

## 베이즈주의와 IBE<sup>†</sup>

여 영 서<sup>‡</sup>

반 프라센은 베이즈주의와 최선의 설명에로의 추론 즉 IBE의 양립불가능성을 논증하여 IBE가 틀렸음을 주장한다. 하지만 대부분의 IBE 옹호자들과 베이즈주의자들은 두 이론 사이의 양립가능성을 모색해 왔다. 최근에 헨더슨은 베이즈주의가 IBE와 양립가능한 이론이 되기 위해 특별히 할 일이 하나도 없다는 입장에서 두 이론의 양립가능성을 검토하고 있다. 본 논문은 헨더슨의 시도가 베이즈주의와 IBE가 근본적으로 동일한 인식적 직관을 공유한다는 지적에 대해서는 동의하지만 베이즈주의에서 IBE가 창발된다는 주장에는 동의하지 않는다. 그 대신 본 논문은 IBE와 베이즈주의가 서로 다른 수준에서 과학적 추론 방식을 제시하고 있는 상호 호혜적인 관계에 있다고 결론짓는다.

【주요어】 베이즈주의, IBE, 헨더슨, 반 프라센

<sup>†</sup> 이 논문은 2014년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2014S1A5A2A01016597). 이 논문은 2014년 1월의 한국논리학회에서 발표한 바 있다. 당시 이 논문에 대해 꼼꼼한 논평을 해 주신 전영삼 선생님과 좋은 질문으로 토론에 참여해 주신 논리학회 회원들께 감사드린다. 이 논문의 심사위원들이 주신 심사평 역시 논문의 입장을 정교하고 명확하게 마무리하는 데에 큰 도움이 되었음을 밝히며 심사위원들께 감사드린다.

<sup>‡</sup> 동덕여자대학교 교양교직학부, yyeo@dongduk.ac.kr.

[A]ll valid rules of scientific inference must be derived from the probability axioms and the rule of conditionalization. It follows that there is nothing left for the methodologists to do in this area. [...] I agree with Feyerabend that there is no Methodology. But my reasons do not stem from an ideology of anarchism or Dadaism; nor do they rely on incommensurability and fellow travelers. A little Bayesianism and a lot of calm reflection are all that is needed.

J. Earman (*Bayes or Bust?* pp. 204-5)

## 1. 들어가기

베이즈주의와 최선의 설명에로의 추론 즉 IBE는 과학적 추론에 관한 서로 다른 이론이다. 베이즈주의는 가설에 대한 믿음의 정도를 확률로 나타내고, 새로 확보한 증거에 따라 그 값을 수정해 나가는 이론이다. 반면 IBE는 주어진 증거에 대해 가장 좋은 설명을 제시하는 가설 즉 최선의 설명을 제시하는 가설을 채택하라는 이론이다. 설명력을 추론의 안내자로 삼는 이론이라고 제시되기도 하는 IBE는 베이즈주의와 달리 확률을 전혀 언급하지 않는다.

베이즈주의와 IBE가 이렇게 서로 다른 이론임에도 불구하고, 반 프라센은 그의 책 *Laws and Symmetry*에서 두 이론의 양립불가능성을 논증하여 IBE가 틀렸음을 주장한다. 반 프라센의 이 논증 이후 베이즈주의와 IBE의 양립가능성 문제는 최근까지도 과학적 추론의 주요 문제로 다뤄지고 있다.

반 프라센의 논증은 IBE가 다음과 같은 믿음 수정 규칙을 제시한다는 가정 아래 전개된다. 즉 새로운 증거가 확보되면 그 새로운 증거까지 포함하여 주어진 모든 증거에 대한 최선의 설명을 제시하는 가설을 채택하라는 규칙이다. 반 프라센은 그 규칙이 베이즈주의가 제시하는 조건화 규칙과 충돌한다는 논증, 즉 두 이론 사이의 양립불가능성 논증으로부터 IBE가 틀렸다는 결론을 이끌어낸다. 하지만 대부분의 IBE

옹호자들과 베이즈주의자들은 반 프라센의 논증을 반박하며 두 이론 사이의 양립가능성을 모색한다. 대표적인 제안을 살펴보자면, 오카사 (Okasha, 2000)는 반 프라센이 IBE를 잘못 제시했다는 입장에서, 립톤 (Lipton, 2004)은 IBE와 베이즈주의가 서로 보완하는 관계라는 입장에서, 그리고 최근에 헨더슨(Henderson, 2013)은 베이즈주의가 IBE와 양립가능한 이론이 되기 위해 특별히 할 일이 하나도 없다는 입장에서 두 이론의 양립가능성을 검토하고 있다.

본 논문은 베이즈주의와 IBE가 근본적으로 동일한 인식적 직관을 서로 다른 수준에서 제시한다는 입장을 내세울 것이다. 이러한 본 논문의 입장은 베이즈주의와 제거적 귀납법 사이의 관계에 대한 졸고 (2004)의 연장선상에 있는 것인데, 여기서는 베이즈주의와 IBE의 양립 가능성에 대한 기존의 논의, 특히 헨더슨의 ‘창발적 양립가능론’(emergent compatibilism)을 비판적으로 검토하고자 한다. 헨더슨은 창발적 양립가능론을 제시하며 베이즈주의가 통계역학이라면 IBE가 열역학이라는 비유를 통해 베이즈주의가 IBE로부터 창발된다는 주장을 펼친다. 헨더슨은 자신의 비유가 베이즈주의가 미시적 이론이라면 IBE가 거시적 이론이라는 점을 보여준다고 한다. 이를 통해 헨더슨이 주장하려는 것이 궁극적으로 베이즈주의와 IBE가 근본적으로 동일한 인식적 직관을 서로 다른 수준에서 제시하는 것이라는 점이라면 본 논문은 헨더슨의 입장에 동의할 수 있다. 하지만 헨더슨이 주장하려는 바가 베이즈주의와 IBE 중 어느 하나가 다른 하나를 창발하는 관계, 그 중에서도 베이즈주의가 IBE로부터 창발하는 관계라는 것이라면 그와 같은 주장에는 동의하지 않는다. 본 논문은 IBE가 상대적으로 불명료한 이론이라는 점, 특히 IBE에서 말하는 설명력의 구체적인 내용이 불명료하다는 점을 지적하되, 베이즈주의보다 과학방법론에 대한 일상적 이해를 훨씬 더 용이하게 한다는 점을 인정한다. 이로써 본 논문은 헨더슨의 주장과 달리 두 이론 사이에 창발 관계를 따질 수 없지만, 헨더슨이 주장하듯이 두 이론이 서로 호혜적인 관계에 놓일 수 있음을 밝힌다.

베이즈주의와 IBE의 양립가능성에 대한 이상과 같은 논의는 두 이

론 모두에 대한 더 정확한 이해를 가능하게 한다. 이는 궁극적으로 과학방법론 논의에서 베이즈주의에 대한 이해를 심화하고 그 적용 영역을 넓히는 것은 과학적 방법에 대한 대부분의 핵심 논제를 포괄할 수 있다는 이어만의 입장<sup>1)</sup>을 뒷받침할 것으로 기대한다.

## 2. 베이즈주의와 IBE의 차이점

베이즈주의와 IBE의 양립 방안을 논의하기 이전에 두 이론 사이의 차이점을 확인하는 것은 불필요한 논란을 피하는 차원에서 유의미하다.<sup>2)</sup>

베이즈주의의 핵심은 믿음의 정도를 확률로 나타낼 수 있고, 합리적

1) 이어만은 그의 책 *Bayes or Bust?* (1992) 서문에서 자신은 월, 수, 금요일에만 베이즈주의자이고, 화, 목, 토요일에는 베이즈주의에 대해 회의적인 입장이라고 쓰고 있지만, 본 논문의 첫 인용문에서 보듯이 과학방법론의 논의는 베이즈주의로 종결됐다면서 베이즈주의를 강력히 옹호한다.

