creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

**저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.**

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

교육학석사 학위논문

# 예비교사의 확률적 상황에 대한 접근 방식 분석

## - 등확률성을 만족하지 않는 상황을 중심으로 -

2023년 2월

서울대학교 대학원

수학교육과

이 종 화

# 예비교사의 확률적 상황에 대한 접근 방식 분석

- 등확률성을 만족하지 않는 상황을 중심으로 -

지도교수 유 연 주

이 논문을 교육학석사 학위논문으로 제출함
2023년 1월

서울대학교 대학원
수학교육과
이 종 화

이종화의 석사 학위논문을 인준함
2023년 1월

위 원 장 권 오 남 (인)

부위원장 이 경 화 (인)

위 원 유 연 주 (인)

# 국문초록

실생활에서 확률은 분야를 막론하고 다양하게 활용되고 있으며 삶에서 마주치는 불확실성을 해석하고 대처하는 도구로서 확률이 쓰이고 있다. 이러한 확률의 중요도 때문에 학교수학에서 확률교육의 중요성은 점점 더 커져가고 있다.

확률의 개념이 출현한 후 역사적으로 다양한 의미로 해석되어 왔다. 수학교육에서 확률은 크게 세 가지 관점인 고전적 확률, 빈도적 확률, 주관적 확률로 나누어진다. 확률교육은 확률적 소양을 갖춤으로써 미래를 예측하고 합리적인 의사결정을 하는 민주 시민으로서의 기본 소양을 기르는 것을 목표로 하고 있고, 이를 위해서는 확률에 대한 세 가지 관점이 조화롭게 다루어져야 한다. 하지만 현재의 학교수학에서는 고전적 관점만이 강조되어 확률의 진정한 의미를 살리지 못하고 있으며, 학생들은 확률을 실생활에 활용하지 못하는 경우가 대부분이다. 수학교사들도 확률이 가지는 본질적인 애매한 특성 때문에 고전적 확률을 다루는 것을 선호하고 있다. 확률교육의 성패는 교사의 질에 달려있다고 해도 과언이 아니므로 교사의 확률에 대한 관점이 어떠한지 살펴보는 연구가 필요하다. 예비교사의 확률에 대한 관점을 살펴보면 정규 교육과정을 통해서 확률교육의 목표에 어느 정도 달성했는지 살펴봄과 동시에 교사가 가지고 있는 관점까지 살펴볼 수 있는 효과를 거둘 수 있다. 수학교육에서 다루는 확률에 대한 관점과 확률이론에서 다루는 확률에 대한 측면을 종합하며 확률에 대한 접근 방식을 분석할 수 있는 틀을 도출해내었다.

본 연구에서는 확률에 대한 예 구성 및 의미를 제시하는 활동을 통해 예비교사들이 기존에 가지고 있던 확률에 대한 관점의 양상이 어떤지 살펴보았다. 특히, 등확률성을 만족하지 않는 상황에서 확률값을

추정하도록 하면서, 빈도적 확률과 주관적 확률에 대한 예비교사의 이해를 중점적으로 살펴보았다. 또한 주어진 확률적 상황에 추가적인 정보가 주어졌을 때 확률을 갱신하여 추정하는 과정을 살펴보았다.

연구의 결과로부터 다음과 같은 점을 파악할 수 있었다. 첫째, 예비교사는 대부분 고전적 확률의 상황을 예로 구성하였으며, 개인마다 달라지는 개인신념형 확률의 예는 거의 존재하지 않았다. 둘째, 등확률성을 만족하지 않는 상황에서 확률값을 추정하는 활동에서 개인이 가진 사전 정보에 따라 다양한 접근 방식을 활용하는 것을 관찰하였다. 셋째, 확률적 상황에서 경험적 데이터를 추가적으로 제공하여 확률값을 갱신하는 활동에서 빈도적 확률에서는 여러 확률 중 개인이 느끼기에 객관적이라고 판단되는 확률일수록 참 확률에 가깝다고 판단하는 경향이 있었다. 반면에 주관적 확률에서는 개인이 느끼기에 객관적인 확률이라고 해서 참 확률에 더 적합하다고 판단하지 않고, 확률적 상황과 관련이 깊다고 생각하는 요소들이 확률값 추정에 많이 고려될수록 참 확률에 가깝다고 판단하였다.

본 연구 결과를 바탕으로 확률교육에 적용할 수 있는 방안에 대한 시사점을 제시하면 다음과 같다. 첫째, 수학교육에서 확률에 대한 관점을 분류하는 새로운 틀을 제시하였고, 그 틀을 실제로 확률적 상황의 예에서 적용해 보았다는 점에서 의의가 있다. 둘째, 수학적 모델 도출 과정의 절차가 참 확률값을 추정하는 과정을 분석하는데 유용한 틀임을 확인하였다. 셋째, 확률이론에서 사용되는 여러 정리를 활용하여 확률에 대한 세 가지 접근 방식에 대한 관계를 분석할 수 있음을 확인하였다.

**주요어 : 고전적 확률, 빈도적 확률, 주관적 확률, 예비교사, 확률교육, 확률에 대한 관점**

**학 번 :** 2021-23023

# 목 차

Ⅰ. 서론 ····· 1
1. 연구의 필요성과 목적 ····· 1
2. 연구 문제 ····· 5
3. 연구의 제한점 ····· 5

Ⅱ. 이론적 배경 ····· 6
1. 수학교육에서 확률에 대한 세 가지 관점 ····· 6
1.1. 고전적 확률 ····· 7
1.2. 빈도적 확률 ····· 8
1.3. 주관적 확률 ····· 9
1.4. 확률의 통계적 측면과 인식론적 측면 ····· 11
2. 확률의 각 관점에서 드러나는 주관적 특성 ····· 16
2.1. 고전적 확률에서의 주관적인 특성 ····· 16
2.2. 빈도적 확률에서의 주관적인 특성 ····· 17
2.3. 주관적 확률에서의 객관성 확보 ····· 19
3. 확률적 소양 ····· 20
3.1. 통계적 소양 ····· 21
3.2. 확률적 소양 ····· 22
4. 확률에 대한 교사 지식 ····· 23
4.1. 확률에 대한 교사의 교과 내용 지식의 구조 ····· 24
5. 확률 모델링 ····· 25

Ⅲ. 연구 방법 ····· 28
1. 연구 방법 및 절차 ····· 28
1.1. 연구 대상 ····· 28
1.2. 조사 방법 ····· 28
1.3. 자료 분석 ····· 29
2. 과제 설계 ····· 30
2.1. <1차 설문지>의 개발의도 ····· 30
2.2. <2차 설문지>의 개발의도 ····· 32
3. 확률에 대한 관점과 접근 방식의 분류 ····· 37

Ⅳ. 연구 결과 ····· 40
1. 예비교사가 기존에 가지고 있는 확률에 대한 관점 · 40
1.1. 예비교사가 구성한 확률의 예와 확률에 대한 관점 분석 · 40
1.2. 예비교사가 형성하고 있는 확률의 의미 ····· 45
2. 등확률성을 만족하지 않는 상황에서의 확률 할당 ··· 51
2.1. 절대로 일어날 수 없는, 반드시 일어나는 사건의 확률 · 51
2.2. 거의 일어나지 않는 상황에 대한 확률 ····· 52
2.3. 거의 대부분 일어나는 상황에 대한 확률 ····· 54
2.4. 윷의 확률과 코로나 감염 확률에 대한 확률 할당 ··· 55
2.5. 확률적 상황에서 드러난 확률에 대한 접근 방식 ····· 57
2.6. 확률에 대한 접근 방식과 모델링 ····· 59
3. 추가 정보가 주어졌을 때의 참 확률에 대한 추정 ··· 62
3.1. 빈도적 확률과 집산 ····· 62
3.2. 주관적 확률과 베이즈 정리 ····· 67
3.3. 빈도적 확률과 가장 좁은 준거집합의 원리 ····· 68
3.4. 주관적 확률과 빈도 원리 ····· 69
3.5. 예비교사가 판단한 주관적 확률의 장단점 ····· 70
3.6. 예비교사의 확률에 대한 관점의 변화 ····· 71

4. 주관적 확률에 대한 교수학적 함의 ····· 72
5. 예비교사의 확률에 대한 오개념 ····· 75

Ⅴ. 결론 및 제언 ····· 77
1. 요약 ····· 77
2. 논의 및 시사점 ····· 80

참고문헌 ····· 84
부록 ····· 89
Abstract ····· 98

# 표 목 차

<표 Ⅱ-1> 대한민국의 2021년 완전 생명표 ····· 8
<표 Ⅱ-2> 확률 측정에 관한 용어 ····· 13
<표 Ⅱ-3> 확률적 소양 - 구성 요소 ····· 22
<표 Ⅲ-1> 설문지 내 확률의 주관적 요소 ····· 36
<표 Ⅳ-1> 연구 참여자 별 예 구성의 개수 ····· 40
<표 Ⅳ-2> 예 구성 활동에서 드러난 확률에 대한 접근 방식 · 41
<표 Ⅳ-3> 확률의 제도적 의미에 따른 분류 ····· 48
<표 Ⅳ-4> 확률의 개인적 의미를 서술하는 데 등장하는 키워드 · 50
<표 Ⅳ-5> <1차 설문지> 문항 2의 응답에 대한 분포 ····· 52
<표 Ⅳ-6> 여러 확률적 상황에 대한 접근 방식별 분류 ····· 57
<표 Ⅳ-7> 접근 방식별 확률값 계산에서 활용되는 정보 ····· 59
<표 Ⅳ-8> 1-1번에서 구한 개인별 빈도적 확률(②번 확률) · 63
<표 Ⅳ-9> 던진 횟수에 대한 판단 근거 ····· 64
<표 Ⅳ-10> 예비교사가 생성한 4가지 종류의 확률 ····· 65
<표 Ⅳ-11> 주관적 확률의 장단점 ····· 70

# 그 림 목 차

[그림 Ⅱ-1] 확률이론에서의 이중분류 ····· 13
[그림 Ⅱ-2] Chernoff의 두 분야에서 확률에 대한 분류도 ····· 15
[그림 Ⅱ-3] 확률을 추정에 대한 세 가지 상호연결된 방법 ···· 25
[그림 Ⅱ-4] 수학적 모델 도출 과정의 세분화 ····· 26
[그림 Ⅲ-1] 확률적 상황에 대한 접근 방식과 확률에 대한 관점 · 39
[그림 Ⅳ-1] 고전적 확률 중 무작위 생성 활동을 이용한 예 ·· 41
[그림 Ⅳ-2] 고전적 확률 중 일상적 상황의 예 ····· 42
[그림 Ⅳ-3] 빈도적 확률의 예 ····· 43
[그림 Ⅳ-4] 개인신념형 확률의 예 ····· 44
[그림 Ⅳ-5] 대인신념형 확률의 예 ····· 45
[그림 Ⅳ-6] 제도적 의미에서의 확률에 대한 각각의 예 ····· 48
[그림 Ⅳ-7] 개인적 의미에서의 확률에 대한 각각의 예 ····· 50
[그림 Ⅳ-8] 윷의 확률에 대한 접근 방식의 예 ····· 56
[그림 Ⅳ-9] 코로나 감염 확률에 대한 접근 방식의 예 ····· 57
[그림 Ⅳ-10] 수학적 모델 도출 과정의 세분화 ····· 59
[그림 Ⅳ-11] 윷의 확률에 대한 응답의 예 ····· 60
[그림 Ⅳ-12] 코로나 감염 확률에 대한 응답의 예 ····· 61
[그림 Ⅳ-13] 베이즈 정리와 확률 갱신 과정에 대한
예비교사의 답변 ····· 68
[그림 Ⅳ-14] <2차 설문지> 2-1번 문항에 대한
두 예비교사의 답변 ····· 73
[그림 Ⅳ-15] Excel로 계산한 두 예비교사의 분할표에
대한 정오 ····· 74
[그림 Ⅳ-16] 조건부확률에 대한 두 예비교사의 기호 표현 ···· 74
[그림 Ⅳ-17] 큰 수의 법칙에 대한 오개념 사례 ····· 75
[그림 Ⅳ-18] 표본의 크기와 표본평균의 분포에 대한 설명 ···· 76

# Ⅰ. 서론

## 1. 연구의 필요성과 목적

확률은 일상생활 흔히 사용되는 수학 용어 중에 하나다. 야구 선수의 타율, 수술의 성공 가능성, 주식 시장에 대한 예측, 일기 예보 등 우리 삶과 관련된 다양한 맥락에서 확률과 관련된 용어가 등장한다. 이처럼 확률은 수학, 의학, 경제, 정치 등 자연과학과 사회과학의 모든 분야에서 활용된다. 또한 모든 성인은 삶의 매 순간마다 확률적인 요소와 마주하게 되며, 다양한 수준의 예측가능성과 불가능성을 해석하고 이에 대처해야 한다(Gal, 2000: Gal, 2017 재인용). 이러한 확률의 중요성 때문에 많은 국가의 수학 교육과정에서 확률이 중요하게 다루어지고 있다(Greer & Mukhopadhyay, 2017; Jones, Langrall, & Mooney, 2007).

확률의 개념이 출현한 이후 학자들은 확률의 의미에 대한 두 가지 주요 관점을 제시하였다(Hacking, 1975). Hacking(1975)은 “확률은 야누스의 얼굴과 같아서 한쪽에는 우연 현상에 대한 통계적 법칙 자체를 다루고자 하는 통계적인 면이 있고, 다른 쪽에는 통계적 배경이라고는 전혀 없는 주장과 관련된 신념의 합리성을 평가하고자 하는 인식론적인 면이 있다.”라고 언급하며 확률을 두 가지 관점으로 구분 지었다. 이렇듯 확률적 사고의 특성을 이해하기 위해 경험을 바탕으로 판단한 확률의 객관적인 측면과 개인의 신념을 바탕으로 판단한 확률의 주관적인 측면의 뒤엉킨 특성을 설명하려는 노력이 역사적으로 계속되어 왔다(우정호, 2017). 이러한 확률의 두 가지 측면을 다루면서 확률의 개념이 발전이 이루어졌고, 확률에 대한 다양한 해석이 등장하게 되었다(Batanero & Diaz, 2007; Batanero, Henry, & Parzysz, 2017; Borovcnik & Kapadia, 2014). 현재 학자들은 확률에 대한 주요 관점을 직관적, 고전적, 빈도적, 성향적, 논리적, 주관적, 공리적 관점으로 나누며 각각의 관점은 상

호보완적이라고 말한다(Batanero & Diaz, 2007).

역사적으로는 위와 같이 확률에 대한 다양한 관점이 존재하지만, 수학교육에서는 확률을 크게 고전적, 빈도적, 주관적 관점으로 해석한다(Borovcnik & Kapadia, 2014; Eichler & Vogel, 2014). 확률에 대한 이러한 세 가지 관점은 나라마다 차이는 있지만 대다수 국가의 교육과정에서 다루어지고 있다. Albert(2003)는 세 가지 해석이 일상생활에서 모두 존재하기 때문에 학생들이 이 세 가지 관점을 모두를 경험해보는 것이 중요하다고 주장했다. 전통적으로는 고전적 확률이 학교 교육과정에서 주로 다뤄졌으며, 최근에는 통계학에 관한 관심의 증가와 기술의 발달로 인해 빈도적 확률이 강조되고 있다(Batanero, Chernoff, Engel, Lee, & Sánchez, 2016). 반면에 주관적 확률은 학교 수업에서 주목받지 못하고 있으며, 주관적 확률에 대한 이해의 부족은 고전적 확률과 빈도적 확률을 적용할 수 없는 상황에서 확률을 사용할 때 학생들에게 혼란을 줄 수 있다(Albert, 2003; Chaput, Girard, & Henry, 2011). 한편, 최근에 빈도적 확률이 강조되고 있다고는 하나 학교수학에서는 여전히 고전적 관점 위주로 다루어지고 있는 현실이다. 이로 인해 현행 학교수학에서는 확률의 개념을 가르치고 있다기보다는 여러 가지 복잡한 사건의 확률을 형식적인 알고리즘의 형태로 계산하는 학습이 여전히 주를 이루고 있다(우정호, 2017). 확률의 고전적 관점만 강조해서 다루게 되면 학생들은 확률 개념을 우연 현상과 거리가 먼 것으로 느끼게 되며(이경화, 1996), 모든 확률적 상황에서 고전적 관점이 적용된다는 착각에 빠질 위험이 있다(Albert, 2003).

맥락과 유리된 고전적 확률만을 가르치는 현상을 지적하며, Stohl(2017)은 확률값을 고전적 확률로만 접근하게 되면 세기 전략에 따라 어떤 사건의 확률과 관련하여 하나의 이론적인 답만을 구하도록 하며, 그 값의 실제적인 해석을 피하게 되는 문제점이 있다고 하였다. 또한, Ahlgren & Garfield(1991)는 “실생활 상황에서의 실제적인 관심사와 관련된 맥락에서 가르치지 않는다면, 확률을 교육과정에 포함시키는 것은 쓸모없는 일이 될 것”이라고 주장하였다. 이에 여러 학자들은 확률교육에서 세 가지 관점을 다룰 수 있는 실제적인 방안에 관하여 다양한 연구 결과를

제시하였다. Shaughnessy(1992)는 확률적 상황에 따라 학생들이 이해하고 사용할 수 있기를 바라는 규범적인 확률론적 아이디어가 있다고 설명하며, 동전 던지기와 같은 상황은 고전적 확률과 빈도적 확률을 연결 지어서 다뤄질 수 있고, 압정 던지기와 같이 이론적인 확률값이 없는 상황은 시뮬레이션을 활용하는 빈도적 확률로 다뤄지는 것이 적합하다고 보았다. Eichler & Vogel(2014)은 정육면체 주사위와 직육면체 주사위를 던지는 상황을 모델링 하여 세 가지 관점 모두를 다룰 수 있는 방안을 제시하였다.

하지만, 학교수학에서 고전적 관점이 다른 두 관점 보다 강조되어 있는 현재 상황에서 빈도적 확률의 가치가 더욱 드러날 수 있도록 하기 위해서는 확률 교육과정에서 수학적 확률이 정의되지 않는 사건에 대해 확률을 추정해 보는 확률적 상황이 첨가되어야 한다(신보미 & 이경화, 2006). 또한 향후 2년 안에 암 치료법이 발견될 확률과 같이 등확률의 선험적 가정에 의하거나 경험에 의한 데이터로는 확률값에 대한 판단이 불가능하고 주관적인 추정만 가능한 상황도 존재한다(Jones at el., 2007). 즉, 빈도적 확률과 주관적 확률이 확률 교육과정에서 의미 있게 다루어지기 위해서는 등확률성이 적용되지 않는 상황에서 확률값을 추정하는 모델링 상황을 통해 확률의 개념을 지도해야 한다.

Shaughnessy(1992)는 확률교육의 성공 여부는 교사에게 달려있다고 강조하며 교사 전문성 개발과 교사 교육에 의미 있는 시사를 주는 확률과 관련된 교사 지식에 대한 연구가 더 많이 필요하다고 주장하였다. Stohl(2017)은 교사가 확률에 대한 고전적 접근과 하나의 답만을 강조하는 결정론적 사고방식에 빠져들지 않도록 부단히 노력해야 한다고 주장하였다. 또한 학자들은 교사들이 수학적인 배경에서 관습적으로 형성되어 온 접근에 따라 통계를 하나의 옳은 방법, 하나의 옳은 답으로 다루려는 시도에서 벗어나야 한다고 주장한다(Hawkins, 1990: Greer & Mukhopadhyay, 2017에서 재인용). Eichler & Vogel(2014)은 교사가 확률교육에서 가져야 할 지식 중 확률에 대한 세 가지 접근 방식 간의 연결에 대한 지식이 필수적이라고 보았다. Jones at el.(2007)은 교사들이 고전적, 빈도적, 주관적 관점의 확률을 종합하여 다루는 활동들이 편안하도

록 느끼는 것이 중요하다고 주장하였다. Borovcnik & Kapadia(2014)는 확률에 대한 역사적이고 철학적인 논쟁 과정을 살펴보는 것이 교사들에게 가치 있는 관점을 제공할 것이라고 주장하였다. 왜냐하면 확률에 대한 세 가지 관점을 알고 있는 교사는 맥락에 따라 확률의 어떠한 접근 방식이 적절한지 학생들이 판단하는 데 도움을 줄 수 있기 때문이다.

종합하면 확률교육에서 고전적, 빈도적, 주관적 관점이 종합적으로 다루어져야 하며 상대적으로 소홀히 다루어지고 있는 두 관점을 다루기 위해서는 등확률성이 적용되지 않는 상황이 도입되어야 한다는 것이 많은 연구자들의 공통된 주장임을 알 수 있다. 확률교육의 성패는 교사에게 달려있으므로 교사가 가지고 있는 확률에 대한 세 가지 관점에 관한 연구가 필수적이다. 하지만 Jones at el.(2007)은 확률에 대한 주관적 접근에 관한 인지 연구를 찾을 수 없다고 설명하며, 후속 연구에 대한 조언으로 주관적 확률과 그것을 개념화하는 방식을 조사하는 것이 시의적절하다고 주장하였다. 교사들의 수업 전문성은 예비교사 시기에 형성된 기초적인 것들로 시작되기 때문에 예비교사에 대한 확률에 대한 관점을 살펴보는 것은 교사의 확률에 대한 이해를 살펴볼 수 있게 한다(Hoaglund, Birkenfeld, & Box, 2014). 또한 확률론을 처음 접하는 예비교사를 대상으로 확률에 대한 관점을 조사하면 확률에 대한 지식의 정도가 정규 교육과정의 고등학교 과정까지 학습한 경험에 기반하므로 교육과정의 목표에 얼마나 도달했는지도 살펴볼 수 있다.

이경화(1996)는 확률교육에서 바람직한 교수학적 변환의 방향을 “고전적 관점에만 치중하지 않고 빈도적 관점의 의미를 보다 살려내며, 주관적 측면을 다소간 고려하는 것”이라고 주장하였다. 이경화, 유연주, & 탁병주(2022)는 교육과정 재구조화 방향을 탐색하면서 데이터 기반 확률교육을 지향해야 한다고 주장하였다. 베이즈 정리는 데이터 기반 확률 계산법의 한 형태로서 주관적 확률과 빈도적 확률을 연결 짓는 고리가 된다. 따라서 2022 개정 교육과정에 맞춰 데이터 기반 확률교육으로 나가기 위한 밑거름으로서 예비교사의 확률에 대한 접근 방식을 살펴보는 것은 충분히 의미 있는 연구가 될 것이다.

## 2. 연구 문제

본 연구의 목적은 다양한 확률적 상황에서 예비교사가 나타내는 확률에 대한 접근 방식을 분석하는 것이다. 우선 확률에 대한 예를 구성하는 활동 및 확률에 대한 의미를 제시하는 활동에서 예비교사가 현재 가지고 있는 확률에 대한 관점을 살펴본다. 이후 등확률성이 만족되지 않는 확률적 상황에 접하였을 때, 예비교사는 확률의 어떠한 접근 방식을 이용하여 확률값을 추정하는지 살펴보는 것이다. 마지막으로 확률에 대한 추가적인 정보가 주어졌을 때 확률값을 조정해나가는 과정에서 드러난 예비교사의 사고 과정을 살펴볼 것이다. 이에 따른 연구 문제를 설정하면 다음과 같다.

첫째, 확률에 대한 뜻과 예를 제시하는 활동에서 드러나는 예비교사의 확률에 대한 접근 방식은 무엇인가?

둘째, 등확률성이 만족되지 않는 상황에서 예비교사는 어떠한 접근 방식을 활용하여 확률을 할당할 것인가?

셋째, 확률에 대한 추가적인 정보가 주어졌을 때, 확률값을 추정해나가는 과정은 어떠한 양상을 띠는가?

## 3. 연구의 제한점

본 연구의 제한점은 다음과 같다.

첫째, 본 연구는 서울특별시에 있는 하나의 대학교에 재학 중인 예비교사들을 대상으로 진행하였다. 또한 표본의 수가 적으므로 설문조사의 신뢰도를 높이기 위해서는 여러 지역에 있는 대학교로 조사 범위를 확대하여 더 많은 수의 표본에 대한 조사가 필요하다.

둘째, 예비교사의 여러 확률적 상황에 대한 접근 방식을 설문지 형식을 통해 조사하였다. 각 문항에 대한 이유를 서술했으나 본인이 가지고 있는 생각 모두가 글에서 다 드러난다고 볼 수는 없으므로 예비교사의 확률에 대한 사고 과정의 일부만 살펴본 연구가 될 것이다.

# Ⅱ. 이론적 배경

이 장에서는 본 연구의 기본이 되는 이론적 배경에 대하여 살펴보고자 한다. 이를 위해 첫째, 수학교육에서 확률에 대해 주로 다루는 세 가지 관점인 고전적, 빈도적, 주관적 확률과 함께 확률의 통계적 측면과 인식론적 측면을 더욱 세분화한 여러 학자들의 관점을 살펴보고, 이를 종합하여 파악해본다. 둘째, 확률의 세 가지 관점에서 드러나는 주관적인 특성을 살펴봄으로써, 확률이 지니는 근본적인 애매성에 대해서 살펴본다. 셋째, 통계적 소양과 확률적 소양을 살펴보고 확률적 소양을 함양하는 데 필요한 요소를 확인할 것이다. 넷째, 확률에 대한 교사 지식의 관점에서 확률에 대한 세 가지 접근의 의미를 살펴볼 것이다. 마지막으로, 등확률성이 만족되지 않는 상황에서 확률값을 추정하는 과정을 설명하는 모델을 살펴볼 것이다.

## 1. 수학교육에서 확률에 대한 세 가지 관점

수학교육의 모든 연구는 탐구하는 대상에 대한 인식론적 성찰에 의해 뒷받침되어야 한다(Batanero et al., 2016). 특히, 학교 교육과정에서 확률을 실제로 가르치는 데 반영되는 확률 개념에는 다양한 관점이 존재하기 때문에 확률 해석에 대한 인식론적 성찰은 더욱 의미가 있다. 확률의 다양한 관점에 대해서는 여러 학자들에 의해 활발히 연구가 되었으며, 대체로 직관적, 고전적, 빈도적, 주관적, 논리적, 성향적, 공리적 해석으로 분류되고 있다(Batanero et al., 2016). 하지만 수학교육 분야에서 주로 다루는 확률에 대한 해석은 고전적, 빈도적, 주관적 확률에 대해서이며(Chernoff & Russell, 2014), 최근에 들어서야 이 세 가지 관점을 바탕으로 확률교육에 관한 다양한 연구들이 진행 중이다(Albert, 2003; Eichler & Vogel, 2014; Kazak & Leavy, 2018). 수학교육에

서의 고전적, 빈도적, 주관적 확률에 대한 서술이 학자들마다 조금씩 차이가 있으므로 이 절에서는 Borovcnik & Kapadia(2014)의 연구에 기반하여 확률에 대한 세 가지 관점의 특징 및 한계에 관하여 서술한다.

### 1.1. 고전적 확률

고전적 확률은 Laplace에 의해 최초로 명시적으로 정의되었다. 사건 $A$에 대한 확률 $P(A)$는 주어진 사건이 일어날 횟수와 일어날 수 있는 모든 사건의 횟수의 비로 정의하였다. 이 정의는 모든 개별 사건은 등확률성을 가진다는 암묵적인 가정을 사용한다. 또한 확률을 구하려는 사건에 해당하는 시행을 경험적으로 실시하지 않아도 확률값을 계산할 수 있다는 점에서 선험적 접근 방식이다. 기하학적 확률이 선험적 접근 방식의 전형적인 예로 볼 수 있다(이경화, 1996).

고전적 확률을 적용하면서 직면하는 문제점은 서로 다른 사건의 등확률성을 판단하는데 주관적인 해석이 개입된다는 점이다(Batanero at el., 2017). Laplace는 '이유 불충분의 원리[1](the principle of insufficient reason)'를 제시하며 각각의 사건에 등확률성을 적용해야 한다고 주장하였다. 그러나 동일한 물리학 실험이라도 여러 가지 서로 다른 대칭을 나타낼 수 있으며, 주사위 던지기와 같이 간단한 상황을 제외하면 고전적 확률을 적용하는 방법이 명확하지 않다는 문제가 있다. 예를 들어 보험회사는 <표 Ⅱ-1>와 같은 생명표를 활용하여 생명 보험 상품을 개발한다. 생명표는 연령마다 사망확률을 다르게 계산한다는 점에서 등확률성을 적용하여 관련된 확률을 구하는 것은 실제와 동떨어진 결과를 낳을 수 있다는 문제점이 존재한다.

1) 한 사건이 다른 사건보다 더 많이 일어난다는 기대할 만한 어떤 이유가 없는 경우에는 가능한 모든 사건에 동일한 확률을 할당해야 한다는 원칙이다(Borovcnik, Bentz, & Kapadia, 1991).

<표 Ⅱ-1> 대한민국의 2021년 완전 생명표[2](통계청, 2022)

| 연 령 Age | 사망확률 ${}_nq_x$ | 생존자수 $l_x$ | 사망자수 ${}_nd_x$ | 정지인구 ${}_nL_x$ | 총생존년수 $T_x$ | 기대여명 $e_x^o$ | 연령 Age |
|---|---|---|---|---|---|---|---|
| 0 | 0.00238 | 100,000 | 238 | 99,805 | 8,361,393 | 83.61 | 0 |
| 1 | 0.00014 | 99,762 | 14 | 99,755 | 8,261,588 | 82.81 | 1 |
| 2 | 0.00012 | 99,747 | 12 | 99,741 | 8,161,833 | 81.83 | 2 |
| 3 | 0.00010 | 99,736 | 10 | 99,731 | 8,062,092 | 80.83 | 3 |
| 4 | 0.00008 | 99,726 | 8 | 99,721 | 7,962,361 | 79.84 | 4 |
| 5 | 0.00007 | 99,717 | 7 | 99,714 | 7,862,640 | 78.85 | 5 |
| 6 | 0.00007 | 99,710 | 7 | 99,707 | 7,762,926 | 77.86 | 6 |
| 7 | 0.00006 | 99,703 | 6 | 99,700 | 7,663,219 | 76.86 | 7 |
| 8 | 0.00006 | 99,697 | 6 | 99,695 | 7,563,519 | 75.86 | 8 |
| 9 | 0.00006 | 99,692 | 6 | 99,689 | 7,463,825 | 74.87 | 9 |
| 10 | 0.00006 | 99,686 | 6 | 99,683 | 7,364,136 | 73.87 | 10 |
| ⋮ (하략) | | | | | | | |

## 1.2. 빈도적 확률

빈도적 확률은 시행을 반복한 후 해당 사건이 관측된 상대도수로부터 구한다. 하지만 이 방법으로는 확률값을 정확하게 구할 수는 없고, 다만 추정만 가능할 뿐이다. 실제로 시행을 해본 다음에 얻은 정보에 기반하기 때문에 실험적이고 경험적인 접근 방식이다. 빈도적 확률의 발전에 가장 크게 이바지한 von Mises의 확률에 대한 정의를 살펴보도록 한다(송하석, 2005).

