 변환하는 이산확률변수, 이산확률변수의 각 점에서의 확률 등의 여러 관련 개념 사이의 관계에 대하여 고찰하고 반성하는 기회를 제공할 필요가 있을 것이다. 이러한 학습 기회가 충분히 제공된다면 이후 연속확률분포에 대한 학습에도 긍정적인 효과를 가져올 수 있을 것이다.

이상의 논의를 종합하면 다음과 같다. 예비수학교사의 이산확률분포에 대한 이해, 나아가 확률분포, 확률과 통계에 대한 이해를 높이기 위해서는 학교 수학 교육과정과 사범대학에서의 예비교사 교육 프로그램에 대한 점검과 연계에 대한 논의가 요구된다. 본 연구의 결과에서 나타난 바를 반영하여 <확률과 통계> 교과서, 교사용 지도서에서는 이산확률분포에 대한 명확한 개념 정의와 상세한 설명을 제공하고, 학생들에게 오개념을 유발할 수 있는 표현은 수정할 필요가 있다. 또한 사범대학에서는 중등교육과정과 연계한 확률과 통계 강좌를 개설하여, 예비수학교사가 대학 수준의 확률과 통계 지식을 바탕으로 중등교육과정의 내용을 살펴보고 개선 방향을 찾는 프로젝트를 수행하도록 하는 등 교사로서의 역량을 기를 수 있는 다양한 기회를 제공해야 할 것이다. 마지막으로 본 연구의 후속연구로 연속확률분포에 대한 예비수학교사의 이해 수준과 중등교육과정을 점검하는 연구를 제안하는 바이다.

참 고 문 헌

Bae, J., Yu, T., Cho, B., Kim, M., Cheon, H., Cho, S., & Byun, D. (2018). *Probability and statistics.* Seoul: Geumsung.

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools.* Doctoral dissertation. Utrecht University: Netherlands.

Ball, D. L., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: what make it special? *Journal of Teacher Education, 59*(5), 389-407. Retrieved Dec. 26, 2019, from https://www.math.ksu.edu/~bennett/onlinehw/qcenter/ballmkt.pdf.

Byun, J. (2005). A *study on the mathematics teachers' knowledge and belief in the high school probability and statistics.* Master's thesis. Korea National University of Education.

Cho, D. (2007). *A study on The understanding of sample mean distribution : a focus on 12th graders.* Master's thesis. Korea National University of Education.

Cho, M. (2011). *An Analysis of Students' Understanding of the Concept of Probability Distributions and Teaching*. Master's thesis. Korea University.

Choi, J., Yun, Y., & Hwang, H. (2014). A study on pre-service teachers' understanding of random variable. *Mathematics School Mathematics, 16*(1), 19-37.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).

Han, H. S. (2016). A study on pre-service mathematics teachers' MKT. *Communications of Mathematical Education, 30*(1), 101-120.

Hong, S., Lee, J., Shin, T., Lee, J., Lee, B., Shin, Y.,…, Kang, M. (2018). *Probability and statistics.* Seoul: Jihaksa.

Hwang, S., Kim, B., Yoon, G., Lee, G., Kim, S., Lee, M .,…, Park, J. (2018). *Probability and Statistics.* Seoul: Miraen.

Kang, J. (2014). *Mathematical knowledge for teaching (MKT) for the concept and mathematical symbols related to random variables*. Master's thesis. Seoul National University.

Kim, J. (2016). *Lecture of probability theory.* Free Academy.

Kim, W., Kim, J., Park, B., Park, S., & Soh, M. (2006). *The modern statistics(4th).* Youngji Publishers: Seoul.

Kim, W., Cho, M., Pang, K., Yun, J., Shin, J., Kim, K.,…, Huh, N. (2018). *Probability and statistics.* Seoul: Visang.

Koh, S., Lee, J., Lee, S., Cha, S., Kim, Y.,…,Cho, S. (2018). *Probability and statistics.* Seoul: Sinsago.

Kwon, O., Shin, J., Jeon, I., Kim, M., Kim, C., Kim, T, …, Hwang, S. (2018). *Probability and statistics*. Seoul: Kyohaksa.

Lee, J. (2016). *Probability and statistics learned as vivid examples.* Hanbit Academy Press.

Lee, J., Choi, B., Kim, D., Jeon, C., Chang, H., Song, Y., & Kim, S. (2018a). *Probability and statistics.* Seoul: Chunjae.

Lee, M., Seo, J., Lee, K., Cho, S., Kim, K., Yu, C.,…, Yoon, K. (2018b). *A Study on content curriculum mapping of Korea in the OECD education 2030 Project.* KICE, RRC 2018-12.

Lee, Y. & Kim, S. (2017). *Understanding statistics*(8th). Yulgokbooks: Seoul.

Lim, J. (2008). *A study of Teaching of Distribution Concept.* Master's thesis. Graduate School of Seoul National University.

Ministry of Education (2015). *Mathematics curriculum. n*otification of the Ministry of Education No. 2015-74. [Vol. 8].

Park, K., Lee, J., Kim, J., Nam, J., Kim, N., Lim, J., …, Yang, J.(2018). *Probability and statistics.* Seoul: Donga.

Ryu, H., Seonwoo, H., Cho, J., Lee, B., Kim, D., Lim, M .,…, Chang, S. (2018). *Probability and statistics.* Seoul: Chunjae.

Son, J. (2009). *A study on probability distribution in the high school curriculum.* Master's thesis. Yonsei University.

Standard Grand Korean Dictionary (2019). Retrieved Oct. 9, 2019, from https://stdict.korean.go.kr/main/main.do

Steigerwald, D. G. (2013). Introduction to measure theory lecture notes. Retrieved Oct. 9, 2019, from http://econ.ucsb.edu/~doug/245a/Lectures/Measure%20Theory.pdf

Wild, C. H. R. I. S. (2006). The concept of distribution. *Statistics Education Research Journal, 5*(2), 10-26.

Woo, J. (2017). *The educational foundation of school mathematics (the third volume).* Seoul National University Publishing Council.

Yun, Y. (2005). *Probability theory.* Jeju National University Publishing Council.