확률 개념은 빈도 개념으로 환원될 수 있는가?

선우환 (서울시립대)

【요약문】 이 논문의 목적은 확률에 대한 빈도 해석(특히 라이헨바하와 새먼에 의해 체계화된 형태)을 비판적으로 고찰하는 데에 있다. 보다 구체적으로, 확률 빈도 해석에 대해 기존에 제기된 반론들을 보다 강화하고 새로운 추가적 반론들을 제기함으로써, 확률 빈도 해석이 종래에 생각되어 왔던 것보다 훨씬 더 심각한 상태에 있음을 보이려 한다. 특히, 빈도 해석이 흔히 여겨지는 바와 달리, 확률 값을 부여하기 위해 여러 국면들에 있어서의 주관적이고 임의적인 선택들에 의존해야 한다는 것이 이 논문에서의 반론들을 통해 드러날 것이다

【주요어】 확률, 빈도 해석, 라이헨바하, 새먼

I. 들어가는 말

확률 해석(probability interpretation)이 하는 일은 하나의 공리 체계인 확률 계산(probability calculus)에 구체적 의미를 부여하는 것이다. 이는 우리의 직관적 확률 개념을 해명(explicate)하는 일이기도 하다. 확률에 대한 빈도 해석(frequency interpretation)은 그 중 가장 널리 받아들여져 온 해석 중의 하나이다. 확률 개념을 상대빈도(의 극한) 개념으로 환원하려고 하는 이 해석은, 확률의 객관성과 비임의성, '삶의 지침'으로서의 역할 등을 잘 설명한다고 여겨지는 경우가 많았다. 뿐만 아니라, 빈도 해석은 확률 진술의 유의미성을 경험주의적으로 이해할 수 있는 유일한 방식으로 여겨지기도 한다.

이 글이 보이고자 하는 것은, 그러한 '장점'들이 단지 의도된 장점들이

며, 실제 빈도 해석은 다른 확률들을 전제해야 하거나 많은 자의적 선택들을 요구한다는 것이다. 또한 그 해석은 우리의 기본적 확률 개념 중 많은 부분들을 설명하지 못한다. 빈도해석에 대해 이미 여러 비판 논변들이 제기된 바 있지만, 이 논문에서 하려는 논의들은 그런 논변들을 더 강화하고 새로운 논변들을 덧붙이는 것을 목표하고 있다. 특히 라이헨바하(Reichenbach)와 새먼(Salmon)에 의해 체계화된 빈도주의를 범례화된 대상으로 해서 논의를 하겠다.

II. 확률에 대한 빈도해석

빈도 해석은 일반적으로 확률을, 사건들의 무한계열(infinite sequence)에 있어서 어떤 속성들이 나타나는 상대빈도(relative frequency)의 극한치라고 정의한다[1]: 우선, 상대빈도 $F^n(B/A)$는

$$F^n(B/A) = \frac{N^n(A \cap B)}{N^n(A)}$$

라고 정의된다. (단, 계열에서 n번째 항까지의 사건들 중 집합 A에 속하는 사건들의 갯수를 $N^n(A)$라고 하고 n까지 중 집합 A와 B에 모두 속하는 사건들의 갯수를 $N^n(A \cap B)$라 한다.) 여기서 흔히 A는 준거집합(reference class), B는 속성집합(attribute class)이라고 불린다. 그러면 확률은

$$P(B/A) = \lim_{n \to \infty} F^n(B/A)$$

라고 정의된다. 여기서는 조건부 확률이 확률 함수의 원초적 형태를 지니게 된다. 또한 확률함수의 두 유입값(input)은 사건 집합이다.

---

1) Reichenbach (1949) p. 68, Salmon (1966) p. 83

확률을 상대빈도 계열의 '극한'으로 취하는데 대해 기본적으로 제기되는 비판(예를 들어 막스 블랙에 의해)은, 경험적으로 주어진 물리적 사건들에 의해 성립하는 상대빈도 계열에 대해 극한 개념이 의미 있느냐는 것이다.[2] 수학에서 사용되는 극한 개념은 계열의 생성함수(generating function)가 주어질 때 그 함수에 대해 적용된다. 즉 라이헨바하 자신이 밝히고 있듯 수학의 극한은 내포적으로(intensionally) 주어진 계열에 대해서만 적용된다. 그렇다면 외연적으로 하나 하나 열거되어 주어지는 물리적 사건들의 계열에 대해 극한 개념이 의미 있게 적용된다고 보장될 수 있는가?