2) 사실 베이즈주의에 대한 공격은 전 방위에서 이루어진다. 반면 베이즈주의 자들 다수는 그러한 공격에 대응하기보다 실제 베이즈주의가 실제 기여할 수 있는 바가 무엇인가의 생산적 논의에 초점을 맞추고 있다. 게다가 굳이 베이즈주의를 46,656가지로 구분하고 다양한 종류의 베이즈주의가 가능하다는 점을 밝히자 일부 비판가들은 그만큼 베이즈주의가 불명확한 입장이며 끊임없이 난점을 회피해가는 이론이라고 비난한다. 이런 베이즈주의의 난점을 지적하면 다른 베이즈주의를 제시하며 그 비판을 피하고 또 저런 난점을 제시하면 또 다른 베이즈주의를 제시하며 베이즈주의자들은 문제 회피만을 일삼는다는 것이다. 사실 그런 측면이 없지 않겠지만 베이즈주의를 비판하는 쪽에서도 마찬가지라고 할 수 있다. 이런 난점에 대답하더라도 또 다른 베이즈주의의 난점을 지적하면서 끊임없이 공격하기 때문이다. 결국 베이즈주의를 다양하게 발전시킬 수 있다는 것이 문제점으로 보일 수도 있겠지만 그것이 문제점이라면 그것은 아직 베이즈주의가 미완의 이론이라는 정도의 소소한 문제점이지 실패한 또는 발전가능성이 없는 이론으로 평가할 문제점이 되지는 않을 것이다.

믿음 체계는 확률 공리와 조건화 규칙에 근거하는 통시적 일관성과 공시적 일관성으로 규정할 수 있다는 것이다.<sup>3)</sup> 과학적 추론에 관한 이론으로서 베이즈주의는 이러한 핵심적인 입장에 덧붙여 가장 큰 확률을 지니는 가설을 추론하라는 채택 규칙이 덧붙여진 것으로 볼 수 있다. 일반적으로 베이즈주의는 채택 규칙이 없는 것으로 이해되지만, 그 경우 IBE와의 비교가 시작부터 무의미해지기 때문에 과학적 추론에 관한 이론으로서 베이즈주의는 채택 규칙이 덧붙여진 것으로 보자는 것이다.<sup>4)</sup> 즉 가설에 대한 믿음의 정도를 확률로 나타내고, 이 확률을 새로 확보한 증거에 따라 수정할 뿐만 아니라 그러한 수정된 확률의 비교를 통해 가장 큰 확률을 지닌 가설을 채택하라는 규칙을 제시하는 이론으로 베이즈주의를 이해하자는 것이다.

IBE의 핵심을 제시하기 위해서는 먼저 IBE가 가추법과 동일한 것인지, 구분이 된다면 어떻게 구분되는 것인지를 살펴볼 수 있다. 퍼스가 제시한 가추법은 그 핵심을 발견의 맥락에 대한 강조에서 찾을 수 있다. 정당화의 맥락보다 발견의 맥락을 중요시하는 것으로 볼 수 있는 퍼스의 가추법은 다음과 같은 추론 형식을 지닌다.

- (1) 놀랄만한 관찰 사실 C가 있다.
- (2) 만약 가설 A가 참이라면 C가 관찰된다.

<sup>3)</sup> 박일호는 베이즈주의자라면 누구나 받아들일만한 원리라며 통시적 일관성과 공시적 일관성을 다음처럼 제시한다. B1. 우리의 믿음은 정도의 문제이고, 합리적 믿음의 정도는 확률 계산 법칙을 만족해야 한다. B2. 조건화(conditionalization)는 합리적인 부분적 믿음 갱신 규칙이다. 하지만 박일호는 베이즈주의를 ‘베이즈주의’라고 부르기보다 Hajek과 Hartmann (2009)의 제안처럼 ‘콜모고로프주의’라고 표상하는 것이 더 적합하다고 본다. 박일호 (2013), p. 156.

<sup>4)</sup> 베이즈주의가 이론 수용의 규칙을 지니고 있는지의 여부는 부분적 믿음과 전적인 믿음 사이의 관계 문제와 연결된다. 즉 얼마만큼의 믿음의 정도를 지닐 때 전적인 믿음을 지니는 것으로 볼 수 있는지의 문제는 가설을 수용할 수 있는 기준이 얼마만큼의 가설에 대한 믿음의 정도가 어느 선을 넘어서야 하는지의 문제와 상통하는 것이다.

(3) 따라서 A가 참이라고 고려해볼 이유가 있다.

퍼스의 가추법과 달리 하만의 IBE는 정당화의 맥락 및 설명력의 역할을 강조하는 특징을 지닌다. 하만의 IBE는 다음과 같은 추론 형식을 지닌다.

- (1) 관찰 사실  $C_1, \dots, C_n$ 이 있다.
- (2) 가설 A는  $C_1, \dots, C_n$ 을 설명한다.
- (3) 어떤 다른 가설도 A만큼  $C_i$ 를 잘 설명하지 못한다.
- (4) 따라서 A는 참이다.

하만의 IBE가 정당화의 맥락을 강조한다고 해서 발견의 맥락에 대해 아무 말도 안한다고 해석한다면 그것은 하만의 IBE를 잘못 해석한 것이다. 하지만 하만의 IBE가 상대적으로 정당화의 맥락에 더 초점을 맞추는 반면 가추법은 상대적으로 발견의 맥락에 더 초점을 맞추고 있다는 것 역시 분명한 사실이다. 게다가 설명력의 역할 역시 퍼스의 가추법보다는 하만의 IBE에서 더 명확하게 제시되고 있다. 이런 측면에서 보면, 반 프라센이 논의하고 있는 IBE는 퍼스의 가추법과 동일한 것이 아니고 하만의 IBE와 같은 것이라 할 수 있다.

이제 베이즈주의와 IBE 사이의 차이점은 다음처럼 정리할 수 있다. 첫째, 베이즈주의는 확률을 사용하여 부분적 믿음에 대해 말하지만, IBE는 확률을 사용하지 않고 믿는다와 믿지 않는다의 두 가지로만 구분되는 절대적인 믿음에 대해 말한다. 둘째, 베이즈주의는 정당화의 맥락에 대해 말하고 발견의 맥락에 대해 말하지 않지만, IBE는 정당화의 맥락뿐만 아니라 발견의 맥락에 대해서도 말한다. 셋째, 베이즈주의는 가설 채택의 문제에 대해서는 말하지 않고 새로운 증거 확보에 따른 믿음의 변화에 대해 말하지만, IBE는 그와 같은 믿음의 변화에 대해서는 말하지 않고 가설 채택의 문제에 대해 말한다.

이와 같은 차이점은 베이즈주의와 IBE가 과학적 추론에 관해 서로 다른 이론임을 분명하게 한다. 반 프라센의 양립불가능성 논증은 이러

한 차이점을 두고, 첫째, IBE도 베이즈주의처럼 확률을 사용하여 부분적 믿음에 대해 말한다는 가정을 하고, 둘째, 정당화의 맥락으로 논의의 영역을 제한하며, 셋째, 베이즈주의도 IBE처럼 가설 채택의 문제에 대해 말한다는 가정 아래 두 이론을 비교가 가능하도록 한다.

이처럼 서로 다른 점들을 빼고 나서 베이즈주의와 IBE가 근본적으로 동일하다고 주장하는 것이 어떤 의미가 있는 것인지, 즉 다른 점은 다른 점이고 같은 점은 같은 점이라고 말하는 것은 무의미한 것이 아닌지 하는 의심의 소지가 있다. 그럼에도 불구하고 반 프라센이 제시하고 이후의 과학철학자들이 검토하고 있는 베이즈주의와 IBE 사이의 양립가능성 논의는 베이즈주의와 IBE 각각에 대한 더 정확한 이해를 추구하는 기회가 될 뿐만 아니라 과학적 추론이 본질에 대한 고민의 기회를 제공한다는 점에서 의의를 찾을 수 있다.