빈도적 확률을 정의하기 위해 von Mises는 집산(collective)이라는 개념을 도입한다. 집산이란 “색깔이나 수 등과 같은 관찰 가능한 속성에 의해서 구별되는 균일한(uniform) 사건이나 시행의 연쇄를 가리킨다(von Mises, 1957: 송하석, 2005에서 재인용).” 즉, 집산은 특정한 ‘임

2) 연령에 대한 계급의 크기가 1인 생명표를 완전 생명표라 한다.

의성'을 가진 '유사한' 사건들이 무한히 나열된 수열로 볼 수 있다(Borovcnik at el., 1991). von Mises는 집산을 경험적 집산(empirical collective)과 수학적 집산(mathematical collective)으로 구분 지었다. 경험적 집산이란 실제 세계에서 특정한 장소와 시각에서 관찰되는 사건들의 나열로서 유한 번의 시행으로 구성된다. 반면에 수학적 집산은 무한한 시행의 연쇄로 구성된다. von Mises는 유한한 경험적 집산을 추상화 또는 이상화하여 무한한 수학적 집산으로 표상된다고 설명한다. 이러한 논의를 바탕으로 von Mises는 개별 사건을 집산에 포함시킴으로써 불확실성을 측정할 수 있다고 보았으며, 상대도수가 향하는 극한으로 확률을 정의한다. 하지만 시행의 횟수가 크더라도 유한인 경험적 집산이 시행의 횟수가 무한인 수학적 집산으로 표상되는 것이 정당한지에 대한 비판이 제기된다(Gillies, 2000).

빈도적 확률의 정의를 적용하는데 직면하는 문제점은 개별 사건이 확률이 같지 않은 서로 다른 집산에 포함될 수 있다는 점이다. 즉, 개별 사건이 집산에 포함되는 과정에서 집산의 선택에 대한 정당화가 필요하게 된다. 보다 근본적으로, 반복적으로 실행을 수행할 수 없는 사건이 많이 존재한다. 특히 확률이 낮은 사건에서 더욱 그러하다. 빈도적 확률에 대한 비판은 크게 세 가지가 있다(Borovcnik & Kapadia, 2014). 첫째, 빈도적 확률은 원칙적으로 계산되지 않는다. 둘째, 빈도적 확률이 실제로 존재하는지 알 수 없다. 마지막으로 잠정적으로 계산된 빈도적 확률값은 확인되거나 거부할 수 없게 된다. 반면에 빈도적 확률은 등확률성을 적용할 수 없는 상황에서 확률을 계산하는 실용적인 수단을 제공한다. 어떤 면에서는 현대 물리학과 생물학과 같이 특정 분야에서는 더욱 과학적으로 여겨진다.

### 1.3. 주관적 확률

주관적인 관점에서 확률을 설명한 Keynes, Ramsey, de Finetti는 확률을 "주어진 결과에 대해 개인이 갖고 있는 정보에 의해 결정되는 개

인적인 신념이라고 설명하였다(Cabria, 1992; Hacking, 1975: Batanero at el., 2017에서 재인용).” 이 정의에 담긴 기본적인 가정은 개인이 의사결정을 내릴 때 암묵적인 선호 패턴에서 파생되는 자신만의 확률을 가지고 있다는 것이다(Borovcnik at el., 1991). 이 아이디어는 도박에서 비롯되었는데, 개인은 도박에 돈을 걸어서 생기는 배당률을 근거로 주관적인 확률을 결정하고 행동한다는 것이다. de Finetti(1974)는 “확률은 존재하지 않는다.”라고 주장하며 객관적인 확률의 존재를 부정한다. 따라서 확률은 이론적인 개념으로서 추정된 확률값은 관찰자의 지식, 관찰되는 상황, 관찰자가 수집할 수 있는 자료와 같이 다양한 변수에 의해 결정된다(Batanero at el., 2017). 개인마다 가지고 있는 확률이 다를 수 있지만 확률의 공리에 위반되지 않는 일관성만 유지한다면 어떠한 확률값을 할당해도 무방하다.

개인이 주관적 확률을 판단할 때 사용하는 정보를 두 가지로 분류할 수 있는데, 경험적 데이터와 독립적으로 개인의 마음에 내재하는 사전 정보와 반복적인 실험의 빈도에서 나온 경험적 정보로 나뉜다. 두 가지 유형의 정보는 베이즈 공식에 의해 결합되어 새로운 확률을 제시한다. 이렇게 확률이 갱신되는 과정을 ‘경험으로부터의 학습’이라고 불린다. 베이즈 공식은 단순한 합리성의 기준과 결합하여 경험과 사전정보를 이용하여 의사결정을 내릴 수 있도록 도와준다. 구체적으로, 사건 $E$에 대한 선험적 확률인 사전 확률 $\mathrm{P}(E)$를 새로운 데이터 $D$에 근거하여 후험적 확률인 사후 확률 $\mathrm{P}(E|D)$로 갱신하는 계산 방법을 베이즈 정리로 표현하면 다음과 같다(이경화 외, 2022).

$$\mathrm{P}(E|D) = \frac{\mathrm{P}(D|E)\mathrm{P}(E)}{\mathrm{P}(D)}$$

고전적 확률과 빈도적 확률은 결점이 있더라도 확률을 계산하는 데 실용적인 방법을 제공하는 반면에, 주관적 확률은 사전 확률을 측정하는 방법에 대한 지침을 제공하지 않는다는 문제점을 가지고 있다. 고전적 확률과 빈도적 확률에서 계산하는 방식을 사전 확률에 적용할 수 있고,

일관성의 원칙이 위배되지 않는 선에서 개인마다 사전 확률을 주관적으로 할당할 수도 있다. 주관적 확률에 대한 본질적인 비판은 객관적인 근거의 결여와 본질적인 주관성이다. 하지만 등확률성이 보장되지 못하는 상황과 동일한 실험을 반복할 수 없는 상황에서 주관적 확률을 적용할 수 있다. 예를 들면, 일부의 지식만 가지고도 판단을 내려야 하는 비즈니스 의사결정 상황이나 금융 위기와 전염병의 확산과 같은 위기에 대한 대처방식에 대해 고려할 때도 적용될 수 있다(Batanero at el., 2016; Eichler & Vogel, 2014).

### 1.4. 확률의 통계적 측면과 인식론적 측면

앞서 살펴본 수학교육에서의 확률에 대한 세 가지 관점은 각각의 장단점을 가지고 있으며 상호보완적인 속성을 지니고 있다. 확률 측정은 철학, 심리학, 수학, 수학교육에서 독자적인 분야로 발달해 왔기 때문에 같은 용어를 사용하더라도 분야마다 다른 정의를 갖게 되었다. 이에 확률이론을 탐구한 Hacking(2001)과 Gillies(2000)의 견해를 살펴보고 현재의 수학교육에서의 확률에 대한 세 가지 관점을 다시 조명해보고자 한다.

Hacking(1975)은 앞서 언급하였듯이 우연 현상에서 확률의 법칙과 관련된 통계적인 측면과 통계적 배경이 없는 명제들에 대한 합리적인 믿음의 정도를 나타내는 인식론적인 측면으로 두 가지 관점을 구분 지었다. 이러한 두 가지 분류는 여러 과학철학자에게서 나타났으며 이를 부르는 용어는 학자마다 달랐지만, Hacking(2001)은 빈도형(frequency-type) 확률과 신념형(belief-type) 확률로 명명하였다. 두 확률을 구분 짓는 기준으로 객관성의 개념이 사용된다. 여기서 말하는 객관성은 두 가지 용도로 사용되는데 첫째는 빈도형 확률과 신념형 확률을 분류하는 기준으로 쓰이고, 둘째는 각각의 유형의 확률을 추가적으로 이중분류할 때 사용된다.

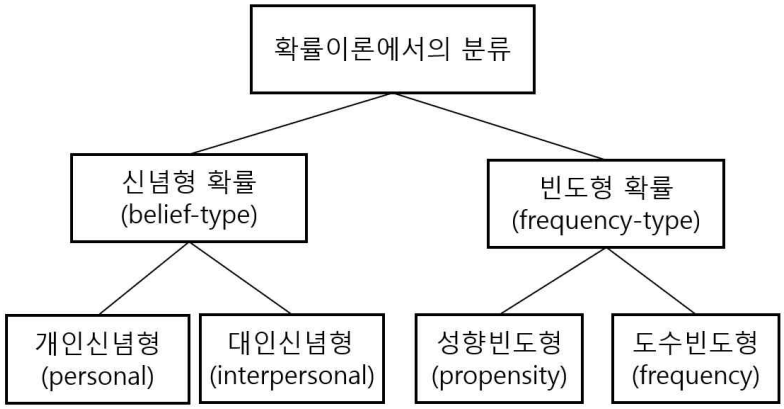
빈도형 확률과 신념형 확률을 분류하기 위해 쓰이는 첫 번째 의미로

서의 객관성은 인간의 신념이나 지식과는 아무런 관련이 없는 객관적인 물질세계의 특징으로 볼 수 있다(Gillies, 2000). 빈도형 확률을 다룰 때 동전의 특성을 살펴보는 것이 객관적인 물질세계의 특징을 살펴보는 예가 된다. 반면에 신념형 확률을 다룰 때는 인간의 신념이나 지식이 사용된다. 따라서 신념형 확률을 측정할 때는 지식의 정도, 합리성의 정도 신념의 정도를 이용한다(Gillies, 2000). 즉, 두 종류의 확률의 분류기준은 확률을 판단할 때 사용되는 근거가 물질에 기인하는지, 인간의 마음이나 사상에서 기인하는지에서 비롯된다. 이제 빈도형 확률과 신념형 확률에서의 각각의 이중분류에서 사용되는 두 번째 의미에서의 객관성을 살펴보기 위하여 먼저 빈도형 확률을 분류하는 객관성의 의미를 살펴본다.

Hacking(2001)은 빈도형 확률이 추가적으로 성향(propensity) 이론과 도수(frequency) 이론에 의해 분류된다고 설명한다. Hacking(2001)은 앞면이 더 잘 나오는 편향된 동전을 예시로 두 확률을 어떻게 구분지을 수 있는지 설명한다. 성향 이론은 동전을 던졌을 때 앞면이 뒷면보다 더 잘 나오는 성향 또는 경향을 지녔다는 등 물리적 성질에 대해 일반적으로 진술한다. 반면 도수 이론은 10번 던졌을 때 앞면이 7번 나오는 상대도수를 살펴보는 등 구체적인 진술을 통해 확률을 구한다. 도수 이론은 성향 이론에 비해 구체적으로 얼마나 더 앞면이 나오는지 알 수 있기 때문에 객관성을 더 지닌다고 볼 수 있다. 이렇듯 빈도형 확률 내의 분류에서의 객관성은 구체적인 수치를 지녔는지에 달려있다.

빈도형 확률과 마찬가지로 Hacking(2001)은 신념형 확률을 추가적으로 개인(personal) 신념형 확률과 대인(interpersonal) 신념형 확률로 이중분류될 수 있다고 보았다. 예를 들어, “아마도 내일 나는 버스를 탈 것이다.”와 “땅에 눈이 쌓여있으니, 아마도 내일 나는 버스를 탈 것이다.”라는 두 진술을 비교해보면, 증거의 유무로 두 문장을 구별할 수 있다. 전자는 마땅한 증거 없이 개인의 신념만으로 판단을 하기 때문에 개인신념형 확률로 분류될 수 있다. 후자는 증거에 대한 논리적 판단이 수반되므로 주어진 증거에 따라 사람들이 비슷한 결론을 내릴 수 있으므로

대인신념형 확률이다. 즉, 주관적 확률에서 대인신념형 확률로 분류되기 위해서는 증거에 기반한 판단이어야 하고, 그 증거가 개인 혼자가 아닌 여러 사람에게 받아들여질 수 있어야 한다. 더 많은 사람들이 비슷한 판단을 내린다는 관점에서 대인신념형 확률이 개인신념형 확률보다 더 객관적이다. Chernoff(2008)는 Hacking(2001)의 네 가지 분류를 [그림 Ⅱ-1]과 같이 도식화하였다.




[그림 Ⅱ-1] 확률이론에서의 이중분류(Chernoff, 2008)

Chernoff(2008)는 수학교육과 확률이론에서 각각의 분야에서 쓰이는 용어가 매우 풍부하다고 서술하면서, 이 분야의 용어를 연결 짓는 작업의 중요성을 강조하며 <표 Ⅱ-2>와 같이 제시하였다.

<표 Ⅱ-2> 확률 측정에 관한 용어(Chernoff, 2008)

| 수학교육 | |
|---|---|
| 선험적, 이론적 | 고전적 |
| 후험적, 실증적, 실험적, 객관적 | 빈도적 |
| 베이즈주의, 직관적, 개인적, 개인의 | 주관적 |

⟷

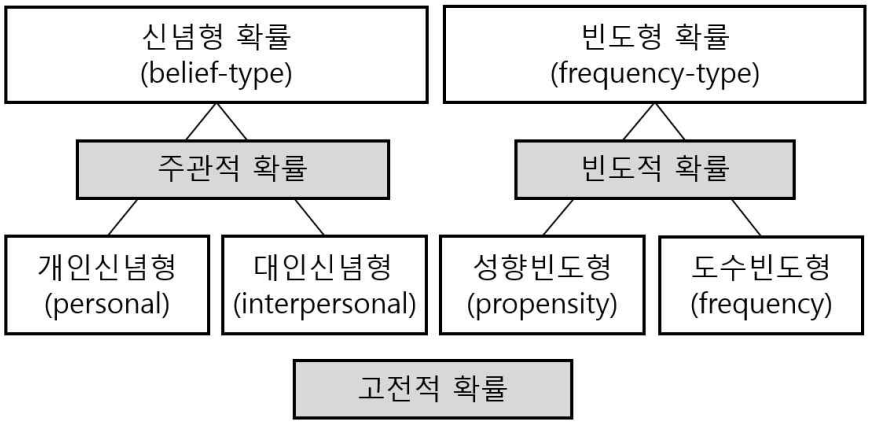
| 확률이론 | |
|---|---|
| 고전적 | 선험적, 이론적 |
| 빈도형 | 후험적, 객관적, 성향빈도, 도수빈도 |
| 신념형 | 주관적, 인식론적인, 개인신념, 대인신념 |

수학교육에서의 고전적, 빈도적, 주관적 확률은 확률이론에서의 고전적, 빈도형, 신념형 확률과 각각 대응할 수 있다. 하지만 확률이론에서 빈도형 확률을 세분화한 성향빈도형 확률과 도수빈도형 확률은 수학교육 분야에서 대응하는 용어는 없고, 확률이론에서 신념형 확률을 세분화한 개인신념형 확률과 대인신념형 확률에 대응하는 용어는 수학교육 분야에서는 없다. Chernoff(2008)는 확률이론에서 빈도형 확률을 세분화하듯이 수학교육에서도 빈도적 확률을 세분화하자고 주장한다.

한편, Jones at el.(2007)는 수학교육 분야에서의 주관적 확률에 대한 통일된 정의가 없다고 주장한다. 주관적 확률이 빈도적 확률보다 복잡한 이유는 확률이론뿐만 아니라 심리학 분야에도 영향을 받기 때문이다. Kahneman & Tversky(1982)는 주관적 확률을 어떠한 사건에 대하여 개인이 확률을 할당하거나, 개인의 행동에서 유추되는 확률의 추정치라고 하였다. 수학교육에서 설명한 주관적 확률을 살펴보면, Jones at el.(2007)은 결과에 대한 개인적인 판단과 정보에 의해서 결정되는 신념도로 보았고, Batanero at el.(2017)은 특정 개개인이 느끼는 불확실성의 정도를 나타내는 것이라 설명하였다. 이러한 정의에 따르면 수학교육에서의 주관적 확률은 개인신념형과 대인신념형을 구분하지 않고, 대인신념형 확률보다는 개인신념형 확률을 지칭하는 것으로 판단된다. 하지만 확률이론에서 존재하는 용어 및 개념이 수학교육에서 존재하지 않으면 주관적 확률은 여전히 혼동되는 상태로 남아있게 되므로 Chernoff(2008)는 수학교육에서의 주관적 확률도 세분화해야 한다고 주장한다.

이러한 Chernoff(2008)의 견해를 받아들여 수학교육에서도 Hacking(2001)이 제시한 객관성을 기준으로 주관적 확률과 빈도적 확률로 나누고, 각각에서의 객관성을 기준으로 개인신념형, 대인신념형, 성향빈도형, 도수빈도형 확률로 세분화할 수 있다. 하지만 이렇게 객관성을 기준으로 확률을 4가지로 분류하면 고전적 확률이 어느 영역에 속하는지에 대한 문제가 생긴다. Gillies(2000)는 고전적 확률에서 확률에 대한 판단 근거는 인간의 정신인 플라톤 세계에 존재하므로 신념형 확률

에 속해야 한다고 보았다. 더 세부적으로는 고전적 확률은 개인마다 다른 신념도를 가지지 않으므로, 개인신념형 확률보다는 대인신념형 확률에 속한다고 보는 것이 바람직하다고 보았다. 하지만 Chernoff(2008)는 Hacking(2001)이 Laplace 시대부터 용어는 존재하지 않았어도 확률의 양면적인 측면을 인식하고 있었다는 지적에 고전적 확률이 빈도형 확률의 범위에도 속할 수도 있다는 의견을 제시한다. 이렇듯 고전적 확률은 대인신념형 확률에만 속하는지, 혹은 대인신념형 확률과 성향빈도형 확률에 걸쳐있는지 논쟁이 있지만, Chernoff(2008)가 주장한 바를 도식화하면 [그림 Ⅱ-2]와 같이 나타낼 수 있다.




[그림 Ⅱ-2] Chernoff의 두 분야에서 확률에 대한 분류도

Chernoff(2008)는 수학교육에서도 확률이론에서의 네 가지 분류를 도입할 것을 두 가지 장점을 제시하며 주장한다. 첫째로, 고전적 확률이 확률이론에서 신념형 확률에 속하는지 빈도형 확률에 속하는지에 대한 논쟁을 이중분류를 도입함으로써 논의를 더욱 명확히 할 수 있다고 설명한다. 둘째로 수학교육에서 용어의 부재로 잘 다루지 못하였던 빈도적 확률의 주관적인 측면인 성향빈도형 확률과 주관적 확률에서의 객관적인 측면인 대인신념형 확률에 대하여 개념을 논의할 수 있다는 점을 들어 수학교육에서의 확률에 대한 논의를 풍부하게 할 수 있을 것이라 기대하였다.

## 2. 확률의 각 관점에서 드러나는 주관적 특성

Chernoff(2008) 객관성을 기준으로 확률을 네 가지로 세분화하는 과정을 살펴보았다. 한편, 이경화(1996)는 확률 개념이 판단 또는 신념이라는 추상적인 개념을 수로 표시하였기 때문에 주관적인 측면과 객관적인 측면이 뒤섞이게 되었고, 이는 확률이 근본적으로 애매성을 가지게 되었다고 설명한다. 이에 여러 철학자가 제시한 확률의 주관적인 특성을 고전적, 빈도적, 주관적 확률의 관점에서 살펴보도록 한다.

### 2.1. 고전적 확률에서의 주관적인 특성

고전적 확률에서 나타나는 주관적인 특성에 대하여 Batanero at el.(2017)은 서로 다른 사건의 등확률성을 판단하는데 주관적인 해석을 토대로 하였다는 점에서 비판받아왔다고 주장한다. Fine(1973)은 두 개의 주사위를 던지는 사건에 대해서 다음과 같은 3가지 모델을 제시한다(Borovcnik & Kapadia, 2014에서 재인용).

Maxwell-Boltzmann 모델: 다음의 36개의 순서쌍은 나올 확률이 모두 같다. (1, 1), …, (1, 6), (2, 1), …, (2, 6), …, (5, 1), …, (5, 6), (6, 1), …, (6, 6); 이 모델에서 순서쌍 (2, 3)과 (3, 2)는 서로 다른 사건으로 본다.

Bose-Einstein 모델: 다음의 21개의 순서쌍은 나올 확률이 모두 같다. (1, 1), …, (1, 6), (2, 2), …, (2, 6), …, (5, 5), (5, 6), (6, 6); 이 모델에서 순서쌍 (2, 3)과 (3, 2)는 같은 사건으로 본다.

Fermi-Dirac 모델: 다음의 15개의 순서쌍은 나온 확률이 모두 같다. (1, 2), …, (1, 6), (2, 3), …, (2, 6), …, (5, 6); 이 모델에서는 같은 수의 눈은 나올 수가 없다.

수학교육에서는 Maxwell-Boltzmann 모델을 기본으로 다루지만, 실제 물리학 분야에서는 이 모델을 따르지 않는 사례가 많이 발견되었다.

본질적으로 구별할 수 없는 짝수의 기본 입자를 포함하는 광자 및 원자의 움직임은 Bose-Einstein 모델에 의해 가장 잘 설명된다. 또한 본질적으로 구별할 수 없고 각 상태가 단일 입자만 포함할 수 있는 전자, 양성자 및 중성자는 Fermi-Dirac 모델에 의해 가장 잘 설명된다 (Borovcnik & Kapadia, 2014). 이렇듯 대칭은 모델링의 목적에 따라 다르며 확률은 실제 대상의 고유한 특성이 아니라 현실을 모델링하려는 노력의 개념적 결과로만 보아야 한다고 설명한다.

## 2.2. 빈도적 확률에서의 주관적인 특성

빈도적 확률은 집산의 개념을 통해 상대도수의 극한으로 확률을 계산하였다. 개별 사건이 집산에 포함되는지를 판단하는 과정에서 주관적인 특성이 나타나게 된다. 송하석(2006)이 제시한 다음 상황을 살펴보며 집산에 따른 빈도적 확률의 주관적인 특성을 살펴보도록 한다. 정상적인 주사위와 6의 눈이 나올 확률이 $\frac{1}{4}$이 나오는 편향된 주사위가 있다고 하자. 편향된 주사위를 500회 이상 던지고, 그 사이에 정상적인 주사위를 10번 던져서 만들어진 수열을 생각해 보자. 이 중에서 정상적인 주사위를 던진 경우를 택해서, 그 경우에 6이 나올 확률이 얼마인가를 묻는다고 하면 von Mises의 집산의 성질에 의해 그 확률은 $\frac{1}{4}$이 되어야 한다. 왜냐하면 그 시행은 6이 나올 확률이 $\frac{1}{4}$인 집산의 한 부분이기 때문이다. 이렇듯 어떤 사람은 같은 주사위 던진 시행이므로 일반적인 주사위와 편향된 주사위가 같은 집산에 포함된다고 판단하여 확률을 $\frac{1}{4}$로 계산하고, 다른 사람은 정상적인 주사위만으로 집산을 구성하여 확률을 $\frac{1}{6}$로 판단할 수 있다. 위의 예시와 같이 집산을 판단하는 기준이 개인마다 다를 수 있다.

빈도적 확률의 또 다른 주관적인 특성을 다루는 원리인 '가장 좁은 준거집합의 원칙(principle of the narrowest reference class)'을 살펴보도록 한다(Gillies, 2000). von Mises의 빈도적 확률은 유사한 사건을 반복적으로 수행해야 하므로 단일 사건에 대한 확률을 판단할 수 없는 한계를 지닌다. 이를 극복하기 위해서 Popper는 '가장 좁은 준거집합의 원칙'을 제안한다. 송하석(2006)이 제시한 다음 예를 통해 이 원리를 살펴보도록 하자. 40세가 된 어떤 남자, 돌이가 41세가 될 때까지 생존할 확률을 계산하는 상황에서 돌이가 단순히 40세가 된 남자라고 판단하거나, 하루에 담배를 한 갑씩 피우는 40세가 된 대한민국 남자로 판단하느냐에 따라 확률값은 달라진다. 후자의 집합이 더 많은 정보를 가지는 부분 집합이기 때문에 더 좁은 준거집합이고 이를 통해 구한 확률이 더 객관적으로 판단하는 원리이다. 하지만 좁은 준거집합을 제시하고 서술할 때 주관적 요소가 개입된다는 문제점이 존재한다.

또 다른 주관적인 특성을 살펴보자. 앞에서 언급하였듯이 Hacking(2001)은 빈도적 확률을 분류할 때, 상대도수와 같이 구체적인 수치에 따라 성향빈도형 확률과 도수빈도형 확률로 세분화할 수 있다고 하였다. Hacking(2001)의 분류는 성향빈도형 확률 내에서 주관적인 특성을 구분하는 기준이 명확하게 주어지지 않는다는 한계가 존재한다. 이에 Gillies(2000)가 빈도형 확률을 세분화하기 위해 도입한 인공적(artefactual) 확률과 형식적 객관적(formal objective) 확률의 개념을 가져와서 빈도적 확률의 주관성을 구분해 볼 수 있다.

Gillies(2000)는 인공적 확률이 물질세계에 존재하는 확률이어서 객관적이지만 인간과 자연의 상호작용의 결과로 나타난 확률이라고 설명한다. 인공적 확률을 도입하게 되면 객관성의 정도에 따라 빈도적 확률을 더 잘 설명할 수 있다고 Chernoff(2008)는 주장한다. 여기에서 말하는 객관성의 정도는 인간과 자연과의 상호작용에 의해서 객관성이 감소하거나, 물리적 혹은 다른 요소에 의해 객관성이 감소한 상황을 말한다.

반면에 형식적 객관적 확률은 인간의 행위가 최대한으로 배제되어 인간과는 독립적인 확률이라고 설명한다. 자연에 존재하는 상태 그대로의

확률이므로 빈도적 확률에서 이상적으로 객관성의 한 극단에 존재할 것이고, 인공적 확률은 인간의 개입에 따른 객관성의 정도에 따라 다양한 스펙트럼을 지니게 된다.

### 2.3. 주관적 확률에서의 객관성 확보

주관적 확률은 같은 상황에 대해서도 다른 확률을 할당할 수 있으므로 본질적으로 주관적이다. 앞서 설명하였듯이 개인신념형 확률인 사전정보와 데이터를 통해 얻은 경험적 정보가 베이즈 정리에 의해 결합되어 새로운 확률을 생성해낼 수 있게 된다. 이렇게 생성된 새로운 확률은 증거가 결합되었으므로 객관성이 확보되는 정도에 따라 대인신념형 확률로 분류되거나 빈도형 확률로 분류될 수도 있다.

베이즈 정리 외에도 주관적 확률에서 객관성을 확보하는 이론을 살펴보자. Hacking(2001)은 주관적 확률과 빈도적 확률을 연결 짓는 원리로서 빈도 원리(frequency principle)를 제시한다. 이 원리를 적용하기 위해선 어떤 사건에 대한 빈도적 확률을 알고 있고, 이 사건에 대한 단일 시행에 관한 결과에 대해서 아는 것이 하나도 없는 상황을 가정한다. 그러면 단일 시행에 대한 사건에 대한 주관적 확률을 이전에 알고 있던 빈도적 확률로 대체할 수 있게 된다는 원리이다. 이를 실생활에 적용하기 위해서는 사건에 관련된 하위 집합을 잘 설정해야 한다. 앞서 제시한 생명표는 생존확률을 빈도 원리에 의해서 판단한 예로 볼 수 있다. 주관적 확률을 빈도적 확률로 전환시켜줄 수 있도록 도와주는 원리라고 볼 수 있다.

Hacking(2001)은 주관적 확률을 분류할 때, 증거의 유무에 따라 개인신념형 확률과 대인신념형 확률로 분류하였다. 그러나 증거가 한 개인의 마음속에만 존재할 때는 위의 기준에 따라 두 확률을 분류하기에는 모호한 부분이 있다. 이를 해결하기 위해 Gillies(2000)가 제안한 개인주관적(intrasubjective) 확률과 상호주관적(intersubjective) 확률을 분류하는 기준을 살펴보자. Gillies(2000)의 개인주관적 확률은

Hacking(2001)의 개인신념형 확률과 같이 확률에 대한 신념이 개인에게만 존재한다. 반면에 상호주관적 확률은 종교집단, 정당과 같은 비슷한 사고방식을 가진 집단은 주어진 확률 상황에 대해서 비슷한 값을 할당한다고 보았다. Gillies(2000)의 주관적 확률에 대한 분류는 확률을 판단할 때 개인 혼자만의 신념인지 집단의 구성원이 비슷한 신념을 바탕으로 판단하는지에 따라 분류하였다. 더 나아가 동일한 증거를 바탕으로 모든 사람이 같은 확률을 제시하는 상황이 상호주관적 확률에서 가장 극단적인 형태라고 설명하였다. 이렇듯 Gillies(2000)는 주관적 확률을 판단할 때, 비슷한 확률값을 내리는 사람이 많아질수록 객관적이라고 판단하였다.

## 3. 확률적 소양

이 절에서는 확률적 소양을 구성하는 요소를 살펴보며 확률적 소양을 함양하기 위해서는 확률에 대한 다양한 접근이 필수적임을 살펴볼 것이다. 2015 개정 교육과정에서 확률과 통계는 현대 정보화 사회의 불확실성을 이해하는 중요한 도구이며, 확률과 통계 학습을 통해 미래를 예측하고 합리적인 의사결정을 하는 민주 시민으로서의 기본 소양을 기르는 것을 목표로 하고 있다(교육부, 2015). Gal(2017)은 '학생들이 확률에 대해 어떤 것을 배워야 하며, 왜 배워야 하는가?'에 대한 질문의 답은 보통 다음의 2가지로 요약된다고 보았다. 첫째, 확률이 수학과 통계의 한 영역으로서 표집과 통계적 유의성과 같은 심화 학습 내용을 학습하거나 다른 과학 분야에서 세부 주제를 이해하는 데 도움이 된다. 둘째, 무작위 사건과 우연 현상은 일상생활에서 늘 마주칠 수 있기에 확률을 배움으로써 학생들이 미래의 삶을 대비하는 데 필수적이다. Gal(2017)은 확률의 외적인 측면인 두 번째 이유에 주목하면서 교육과정의 설계, 교사 연수, 교육 내용 및 평가 등을 고찰할 때 가치 판단의 지침으로서 확률적 소양을 제시한다.

