논리연구 16-2(2013) pp. 155-187

조건화와 입증: 조건화 옹호 논증* **

박 일 호

【국문요약】 본 논문의 목적은 베이즈주의 믿음 갱신 규칙인 조건화를 옹호하는 것이다. 이를 위해서 필자는 경험과 무관한 믿음은 바뀌지 말아야 한다는 무관성 원리를 도입한다. 이 원리를 엄밀하게 정식화한 뒤, 무관성 원리와 조건화가 동치라는 것이 증명된다. 그리고 이 무관성 원리를 베이즈주의 입증 이론을 이용해서 옹호한다. 이를 위해서 필자는 베이즈주의 입증 이론가들이라면 받아들여야 하는 몇 가지 논제들을 제시하고, 무관성 원리를 위반한다면 그 논제들이 만족될 수 없다는 것을 보여준다.

【주요어】 베이즈주의, 조건화, 베이즈주의 입증이론, 무관성 원리

접수일자: 2013.02.05 심사 및 수정 완료일: 2013.04.01 게재확정일: 2013.04.10

* 이 논문은 2010년도 정부재원(교육과학기술부 인문사회연구역량강화사업)으로 한국연구재단의 지원을 받아 연구되었음.[NRF-2010-352-A00009]

** 중요하고 엄격한 논평을 통해 몇 가지 흥미로운 문제를 지적해 주신 익명의 심사위원들에게 감사의 말씀을 드린다.

1. 서론: 베이즈주의 인식론의 두 가지 원리

1970년대 이후 베이즈주의는 인식론, 특히 과학철학의 다양한 인식론적 문제들에 대한 흥미로운 해결책을 제시해왔다. 더 나아가 최근 베이즈주의 인식론의 연구 영역은 전통 인식론적 주제와 새로운 인식론적 주제로 확장되고 있다. 가령, 최근 밝혀진 지각적 독단론과 베이즈주의와의 충돌은 전통 인식론자들과 베이즈주의 인식론자들로 하여금 서로의 영역에 관심을 가지게 하였다. 뿐만 아니라, 근래 많은 학자들의 관심을 끌고 있는 사회 인식론에서도 베이즈주의 인식론은 중요한 역할을 해내고 있다.[1] 사회 인식론은 개별적인 믿음들로부터 집단의 믿음을 형성하는 합리적인 방법, 다른 이의 믿음이 나의 믿음 형성에 미치는 영향 등을 탐구한다. 확률을 이용하는 베이즈주의의 형식적-수학적 방법은 사회 인식론의 중요한 탐구 도구로 간주되고 있다.

이런 베이즈주의 인식론의 정체성은 다양한 방식으로 규정될 수 있다. 각자가 받아들이는 논제들에 따라서 여러 종류의 베이즈주의자들이 존재한다. 하지만 대부분의 베이즈주의자들이 받아들이는 최소 원리들도 있다. 특히 베이즈주의자라면 다음 두 원리를 받아들일 수밖에 없다.

> B1. 우리의 믿음은 정도의 문제이고, 합리적 믿음의 정도는 확률 계산 법칙을 만족해야 한다.
>
> B2. 조건화(conditionalization)는 합리적인 부분적 믿음 갱신 규칙이다.

---

1) 지각적 독단론과 베이즈주의 사이의 관계에 대한 대표적인 논의는 White (2006)에서 발견할 수 있다. 그리고 사회 인식론에 관해서는 Goldman and Whitcomb (2011)을 보라.

만약 우리의 믿음이 정도의 문제가 아니라, 전부 아니면 전무(all or nothing)의 문제라고 생각하는 사람은 베이즈주의자일 수 없다. 모든 베이즈주의자들은 내일 비가 온다는 것을 80%정도 믿는 것과 같은 부분적 믿음(partial belief)의 존재를 인정하고 있으며, 그 부분적 믿음의 합리성을 고찰하고 있다. 그들은 부분적 믿음들의 합리성을 확률 계산 규칙, 즉 확률의 수학적 공리를 이용해 규명하려고 한다. 흔히 우리들은 논리학 규칙을 만족하지 않는 믿음들은 합리적이지 않다고 생각한다. 이와 마찬가지로 베이즈주의자들은 확률 계산 규칙을 만족하지 않는 부분적 믿음은 합리적이지 않다고 생각한다.[2] 바로 이 점에서 B1을 거부하는 사람들은 베이즈주의자라고 불릴 수 없다.

여기서 B1은 부분적 믿음에 대한 공시적 논제라는 것에 주목할 필요가 있다. 즉 B1은 특정 시간의 부분적 믿음들에 대한 이야기만을 하고 있을 뿐이다. 이와 달리 B2는 통시적 논제이다. 베이즈주의자들은 특정 시간의 부분적 믿음뿐만이 아니라, 부분적 믿음의 시간에 걸친 변화들도 다루고 있다. 가령, t 시점에 있는 어떤 사람을 생각해보자. 그가 t+1 시점에 무언가를 경험하게 되었다고 하자. 그럼 그 경험은 그의 부분적 믿음을 수정하도록 만들 것이다. B2는 조건화에 따라서 믿음을 수정하는 것은 합리적이라고 말한다. 다음은 가장 단순한 형태의 조건화 규칙이다.

2) 부분적 믿음의 존재를 인정하지만, 확률을 거부하는 입장도 있다. 몇몇 학자들은 확률이 아니라 다른 형식적 도구를 이용해서 부분적 믿음을 표상하려고 한다. 가령, Zadeh (1978)의 fuzzy sets, Dempster (1967)과 Shafer (1967)의 evidence functions, Shackle (1949)의 potential surprise functions가 대표적이다. 사실 '베이즈주의'는 베이즈주의의 핵심적인 특징을 포착하지 못하는 이름이다. Hájek and Hartmann (2009)는 확률, 특히 콜모고로프(Kolmogorov)의 확률 이론을 이용해서 믿음의 정도를 표상하려고 한다는 의미를 포착하기 위해서는 '베이즈주의'라는 이름보다는 '콜모고로프주의'가 더 어울린다고 말한다.

> **단순 조건화(Simple Conditionalization, SC).**
> 경험에 의해서 명제 E를 확신하게 되었을 때, 즉 $p(E) \neq q(E)=1$일 때, 모든 명제 X에 대해서 다음이 성립한다: $q(X)=p(X|E)$.

여기서 p는 경험 이전에 가지고 있었던 부분적 믿음 함수이고, q는 경험 이후에 가지게 된 부분적 믿음 함수이다. 흔히, 우리는 p를 사전 함수, q를 사후 함수라고 부른다. 여기서 우리는 주의할 것이 있다. B2는 모든 부분적 믿음의 갱신이 조건화를 따라야 한다고 말하지 않는다. 단지 조건화가 합리적 믿음 갱신 규칙이라고 말할 뿐이다.

사실, 베이즈주의자들은 모든 부분적 믿음의 갱신이 조건화를 따라서 이루어져야 한다고 주장할 수 없다. 왜냐하면 조건화의 적용 범위는 제한적이기 때문이다. 특히, 조건화는 경험에 직접적으로 영향을 받아 일어나는 부분적 믿음의 변화에 대해서 말하고 있지 않다. SC를 살펴보자. 이 규칙은 *경험에 의해서 명제 E를 확신하게 되었을 때* 다른 믿음들이 어떻게 변해야 하는 지 말해줄 뿐이다. 조건화는 우리는 왜 경험을 통해서 명제 E*가 아닌 명제 E를 확신하게 되었는가에 대해서 침묵하고 있다. 즉 조건화는 경험이 어떤 명제에 영향을 주어야 하는지 말해주고 있지 않다.

이 뿐만이 아니다. SC는 경험에 의해서 확신하게 되는 명제가 있는 경우에만 적용될 수 있다. 하지만 경험이 항상 무언가를 확신하게 만들어주는 것은 아니다. 붉은 빛 조명이 켜진 방에 있는 어떤 사람을 생각해보자. 그 사람은 방에 있는 어떤 헝겊을 관찰하고 있다. 이 관찰 이후에 그는 헝겊의 색깔에 대한 자신의 부분적 믿음을 수정하였다고 하자. 가령 그는 그 헝겊이 붉은 색이라는 것에 대한 믿음의 정도를 수정하였다고 하자.[3] 이 경우, 그의 관찰에 의

해서 확신하게 된 명제가 있는가? 아마도 반드시 그런 명제가 있다고 말할 수는 없을 것이다. 그는 그 헝겊이 붉은 색이라는 것을 강하게 믿게 되겠지만, 그것을 확신하지는 못할 수도 있다. 이와 같이 경험이 우리를 항상 어떤 명제에 대한 확신으로 이끄는 것은 아니다. 그럼 SC를 살펴보자. 이 규칙은 경험에 의해서 어떤 명제 E를 확신하게 된 경우에만 적용될 수 있다. 따라서 SC는 위의 헝겊의 사례에는 적용될 수 없다. 물론 베이즈주의자들은 헝겊의 사례에 적용할 수 있도록 조건화를 일반화시키기도 한다. 다음 제프리 조건화가 대표적이다.