### 3. 반 프라센의 양립불가능성 논증

반 프라센은 베이즈주의와 IBE가 양립불가능하다고 논증하여 IBE가 틀렸음을 주장한다. 그 논증의 핵심은 베이즈주의의 믿음 수정 규칙과 IBE의 믿음 수정 규칙이 충돌한다는 것이다. 베이즈주의의 믿음 수정 규칙은 조건화 규칙이다. 하지만 IBE에는 원래 믿음 수정 규칙이 없다. 여기서 반 프라센은 자신의 양립불가능성 논증을 위해 IBE가 확률을 사용하는 믿음 수정 규칙을 지닌다고 가정한다. IBE의 입장에 들어맞을 수 있는 믿음 수정 규칙으로 확률을 사용하지만 믿음의 수정 과정에서 설명력의 역할을 강조하는 것으로 가정하는 것이다. 즉 성공적인 설명을 제시하는 가설에 대해 기준에 부여했던 확률보다 더 큰 확률을 부여해야 한다는 규칙이다. 이 정도의 내용을 지닌 규칙이라면 IBE의 입장에서 부정하기 어렵다. 반 프라센의 양립불가능성 논증은 그와 같은 종류의 규칙, 즉 IBE가 믿음 수정 규칙을 지닌다면, 성공적인 설명을 제시하는 가설에 좀 더 큰 확률을 부여하는 종류의 규칙을 지니게 될 텐데 그런 어떤 규칙도 베이즈주의의 믿음 수정 규칙인 조

전화 규칙과 충돌한다는 것이다.

반 프라센의 논증은 구체적으로 베이즈주의자인 피터와 IBE를 옹호하는 전도사 사이의 대화로 전개된다. 먼저 베이즈주의자인 피터는 외계주사위를 발견한 후, 그 주사위가 1이 나올 확률을  $i/n$ 로 부여하는 가설  $H$ 를 설정했다. 피터는 가설  $H$  이외에도  $n-1$ 개의 가설을 설정하였고, 각 가설에 대해 적절한 사전확률을 부여하였다. 피터는 또 그 외계주사위를 여러 번 던진 결과는 계속해서 1이 나온 상황을 증거로 확보했다. 피터는 이와 같은 결과를 증거로 삼아  $H$ 의 확률을 조건화 규칙에 따라 갱신한다. 그 결과  $H$ 의 확률은  $i/n$ 의 확률보다 높아지게 되었다. IBE를 옹호하는 전도사는 피터의 이런 작업 과정을 지켜 본 후, 피터에게  $H$ 는 1이 계속 나왔다는 증거를 성공적으로 설명한다는 점을 들어  $H$ 에 대해 기존에 부여했던 확률보다 더 큰 확률을 부여해야 한다. 즉 성공적인 설명을 제시한 만큼 그 가설을 더 많이 신뢰해야 한다고 주장한다. 이 때 피터는 IBE 전도사에게 다시 한 번 묻는다. “뭐라고요? 내가 부여하려는 것보다 더 많이 부여해야 한다고요?” 피터는 자신이 이미 1이 계속 나왔다는 증거를 조건화 규칙에 따라 반영하여  $H$ 의 확률을 올렸다. 하지만 IBE 전도사는 그렇게 올린 정도가 불충분하다는 것이다. 성공적인 설명을 제시했다는 점은 추가적으로  $H$ 의 확률을 올릴 이유가 된다는 것이다. 이에 IBE 전도사는 피터가 부여하려는 것보다 더 큰 확률을  $H$ 에 대해 부여해야 한다고 답변한다.

IBE 전도사는 설명에 성공했다는 점을 들어 확률을 추가로 올려줘야 한다고 강조하지만 IBE 전도사에 따라 성공적 설명을 제시하는 가설의 확률을 추가로 더 올리게 되면, 베이즈주의의 조건화 규칙을 어기게 되고 결국 더치복에 빠지게 된다는 것이 반 프라센의 논증이다. 성공적으로 설명했다는 점이 가설의 사후확률을 결정하는 데에 추가로 반영되어야 하는 요소라면 IBE는 베이즈주의와 양립불가능하다는 것이다. 나아가 베이즈주의와 IBE 사이의 양립불가능성이 함의하는 바는 IBE가 틀렸다는 것이다.

반 프라센의 논증은 다음과 같이 정리할 수 있다.

- (1) 베이즈주의의 조건화 규칙과 IBE의 믿음 수정 규칙이 충돌한다.
- (2) 베이즈주의와 IBE는 양립 가능하지 않다.
- (3) 베이즈주의가 맞다.
- (4) 따라서 IBE가 틀렸다.

이러한 반 프라센의 논증에 대해 우리는 네 가지의 가능한 반박 전략을 생각할 수 있다.<sup>5)</sup> 첫째, 베이즈주의 자체의 문제점으로 인해 베이즈주의가 옳다고 추론할 수 없다는 반박이 가능하다. 둘째, 카네먼과 티베르스키가 보여주듯이 사람들은 확률을 사용하는 추리를 잘 못한다. 따라서 베이즈주의는 사람들이 실제로 추리하는 방식을 나타내지 못한다. 그러니 베이즈주의가 옳다고 할 수 없다는 반박이 가능하다. 셋째, 연역논리학이 논리적 일관성에 대해 말하는 것처럼 베이즈주의는 확률적 일관성에 대해 말할 뿐이기 때문에 IBE와 충돌할 일이 없다는 반박이 가능하다. 연역논리학의 추론 규칙은 논리적 일관성을 강요할 뿐이기 때문에 결론이 마음에 들지 않으면 전제를 바꿀 수 있도록 허용하는 것처럼 베이즈주의도 확률적 일관성을 강요할 뿐이기 때문에 마음에 들지 않는 사후확률이 나오면 사전확률과 가능도를 바꾸면 된다는 것이다.<sup>6)</sup> 넷째, 베이즈주의와 IBE는 양립 가능하기 때문에 반 프라센의 논증은 따라 나오지 않는다는 반박이 가능하다.

이상의 네 가지 반박 전략 중에서 본 논문이 따져 보고자 하는 것은 마지막의 네 번째 전략이다. 이와 관련된 다양한 양립 가능론의 입장을 다음 절에서부터 살펴보기로 한다.

5) Lipton (2004), pp. 104-7.

6) 립톤은 하우슨(Howson 2000, ch. 7)이 이러한 입장을 취하는 것으로 보고 있는 듯하다. 하우슨은 베이즈주의를 논리 체계로 보기 때문에 사전확률의 결정 문제라는 부담을 덜 수 있다고 생각하는 것이다. 하지만 하우슨이 사전확률을 임의로 바꿀 수 있다고 보는지는 의심스럽다. 연역논리학이 개별 명제의 참, 거짓을 경험과학자의 판단에 맡기고 그 판단을 존중하는 것처럼 베이즈주의도 개별 확률 부여자의 판단을 존중하는 것이지 아무렇게나 바꿀 수 있다고 주장하는 것은 아니기 때문이다.

#### 4. 기존의 양립가능론

베이즈주의와 IBE를 결합시키는 방법 중 하나는 베이즈주의가 이미 가설의 설명력을 고려하여 확률을 부여하기 때문에 성공적 설명을 제시하는 가설의 확률을 추가로 더 올릴 필요는 없다고 대답하는 것이다. 이 방안을 베이즈주의 중심의 양립가능론이라 할 수 있다. 이 방안은 베이즈주의가 사전확률과 가능성도를 결정함에 있어 가설의 설명력을 이미 고려하여 반영한다고 주장함으로써 반 프라센의 양립불가능론 논증을 반박하는 것이다.

이와 같은 방식으로 베이즈주의와 IBE를 결합시키는 것은 두 이론 모두에게 이득이 될 수 있다. IBE의 측면에서는 과학적 추론에서 설명력의 역할을 인정받는 것과 동시에 베이즈주의가 제공하는 강력하고 정교한 형식적 틀을 사용할 수 있다는 이득이 있다. 또 베이즈주의의 측면에서도 사전확률과 가능성도를 결정하는 과정을 IBE의 설명력을 도입함으로써 좀 더 직관에 부합하게 제시할 수 있다는 이득이 있다.