Gal(2017)은 확률적 소양을 실제 세계에서 접하는 확률 문제 상황과

관련하여 학생들이 적절한 소양을 갖추도록 하는데 필요한 지식과 성향(disposition)이라고 보았다. 확률적 소양은 통계적 소양과 밀접하게 관련되어 있지만 확률적 소양보다 통계적 소양이 먼저 주목받기 시작했다. 따라서 확률적 소양을 살펴보기에 앞서 통계적 소양에 대하여 살펴보도록 한다.

### 3.1. 통계적 소양

Wallman(1993)은 미국 통계 협회(American Statistical Association)의 협회장 취임 연설에서 통계적 소양을 다음과 같이 제시한다.

> '통계적 소양'은 일상생활에 녹아있는 통계적 자료를 이해하고 비판적으로 평가하는 능력으로 공적이거나 사적인 의사결정, 전문적이거나 개인적인 의사결정을 내리는 데 통계적 사고가 기여하는 바를 인식하는 능력과 결부된다(Wallman, 1993).

이후 여러 연구자들이 통계적 소양을 기술하는 역량 모델을 제시하는데 대표적으로 Gal(2002)의 모델을 간략히 소개한다. Gal(2002)은 통계를 학습한 성인들이 갖춰야 하는 소양으로서 통계적 소양을 다음과 같이 정의하였다.

> 통계적 소양은 다음의 두 가지 요소가 서로 연관되어 있다. (1) 다양한 맥락에서 접하게 되는 통계적 정보, 데이터와 관련된 주장, 또는 확률적 현상들을 해석하고 비판적으로 평가하는 개인의 능력, (2) 정보의 의미를 이해하거나 정보가 가지는 함의에 대한 의사 표시 또는 제시된 결론을 수용하는 것에 대한 관심 등과 같은, 통계적 정보에 대해 토론하거나 의사소통하는 능력(Gal, 2002).

Gal(2002)은 통계적 소양을 구성하는 모델을 제시하였는데 크게 지

식 요소와 성향 요소로 분류할 수 있다. 지식 요소는 소양적 능력(literacy skills), 통계적 지식, 수학적 지식, 맥락적 지식, 비판적 질문으로 세분화하였고, 성향적 요소에는 신념과 태도, 비판적 자세가 포함된다.

성인 대부분은 통계적 정보의 생산자라기보다는 소비자이기 때문에 의사결정이 필요한 많은 실생활 상황에 대처하기 위해서는 확률적 소양을 지닐 필요가 있다(Gal, 2017). 통계적 소양을 확장하여 Gal(2017)이 제시한 확률적 소양에 관한 모델을 살펴본다.

## 3.2. 확률적 소양

Gal(2017)은 통계적 소양을 구성하는 모델과 비슷하게 확률적 소양에 관한 모델을 지식 요소와 성향 요소로 분류하였다. Gal(2017)은 지식 요소와 성향 요소에 대한 세부 구성 요소를 <표 Ⅱ-3>과 같이 범주화하였다.

<표 Ⅱ-3> 확률적 소양 - 구성 요소(Gal, 2017)

| |
|---|
| **지식 요소** |
| 1. 핵심 개념(Big ideas) : 변이(Variation), 무작위성, 독립, 예측가능성/불확실성 |
| 2. 확률 계산 : 사건에 대한 확률을 추정하거나 찾는 방법 |
| 3. 언어 : 가능성에 대하여 의사소통할 때 사용하는 용어나 방법 |
| 4. 맥락 : 다양한 맥락 속 공적이든 사적인 담화에 등장하는 확률적인 사안과 메시지의 역할 및 암시에 대한 이해 |
| 5. 비판적 질문 : 어떤 사안에서 확률을 다룰 때 행해지는 반성 |
| |
| **성향 요소** |
| 1. 비판적 자세 |
| 2. 신념과 태도 |
| 3. 불확실성과 위험에 대한 개인적인 감정(예, 위험에 대한 반감) |

Gal(2017)은 확률적 소양을 구성하는 요소를 <표Ⅱ-3>과 같이 구분하였지만, 학습 과정이나 실제 상황에서는 복잡한 방식으로 상호작용한다. 따라서 학생들이 확률적 소양을 갖추기 위해서는 한두 가지 요소만 집중하여 지도하는 것을 경계해야 한다고 주장하였다.

앞서 언급했듯이 확률적 소양은 확률교육의 지침으로 작용할 수 있다. Gal(2017)은 대중들이 다양한 확률적 상황에 대처하기 위해서는 확률적 소양을 지닐 필요가 있다고 주장하였으며, 현재 대부분의 학교 교육과정에서 확률을 고전적 관점 및 빈도적 관점과 관련된 형식적 지식만을 전달하는데 집중하는 경향을 반대하였다. 또한 교육부에서 제시한 2022 개정 교육과정 총론 시안에 따르면 다양한 위기 상황에 대한 대처와 같은 예측 불가능한 변화에 대한 대응능력이 필요함을 첫 번째로 내세우고 있다(교육부, 2021). 위기 상황은 실생활과 밀접하게 연관되어 있으며, 선험적으로 등확률성을 적용할 수 없는 상황이 대다수이다. 따라서 빈도적 확률과 주관적 확률을 더욱 비중 있게 다루는 것은 2022 개정 교육과정의 목표와도 부합한다.

## 4. 확률에 대한 교사 지식

이 절에서는 확률교육에서 교사 지식이 중요하며, 확률에 대한 교사 지식에서 확률에 대한 접근 방식이 주요 요소임을 살펴볼 것이다. Shaughnessy(1992)는 확률과 통계에 대한 지난 연구들을 분석하며 후속 연구에 대한 조언으로 확률에 대한 교사의 인식에 관한 연구가 필요하다고 주장하였다. 또한, Stohl(2017)은 학생들의 확률적 추론을 향상시키기 위한 확률 교육과정의 성공 여부는 확률 자체를 교사가 어떻게 이해하고 있느냐에 달려있다고 주장하였다. Stohl(2017)은 교사의 확률에 대한 지식과 확률을 가르치기 위한 지식을 다음의 네 가지 범주로 분류하였다.

(1) 확률에 대한 교사의 신념과 내용 지식
(2) 학생들이 확률을 이해하는 방식에 대한 교사의 이해
(3) 교사들의 확률 수업의 실행
(4) 교사들이 스스로 학습하거나 가르칠 때 사용하는 시뮬레이션 도구의 활용도

Stohl(2017)은 확률에 대한 교사 교육에서 무엇보다도 확률에 대한 내용 지식이 중요하다고 주장하였다. 이에 Kvatinsky & Even(2002)이 제시한 확률과 관련하여 교사들이 알아야 하는 교과 내용 지식의 구조를 살펴본다.

## 4.1. 확률에 대한 교사의 교과 내용 지식의 구조

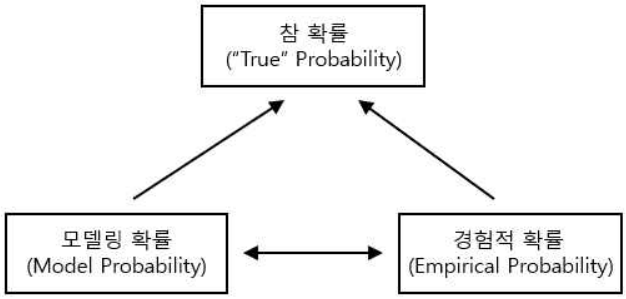
Kvatinsky & Even(2002)은 확률에 대한 교사의 교과 내용 지식을 다음의 7가지 측면으로 서술하였다.

(1) 확률의 본질적 특성 : 비결정론적 현상을 다루는 수학적 방법으로서의 확률의 특성, 확률에 대한 세 가지 접근 방식의 이해
(2) 확률의 영향력 : 확률이 일상적인 상황에서부터 양자 역학에 이르기까지 확률의 지닌 혁명적인 역할에 대한 인식
(3) 다양한 표상 및 모델 : 확률을 구하고 해석하는 데 필요한 표, 벤 다이어그램, 수형도 등의 표상과 모델을 다룰 줄 아는 능력
(4) 상황에 따른 대안적인 접근 방식 : 확률의 세 가지 접근 방식을 상황에 맞게 적용
(5) 전형적인 활동의 예시 : 독립을 설명할 때 사용되는 주사위를 던지는 상황, 조건부확률에서의 몬티홀 문제와 같은 전형적인 예시의 제공
(6) 직관과 확률이론이 다를 수 있음을 이해하는 데 필요한 다양한 지식과 이해
(7) 수학에 대한 기본 지식 : 확률의 공리, 큰 수의 법칙에서의 극한의 의미 등에 대한 이해

Kvatinsky & Even(2002)이 제시한 7가지 측면은 각각이 독립된 요소가 아니라 서로 밀접하게 연관되어 확률에 대한 지식을 형성한다. 측면 (1)인 확률의 본질적인 특성은 확률의 각 접근 방식의 장점과 한계와 연결되어 있고, 측면 (4)인 대안적인 접근 방식은 교사가 각 접근 방식을 적절하게 사용할 수 있는 능력과 관련된다. 이는 교사의 확률에 대한 세 가지 접근 방식에 대한 연구가 필요함을 역설한다.

## 5. 확률 모델링

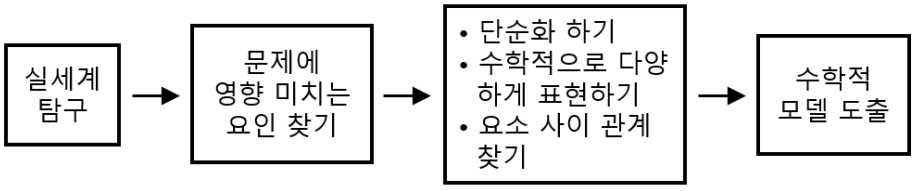
예비교사의 확률에 대한 관점을 탐구하기 위해서는 확률 모델링 개념을 살펴보아야 한다. 확률교육을 모델링의 관점에서의 접근을 강조하는 여러 연구가 수행되었다(Chaput at el., 2011; Eichler & Vogel, 2014). 확률 모델링은 그 자체로서 수학적 모델링 과정일 뿐만 아니라 확률의 주요 관심사로서 확률 지식의 또 다른 본질이기 때문에 주요하다(Batanero & Sánchez, 2005; Pfannkuch, 2005: 신보미 & 이경화, 2008에서 재인용). 또한 Batanero at el.(2017)은 확률 모델링으로 무작위 상황을 설명하는 것은 고전적, 빈도적, 주관적 접근 사이에 존재하는 논쟁을 극복하는데 기여한다고 주장하였다. 서론에서 빈도적 확률과 주관적 확률이 더욱 의미있게 다뤄지기 위해서는 등확률성을 만족하지 않는 상황을 예비교사에게 제시해야 한다고 주장하였는데 이는 Pfannkuch & Ziedins(2014)가 제시한 모델인 [그림 Ⅱ-3]에서도 드러난다.




[그림 Ⅱ-3] 확률의 추정에 대한 세 가지 상호연결된 방법

여기서 말하는 참 확률이란 정확한 값을 알 수는 없지만 확률적 상황에 대해서 존재하는 확률값이다. 모델링 확률은 확률적 상황을 이론적으로 모델을 통해 구한 확률이고, 경험적 확률은 시행 데이터를 통해 구한 확률인데 이 둘은 모두 참 확률을 추정하기 위한 확률이다. Pfannkuch & Ziedins(2014)가 제시한 세 확률의 관계로부터 확률값을 추정할 때 모델의 참 확률에 얼마나 비슷하게 추정하는지에 따라 "좋은 모델(good model)", "나쁜 모델(poor model)", "모델 없음(no model)"으로 분류하였다. 세 가지 모델에 대한 예로 '공정한 단일 주사위 던지기', '농구 선수의 자유투 성공률', '중환자실 병상 수 모델링'을 예로 들었는데, 이 세 가지 상황은 각각 고전적 확률, 빈도적 확률, 주관적 확률의 관점을 바탕으로 확률값을 추정하는 것과 비슷한 양상을 보인다.

참 확률을 구하기 위해서는 모델링 확률과 경험적 확률의 상호작용을 통해 구할 수 있다. 경험적 확률은 시행을 반복해보거나 통계 자료를 수집함으로써 참 확률을 추정하게 된다. 반면 모델링 확률은 이론적 모델을 세워서 참 확률을 추정해야 하는데 이론적 모델을 세우는 과정은 정혜윤, 이경화, 백도현, 정진호, & 임경석(2018)이 제시한 모델을 적용하고자 한다.




[그림 Ⅱ-4] 수학적 모델 도출 과정의 세분화(정혜윤 외, 2018)

[그림 Ⅱ-4]는 실생활과 관련 있는 문제에서 수학적 모델을 도출해내는 과정을 나타낸 것으로 다양한 수준과 환경의 학습자들에게 활용할 수 있는 모델이다(정혜윤 외, 2018). 문제에 영향을 미치는 요인을 찾는 것이 수학적 모델링 과정의 첫걸음이 되며, 그다음의 세 가지 전략은 상황에 따라 쓰이지 않을 수도 있다. 이 모델은 기초적인 형태로서 수준

과 상황에 따라 변형시켜서 활용할 수 있다. 이러한 범용성 덕분에 Pfannkuch & Ziedins(2014)의 모델링 확률을 구할 때 쓰이는 이론적 모델에서도 적용할 수 있다.

지금까지의 논의를 종합하면 확률교육에서는 확률에 대한 세 가지 관점이 종합적으로 다루어져야 하는데, 현재의 수학교육에서는 고전적 확률만 지나치게 강조하여 기계적인 확률 계산에 집중하는 경향이 있다. 고전적 확률만을 강조하여 가르치게 되면, 확률과 관련된 상황에서 하나의 이론적인 답만이 존재한다고 믿으며, 그 값의 실제적인 해석을 소홀히 하게 된다. 확률적 소양을 기르는 교육과정을 수행하기 위해서는 교사는 확률에 대한 세 가지 관점에 대한 이해와 적용 능력을 지니고 있어야 한다. 이에 예비교사에게 확률에 대한 예 구성 활동을 제시하여 현재 지니고 있는 확률에 대한 관점을 살펴볼 것이다.

상대적으로 소홀히 다루어지고 있는 빈도적 확률과 주관적 확률을 강조하기 위해서는 등확률성을 만족하지 않는 상황을 다루어서 확률값을 추정하는 경험이 필요하다. 고전적 확률을 적용할 수 없는 상황이기에 예비교사는 빈도적 확률과 주관적 확률을 어떻게 활용하는지 살펴볼 수 있다. 이는 Pfannkuch & Ziedins(2014)에서 모델링 확률을 통해 참 확률을 추정하는 상황이며, 정혜윤 외(2018)의 모델을 적용해 볼 수 있다.

마지막으로 확률이론에서 확률의 객관성을 높이는 여러 정리를 적용하는 모델링 상황을 제시하여 모델링 확률과 경험적 확률이 어떻게 상호작용하여 참 확률값을 추정해나가는지 살펴볼 수 있다.

# Ⅲ. 연구 방법

## 1. 연구 방법 및 절차

### 1.1. 연구 대상

본 연구는 서울특별시에 있는 대학교 수학교육과에 재학 중인 예비교사를 대상으로 수행되었다. 예비교사는 최근까지 학생으로서 확률교육을 받아온 교육의 수혜자이자 앞으로 확률교육을 실행할 교육의 제공자로서 본 연구에 가장 적합한 대상이다. 앞서 언급하였듯이 주관적 확률은 베이즈 정리에 의해서 사전 정보와 경험 정보가 결합하여 새로운 확률을 제시한다. 따라서 주관적 확률을 더욱 잘 이해하기 위해서는 베이즈 정리에 대한 학습이 선행되어야 한다. 이에 본 연구는 베이즈 정리를 학습하게 되는 과목인 2학년 확률론 전공 수업을 듣는 예비교사를 대상으로 연구를 진행한다. 또한 확률론 수업은 확률과 통계에 대한 대학교 수준의 이론을 처음으로 접하게 되는 과목으로서 확률과 통계에 대한 지식은 정규 교육과정의 고등학교 과정까지 학습한 경험에 기반한다.

연구 참여자는 확률론 수업에서 설문에 관련된 개념을 학습한 뒤 1차 설문을 진행하였다. 확률의 정의와 관련해서는 고전적 정의, 빈도적 정의, Kolmogorov가 제시한 확률의 공리적 정의에 대해서 학습하였지만, 주관적 확률에 관한 내용은 학습하지 않았다. 또한 앞서 언급한 베이즈 정리까지 학습하여 주관적 확률을 이해하기 위한 선행 조건은 갖추었다.

### 1.2. 조사 방법

서울특별시에 있는 대학교 수학교육과의 2학년 전공과목인 확률론 수업을 마친 뒤 IRB에 동의한 학생을 대상으로 1차 설문을 실시하였다.[3)]

1차 설문은 24명이 참여하였고 30여 분간 진행되었다. 이후 일주일 뒤에 2차 설문을 실시하였다. 1차 설문에 참여한 24명 중 22명이 2차 설문에 참여하였으며 무응답이 많은 2명의 데이터를 제외한 20명의 데이터를 분석하였다. 2차 설문은 1시간가량 소요되었다.

고전적 확률, 빈도적 확률, 주관적 확률이 생소한 예비교사가 많을 것으로 판단되어 2차 설문 시작 전에 확률에 대한 세 가지 관점인 고전적 확률, 빈도적 확률, 주관적 확률에 대하여 개념 설명을 실시하였다. 2차 설문은 2가지 모델링 상황과 설문지 활동을 통한 확률적 상황에 대한 예비교사의 접근 방식을 묻는 문항으로 구성되어 있다. 윷놀이와 코로나 감염 확률에 대한 모델링 상황은 단계별로 함께해야 하는 활동이 존재하므로 20분간 단계별로 간단한 설명을 제시하며 확률값을 추정하는 모델링 상황을 해결해 나가도록 도움을 주었다. 남은 시간은 각 모델링 상황에서의 교수학적인 질문과 확률에 대한 생각을 묻는 질문에 응답을 하였다.

### 1.3. 자료 분석

자료의 분석 과정은 자료의 코딩을 통해 이루어졌다. 설문 문항에 따라 코딩은 에틱 코딩(etic coding)과 에믹 코딩(emic coding)을 혼합하여 사용하였다. 연구자의 관심에 따라 분류하는 에틱 코딩의 방식은 예비교사의 확률에 대한 예시를 묻는 문항에서는 드러나는 교사의 확률에 대한 관점이 고전적, 빈도적, 주관적 확률인지 분류하는데 쓰였다. 확률의 뜻을 제시하는 문항과 모델링 상황에서 참 확률을 구하기 위한 단계별 문항마다 서술한 예비교사의 판단 근거에 대한 분석은 에믹 코딩의 방식으로 분류하였다. 에믹 코딩의 방식으로 범주화하였더라도 기존에 있는 이론과 부합하는 범주화 자료가 있으면 에틱 코딩과 결합하여 자료를 분석하였다.

---

3) 본 연구는 서울대학교 생명윤리위원회의 승인을 받고 진행하였다(승인 번호 IRB No. 2209/001-024).

## 2. 과제 설계

본 연구를 위한 설문지는 예비교사가 현재 지니고 있는 확률에 대한 관점을 확인하고, 등확률성을 만족하지 못하는 상황에서 확률값을 추정하는데 어떠한 사고과정을 거쳐서 판단하는지 살펴보기 위한 내용들로 구성하였다. 세부적으로, 예비교사가 가지고 있는 확률에 대한 관점을 살펴보기 위해 Jolfaee(2015)가 개발한 확률에 대한 예와 뜻을 구성하는 활동을 변형하여 문항을 제작하였다. 등확률성을 만족하지 못하는 상황에서 확률값을 추정하는 데 사용하는 접근 방식을 탐구하기 위해 2가지 모델링 상황인 윷놀이와 코로나 감염 확률을 제시하였는데, 코로나 감염 확률은 Batanero & Borovcnik(2016)가 제시한 활동에 질병관리청에서 제공하는 수치와 코로나 검사의 맥락을 결합하여 변형하였다. 두 모델링 상황에서의 참 확률을 구하기 위하여 여러 경험적 데이터를 제공한 뒤, 확률이론에서 확률의 객관성을 높이기 위한 여러 정리들과 결합하여 단계별로 참 확률에 더욱 가깝게 추정하는 사고 과정이 드러나도록 구성하였다. 전문가의 감수를 거쳐 문항의 신뢰도를 높였다. 설문지는 1차와 2차 설문으로 구성되어 있으며 [부록]에 수록하였다.

### 2.1. <1차 설문지>의 개발의도

<1차 설문지>는 연구 참여자가 지니고 있는 확률에 대한 현재의 관점을 확인하고, 등확률성을 만족하지 못하는 확률적 상황에 대한 접근 방식을 살펴보는 문항으로 구성되어 있다(부록 참조). <1차 설문지>의 문항별 개발 의도는 다음과 같다.

1-1번과 1-2번 문항의 과제는 Jolfaee(2015)가 사용한 질문을 변형하여 제시한 문항이다. Jolfaee(2015)는 확률에 대한 의미를 먼저 제시하고, 그에 해당하는 예를 1가지만 구성하도록 하는 문항이었다. 하지만 본 연구에서는 다양한 예를 구성하는 활동을 통해 연구 참여자가 가지고 있는 확률에 대한 예 공간[4]을 살펴보고, 이를 통해 확률에 대한

다양한 의미를 서술할 수 있도록 구성하였다. 시간의 제약이 있어 예를 구성하는 활동의 개수는 최대 5개로 제한하였고, 확률에 대한 다양한 의미를 제시하도록 확률의 뜻을 1개 이상 응답하도록 설계하였다.

2-1부터 2-5번 문항의 과제 역시 Jolfaee(2015)의 질문을 변형하였다. Jolfaee(2015)의 원래 문항에서는 확률이 0, 거의 0, 거의 1, 1인 상황에 대한 예시를 각각에 대해 한 가지씩 구성하도록 하였다. 하지만 본 설문에서는 문항 1-1에서 이미 많은 예를 구성하기도 하였고, 예비교사가 구성한 예는 고전적 확률의 예가 대부분일 것으로 예상되었다. 따라서 Jolfaee(2015)와 반대로 거의 일어나지 않거나 항상 일어나는 상황을 제시하고 연구 참여자가 확률값을 할당하는 문항으로 구성하였다. 확률값을 선택하도록 문항을 변형함으로 인해 예비교사 각자가 확률값이 고정된 수가 아니라 개인마다 다를 수 있음을 간접적으로 파악할 수 있을 것이라 예상하였다. 2-1번부터 2-3번 문항은 거의 일어나지 않는 상황을 제시하였고, 2-4번과 2-5번 문항은 항상 일어나는 상황을 제시하였다.

각 문항별로 살펴보면 2-1번 문항은 동전을 던지는 주체가 개인이 되면 확률을 다르게 판단하는지 묻는 문항이다. 2-2번 문항은 2-1번 문항 보다 더 일상적인 상황으로 제시하였다. 2-3번 문항은 드물게 일어나는 위험의 상황을 어떻게 판단하는지 묻는 문항이다. 2-4번 문항은 귀납 논증의 반례로 유명한 백조에 관한 문항이다. 2-5번 문항은 과학적 진리에 관한 판단을 확인하기 위한 문항이다.

3-1번과 3-2번은 등확률성을 만족하지 않는 상황 2가지를 제시한 문항이다. 3-1번은 등확률성을 만족하지 않는다고 쉽게 판단할 수 있는

---

4) Watson & Mason(2015)이 제시한 개념으로 하나의 예가 아니라 여러 예로 이루어진 장소의 의미를 내포한다. Watson & Mason(2015)은 예 공간이 수학적인 주제와 개념에 기인하여 만들어지며, 거꾸로 예 공간에 의하여 수학적인 주제와 개념이 형성되기도 한다고 보았다. 예 공간을 보면 수학적인 주제와 개념에 대하여 얼마나 이해하고 있는지 또는 이해할 수 있는지를 알 수 있으므로 본 논문에서는 예 공간을 탐색하여 예비교사들의 기존에 가지고 있는 확률에 대한 관점을 살펴보고자 한다.

상황인 윷을 던지는 상황이다. 3-2번은 코로나에 감염될 확률을 기존에 알고 있던 정보를 바탕으로 개인의 추측을 통해 확률값을 할당하라고 제시하였다. 이미 2-1번에서 2-5번까지의 문항에서 확률이 사람마다 다를 수 있음을 확인하였기에 개인의 판단을 확률값으로 할당하는 데 거부감이 적을 것으로 예상하였다. 3-1번과 3-2번 두 문항은 개인이 확률값을 추정할 때 사전 정보로 어떠한 정보를 활용하는지 확인하는 문항이다. 이 문항은 <2차 설문지>에서 경험적인 정보가 추가되는 문항과 연결되어 사전 정보와 경험적 정보가 결합하여 생기게 되는 새로운 확률을 예비교사가 어떻게 다루는지 살펴보게 된다.

## 2.2. <2차 설문지>의 개발의도

<2차 설문지>는 앞서 말했듯이 <1차 설문지>의 3-1번, 3-2번 문항과 연관된 2가지 모델링 활동이 제시된다(부록 참조). 또한 마지막에 활동을 다 마치고 확률에 대한 관점의 변화가 있는지 묻는 문항도 포함되어 있다. <2차 설문지>의 문항별 개발 의도는 다음과 같다.

활동에 들어가기에 앞서 상대적으로 생소한 주관적 확률을 설명하기 위해 고전적 확률, 빈도적 확률, 주관적 확률의 특징 및 한계에 대한 개념을 간단히 설명하고 진행하였다.

모델링 상황 1은 <1차 설문지>의 3-1번 문항인 윷놀이 상황과 관련이 있다. <2차 설문지>의 1-1번 문항은 주관적 사전 확률과 비교하여 실제로 윷을 10번과 50번 사이에 원하는 만큼 던지도록 하였다. 윷을 많이 던질수록 빈도적 확률이 점점 더 안정화되겠지만 예비교사가 빈도적 확률을 구하기에 충분하다고 판단하는 시행의 최소 횟수를 고민하도록 한 문항이다. 1-2번 문항은 모든 연구 참여자가 시행한 윷놀이 데이터가 빈도적 확률을 구하기에 적절한지 판단하는 문항이다. 빈도적 확률을 구하기 위해서는 von Mises가 주장하였듯이 같은 집산에 속해야 한다. 따라서 뒤집힌 윷의 개수, 던진 횟수, 확률값 등을 참고해 연구 참여자들이 생성한 데이터가 자신이 생성한 집산과 같은 '임의성'을 지니는

지를 판단하게 된다. 모든 연구 참여자가 던진 데이터를 수집한 엑셀 데이터를 보여주고, 이 중에서 같은 집산에 포함되기에 부적합한 연구 참여자의 데이터를 제외하도록 설계하였다. 1-3번 문항은 제외하고 남은 데이터들 중에서 사람마다 던지는 횟수를 일치시킬지 묻는 문항이다. 많이 던진 사람의 데이터와 적게 던진 사람의 데이터가 동등한 비중으로 다뤄야 하는지를 판단하는 문항이다.

1-4번은 엑셀에서 자동으로 계산되는 확률 4개를 적는 문항이다.
①은 <1차 설문지>의 3-1번 문항에서 응답한 주관적 확률,
②는 1-1번 문항에서 개인이 던진 데이터로만 구한 빈도적 확률,
③은 1-2번, 1-3번 문항에서 선별된 데이터를 통해 구한 빈도적 확률,
④은 모든 연구 참여자의 데이터를 선별 없이 구한 빈도적 확률이다.
데이터를 필터링하지 않고, 개인별로 던진 횟수를 일치시키지 않은 연구 참여자는 확률 ③과 ④가 일치하게 된다. 확률값이 일치하더라도 두 확률을 구분하기 위해 1-5번과 1-6번 문항을 응답할 때, 최소 하나 이상의 데이터를 제외한 상황을 가정하고 응답하도록 요구하였다. 1-5번은 2장에서 설명하였듯이 객관성으로 확률을 세부 분류할 수 있으므로 각각의 확률의 객관성을 순위대로 적는 문항이다. 1-6번은 각각의 확률이 윷이 뒤집힐 확률을 나타내기에 적합하다고 판단하는 순위를 적는 문항이다. 1-7번은 위의 활동에서 얻을 수 있는 교수학적인 의의를 적는 문항이다.

모델링 상황 2는 <1차 설문지>의 3-2번 문항인 코로나 감염 확률과 관련이 있다. 3-2번에서 응답한 주관적 확률을 코로나 감염에 대한 사전 확률로 제시하고, 민감도와 특이도의 개념을 이용해 조건부확률로서 다루도록 설계하였다. Batanero & Borovcnik(2016)가 제시한 유병률, 민감도, 특이도를 통해 감염 확률을 분할표에 채우는 예제 문항을 실제 코로나 데이터와 결합하여 제작한 문항이다. 여기에서 제시하는 민감도와 특이도는 질병관리청에서 제시하는 수치이다.

실제 상황을 모델링하여 연구 참여자가 코로나 검진을 위해 신속항원검사를 실시하였더니 양성이 나왔고, 정확한 감염 유무를 판단하기 위해

PCR 검사를 추가적으로 실시하였더니 음성이 나오는 상황을 제시하였다.

<2차 설문지>의 문항 2-1번부터 2-3번은 각각 감염될 확률, 민감도, 특이도를 이용하여 표를 단계별로 채우는 문항이다. 2-4번 문항은 분할표를 완성하고, 신속항원검사에서 음성이 나왔을 때 코로나 감염 확률을 기호로 나타내도록 하였다. 2-5번 문항은 추가로 실시한 PCR 검사의 민감도와 특이도를 이용해 분할표를 작성하는 문항이다. 2-6번 문항은 신속항원검사에서는 양성이 나오고 PCR 검사에서는 음성이 나왔을 때, 코로나 감염 확률을 묻고 기호로 표시하도록 요구하는 문항이다. 2-7번 문항은 사전 확률인 주관적인 코로나 감염 확률이 신속항원검사 결과 및 PCR 검사 결과와 같은 새로운 증거가 추가됐을 때, 각각이 확률이 어떻게 변하고 어떠한 의미를 가지는지 묻는 문항이며, 교수학적인 의의를 추가적으로 묻는 문항이다.