제프리 조건화 (Jeffrey Conditionalization, JC).
경험에 의해서 어떤 분할 집합 $\mathbf{E}=\{E_1,\cdots E_n\}$의 각 원소들에 대한 부분적 믿음이 수정되었을 때, 즉 모든 $E_i \in \mathbf{E}$에 대해서 $p(E_i) \neq q(E_i) < 1$일 때, 모든 명제 X에 대해서 다음이 성립한다: $q(X)=\Sigma_i q(E_i)p(X|E_i)$.

여기서 분할 집합이란 원소들이 서로 양립불가능하고, 모든 원소들을 선언지로 가진 선언문이 논리적 진리라는 것을 뜻한다. 그럼 헝겊의 예를 다시 생각해보자. 붉은 빛의 조명이 켜진 방에서 헝겊을 관찰했더니 그 헝겊이 붉은 색이라는 것을 0.8정도로 믿게 되었다고 하자. 여기서 그 헝겊의 색이 붉다는 명제를 E라고 하자. 그럼 분할 집합 {E, ~E}의 원소들에 대한 부분적 믿음이 수정되었기 때문에 JC를 적용할 수 있다. E에 대한 새로운 부분적 믿음은 0.8이고 ~E에 대한 새로운 부분적 믿음은 0.2이기 때문에, JC는 모든 명제 X에 대해서 다음이 성립해야 한다고 말한다:

3) 이 글에서 '부분적 믿음(partial belief)'과 '믿음의 정도(degree of belief)'는 동의어로 간주된다. 아래에서 필자는 이 둘 모두를 사용할 것이다.

q(X)=0.8p(X|E)+0.2p(X|~E). 이렇듯 SC와 달리 JC는 헝겊의 예를 성공적으로 다룰 수 있다.

JC는 SC를 일반화한 것이다. 즉 JC는 SC를 함축한다.[4] 하지만 JC가 모든 사례를 다룰 수 있는 것은 아니다. 경험 이후에 명제 E, F에 대한 믿음의 정도가 수정되었다고 하자. 이 경우, JC를 적용하기 위해서는 이 명제들의 집합 {E,F}가 분할 집합이어야 한다. 그렇지 않은 경우에는 JC를 적용할 수 없다. 따라서 JC를 선호하는 베이즈주의자들 역시 모든 합리적 믿음 갱신이 조건화를 따라야 한다고 주장하지는 않는다. 베이즈주의자들이 주장하는 것은 단지 조건화가 합리적 믿음 갱신 규칙이라는 것뿐이다.

필자는 지금껏 베이즈주의 인식론의 핵심적인 논제를 설명했다. 필자는 이 논제들을 받아들이지 않는 베이즈주의자들은 극히 드물 것이라고 생각한다. 이 두 가지 논제들 중에서 필자가 관심을 가지는 것은 두 번째 논제, 즉 B2이다. 앞에서 언급했듯이 조건화는 모든 믿음 갱신에 적용될 수 있는 것은 아니다. 하지만 조건화의 적용 범위는 본 논문의 주제가 아니다.[5] 그보다 필자는 본 논문에서 조건화 그 자체에 관심을 기울일 것이다. 특히 조건화의 본성을 탐구하고 그것을 다소 간단한 방식으로 옹호(vindicate)하려고 시도할 것이다. 이 목표를 달성하기 위해 본 논문은 우선 부분적 믿음의 갱신에 관한 아주 직관적인 원리와 조건화가 동치라는 것을 증명할 것이다 (2절). 그리고 그 증명을 이용해서 과학철학의 입증의 문제를 다루는 데 있어 조건화만큼 좋은 믿음 갱신 규칙은 없다는 것

4) 관련된 증명은 여러 베이즈주의 교과서에서 찾아볼 수 있다. 가령, Howson and Urbach (1993), Jeffrey (1983)을 보라.

5) 조건화의 적용범위에 대한 논의를 위해서는 Bradley (2005), Douven and Romeijin (2011), Grove and Halpern(1997), Harper and Kyburg (1968), Jeffrey (1970), Levi (1967), Park (forthcoming), van Fraassen(1981, 1989)를 보라.

을 보일 것이다 (3-4절).

2. 조건화와 무관성 원리

무언가를 경험했다고 가정하자. 그리고 그 경험에 의해서 어떤 명제에 대한 믿음의 정도가 수정되었다고 하자. 그럼 다른 명제들에 대한 믿음의 정도는 어떻게 되어야 하는가? 이 질문과 관련해서, 우리가 당연하게 여기는 것이 하나 있는 듯하다. 가령, 밖을 관찰하기 전에 서울에 눈이 온다는 것을 30% 정도로 믿고 있었다고 생각해 보자. 그런데 밖을 관찰한 이후 서울에 눈이 온다는 것을 90% 정도로 믿게 되었다고 하자. 이 경우, 서울의 교통은 혼잡하다는 것에 대한 믿음 역시 수정되어야 할 것 같다. 왜냐하면 서울에 눈이 온다는 것과 교통이 혼잡하다는 것은 서로 무관하지 않기 때문이다. 그럼 미국의 증시가 하락한다는 것에 대한 믿음은 어떠한가? 아마도 많은 사람들은 서울의 날씨와 미국의 증시는 아무런 상관이 없다고 여길 것이다. 따라서 서울에 눈이 온다는 것에 대한 믿음의 변화는 미국의 증시가 하락한다는 것에 대한 믿음에 어떤 영향도 가하지 못할 것이며, 이에 서울에 눈이 온다는 믿음이 바뀌었더라도 미국의 증시가 하락한다는 것에 대한 믿음은 바뀌지 말아야 할 것이다. 정리하자면 경험에 의해서 어떤 명제에 대한 믿음의 정도가 바뀌었을 때, 그 명제와 관련이 없는 명제의 믿음의 정도는 바뀌지 말아야 한다는 것은 무척 당연한 듯이 보인다. 나는 이것을 무관성 원리(Irrelevance Principle)라고 부를 것이다. 아주 간단하게 말해, 이 원리는 경험과 무관한 믿음은 바뀌지 말아야 한다는 것이다.

1) 무관성 원리의 정식화

그럼 무관성 원리는 어떻게 정식화할 수 있을까? 특히 두 명제

가 무관하다는 것은 어떻게 나타낼 수 있을까? 만약 베이즈주의자들처럼 확률을 이용해 우리의 부분적 믿음을 나타낸다면, 어렵지 않게 무관성 원리를 정식화할 수 있다. 확률을 이용했을 때, 두 명제 X, Y 사이의 무관성은 다음의 확률적 독립성으로 정식화된다: $p(XY)=p(X)p(Y)$.[6] 그럼 앞에서 언급한 무관성 원리는 다음과 같이 나타낼 수 있다.

IP. 경험에 의해서 명제 E를 확신하게 되었을 때, 즉 $p(E) \neq q(E)=1$일 때, 모든 X에 대해서 다음이 성립한다: 만약 $p(XE)=p(X)p(E)$라면, $p(X)=q(X)$이다.

물론 이것은 다양한 방식으로 일반화될 수 있다. 우선 조건부 믿음의 정도를 포괄할 수 있도록 IP를 수정할 수 있다. 명제 Z라는 조건 아래에서 명제 X와 Y 사이의 독립성은 다음과 같이 정식화된다: $p(XY|Z)=p(X|Z)p(Y|Z)$. 그럼 이 조건부 독립성을 이용하여 IP는 다음과 같이 일반화될 수 있다.

단순 무관성 원리(Simple Irrelevance Principle, SIP). 경험에 의해서 명제 E를 확신하게 되었을 때, 즉 $p(E) \neq q(E)=1$일 때, 모든 명제 X와 Y에 대해서 다음이 성립한다: 만약 $p(XE|Y)=p(X|Y)p(E|Y)$라면, $p(X|Y)=q(X|Y)$이다.

필자가 IP보다 더 일반적인 위 원리에 '단순'이라는 표현을 붙인

---

6) 이 논문에서 "XY"는 X와 Y로 이루어진 연언문을 나타낸다. 확률적 독립성은 다양한 방식으로 나타낼 수 있다. 가령 $p(X|Y)=p(X)$와 같이 확률적 독립성을 나타낼 수 있다. 하지만 이렇게 나타내는 경우 $p(Y)>0$이라는 것이 가정되어야 한다. 왜냐하면 $p(Y)=0$인 경우 $p(X|Y)$는 정의되지 않기 때문이다. 그러나 위의 방식대로 나타내는 경우 $p(Y)>0$이 가정될 필요는 없다.