이 비판에 대해 새먼은, 극한을 정의하는데 사용되는 '모든', '존재한다', '보다 작다' 등의 표현들이 물리적 사건들에 적용되는 데 있어 모두 유의미하므로 이런 표현들을 결합해 구성한 술어를 물리적 사건들에 적용하는 것은 문제가 없다고 대답한다.[3] 그러나 그 대답을 인정한다고 하더라도, 외연적으로 주어진 물리적 계열에 있어 극한이 존재하느냐는 것은 우연적인 문제이고, 또한 그런 계열의 유한한 부분으로부터 극한치를 추론해내기 위해서는 귀납이 요구된다. 그래서 라이헨바하와 새먼은 계열의 관찰된 유한한 첫 부분으로부터 극한치를 추론하기 위해 다음의 귀납 규칙을 제안한다.[4]

$$F^n(B/A)=m/n \text{ 이라고 주어지면, } \lim_{n\to\infty} F^n(B/A) = m/n\text{이라고 추론하라.}$$

'매거에 의한 귀납 규칙(rule of induction by enumeration)이라 불리는 이 규칙은 계열의 유한한 첫 부분에서 얻을 상대빈도를 그대로 무한계열의 전체의 상대빈도로 투사한다. (이 규칙은 m/n=1인 특수한 경우, 우리가 일상적으로 알고 있는 매거에 의한 귀납이 된다.)

---

2) Black "Comments" in *Induction: Some Current Issues* (eds. by H. E. Kyburg & E. Nagel)

3) Salmon (1966) p. 84

4) Reichenbach (1949) pp. 339-340, Salmon (1966) ibid

빈도 해석에 대한 기존의 비판들 중 몇 가지가 이미 언급되거나 시사되었다. 첫째, 개별적 사건의 확률이 빈도 해석에 의해서는 의미를 부여받지 못한다는 것이다. (이는 단칭 경우(single case)의 문제라고 불린다.) 둘째, 상대빈도 계열에 있어서의 극한 개념이 유의미한가 하는 것이다. 셋째, 관찰 가능한 유한한 계열로부터 극한값이 알려질 수 있는가에 대한 것이다. 이 중 둘째 비판에 대해서는 이미 대답된 것으로 간주하고, 첫째와 셋째 비판과 관련해서는 그 비판들에 대한 빈도주의적 해결 시도에 대해 다시 비판을 제기하겠다.

III. 단칭 경우 문제에 대한 빈도주의자의 해결책이 지닌 문제들

1.단칭 경우 문제에 대한 새먼의 해결시도

빈도 해석에 의해 정의된 확률은 일차적으로 두 집합 사이의 관계이다. 그러나 확률 판단을 하는 대부분의 (혹은 최소한 상당히 많은) 경우 우리는 특정한 개별 사건에 대해 확률을 부여한다. '이번에 이 주사위를 던져 1의 면이 나올 확률은 1/6이다'나 '내일 비가 올 확률은 40%이다' 등의 진술들이 모두 이런 종류의 확률 판단을 담고 있다(단칭 경우 문제). 흔히 빈도주의자들은, 준거집합에 부여된 확률을 그 집합에 속한 각각의 개별 사건들에 할당하는 방식으로, 이를 해결하려 했다. 이것이 해결책이 되지 못함은 분명하다. 개별적 사건들은 일반적으로 하나 이상의 준거집합에 속할 수 있기 때문이다(준거집합의 문제). 결국 이들 빈도주의자의 방식을 따른다면, 한 개별 사건에 대해 그 사건을 어떻게 기술하는가에 따라 서로 다른 확률 값이 부여될 수 있다.

이에 대한 새먼의 해결책은 개별사건이 속하는 '가장 광범위한 동질적 준거집합(the broadest homogeneous reference set)'을 찾아, 그 준거집합에서의 상대빈도를 그 개별 사건의 확률로 삼는 것이다.[5)]

새먼에 의하면 집합 A가 속성들 B의 발생에 관해 동질적(homogeneous)이라는 것은, B의 발생에 통계적으로 유관한(statistically relevant) 속성C에 의해 A를 분할할 수 없을 경우를 말한다. 속성 C가 (A 속에서) B의 발생에 통계적으로 유관하다고 하는 것은, 속성 C가

$$P(B/A\cap C) \neq P(B/A)$$

이게끔 A를 A∩C와 A∩C′의 두 부분으로 나뉘어지게 하는 경우를 말한다.[6] 이러한 기준은 이를테면 보험회사에서 특정한 운전자에 대해 사고를 낸 확률을 부여할 때 사용하는 방법이다. 운전면허를 지닌 모든 운전자들의 집합은 사고를 내는 일에 대해 동질적이지 않다. 이 집합은 연령, 성, 운전경력 등의 속성에 의해 분할될 수 있다. 이런 속성들에 의해 분할해 나가서 더 이상 분할할 수 없게 동질적인 (그러면서 가능한 한 가장 큰) 집합을 준거집합으로 삼는다는 것이다.