2-8번은 질병관리청에서 제공한 실제 코로나 감염자수를 제시하여 유병률을 구하는 문항이다. 유병률은 모집단을 정할 때마다 다른 수치를 제공하므로, 유병률의 모집단을 설정할 때 자신을 얼마나 잘 표현하는 집단을 선택하는지 확인하는 문항이다. Popper가 제시한 '가장 좁은 준거집합의 원칙'에 따라 더 많은 정보를 포함하는 집합이 상대적으로 더욱 객관적으로 보았으므로(송하석, 2006), 예비교사들이 코로나 감염 확률을 추정할 때 본인과 관련된 속성을 더 많이 포함한 유병률의 모집단을 활용하여 확률값을 결정하는지 확인하는 문항이다. 2-9번은 2-1번에서 사용했던 코로나 감염에 대한 주관적 확률과 2-8번에서 유병률을 통해 추정한 코로나 감염 확률을 비교하는 문항이다. 두 확률은 Hacking(2001)이 설명한 빈도 원리로 연결 지어질 수 있다. 코로나 감염 확률을 추정하는 주관적 확률과 유병률을 통해 구한 빈도형 확률을 비교해보는 문항이다.

3-1번은 기존에 주관적 확률에 대해서 접해본 경험을 묻는 문항이다. 이후 3-2번 문항을 묻기에 앞서 <1차 설문지>의 2-1번부터 3-2번 문항에 대한 연구 참여자의 응답을 정리한 표를 제공하고, 연구가 진

행되는 동안 응답했던 활동을 나열하였다. 이를 바탕으로 3-2번 문항에서 주관적 확률의 장단점을 적도록 하였다. 3-3번 문항은 확률에 대한 개념의 변화를 묻는 문항이다. 빈도적 확률에서의 주관적인 특성, 주관적 확률에서 객관성을 추가했던 활동이 확률에 대한 생각의 변화가 일어났는지 살피도록 의도하였다. 3-4번 문항은 연구에 참여하면서 수행했던 활동들이 확률을 가르칠 때 반영할지를 묻기 위한 문항이다. 이를 확인하기 위해 중학생 또는 고등학생에게 가르칠 의향이 있는 질문으로 설계하였다.

설문지의 각 문항에 해당하는 확률의 주관적(객관적) 요소를 정리하면 <표 Ⅲ-1>과 같다

<표 Ⅲ-1> 설문지 내 확률의 주관적 요소

| 차수 | 문항 | 조사내용 | 주관적 요소 |
|---|---|---|---|
| 1차 | 1-1 | 확률의 예 구성 활동 | |
| | 1-2 | 확률의 뜻 | |
| | 2-1, 2-2, 2-3, 2-4, 2-5 | 확률이 0과 1인 사건 | 주관적 확률의 할당 |
| | 3-1, 3-2 | 주관적 확률 할당 | 주관적 확률의 할당 |
| 2차 | 0-1, 0-2, 0,3 | 확률에 대한 세 가지 접근 방식 개념 설명 | 고전적, 빈도적, 주관적 확률의 주관적인 특성 |
| | 1-1, 1-2, 1-3 | 윷의 빈도적 확률 | 빈도적 확률의 주관적인 특성(집산 선택의 주관성) |
| | 1-5, 1-6 | 빈도적 확률의 객관성 및 적합성 | |
| | 1-7 | 윷놀이 활동에서의 교수학적인 의의 | |
| | 2-1, 2-2, 2-3, 2-4 | 신속항원검사와 코로나 감염 확률 | 사전정보와 경험정보의 결합(베이즈 정리) |
| | 2-5, 2-6 | PCR검사와 코로나 감염 확률 | |
| | 2-7 | 증거가 추가될 때의 확률의 변화 및 교수학적인 의의 | |
| | 2-8 | 코로나 유병률 | 빈도적 확률의 주관적인 특성(가장 좁은 준거집합의 원리) |
| | 2-9 | 주관적 확률과 빈도적 확률 | 주관적 확률과 빈도적 확률의 연결(빈도 원리) |
| | 3-1 | 주관적 확률의 경험유무 | |
| | 3-2 | 주관적 확률의 장단점 | |
| | 3-3 | 확률에 대한 인식 변화 | |
| | 3-4 | 주관적 확률과 교육과정 | |

## 3. 확률에 대한 관점과 접근 방식의 분류

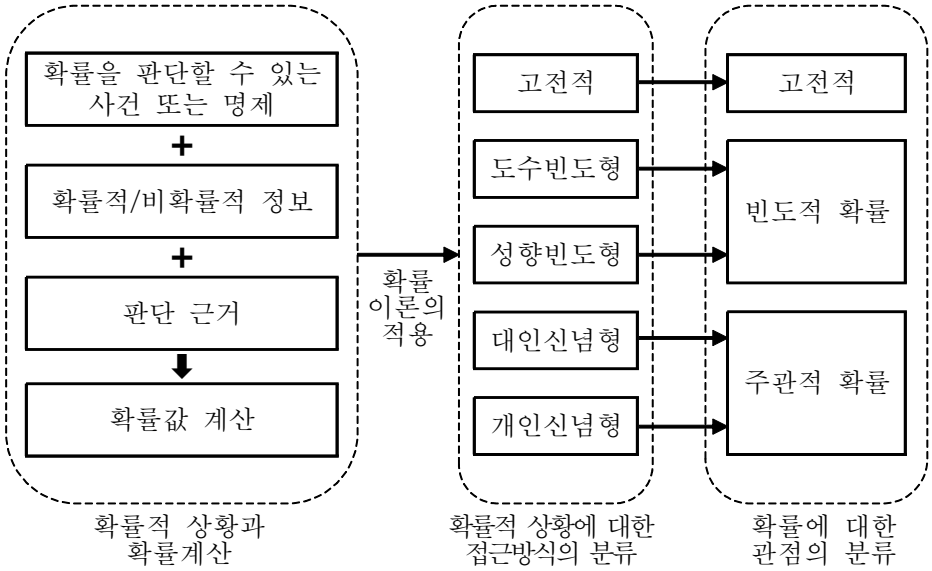
Chernoff(2008)는 수학교육에서 빈도적 확률과 주관적 확률이 어떨 때는 확률의 개념을 설명하는 이론으로 쓰이고, 어떨 때는 확률이 속하는 유형을 분류하는 역할로 쓰이는 등 두 가지 의미를 혼용해서 사용한다고 주장하였다. 이에 Chernoff(2008)는 수학교육에서의 빈도적 확률과 주관적 확률은 확률을 분류하는 용어로 계속 사용하고, 확률이론에서 쓰이는 개인신념형, 대인신념형, 성향빈도형, 도수빈도형 확률을 도입하여 수학교육에서 다뤄지는 확률의 개념을 더욱 확장하자고 주장하였다. 이에 본 논문에서도 확률이론에서 쓰이는 4가지 개념을 바탕으로 예비교사가 제시한 확률적 상황에 대한 확률값을 구하기 위해 어떠한 방식으로 접근하는지 분석하도록 한다.

확률적 상황에 대한 접근 방식을 살펴보기 위하여 Gal(2017)이 제시한 확률적 소양의 지식 요소 중 하나인 확률 계산의 과정에 대한 설명을 살펴보자. Gal(2017)은 학문적이지 않은 분야에서 확률을 계산할 때 비확률적인 정보까지 포함하는 다양한 정보가 활용되며, 보다 복잡한 판단 과정을 거쳐서 종합된 결과가 확률이라고 보았다. 또한 확률을 판단할 때 사용되는 근거를 살펴보아야 하며, 근거의 질에 대한 판단도 함께 이루어져야 한다고 보았다. 이에 본 논문에서는 주어진 확률적 상황에 대하여 확률값을 구하기 위해 사용되는 근거와 절차를 확률이론에 비추어 분류하는 과정을 확률에 대한 접근 방식이라고 부르도록 한다. 또한 확률적 상황에 대한 접근 방식에 대한 분류를 통해 예비교사가 확률에 대한 3가지 관점인 고전적, 빈도적, 주관적 확률을 지니고 있는지 판단하도록 한다.

하지만 이전 장에서도 설명하였듯이 고전적 확률이 주관적 확률과 빈도적 확률에 중복으로 해석되는 영역이 존재하므로, 이를 구분하는 기준이 필요하다. 이에 Hacking(2001)이 제시한 확률적 상황에 대한 판단을 살펴보자. Hacking(2001)은 1000개의 숫자가 적힌 공에서 1등을 뽑는 사건에 대하여, 확률이론에서의 4가지 서로 다른 접근 방식을 바탕

으로 확률을 계산하여도 확률값은 모두 $\frac{1}{1000}$로 동일하게 나온다고 설명한다. 이는 등확률성을 만족하는 상황에서는 확률을 신념형으로 접근하든 빈도형으로 접근하든 확률은 같은 값을 지니게 된다는 뜻이다. 결국, 확률값을 측정할 때 접근 방식에 따라 달라지는 경우는 등확률성을 만족하지 못하는 상황에서 발생한다고 볼 수 있다.

이에 본 논문에서는 표본공간을 명확히 제시할 수 있고, 각각의 사건이 등확률성을 보장하는 경우에만 고전적 확률로 분류하고, 등확률성을 보장하지 않는 경우에만 확률값을 계산하는 정보와 방식에 따라 개인신념형, 대인신념형, 성향빈도형, 도수빈도형으로 나누도록 한다. 이 4가지 확률이론에 대한 분류는 앞서 살펴본 객관성의 개념을 따르도록 한다. 판단에 대한 근거가 물질세계에 있으면 빈도형 확률, 사람의 마음속에 있으면 신념형 확률로 분류한다. 빈도형 확률에서의 이중분류는 구체적인 수치로 표현되는 객관성을 지니면 도수빈도형 확률로 분류하고, 물리적인 특성과 같이 상대적으로 주관적인 판단을 통해 확률을 계산한 경우에는 성향빈도형으로 분류한다. 신념형 확률에서의 이중분류는 근거가 여러 사람에게 받아들여지는 객관성을 지니고 있으면 대인신념형으로 분류하고, 근거가 개인의 마음에만 존재하여 사람마다 다를 경우는 개인신념형으로 분류한다. 확률을 계산하는 과정, 확률적 상황에 대한 접근 방식의 분류, 확률에 대한 관점의 분류에 대한 관계를 표현하면 [그림 Ⅲ-1]과 같다.




[그림 Ⅲ-1] 확률적 상황에 대한 접근 방식과 확률에 대한 관점

확률적 상황에 대한 예를 통해 위의 도식을 살펴보자. 태아가 남자일 확률이 “남자와 여자 두 가지 경우만 가능하므로 확률은 0.5이다.”라고 확률값을 계산하면 이는 고전적 확률의 접근 방식을 따른 것이고 고전적 확률에 대한 관점을 지녔다고 볼 수 있다. 반면, “보통 출생시의 자연적인 남녀 성비는 105:100이므로 확률은 0.51이다.”라고 확률값을 계산했으면 도수빈도형 접근 방식에 따른 것이고 빈도적 확률에 대한 관점을 지녔다고 볼 수 있다. 이렇듯 확률적 상황, 확률값, 판단 근거 등을 종합적으로 고려해서 확률에 대한 접근 방식을 판단하고 이를 통해 예비교사가 지니고 있는 확률에 대한 관점을 살펴보도록 한다.

# Ⅳ. 연구 결과

## 1. 예비교사가 기존에 가지고 있는 확률에 대한 관점

예비교사들이 기존에 가지고 있는 확률에 대한 관점을 파악하기 위해 <1차 설문지>의 1-1번, 1-2번 문항의 답변들을 분석하였다.

### 1.1. 예비교사가 구성한 확률의 예와 확률에 대한 관점 분석

문항 1-1번은 확률에 대한 예를 구성하는 문항이다. 하나의 예가 아니라 최대 5가지 예를 구성하라고 했으므로 Watson & Mason(2015)이 소개한 개념인 예 공간을 구성했다고 볼 수 있다. 또한 예를 구성할 때 최대한 다른 상황을 제시하라는 조건이 있었으므로 비슷한 예를 무의미하게 생성해 나가는 것이 아니라 연구 참여자가 기존에 가지고 있는 예 공간에 속해있는 예를 충분히 추출해냈을 것으로 볼 수 있다.

<1차 설문지>에 응답한 24명의 연구 참여자는 총 71개의 예를 구성하였고 참여자 별로 구성한 예의 개수는 <표 Ⅳ-1>과 같다.

<표 Ⅳ-1> 연구 참여자 별 예 구성의 개수

| 예 구성의 개수 | 인원(명) |
|---|---|
| 1 | 1 |
| 2 | 8 |
| 3 | 10 |
| 4 | 1 |
| 5 | 4 |
| 합계 | 24 |

71개의 예를 5가지 확률에 대한 접근 방식인 고전적, 성향빈도형, 도

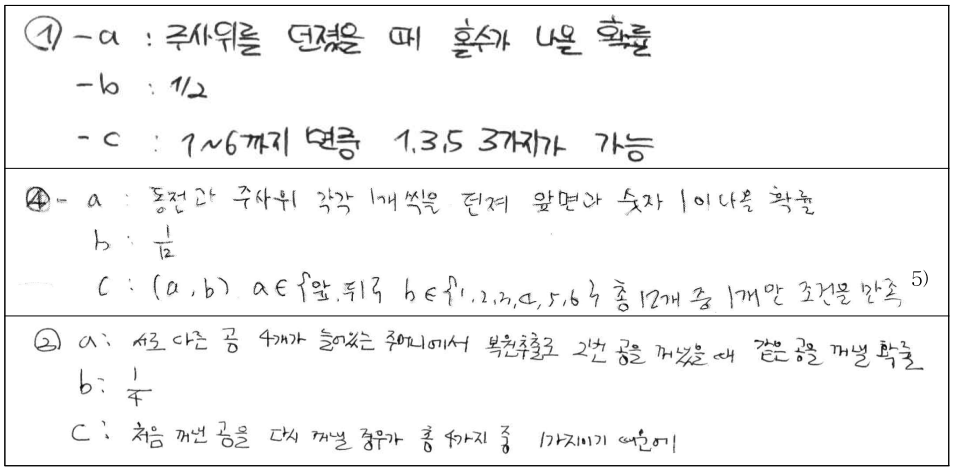
수빈도형, 개인신념형, 대인신념형으로 구분한 것의 개수는 <표 Ⅳ-2>와 같다.

<표 Ⅳ-2> 예 구성 활동에서 드러난 확률에 대한 접근 방식

| 접근 방식 | 인원(명) |
|---|---|
| 고전적 | 64 |
| 성향빈도형 | 0 |
| 도수빈도형 | 2 |
| 개인신념형 | 2 |
| 대인신념형 | 3 |
| 합계 | 71 |

### 1.1.1. 고전적 확률의 사례

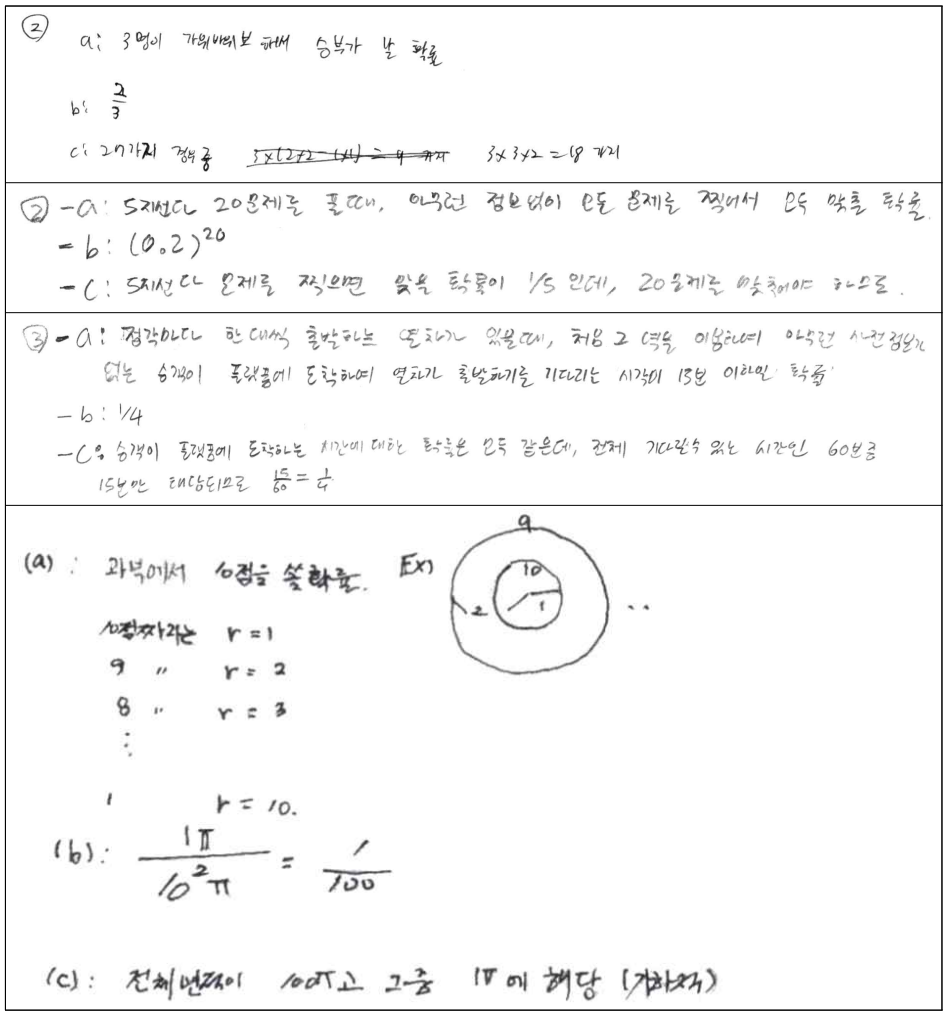
고전적 확률은 모두 등확률성을 만족하지만 이를 범주화하면 두 종류로 나눌 수 있다. 문항에서 예시로 들은 동전 던지기와 같이 수학 교과서에서 쓰이는 무작위 생성 활동인 주사위, 카드, 주머니 속 구슬 등의 소재가 포함된 예의 개수는 34개가 있다. 이러한 예비교사의 답변은 [그림 Ⅳ-1]과 같다.




[그림 Ⅳ-1] 고전적 확률 중 무작위 생성 활동을 이용한 예

5) 연필로 작성한 응답은 인식하기 쉽도록 그림과 같이 볼펜으로 덧칠하였다.

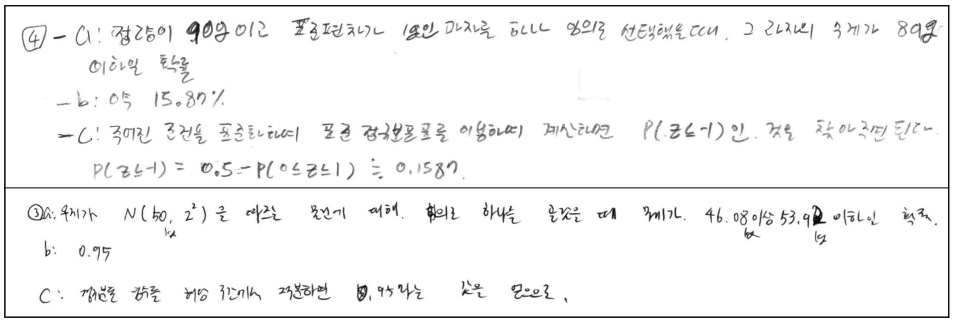
두 번째 유형은 일상적 상황을 묘사했지만, 가정에 등확률성을 적용하여 고전적 확률로 판단한 예들이 있다. 가위바위보, 줄 세우기와 같이 상황 자체에 등확률성이 내포한 예가 있는 반면, 오지선다형 문제와 같이 각 문항이 선택되는 확률이 같다는 것을 제시하여 등확률성을 만족하도록 제시한 상황 등이 있다. 또한 고전적 확률에 포함되는 기하학적 확률을 이용한 예 또한 여럿 있었다. 이러한 예비교사의 답변 30개가 있었고, 그중 일부는 [그림 Ⅳ-2]와 같다.



[그림 Ⅳ-2] 고전적 확률 중 일상적 상황의 예

### 1.1.2. 빈도적 확률의 사례

빈도적 확률은 확률에 대한 판단 근거가 물질세계에 존재해야 한다. 빈도적 확률과 관련된 예는 2가지 밖에 없었는데 두 가지 예 모두 표준정규분포를 이용하여 계산할 수 있는 상황을 제시하였다. 표준정규분포를 이용하는 상황이 되려면 뽑히는 상대도수가 높은 수여야 가능하고 이는 구체적인 데이터인 상대도수와 관련된 개념이므로 도수빈도형 확률로 접근하였다고 볼 수 있다. 이러한 예비교사의 답변은 [그림 Ⅳ-3]과 같다.

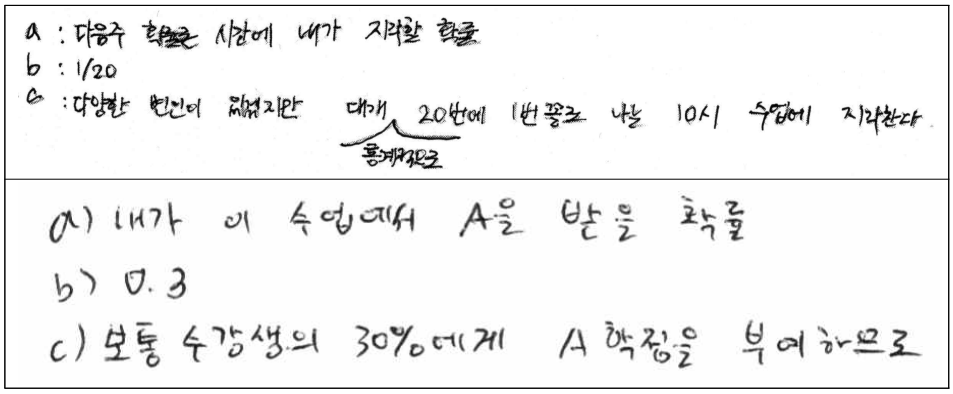

④ - a: 정량이 90g 이고 표준편차가 1g인 마카롱 하나 ㅎ의를 선택했을 때, 그 마카롱의 중제가 89g 이하일 확률
- b: 0.5 15.87%
- c: 주어진 조건을 표준화하여 표준 정규분포를 이용하여 계산하면 P(z≤-1)인 것을 찾아주면 된다.
P(z≤-1) = 0.5 - P(0≤z≤1) ≒ 0.1587.

③ a: 중제가 N(50, 2²)을 따르는 것에 이해. 하나를 뽑았을 때 중제가 46.08 이상 53.92 이하인 확률.
b: 0.95
c: 개념을 공부를 제대로 하였거나 적분하면 0.95라는 것을 알므로.

[그림 Ⅳ-3] 빈도적 확률의 예

### 1.1.3. 주관적 확률의 사례

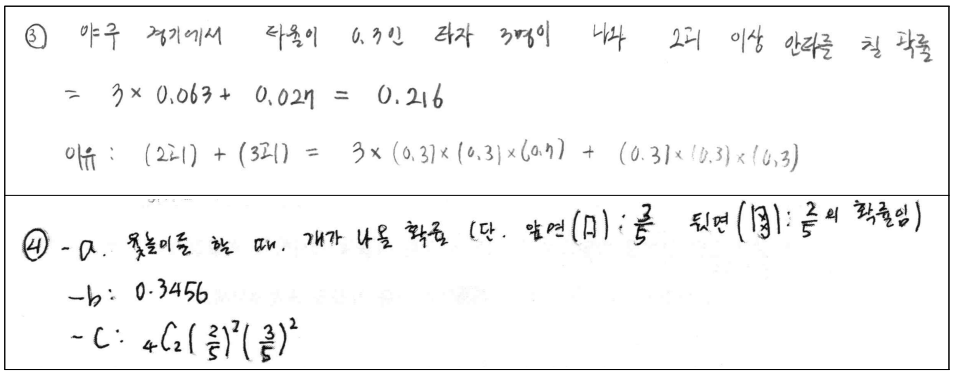
주관적 확률은 인간의 마음속에 있는 확률로서 확률에 대한 판단 근거가 대다수가 동의하는지에 따라 개인신념형과 대인신념형으로 세분화된다. 개인신념형 확률은 주어진 확률 상황에 대하여 응답자마다 다른 값을 제시할 수도 있는 성질을 지닌다. 이러한 예를 제시한 경우는 2개가 있었으며 그 예는 [그림 Ⅳ-4]와 같다.


a : 다음주 확률로 시장에 내가 지각할 확률
b : 1/20
c : 다양한 변인이 있겠지만 대개 20번에 1번꼴로 나는 10시 수업에 지각한다
통계적으로

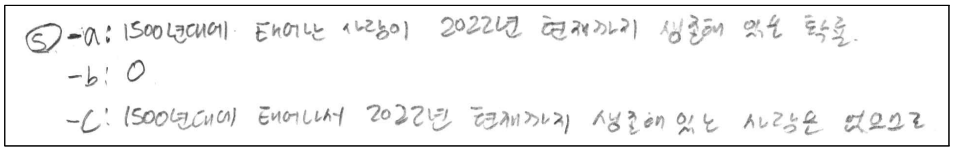
a) 내가 이 수업에서 A을 받을 확률
b) 0.3
c) 보통 수강생의 30%에게 A 학점을 부여하므로

[그림 Ⅳ-4] 개인신념형 확률의 예

대인신념형 확률은 판단에 대한 근거 여러 사람이 동의해야 한다. 이 유형에 속하는 예는 총 3개가 있다. 그중 2개는 등확률성을 만족하진 않지만 문제 상황에 직접 제시하여 마치 선험적으로 근거가 모든 사람에게 공유되는 상황이다. 이러한 이유에 따라 대인신념형 확률로 분류하였다. 남은 1가지 예는 “1500년대 태어난 사람이 2022년 현재까지 생존해 있을 확률”이다. 이 예는 현재까지 생존한 사람이 존재했다는 증거를 찾을 수 없으므로 근거가 존재한다고 볼 수 있다. 또한 Kaynes의 유명한 말인 “우리 모두는 결국에는 죽는다.”와 같은 진술을 Hacking(2001) 또한 신념형 확률로 분류하였다. 위에서 설명한 3가지 예는 [그림 Ⅳ-5]와 같다.


③ 야구 경기에서 타율이 0.3인 타자 3명이 나와 2리 이상 안타를 칠 확률
$= 3 \times 0.063 + 0.027 = 0.216$
이유 : (2리) + (3리) $= 3 \times (0.3) \times (0.3) \times (0.7) + (0.3) \times (0.3) \times (0.3)$

④ - a. 못놀이를 할 때, 개가 나올 확률 (단, 앞면(A) : $\frac{2}{5}$ 뒷면(B) : $\frac{2}{5}$의 확률임)
- b : 0.3456
- c : ${}_4C_2 \left(\frac{2}{5}\right)^2 \left(\frac{3}{5}\right)^2$


⑤ -a: 1500년대에 태어난 사람이 2022년 현재까지 생존해 있을 확률.
-b: 0
-c: 1500년대에 태어나서 2022년 현재까지 생존해 있는 사람은 없으므로

[그림 Ⅳ-5] 대인신념형 확률의 예

앞에서 살펴보았듯이 예비교사들이 제시한 71가지의 확률에 대한 예에서 고전적 확률이 차지하는 비중은 90.1%로 매우 높다. 2가지의 빈도적 확률의 예도 수학 교과서에서 자주 접하는 통계적 확률의 전형적인 예이다. 반면 주관적 확률은 5로 비중은 7%이다. 일상생활과 가장 관련 깊은 주관적 확률이 예비교사의 예 구성 활동에서 이렇게 낮은 비중을 차지하는 상황은 바람직하지 않아 보인다. 왜냐하면 2022 개정 교육과정이 지향하는 예측 불가능한 변화에 대한 대응능력을 키우기 위해서는 확률적 소양을 갖춰야 하는데, 실생활에서 다루는 상황들은 대부분 등확률성을 만족하지 않고 개인의 상황과 관련된 주관적 확률이기 때문이다.

또한 개인신념형 확률이 2개밖에 없다는 사실은 사람들마다 확률값이 다를 수 있는 상황의 예가 2개뿐이라는 의미이다. 바꿔 말하면 69개의 예에서는 확률적 상황에 대한 확률값이 모든 사람이 똑같은 값을 할당하게 된다는 뜻이다. 이러한 응답의 결과로 판단하건대 예비교사들이 확률을 설명할 때, 수학적인 대상으로서 모든 사람이 동일한 값을 가지는 상황만으로 한정하게 될 가능성이 높다고 볼 수 있다.

### 1.2. 예비교사가 형성하고 있는 확률의 의미

문항 1-2는 본인이 생각하는 확률의 뜻을 1개 이상 답변하도록 하는 문항이다. 예비교사들에게 형성된 확률에 대한 의미를 분석하기에 앞서 Batanero & Diaz(2007)가 제시한 확률의 의미와 관련된 5가지 요소 모델을 살펴보도록 한다. Batanero & Diaz(2007)는 확률교육에 대한 연구 프로그램을 개발하면서 확률의 다양한 의미를 분석하기 위하여 5가지 상호 관련된 요소를 추출하였다. Batanero & Diaz(2007)는 확률

에 대한 균형 잡힌 이해를 갖추기 위해서는 다음의 5가지 요소가 포함되어야 한다고 제안한다.

1) 확률의 개념이 발생하면서 생겨난 문제 : Galileo가 1620년에 제시한 세 개의 주사위 눈의 합에 관한 문제와 같이 확률의 초기 개념 형성을 위해 사용됐던 우연 게임과 관련된 여러 문제들을 이해해야 한다.
2) 확률의 표현 : 확률은 추상적인 개념이므로 퍼센트, 분수, 표, 밀도 함수, 대수적 표현 등과 같이 명시적으로 확률을 표현해야 한다.
3) 확률과 관련된 문제 및 확률값을 계산하기 위해 사용되는 알고리즘 또는 풀이과정 : 조합론, 해석, 대수, 기하와 같이 확률 문제를 해결하는데 유용한 수학적 도구를 가지고 있으며, 계산기와 컴퓨터 또한 기술적으로 확률 문제를 푸는 데 도움을 준다.
4) 확률의 정의 : 확률의 다양한 정의, 확률의 기본 공리, 기댓값, 상대도수, 우도, 중심극한정리와 같은 확률의 개념과 성질이 모두 포함된다.
5) 확률의 성질 또는 문제를 해결하는 과정에서 논증의 타당성을 입증하기 위해 사용되는 증명과 논거 : Galileo의 초기 확률 문제를 정당화할 때 컴퓨터 시뮬레이션으로 설명하거나, 표본을 나열하면서 설명할 수 있듯이 이러한 논증에서 사용되는 사고방식을 말한다(Batanero & Diaz, 2007).