이유는 단순 조건화 SC와 관련이 있다. 다음 절에서 분명해지겠지만 SIP는 SC와 동치이다. 이에 표현상의 일관성을 위해 필자는 IP가 아니라 위 무관성 원리에 '단순'이라는 수식어을 붙였다. 하지만 그 본성을 고려할 때, SIP가 말하는 바는 IP가 말하는 바와 같다. 경험에 의해서 명제 E를 확신하게 되었다고 가정해보자. IP는 그 명제 E와 무관한 명제에 대한 *비조건부* 믿음이 그대로 유지되어야 한다고 말하고 있는 반면, SIP는 E와 *조건부로* 무관한 명제들에 대한 *조건부 믿음*이 그대로 유지되어야 한다고 말하고 있을 뿐이다. 비조건부 믿음과 조건부 믿음 사이의 차이일 뿐 IP와 SIP는 본질적으로 같은 직관을 포착하고 있다. SIP가 IP를 함축한다는 것을 고려할 때 이 둘이 본질적으로 같은 직관을 포착한다는 것은 더욱 분명해진다.[7]

무관성 원리는 더욱 일반화될 수 있다. IP와 SIP는 모두 경험에 의해서 명제 E를 확신하게 되는 경우만을 다루고 있다는 점에 주목하라. 따라서 SC가 아니라 JC가 적용될 수밖에 없는 상황, 가령 경험에 의해서 어떤 명제를 확신하게 되는 것이 아니라 단지 어떤 분할 집합 $\mathbf{E}=\{E_1,\cdots E_n\}$의 각 원소들에 대한 부분적 믿음이 수정된 상황에 대해서 SIP는 침묵하고 있다. 이런 상황까지 포괄할 수 있도록 무관성 원리를 확장하고 싶다면, 우리는 분할 집합 **E**와 임의의 명제 X 사이의 확률적 독립성 관계를 정식화할 수 있어야 한다. 앞에서 제시된 확률적 독립성은 모두 명제들 사이의 관계였다는 것을 기억하라. 하지만 지금 우리에게 필요한 것은 분할 집합과 명제 사이의 확률적 독립성 관계다. 이 관계는 다음과 같이 제시될 수 있다:[8]

---

7) SIP는 모든 X와 Y에 대한 무관성 원리이다. 따라서 Y가 논리적 진리일 때도 SIP는 성립해야 한다. Y가 논리적 진리일 때, SIP는 IP와 동치가 된다. 따라서 SIP는 IP를 함축한다.

8) 이런 정의는 Diaconis and Zabel (1982)에서 찾아볼 수 있다.

분할집합과 명제 사이의 확률적 비조건부 독립성.
명제 X가 분할 집합 $\mathbf{E}=\{E_1,\cdots E_n\}$와 확률적으로 독립적인 경우에, 그리고 그런 경우에만 다음이 성립한다: 모든 $E_i \in \mathbf{E}$에 대해서 $p(XE_i)=p(X)p(E_i)$이다.
분할집합과 명제 사이의 확률적 조건부 독립성.
명제 Y라는 조건 아래에서 명제 X가 분할 집합 $\mathbf{E}=\{E_1,\cdots E_n\}$와 확률적으로 독립적인 경우에, 그리고 그런 경우에만 다음이 성립한다: 모든 $Ei \in \mathbf{E}$에 대해서 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$이다.

이제 위 확률적 독립성에 대한 정의를 이용하면, SIP는 다음과 일반화될 수 있다.

일반 무관성 원리(Generalized Inirrelevance Principle, GIP).
경험에 의해서 어떤 분할 집합 $\mathbf{E}=\{E_1,\cdots E_n\}$의 각 원소들에 대한 부분적 믿음이 수정되었을 때, 즉 모든 $E_i \in \mathbf{E}$에 대해서 $p(E_i)\neq q(E_i)<1$일 때, 모든 명제 X, Y에 대해서 다음이 성립한다: 만약 모든 $E_i \in \mathbf{E}$에 대해서 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$라면, $p(X|Y)=q(X|Y)$이다.[9]

물론 GIP는 SIP와 IP를 함축한다.[10] 이 무관성 원리를 통해서

---

9) 여기서 한 가지를 주의해야 한다. 특히 "만약 모든 $Ei \in \mathbf{E}$에 대해서 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$라면, $p(X|Y)=q(X|Y)$이다."를 "모든 $E_i \in \mathbf{E}$에 대해서, 만약 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$라면 $p(X|Y)=q(X|Y)$이다."로 혼동해서는 안 된다. 기호화해서 표현한다면, 앞에 진술은 "$\forall E_i \in \mathbf{E} p(XE_i|Y)=p(X|Y)p(E_i|Y) \supset p(X|Y)=q(X|Y)$"이지만 뒤에 진술은 "$\forall E_i \in \mathbf{E}[p(XE_i|Y)=p(X|Y)p(E_i|Y) \supset p(X|Y)=q(X|Y)]$"이다. 이 두 진술에서 전칭 양화사의 범위에 주의해야 한다.

10) 앞의 주석 6)에서 SIP가 IP를 함축한다는 것을 보였다. GIP가 SIP를 함축한다는 것도 쉽게 보일 수 있다. 경험 이후 $\mathbf{E}=\{E, \sim E\}$라고 하자. 그럼 GIP로

필자가 포착하고자 하는 직관이 무엇이었는지 다시 생각해보자. 거칠게 말하자면, 어떤 경험과 무관한 믿음은 그 경험 이후에도 바뀌지 말아야 한다는 것이었다. 만약 경험을 어떤 명제 E에 대한 확신으로 나타낼 수 있다면, 무관성 원리는 E와 확률적으로 무관한 명제들에 대한 (조건부) 믿음은 그대로 유지되어야 한다고 말한다. 이것은 SIP(과 IP)에 의해서 정식화되었다. 또 경험을 어떤 분할 집합의 원소들에 대한 믿음의 변화로 나타낼 수 있다면, 무관성 원리는 그 분할 집합과 무관한 명제들에 대한 믿음은 그대로 유지되어야 한다고 말한다. 이것은 GIP에 의해서 정식화되었다.

사실 필자에게 무관성 원리는 특별한 정당화가 필요 없을 정도로 자명해 보인다.[11] 무언가를 경험한 이후 그 경험과 상관없는 믿음이 바뀐다면, 우리의 믿음 체계는 큰 결함을 가지게 되는 것처럼

부터 다음이 도출된다:

(*) 경험에 의해서 어떤 분할 집합 **E**={E,~E}의 각 원소들에 대한 부분적 믿음이 수정되었을 때, 모든 명제 X, Y에 대해서 다음이 성립한다: 만약 $p(XE|Y)=p(X|Y)p(E|Y)$이고 $p(X{\sim}E|Y)=p(X|Y)p({\sim}E|Y)$라면, $p(X|Y)=q(X|Y)$이다.

확률 계산 규칙에 따르면, $p(XE|Y)=p(X|Y)p(E|Y)$와 $p(X{\sim}E|Y)=p(X|Y)p({\sim}E|Y)$는 동치이다. 따라서 GIP로부터 다음이 도출된다.

(**) 경험에 의해서 어떤 분할 집합 E={E,~E}의 각 원소들에 대한 부분적 믿음이 수정되었을 때, 모든 명제 X, Y에 대해서 다음이 성립한다: 만약 $p(XE|Y)=p(X|Y)p(E|Y)$라면, $p(X|Y)=q(X|Y)$이다.

그리고 분할 집합 {E,~E}의 원소들 중에서 E에 대한 믿음의 정도가 1이 되는 경우 (**)는 SIP와 동치가 된다. 따라서 GIP는 SIP를 함축한다.

11) 이 무관성 원리가 무척 직관적이라는 점에 동의하신 익명의 심사위원 중에 한 분은 이 원리가 확률 공리로부터 도출되며, 심지어 동어반복이라고 지적하셨다. 하지만 이런 지적은 분명 잘못이다. 무관성 원리는 확률 공리로부터 도출되지 않는다. 확률 공리는 하나의 확률 함수에 대한 것이다. 그것은 확률 함수들 사이의 관계에 대해서는 언급하지 않는다. 한편 무관성 원리는 두 개의 확률 함수 사이의 관계에 대한 것이다.

보이기 때문이다. 무관성 원리가 성립하지 않는 믿음의 변화는 과학자가 수상술(手相術)에 따라서 자신의 믿음을 바꾸는 것에 비유할 수 있다. 철수는 의사이다. 그는 손금의 모양과 노년의 건강이 아무런 관련이 없다고 생각한다. 그럼에도 불구하고 그는 자신의 손금 모양을 보고 자신이 노년에 큰 질병을 앓게 될 것이라는 명제에 대한 믿음을 수정한다. 이런 철수의 믿음 변화는 어떤 문제를 가지고 있다는 것은 분명하며, 이에 대한 어떤 특별한 논증도 필요 없어 보인다.[12] 하지만 필자는 3절에서 무관성 원리를 옹호하는 논증을 제시할 것이다. 그리고 그 논증을 이용해서 조건화를 옹호할 것이다. 그 전에 우리는 무관성 원리와 조건화, 특히 GIP와 JC 사이의 관계를 먼저 살펴보아야 한다.

2) 무관성 원리와 조건화

앞에서 언급했듯이 JC는 SC를 일반화한 것이다. GIP 역시 IP와 SIP를 일반화한 것이다. 따라서 조건화와 무관성 원리 사이의 일반적인 관계를 파악하기 위해서는 JC와 GIP 사이의 관계를 다루는 것이 합당하다. 아래에서 나는 이 둘 사이의 관계에 보다 집중할 것이다.