이와 같은 베이즈주의 중심의 양립가능론이 올바른 답변이려면, 사전확률과 가능성도가 어떤 방식으로 가설의 설명력을 고려하고 있는지를 구체적으로 밝혀낼 수 있어야 한다. 하지만 사전확률과 가능성도가 어떻게 결정되는 것인지에 대해서는 베이즈주의자들도 서로 의견이 다르다. 베이즈주의의 다양한 입장에는 사전확률과 가능성도가 단 하나의 올바른 값을 지닌다는 극단적 객관주의자의 시각에서부터 어떤 값이라도 허용될 수 있다는 극단적 주관주의자의 시각도 있기 때문이다. 그래서 가설의 설명력이 사전확률과 가능성도를 결정하는 데에 어떤 역할을 어떻게 하는가를 밝혀내는 것은 쉬운 일이 아니다.

사전확률과 가능성도의 결정 방식이 기술적인지 즉 사람들이 실제로 가설의 설명력을 고려하여 사전확률과 가능성도를 결정하기 때문인지 아니면 규범적인지 즉 가설의 설명력을 고려하여 사전확률과 가능성도를 결정하는 것이 올바르기 때문인지에 대해서도 분명한 입장을 취해야 한다. 립톤은 기술적 결정 방식을 제안한다. 립تون에 따르면, 사람들은 베이즈주의에서 말하는 확률 부여를 잘하지 못한다. 카네먼과 티베르

스키가 제시한 여러 실험 결과가 그렇다는 것이다. 그에 비해 사람들은 가설의 설명력에 따라서는 사전확률과 가능성도를 상대적으로 잘 결정한다. 따라서 인간은 베이즈주의에서 말하는 방식으로 추론하지 않는다는 공격에 대해 베이즈주의는 IBE의 도움을 얻어 옹호하는 방안을 찾을 수 있다는 것이 립톤의 평가이다. 이와 달리 바이스버그는 규범적 방식을 통한 결합을 제안하는데, 가설의 설명력이 구체적으로 어떻게 사전확률과 가능성도를 제한하는지에 대한 논의는 좀 더 구체적으로 발전될 필요가 있는 상황이다.<sup>7)</sup> 이상에서 볼 때, 설명력을 도입하면 사전확률과 가능성도를 결정하는 과정에 대해 좀 더 직관적인 이해가 가능하다는 베이즈주의 중심의 양립가능론은 아직 논의가 더 필요한 상황이라 할 수 있다.

베이즈주의와 IBE를 결합시키는 다른 방법은 IBE에 맞게 베이즈주의를 수정하는 것이다. 이 방안을 IBE 중심의 양립가능론이라 할 수 있다. IBE의 제안처럼 성공적 설명을 제시하는 가설의 확률을 추가로 더 올릴 필요가 있다고 인정하고, 조건화 규칙을 수정하는 방안을 모색하는 것이다. 베이즈주의를 수정하는 방안을 고려하는 것이다. 하지만 IBE를 살리기 위해서 이처럼 베이즈주의를 수정하는 방안은 아직 논리적 가능성일 뿐인 듯하다. 물론 베이즈주의 중심의 양립가능론도 IBE를 더 이상 IBE가 아닌 것으로 만드는 방안이라고 비판하며, IBE 중심의 양립가능론에 대해 베이즈주의를 베이즈주의가 아닌 것으로 만드는 방안이라는 부정적인 평가에 대응할 수 있다.<sup>8)</sup> 그렇게 IBE 중심의 양립가능론도 고민해 볼 수 있는 선택지로 내세울 수 있다. 하지만 IBE 자체가 불명료한 상황이어서 그런지 아직 조건화 규칙의 대체 방식 예를 들어 최대 엔트로피 법칙 같은 것이 IBE와 직접 연결된 것으로 제시되지도 않았다. 따라서 조건화 규칙을 수정하는 방식으로 IBE 중심의 양립가능론이 구체적으로 제안된 바는 없는 것으로 보인다.

이처럼 기존의 양립가능론은 베이즈주의의 부족한 측면을 IBE가 메워줄 수 있다는 입장을 중심으로 제시되어 왔다. 베이즈주의 중심의

<sup>7)</sup> Weisberg (2009).

<sup>8)</sup> Ibid.

양립가능론조차도 IBE의 설명력을 도입해서 사전확률과 가능도의 결정 과정을 명확하게 만들려고 하는 것이다. 립톤이 “베이즈주의가 믿음 수정의 메커니즘을 제공한다 할지라도 IBE는 그 메커니즘의 심리를 밝혀줄 수 있을 것”<sup>9)</sup>이라고 기대하는 것처럼 IBE가 베이즈주의에 기여할 수 있는 부분을 특별히 강조하는 것이 기존의 양립가능론의 주된 접근 방식이다.

헨더슨은 이러한 기존의 양립가능론이 반 프라센의 비판에 대한 답변일 뿐이고 베이즈주의와 IBE가 결합해야 할 다른 동기를 충분히 제시하고 있지 못하다고 비판한다. 베이즈주의가 양립가능한 이유가 단지 두 이론을 인위적으로 결합시키는 것이 가능하기 때문이라면 제대로 된 양립가능론이 아니라는 것이다. 억지 결혼식 스타일의 양립가능론은 베이즈주의와 IBE의 어느 쪽에도 큰 의미가 없다는 헨더슨의 판단은 자연스러운 양립가능론이 성립할 때 더 반박하기 어렵다. 그럼 헨더슨이 제시하는 양립가능론이 자연스러울 수 있는지를 살펴보기로 한다.

## 5. 헨더슨의 창발적 양립가능론

헨더슨의 창발적 양립가능론은 IBE가 베이즈주의의 사전확률과 가능도를 규정하는 데에 사용되는 방식으로 두 이론이 결합하는 것이 아니라 IBE가 베이즈주의로부터 창발된다는 입장이다. 궁극적으로 헨더슨은 베이즈주의와 IBE가 모두 상당히 유사한 인식적 직관을 포착하고 있다고 주장하며, 그와 같은 점에서 창발적 양립가능론은 기존의 양립가능론과 달리 두 이론 사이의 자연스러운 결합을 가능하게 한다는 것이다.

헨더슨의 창발적 양립가능론이 기존의 양립가능론과 다른 점은 IBE의 두 개념 즉 가설의 그럴듯함(likeliness)과 가설의 사랑스러움

---

9) Lipton (2004), p. 108.

(loveliness)이 각각 베이즈주의의 사전확률과 가능도에서 창발된 것이라는 입장이다. 기존의 양립가능론에서는 대부분 이 두 개념이 사전확률과 가능도를 결정하거나 결정하는 데에 중요한 역할을 하는 것으로 보았는데, 헨더슨은 이러한 입장에서 한 걸음 더 나아가 베이즈주의의 사전확률과 가능도에서 가설의 그럴듯함과 가설의 사랑스러움이 창발되었음을 주장하는 것이다.

IBE를 가설의 그럴듯함과 가설의 사랑스러움이란 두 개념으로 분석하는 것은 립تون을 따른 것이다. 립تون은 IBE가 최선의 설명을 제시하는 가설을 고르는 추론일 때, 무엇이 최선의 설명인지에 대해 논의하면서 그것이 단순히 참이 될 확률이 높은 것이라고 해서는 설명력의 역할을 제대로 인정하지 않는 것이라는 점을 들어 IBE가 단순히 가장 그럴듯한 가설에로의 추론일 수 없다고 본다. 립تون은 그 대신 최선의 설명을 제시하는 가설은 이해를 가장 잘 할 수 있게 하는 것인데, 이것은 가설이 지니는 설명력의 역할을 강조하는 것으로 가장 사랑스러운 가설이 추구하는 바라는 입장이다.