위에서 제시한 5가지 요소가 복합적으로 작용하면서 확률에 대한 균형 잡힌 이해를 돕는다. Batanero & Diaz(2007)는 위의 5개의 각각의 요소가 개인과 교육기관에 따라 의미가 달라진다는 점에 주목하였다. 예를 들어 초등학생에게는 확률에 대한 직관적인 개념과 간단한 계산만 다뤄도 되지만, 확률적 소양을 갖춘 시민에게는 주식 시장, 의료 진단과 같은 의사결정에서 확률이 사용되는 과정을 이해시키는 과정이 필요하다. 과학 분야나 대학 수준에서는 확률분포, 중심극한정리와 같은 고등 개념이 필요할 수도 있다. 이렇듯 동일한 요소에 속하는 개념이라도 개인과 교육기관에 따라 달라지는 점에 착안하여 Batanero & Diaz(2007)는 확률의 의미를 제도적(institutional) 의미와 개인적(personal) 의미로 구별지었다. 제도적 의미는 교육기관이나 교과서에서

제공한 의미이고, 개인적 의미는 특정 개인이 직관에 의해 형성된 의미이다. 제도적 의미와 개인적 의미를 구분할 때 모호한 점이 존재하지만, 제도적 의미를 다루는 교육기관에서는 확률을 계산하는 과정 및 확률의 성질을 다루는 데 중점을 둔다. 따라서 확률을 구하는 방식을 구체적으로 제시하거나, 확률의 공리, 조건부확률과 같이 확률의 성질을 다루는 데 필요한 요소가 포함되면 제도적 의미로 해석한다. 또한 제도적 의미는 교과서에서 다뤄지기 때문에 수학적 확률, 통계적 확률과 같이 교과서에서 개념을 설명했던 용어와 예비교사가 응답한 의미를 서로 연관 지을 수 있다. 이러한 요소가 없으면 개인의 직관에 의해서 받아들여진 개념으로 판단해 개인적 의미로 분류한다.

문항 1-2에 대하여, 24명의 예비교사가 36가지의 확률의 의미에 대한 설명을 제시하였다. 4개의 응답을 한 1명을 제외하면 23명은 모두 1개 또는 2개의 응답을 제시하였다. 36개의 응답 중 제도적 의미는 21개, 개인적 의미는 15개로 분류된다.

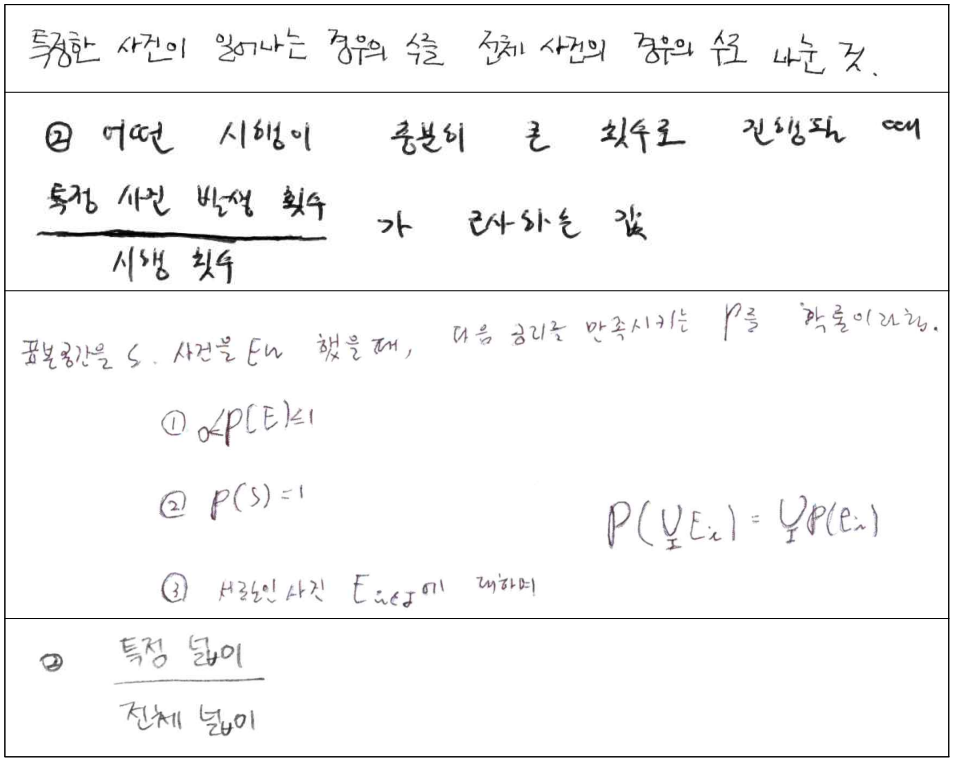
### 1.2.1. 제도적 의미

앞서 언급하였듯이 제도적 의미는 교과서에서 서술된 내용에 기반하므로 확률의 의미에 대한 용어가 존재한다. 따라서 제도적 의미에 속하는 확률의 의미를 교과서의 확률에 대한 용어와 대응시켜 분류하고자 한다. 추가적으로 제도적 의미로 분류하는데 고려한 사항으로서 문항 1-2에서 확률의 정의가 아닌 확률의 뜻을 서술하라고 요구하였기 때문에 정의가 갖춰야 할 엄밀성이 부족하더라도 의미가 통하면 제도적 의미에 해당된다고 판단하였다. 예를 들면 “특정 사건의 측도(전체 측도가 1)”라고 응답한 예가 있었는데 아직 확률의 엄밀한 정의를 학습하지 않은 상태지만 전사건의 확률이 1인 성질과 확률을 구하는 방식에 대해서 간략하게나마 서술하였으므로 제도적인 의미로 분류하였다. 따라서 제도적 의미를 분류하면 <표 Ⅳ-3>과 같다.

<표 Ⅳ-3> 확률의 제도적 의미에 따른 분류

| 용어 | 개수 |
|---|---|
| 수학적 확률 | 15 |
| 통계적 확률 | 3 |
| 공리적 확률 | 2 |
| 기하학적 확률 | 1 |
| 합계 | 21 |

24명의 예비교사 중에 15명이 수학적 확률을 제시하였으므로 많은 예비교사가 지니고 있는 확률의 의미에서 수학적 확률의 개념을 포함한다고 볼 수 있다. 반면 수학적 확률을 제외한 확률에 대한 제도적 의미를 제시한 응답은 6개로 적게 나타났다. <표 Ⅳ-3>에서 제시한 4가지 용어에 대한 대표적인 예 1가지씩을 나타내면 [그림 Ⅳ-6]과 같다.




[그림 Ⅳ-6] 제도적 의미에서의 확률에 대한 각각의 예

[그림 Ⅳ-6]에서 3번째 예는 확률의 공리적 정의를 정확히 서술하였는데 그 원인을 살펴보면 설문을 실시하기 일주일 전에 관련 개념을 학습하였기 때문으로 보인다. 이는 최근에 어떤 수학적인 경험을 했는가에 따라 예 공간과 그 구조를 다르게 구성한다고 주장한 Watson & Mason(2015)의 설명에 부합하는 사례로 보인다.

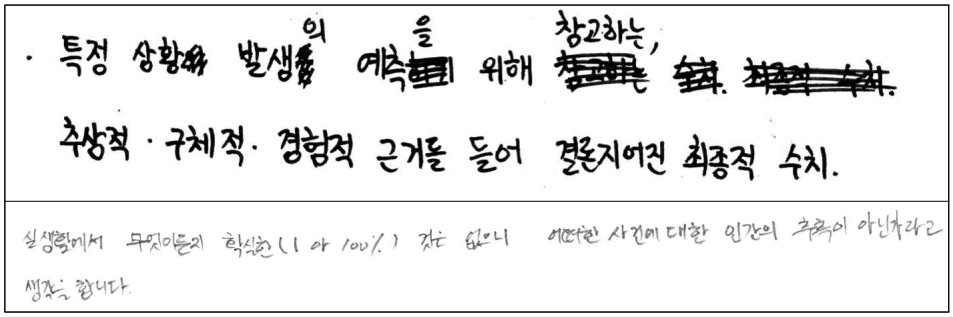
### 1.2.2. 개인적 의미

Gal(2017)의 제시한 확률적 소양의 구성 요소 중 가능성에 대해 의사소통하는데 필요한 용어와 전략인 언어 요소를 2가지로 세분화하여 설명한다. 하나는 추상적 구조와 관련되는 용어와 문구에 대한 이해를, 다른 하나는 실제로 일어날 수 있는 사건의 가능성을 표현하거나 설명하는 다양한 방법에 대한 이해이다(Gal, 2017). 예비교사의 확률의 뜻에 대한 응답에서도 언어에 대한 두 가지 세부 요소로 표현되었다. 개인적 의미에 속하는 응답의 대부분은 확률은 수이기 때문에 이 값을 어떻게 구하는 지 방법에 대한 설명하였고, 이는 실제로 일어날 수 있는 사건의 가능성을 표현하거나 설명하는 다양한 방법에 속한다. 또 이 수가 지니는 의미에 대한 설명은 추상적 구조와 관련된 문구에 해당한다. 예를 들어 “어떤 사건이 일어날 가능성의 비율”을 살펴보면 확률에 대한 표현 방식은 비율로 나타낼 수 있고, 그 확률이 지니는 추상적 구조는 사건이 일어날 가능성을 설명해주는 수치라고 볼 수 있다. 이 두 가지 설명 요소를 기준으로 키워드를 추출하였고 이를 분류하면 <표 Ⅳ-4>와 같다.

<표 Ⅳ-4> 확률의 개인적 의미를 서술하는 데 등장하는 키워드[6]

| 요소 | 키워드 | 개수 |
|---|---|---|
| 추상적 구조 | 가능성 | 7 |
| | 기대 | 4 |
| | 근거 | 2 |
| | 추측 | 1 |
| | 예측 | 1 |
| 표현 방법 | 비율 | 2 |
| | 정도 | 2 |
| | 확률 | 1 |

<표 Ⅳ-4>에서 마지막에 등장한 확률이 키워드로 포함된 한 가지 예를 살펴보면, “통계적이든 수학적이든지 간에 상관없이 어떠한 사건이 일어날 확률이라고 생각합니다.”로 표현한 것으로 보아 가능성을 나타내는 수치를 확률로 서술한 것으로 보이지만 키워드대로 확률로 셈을 하였다. 확률의 추상적 구조를 서술하는 키워드는 가능성, 기대, 근거, 추측, 예측으로 볼 수 있고, 확률에 대한 수치를 표현하는 방법을 설명한 키워드는 비율, 정도, 확률이라고 볼 수 있다. 그중에서 추측 및 예측과 관련된 키워드가 포함된 예 2가지를 살펴보도록 한다. 예비교사가 작성한 내용을 발췌하면 [그림 Ⅳ-7]과 같다.




[그림 Ⅳ-7] 개인적 의미에서의 확률에 대한 각각의 예

6) 하나의 예에서 두 개의 키워드가 등장하면 따로 셈을 하였다.

위의 두 예는 모두 주관적 확률의 성질을 잘 묘사하는 예로 볼 수 있다. 이 두 예비교사는 일상적으로 사용하는 확률이 확률적 판단의 근거로 사용될 수 있다는 것을 인식하고 있다고 판단된다. 한편, 위 두 예를 제도적 확률로 분류하지 않은 이유는 주관적 확률을 교육기관에서 배운 경험이 없으므로 두 예비교사가 본인의 직관에 의해 구성한 의미로 보아 개인적 확률로 분류하였다.

문항 1-2에서 응답한 예비교사 24명 중 15명이 확률의 의미에 대하여 수학적 확률에 해당하는 설명을 제시하였다. 기하학적 확률도 수학적 확률의 특수한 사례로 볼 수 있으므로 수학적 확률을 제외한 제도적 정의는 5개밖에 드러나지 않았다. 이는 학교수학에서 확률이 고전적 관점에서만 주로 다뤄져 왔다는 기존의 비판과 부합하며, 빈도적 확률에 대한 의미를 더욱 강조해야 할 필요성을 말해 준다.

확률의 개인적 의미를 살펴보면 확률은 대부분 사건이 일어나는 가능성을 나타내는 수치라고 인식한다고 볼 수 있다. 그중에서도 주관적 확률과 관련된 의미를 제시한 예비교사는 2명으로, 앞선 절에서 개인신념형 확률에 대한 예가 1개인 것과 마찬가지로 예비교사들은 전반적으로 주관적 확률에 대한 인식이 부족하다고 판단할 수 있다.

## 2. 등확률성을 만족하지 않는 상황에 대한 확률 할당

이 절에서는 등확률성을 만족하지 않는 다양한 상황에 대하여 예비교사들이 확률을 할당하는 데 어떤 접근 방식을 사용하는지 분석하고자 한다. 분석에 사용된 문항은 <1차 설문지>의 2번과 3번 문항이다.

### 2.1. 절대로 일어날 수 없는, 반드시 일어나는 사건의 확률

<1차 설문지>의 문항 2번에서는 거의 일어나지 않는 사건 또는 거의 일어나는 사건의 확률을 다룬다. 확률이 0 또는 1인 사건의 중요성을 반영하여 중

학교 2학년 수학 교과서에서는 절대로 일어날 수 없는 사건의 확률과 반드시 일어나는 사건의 확률로 다뤄지고 있다. 중학교 2학년 과정에서는 확률이 고전적 확률로 다뤄지기 때문에 표본공간에 사건이 속해있는지에 따라 각각의 확률을 0과 1로 할당한다. 확률이 0 또는 1인 사건을 고전적 확률로 다루게 되면, 확률에 관한 문제라기보다는 집합의 포함 관계를 확인하는 문제에 더 가깝게 되어 확률값을 형식적으로만 다루게 되는 위험이 있다. 이에 문항 2-1에서 문항 2-5까지 다섯 개의 문항은 확률에 대한 여러 접근 방식을 적용해 볼 수 있는 상황 중 확률이 거의 0인 상황과 확률이 거의 1인 상황을 제시하였다. 각 문항에 대한 예비교사의 응답에 대한 분포는 <표 Ⅳ-5>와 같다.

<표 Ⅳ-5> <1차 설문지> 문항 2의 응답에 대한 분포

| 문항 | 확률 상황 | 0 | 거의 0 | 거의 1 | 1 |
|---|---|---|---|---|---|
| 2-1 | 내가 보통의 백 원짜리 동전을 100번 던졌을 때, 100번 모두 동전의 앞면이 나올 확률 | 0명 (0%) | 24명 (100%) | | |
| 2-2 | 내가 로또 1등에 당첨될 확률 | 5명 (20.8%) | 19명 (79.2%) | | |
| 2-3 | 내가 길을 가다가 벼락에 맞을 확률 | 1명 (4.2%) | 23명 (95.8%) | | |
| 2-4 | 백조가 흰 색일 확률 | | | 18명 (75%) | 6명 (25%) |
| 2-5 | 내일 아침에 해가 뜰 확률 | | | 11명 (45.8%) | 13명 (54.2%) |

## 2.2. 거의 일어나지 않는 상황에 대한 확률

문항 2-1번은 동전을 던지는 예로 등확률성을 만족하는 확률적 상황이다. 등확률성을 내포하고 있지만 “내가 보통의 백 원짜리 동전을 100번 던졌을 때, 100번 모두 동전의 앞면이 나올 확률”과 같이 제시하였다. 행위의 주체가 개인임을 확인할 수 있는 주어를 포함하여 예비교사가 일상 경험에서 동전을 100번 던지는 상황을 실행하지 않을 것으로 판단하면 확률을 0으로 응답할 수도 있다. 또한 뒤에 나오는 2-2번 문항과 비교하는 대조군의 역할도 수행한다.

응답 결과를 분석하면 24명 모두 확률이 거의 0이라 응답하였다. 앞면이 나올 확률을 잘못 판단하여 $\left(\frac{1}{6}\right)^{100}$로 응답한 1명을 제외하면, 모두 확률값을 $\left(\frac{1}{2}\right)^{100}$, $0.5^{100}$, 분모가 $2^{100}$ 등과 같이 정확히 계산한 수치를 제시하였다. 이렇게 구한 확률값은 0이 아닌 양수이므로 확률이 거의 0이라 판단한다고 응답하였다. 모든 예비교사가 동전을 던지는 상황을 고전적 확률로 접근하여 응답한 이유를 분석해보면, 그중 하나는 동전 던지기 상황이 Watson & Mason(2015)이 설명한 '관례적인 예'이기 때문으로 볼 수 있다. 동전 던지기는 확률 개념이 등장했을 때부터 함께 다뤄져 왔고, 교과서에서 확률을 다룰 때 빈번히 다뤄져 왔다. 이에 개개인은 확률을 계산하는 데 매우 익숙한 상황이므로 확률에 대한 깊은 판단 없이 기계적으로 확률을 계산한 결과로 보인다. 또 다른 이유로는 "내가 보통의 백 원짜리 동전을 100번 던졌을 때,"는 조건부확률의 조건절로 해석할 수 있다. 이렇게 해석하면 이미 내가 동전을 100번 던진 상황을 전제로 하기 때문에 행위의 주체인 '내가' 동전을 던질지 말지 결정하는 내용은 고려되지 못한다. 이러한 요소를 고려하여 문항 2-1를 "내가 던진 보통의 동전이 100번 모두 앞면이 나올 확률"과 같이 변형하면 관습적으로 주어지던 조건부확률의 표현과 달라져 다른 응답이 나올 수도 있을 것이다.

문항 2-2는 "내가 로또 1등에 당첨될 확률"을 묻는 문항이다. 로또의 당첨 원리가 문항에 설명되어 있지 않지만, 국내 로또 당첨 방식은 45개의 공 중에서 6개의 공의 숫자가 일치하면 1등에 당첨된다. 이 상황은 공 45개가 선택될 확률이 모두 동등하므로 등확률성을 만족한다고 볼 수 있다. 등확률성을 만족한다는 면에서는 문항 2-1과 비슷하지만, 일상생활과 더욱 밀접한 상황을 제시한다.

예비교사의 응답을 살펴보면 확률이 0이라고 응답한 예비교사는 총 5명으로 모두 로또를 구입하지 않기 때문에 사건이 발생하지 않는다고 응답하였다. 이는 로또에 당첨될 확률이 존재하지만, 개인의 행동 및 신념과 관련되어 0으로 응답하였으므로 주관적 확률 중에서 개인신념형 확률로 분류할 수 있다. 확률이 거의 0이라고 응답한 예비교사는 19명이다. 그중 고전적 확률로 분류된 예비교사의 응답은 10개가 해당되는데, 예비교사가 제시한 이유는 확률이 분자가 1, 분

모가 ${}_{45}C_6$ 또는 800만과 같이 계산가능한 값이기 때문에 확률값이 0이 아니라고 설명하였다. 예비교사 중 1명은 “로또 1등에 당첨된 사람들이 있으니깐”을 이유로 확률이 거의 0이라고 응답하였다. 이는 로또 1등에 당첨된 수는 상대도수를 구할 때, 정확한 분모의 수치는 모르지만 분자의 빈도에 대한 정보를 이용하여 판단하였으므로 도수빈도형 확률에 속한다. 확률값을 계산하지 않고 “거의 작지만 0은 아니다.”라고 응답한 예비교사는 8명으로 조합을 이용한 구체적인 계산이 아닌 막연히 가능하다고 판단하였으므로 개인신념형 확률로 분류한다.

문항 2-3은 벼락에 맞을 확률을 구하는 문항이다. 확률이 0이라 응답한 예비교사는 1명으로 “연속확률 변수에서, 예를 들어 $P(Z=k)=0$ $(k \in R)$인 것과 같은 이유로 생각한다.”와 같이 연속확률변수에서 특정한 값이 가지는 확률이 0인 성질을 사용하여 응답하였다. 1절에서의 확률에 대한 예에서도 살펴보았듯이 표준정규분포를 통한 확률의 판단은 도수빈도형에 속한다고 볼 수 있다. 다른 23명의 예비교사는 모두 확률이 거의 0으로 응답하였다. 빈도적 확률에 속하는 성향빈도형 확률과 도수빈도형 확률에 속하는 응답은 각각 2명과 8명이다. 주관적 확률에 속하는 응답은 개인신념형 응답만 존재하였으며 13명이 해당된다. 성향빈도형 확률은 번개가 치는 물리적 상황인 면적, 저항과 같은 근거로 판단을 내렸고 도수빈도형 확률은 벼락맞은 사례를 제시하여 확률에 대한 판단을 하였다. 13명 중 11명은 특별한 근거 없이 가능한 사건이라고 응답한 반면, 2명은 개인이 번개가 치는 공간에 존재할 확률도 고려해야 한다고 봤다. 앞선 2장에서 설명하였듯이 개인신념형 확률에서 아무런 근거 없이 주관적으로 확률을 할당하는 경우가 있고, 개인과 관련된 여러 요소를 고려하여 확률을 할당하는 경우가 있음을 확인하였다.

## 2.3. 거의 대부분 일어나는 상황에 대한 확률

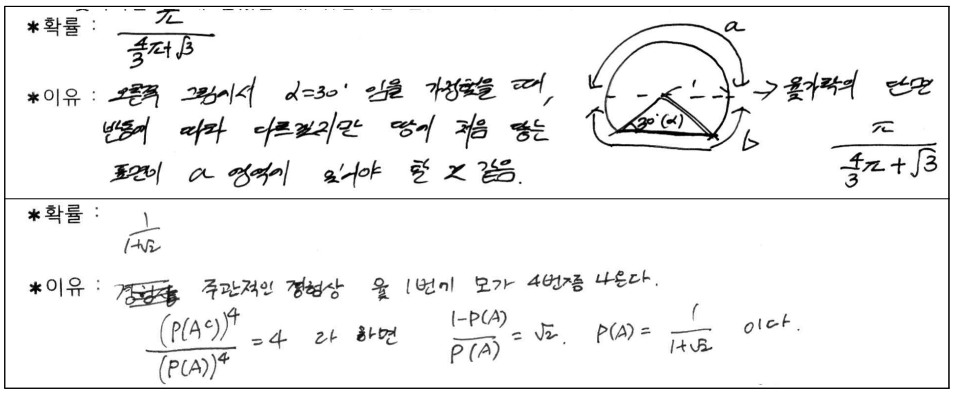
문항 2-4는 “백조가 흰색일 확률”을 묻는 문항이다. 확률이 1이라고 응답한 예비교사 6명의 응답을 분석해보면, 모두 백조의 정의에서 흰색임을 내포하고 있다고 보았다. 등확률성을 만족하는 상황이 아니고, 백조와 흰색의 정의 및 두 개념의 관계에 대한 판단만으로 확률을 할당였기 때문에, 판단의 근거는 인

간의 마음속에 있고 이는 신념형 확률에 속한다. 또한 개념의 내포관계를 이용하여 판단을 내리는 개개인은 모두 같은 판단을 내릴 것이므로 대인신념형 확률에 속한다. 확률이 거의 1이라고 응답한 18명의 예비교사의 응답을 분석해보면 모두 돌연변이의 존재를 언급하고 있다. 이 중 16명은 돌연변이가 생길 수도 있다는 막연한 가능성만 언급하므로 개인신념형 확률이고, 다른 2명은 "검은 백조"라는 반례의 존재를 언급하므로 도수빈도형 확률로 분석할 수 있다.

문항 2-5는 "내일 아침에 해가 뜰 확률"을 묻는 문항이다. 확률이 1이라고 응답한 13명의 예비교사의 응답을 분석해보면 2개의 도수빈도형, 2개의 대인신념형, 9개의 개인신념형으로 분류할 수 있다. 도수빈도형은 과거의 관측에서 해가 뜨지 않은 적이 없음을 근거로 들었고, 대인신념형은 아침의 용어가 해가 뜬다는 의미가 포함되어 있음을 근거로 제시하였다. 개인신념형은 지구가 멸망할 가능성이 0으로 제시하거나, 해가 뜬다는 과학적 진리는 항상 참이므로 확률이 1이라 응답하였다. 확률이 거의 1이라고 응답한 11명의 예비교사는 모두 개인신념형 확률로 문항 2-5를 판단하였다. 판단의 근거로는 해를 뜨는 현상에 대한 관측의 오류 가능성 및 천문학적 현상에 의해 해가 뜨지 않을 가능성이 존재하므로 확률이 거의 1이라 응답하였다. 이러한 가능성에 대한 판단 근거는 개인의 믿음에 달려있으므로 개인신념형 확률로 분류하였다.

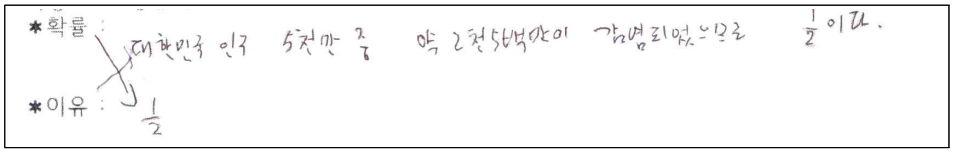
## 2.4. 윷의 확률과 코로나 감염 확률에 대한 확률 할당

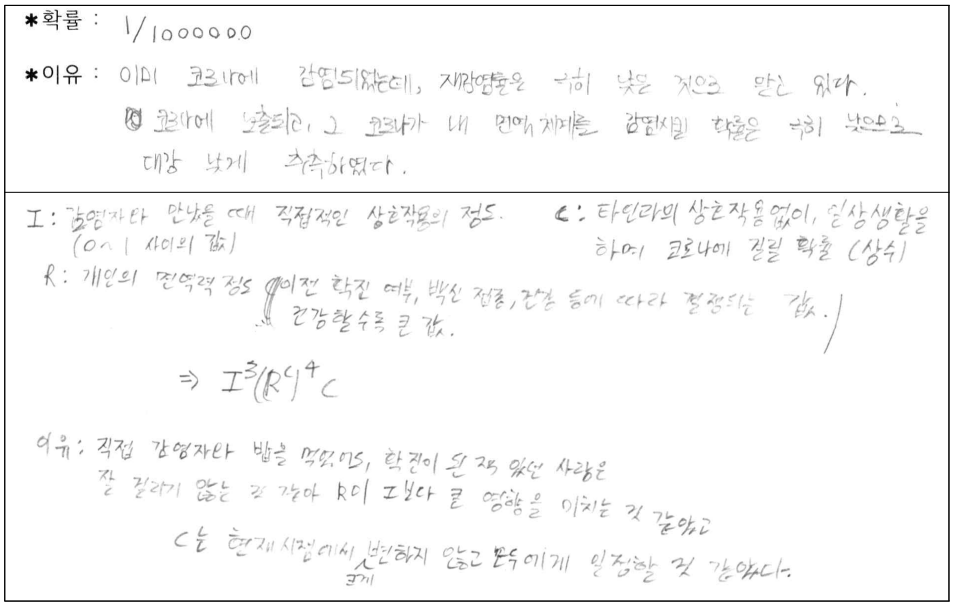
문항 3-1은 윷을 던졌을 때 뒤집힌 면이 나올 확률을 추측해보는 상황이다. 윷이 뒤집힐 확률에 대한 분포는 확률을 정확한 수로 제시한 응답이 19개이며, 5개의 응답은 확률을 수로 제시하기보다는 "0.5보다 크다." 혹은 "0.5보다 작다."와 같은 상대적 비교, 추측된 수식, "알 수 없다" 등의 응답이 있었다. 확률에 대한 접근 방식으로 분류해보면 성향빈도형은 18개, 개인신념형은 6개로 분류되었다. 성향빈도형은 물리적인 성질인 길이, 면적, 무게중심, 낙하상황, 던지는 방식 등의 근거를 들며 확률을 할당하였다. 개인신념형은 과거 윷놀이를 해온 기억에 의존하여 확률을 할당하였다. 대표적인 예를 나타내면 [그림 Ⅳ-8]과 같다.




[그림 Ⅳ-8] 윷의 확률에 대한 접근 방식의 예

문항 3-2는 본인이 코로나에 감염될 확률을 추측해보는 상황이다. 코로나 감염 확률을 정확한 확률값으로 제시한 예비교사는 17명이다. 반면, 감염확률에 대한 모델링 상황을 확률값 없이 수식의 형태로만 표현하거나, “거의 0”, “거의 1”, “알 수 없다” 등으로 응답한 예비교사는 7명이다. 다른 확률적 상황과는 다르게 확률을 판단할 때 다양한 요소가 작용하기 때문에 “확률값을 측정하기 어렵다.”라고 응답하며 확률값을 제시하지 않은 예비교사가 3명이 존재하였다. 확률값을 대략적으로라도 추정하지 않은 3명을 제외하고 확률에 대한 접근 방식을 분류해보면 1명만 도수빈도형이고 나머지 20명은 개인신념형이다. 도수빈도형은 평소 매체에서 접한 정보인 누적확진자수가 2500만 명을 근거로 확률을 0.5라 응답하였다. 개인신념형 확률을 분류하면 개인의 감염여부에 기반하여 판단한 예비교사는 10명이고, 코로나 감염 상황에 미치는 요소를 개인적인 판단에 의해 모델링하여 계산한 예비교사는 10명이다. 도수빈도형, 개인경험에 의한 개인신념형, 모델링 상황을 나타낸 개인신념형에 대한 각각의 대표적인 예를 나타내면 [그림 Ⅳ-9]과 같다.





[그림 Ⅳ-9] 코로나 감염 확률에 대한 접근 방식의 예

## 2.5. 확률적 상황에서 드러난 확률에 대한 접근 방식

앞서 설명한 내용을 바탕으로 문항 2-1부터 3-2에서 나타난 확률적 상황에 대한 접근 방식을 정리하면 <표 Ⅳ-6>와 같다.

<표 Ⅳ-6> 여러 확률적 상황에 대한 접근 방식별 분류

| 접근 방식 | | 2-1 | 2-2 | 2-3 | 2-4 | 2-5 | 3-1 | 3-2 |
|---|---|---|---|---|---|---|---|---|
| 고전적 | | 24 | 10 | | | | | |
| 빈도적 | 도수빈도형 | | 1 | 9 | 2 | 2 | | 1 |
| | 성향빈도형 | | | 2 | | | 18 | |
| 주관적 | 대인신념형 | | | | 6 | 2 | | |
| | 개인신념형 | | 13 | 13 | 16 | 20 | 6 | 20 |
| 확률값 미응답 | | | | | | | | 3 |

고전적 확률은 등확률성을 만족하는 상황에서만 분류되기 때문에 문항 2-1번과 2-2번 문항에서만 가능한 상황이다. 2-1번에서 제시된 동전과 같은 익숙한 확률 생성 도구를 이용한 예에서는 예비교사 모두가 고전적 확률을 통해

계산하는 양상을 보였다. 로또 당첨에서는 공이 뽑히는 상황을 수학적 상황으로 해석한 경우에서만 고전적 확률을 통해 접근하였다.