JC는 다양한 방식으로 재서술될 수 있다. 즉 JC는 다양한 동치문을 가지고 있다.[13] 그 중에서 가장 많이 알려진 것은 다음 고정

12) 여기서 필자는 *의사*인 철수가 수상술을 신뢰하는 것이 문제라고 말하고 있지 않다. 필자의 지적은 *노년의 건강과 손금의 모양이 무관하다는 것을 알고 있는* 철수가 손금의 모양에 따라서 자신의 믿음을 바꾸는 것이 문제라는 것이다. 필자의 논의 맥락에서, 건강과 손금의 모양이 관련 있다고 생각하는 의사가 손금의 모양에 따라서 자신의 믿음을 바꾸는 것은 전혀 문제가 되지 않는다.

13) 고정성 이외에서 JC와 동치인 명제들이 있다. 그런 명제들은 Jeffrey (2004), pp.56-57에서 찾아볼 수 있다.

성(Rigidity)이라고 불리는 것이다.

> **고정성.**
> 경험에 의해서 어떤 분할 집합 $\mathbf{E}=\{E_1,\cdots E_n\}$의 각 원소들에 대한 부분적 믿음이 수정되었을 때, 즉 모든 $E_i \in \mathbf{E}$에 대해서 $p(E_i) \neq q(E_i) < 1$일 때, 모든 명제 X에 대해서 다음이 성립한다: 모든 $E_i \in \mathbf{E}$에 대해서, $p(X|E_i)=q(X|E_i)$이다.

고정성은 경험에 의해서 그 믿음의 정도가 수정된 명제를 조건으로 하는 조건부 믿음의 정도는 바뀌지 말아야 한다는 것이다. 그리고 고정성과 JC가 동치라는 것은 널리 알려진 사실이다. 그럼 고정성과 GIP 사이의 관계는 어떠한가? 만약 우리가 이 둘이 서로 동치라는 것을 보인다면, GIP와 JC가 서로 동치라는 것을 증명하는 셈이 된다. 과문한 탓인지는 모르겠으나, 필자는 JC가 GIP와 동치라는 사실을 지적하는 글을 아직 본 적이 없다. 만약 필자의 증명이 옳다면, GIP라는 아주 직관적인 원리를 통해서 JC를 옹호할 수 있을 것이다.

그럼 이제, 고정성이 GIP와 동치라는 것을 증명해보자. 아래의 증명은 두 가지로 구성된다. 필자는 먼저 GIP가 고정성을 함축한다는 것을 보일 것이다. 그 다음 필자는 고정성이 GIP를 함축한다는 것을 보일 것이다.

가. GIP는 고정성을 함축한다

GIP와 고정성의 전건을 가정하자. 즉 모든 $E_i \in \mathbf{E}=\{E_1,\cdots E_n\}$에 대해서 $p(E_i) \neq q(E_i) < 1$라고 가정하자. 그럼 GIP에 따르면, 임의의 명제 X, Y에 대해서 다음이 성립한다:

(2a) 만약 모든 $E_i \in \mathbf{E}$에 대해서 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$라면, $p(X|Y)=q(X|Y)$이다.

따라서 임의의 명제 X와 **E**의 임의의 원소 E_j에 대해서도 (2a)가 성립한다. 따라서 다음이 도출된다.

(2b) 만약 모든 $E_i \in \mathbf{E}$에 대해서 $p(XE_i|E_j)=p(X|E_j)p(E_i|E_j)$라면, $p(X|E_j)=q(X|E_j)$이다.

이제 (2b)의 전건, '모든 $E_i \in \mathbf{E}$에 대해서 $p(XE_i|E_j)=p(X|E_j)p(E_i|E_j)$이다'를 생각해보자. 만약 $i=j$라면 $p(XE_i|E_j)=p(X|E_j)p(E_i|E_j)$는 논리적 진리인 $p(X|E_i)=p(X|E_i)$로 변형된다.[14] 만약 $i \neq j$라면 $p(XE_i|E_j)=p(X|E_j)p(E_i|E_j)$는 논리적 진리인 0=0으로 변형된다.[15] 따라서 (2b)의 전건은 논리적 진리이다. 따라서 (2b)는 (2b)의 후건 '$p(X|E_j)=q(X|E_j)$'을 함축한다. 따라서 임의의 명제 X와 **E**의 임의의 원소 E_j에 대해서, 다음이 성립한다: $p(X|E_j)=q(X|E_j)$. 그러므로 GIP는 고정성을 함축한다.

나. 고정성은 GIP를 함축한다

GIP와 고정성의 전건을 가정하자. 즉 모든 $E_i \in \mathbf{E}=\{E_1,\cdots E_n\}$에 대해서 $p(E_i) \neq q(E_i) < 1$라고 가정하자. 그럼 고정성에 따라, 임의의 명제 X에 대해서 다음이 성립한다.

---

14) 확률 계산 규칙에 따르면 $i=j$일 때 $p(XE_i|E_j)=p(XE_i|E_i)=p(X|E_i)$이며 $p(E_i|E_j)=p(E_i|E_i)=1$이다. 따라서 $p(XE_i|E_j)=p(X|E_j)p(E_i|E_j)$는 $p(X|E_i)=p(X|E_i)$로 변형된다.

15) $i \neq j$일 때, E_i와 E_j는 양립불가능하다. 따라서 확률 계산 규칙에 따르면 $i \neq j$일 때, $p(XE_i|E_j)=p(E_i|E_j)=0$이다. 그러므로 $p(XE_i|E_j)=p(X|E_j)p(E_i|E_j)$는 0=0으로 변형된다.

(2c) 모든 $E_i \in \mathbf{E}$에 대해서, $p(X|E_i)=q(X|E_i)$이다.

이제 GIP를 증명하기 위해, (2a)의 전건을 가정하자. 즉 다음이 성립한다.

(2d) 모든 $E_i \in \mathbf{E}$에 대해서 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$이다.

그럼 GIP에 대한 증명은 (2c)과 (2d)로부터 $p(X|Y)=q(X|Y)$를 도출하는 것으로 충분하다. 먼저 확률 계산 규칙에 의해서 다음이 성립한다는 것에 주목하자.

(2e) $q(X|Y) = \dfrac{q(XY)}{q(Y)} = \dfrac{\Sigma_i q(E_i)q(XY|E_i)}{\Sigma_i q(E_i)q(Y|E_i)}$

그리고 임의의 명제에 대해서 성립하는 (2c)에 의해서 (2e)는 다음과 같이 변형된다.

(2f) $q(X|Y) = \dfrac{\Sigma_i q(E_i)p(XY|E_i)}{\Sigma_i q(E_i)p(Y|E_i)} = \dfrac{\Sigma_i q(E_i)p(X|YE_i)p(Y|E_i)}{\Sigma_i q(E_i)p(Y|E_i)}$

한편 임의의 명제 X, Y에 대해서 성립하는 (2d)에 의해서 다음이 성립한다.[16)]

모든 $E_i \in \mathbf{E}$에 대해서 $p(X|YE_i)=p(X|Y)$이다.

따라서 (2f)은 다음과 같이 변형된다.

---

16) 확률 계산 규칙에 따르면 $p(XE_i|Y)=p(X|E_iY)p(E_i|Y)$가 성립한다. 그럼 $p(XE_i|Y)=p(X|Y)p(E_i|Y)$라는 가정에 의해서 $p(X|E_iY)=p(X|Y)$가 성립한다.

$$q(X|Y)=\frac{p(X|Y)\Sigma_i q(E_i)p(Y|E_i)}{\Sigma_i q(E_i)p(Y|E_i)}=p(X|Y).$$

그러므로 (2c)과 (2d)에 의해서 p(X|Y)=q(X|Y)가 도출된다. 따라서 고정성은 GIP를 함축한다.

위의 증명을 통해서 우리는 고정성이 GIP와 동치라는 것을 알게 되었다. 앞에서 언급했듯이 고정성은 JC와 동치이다. 따라서 GIP 역시 JC와 동치이다. 이와 유사한 방식으로, 단순 무관성 원리 SIP가 단순 조건화 SC와 동치라는 것도 어렵지 않게 증명할 수 있다.[17)]

앞에서 언급했듯이 무관성 원리는 특별한 증명이 필요 없을 정도로 당연해 보인다. 따라서 무관성 원리를 정식화하고 있는 GIP가

---

17) 경험 이후에 명제 E를 확신하게 되었다고 하자. 즉 q(E)=1이라고 하자. 그럼 SIP에 따르면, 모든 명제 X, Y에 대해서 다음이 성립한다:

(i) 만약 p(XE|Y)=p(X|Y)p(E|Y)라면, p(X|Y)=q(X|Y)이다.

(i)는 모든 명제에 대해서 성립하기 때문에, Y가 E일 때도 성립한다. 따라서 모든 X에 대해서 다음이 성립한다.

(ii) 만약 p(XE|E)=p(X|E)p(E|E)라면, p(X|E)=q(X|E)이다.