하지만 IBE를 단순히 가장 그럴듯한 가설에로의 추론이 아니라 가장 사랑스러운 가설에로의 추론으로 이해하는 것은 오해의 소지가 있다. 역시 이 점에 동의하는데, 헨더슨은 그러한 오해의 소지를 보여주기 위해 구체적으로 음모론에 대해 생각해 볼 것을 제안한다.<sup>10)</sup> 보통 음모론은 여러 가지 측면을 멋지게 설명하는 가설이지만 너무나 예외적인 원인을 지적하는 가설이어서 말이 안 되는 가설이라고 판단되는 것이다. 즉 음모론은 가설의 그럴듯함이라는 측면에서는 매우 나쁜 평가를 받는 것이다. 하지만 IBE를 가장 사랑스러운 가설에로의 추론으로 이해하면 음모론이 가장 사랑스러운 가설로 평가될 수 있기에 제대로 된 IBE라면 음모론을 배제할 수 있는 방안을 갖춰야 하는 것이다.

립تون 역시 음모설이나 아주 기묘한 가설도 괜찮은 설명을 제시할 수 있고 설명력의 측면에서만 보면 그런 가설이 가장 사랑스러운 가설로 뽑힐 가능성이 없지 않다는 점을 인정한다. 이에 대해 립تون은 최선의 설명을 제시하는 가설의 후보를 일차적으로 가설의 그럴듯함에 따라

---

<sup>10)</sup> Henderson (2014), p. 699.

선별하고, 그 뒤 그 중에서 가장 사랑스러운 가설을 고르는 방식이 있을 수 있다고 말하며, 이때 음모설이나 아주 기묘한 가설은 그럴듯함이 부족하다는 점에서 초기에 배제할 수 있게 된다고 주장한다.<sup>11)</sup> 헨더슨 역시 이런 종류의 가설을 초기에 배제하는 방안이 없다면 최선의 설명에로의 추론은 난관에 빠질 수 있다고 주장하며, 립톤과 마찬가지로 그 해결 방안을 가설의 그럴듯함에서 찾고자 한다. IBE는 가장 그럴듯한 가설에로의 추론이라는 측면과 함께 제시될 때 설득력 있는 과학적 추론의 방법으로 해석될 수 있다는 것이다. 이렇게 보면 IBE는 일차적으로 그럴듯한 가설들을 골라 그 중에서 가장 사랑스러운 가설에로 추론하는 것으로 보는 것이 하나의 방안이다. IBE를 이런 방식으로 이해하는 것의 문제점은 이 두 가지 측면, 즉 그럴듯함과 사랑스러움이 정확히 어떤 기준과 관계로 각기 역할을 하는 것인지에 관해 일반적인 업무 분담 규정이 없다는 것이다.<sup>12)</sup>

헨더슨에 따르면, 가설의 그럴듯함은 베이즈주의의 사전확률에서 나온다. 사전확률은 “새로운 정보를 입수하기 이전의 배경 지식에 근거한 것이다. 즉 조건화 규칙에 따라 새로운 정보를 반영하는 의견의 수정 과정을 거치기 이전 상태에 과학자가 지닌 가설에 대한 의견”<sup>13)</sup>으로 “과학자들이 관심을 가진 가설에 대해 [...] 가설이 얼마나 그럴듯한지의 판단 결과”<sup>14)</sup>이다. 이 때 헨더슨은 사후확률이 사전확률과 비

11) Lipton (2004), pp. 60-1.

12) Henderson (2014), p. 699. IBE 옹호자들은 아직까지 이 점에 대해 명료한 입장을 제시하지 못하는 것으로 보이며, 헨더슨 역시 이 문제를 직접 거론하면서도 혼란스러운 표현으로 불명확한 입장을 드러내고 있다. 단순히 가장 사랑스러운 가설에로의 추론이라는 측면과 가장 그럴듯한 가설에로의 추론이라는 측면을 모두 만족해야 한다고 요구하는 것은 음모론이나 과학 혁명기의 추론과 같은 상황에서는 IBE를 사용하기 어렵게 만드는 문제가 있고, 둘 중 한 가지 측면만을 만족해도 된다고 요구하는 것은 기상천외한 음모론이나 폐기될 것처럼 보이는 과거 패러다임의 이론 추론이 용인될 수 있다는 문제가 있다.

13) 여영서 (2004).

례 관계에 있다는 점에 초점을 맞추는 것으로 보인다. 사전확률이 낮다면 참일 가능성도 낮기 때문이다. 가설의 그럴듯함과 마찬가지로 사전확률이 음모설이나 아주 기묘한 가설을 초기에 배제하는 역할을 할 수 있다면, 헨더슨이 주장하는 것처럼 가설의 그럴듯함이 사전확률에서 나온 것이라 볼 수 있다는 것이다.

한편 가설의 사랑스러움은 베이즈주의의 가능도에서 나온다는 것이 헨더슨의 생각이다. 가능도는 배경 지식 및 가설이 주어졌을 때 증거에 대한 확률이다. 이 때 헨더슨은 가능도가 가설과 증거 사이의 관계에 대한 값이라는 점에 초점을 맞추는 것으로 보인다. 사전확률은 가설과 증거 사이의 관계에 대한 값이 아니기 때문이다. 가설의 사랑스러움은 증거에 대해 이해를 잘 할 수 있게 하는 것으로 가설이 증거에 대해 지니는 설명력의 역할을 무엇보다 강조하는 것이다. 이에 립톤은 IBE가 적절한 조건을 덧붙이더라도 결국 가장 사랑스러운 가설을 추론하는 것이 되어야 한다고 본다. 이에 덧붙여 헨더슨은 가장 사랑스러운 가설을 추론할 수 있기 위해서는 경쟁하는 가설들이 지니고 있는 가설의 사랑스러움을 비교할 수 있어야 하는데, 이를 위해서는 증거를 가설이 얼마나 단순하게 설명할 수 있는지와 얼마나 통일성을 가지고 설명할 수 있는지 등이 기준이 될 수 있다고 주장한다. 헨더슨은 특히 가설의 구조가 단순한가 아니면 복잡한가에 따라 가설의 사랑스러움이 달리 측정되어야 한다고 주장하며, 타가드에 따라 가설의 구조가 단순하면 단순할수록 그 핵심가설에 의존해서 증거가 얼마나 기대할만한 일이었는지를 따지고, 가설의 구조가 복잡하면 복잡할수록 핵심가설 이외의 보조가설에 의존해서 증거가 얼마나 기대할만한 일이었는지를 따지는 방안을 제시한다.<sup>14)</sup>

<sup>14)</sup> Ibid.

<sup>15)</sup> 실제 과학에서의 이론 구조는 복잡하여 핵심가설과 보조가설을 구분할 필요가 있다. 헨더슨은 설명의 과정에서 보조가설의 역할보다 핵심가설의 역할이 크면 클수록 단순하고 통일적인 설명이 더 용이해진다는 점에 주목하여 단순성과 통일성이 어떻게 증거와 가설 사이의 관계를 나타내고 있는 것인지를 제시한다. 즉 핵심가설의 역할이 크면 클수록 가설과 증거 사이

하지만 가설의 사랑스러움은 가설 자체가 지니는 특성을 평가한 것일 수 있다. 단순성이나 통일성 등이 가설 자체가 지니는 특성으로 대표적인 것인데, 립톤은 이미 하만을 언급하며 단순성이나 통일성 등이 꼭 증거와의 관계에서 파악되어야 하는 것은 아니라고 밝힌 바 있다. 가설의 사랑스러움이 증거와 독립적으로 가설 자체가 지니는 특성을 고려한다면, 그것은 가설의 사랑스러움이 사전확률과 연관을 가지는 것으로 봐야 하는 것이다. 이렇게 가설의 사랑스러움과 사전확률 사이에 연관이 있다는 주장은 가능성도에서 가설의 사랑스러움이 나온다는 핸더슨의 주장과 충돌한다. 가능성도에서 가설의 사랑스러움이 나온다는 주장은 가설의 사랑스러움과 사전확률 사이의 무관성을 함축하기 때문이다.