빈도적 확률과 주관적 확률을 구분하는 기준은 앞에서도 언급하였듯이 확률에 대한 판단 근거가 개인의 마음속에 존재하는지, 물질세계에 존재하는지 여부를 판단하여 분류하였다. 그중에서 과거의 경험을 기반으로 판단할 때 경계가 모호한 예가 존재하였다. 예를 들면 3-1번 문항에서 "주관적인 경험상 윷 1번에 모가 4번쯤 나온다."라는 근거를 들은 예는 주관적 확률로 분류하였고, 2-3번 문항에서 "번개에 맞는 사례가 존재한다."고 응답한 예는 빈도적 확률로 분류하였다. 윷의 사례는 개인에게 존재하는 기억에 의한 근거이므로 주관적 확률로 분류하였고, 번개의 사례는 과거의 경험이 뉴스, 기사, 책 등의 자료에 의한 근거이므로 빈도적 확률로 분류하였다. 이렇듯 과거의 경험이라도 개인의 경험에 대한 기억과 같이 정보가 마음속에 존재하면 주관적 확률로 분류하였고, 기사, 뉴스 등 빈도에 대한 정보가 물질세계에 존재하면 빈도적 확률로 분류하였다.

빈도적 확률에서 도수빈도형은 상대도수를 구하는 경우가 일반적이지만, 질문 문항이 실제 시행을 할 수 없는 확률적 상황들로만 구성되어 있으므로 위의 예와 같이 과거의 경험 빈도에 대한 정보가 물질세계에 존재한 예들만 있었다. 성향빈도형 확률은 윷놀이와 벼락 맞을 확률에서만 등장했는데, 이 두 사례는 예비교사가 물리적 성질에 대한 지식을 어느 정도 갖춘 예이기 때문으로 판단된다.

주관적 확률에 대한 분류를 살펴보면 대인신념형 확률은 논리적인 판단 관계로 확률을 1로 할당한 문항 2-4번과 2-5번에서만 등장했다. 논리적인 판단은 여러 사람들이 이미 합의한 판단 규칙이기 때문에 대인신념형 확률에 속한다고 볼 수 있다. 개인신념형 확률은 예비교사가 특별한 근거를 제시하지 않아서 분류가 되는 경우도 있었지만, 개인이 마음속에 있는 판단의 증거를 서술한 경우도 개인신념형 확률로 판단하였다.

개인신념형 확률을 응답한 예를 살펴보면 주어진 확률에 대한 진술에서 본인이 아는 정보가 많지 않아 막연히 확률을 할당한 사례가 많았다. 특히 극단적인 상황에 대한 확률을 묻는 문항인 2-1번부터 2-5번에서 사건이 일어나거나 일어나지 않는 상황에 대한 막연한 가능성이 존재한다는 응답이 빈번하게 등장하

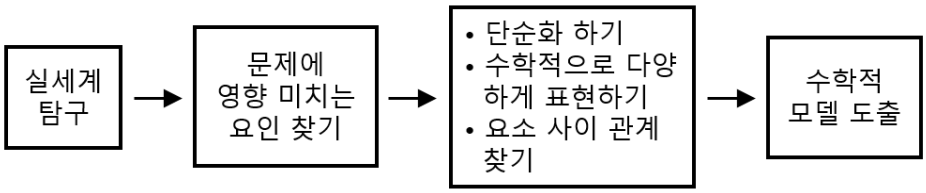
였다. 반면 문항 3-1번과 3-2번에서는 확률값을 구체적인 수로 할당해야 하기 때문에 개인신념형 확률 중에서도 예비교사 각자가 개별적으로 확률적 상황에 대한 수학적 모델을 제시하거나 개인의 경험에 기반하여 확률값을 판단하는 등 자신만의 판단 근거를 활용하는 모습을 보이는 사례가 더욱 많이 나타났다. 문항 2-1번부터 3-2번까지 7가지의 확률적 상황에 대하여 예비교사가 확률값을 추정하는 데 활용되는 정보를 범주화하여 정리하면 <표 Ⅳ-7>과 같다.

<표 Ⅳ-7> 접근 방식별 확률값 계산에서 활용되는 정보

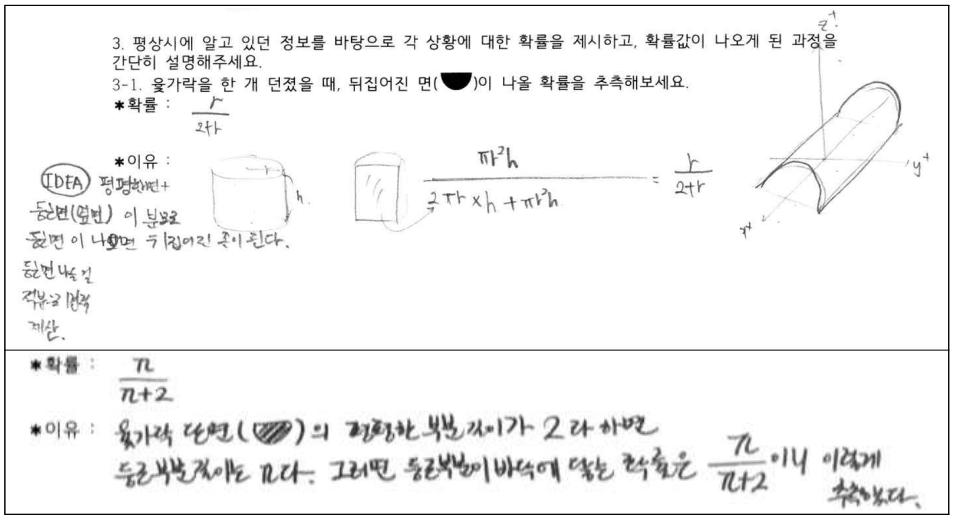
| 관점 | 접근 방식 | 확률값 계산에서 활용되는 정보 |
|---|---|---|
| 고전적 | 고전적 | 등확률성 모델링 |
| 빈도적 | 도수빈도형 | 과거의 사례에 대한 빈도 정보 |
| | 성향빈도형 | 물리적인 성질을 활용 |
| 주관적 | 대인신념형 | 논리적 포함관계, 전제(지식) |
| | 개인신념형 | 막연한 개인의 추측, 확률값을 추정하는 데 관련되는 요소에 대한 판단 |

## 2.6. 확률에 대한 접근 방식과 모델링

문항 3-1번과 3-2번은 윷을 던지는 상황과 코로나에 감염될 상황에 대한 확률을 추정하는 문항이다. 경험적인 정보는 주어지지 않은 채 개인이 가지고 있는 사전 정보만을 가지고 확률을 추정할 수밖에 없는 상황이다. 이는 Pfannkuch & Ziedins(2014)가 제시한 확률 추정의 방법 중에서 모델링 확률을 구하는 상황과 일치한다. 예비교사가 참 확률을 추정하기 위해서 활용한 접근 방식이 정혜윤 외(2018)가 제시한 모델에 적용 가능한지 예를 통해 살펴보고자 한다.



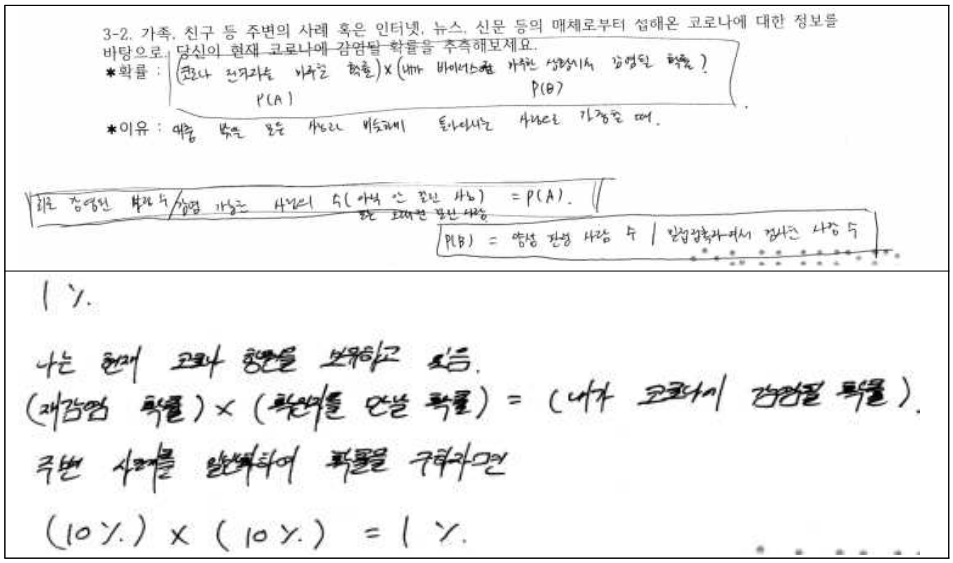

[그림 Ⅳ-10] 수학적 모델 도출 과정의 세분화(정혜윤 외, 2018)




[그림 Ⅳ-11] 윷의 확률에 대한 응답의 예

[그림 Ⅳ-11]의 첫 번째 예에서 ‘문제에 영향을 미치는 요인 찾기’ 단계에서는 윷이 바닥에 닿는 면의 넓이가 확률에 영향을 미친다고 판단하였다. ‘단순화하기’에서는 윷이 반원 모양을 가진다고 가정을 하였고, ‘수학적으로 다양하게 표현하기’에서는 넓이를 구하기 위해서 3차원 좌표계를 도입하였다. 이 과정을 통해 윷이 뒤집힐 확률이 $\frac{r}{2+r}$임을 추정하였다.

두 번째 예를 살펴보면 ‘문제에 영향을 미치는 요인 찾기’ 단계에서는 윷이 바닥에 닿는 둘레의 길이가 확률에 영향을 미친다고 판단하였다. ‘단순화하기’에서는 윷이 반원 모양을 가진다고 가정을 하여, 확률이 $\frac{\pi}{\pi+2}$임을 추정하였다.




[그림 Ⅳ-12] 코로나 감염 확률에 대한 응답의 예

[그림 Ⅳ-12]의 첫 번째 예에서 ‘문제에 영향을 미치는 요인 찾기’ 단계에서는 코로나 전파자를 만나는 상황과 바이러스를 마주했을 때 실제로 감염되는 상황이 해당된다고 보았다. ‘단순화하기’에서는 두 상황에 대한 확률을 구할 때 필요한 분자와 분모를 간단한 상황으로 제한하였다. ‘요소 사이 관계 찾기’에서는 두 확률이 독립이라고 가정하여 두 확률을 곱함으로써 구할 수 있다고 보았다. 하지만 분자와 분모에 대한 데이터를 현재는 알지 못하므로 참 확률의 추정값을 도출해내지는 못하였다.

두 번째 예를 살펴보면 ‘문제에 영향을 미치는 요인 찾기’ 단계에서는 자신은 코로나에 감염되었던 경험이 있으므로, 재감염 확률과 확진자를 만나는 상황이 해당된다고 보았다. ‘요소 사이 관계 찾기’에서는 두 확률이 독립이라고 가정하였다. 두 확률을 대략 10%로 가정해서 참 확률이 0.01이 될 것으로 추정하였다.

위의 네 가지 예를 살펴봄으로써 정혜윤 외(2018)의 모델로 주관적으로 확률을 추측하는 과정을 분석할 수 있음을 확인하였다. 몇몇 예비교사들은 확률적 상황에서 주관적으로 확률값을 추정하는 데 어려움을

겪었었는데, 정혜윤 외(2018)의 모델을 먼저 제시하고 모델에서 제시하는 단계를 하나하나씩 밟아나가면 확률값을 추정하는 데 도움을 줄 수 있을 것으로 보인다.

## 3. 추가 정보가 주어졌을 때의 참 확률에 대한 추정

이 절에서는 확률적 상황에 대한 추가적인 경험적 정보가 제시되었을 때 예비교사의 응답을 통하여 확률값을 어떠한 방식으로 조정해나가는지 살펴보도록 한다.

### 3.1. 빈도적 확률과 집산

<1차 설문지>의 문항 3-1번은 예비교사가 가지고 있는 사전 정보를 활용하여 Pfannkuch & Ziedins(2014)가 제시한 모델링 확률을 계산하였다. 이 확률을 ①번 확률로 예비교사에게 제시하였다. <2차 설문지>의 윷놀이 상황은 윷을 실제로 던짐으로써 경험적 확률과 모델링 확률의 상호작용을 살펴볼 수 있게 된다. 특히 von Mises가 제시한 집산의 개념을 활용하여 예비교사 개인마다 추가적으로 3가지 종류의 경험적 확률을 얻게 되고 각각의 확률에 대한 객관성과 적합성을 판단하게 된다.

<2차 설문지>의 1-1번 문항은 예비교사가 윷이 뒤집힐 확률을 구하기에 적합한 최소 횟수만큼 실제로 윷을 던지도록 요구하는 문항이다. 분석에 활용된 20명의 윷을 던지는 시행에 대한 결과는 표 <표 Ⅳ-8>과 같다. 여기서 구해진 경험적 확률을 ②번 확률이라고 예비교사에게 제시하였다.

<표 Ⅳ-8> 1-1번에서 구한 예비교사 개인별 빈도적 확률(②번 확률)[7]

| 예비교사 | 던진 횟수 | 응답데이터 | 뒷면 횟수 | ②번 확률 |
|---|---|---|---|---|
| S02 | 25 | 1010101010000000000000000 | 5 | 0.2 |
| S03 | 25 | 1111011111001000011001111 | 16 | 0.64 |
| S04 | 50 | 00010001101010010000001010101000000011001000000001 | 14 | 0.28 |
| S05 | 20 | 00111111011101101101 | 14 | 0.7 |
| S06 | 20 | 00011100100001100101 | 8 | 0.4 |
| S07 | 20 | 00001000001001010001 | 5 | 0.25 |
| S08 | 25 | 0001011010011101110011110 | 14 | 0.56 |
| S10 | 50 | 10110000101101111011011110111101000001101001000011 | 28 | 0.56 |
| S11 | 30 | 000110100000010110000101000001 | 9 | 0.3 |
| S12 | 25 | 1000000011100000000111101 | 9 | 0.36 |
| S13 | 30 | 010001001111100000000101100100 | 11 | 0.37 |
| S15 | 30 | 011000100101110110010110101101 | 16 | 0.53 |
| S16 | 20 | 01011100000001011011 | 9 | 0.45 |
| S17 | 15 | 101101110010011 | 9 | 0.6 |
| S18 | 25 | 1000100100001000001001101 | 8 | 0.32 |
| S20 | 33 | 011010010100101110110010101001100 | 16 | 0.48 |
| S21 | 44 | 01001100111010111010100111001011111011110111 | 28 | 0.64 |
| S22 | 20 | 01000110110001001010 | 8 | 0.4 |
| S23 | 16 | 1011001010101100 | 8 | 0.5 |
| S24 | 31 | 1000110111000001111111111011010 | 19 | 0.61 |

확률과 통계 교과서에서 등장하는 개념인 모비율의 추정을 이용하면 모비율에 대한 신뢰구간을 구함으로써 확률값을 추정할 수 있다. 하지만 던진 횟수를 판단하는 기준은 학습한 경험이 없으므로 문항 1-1번을 통해 개인마다 어떤 근거로 던진 횟수를 결정하는지 확인하는 문항이다. 던진 횟수에 대한 판단 근거를 분류하면 <표 Ⅳ-9>와 같다.

7) 예비교사에 대한 번호 S01~S24는 <1차 설문지>의 응답을 기준으로 부여하였다. <2차 설문지>의 분석에서 빠진 예비교사 S01, S09, S14, S19를 제외한 20명의 데이터를 표로 나타낸 것이다.

<표 Ⅳ-9> 던진 횟수에 대한 판단 근거

| 판단 근거 | 인원(명) |
|---|---|
| 이 정도 횟수면 적당하다고 판단해서 | 7 |
| 패턴이 어느 정도 반복되는 경향을 발견함 | 4 |
| 분모가 25면 확률이 유한소수로 계산되므로 | 2 |
| 50번이 최대이므로 | 2 |
| 교과서에서 $n$이 30이상이면 충분하다고 배움 | 2 |
| <1차>에서 예측한 확률 ①과 비슷한 값을 얻음 | 1 |
| 시간의 제약 때문에 충분히 던지지 못함 | 1 |
| 무응답 | 1 |
| 합계 | 20 |

<표 Ⅳ-9>에서 나타난 판단 근거를 살펴보면, 정해진 기준이 없으므로 개인이 적당하다고 생각한 횟수만큼 시행했다고 응답한 예비교사는 7명으로 가장 많았다. 이외에 실제로 판단에 사용된 근거는 데이터의 패턴, 계산의 편의성, 최대 시행 횟수, 시행 횟수에 대한 학습의 경험, 주관적 확률값과의 비교 등이 있었다.

<2차 설문지>의 문항 1-2번과 1-3번은 설문에 참여한 예비교사가 던진 윷놀이 데이터를 모두 합쳤을 때 그 데이터가 윷이 뒤집힐 확률을 구하는 데 적합한 집산인지 판단하도록 설계한 문항이다. 문항별로 살펴보면 1-2번 문항에서는 예비교사 각자가 윷을 원하는 횟수만큼 던져서 생성한 데이터를 종합하여 제시한 뒤, 윷이 뒤집힐 확률을 구하기에 부적합한 예비교사의 집산이 있으면 제외하도록 하는 문항이다. 20명의 응답 중 12명은 제외할 집산이 없다고 응답하였고, 제외할 집산이 있다고 응답한 예비교사는 8명이다. 제외하는 이유를 분석해보면 8명 중에서 3명은 최소 횟수가 상대적으로 적었기 때문에 데이터를 제외하였고, 2명은 확률값이 본인의 기준과 차가 크기 때문에 데이터를 제외하였다. 위의 두 이유를 모두 들어 제외한 인원은 1명이었고, 2명은 데이터의 분포가 편향된 집산을 제외하였다.

1-3번 문항에서는 1-2번에서 예비교사가 각자 선별하고 남은 데이터 중에서 예비교사 개개인이 윷을 던진 횟수를 모두 똑같이 일치시켜야 하는지 묻는 문항이다. 20명의 응답 중 4명이 윷을 던진 횟수를 일치시키는 것이 윷이 뒤

집힐 확률을 구하기에 더 적합하다고 판단하였다. 4명 중 3명은 최소 횟수만큼 일치시켜야 한다고 응답했으며, 1명은 자신이 주관적으로 패턴을 분석한 결과로 판단한 횟수로 일치시켜야 한다고 응답하였다.

문항 1-4번에서는 예비교사가 개인별로 구한 4가지 종류의 윷이 뒤집힐 확률값을 적도록 하였으며, 각 확률에 대한 설명은 다음과 같다. ① <1차 설문지> 3-1번에서 응답한 주관적 확률, ② <2차 설문지>의 1-1번에서 구한 혼자서 윷을 던져 구한 빈도적 확률, ③ 모든 예비교사가 생성한 윷의 데이터를 <2차 설문지>의 1-2번, 1-3번 활동을 통해 개인별로 선별한 데이터로 집산을 구성한 빈도적 확률, ④ 모든 예비교사가 생성한 윷의 데이터로 구한 빈도적 확률이다. 20명의 예비교사의 4가지 확률을 나열하면 <표 Ⅳ-10>과 같다.

<표 Ⅳ-10> 예비교사가 생성한 4가지 종류의 확률

| 예비교사 | ①번 확률 | ②번 확률 | ③번 확률[8] | ④번 확률 |
|---|---|---|---|---|
| S02 | 0.41 | 0.2 | 0.43 | 0.46 |
| S03 | 0.53 | 0.64 | 0.46 | |
| S04 | $\frac{r}{2+r}$ | 0.28 | 0.46[9] | |
| S05 | 0.6 | 0.7 | 0.48 | |
| S06 | 0.5이상 | 0.4 | 0.41 | |
| S07 | 0.75 | 0.25 | 0.46 | |
| S08 | 0.4 | 0.56 | 0.45 | |
| S10 | 0.61 | 0.56 | 0.46 | |
| S11 | 0.5이하 | 0.3 | 0.456 | |
| S12 | 0.25 | 0.36 | 0.46 | |
| S13 | 0.61 | 0.37 | 0.46 | |
| S15 | 0.5 | 0.53 | 0.45 | |
| S16 | 0.6 | 0.45 | 0.46 | |
| S17 | 0.6 | 0.6 | 0.46 | |
| S18 | 0.5 | 0.32 | 0.46 | |
| S20 | 0.5이상 | 0.48 | 0.51 | |
| S21 | 0.67 | 0.64 | 0.43 | |
| S22 | 0.61 | 0.4 | 0.46 | |
| S23 | $\frac{180+2\tan^{-1}\left(\frac{i}{n}\right)}{360}$ | 0.5 | 0.46 | |
| S24 | 0.7 | 0.61 | 0.46 | |

문항 1-2번에서 데이터를 선별하지 않은 12명 중에서 예비교사 S21만 문항 1-3번에서 개개인이 던진 횟수를 일치시켜야 한다고 응답하였다. 남은 11명은 문항 1-3번에서 개개인이 던진 횟수를 일치시키지 않도록 응답하였으므로 확률 ③과 ④의 확률값이 서로 같다. 확률 ③과 ④의 값이 다른 9명의 예비교사가 생성한 확률 ③은 <표 Ⅳ-10>의 해당되는 칸에 음영 처리를 하여 강조하였다. 같은 확률값이 나온 11명의 예비교사에게도 확률 ③과 ④를 구별짓기 위해 확률 ③에서 최소 1개 이상의 데이터가 선별된 집산으로부터 구한 확률로 간주하고 응답하도록 설문지를 구성하였다. 하지만 설문의 의도를 파악하지 못하여 11명 중 4명은 확률 ③과 ④를 같은 확률로 취급하였다. 이들 4명의 응답은 확률 ③을 제외하고 ①, ②, ④의 세 종류의 확률만으로 문항 1-5번과 1-6번을 분석하였다.

문항 1-5번은 4종류의 확률 중에서 어느 확률이 가장 객관적인지 판단하는 문항이다. 미응답자 1명을 제외한 19명의 가장 객관적인 확률을 고르는 응답의 분포는 ①은 3명, ②는 1명, ③은 1명, ④는 14명이다. 확률 ④가 가장 객관적이라고 응답한 이유를 살펴보면 주관에 의해 데이터를 선별하지 않았고, 데이터의 수가 가장 많기 때문으로 응답하였다. 확률 ①이 가장 객관적이라고 응답한 이유는 이론적으로 예측한 수학적 확률이라고 응답한 예비교사가 2명이 있었고, 다른 1명은 전체 시행 횟수가 충분치 않아 ④의 빈도적 확률보다 더 객관적이라고 판단하였다.

문항 1-6번은 4종류의 확률 중에서 어느 확률이 가장 윷이 뒤집힐 확률을 나타내기에 적합한지 판단하는 문항이다. 미응답자 1명을 제외한 19명의 가장 적합한 확률을 고르는 응답의 분포는 ①은 4명, ②는 1명, ③은 4명, ④는 10명이다. 확률 ④가 가장 적합하다고 응답한 이유는 표본 수가 많고 가장 객관적인 확률이기 때문이라고 설명하였다. 확률 ③은 적절한 데이터만을 선별하였기 때문이라고 응답하였다. 확률 ②로 응답한 1명은 자신이 예측한 확률 ①과 가장 비슷한 확률값이 나왔기 때문이라고 응답하였다. 확률 ①은 이론상으로

8) 음영 처리된 확률은 한 개 이상의 데이터를 제외해야 한다고 응답한 9명의 ③번 확률이다.

9) 문항 1-2번에서 제외할 집산이 있다고 응답하였지만, 엑셀에서 필터링 기능을 저장하지 않아서 필터링되지 않은 확률인 0.46으로 계산되었다.

구한 확률이고, 총 시행 횟수가 충분치 않기 때문이라고 응답하였다.

문항 1-5번과 1-6번을 연관 지어 확률의 객관성과 적합성을 살펴보도록 하자. 1-5번과 1-6번의 응답이 일치하는 응답은 19개 중 15개로 객관적인 확률이 윷이 뒤질힐 확률에 적합하다고 판단하였다. 확률의 객관적인 정도와 적합성이 다른 4명 중 3명은 가장 객관적인 확률은 확률 ④로 응답하고, 가장 적합한 확률은 확률 ③으로 응답하였다. 이들은 적절한 인간의 개입이 더 적합한 확률을 구할 수 있다고 판단한다고 볼 수 있다. 남은 1명은 가장 객관적인 확률은 확률 ④로 응답하고, 가장 적합한 확률은 확률 ①로 응답하였는데, 확률 ①이 여러 번의 시행을 거치지 않아도 판단되는 사고 실험이기 때문이라고 서술하였다. 이는 빈도적 확률에서 도수빈도형 확률보다 성향빈도형 확률이 더 적합하다고 판단한 것으로 해석할 수 있다.

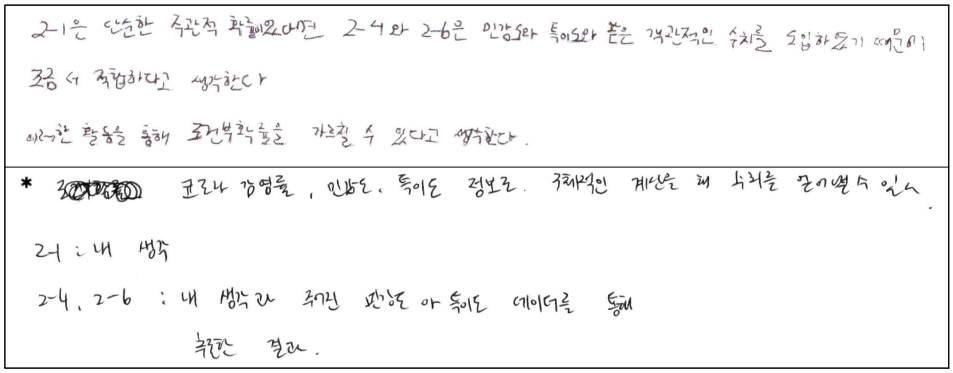
## 3.2. 주관적 확률과 베이즈 정리

<2차 설문지>의 문항 2-1번부터 2-6번까지는 <1차 설문지>에서 추측했던 본인이 코로나에 감염될 확률을 검사키트의 양성 여부에 따라 확률이 갱신되는 과정을 모델링한 상황이다. 신속항원검사와 비인두도말 PCR 검사 도구의 민감도와 특이도는 질병관리청에서 제공한 수치를 사용하여 실제 상황과 가깝게 설계하였다.

세부적으로 문항 2-1번은 주관적 확률로부터 10만 명당 코로나에 감염된 사람 수를 계산하는 문항이다. 문항 2-2번부터 문항 2-4번 문항은 신속항원검사에서 양성으로 나타났을 때, 코로나에 감염됐을 확률을 구하는 상황이다. 문항 2-5번과 문항 2-6번은 추가적으로 실시한 PCR 검사에서 음성으로 나타났을 때, 코로나에 감염됐을 확률을 구하는 상황이다. 이렇듯 일련의 활동에서 코로나 감염 확률에 관한 3가지 확률이 구해진다. 조건부확률과 관련된 상황으로 각각의 확률값은 분할표를 통해 계산된다.

문항 2-7번은 위의 세 확률을 비교하는 질문과 교수학적으로 어떤 의미를 지니는지 묻는 질문이 혼합되어 있다. 20명의 예비교사 중 8명은 객관적인 증거인 검사 결과에 대한 정보가 조건부확률로 결합하여 확률이 계산됨을 설명하였다. 이

는 개인신념형 확률인 코로나 감염 확률인 검사 결과의 양성 유무인 경험적 정보와 결합하여 갱신된 확률은 대인신념형 확률이 되는 과정에 대한 이해를 갖춘 응답으로 해석할 수 있다. 대표적인 예비교사의 응답은 [그림 Ⅳ-13]와 같다.


2-1은 단순한 주관적 확률이었다면 2-4와 2-6은 민감도와 특이도와 같은 객관적인 수치를 도입하였기 때문에 조금 더 적합하다고 생각한다
이러한 활동을 통해 조건부확률을 가르칠 수 있다고 생각한다.

* 코로나 감염률, 민감도, 특이도 정보로 구체적인 계산을 해 수치를 구해볼 수 있다.
2-1 : 내 생각
2-4, 2-6 : 내 생각과 주어진 민감도 와 특이도 데이터를 통해 유추한 결과.

[그림 Ⅳ-13] 베이즈 정리와 확률 갱신 과정에 대한 예비교사의 답변

### 3.3. 빈도적 확률과 가장 좁은 준거집합의 원리

문항 2-8번은 유병률을 이용하여 코로나에 감염될 확률을 추정해보는 활동이다. Popper는 가장 좁은 준거집합의 원리를 적용하면 같은 도수빈도형 확률이더라도 더욱 객관적인 확률을 얻을 수 있다고 판단하였으므로 예비교사의 응답도 이에 해당하는지 확인하는 문항이다. 문항 2-8번에서 제공한 인구현황은 성별(남녀), 지역(전국과 서울), 연령(20대 유무)로 세분화되었으며, 유병률의 분모에 해당하는 인구집단을 선택할 수 있는 방법의 수는 12개가 있다. 반면 유병률의 분자에 해당하는 누적 확진자수는 전국, 서울시, 전국 20대, 전국 남성, 전국 여성으로 총 5가지 데이터가 제공되었다. 누적 확진자수에 대응하는 인구집단을 모집단으로 설정하면 분자와 분모의 정확한 수치를 모두 알고 있으므로 한 번의 계산으로 유병률을 구할 수 있다. 반면 누적 확진자수가 제공되지 않은 7개의 인구집단을 모집단으로 설정하면 누적 확진자수의 비율에 따라 유추해서 유병률을 계산해야 한다. 12개의 유병률 중에서 가장 넓은 준거집합은 전국 인구수를 모집단으로 구한 유병률이고, 가장 좁은 준거집합은 서울에 거주

하는 20대 남자 혹은 여자 인구수를 모집단으로 구한 유병률이다.