(ii)의 전건은 논리적 진리이다. 따라서 (ii)는 p(X|E)=q(X|E)를 함축한다. q(E)=1이라는 것이 가정되었다. 그럼 확률 계산 규칙에 의해서 p(X|E)=q(X)가 도출된다. 따라서 SIP로부터 SC가 도출된다. 역에 대한 증명도 어렵지 않다. SC를 가정하자. 그리고 경험 이후에 명제 E를 확신하게 되었다고 하자. 그럼 SC에 따르면, 모든 명제 X, Y에 대해서 다음이 성립한다:

(iii) q(X|Y)=p(X|YE).

이제 SIP의 전건을 가정하자. 즉 다음을 가정하자.

(iv) p(XE|Y)=p(X|Y)p(E|Y).

그럼 확률 계산 규칙에 의해서 (iv)로부터 p(X|YE)=p(X|Y)가 도출된다. 따라서 (iii)으로부터 q(X|Y)=p(X|Y)가 도출된다. 그러므로 모든 명제 X, Y에 대해서 다음이 성립한다: 만약 p(XE|Y)=p(X|Y)p(E|Y)라면, p(X|Y)=q(X|Y)이다. 이것이 바로 SIP이다.

JC와 동치라는 것은 JC에 대한 정당화 논변으로 간주될 수도 있을 것이다. 사실 JC 자체가 의미하는 바가 무엇인지 직관적으로 파악하기란 그리 쉽지 않다. 하지만 GIP를 염두에 둔다면, JC를 좀 더 직관적인 수준에서 이해할 수 있게 된다. 즉 우리는 JC가 아주 직관적인 원리, 즉 경험과 무관한 믿음은 바뀌지 말아야 한다는 원리일 뿐이라는 것을 알게 되었다. JC를 따르면 우리의 부분적 믿음의 변화는 그 원리를 만족할 것이며, JC를 따르지 않는다면 우리의 부분적 믿음의 변화는 그 원리를 위반하게 될 것이다. 이보다 더 직관적인 방식으로 JC를 옹호할 수 있을까?[18]

물론, 무관성 원리에 대한 정당화를 요구하는 독자가 있을 수도 있다. 심지어 무관성 원리의 직관적인 면을 부정하고 그 원리보다 더 직관적인 원리를 제시하라고 요구하는 독자가 있을 수도 있다. 그러나 필자는 그 원리보다 더 직관적인 원리가 무엇이 있을지 의심스럽다. 따라서 필자는 무관성 원리를 정당화하기 위해 더욱더 직관적인 원리를 도입하지는 않을 것이다. 그 대신 필자는 무관성 원리를 위반하는 경우 베이즈주의 인식론이 어떤 문제에 봉착하게 되는 지 보여줄 것이다.

18) 나는 여기에서 익명의 심사위원 중에 한 분이 지적한 '인식적 보수주의(epistemic conservatism)'에 대해서 언급할 필요가 있다. 심사위원의 지적대로, 크리스텐슨(Christensen, 1994)은 베이즈주의 인식론이 '인식적 보수주의'를 함의하고 있다고 말한다. 그리고 이러한 보수성과 관련된 사실로 조건화가 고정성과 동치라는 것, 즉 조건화를 이용해서 믿음을 수정하는 경우에 특정한 조건부 믿음의 정도가 수정되지 않는다는 것이 언급된다. 이와 관련해, 본 논문은 그렇게 수정되지 않는 믿음의 정도들에는 무엇이 있는 지를 좀 더 분명히 했다고 할 수 있으며, 이에 베이즈주의의 보수적 성격을 규명하는 데 있어 본 논문이 도움이 될 것이라고 생각한다. 이런 사실을 알려주신 심사위원에게 감사의 말씀을 전한다.

3. 베이즈주의 입증 이론

베이즈주의 인식론은 가설과 증거 사이의 관계를 다루는 입증의 문제들에 대해서 몇몇 흥미로운 해결책들을 제시해왔다. 필자는 아래 4절에서 무관성 원리가 위반되는 경우 베이즈주의 입증 이론이 직면하게 될 문제점을 제시할 것이다. 그 전에 베이즈주의 입증 이론과 그 지지자들이 받아들일 수밖에 없는 몇 가지 가정들을 살펴보도록 하자.

1) 조건화가 가정되지 않은 베이즈주의 입증 이론

본격적인 논의에 앞서, 베이즈주의 입증 이론의 기본적인 특징을 생각해보자. 흔히 베이즈주의 입증 이론(Bayesian Confirmation Theory)은 다음과 같이 제시된다.[19)]

BCT 만약 $p(H)<p(H|E)$이면 E는 H를 입증한다.
만약 $p(H)=p(H|E)$이면 E는 H에 중립적이다.
만약 $p(H)>p(H|E)$이면 E는 H를 반입증한다.

---

19) 일반적으로 증거 E와 가설 H 사이의 입증 관계는 배경지식에 대해서 상대적이다. 하지만 아래 BCT에는 이 배경지식에 대한 고려가 포함되어 있지 않았다. 배경지식을 고려할 때 베이즈주의 입증 이론은 다음과 같이 나타낼 수 있다.

만약 $p(H|K)<p(H|EK)$이면 E는 배경지식 K에 상대적으로 H를 입증한다.
만약 $p(H|K)=p(H|EK)$이면 E는 배경지식 K에 상대적으로 H에 중립적이다.
만약 $p(H|K)>p(H|EK)$이면 E는 배경지식 K에 상대적으로 H를 반입증한다.

배경지식은 입증이론의 여러 주제에서 중요한 역할을 하지만 본 논문의 맥락에서는 별로 큰 역할을 하지 못한다. 따라서 본 논문은 배경지식을 고려하지 않는다.

이 이론은 왜 ‘베이즈주의’ 입증 이론이라고 불리는가? 이 질문에 답하기 위해서는 정량적인 방식으로 증거 E와 가설 H 사이의 관계를 규명하려는 시도들이 가진 일반적인 특징들을 생각해볼 필요가 있다. (베이즈주의 입증 이론을 포함하여) 정량적인 입증 이론(Quantitative Confirmation Theory, QCT)의 기본적인 생각은 E를 배우기 전의 H에 대한 믿음의 정도와 E를 배운 이후의 H에 대한 믿음의 정도 사이의 비교를 통해서 E와 H 사이의 증거적 관계가 규명될 수 있다는 것이다. B(H)가 E를 배우기 전 H에 대한 믿음의 정도를 나타낸다고 하자. 그리고 $B_E(H)$는 E를 배운 이후의 H에 대한 믿음의 정도를 나타낸다고 하자. 그럼 정량적인 입증 이론의 기본 생각은 다음과 같이 나타낼 수 있다.

QCT 만약 $B(H)<B_E(H)$이면 E는 H를 입증한다.
만약 $B(H)=B_E(H)$이면 E는 H에 중립적이다.
만약 $B(H)>B_E(H)$이면 E는 H를 반입증한다.

앞에서 언급했듯이 베이즈주의자들은 우리의 믿음의 정도는 확률 계산 규칙을 따라야 한다고 주장한다. 즉 그들은 우리의 부분적 믿음이 확률 함수에 의해서 표상될 수 있다고 생각한다. 따라서 베이즈주의자들에게 B와 B_E는 모두 확률 함수가 될 것이다. 만약 p가 E를 배우기 전 믿음의 정도를 나타내는 확률 함수이고, q가 E를 배운 이후의 믿음의 정도를 나타내는 확률함수라면, 베이즈주의자들은 QCT를 다음과 같이 변형할 것이다.

CT 만약 $p(H)<q(H)$이면 E는 H를 입증한다.
만약 $p(H)=q(H)$이면 E는 H에 중립적이다.
만약 $p(H)>q(H)$이면 E는 H를 반입증한다.

이제 전형적인 베이즈주의자들처럼 조건화가 합리적인 믿음 갱신 규칙이라고 가정해보자. CT에 나타나는 q는 E를 배운 이후의 부분적 믿음 함수이다. 그럼 조건화, 특히 SC에 의해서 q(H)=p(H|E)가 성립한다. 이제 CT는 BCT로 변형된다. 정리하자면, 베이즈주의 입증 이론 BCT는 일반적인 정량적 입증 이론 QCT에 두 개의 가정이 추가된 이론이다. 첫 번째는 우리의 믿음의 정도는 확률 함수를 이용해서 표상되어야 한다는 것이며, 두 번째는 우리의 믿음의 정도의 변화는 조건화 규칙을 통해서 이루어져야 한다는 것이다. 이런 가정이 추가되었다는 것이 바로 BCT가 '베이즈주의' 입증 이론이라고 불리는 이유다.

앞서 언급했듯이 이 논문은 두 번째 가정만을 다루고 있다. 필자는 우리의 믿음의 정도가 확률 함수에 의해서 표상되어야 한다는 것을 받아들인다. 이 글이 다루고 있는 것은 믿음의 정도가 조건화를 통해서 수정되어야 하는지 여부이다. 따라서 조건화가 가정된 BCT를 탐구의 대상으로 삼는다면 필자의 논의는 순환성에 빠질 수 있다. 이미 조건화가 가정된 BCT가 조건화를 받아들여야 한다고 말하는 것은 동어반복에 불과하다. 이런 이유에서 필자는 믿음의 정도는 확률 함수를 이용해 표상되어야 한다는 것은 가정되지만 조건화는 가정되지 않은 입증 이론, 즉 CT가 어떤 문제에 직면하게 되는 지 살펴볼 것이다.