2.새먼의 기준에 대한 비판

해결책으로 내세운 새먼의 기준이 받아들일 만 하지 않다는 것을 보기 위해 다음의 예를 살펴보겠다. 정확히 5분 간격으로 전철이 오는 (그리고 내가 이 사실을 알고 있는) 전철역이 있다. 전철역으로 들어가면서 (그 때가 오후 3시 17분이었다), 내가 전철을 3분 이상 기다릴 확률이 2/5라고 생각하는 것은 합리적이다. 그런데 내가 아는 것은 전철이 5분 간격으로 온다는 사실뿐이지만 실제로는 전철이 오는 정확한 시각까지도 매일 일정하다고 하자. 그리고 실제로 전철은 오후 3:13, 오후 3:18, 오후 3:23 등에 온다. 빈도 해석론자들은 지금 내가 전철을 3분 이상 기다릴 확률을 구하기 위해, 지금과 유사한 상황에서 이를테면 사람들이 전철을 타러 나갔을 때의 경우들을 무수히 살펴보아 그 중 전철을 3분 이상 기다리는 상대빈

5) Salmon (1966) p. 91

6) Reichenbach (1949) p. 374, Salmon (1966) p. 91

도가 얼마나 되는지 조사할 수도 있다. 그런데 이 때 4분 간격으로 전철이 오는 전철역에서의 경우까지 조사하려 하지는 않을 것이다. 둘 다 어떤 전철역에 들어가 전철을 기다리는 사건들이지만 두 사건 집합간에 통계적으로 유관한 속성의 차이가 있다고 말해질 것이다. 그러면서도 오후 3:17에 이 전철역에 들어간 사람들의 경우 뿐 아니라 다른 시각에 이 전철역에 들어간 사람들의 경우들까지 조사하려 하는 이유는 무엇인가? 만약 오후 3:17에 이 전철역에 들어간 경우들만을 조사한다면 3분 이상 기다리는 경우들의 상대빈도는 0일 것이다. 그리고 전철역으로 들어가는 시각이 오후 3:17이었다는 속성(C라 하자)은 새먼의 기준에 의해 3분 이상 기다린다는 속성(A)에 대해 통계적으로 유관하다. 왜냐하면 (B를 이 전철역에 들어가는 사건들의 집합이라 할 때) $P(B/A \cap C)=0$인데 반해 $P(B/A)=2/5$이므로, $P(B/A \cap C) \neq P(B/A)$이기 때문이다.

그렇다면, 무엇이, 어떤 것들(전철이 4분 간격으로 오는 전철역의 경우)을 지금 우리의 상황과 본질적으로 다른 것이게끔 하고, 다른 것들(전철역에 오후 3:19에 들어간 경우)을 우리의 상황과 함께 고려될 유형의 것이게끔 하는가? 만약 내가 전철이 5분 간격으로 온다는 사실까지도 몰랐다고 하자. 그 경우 4분 간격으로 전철이 오는 전철역의 경우를 통계에서 배제해야 한다는 것은 미심쩍어지게 될 것이다. 이는 각 사건의 확률을 부여하기 위해 준거집합을 정하는 일 자체가 우리가 아는 바 혹은 믿는 바에 의존해 있다는 것을 보여주고 있다.

이에 대해 빈도 해석론자는 '인식적으로 동질적인'이나 '실천적으로 동질적인'의 개념을 끌어들여 대답하려 할지도 모른다. 예를 들어 이 주사위를 이번에 던져서 1의 면이 나올 확률이 1/6이지만 이 사건의 초기 조건을 충분히 기술하면 (거시 계에 대한 결정론적 법칙이 알려져 있다고 가정했을 때) 이를 아는 한에서 1의 면이 나올지 나오지 않을지를 계산할 수 있을 것이다. 그러나 초기 조건을 이와 같이 상세하게 아는 것은 실천적으로 가능하지 않거나 바람직하지 않으며, 이런 특정 초기 조건들이 통계적으로 유관한 속성이기는 하지만 이것은 고려되지 않은 한에서의 동질적인 집합도 실천적으로 동질적인 집합은 된다고 할 수 있다. 빈도 해석론자는 이런

'실천적으로 동질적인' 집합이 준거집합 역할을 할 것으로 볼 수 있다.