20명의 예비교사가 선택한 모집단을 분류해보면 전국 인구수는 8명, 서울 인구수는 1명, 전국 20대 인구수는 3명, 전국 20대 남(여)성 인구수는 2명, 서울 20대 인구수는 1명, 서울 20대 남(여)성 인구수는 4명, 기타는 1명으로 분류되었다. 전국 인구수가 8명으로 가장 많은 응답이 제시된 결과로 판단하건대, 예비교사들은 코로나 유병률을 구할 때 가장 좁은 준거집합의 원리을 따르지 않는다고 볼 수 있다. 예비교사들이 모집단으로 자신이 해당되는 서울 20대 남(여)성 인구를 고르지 않고 다양한 모집단을 선택한 이유를 살펴보면 개인마다 생각하는 코로나 감염에 관련된 속성이 다르기 때문으로 볼 수 있다. 어떤 예비교사는 나이가 면역력에 관련된 요소로 보고, 다른 예비교사는 성별은 코로나 감염에 무관하다고 판단하기도 하였다. 또한 서울 이외의 지역을 많이 돌아다니는 예비교사들은 서울이 아닌 전국을 모집단으로 선택하였다. 이렇듯 단순히 가장 좁은 준거집합을 선택하는 것이 적합한 확률을 구할 수 있을 것으로 판단하지 않고, 개개인마다 코로나 감염과 관련된 요소를 선별하여 판단하는 것을 알 수 있었다.

### 3.4. 주관적 확률과 빈도 원리

문항 2-9번은 2-1번에서 활용한 주관적 확률과 2-8번에서 유병률을 통해 추측한 코로나 감염 확률을 비교하는 문항이다. 코로나에 감염될 확률을 구하기 위해 2-1번에서 사용된 확률은 개인신념형 확률로 볼 수 있고, 2-8번에서 사용된 확률은 도수빈도형 확률이라고 판단할 수 있다. Hacking(2001)이 설명한 빈도 이론에 따르면 코로나에 감염될 확률에 대하여 아는 정보가 누적 확진자수 밖에 없다면 코로나에 감염될 주관적 확률을 유병률로 계산한 확률로 대체해도 된다는 원리이다.

20명의 예비교사 중 무응답을 제외한 18명의 답변으로 두 확률의 특성을 살펴보도록 한다. 기본적으로 “전자는 추측에 의한 주관적 판단이고, 후자는 데이터에 의한 객관적 판단이다.”라는 응답처럼 대부분 전자를 주관적 확률, 후자를 빈도적 확률로 해석하였다. 주관적 확률의 특징

으로는 “개인의 경험을 고려한 추측이 가능하다.”, “미래에 대한 추측이다.”, “주관적이어서 모호하다.” 등을 서술하였다. 빈도적 확률의 특징으로는 “통계 데이터를 활용하여 객관적이다.”, “모든 사람의 경험이 반영된 확률이다.”, “내가 코로나에 걸릴 확률을 직접적으로 말해주기는 어렵다.” 등을 서술하였다. 정리하면 주관적 확률은 본인이 코로나에 걸릴 확률을 설명하기에 더 적합하다고 판단하는 예비교사들이 있었고, 빈도적 확률은 통계 데이터를 활용하기 때문에 객관적인 확률로 더 적합하다고 판단하는 예비교사들이 있었다.

### 3.5. 예비교사가 판단한 주관적 확률의 장단점

<2차 설문지>의 문항 3-2번에서는 설문 과정에서 활동했던 내용을 바탕으로 주관적 확률의 장단점을 서술하는 문항이다. 예비교사의 응답을 바탕으로 주관적 확률의 장단점을 서술하면 <표 Ⅳ-11>와 같다.

<표 Ⅳ-11> 주관적 확률의 장단점

| 장점 | 단점 |
|---|---|
| • 개개인의 생각을 수치로 나타낼 수 있다.<br>• 일상생활 속 여러 상황을 수학적으로 생각해 볼 수 있는 사고력을 기를 수 있다.<br>• 등확률 가정이 없거나 반복해서 시행할 수 없는 사건의 확률을 주관적으로 정의 할 수 있다.<br>• 각 개인은 처한 상황이 모두 다르기에 개인에 따라 적절한 확률을 제시할 수 있다.<br>• 확률값을 빠르고 쉽게 구할 수 있다.<br>• 근거가 충분하다면 주관적 확률이 일상생활에서 꽤 유용할 것이다. | • 확률의 정의, 의미에 대해 오개념이 생길 수 있다.<br>• 사람마다 동일한 사건에 대한 확률을 다르게 계산할 수 있다.<br>• 개인의 잘못된 신념이나 편견이 적용될 수 있다.<br>• 정확성이 보장될 수 없으며, 실제 확률과 차이가 크게 날 수 있다.<br>• 신뢰하기 어렵다.<br>• 구할 때마다 달라질 수 있다. |

주관적 확률의 장점은 고전적, 빈도적 확률로 계산할 수 없는 다양한 상황에 대한 확률을 제공하는 점에서 나타나는 특성이 주를 이루었다. 이는 일상생활에 있는 상황을 해결하거나 개인에 관한 판단을 할 때 도움이 된다는 응답이 여럿 존재하였다. 또한 주관적 확률을 구하는 데 시간이 가장 적게 들었다는 경험을 바탕으로 주관적 확률을 빠르고 쉽게 구할 수 있다고 판단하였다.

반면에 단점으로는 대부분 부정확성을 들었다. 동일한 사건이라도 개인의 편견이 작용하여 실제 확률과 다를 수 있으며, 개개인마다 편차가 클 수 있다는 위험이 있다고 판단하였다. 또한 같은 개인이라도 계산할 때마다 다를 수 있다고 응답한 내용은 주관적 확률의 신뢰성에 대한 의문을 가졌다고 볼 수 있다. 학습 면에서는 학생들에게 확률에 대한 오개념을 형성할 수 있다는 점에 대해 우려를 표했다.

### 3.6. 예비교사의 확률에 대한 관점의 변화

<2차 설문지>의 3-3번 문항은 설문 과정에서 활동했던 내용을 바탕으로 빈도적 확률과 주관적 확률에 대한 관점의 변화가 있는지를 묻는 문항이다. 관점의 변화가 있다고 응답한 예비교사는 14명이고, 6명은 변화가 없다고 응답하였다. 변화가 없다고 응답한 6명의 이유를 살펴보면 고전적 확률 및 빈도적 확률이 확률을 나타내는 개념이라는 생각에 변화가 없었다. 주관적 확률은 실제 확률은 아니고 상황을 모델링하고 예측하는 도구의 역할 정도만 수행한다고 보았다.

확률에 대한 관점에서 변화가 있다고 응답한 예비교사의 응답을 분석해보면, 주관적 확률도 확률로서 다뤄질 수 있으며 일상생활에서 자주 사용되고 있음을 깨닫게 되었다고 설명하였다. 또한 확률이란 객관적인 수치가 아니라 사람들마다 다를 수 있음을 깨닫게 되었고 오히려 모든 사람에게 동일한 값을 적용하는 확률이 부정확할 수 있다고도 판단하였다. 이들은 주관적 확률을 인정하긴 하지만 개인의 막연한 생각을 확률로 할당하기보다는 앞선 활동해서 해왔던 것처럼 주관적 확률이 객관성

을 갖추기 위해서 노력해야 한다고 서술하였다.

이러한 확률에 대한 세 가지 관점에 대한 생각의 변화를 바탕으로 <2차 설문지>의 3-4번 문항은 중학생 및 고등학생에게 주관적 확률을 가르칠 의사가 있는지를 선택하는 질문이다. 중·고등학생 모두에게 가르쳐야 한다는 응답은 8명, 중학생만 가르쳐야 한다는 응답은 3명, 고등학생만 가르쳐야 한다는 응답은 5명, 중·고등학생 모두에게 가르치지 말아야 한다는 응답은 4명으로 학생들에게 주관적 확률을 가르쳐야 한다고 생각하는지에 대한 판단은 의견이 분분하였다.

주관적 확률을 가르쳐야 된다는 의견으로는 "일상생활에서 판단을 내릴 때 도움이 된다.", "고전적 확률만 확률이라고 생각하는 것을 방지할 수 있다.", "확률에 대한 폭넓은 시야를 갖게 한다.", "고전적, 빈도적 확률로 구할 수 없는 확률을 계산하게 한다." 등이 있었다.

반면 주관적 확률을 가르치지 말아야 한다는 의견으로는 "확률에 대한 오개념이 생길 것 같다.", "평가 방식에 대한 고민이 필요하다.", "확률에 대한 개념에 혼동이 올 것 같다.", "난이도가 높아서 학생들이 이해하기 어렵다" 등이 있었다.

## 4. 주관적 확률에 대한 교수학적 함의

1, 2차 설문지의 문항과 응답을 참고하여 중·고등학교 교육과정에 적용할 수 있는 방안을 살펴보도록 한다. <1차 설문지>의 2-1번에서 2-5번까지의 문항은 확률이 거의 0 또는 거의 1인 상황을 제시하고 확률값을 할당해보는 문항이었다. 앞서 언급하였듯이 중학교 2학년 확률 단원에서 절대로 일어나지 않는 사건과 반드시 일어나는 사건의 확률을 각각 1과 0임을 학습한다. 그리고 예제로 나오는 문항은 주사위의 두 눈의 합이 1일 확률, 주사위의 두 눈의 합이 2 이상일 확률과 같이 경우의 수가 표본공간에 속하지 않으면 확률이 0, 사건과 표본공간이 같은 상황이면 확률이 1로 설명한다. 이는 확률이 0과 1이라는 성질을 단순히 공식에 확인하는 절차밖에 되지 않는다.

반면에 2-3번과 2-5번과 같이 '내가 길을 가다가 벼락에 맞을 확률', '내일 아침에 해가 뜰 확률'을 학생들에게 추측해보라고 하고 서로 토의하는 활동을 함으로써 실생활에서 확률이 0과 1이 되는 상황을 생각해보면서 간단하게 주관적 확률을 도입할 수 있다.

또한, 이경화(1996)는 주관적 확률이 형식화, 알고리즘화가 불가능하며, 평가 방법도 모호하기 때문에 주관적 확률을 수업 시간 및 교과서에서 다루기 힘들다고 하였다. <2차 설문지>의 두 번째 모델링 활동인 코로나 감염 확률에 대해서 살펴보자. 2-1번 문항에서 각자 주관적으로 추측한 코로나 감염 확률에 기반하여 신속항원검사에서 양성이 나온 상황에 대한 분할표를 채우는 활동을 실시하였다. 민감도 0.9, 특이도 0.96인 신속항원검사 키트에서 양성이 나왔을 때, 분할표를 채운 두 예비교사의 응답의 예시는 [그림 Ⅳ-14]과 같다.

모델링 상황 2 : 코로나에 감염됐을 확률은?

| | 코로나 감염($A$) | 코로나 비감염($A^C$) | 계 |
|---|---|---|---|
| 신속항원검사 양성($B_1$) | 72000 | 800 | 72800 |
| 신속항원검사 음성($B_1^C$) | 8000 | 19200 | 27200 |
| 계 | 80000 | 20000 | 100,000명 |

**<표1. 신속항원검사와 코로나 감염 분할표>**

2-1. 지난 설문의 3-2에서 응답한 코로나에 감염됐을 확률( 0.8 )을 이용하여 다음의 분할표에서 코로나에 감염된 사람 수를 빨간색으로 채워보세요.

모델링 상황 2 : 코로나에 감염됐을 확률은?

| | 코로나 감염($A$) | 코로나 비감염($A^C$) | 계 |
|---|---|---|---|
| 신속항원검사 양성($B_1$) | 1800 $B_1 \mid A$ | 4900 | 6700 |
| 신속항원검사 음성($B_1^C$) | 200 | 93100 $B_1^C \mid A^C$ | 93300 |
| 계 | 2000 | 98000 | 100,000명 |

**<표1. 신속항원검사와 코로나 감염 분할표>**

2-1. 지난 설문의 3-2에서 응답한 코로나에 감염됐을 확률( 0.02 )을 이용하여 다음의 분할표에서 코로나에 감염된 사람 수를 빨간색으로 채워보세요.

[그림 Ⅳ-14] <2차 설문지> 2-1번 문항에 대한 두 예비교사의 답변

두 예비교사가 응답한 코로나에 감염됐을 확률이 각각 0.8과 0.02로 다르기 때문에 분할표에 채워진 숫자가 다르다.

| | | | | |
|---|---|---|---|---|
| 코로나 유병률 | 0.8 | | | |
| | | | | |
| | 코로나 감염 | 코로나 비감염 | 계 | |
| 신속항원검사 양성 | 72000 | 800 | 72800 | |
| 신속항원검사 음성 | 8000 | 19200 | 27200 | |
| 계 | 80000 | 20000 | 100000 | |
| 신속항원검사에서 양성으로 나왔을 때 코로나에 감염됐을 확률 | | | | 0.98901 |

| | | | | |
|---|---|---|---|---|
| 코로나 유병률 | 0.02 | | | |
| | | | | |
| | 코로나 감염 | 코로나 비감염 | 계 | |
| 신속항원검사 양성 | 1800 | 3920 | 5720 | |
| 신속항원검사 음성 | 200 | 94080 | 94280 | |
| 계 | 2000 | 98000 | 100000 | |
| 신속항원검사에서 양성으로 나왔을 때 코로나에 감염됐을 확률 | | | | 0.31469 |

[그림 Ⅳ-15] Excel로 계산한 두 예비교사의 분할표에 대한 정오

하지만 [그림 Ⅳ-15]를 살펴보면 Excel에서 자동으로 계산되는 표를 활용하여 채점한 결과 위의 예비교사는 정답이고 아래 예비교사는 오답으로 채웠음을 알 수 있다. 이처럼 Excel을 이용하여 초기의 코로나 감염 확률만 입력하면 개개인마다 수치가 다른 분할표를 제시하더라도 올바르게 분할표를 채웠는지 평가할 수 있게 된다.

2-4. 합계를 이용하여 나머지 표의 빈칸을 채워보고, 신속항원검사에서 양성으로 나왔을 때 실제 코로나에 감염됐을 확률을 구하고, 확률 기호를 이용하여 나타내보세요.

*확률 : 4800/8800 = 0.542...

*기호 : $P(A \mid B_i)$

2-4. 합계를 이용하여 나머지 표의 빈칸을 채워보고, 신속항원검사에서 양성으로 나왔을 때 실제 코로나에 감염됐을 확률을 구하고, 확률 기호를 이용하여 나타내보세요.

*확률 : 900/4860 = 약 18.5%

*기호 : $P(B_i \mid A) / (P(B_i \mid A) + P(B_i \mid A^c))$

[그림 Ⅳ-16] 조건부확률에 대한 두 예비교사의 기호 표현

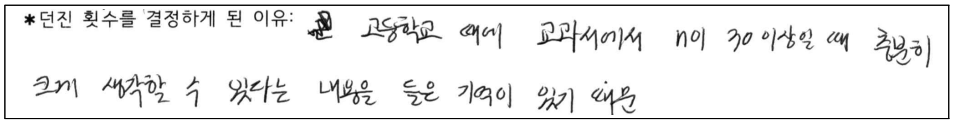
또한 [그림 Ⅳ-16]과 같이 글로 서술된 주관적 확률에 대한 상황을 기호로 표기하는 것을 확인하면서 평가하는 방법 또한 존재한다. [그림

Ⅳ-16]의 첫 번째 응답은 코로나에 감염되는 사건 $A$와 신속항원검사에서 양성이 나오는 사건 $B_1$를 조건부확률로 올바르게 표현한 사례고, 두 번째 응답은 $\mathrm{P}(A|B_1)$와 $\mathrm{P}(B_1|A)$을 혼동하여 잘못된 기호로 표현한 사례다.

<2차 설문지>의 3-4번 중·고등학생에게 주관적 확률을 가르쳐야 한다고 생각하는지 묻는 문항에서 중학생에게 가르쳐야 한다는 응답은 11명, 고등학생에게 가르쳐야 한다는 응답은 13명이 나왔다. 예비교사들도 학교수학에 주관적 확률의 도입이 필요하다는 점에 공감하고 있음을 알 수 있다.

## 5. 예비교사의 확률에 대한 오개념

이 장에서는 설문의 응답에서 드러난 예비교사의 확률에 대한 오개념을 살펴보도록 한다. 크게 두 가지 오개념이 설문에서 등장하였는데, 우선 윷을 던지는 횟수를 판단하는 근거를 살펴보는 <2차 설문지>의 문항 1-1번 문항의 응답의 한 사례인 [그림 Ⅳ-17]을 살펴보자.




[그림 Ⅳ-17] 큰 수의 법칙에 대한 오개념 사례

[그림 Ⅳ-17]의 사례처럼 “교과서에서 $n$이 30이상이면 충분하다고 배웠던 기억이 있다.”와 같이 응답한 예비교사는 2명이 있었는데, 교과서에서 $n$의 크기가 숫자 30과 관련지어 설명된 내용은 표본평균을 학습할 때 등장한다. [그림 Ⅳ-18]는 고등학교 확률과 통계 교과서에서 서술된 내용을 발췌한 것이다(권오남 외, 2019).

| | |
|---|---|
| 보통 $n$이 30보다 크면 충분히 큰 것으로 본다. | 또 모집단의 분포가 정규분포가 아닌 경우에도 $n$이 충분히 크면 표본평균 $\overline{X}$는 근사적으로 정규분포 $\mathrm{N}\left(m,\ \frac{\sigma^2}{n}\right)$을 따른다고 알려져 있다. |

[그림 Ⅳ-18] 표본의 크기와 표본평균의 분포에 대한 설명

[그림 Ⅳ-18]과 같이 모집단의 분포 형태에 상관없이 무작위로 추출된 표본의 크기가 충분히 크면 표본평균의 분포가 근사적으로 정규분포를 따른다는 사실은 중심극한정리와 관련이 있다(우정호, 2017). 하지만 1-1번 문항에서 집산의 크기를 통해 빈도적 확률을 추정하는 원리는 큰 수의 법칙과 관련이 있으므로 $n$이 30과 관련된 개념은 큰 수의 법칙과 중심극한정리를 혼동한 사례로 볼 수 있다.

두 번째 오개념은 <2차 설문지>의 문항 2-6번에서는 “신속항원검사와 달리 PCR 검사에서는 음성으로 나왔을 때 코로나에 감염됐을 확률”을 기호로 적게 하는 문항에서 등장하였다. 이 문항은 조건부확률의 두 가지 복합 사건이 조건부확률에 쓰이는 문항으로 정규 고등학교 과정에는 없는 복잡한 상황이다. 정답인 기호 $\mathrm{P}(A|B_1 \cap B_2^C)$를 응답한 인원은 2명뿐이고, 남은 18명 중에서 16명은 모두 동일한 오답인 $\mathrm{P}(A \cap B_1|B_2^C)$로 제시하였다. 복합 사건에 대한 조건부확률을 기호로 표시하는데 많은 예비교사들이 어려움을 겪고 있음을 알 수 있다.

한편, 이러한 응답의 결과가 난이도 때문일 수도 있겠지만, 확률적 상황에 대한 표현이 명확하지 않았던 요인도 작용했을 수 있다. 예비교사가 복합 사건에 대한 조건부확률의 개념에 대하여 어떠한 어려움을 겪는지 보다 더 명확하게 판단하기 위해서는 2-6번 문항에 대한 표현을 “신속항원검사에서는 양성이 나오고 PCR 검사에서는 음성으로 나왔을 때, 코로나에 감염됐을 확률”과 같이 수정하는 것이 도움이 될 것으로 판단된다.

# Ⅴ. 결론 및 제언

## 1. 요약

본 연구는 학교수학에서 확률교육이 고전적 확률 위주로 다뤄지고 있는 점에 대한 문제의식에서 출발하였다. 확률에 대한 고전적 관점만 강조해서 다루게 되면 학생들은 확률 개념이 우연 현상과 거리가 멀다고 느끼게 되며, 모든 확률적 상황에서 확률값이 한 가지 정답만 가진다는 착각에 빠질 위험이 있다. 미래를 대비하고 합리적인 의사결정을 하는 민주 시민으로 교육하기 위해서는 학교수학에서 다양한 문제해결 상황에 확률적 사고를 활용할 수 있도록 확률적 소양을 길러야 한다. 확률교육이 이러한 취지에 맞게 실행되기 위해서는 교사가 확률에 대한 다양한 관점을 이해하고 적용할 수 있어야 한다. 이를 위해 예비교사의 확률에 대한 관점을 살펴보는 연구를 진행하였다. 각 장에서 다룬 주요 내용은 다음과 같다.

Ⅱ장에서는 확률에 대한 관점과 관련된 선행 연구를 분석하였다. 첫째로 수학교육에서 주로 다루어지는 확률에 대한 세 가지 관점인 고전적, 빈도적, 주관적 확률의 특징 및 한계를 살펴보았다. 이후 확률의 두 가지 측면인 신념형 확률과 빈도적 확률에 대한 이론을 살펴보면서 객관성을 기준으로 개인신념형, 대인신념형, 성향빈도형, 도수빈도형으로 구분하였다. 이후 위의 네 가지 확률 개념을 수학교육에 도입했을 때 나타나는 장점을 살펴보면서 주관적 확률과 빈도적 확률의 세부 분류의 필요성을 확인하였다. 둘째로 확률의 각 관점에서 드러나는 주관적인 특성을 살펴보면서 확률 개념은 객관적인 측면과 주관적인 측면이 뒤섞여 있어 근본적으로 애매성을 가짐을 확인하였다. 셋째로 확률적 소양의 구성 요소를 살펴보면서 확률교육의 목표를 달성하기 위해서는 학교수학에서 세 가지 관점의 확률을 조화롭게 다뤄야 함을 확인하였다. 넷째로 확률과 관련된 교사 지식의 구성 요소를 살펴보면서 확률에 대한 여러 관점의

이해가 필수적임을 확인하였다. 마지막으로 등확률성을 만족하지 못하는 상황에서 개인이 가지고 있는 정보를 바탕으로 판단한 모델링 확률과 경험적 정보를 바탕으로 한 경험적 확률이 상호작용하면서 참 확률에 다가가는 과정을 설명한 모델을 살펴보았다.

Ⅲ장에서는 연구 방법을 설명하였다. 연구대상은 서울특별시에 있는 대학교 수학교육과에 재학 중인 예비교사로 선정하였다. 설문지는 총 2회 실시하였으며 예비교사들의 확률에 대한 관점을 알아보기 위한 문항으로 설계하였다. 마지막으로 Gal(2017)의 확률 계산의 절차, 확률적 상황에 대한 접근 방식에 대한 논의를 통해 예비교사의 확률에 대한 관점을 살펴보는 틀을 제시하였다.

Ⅳ장에서는 1차와 2차 설문 문항의 응답을 분석하여 세 가지 연구문제의 결과를 제시하였다. 다음과 같이 연구 결과를 요약할 수 있다.

첫째, 예비교사가 구성한 확률에 대한 예를 분석한 결과 예비교사가 현재 가지고 있는 확률에 대한 관점은 고전적 확률에 치우쳐 있음을 확인할 수 있었다. 또한 예비교사가 제시한 확률의 의미를 제도적 의미와 개인적 의미로 분류하였다. 교육기관이나 교과서에서 제공하는 의미인 제도적 의미는 대부분 고전적 확률만 등장하였으며, 통계적 확률이나 공리적 확률에 대한 의미는 거의 등장하지 않았다. 개인의 직관에 의해 형성된 의미인 개인적 의미를 키워드로 분석한 결과 확률을 대부분 “가능성을 표현하는 수”로 인식하는 것을 알 수 있었다. 주관적 확률과 관련된 키워드는 “추측 또는 예측”임을 알 수 있었다. 여러 학자들이 문제제기해왔던 확률교육이 고전적 확률에 치우쳐 있음을 확인할 수 있었다.

둘째, 등확률을 만족하지 않는 상황에서 예비교사가 확률값을 추정하는 과정을 분석하면서 각각의 접근 방식에서 활용되는 정보의 유형을 범주화할 수 있었다. 도수빈도형 확률에 대한 접근에서는 반복되는 시행을 실시할 수 없는 상황이었으므로 과거의 사례에 대한 빈도 정보에 대한 기억을 바탕으로 확률값을 추정하였다. 성향빈도형 확률에 대한 접근에서는 물리적인 성질에 대한 정보를 바탕으로 확률값을 추정하였다. 대인신념형 확률에 대한 접근에서는 확률적 상황에서 나타나는 논리적 포함

관계를 분석함으로써 확률값을 추정하는 모습을 보였다. 개인신념형 확률에 대한 접근에서는 확률적 상황에 대하여 가지고 있는 정보가 적은 경우에는 막연하게 추측한 개인의 신념을 바탕으로 확률을 추정하였고, 확률적 상황과 관련된 요소를 도출해낼 수 있는 경우에는 모델링 과정을 통해 확률값을 추정하는 양상을 드러내었다. 또한 정혜윤 외(2018)가 제시한 수학적 모델의 도출 절차를 통해 확률값을 추정하는 상황을 분석할 수 있었다.

셋째, 참 확률을 추정하는데 객관성이 어떠한 기준으로 쓰이는지 확인할 수 있었다. Pfannkuch & Ziedins(2014)는 모델링 확률과 경험적 확률이 상호작용하여 참 확률을 추정할 수 있다고 보았다. 이를 확인하기 위해 윷놀이 상황과 코로나 상황에서 개인마다 주관적으로 모델링 확률을 추측하고, 여러 경험적인 정보를 제공하면서 예비교사가 어떻게 확률값을 갱신해나가는지 살펴보았다. 빈도적 확률과 관련된 윷의 확률에서는 각자가 생성한 네 가지 종류의 확률에 대하여 대부분의 예비교사는 시행 횟수가 가장 많은 집산을 통해 구한 빈도적 확률이 가장 객관적이라고 판단하였다. 하지만 몇몇 예비교사는 모든 사람의 시행 결과를 합쳐도 통계적 확률을 구하기에는 시행 횟수가 부족하므로 물리적인 성질을 바탕으로 추측한 모델링 확률이 더 객관적이라고 판단하였다. 참 확률을 추정하기 위해서 확률의 적합성을 묻는 문항에서는 대부분의 예비교사가 가장 객관적이라고 판단한 확률이 가장 적합한 확률이라고 판단하였다. 하지만 몇몇 예비교사는 빈도적 확률을 구하기에 부적절한 데이터를 제외한 집산으로 구한 빈도적 확률이 더 적합하다고 판단하였다. 이는 적당한 인간의 개입은 확률의 객관성을 감소시키더라도 참 확률을 추정하는데 더욱 적합하다고 판단한 것으로 볼 수 있다. 주관적 확률과 관련된 코로나 감염 확률에서는 개인이 주관적으로 추측한 모델링 확률에 코로나 진단 키트의 결과인 경험적인 정보가 조건부확률로 결합하여 갱신한 확률이 코로나 감염 확률을 추정하기에 더 객관적이므로 대다수의 예비교사가 더 적합한 확률이라고 판단하였다. 하지만 유병률을 통해서 구한 빈도적 확률은 객관적이지만 본인과 관련된 코로나 감염 요인을

반영하지 못하는 점을 들어 주관적 확률에 비해 적합하지 않다고 판단한 예비교사가 다수 존재하였다. 윷의 확률은 객관성이 확률값을 추정하는데 가장 중요한 요소였다면, 코로나 감염 확률과 같이 관련된 정보에 대한 다양한 의견이 존재하는 상황에서는 개인별로 중요하다고 판단되는 정보에 따라 참 확률 추정에서 다양한 양상을 보임을 확인할 수 있었다.

## 2. 논의 및 시사점

본 연구 결과를 바탕으로 예비교사가 가지고 있는 확률에 대한 관점 및 확률교육에 적용할 수 있는 방안에 대한 시사점을 제시하면 다음과 같다.

첫째, 본 연구는 수학교육에서 확률에 대한 관점을 분류하는 새로운 틀을 제시하고, 그 틀을 실제로 확률적 상황의 예에서 적용해 보았다는 점에서 의의가 있다. Ⅱ장에서 설명하였듯이 Chernoff(2008)는 수학교육에서 고전적, 빈도적, 주관적 확률로의 분류는 확률이론에서 사용되는 용어 및 개념의 부재로 빈도적 확률의 주관적인 측면인 성향빈도형 확률과 주관적 확률의 객관적인 측면인 대인신념형은 잘 다루어지지 않았었다고 지적하였다. 본 연구에서는 고전적 확률에 추가로 도수빈도형, 성향빈도형, 대인신념형, 개인신념형 확률로 분류하는 기준을 세우고 확률적 상황의 예에 적용해 보면서, 수학교육에서 확률에 대한 관점을 더 넓고 세심하게 다룰 수 있음을 확인하였다. 더 나아가 예비교사가 구성한 다양한 확률적 상황의 예를 분석하면서 예비교사가 지니고 있는 확률에 대한 관점을 정교하게 파악할 수 있었다. 본 연구에서 제시한 확률에 대한 분류기준을 활용하면 그동안 수학교육에서 크게 주목받지 못했던 확률에 대한 주관적인 특성을 학생들이 어떻게 개념화하는지 살펴볼 수 있을 것이다.

둘째, 이론적 확률이 존재하지 않는 확률적 상황에 대한 확률값을 찾는 과정을 수학적 모델 도출 과정의 단계로 분석하였다는 점에서 의의가 있다. 예비교사들은 등확률성을 만족하지 않는 상황에서 각자가 사전에

가지고 있던 정보를 활용하여 확률값을 추정해냈다. 확률을 추정해나가는 과정을 정혜윤 외(2018)의 수학적 모델 도출 과정의 세부 단계의 요소를 적용하여 분석할 수 있음을 확인하였다. 구체적으로 확률값에 영향을 미치는 사전 정보의 유형에 따라 확률에 대한 접근 방식을 분류할 수 있으며, 그다음 단계인 '단순화 하기', '수학적으로 다양하게 표현하기', '요소 사이 관계 찾기'가 확률값을 계산하는 실제적인 전략으로 사용됨을 확인하였다. 이는 정혜윤 외(2018)의 모델이 확률값을 추정하는 과정을 분석하는데 유용한 분석 도구가 될 수 있음을 보여준다. 한편, 몇몇 예비교사들은 다양한 정보가 영향을 미치는 확률적 상황에 대해서는 확률값을 추정하는 데 어려움을 겪는 모습을 드러내었다. 정혜윤 외(2018)의 모델이 확률값을 추정하는 단계를 분석하는데 유용한 도구로 쓰였으므로, 역으로 확률값을 추정하는 데 어려움을 겪는 학생들에게 정혜윤 외(2018)의 모형의 절차를 제시하고 각 접근 방식에 알맞은 정보를 제공하면, 여러 확률적 상황에서 확률값에 대한 추정을 수월하게 이끌 수 있을 것으로 예상된다.