2) 확률-입증 논제

그럼 먼저 몇 가지를 추가로 가정해보자. 앞서 언급했듯이 조건화가 합리적 믿음 갱신 규칙이라는 것을 받아들이든 그렇지 않든지 간에, 베이즈주의 입증 이론가들은 확률이 우리의 믿음의 정도를 표상한다는 것을 받아들인다. 당연히 그들은 확률에 의해서 가설과 증거 사이의 입증 관계가 규명된다고 생각할 것이다. "확률에 의해

서 가설과 증거 사이의 입증 관계가 규명된다."는 말은 무엇을 의미하는가? 여러 가지 의미가 있을 수 있겠지만, 이 말은 최소한 "어떤 두 명제 사이의 확률적 관계가 다른 두 명제 사이의 확률적 관계와 같다면 앞의 두 명제 사이의 입증 관계는 뒤의 두 명제 사이의 입증 관계와 같다."는 것을 함축하는 것 같다. 다음 두 사례를 생각해보자.

> **사례 1.** 여기 어떤 항아리가 있다. 이 항아리에는 100개의 공이 있다. 공의 분포에 대한 두 가지 가설이 있다. 첫 번째 가설 A는 100개의 공이 모두 검은 색이라는 것이고, 두 번째 가설 A*는 50개는 검은 색, 50개는 흰 공이라는 것이다.
>
> **사례 2.** 여기 어떤 친목모임이 있다. 이 친목 모임의 회원들은 100명이다. 회원들의 직업에 대한 두 가지 가설이 있다. 첫 번째 가설 X는 100명이 모두 의사라는 것이고, 두 번째 가설 X*는 50명은 의사, 50명은 간호사라는 것이다.

사례 1과 사례 2는 무척 유사하다. 사례 1의 공은 사례 2의 회원, 사례 1의 검은 색은 사례 2의 의사, 사례 1의 흰 색은 사례 2의 간호사와 대응한다. 뿐만 아니라 사례 1의 공의 색깔에 대한 가설은 사례 2의 회원들의 직업에 대한 가설에 대응한다. 따라서 가설 A와 항아리에서 뽑은 하나의 공이 검은 색이라는 증거 사이의 확률적 관계는 가설 X와 조사된 한 회원이 의사라는 증거 사이의 확률적 관계와 일치한다. 만약 베이즈주의자들처럼 확률적 관계를 이용해서 입증 관계를 규명하고자 한다면, 가설 A와 항아리에서 뽑은 하나의 공이 검은 색이라는 증거 사이의 입증 관계는 가설 X와 조사된 한 회원이 의사라는 증거 사이의 입증 관계와 같다고

말해야 한다는 것은 지극히 당연해 보인다.[20] 정리하자면, 확률을 이용해서 가설과 증거 사이의 입증 관계를 규명하고자 한다면 최소한 다음 확률-입증 논제를 받아들여야 한다.

확률-입증 논제 (Probability-Confirmation Thesis, PC)
임의의 네 명제 A, B, X, Y에 대해서, A와 B 사이의 확률적 관계가 X와 Y 사이의 확률적 관계와 같다면 A와 B 사이의 입증 관계는 X와 Y 사이의 입증 관계와 같다.

이 논제를 분명히 하기 위해서 우리는 두 가지 동일성을 정의해야 한다. 첫 번째는 확률적 관계의 동일성이고, 두 번째는 입증 관계의 동일성이다. 우선 확률적 관계의 동일성을 먼저 생각해보자. A와 B 사이에 성립하는 확률적 관계는 하나의 명제로 서술될 수 있을 것이다. 예를 들어 다음 명제가 그러한 명제라고 가정하자.

(3a) A의 확률은 B의 확률보다 크다.

이 명제는 A와 B 사이에 성립하는 확률적 관계를 표현하고 있다. 그럼 다음 명제를 생각해보자.

(3b) ~X의 확률은 ~Y의 확률보다 작다.

---

20) 익명의 심사위원 중에 한 분이 지적하신 것과 같이, 필자는 확률 관계가 같다고 하더라도 항상 같은 의사 결정을 내리는 것은 아니라고 생각한다. 그러나 어떤 가설을 채택하고 거부하는 의사 결정과 상관없이, 증거가 가설을 입증하는지 여부는 확률에 의해서 결정될 수 있다고 추측한다. 증거와 가설 사이의 확률적 관계만으로는 가설의 (채택 여부가 아니라) 입증 여부를 온전히 결정할 수 없다고 생각하는 베이즈주의자가 있을까? 현재 필자는 이 질문에 대해서 부정적인 입장을 가지고 있지만 좀 더 숙고할 필요가 있다. 이런 문제를 지적해주신 심사위원에게 감사의 인사를 드린다.

X와 Y 사이의 확률적 관계는 이 명제에 의해서 서술된다고 가정하자. 여기서 (3a)에 등장하는 A와 B 각각을 X와 Y로 대체한다면 우리는 다음 명제를 만들 수 있다:

(3a') X의 확률은 Y의 확률보다 크다.

이제 (3a')과 (3b)를 비교하자. 간단한 수학 계산을 통해서 (3b)와 (3a')이 동치라는 것을 확인할 수 있다. 나는 이런 경우에 A와 B 사이에 성립하는 확률적 관계는 X와 Y 사이에 성립하는 확률적 관계와 같다고 말할 것이다. 요약하자면, 확률적 관계의 동일성은 다음과 같이 정의된다.

> **확률적 관계의 동일성.**
> A와 B 사이의 확률적 관계가 X와 Y 사이의 확률적 관계와 같은 경우, 그리고 그런 경우에만 다음이 성립한다: A와 B 사이의 확률적 관계를 기술하는 명제에 등장하는 모든 A와 B 각각을 X와 Y로 바꾸어 만들어진 명제는 X와 Y 사이의 확률적 관계를 기술하는 명제와 동치이다.

이제 입증 관계의 동일성을 생각해보자. 필자는 임의의 명제 A와 B 사이에 성립하는 입증 관계를 다음 6가지로 제한한다.

A가 B를 입증한다; A가 B에 중립적이다; A가 B를 반입증한다;
B가 A를 입증한다; B가 A에 중립적이다; B가 A를 반입증한다.

그럼 쉽게 입증 관계의 동일성을 파악할 수 있다. 가령, A가 B를 입증하지만 X는 Y에 중립적인 경우, A와 B 사이의 입증 관계

는 X와 Y 사이의 입증 관계와 다르다고 말해야 한다. 이제 앞에서 제시한 확률-입증 논제, 즉 PC의 의미는 분명해졌다. 확률을 이용해서 입증의 문제를 해결하려는 베이지주의자들이 PC를 거부하는 것은 쉽지 않을 듯하다. 마지막으로 입증 이론의 한 가지 특징을 더 고려해보자.

3) 가설 대칭성 논제

무언가를 경험한 이후에 H에 대한 믿음의 정도가 증가했다고 가정해보자. 그리고 믿음의 정도는 확률 계산 규칙을 만족한다고 생각하자. 확률 계산 규칙에 따라서 H에 대한 믿음의 정도와 ~H에 대한 믿음의 정도의 합은 1로 일정하다. 따라서 H에 대한 믿음의 정도가 증가했을 때, ~H에 대한 믿음의 정도는 감소할 수밖에 없다.

이제 어떤 증거 E가 H를 입증한다고 하자. 즉 E를 배우기 전 H에 대한 믿음의 정도보다 E를 배운 이후 H에 대한 믿음의 정도가 더 크다고 하자. 그럼 당연히 E를 배우기 전 ~H에 대한 믿음의 정도보다 E를 배운 이후 ~H에 대한 믿음의 정도는 더 작을 수밖에 없다. 달리 말하자면 다음 가설 대칭성 논제는 무척 직관적인 듯이 보인다.

> **가설 대칭성 논제 (Hypothesis Symmetry Thesis, HS).**
> E가 H를 입증한다면, 그리고 그런 경우에만 E는 ~H를 반입증한다.

입증 이론가들은 가설과 증거 사이에 성립할 수 있는 다양한 종류의 대칭성에 대해서 연구를 해왔다. 가령, (i) E가 H를 입증한다면 ~E가 H를 반입증한다, (ii) E가 H를 입증한다면 H가 E를 입증한다, (iii) E가 H를 입증한다면 ~E가 ~H를 입증한다 등이 대표적

이다.[21] 물론 입증 이론가들이 이러한 대칭성을 모두 받아들이는 것은 아니다. 특히 엘러리 엘스(Ellery Eells)와 브랜든 화이텔슨(Branden Fitelson)은 방금 언급한 (i)-(iii)을 거부한다. 하지만 그들은 HS에 대해서는 긍정적이다. HS에 대한 긍정적인 입장은 정량적인 입증 이론에만 국한된 것은 아니다. 예를 들어, 가설-연역적 입증 이론을 제시한 칼 헴펠(Carl Hempel) 역시 HS를 받아들인다.[22] HS에 대해서 필자는 이들과 의견을 같이 한다. 특히 우리의 믿음의 정도가 확률에 의해서 표상된다고 생각하는 베이지주의자들은 HS를 반드시 받아들여야 한다고 생각한다. 왜냐하면 앞에서 말한 것처럼 우리의 믿음의 정도가 확률에 의해서 표상된다면, H에 대한 믿음의 정도가 증가했을 때 ~H에 대한 믿음의 정도는 감소할 수밖에 없기 때문이다.