그러나 빈도 해석론자가 이렇게 대답하려 한다면, 그는 개별 사건들에 확률을 부여하는 문제에 있어서, 객관성을 포기하는 것이다. 개별 사건의 확률 부여는 단순한 실천적인 적용의 융통성의 문제로 되어 버린다.

IV. 경험주의와 빈도 극한개념의 문제

1. 빈도 해석은 경험주의적 의의를 지니는가?

흔히 빈도 해석의 가장 큰 장점이라고 여겨지는 것은, 이 해석이 확률을 경험주의적으로 다룰 수 있게 해 준다는 점이다. '이번에 주사위를 던져 1의 면이 나올 확률은 1/6이다'라는 진술은 경험주의자들을 당황스럽게 만드는 경향이 있다. 이 진술은 실제로 이번에 1의 눈이 나오더라도 나오지 않더라도 검증되거나 반증되지 않는다. 어떻게 이런 진술이 의미 있게 이루어질 수 있는 것일까? 이런 진술을 경험주의자가 의미 있게 받아들일 수 있는 유일한 방식은 확률에 대한 빈도해석이라 생각되는 경우가 없지 않았다. 즉 이 진술은 주사위를 계속 던져 전체 계열 속에서 1의 면이 나오는 비율이 1/6이라는 것을 의미하는 것으로 보아야 한다는 것이다. 그러나 유한 계열에서는 -그것이 아무리 긴 계열이든- 1의 면이 나올 확률이 각 시행에서 1/6이었음에도 불구하고 그 계열이 상대 빈도가 1/6과 매우 다르게 실현되는 것은 항상 가능하다. 물론 그 유한 계열의 길이가 길수록 그 계열에서의 상대빈도가 1/6과 매우 다르게 실현될 확률은 작아진다. 따라서 꽤 긴 계열에서의 상대빈도가 1/6과 상당히 다르다면 (예를 들어 그 계열 전체에서 1이 한번도 나오지 않았다면) 그 주사위의 1의 면이 나올 확률이 1/6이라는 가정은 그다지 그럴듯하지 않게 된다. 이 경우 계열이 그렇게 성립했다는 사건이, '주사위의 1의 면이 나올 확률이 1/6이다'라는 가정 하에서는 논리적으로 불가능하다는 점에서 그 가정을 논리적으로 반증하지는 않지만, 그 가정 하에서 일어날 확률이 작다는 점에서

그 가정을 개연적이지 않게 만들기는 한다.(이를 '개연적으로 반증되었다'라고 부르기로 하자) 그러나 이 점에 관한 한 그것은 단칭적인 확률 진술과 본질적으로 다르지 않다. 중요한 것은 단칭 시행인가 시행들의 계열인가가 아니라 확률 부여가 1이나 0에 얼마나 가까운가이다. 예를 들어 어떤 시행이 어떤 속성을 지닐 확률이 0에 가깝게 부여되었을 수록, 그 시행이 실제로 그 속성을 지녔다는 관찰이 그 확률 부여를 개연적으로 반증하는 정도가 더 크다. '내일 비가 올 확률이 30%이다'라는 진술 대신 내일과 똑같은 조건의 날들(내일 비가 올 확률과 똑같은 정도의 비가 올 확률을 지닌 날들)의 n계열에서 실제로 비가 오는 날들의 상대빈도가 30% $\pm\sigma$의 구간 내에 있을 확률에 대한 진술을 고려하는 이유는 후자의 확률부여는 1에 더 가까워서 실제로 그 구간을 벗어나게 상대빈도가 주어졌을 경우 개연적 반증의 정도가 더 크기 때문이다. 그러나 후자의 진술 역시 확률에 대한 진술임에 주의할 필요가 있다. 위의 작업은 단지 확률 계산 내에서 한 확률 진술로부터 다른 확률 진술을 끌어내는 것에 불과하다. 이는 확률 계산에서의 정당한 추론이며, 이 추론을 위해서는 확률이 상대빈도를 '의미'한다는 해석은 전연 요구되지 않는다.