셋째, 확률이론에서 사용되는 여러 정리들을 활용하여 확률에 대한 세 가지 접근 방식에 대한 관계를 분석할 수 있음을 확인하였다. Eichler & Vogel(2014)은 교사가 확률교육에서 가져야 할 지식 중 확률에 대한 세 가지 접근 방식 간의 연결에 대한 지식이 필수적이라고 보았다. 이에 확률이론에서의 각 정리들이 어떠한 접근 방식들과 서로 관련되는지 살펴보자. 첫째로 큰 수의 법칙은 고전적 확률과 빈도적 확률을 연결짓는 원리로 볼 수 있다. 둘째로 빈도 원리는 주관적 확률을 판단하는데 통계자료를 활용하여 빈도적 확률로 사건의 확률을 추정할 수 있는 원리로 주관적 확률과 빈도적 확률을 연결 지어 준다. 셋째로 베이즈 정리는 조건부확률을 활용해 개인신념형 확률이 정보의 양과 종류에 따라 대인신념형 확률 또는 도수빈도형 확률로 전환될 수 있다. 마지막으로 집산의 성질은 집산을 구성하고 있는 사건들을 확인함으로써 도수빈도형 확률 내에서 객관성의 정도를 비교할 수 있게 한다. 이렇듯 확률이론에서 쓰인 여러 정리가 확률에 대한 세 가지 접근 방식의 관계를 살펴보는 데

다양하게 활용될 수 있다. 따라서 확률에 대한 세 가지 접근 방식과 관련된 교사 지식을 갖추기 위해서는 교사들은 철학에서 확률이 역사적으로 어떻게 개념이 정교화되어 왔는지 살펴봐야 하고, 확률이론에서의 여러 정리들을 이해하는 것이 중요하다는 사실을 알 수 있었다.

마지막으로 본 연구의 제한점을 제시하고, 후속 연구를 제언하고자 한다. 본 연구에서는 확률이론에서 다루는 신념형 확률의 세부 분류인 개인신념형과 대인신념형 확률, 빈도형 확률의 세부 분류인 성향빈도형과 도수빈도형 확률에 대한 개념을 도입함으로써 기존의 수학교육에서 다루어졌던 빈도적 확률과 주관적 확률의 범위가 더욱 확장되고 세심하게 다뤄질 수 있음을 확인하였다. 하지만 고전적 확률이 빈도적 확률 및 주관적 확률과 중복되어 해석되는 지점이 존재하는 문제를 해결하기 위해서 본 연구에서는 고전적 확률을 등확률성을 지니는 상황만으로 한정해서 다루었다. 이는 세 가지 관점의 확률에 대한 명확한 분류기준을 제시한다는 장점이 있으나, 고전적 확률이 다룰 수 있는 상황을 극히 제한한다는 점을 확인하였다. 이에 후속 연구를 수행하여 세 확률의 관계를 더욱 잘 나타내는 모델을 탐구하거나, 더욱 다양한 확률적 상황의 예를 분류해보면서 세 확률의 관계에 대하여 세심하게 논의할 필요가 있다.

등확률성을 만족하지 않으면서 우리에게 친숙한 상황인 윷을 던지는 상황과 코로나 감염에 관련된 상황을 각각 빈도적 확률과 주관적 확률을 대표하여 확률을 추정하는 활동을 하였다. 확률의 객관적인 특성을 높이는 방식 말고도 참 확률을 추정하는 여러 모델링 상황을 고안할 필요가 있다. 또한 윷을 던지는 상황과 코로나 감염과 같은 구체적인 상황에서 확률을 추정하는 과정을 살펴보았는데, 모델링의 특성상 개인이 사전에 알고 있는 정보에 따라 다른 양상을 띨 수 있으므로 이 결과만 가지고 일반화하기에는 무리가 있다. 수학을 다루는 데 익숙한 예비교사들을 대상으로 실시한 연구 결과로써 학생들에게 동일한 활동을 실시하더라도 전혀 다른 결과가 나타날 수도 있다. 참 확률을 찾을 수 있는 명확한 방법이 존재하지 않으므로 학생들에게 다룰 때는 모델링 상황을 더욱 정교화해서 다룰 필요가 있다.

주관적 확률이 가지는 장점에도 불구하고 수업과 교과서에서 다루기 어려운 점이 다수 존재하기 때문에 주관적 확률은 학교수학에서 외면받아 왔다. 데이터의 종류가 다양해지고 양이 많아질수록 주관적 확률을 다뤄야 하는 사회적 필요성을 점점 더 무시하기 어려워질 것이다. 학교수학에서 주관적 확률을 체계적으로 다룰 수 있는 방법에 대한 후속연구가 필요하다.

# 참 고 문 헌


교육부(2015). **수학과 교육과정**. 교육부 고시 제2015-74호 [별책8].

교육부(2021). **2022 개정 교육과정 총론 주요사항(시안)**. 세종: 교육부

권오남, 신준국, 전인태, 김미주, 김철호, 김태홍, 박재희, 박정숙, 박지현, 박찬호, 박효근, 오국환, 조경희, 조상현, & 황성문(2019). **고등학교 확률과 통계**. 서울: 교학사

송하석(2005). 확률의 상관 빈도이론과 포퍼. **논리연구, 8**(1), 23-46.

송하석(2006). 포퍼의 확률의 성향 이론. **논리연구, 9**(1), 31-62.

신보미, & 이경화(2006). 컴퓨터 시뮬레이션을 통한 통계적 확률 지도에 대한 연구. **수학교육학연구, 16**(2), 139-156

신보미, & 이경화(2008). 시뮬레이션을 활용한 확률 지식의 교수학적 변환. **수학교육학연구, 18**(1), 25-50.

우정호(2017). **학교수학의 교육적 기초** (하)(개정판). 서울: 서울대학교 출판문화원.

이경화(1996). **확률 개념의 교수학적 변환에 관한 연구**. 서울대학교 대학원 박사학위 논문.

이경화, 유연주, & 탁병주(2022). 데이터 기반 확률교육을 위한 수학과 교육과정 재구조화 방향 탐색. **학교수학, 24**(1), 89-117.

정혜윤, 이경화, 백도현, 정진호, & 임경석(2018). 수학적 모델링 관점에 의한 <수학과제 탐구> 과목용 과제의 설계. **학교수학, 20**(1), 149-169.

통계청(2022. 12. 6.). 2021년 생명표 작성 결과 [보도자료]. https://kostat.go.kr/portal/korea/kor_nw/1/1/index.board?bmode=read&aSeq=422107

Ahlgren, A., & Garfield, J. (1991). Analysis of the probability curriculum. In *Chance encounters: probability in education* (pp. 107-134). Springer, Dordrecht.

Albert, J. H. (2003). College students' conceptions of

probability. *The American Statistician, 57*(1), 37-45.

Batanero, C., & Borovcnik, M. (2016). *Statistics and probability in high school*. Springer.

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). *Research on teaching and learning probability*. Springer Nature.

Batanero, C., & Díaz, C. (2007). The meaning and understanding of mathematics. In *Philosophical dimensions in mathematics education* (pp. 107-127). Springer, Boston, MA.

Batanero, C., Henry, M., & Parzysz, B. (2017). 우연(chance)과 확률의 본성. In Jones, G. A. (Ed.). **확률, 왜 그리고 어떻게 가르칠까?** (신보미, 이경화 역) (pp. 3-26). 서울: 경문사. (영어 원작은 2005년 출판).

Borovcnik, M., Bentz, H. J., & Kapadia, R. (1991). A probabilistic perspective. In *Chance encounters: Probability in education* (pp. 27-71). Springer, Dordrecht.

Borovcnik, M., & Kapadia, R. (2014). A historical and philosophical perspective on probability. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking: Presenting plural perspectives* (pp. 7-34). Dordrecht: Springer.

Chaput, B., Girard, J. C., & Henry, M. (2011). Frequentist approach: Modelling and simulation in statistics and probability teaching. In *Teaching Statistics in school mathematics-Challenges for teaching and teacher education*(pp. 85-95). Springer, Dordrecht.

Chavoshi Jolfaee, S. S. (2015). *Objective and subjective probability: Undergraduate students' descriptions, examples, and arguments* (Doctoral dissertation, Education: Education).

Chernoff, E. J. (2008). The state of probability measurement in mathematics education: A first approximation. *Philosophy of Mathematics Education Journal*, *23*, 1-23.

Chernoff, E. J., & Russell, G. L. (2014). Preface to perspective I: Mathematics and philosophy. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking: Presenting plural perspectives* (pp. 3-5). Dordrecht: Springer.

De Finetti, B. (1974). *Theory of Probability : A Critical Introductory Treatment / Bruno De Finetti ; Traslated by Antonio Machi and Adrian Smith.*

Eichler, A., & Vogel, M. (2014). Three approaches for modelling situations with randomness. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking: Presenting plural perspectives* (pp. 75-99). Dordrecht: Springer.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review, 70*(1), 1-25.

Gal, I. (2017). 대중을 위한 "확률적 소양"에 대하여: 구성요소와 교육적 딜레마. In Jones, G. A. (Ed.). **확률, 왜 그리고 어떻게 가르칠까?** (신보미, 이경화 역) (pp. 27-50). 서울: 경문사. (영어 원작은 2005년 출판).

Gillies, D. (2000). *Philosophical theories of probability.* New York: Routledge.

Greer, B., & Mukhopadhyay, S. (2017). 불확실성의 수학화에 대한 교수-학습: 역사적·문화적·사회적·정치적 맥락에서. In Jones, G. A. (Ed.). **확률, 왜 그리고 어떻게 가르칠까?** (신보미, 이경화 역) (pp. 281-309). 서울: 경문사. (영어 원작은 2005년 출판).

Hacking, I. (1975). *The Emergence of Probability : A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference. Cambridge,* MA: Cambridge University

Press.

Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge: Cambridge University Press.

Hoaglund, A., Birkenfeld, K., & Box, J. (2014). Professional learning communities: Creating a foundation for collaboration skills in pre-service teachers. *Education, 134*(4), 521-528.

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. *Second handbook of research on mathematics teaching and learning, 2,* 909-955.

Kahneman, D., & Tversky, A. (1982). Subjective probability: A judgment of representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* Cambridge, New York: Cambridge University Press.

Kazak, S., & Leavy, A. M. (2018). Emergent reasoning about uncertainty in primary school children with a focus on subjective probability. In *Statistics in early childhood and primary education*(pp. 37-54). Singapore: Springer.

Kvatinsky, T., & Even, R. (2002). Framework for teacher knowledge and understanding about probability. In *Proceedings of the Sixth International Conference on Teaching Statistics (CD). Cape Town, South Africa: International Statistical Institute.*

Pfannkuch, M., & Ziedins, I. (2014). A modelling perspective on probability. In *Probabilistic Thinking*(pp. 101-116). Springer, Dordrecht.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions.

Stohl, H. (2017). 교사 교육과 전문성 개발에서의 확률. In Jones, G. A. (Ed.). **확률, 왜 그리고 어떻게 가르칠까?** (신보미, 이경화 역) (pp. 331-352). 서울: 경문사. (영어 원작은 2005년 출판).

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1-8.

Watson, A., & Mason, J. (2015). **색다른 학교수학**. (이경화 역). 서울: 경문사. (영어 원작은 2005년 출판).

# <부록>

## <1차 설문지>

학번 :

1-1 (a) 확률적 상황의 예를 서술하고 (b) 그에 해당하는 확률값을 적어주세요. 또한 (c) 확률값이 나오게 된 이유를 간단히 설명해주세요.(가능한 한 다양한 상황을 제시해주시되 최대 5개까지 작성해주세요.)

예시 ①-a : 동전을 한 번 던졌을 때 앞면이 나올 확률
-b : 0.5
-c : 앞면 또는 뒷면이 가능한데 2가지 경우 중 앞면은 1가지 경우만 해당되므로

1-2 본인이 생각하는 확률의 뜻을 서술해주세요.(한 개 이상 답변해도 됩니다.)

2. 다음의 확률 상황의 예시를 읽고 확률값이 0인지, 거의 0에 가까운지, 거의 1에 가까운지, 1인지 해당되는 칸에 체크하고, 간단한 이유를 제시해주세요.

| 순 | 확률 상황 | 0 | 거의 0 | 거의 1 | 1 |
|---|---|---|---|---|---|
| 1 | 내가 보통의 백 원짜리 동전을 100번 던졌을 때, 100번 모두 동전의 앞면이 나올 확률 | | | X | X |
| 이유 | | | | | |
| 2 | 내가 로또 1등에 당첨될 확률 | | | X | X |
| 이유 | | | | | |
| 3 | 내가 길을 가다가 벼락에 맞을 확률 | | | X | X |
| 이유 | | | | | |
| 4 | 백조가 흰 색일 확률 | X | X | | |
| 이유 | | | | | |
| 5 | 내일 아침에 해가 뜰 확률 | X | X | | |
| 이유 | | | | | |

3. 평상시에 알고 있던 정보를 바탕으로 각 상황에 대한 확률을 제시하고, 확률값이 나오게 된 과정을 간단히 설명해주세요.

3-1. 윷가락을 한 개 던졌을 때, 뒤집어진 면( )이 나올 확률을 추측해보세요.

∗확률 :

∗이유 :

3-2. 가족, 친구 등 주변의 사례 혹은 인터넷, 뉴스, 신문 등의 매체로부터 접해온 코로나에 대한 정보를 바탕으로, 당신이 현재 코로나에 감염될 확률을 추측해보세요.

∗확률 :

∗이유 :

<2차 설문지>

확률의 3가지 접근 방식(개념 설명)

0-1. 고전적 확률(Classical a Priori Theory) : Laplace에 의해 최초로 명시적으로 정의된 확률

•$P(A) = \dfrac{\text{사건 } A\text{가 일어날 횟수}}{\text{일어날 수 있는 모든 사건의 횟수}}$

•특징 : 불충분한 이유의 원리를 제시하며 각각의 사건에 등확률을 적용해야 한다고 주장

*불충분한 이유의 원리 : 한 사건이 다른 사건보다 더 많이 일어난다는 기대할 만한 어떤 이유가 없는 경우에는 가능한 모든 사건에 동일한 확률을 할당해야 한다는 원칙

•한계 : 간단한 상황(ex. 주사위 던지기)을 제외하면 실제 경험에서 고전적 확률을 적용할 수 없는 상황이 대다수

0-2. 빈도적 확률(Frequentist Theory) : 개별 사건을 집산(collective)에 포함시킴으로써 불확실성을 측정할 수 있으며 상대도수가 향하는 극한을 확률로 봄(von Mises)

※ 집산 : 특정한 '임의성'을 가진 '유사한' 사건들의 무한한 사건의 연쇄

•한계 : 동일한 사건이라도 서로 다른 집산에 포함될 수 있음. 확률을 반복적으로 수행할 수 없는 사건이 다수 존재, 빈도적 확률이 실제하는지, 계산된 값이 참 값인지 알 수 없음.

•특징 : 등확률을 적용할 수 없는 상황에서 확률을 계산하는 실용적인 수단을 제공

0-3 주관적 확률(Subjectivist Theory) : 주어진 결과에 대해 개인이 갖고 있는 정보에 의해 결정되는 개인적인 신념을 확률로 봄(de Finetti), 개인마다 확률이 다를 수 있지만 확률의 공리에 위반되지 않는 일관성만 유지한다면 모두 허용됨.

•한계 : 사전 확률을 측정하는 방법에 대한 지침이 없음

•특징 : 1. 고전적(등확률성의 결여), 빈도적(반복 불가능한 사건) 확률을 적용하지 못하는 상황에서 적용가능

2. 사전 정보(개인의 신념) & 경험적 정보(반복된 실험에서 얻어진 데이터)가 베이즈 공식에 의해 결합되어 새로운 확률을 제시

모델링 상황 1 : 윷가락이 뒤집힐 확률은?

1-1. 지난 설문에서 응답한 윷가락이 뒤집힐 확률은 (_____)입니다. 이제는 실제로 윷가락을 던져서 확률을 계산해보고자 합니다. 윷가락을 한 번 던질 때마다 나온 결과를 뒤집어진 면( )은 1, 반대면은 0이라고 기록합니다. 던져서 나온 데이터를 살펴보면서 윷가락을 빈도적 확률을 계산하기에 이 정도 횟수면 충분하다고 생각하는 최소 횟수(10회 이상 50회 이하)만큼 던집니다. 그 결과를 구글 설문지에 응답하세요. (응답 예시: 12번, 011001111011)

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |

∗수열
∗횟수:
∗던진 횟수를 결정하게 된 이유:

1-2. 당신과 친구들이 윷가락을 던진 횟수와 각각 구한 확률에 데이터가 여기 있습니다. 윷가락이 뒤집힐 확률을 구하기에 같은 집산에 포함되기에 부적절하다고 판단되는 사람의 데이터가 있으면, 스프레드시트의 필터 기능을 이용하여 그 사람의 데이터를 제외하고,
그 이유를 서술하세요.
□ 제외하는 데이터가 없다. □ 제외하는 데이터가 있다 → 데이터의 개수는?
∗제외를 하(지않)는 이유 :

1-3. 1-2에서 필터링한(제외한) 데이터들을 모두 종합하여 빈도적 확률을 계산하고자 합니다. 각각의 데이터에서 윷가락을 던진 횟수가 사람마다 각자 다른 상황입니다. 각각의 데이터에서 윷가락을 던진 횟수가 다른 상태를 보존할 경우 스프레드시트 빈 칸에 0을 입력하고, 던진 횟수를 모두 같은 횟수로 일치시키길 희망할 경우 스프레드시트에 본인이 생각하는 통일된 던진 횟수를 입력하세요.
□ 던진 횟수를 그대로 둔다 → 숫자 0 입력
□ 일치시켜야 한다 → 통일해야 되는 횟수 : _____ 입력
∗던진 횟수를 일치시킨(시키지 않은) 이유 & 이 숫자를 선택한 이유 :

1-4. 각각에 해당되는 윷가락이 뒤집힐 확률을 기록하세요.
① 지난 설문에서 응답했던 주관적 확률: _____
② 1-1에서 나온 본인의 데이터로 구한 빈도적 확률: _____
③ 1-3에서 나온 전체 윷놀이 데이터 중에서 본인의 판단 하에 필터링한 빈도적 확률: _____
④ 본인과 친구들의 데이터를 필터링하지 않고 모두 다 합쳐서 구한 빈도적 확률 : _____

※ [1-5, 1-6] 응답자에 따라 ③, ④의 확률값이 같을 수도 있지만, 1-5, 1-6에서 응답할 때,
③은 일부 데이터가 같은 집산에 포함하지 않다고 판단되어 한 개 이상의 데이터가 제외되어 구한 빈도적 확률로 보고, ④는 모든 사람의 데이터가 같은 집산에 속하므로 모든 데이터를 포함하여 구한 빈도적 확률로 구분하여 응답해주세요.

1-5. 각각의 확률 ①, ②, ③, ④에서 가장 객관적이라고 생각하는 확률의 순위를 매기고, 그렇게 판단한 이유를 서술해주세요.

∗순위 :

| 1위 | | 2위 | | 3위 | | 4위 | |
|---|---|---|---|---|---|---|---|

∗이유 :

1-6. 각각의 확률 ①, ②, ③, ④에서 윷가락이 뒤집힐 확률에 가장 적합하다고 생각하는 확률의 순위를 매기고, 그렇게 판단한 이유를 서술해주세요.

∗순위 :

| 1위 | | 2위 | | 3위 | | 4위 | |
|---|---|---|---|---|---|---|---|

∗이유 :

1-7. 윷놀이 활동에서 학생들에게 확률과 관련하여 무엇을 가르칠 수 있을까요? 최대한 자세히 서술해주세요.

∗

모델링 상황 2 : 코로나에 감염됐을 확률은?

| | 코로나 감염($A$) | 코로나 비감염($A^C$) | 계 |
|---|---|---|---|
| 신속항원검사 양성($B_1$) | | | |
| 신속항원검사 음성($B_1^C$) | | | |
| 계 | | | 100,000명 |

<표1. 신속항원검사와 코로나 감염 분할표>

2-1. 지난 설문의 3-2에서 응답한 코로나에 감염됐을 확률(　　)을 이용하여 다음의 분할표에서 코로나에 감염된 사람 수를 빨간색으로 채워보세요.

2-2. 질병관리청에서 신속항원검사법에 대한 민감도(90%)와 특이도(96%)에 대한 수치를 제공하였습니다. 민감도(Sensitivity)는 코로나19에 걸린 사람을 검사했을 때, 결과가 양성으로 나타날 확률입니다. 민감도를 확률 기호를 이용하여 나타내보고, 해당되는 표의 칸에 알맞은 인원수를 파란색으로 채워보세요.

∗기호 :

2-3. 특이도(Specifity)는 코로나19에 걸리지 않은 사람을 검사했을 때, 결과가 음성으로 나타날 확률입니다. 특이도를 확률 기호를 이용하여 나타내보고, 해당되는 표의 칸에 알맞은 인원수를 녹색으로 채워보세요.

∗기호 :

2-4. 합계를 이용하여 나머지 표의 빈칸을 채워보고, 신속항원검사에서 양성으로 나왔을 때 실제 코로나에 감염됐을 확률을 구하고, 확률 기호를 이용하여 나타내보세요.

∗확률 :

∗기호 :

2-5. 신속항원검사에서 양성으로 나왔으므로, 질병관리청에서는 비인두도말 PCR 검사를 받으라고 권고하였습니다. 비인두도말 PCR 검사의 민감도와 특이도는 각각 98%, 100%입니다. <표1>의 데이터와 민감도, 특이도를 이용하여 <표2>를 채워보세요.

| | 신속항원검사 양성이면서 코로나 감염($A \cap B_1$) | 신속항원검사 양성이면서 코로나 비감염($A^C \cap B_1$) | 계 |
|---|---|---|---|
| PCR 양성($B_2$) | | | |
| PCR 음성($B_2^C$) | | | |
| 계 | | | |

<표2. PCR 검사와 코로나 감염 분할표>

2-6. 신속항원검사와 달리 PCR 검사에서는 음성으로 나왔을 때 코로나에 감염됐을 확률을 구하고, 확률 기호를 이용하여 나타내보세요.

∗확률 :

∗기호 :

2-7. 질문 2-1, 2-4, 2-6에서 나온 확률은 어떠한 차이가 있을까요? 이 활동에서 학생들에게 무엇을 가르칠 수 있을까요?

∗

2-8. 당신은 정부에서 공개한 8월 3일자 인구자료와 연령별, 지역별 코로나 누적확진자 수 데이터의 일부를 접하게 되었습니다. 코로나가 얼마나 퍼져있는지 알려주는 지표인 코로나의 유병률을 통해 코로나의 감염 확률을 판단하고자 합니다. 코로나 유병률은 인구집단을 설정하는 방식에 따라 달라집니다. 당신이 코로나에 감염됐는지 판단하기에 가장 적합다고 생각하는 코로나 유병률을 구해보세요.
(이번 설문에서 적용하는 코로나 유병률은 코로나 발생시점부터 8월 3일까지의 특정 지역의 인구집단 중에서 코로나에 확진된 인구수의 비율입니다.)
<인구현황>

| | 남자 인구수 | 여자 인구수 | 총 인구수 |
|---|---|---|---|
| 전국 | 25,708,650 | 25,865,796 | 51,574,446 |
| 서울 | 4,604,405 | 4,888,806 | 9,493211 |

<20대 인구현황>

| | 20대 남자 인구수 | 20대 여자 인구수 | 20대 총 인구수 |
|---|---|---|---|
| 전국 | 3,423,973 | 3,107,663 | 6,531,636 |
| 서울 | 676,526 | 734,916 | 1,411,442 |

<확진자수 데이터(8월 3일 00시 기준)>

| |
|---|
| 누적 확진자수 : 20,052,305명<br>서울시 누적 확진자수 : 4,032,709명 (20.11%)<br>전국 20대 누적 확진자수 : 2,987,809명 (14.90%)<br>전국 누적확진자 성별 비율 : 남성 9,414,916(46.95%), 여성 10,637,389명(53.05%) |

*유병률 : _____
*구하게 된 과정 :

2-9. 2-1에서 사용했던 코로나에 감염될 확률(전자)과, 2-8에서 유병률을 통해 판단한 코로나에 감염된 확률(후자)을 두 확률의 특성을 각각 서술하고 유사점과 차이점을 서술해주세요.
*

3-1 이번 확률론 강좌 이전에 주관적 확률에 관련된 내용을 접해본 적이 있습니까?
□ 예 □ 아니오

3-2. 다음은 지금까지 응답했던 설문 문항에 대한 간략한 설명입니다.

*1차 설문지 : ① 확률적 상황 예로 들기(1-1), ② 확률의 뜻 서술하기(1-2),
③ 일상생활에서 거의 일어나지 않거나 대부분 일어나는 상황에 대한 확률 판단(2),
④ 윷이 뒤집히는 상황 및 코로나에 감염될 확률에 대한 주관적 확률 할당(3)
*수업 내용 : ⑤ 확률에 대한 접근 방식 3가지
*모델링 상황 1 : ⑥ 윷가락이 뒤집힐 확률은?
*모델링 상황 2 : ⑦ 코로나에 감염됐을 확률은?

다음은 지난 번 1차 설문 문항 2번과 3번의 결과를 요약한 표 및 줄기와 잎그림입니다. (응답의 예시, 실제 설문 조사 후 결과 반영하여 제시할 예정입니다.)

| 순 | 확률 상황 | 0 | 거의 0 | 거의 1 | 1 |
|---|---|---|---|---|---|
| 1 | 내가 보통의 백 원짜리 동전을 100번 던졌을 때, 100번 모두 동전의 앞면이 나올 확률 | 0명 (0%) | 24명 (100%) | | |
| 2 | 내가 로또 1등에 당첨될 확률 | 5명 (20.8%) | 19명 (79.2%) | | |
| 3 | 내가 길을 가다가 벼락에 맞을 확률 | 1명 (4.2%) | 23명 (95.8%) | | |
| 4 | 백조가 흰 색일 확률 | | | 18명 (75%) | 6명 (25%) |
| 5 | 내일 아침에 해가 뜰 확률 | | | 11명 (45.8%) | 13명 (54.2%) |

(2|5는 0.25)

| 줄기 | 잎 |
|---|---|
| 2 | 5 |
| 3 | |
| 4 | 0 1 5 |
| 5 | 0 0 3 |
| 6 | 0 0 0 0 0 1 1 1 1 7 |
| 7 | 0 5 |

| 0.5미만 | 0.5 | 0.5초과 |
|---|---|---|
| 5 | 2 | 15 |

윷이 뒤집힐 확률

1) 코로나 감염 확률을 크기 순으로 나열(15개)

| | | | | |
|---|---|---|---|---|
| 0 | 0 | 0.000001 | 0.001 | 0.001 |
| 0.0056 | 0.01 | 0.01 | 0.01 | 0.02 |
| 0.05 | 0.2 | 0.5 | 1 | 1 |

2) 숫자 이외의 값을 답변한 인원 수(9명)

| 응답 | 인원 |
|---|---|
| 거의 0 | 3명 |
| 거의 1 | 2명 |
| 알 수 없다 | 2명 |
| 기타 | 2명 |

코로나 감염될 확률

위의 설문 결과와 지금까지 해왔던 활동을 바탕으로 주관적 확률의 장단점을 1,2차 설문에서 했던 활동 ①~⑦과 관련지어 서술해주세요.

3-3. ①~⑦의 활동을 하고 난 뒤 확률에 대한 생각의 변화가 있습니까?

□ 예 □ 아니오

있으면 변화에 대하여 간략히 서술하고, 없으면 기존의 확률 개념이 어떠했는지 1,2차 설문에서 했던 활동 ①~⑦과 관련지어 간략히 서술해주세요.(빈도적 확률과 주관적 확률과 관련된 확률의 개념의 변화를 서술해 주세요.)

*

3-4. 중학생 또는 고등학생들에게 주관적 확률과 관련된 내용을 가르쳐야 한다고 생각합니까?

□ 중&고등학생 모두 가르쳐야 한다.

□ 중학생만 가르쳐야 한다.

□ 고등학생만 가르쳐야 한다.

□ 중&고등학생 모두 가르치지 말아야 한다.

위와 같이 응답한 이유를 서술해주세요.

*

## Abstract

# Analysis of Pre-Service Teachers' Approach to Probabilistic Situations

## : Focusing on Non-Equiprobable Situations



Lee, Jonghwa

Department of Mathematics Education

The Graduate School

Seoul National University



Probability is widely used in real-life situations to deal with uncertainties. As a result, the subject of probability has become a crucial topic in school mathematics education.

Throughout history, probability has been interpreted in various ways. Probability is largely divided into three perspectives in mathematics education: classical probability, frequentist probability, and subjective probability. The goal of probability education is to equip students with the skills to predict the future and make rational decisions through probability literacy. However, school mathematics primarily focuses on the classical perspective of probability, resulting in a limited understanding of the true nature of probability among students. This is largely due to the ambiguity surrounding the subject, which leads mathematics teachers to also prefer teaching classical probability. It is crucial to study the perspectives of

teachers on probability, as the quality of the education they provide has a significant impact on students' success in grasping the concept. By examining the perspectives of pre-service teachers on probability, it is possible to assess both the teachers' understanding of probability and the effectiveness of probability education in the regular curriculum. The study synthesized the perspectives on probability in mathematics education and probability theory and developed a framework for analyzing the approach to probability.

In this study, the perspectives of pre-service teachers on probability were assessed through activities involving the construction of examples and presentation of the meaning of probability. The examination included the process of estimating probabilities in non-equiprobable situations, as well as the understanding of frequentist and subjective probability. Additionally, the study evaluated the ability of pre-service teachers to infer the true probability in a given probabilistic situation when presented with additional empirical information.

From the results of the study, the following points were found. First, most of the pre-service teachers constructed examples of classical probability and there were very few examples of personal belief-type probability that varied from person to person. Second, in the activity of estimating the probability value in non-equiprobable situations, it was observed that different approaches were used depending on the individual's prior information. Third, in the activity of updating the probability value by providing additional empirical data in a probabilistic situation, there was a tendency to believe that a more objective probability was closer to the true probability in the frequentist probability approach. However, in subjective probability, the true probability was not judged to be more accurate based on its objective nature, but rather estimated by considering the factors that individuals deemed relevant to the probabilistic situation.

Based on the results of this study, the following suggestions for probability education can be made. First, the new framework for classifying the perspective on probability in mathematics education can be applied. Second, the process of deriving a mathematical model can be used as a useful framework for analyzing the process of estimating the true probability value. Third, it was confirmed that the relationship between the three approaches to probability can be analyzed through the use of various theorems in probability theory.

**keywords : classical probability, frequentist probability, subjective probability, pre-service teacher, probability education, perspective on probability**

***Student Number*** **: 2021-23023**