이 절에서 필자는 세 가지를 도입했다. 첫 번째는 조건화가 가정되지 않은 베이즈주의 입증 이론 CT이며, 두 번째는 확률과 입증 사이의 관계를 나타내는 확률-입증 논제 PC이다. 그리고 마지막 세 번째는 가설과 증거 사이의 대칭성을 나타내는 가설 대칭성 논제 HS이다. 다음 절에서 필자는 CT, PC, HS를 모두 받아들이지만 무관성 원리를 받아들이지 않으면 모순에 직면한다는 것을 보일 것이다. 필자의 논증이 설득력이 있다면, 이 논증은 무관성 원리와 동치인 조건화에 대한 옹호 논증으로 간주될 수 있을 것이다.

---

21) 이런 대칭성과 관련된 논의를 위해서는 Carnap (1962), §67, Eells and Fitelson (2001)를 보라. 위의 '가설 대칭성 (Hypothesis Symmetry)'이라는 이름은 Eells and Fitelson (2001)에서 가져온 것이다.

22) Hempel (1965), p.37을 보라. 같은 내용을 Hempel (1965)의 한글 번역본 p. 73에서 찾을 수 있다.

4. 베이즈주의 입증 이론과 무관성 원리

우리의 목표는 무관성 원리와 3절에서 논의한 CT, PC, HS 사이의 관계를 파악하는 것이다. 특히, 필자는 무관성 원리가 위반된다면 CT, PC, HS는 모순을 함축한다는 것을 보일 것이다. 이를 위해서 먼저 입증 이론과 관련해서 고려해야 할 무관성 원리는 무엇인지 살펴보자.

3.1절에서 제시된 정량적 입증 이론들은 모두 E를 배우기 전의 H에 대한 믿음의 정도와 E를 배운 이후의 H에 대한 믿음의 정도를 비교하고 있다. 따라서 그러한 입증 이론과 무관성 원리 사이의 관계를 고려하기 위해서는 경험에 의해서 E를 배운 경우, 즉 경험에 의해서 E를 확신하게 된 경우를 살펴보아야 한다. 또한 3.1절에서 제시된 정량적 입증 이론들은 모두 H에 대한 *비조건부* 믿음의 정도를 고려하고 있다. 따라서 우리는 H에 대한 비조건부 믿음의 정도와 관련된 무관성 원리를 살펴보아야 한다. 경험에 의해서 E를 확신하였을 때 단순 무관성 원리 SIP는 H에 대한 비조건부 믿음의 정도에 대해서 다음을 함축한다.

(4a) 만약 $p(HE)=p(H)p(E)$라면 $p(H)=q(H)$이다.

바로 이 (4a)는 IP의 후건이다. 그리고 이것이 입증 이론과 관련해서 우리가 고려해야 하는 무관성 원리이다. 이제 (4a)와 3절에서 논의된 CT, PC, HS 사이의 관계를 파악해보자. 즉, (4a)가 위반되었을 때, CT, PC, HS가 어떤 문제에 직면하게 될 지 생각해보자.

일단 (4a)의 부정을 가정해보자. 즉 $p(HE)=p(H)p(E)$이지만 $p(H) \neq q(H)$라고 가정하자. 그럼 두 가지 경우가 가능하다. 첫 번째는 $p(H)<q(H)$인 경우이고, 두 번째는 $p(H)>q(H)$인 경우이다. 첫 번째

경우를 먼저 생각해보자. CT에 따르면 $p(H)<q(H)$일 때 E는 H를 입증한다. 따라서 우린 첫 번째 경우 E는 H를 입증한다고 말해야 한다. 여기서 우리는 $p(HE)=p(H)p(E)$가 $p(\sim HE)=p(\sim H)p(E)$와 동치라는 것에 주목해야 한다. 따라서 H와 E 사이의 확률적 관계는 ~H와 E 사이의 확률적 관계와 같다. 그럼 PC에 의해서 H와 E 사이의 입증 관계는 ~H와 E 사이의 입증 관계와 같다. 앞에서 E가 H를 입증한다고 했다. 따라서 E도 ~H를 입증한다. 그러나 이것은 가설 대칭성 논제 HS와 충돌한다. 따라서 $p(HE)=p(H)p(E)$이지만 $p(H)<q(H)$인 경우 모순이 발생한다. 두 번째로 $p(H)>q(H)$인 경우도 생각해보자. 이 경우도 마찬가지다. $p(H)>q(H)$이기 때문에 E는 H를 반입증한다. 그리고 $p(HE)=p(H)p(E)$가 $p(\sim HE)=p(\sim H)p(E)$와 동치이고, 따라서 H와 E 사이의 확률적 관계는 ~H와 E 사이의 확률적 관계와 같다. 그럼 PC에 의해서 H와 E 사이의 입증 관계는 ~H와 E 사이의 입증 관계와 같다. 그러므로 E는 ~H를 반입증한다. 하지만 이것은 HS와 충돌한다. 따라서 $p(HE)=p(H)p(E)$이지만 $p(H)>q(H)$인 경우 모순이 발생한다. 결국 (4a)의 부정, CT, PC, HS는 모순을 함축한다. 따라서 CT, PC, HS를 받아들이는 베이즈주의자들은 반드시 (4a)를 받아들여야 한다. (4a)는 IP의 후건이라는 것을 기억하자. 그리고 CT, PC, HS는 모두 경험에 의해서 E를 배우는 경우만, 즉 IP의 전건이 성립하는 경우만을 고려하고 있다는 것도 기억하자. 그럼 우리는 CT, PC, HS를 받아들이는 베이즈주의자들은 반드시 IP도 받아들여야 한다고 결론내릴 수 있다.

그러나 여기서 주의할 것이 있다. 단순 무관성 원리 SIP는 IP를 함축한다. 그렇지만 그 역은 성립하지 않는다. 그래서 CT, PC, HS를 받아들이는 베이즈주의자들은 반드시 IP를 받아들여야 한다는 결론으로부터 CT, PC, HS를 받아들이는 베이즈주의자들은 반드시 SIP를 받아들여야 한다는 결론이 연역적으로 도출되는 것은 아니

다. 하지만 앞에서 언급했듯이 그 본성을 고려할 때, SIP가 말하는 바는 IP가 말하는 바와 같다. 이 둘은 모두 경험과 관련이 없는 명제에 대한 믿음의 정도는 경험 이후에도 그대로 유지되어야 한다고 말하고 있다. SIP는 조건부 믿음의 정도와 비조건부 믿음의 정도 모두를 포괄하지만 IP는 비조건부 믿음의 정도만 다룰 수 있을 뿐이다. 그 둘의 차이는 단지 적용 범위의 문제일 뿐이다. 따라서 위의 결론으로부터 CT, PC, HS를 받아들이는 베이즈주의자들은 SIP를 반드시 받아들여야 한다는 결론이 연역적으로 도출되지 않지만 어느 정도 옹호(vindicate)될 수는 있다. 마찬가지로 SIP와 동치인 단순 조건화 SC에 대해서도 비슷한 결론을 내릴 수 있다. 즉 위 논증을 통해서 CT, PC, HS를 받아들이는 베이즈주의자들은 SC를 받아들여야 한다는 결론이 옹호될 수 있다.

5. 결론을 대신하여

본 논문에서 필자는 베이즈주의 인식론의 핵심적인 요소 중에 하나인 조건화를 옹호하는 논변을 제시했다. 이 목적을 위해서 그동안 잘 알려지지 않았던 사실, 즉 무관성 원리와 조건화가 동치라는 것을 증명했다. 단순 무관성 원리는 단순 조건화와 동치였으며, 일반 무관성 원리는 제프리 조건화와 동치였다. 그리고 이러한 무관성 원리의 한 가지 귀결, 즉 IP가 만족되지 않을 경우, 베이즈주의 입증 이론은 심각한 문제에 직면할 수 있다는 것을 보였다. 특히, 베이즈주의자들이라면 누구나 받아들일 듯이 보이는 가정들－CT, PC, HS－로부터 IP가 도출된다는 것을 증명하였다. 이를 이용해서 필자는 무관성 원리와 조건화가 옹호될 수 있다고 주장하였다.

합리적인 믿음의 정도 갱신 규칙으로 조건화가 처음 제시된 것

은 1920년대이다. 그 후 1960년대에 JC가 제안되었다.[23] 한편 조건화를 정당화하는 작업은 70년대에 본격적으로 시작되었다. 그 후 80년대와 90년대 그리고 2010년대 초반인 지금까지도 조건화에 대한 새로운 정당화 논변이 꾸준히 제시되고 있다.[24] 조건화가 처음 등장한 이후 짧지 않은 시간이 지났고 적지 않은 수의 정당화 논변이 제안되었음에도 불구하고, 여전히 연구가 진행되고 있다. 이런 사실은 조건화에 대한 합의된 정당화가 아직 마련되지 않았으며 이런 연구가 여전히 철학적으로 유의미하다는 것에 대한 방증이라고 할 수 있다.