핸더슨은 단순성이나 통일성 등은 단지 과거의 증거가 아니라 미래의 증거까지 포함하는 전체 증거를 고려하여 평가되는 것임을 지적한다. 최선의 설명을 제시하는 가설이라는 것은 과거의 증거만이 아니라 미래의 증거까지도 포함하는 전체 증거를 고려한 결과라는 것이다. 이로부터 핸더슨은 “립تون이 제안하듯이, 전체 증거에 대한 최선의 설명을 결정하는 데에 포함된 단순성과 통일성 같은 설명적 고려를 통해 사람들이 사전확률을 제약하는지 또는 제약할 수 있는지는 분명하지 않다.”<sup>16)</sup>고 평가한다.

이러한 핸더슨의 평가는 베이즈주의에서 사전확률이 어떻게 결정되는지 뿐만 아니라 단순성과 통일성 같은 설명적 고려가 정확히 무엇을

---

의 단순성과 통일성의 힘은 커지는 것이다. 또 그럴수록 더 사랑스런 가설이 된다는 것이다. 천동설과 지동설 등 실제 과학에서의 일반 이론은 각각 핵심가설과 보조가설로 구성되어 있는데 이제 IBE를 통해 가장 사랑스러운 가설로 추론한다는 것은 천동설과 지동설 각각의 핵심가설이 주어진 증거를 설명하는 데에 있어 얼마나 주도적인 역할을 하는가에 따라 구분된다 는 것이다. 이것은 단지 주어진 증거를 천동설이나 지동설이 얼마나 확률적으로 그럴듯한 것으로 만드느냐의 문제가 아니다. 주어진 증거를 설명하는 데에 있어 핵심가설의 역할이 얼마나 큰지에 따라 가설의 아름다움이 결정되는 것이다.

<sup>16)</sup> Henderson (2014), p. 696.

따지는 것인지를 불명료하다는 점에 편승한 것이다. 미래의 증거까지도 포함하는 전체 증거에 대한 최선의 설명이 무엇인지를 결정하는 데에 단순성과 통일성 같은 설명적 고려를 하지 않을 이유가 없고 하지 못할 이유도 없다. 게다가 단순성과 통일성이 아직 제시되지 않은 미래의 증거와 어떤 관계를 지니는 것인지를 따져야 하는 것이라면, 그런 단순성과 통일성은 가설 자체가 지니는 특성이라고 보는 것이 옳다. 아직 제시되지 않은 미래의 증거가 어떤 것인 줄 알고 그 관계를 분석할 수 있겠는가? 단지 가설 자체가 지니는 특성을 놓고 미래의 증거와 어떤 관계를 지니는 것인지를 추측할 수 있을 뿐이다. 따라서 헨더슨이 제시하듯 최선의 설명이 무엇인지를 결정하는 데에 미래의 증거까지도 포함하는 전체 증거를 따져야 한다는 것은 바로 가설의 사랑스러움이 가설 자체가 지니는 특성에 의해 제약되는 측면이 있음을 뒷받침한다. 일단 가능성에서 가설의 사랑스러움이 나온다는 헨더슨의 주장이 가설의 사랑스러움과 사전확률 사이의 무관성을 함축한다면, 전체 증거의 범위가 미래의 증거를 포함한다고 하더라도 가설의 사랑스러움이 가설 자체가 지니는 특성을 평가한 것일 수 있기에 헨더슨의 입장은 성립하기 어렵다.

가능도에서 가설의 사랑스러움이 나온다는 헨더슨의 주장은 가능성과 가설의 사랑스러움이 정확히 어떻게 연결되어 있는지를 모르는 상황에 근거한 것이다. 이 점은 헨더슨과 반대의 입장에서도 마찬가지이다. 가능성과 가설의 사랑스러움이 정확히 어떻게 연결되어 있는지를 모르는 상황에 근거해서 둘 사이에 연결 관계가 있다고 주장하는 것도 크게 설득력을 얻기 어렵기 때문이다. 가설의 그럴듯함과 사전확률 사이의 관계에 있어서도 마찬가지이다. 가설의 그럴듯함이 사전확률과 정확히 어떻게 연결되어 있는지에 대해서는 아직 정밀한 분석이 없는 상황이다. 이러한 무지의 상황에서 헨더슨은 가설의 그럴듯함과 가설의 사랑스러움이 각각 사전확률과 가능성에서 나온다고 주장하고, 바로 그렇기 때문에 베이즈주의는 IBE와의 양립가능성을 위해 특별히 다르게 할 일이 없다고 주장하는 것이다.

사실 가설의 그럴듯함과 가설의 사랑스러움을 각각 사전확률과 가

능도에 연결시키려는 시도는 기존의 양립가능론에서도 여러 번 검토된 바 있다. 여기서 헨더슨은 자신의 주장이 이전의 양립가능론과 다르다고 주장하는데, 그 다른 점은 가설의 그럴듯함과 가설의 사랑스러움이 각각 베이즈주의의 사전확률과 가능도에서 나온 것이라고 주장하는 것이다. 그 나오는 관계라는 것만으로 헨더슨의 입장이 기존의 양립가능론과 다르다고 말하기는 간단하지 않다. 이 점에서 볼 때, 헨더슨이 자신의 입장을 창발적 양립가능론이라고 이름붙인 점을 고려할 필요가 있는 듯 싶다. 그럼 헨더슨이 가설의 그럴듯함과 가설의 사랑스러움이 각각 베이즈주의의 사전확률과 가능도에서 “나온” 것이라고 주장하는 것은 “창발된” 것이라고 주장하는 것으로 생각하는 것이 그럴 듯하다. 단지 가설의 그럴듯함과 가설의 사랑스러움이 사전확률과 가능도를 제약하거나 결정하는 데에 중요한 역할을 한다면서 그들을 연결하려는 시도는 헨더슨 이전에도 여러 번 있었기 때문이다.

예를 들어 오카샤가 이미 가설을 중심으로 제시되는 설명의 그럴듯함은 가설이 지니는 전반적인 신뢰도를 반영하는 것으로 가설의 사전확률  $P(H)$ 에 해당하고, 가설을 중심으로 제시하는 설명의 사랑스러움은 가설이 참일 경우 증거가 더 잘 이해할 수 있는 일이 된다는 것이어서 가능도  $P(e|H)$ 에 해당한다고 주장한 바 있다. 이때 오카샤는 가설이 더 좋은 설명을 제공할수록 가설의 사전확률이나 가능도 또는 둘다의 확률이 커지는 것이라며 가설의 그럴듯함과 가설의 사랑스러움을 각각 사전확률과 가능도에 연결시키려는 자신의 제안을 정당화했다.

헨더슨은 이러한 시도를 포함하여 기존의 모든 양립가능론이 모두 IBE와 베이즈주의를 인위적으로 결합시키는 것에 불과하고 반 프라센의 비판에 대한 답변이라는 측면 이외에 충분한 동기 없이 제안되어 왔다고 부정적으로 평가한다. 기존의 양립가능론이 모두 인위적인 결합이라는 비판을 뒷받침하기 위해 헨더슨이 제시하는 근거의 하나는 기존의 양립가능론과 달리 헨더슨의 양립가능론은 IBE와 베이즈주의의 양쪽이 서로 호혜의 관계에 있게 된다는 주장이다. 즉 “IBE는 베이지안 추론의 진행 과정을 질적 용어로 표현해내는 방법을 제시하고, 베이즈주의는 IBE에서 어떻게 설명의 덕목들이 선호되고 있는 것인지

를 체계적으로 해설하는 방법을 제시한다.”<sup>17)</sup>는 것이다. 그러나 이러한 측면은 표현이 다를 뿐 립톤이 이미 제시한 바 있다. 립تون은 그의 양립가능론에서 둘 사이의 보완 관계, 즉 IBE는 사전확률과 가능도를 결정하면서 베이즈주의를 ‘실현’시키고 베이즈주의는 믿음 수정의 메커니즘을 제공할 수 있다고 주장한 바 있다.