기존 시도들은 대체적으로 하나의 문제점을 공유한다. 그 문제점은 각 정당화가 의존하는 각각의 특별한 원리들 때문에 발생한다. 예를 들어 생각해보자. P. M. 윌리암스(P. M. Williams)는 최소 정보 원리(Principle of Minimal Information)을 제안하고 그것으로부터 SC와 JC를 도출함으로써 조건화를 정당화한다. 여기서 문제는 과연 최소 정보 원리가 합리적이냐는 것이다. 거칠게 말해서, 이 원리는 행위자가 믿음의 정도를 수정할 때 기존 믿음의 정도에서 가장 덜 벗어나야 한다는 것이다. 이 원리는 합리적인가? 이 질문

---

23) 조건화를 명시적인 믿음 갱신 규칙으로 제시한 사람은 프랭크 램지(Frank Ramsey)이다. Ramsey (1926)을 보라. 그리고 JC는 1960년대 처음으로 출판된 Jeffrey (1983)에서 제시되었다.

24) 조건화를 정당화하는 논문들은 무척 많다. 몇 가지 논증들은 다음 참고문헌에서 찾을 수 있다. 더치 전략 논증(Dutch Strategy Argument)을 위해서는 Teller (1973), Armendt (1980), Skyrms (1987)을, 대칭성 논증(Symmetry Argument)을 위해서는 Hughes and van Fraassen (1984), van Fraassen (1986; 1989)을, 반영 원칙으로부터의 논증(Argument from the Principle of Reflection)을 위해서는 van Fraassen (1999)와 Maher (1993)를, 최소 정보 원칙으로부터의 논증(Argument from the Principle of Minimal Information)을 위해서는 Williams(1980)을, 인식적 기대 효용 최대화 원칙으로부터의 논증(Argument from the Principle of Maximization of Expected Utility)을 위해서는 Greaves and Wallace(2006)를 보라.

에 대한 답이 제시되지 않는다면, 윌리암스의 정당화는 한계가 있다. 몇몇 다른 정당화들도 비슷하다. 특별한 원리를 세우고 그 원리로부터 조건화를 도출한다. 그러나 그 원리의 합리성에 대한 추가 논변이 제시되지 않는다면 정당화 역시 만족스럽지 않을 것이다.

그럼 본 논문에서 제시된 무관성 원리는 어떠한가? 이 원리 역시 조건화를 함축한다. 이 원리들은 조건화를 함축하는 다른 원리들만큼 정당화가 필요한가? 일단 필자가 보였듯이 무관성 원리는 입증이라는 개념을 이용해서 옹호될 수 있다는 사실을 지적할 필요가 있다. 또한 무관성 원리와 조건화를 함축하는 다른 원리들 사이의 차이도 생각해보아야 한다. 특히, 무관성 원리의 이론적 위치를 생각해볼 필요가 있다. 앞에서 언급한 최소 정보 원리는 조건화를 함축하지만 그 역은 성립하지 않는다. 하지만 필자가 제시한 무관성 원리는 조건화와 동치이다. 따라서 최소 정보 원리가 만약 정당화된다면, 무관성 원리와 조건화가 모두 정당화되는 것이다. 조건화를 함축하는 다른 원리들도 마찬가지다. 각 원리들이 정당화된다면, 그것이 함축하는 조건화와 그 동치 무관성 원리가 모두 정당화된다. 따라서 다음과 같은 결론을 내릴 수 있다: 무관성 원리는 적어도 조건화를 함축하는 다른 원리들만큼은 정당화될 수 있다.

참고 문헌

Armendt, B. (1980), "Is There a Dutch Book Argument for Probability Kinematics?", *Philosophy of Science*, 47, pp. 583-589.

Bradley, R. (2005), "Radical Probabilism and Bayesian Conditioning", *Philosophy of Science*, 72, pp. 342-364.

Christensen, D. (1994), "Conservatism in Epistemology", *Noûs*, 28, pp. 69-89.

Carnap, R. (1962), *Logical Foundations of Probability*, Second ed., Chicago: University of Chicago Press.

Dempster, A. P. (1967), "Upper and lower probabilities induced by a multivalued mapping", *Annals of Mathematical Statistics*, 38: pp. 325-339.

Diaconis, P. and Zabell, S. (1982), "Updating Subjective Probability", *Journal of the American Statistical Association*, 77, pp. 822-830.

Douven, I. and Romeijn, J. (2011), "New Resolution of the Judy Benjamin Problem", *Mind*, 120, pp. 637-670.

Eells, E. and Fitelson, B. (2002), "Symmetries and Asymmetries in Evidential Support", *Philosophical Studies*, 107, pp. 129-142.

Goldman, A. and Witcomb, D. (eds.) (2011), *Social Epistemology: Essential Readings*, Oxford Unversity Press.

Greaves, H. and Wallace, D., (2006), "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility" *Mind*, 115, pp.607-632.

Grove, A. and Halpern, J. (1997), "Probability Update: Conditioning vs. Cross-Entropy", *Proceedings of the Thirteenth Annual*

Conference on Uncertainty in Artificial Intelligence, Providence, Rhode Island, August, pp. 1-3.

Hájek, A. and Hartmann S. (2010), "Bayesian Epistemology", in J. Dancy et al. (eds.), *The Companion to Epistemology*, Blackwell Press, pp.93-105.

Harper, W. and Kyburg, H. (1968), "The Jones Case", *The British Journal for the Philosophy of Science*, 19, pp. 247-251.

Hempel, C. (1965), *Aspects of Scientific Explanation and other Essays in the Philosophy of Science,* The Free Press. (번역본: 칼 구스타프 헴펠 지음, 전영삼 · 여영서 · 이영의 · 최원배 옮김 (2011) 『과학적 설명의 여러 측면: 그리고 과학철학에 관한 다른 논문들』, 서울: 나남.)

Howson, C. and Urbach, P. (1993), *Scientific Reasoning: The Bayesian Approach, 2nd ed*. Open Court.

Hughes, R.I.G and van Fraassen, B. (1984), "Symmetry Arguments in Probability Kinematics", *Proceedings of the 1984 Biennial Meetings of the Philosophy of Science Association*, pp. 851-869.

Jeffrey, R. (1983), *The Logic of Decision*, University of Chicago Press.

Jeffrey, R. (1992), *Probability and the Art of Judgment*, Cambridge University press.

Jeffrey, R. (2004), *Subjective Probability: The real thing,* Cambridge University Press.

Levi, I. (1967), "Probability Kinematics", *The British Journal for the Philosophy of Science*, 18, pp. 197-209.

Maher, P. (1993), *Betting on Theories*, Cambridge University Press.

Park, Ilho (forthcoming), "Simultaneous Belief Updates via Successive Jeffrey Conditionalization", *Synthese*, Doi:

10.1007/s11229-012-0207-7.
Ramsey, F. (1926), “Truth and Probability” in Henry K. Jr. and Howard S., (eds.), (1980) *Studies in Subjective Probability,* New York: Krieger Publishing, pp.23-52.
Shackle, G. (1949), *Expectation in Economics*, Cambridge University Press.
Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton University Press.
Skyrms, B. (1987), “Dynamic Coherence and Probability Kinematics”, *Philosophy of Science*, 54, pp. 1-20.
Teller, Paul, (1973), “Conditionalization and Observation”, *Synthese*, 26, pp. 218-258.
van Fraassen, B. (1981), “A Problem for Relative Information Minimizers in Probabilistic Kinematics”, *British Journal for the Philosophy of Science*, 32, pp. 375-379.
van Fraassen, B. (1986), “A Demonstration of the Jeffrey Conditionalization Rule”, *Erkenntnis*, 24, pp. 17-24.
van Fraassen, B. (1989), *Laws and Symmetry*. Oxford: Clarendon Press.
van Fraassen, B. (1999), “Conditionalization, A New Argument for”, *Topoi*, 18, pp. 93-96.
White, R. (2006), “Problems for Dogmatism”, *Philosophical Studies* 131, pp. 525-57.
Zadeh, L. A. (1978), “Fuzzy sets as a basis for a theory of possibility”, *Fuzzy Sets and Systems*, 1, pp. 3-28.

경희대학교
Kyung Hee University
ipark.phil@gmail.com

Conditionalization and Confirmation: A Vindication of Conditionalization

Ilho Park

The main objective of this paper is to vindicate the Bayesian belief updating rule, i.e. conditionalization. For this purpose, I introduce first what I call *Irrelevance Principle*, and show that this principle is equivalent to conditionalization. In turn, the principle is vindicated by means of Bayesian confirmation theory. That is, I suggest some theses that Bayesian confirmation theorists should accept, and prove that if Irrelevance Principle is violated, the theses cannot holds.

Key Word: Baysianism, Conditionalization, Bayesian Confirmation theory, Irrelevance principle