기존의 양립가능론을 비판하기 위해 헨더슨이 제시하는 또 다른 근거는 IBE를 이용해서 베이즈주의의 사전확률과 가능도를 제약하려는 시도가 결국 베이즈주의보다 우선해서 IBE에 규범성을 부여하는 것이 되어 버린다는 점이다. 그렇게 IBE를 베이즈주의보다 규범적으로 우선 할 수 있느냐 하는 것이다. 이 점만으로는 기존의 양립가능론을 한꺼번에 반박할 수 없다. 기존의 양립가능론 중에는 IBE를 중심으로 하는 것도 있기 때문이다. 게다가 IBE를 이용해서 베이즈주의의 사전확률과 가능도를 제약하려는 시도가 결국 베이즈주의보다 우선해서 IBE에 규범성을 부여하는 것인지에 대한 판단도 쉽지 않다. 앞에서 논의한 바처럼 사전확률과 가능도 그리고 가설의 그럴듯함과 가설의 사랑스러움이 정확히 무엇이고 어떻게 결정되는 것인지에 대해 아직 정밀한 분석이 없는 상황이기 때문이다.

무엇보다도 가설의 그럴듯함과 가설의 사랑스러움이 각각 사전확률과 가능도에서 ‘창발된다’고 주장할 수 있으려면 헨더슨은 그 ‘창발된다’는 의미를 분명하게 해야 할 것이다. IBE와 베이즈주의 사이의 관계에 대한 헨더슨의 주장이 지니는 함의 나아가 IBE와 베이즈주의 사이의 관계 자체에 대해 더 정확한 이해가 가능할 수 있으려면, 그 창발의 의미에 대해 고민할 필요가 있다. 이를 다음 절에서 시도해 보기로 한다.

---

<sup>17)</sup> Henderson (2014), p. 712.

## 6. 베이주주의와 IBE

헨더슨의 양립가능론에 따르면, 가설의 그럴듯함과 가설의 사랑스러움은 각각 사전확률과 가능도에서 창발되는 것이다. 여기서 ‘창발된다’는 것은 무슨 뜻인가? 헨더슨은 “만약 이론이 더 그럴듯하면 더 높은 사전확률을 부여받고, 이론이 더 사랑스러우면 더 높은 가능도를 부여받는다.”<sup>18)</sup>라고 말하면서도 가설의 그럴듯함과 사전확률이 그리고 가설의 사랑스러움과 가능도가 각각 일대일의 대응을 이루는 것은 아니라고 강조한다. 헨더슨은 단지 가설의 그럴듯함이 사전확률에 대해 그리고 가설의 사랑스러움이 가능도에 대해 충분조건이 된다고 주장하는 것이다. 그러나 자신의 양립가능론을 ‘창발된다’는 단어를 사용해서 ‘창발적 양립가능론’이라고 부르면서까지 강조한 것과 달리 헨더슨은 그 ‘창발된다’의 의미를 충분조건이 된다는 것 이상으로 명확하게 밝히지 않는다.

사실 가설의 사랑스러움과 가설의 그럴듯함이란 용어 자체는 IBE를 명료하게 제시하기 위해 립톤이 처음 도입한 것이고, 설명력의 요소들을 가설의 사랑스러움과 가설의 그럴듯함으로 분석하여 베이즈주의의 사전확률과 가능도에 연결하려는 시도는 헨더슨 이전에 제시된 기존의 양립가능론에서 찾아볼 수 있는 것이다. 기존의 양립가능론에서는 가설의 그럴듯함과 가설의 사랑스러움이 사전확률과 가능도를 제약하거나 결정하는 데에 중요한 역할을 한다는 입장이다. 이에 비해 헨더슨은 가설의 사랑스러움과 가설의 그럴듯함에 대해 그런 역할을 부여하기를 거부한다. 헨더슨의 새로운 양립가능론은 가설의 그럴듯함과 가설의 사랑스러움이 베이즈주의의 사전확률과 가능도에서 창발된 것이라고 주장하는 것이다. 그러나 가설의 그럴듯함과 가설의 사랑스러움이란 용어 자체는 베이즈주의의 사전확률과 가능도라는 개념보다 늦게 나타난 것이다. IBE 자체는 그 기원을 베이즈주의 이전으로 거슬러 올라갈 수 있을지 몰라도 가설의 사랑스러움과 가설의 그럴듯함이란 용

---

<sup>18)</sup> Ibid.

어는 현대의 철학자인 립تون에 의해 처음 사용된 것이기 때문이다. 그렇다면 창발이 무슨 의미를 지니던 가설의 사랑스러움과 가설의 그럴듯함이 사전확률과 가능성도에서 창발된 것이라 보기 어렵지 않을까?

가설의 그럴듯함과 가설의 사랑스러움이란 개념이 훨씬 더 직관적이고 일상적으로 이해가 쉬운 내용이라는 것은 의심의 여지가 없다. 또 베이즈주의의 사전확률과 가능성도라는 개념이 더 분석적이고 전문적이라는 점 역시 분명하다. 이런 상황에서 가설의 그럴듯함과 가설의 사랑스러움이 사전확률과 가능성도에서 창발된 것이라 할 수 있을까? 일반적으로 창발이라고 하면 기준에 없던 새로운 내용이 추가된 것이다. 하지만 가설의 사랑스러움과 가설의 그럴듯함이 사전확률과 가능성도에 없던 새로운 내용을 추가한 것인가? 그보다 사전확률과 가능성도는 가설의 사랑스러움과 가설의 그럴듯함이 어떻게 확률적 추론 과정에서 적용될 수 있는지를 구체화하는 장치가 아닐까? 이런 상황에서 창발이 무슨 의미를 지니던 가설의 사랑스러움과 가설의 그럴듯함이 사전확률과 가능성도에서 창발된 것이라 보기는 어렵다.

창발이라는 관계를 버리고 나면 IBE와 베이즈주의의 관계에 대한 헨더슨의 주장은 꽤 그럴듯하다. 그것은 헨더슨이 IBE와 베이즈주의 사이의 관계를 서로 다른 수준에서 제시된 이론이라는 주장을 펼치기 때문이다. 헨더슨은 이 주장을 IBE와 베이즈주의의 관계가 열역학과 통계역학 사이의 관계와 유사하다는 다음과 같은 비유를 통해 제시한다. 즉 “열역학은 열 현상을 거시적 관점에서 기술하는 반면 통계역학은 열 현상을 분자구조의 미시적 관점에서 기술한다. 열역학에서 사용하는 열, 엔트로피, 에너지 등의 개념은 통계역학에서 중복된 잉여개념이 되지 않는다. 오히려 그런 개념은 분자집합이 지닌 특성을 가리키는 데 사용한다.”<sup>19)</sup> 여기서 헨더슨은 열역학이 IBE라면 통계역학이 베이즈주의라고 비유하는 것이다. 헨더슨이 이 비유를 통해 IBE와 베이즈주의 사이의 관계에 대해 강조하고자 한 점은 아마도 다음과 같은 측면일 것이다. 즉 IBE와 베이즈주의가 서로 상당한 인식적 직관을 공유한다는 점, 즉 이론의 단순성과 통일성 같은 설명적 덕목들이 이론

---

<sup>19)</sup> Ibid., p. 711.

선택의 문제에서 맡는 역할에 대해 서로 유사한 평가를 하는데, 그러한 측면을 베이즈주의는 확률을 중심으로 정교한 이론 체계를 세워 제시하고 있고 IBE는 설명력이란 직관적 내용을 중심으로 알기 쉽게 제시하고 있는 것이고, IBE에서 사용하는 개념들과 연결되는 베이즈주의의 정밀한 개념들이 IBE의 개념들을 불필요한 것으로 만들지 않는다는 것이다. IBE는 불명료한 이론일 수 있지만 일상적인 논의에서 사용하기에는 쉽고 직관적이어서 오히려 적절할 수 있는 것이다. 그런 특징 때문에 IBE는 널리 사용되는 추론 방식일 수 있다. 또 그렇기 때문에 IBE의 지위가 베이즈주의에 비해 격하되는 것도 아니다.

IBE와 베이즈주의가 서로 호혜의 관계에 있을 수 있다는 점을 강조하는 헨더슨의 접근방식은 설득력을 가진다. IBE의 두 개념, 가설의 초기 그럴듯함과 가설의 사랑스러움은 각각 베이즈주의의 사전확률과 가능도에 의해 좀 더 명확한 이해가 가능하고 또 거꾸로 좀 더 명확한 이해를 가능하게 한다. 물론 가설의 초기 그럴듯함이란 개념이 과연 IBE를 구성하는 핵심 개념에 해당하는가 하고 의심할 수 있다. 사실 퍼스의 가추법이나 하만의 IBE 어디에서도 가설의 초기 그럴듯함이란 개념을 찾아볼 수 없기 때문이다. 그렇지만 IBE가 경쟁하는 가설을 제거하는 과정의 최종적인 결과로 최선의 가설에로 추론한다는 측면 역시 퍼스의 가추법이나 하만의 IBE에서 찾아볼 수 있는 부분은 아니다. 그럼에도 불구하고 여영서(2008)가 주장하듯이 IBE가 제거적 귀납법과 상통한다고 말할 수 있다면, IBE가 가설의 초기 그럴듯함이란 개념에 의존한다고 말하는 것은 잘못이 아닐 것이다.<sup>20)</sup> 마찬가지로 가능도가 과연 IBE에서 가설의 사랑스러움이란 개념에 정확히 대응하는가 하고 의심할 수 있다. 사실 가능도 개념도 가설의 사랑스러움이란 개념처럼 증거와 가설 사이의 관계에 대한 개념이지만 그 정확한 의미에 대해서는 좀 더 논의를 필요로 하는 불분명한 측면이 있는 개념이다. 가능도 개념으로 새로운 가설을 찾아내는 발견의 맥락까지도 말할 수 있다는 주장도 전혀 불가능하지는 않다.

---

20) 베이즈주의와 제거적 귀납법 사이의 관계에 대해서는 여영서 (2007)와 전영삼 (2013) 참조.

그러나 IBE와 베이즈주의가 서로 호혜의 관계에 있고 서로의 이론을 더 명확히 밝히는 데에 도움이 되는 관계라고 해서 각각의 이론이 여전히 불명료한 측면을 지닌다는 점을 부정할 수 있는 것은 아니다. 이에 무엇보다 논의가 필요한 부분은 주관적 베이즈주의, 즉 어떤 확률도 허용된다는 입장과 객관적 베이즈주의, 즉 올바른 확률은 하나라는 입장의 정확한 이해 및 구분이다. 헨더슨은 IBE와 객관적 베이즈주의의 결합은 가능할 수 있지만 IBE와 주관적 베이즈주의의 결합은 형식적 결합에 불과하기 때문에 발전할 수 없다고 주장한다. 이는 주관적 베이즈주의에 대한 오해 때문이다. 종종 주관적 베이즈주의는 사전 확률과 가능도를 규정하는 방법이 아무 것도 없다는 입장으로 해석되는데, 헨더슨은 그런 베이즈주의가 초기 그럴듯함이라는 개념으로 아주 예외적인 가설을 배제하는 방안을 지니는 IBE와 동일하다고 할 수 없었던 것이다. 하지만 여영서(2007)가 주장하듯이 주관적 베이즈주의 조차도 사전확률과 가능도를 그렇게 임의적으로 부여하는 것이 아니라 점을 고려한다면, 과학적 추론의 방법을 제시하는 베이즈주의는 이미 과학의 방법이 요청하는 과학적 객관성을 확보하기 때문에 헨더슨이 우려하듯 베이즈주의가 아주 예외적인 가설을 배제하는 방안을 가지지 못한 것이 아니다.<sup>21)</sup> IBE가 객관적인 과학적 추론 방식을 제시하는 만큼 베이즈주의도 충분히 객관적인 과학적 추론 방식을 제시하고 있다.

마지막으로 IBE와 베이즈주의의 올바른 관계를 정립하기 위해 놓치지 말아야 할 점으로 IBE에 대한 일반적 착각을 지적할 필요가 있다. 그것은 IBE가 설명을 추론의 안내자로 삼아야 한다는 입장이지만 IBE의 목표는 설명을 제시하는 것이 아니라 진리를 추론하는 방법이라는 점이다. 립톤은 진리가 다 좋은 설명을 제공하는 것은 아니라고 지적하는 것과 마찬가지로 설명을 제공할 수 있다고 해서 다 진리는 아니다. IBE가 설명을 중시하지만 설명이라고 하는 것은 진리가 지니는 힘 일 뿐이다. 특히 진리가 아니어도 그럴듯한 설명을 제시할 수 있고, 그럴듯한 설명을 제시한다고 해서 진리라고 추론하는 것은 잘못된 판단

---

<sup>21)</sup> 여영서 (2007) 참조.

이다. 이런 오판을 피할 필요가 있다면 IBE와 베이즈주의는 모두 진리 탐구의 추론 방식이라는 점을 다시 한 번 명확히 인식해야 할 것이다. 설명은 IBE의 최종 목적도, 베이즈주의의 최종 목표도 아니다. 설명은 그 자체로 커다란 유용성을 지니지만 진리 추론의 과정에서는 하나의 단서이고 도구이다. IBE와 베이즈주의는 모두 과학적 추론의 방법을 제시하고 있다.

## 참고문헌

- 박일호 (2013), 「조건화와 입증: 조건화 응호 논증」, 『논리연구』 16집 2호, pp. 155-87.
- 여영서 (2004), 「베이즈주의와 제거적 귀납주의」, 『논리연구』 7집 2호, pp. 121-46.
- \_\_\_\_\_ (2007), 「베이즈주의의 사전확률과 과학적 객관성」, 『철학탐구』 22집, pp. 147-71.
- 전영삼 (2013), 『귀납, 우리는 언제 비약할 수 있는가』, 아카넷.
- \_\_\_\_\_ (2015), 「최선의 설명에로의 추론과 귀납적 체계화」, 『과학철학』 18권 2호, pp. 65-99.
- Douven, I. (2011), “Abduction”, *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.), URL=<<http://plato.stanford.edu/archives/spr2011/entries/abduction/>>.
- Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: MIT Press.
- Harman, G. (1970), “The Inference to the Best Explanation”, *Philosophical Review* 74: pp. 88-95.
- Henderson, L. (2014), “Bayesianism and Inference to the Best Explanation”, *BJPS* 65: pp. 687-715.
- Lipton, P. (2004), *Inference to the Best Explanation*, Routledge.
- Okasha, S. (2000), “Van Fraassen's Critique of Inference to the Best Explanation”, *Studies in History and Philosophy of Science* 31: pp. 691-710.
- Weisberg, J. (2009), “Locating IBE in the Bayesian Framework”, *Synthese* 167: pp. 125-44.

논문 투고일	2016. 07. 03.
심사 완료일	2016. 07. 12.
게재 확정일	2016. 07. 20.

## Bayesianism and IBE

Yeongseo Yeo

---

Van Fraassen argued that Bayesianism and IBE are incompatible and concludes that IBE is wrong. However, there has been various attempts to rebut van Fraassen's argument and present a way to make them compatible. Recently, Henderson argued that there is nothing Bayesianism has to do in order to make two theories compatible. This paper agrees with Henderson that Bayesianism and IBE shares an epistemic intuition but disagrees with Henderson that IBE emerges from Bayesianism. Instead, this paper claims that the relation between IBE and Bayesianism is not only that they are compatible but also that they are beneficial to each other.

**Keywords:** Bayesianism, IBE, Henderson, van Fraassen