빈도 해석이 주장하는 것은 단일 시행의 확률을 시행 계열의 **상대 빈도에 대한 확률**이 아닌 **상대빈도 자체**와 동일시하자는 것이다. 물론 유한 계열에 있어선 이런 동일시는 전연 그럴듯하지 않으므로 계명된 빈도주의자들은 이를 무한 계열에서의 상대빈도와 동일시한다. 그러나 그 자체는 아무런 경험적 의의를 지니지 않는데, 이는 어차피 무한 계열이 다 관찰될 수는 없기 때문이다. 빈도주의자들은, 그래도 확률을 무한 계열에서의 상대빈도와 동일시함으로써 경험적으로 접근 가능한 유한계열에서의 상대빈도로부터 개연적으로나마 확률을 추론할 수 있다고 말할지도 모른다. 그러나 위에서 보았듯 유한계열의 상대빈도 자체가 아닌 그 상대빈도의 확률은 확률 계산 내에서 각 단일 시행의 확률로부터 추론될 수 있고, 따라서 어떤 특정한 유한계열의 상대빈도로부터 **확률적으로** 단일 시행의 확률에 대한 가설을 긍정하거나 부정하는 것은 확률 계산을 어기지 않는 어떠한 해석에서도 허용된다. 결국 상대 빈도와 관련해 경험적으로 다뤄질 수 있

는 부분은 빈도 해석을 요구하지 않으며 빈도 해석에 의해 잉여로 추가되는 부분은 아무런 경험적 의의를 지니지 않는다.

2. 베르누이의 정리와 빈도 해석

그럼에도 불구하고 빈도주의자들로 하여금 확률을 상대빈도의 극한과 동일시하게 만드는 것이 있다면 그것은 베르누이의 정리(the theorem of Bernoulli)일 것이다.[7] 베르누이의 정리는 다음과 같이 간단히 표현된다.[8]

$$P(\lim_{n \to \infty} \frac{S_n}{n} = p) = 1$$

여기서 S_n/n은 n번의 독립적인 베르누이 시행(두 가지 결과, 이를테면 성공과 실패가 가능한 실험)에서의 상대빈도이다. P는 각 시행에서 문제의 속성을 가질(성공을 할) 확률이다. 일견 보기에 이 정리는 빈도 해석을 정당화해 줄만큼 충분히 강력한 내용을 말하는 것 같다. 왜냐하면 이 정리는 상대 빈도의 극한이 확률과 동일할 확률이 1이라고 말하는 것처럼 보이기 때문이다.

문자 그대로는 그것은 옳다. 그러나 '1의 확률을 지닌다'는 것에 대해서도 (빈도주의가 옳다면) 빈도 해석에 의해 의미가 부여되어야 한다. 빈도 해석에 의해 대문자 'P'의 의미가 주어졌을 때, 베르누이의 정리가 말하는 바는 단지, 계열들의 계열(메타 계열) 속에서 '상대 빈도가 p로 수렴하는' 계열들이 차지하는 상대 빈도의 극한치가 1이라는 것이다. 그러나 이는 계열들의 계열 속에 '상대빈도가 p로 수렴하는' 계열들만 존재한다는 것을 의미하는 것이 아니다. 오히려 '상대빈도가 p로 수렴하지 않는' 계열들이 무한히 많이 있을 수 있다. (자연수들의 계열 속에서 제곱수는 자연수의 갯수만큼 많이 있지만 이 계열 속에서 제곱수의 상대빈도의 극한치는 0이라는 것을 상기해 보는 것이 도움이 될 것이다). 결국 빈도주의

7) 예를 들어, Reichenbach (1949) sec. 49

8) 이것의 증명에 대해서는 Kneale (1949) pp. 136-138, Reichenbach (1949) Sec. 49 참조.

적으로 해석된 한에 있어 확률이 0이라는 것은 불가능성을 의미하지 않는다. 실제로 메타계열 속에 '상대빈도가 p로 수렴하지 않는' 계열이 무수히 있으면서도 그런 계열이 존재할 확률이 0일 수 있다!

따라서 우리가 문제삼는 특정 계열이 '상대빈도가 수렴하는' 계열이란 보장은 없다. 그러므로 베르누이의 정리에 의해, 상대빈도의 극한치와 확률을 동일시하는 것은 정당화되지 않는다. 결국 베르누이의 정리 역시 확률 계산 내의 다른 정리들과 마찬가지로 한 확률로부터 다른 확률을 끌어내도록 허용해 줄뿐이다. 빈도 해석이 그럴듯해지기 위해선 확률 계산의 정리 이상의 어떤 직관이 요구되는 것 같다.

3. 충만의 원리와 상대빈도의 극한

그런데 그런 직관이 첨가될 때 무한 빈도해석이 봉착할 또 다른 흥미로운 문제점을 지적하겠다. 공정한 주사위를 던졌을 때 1이 나올 확률이 1/6이라고 하지만, 그런 주사위를 던진 사건들의 유한한 계열은 어떤 형태의 것도 다 가능하다는 것에 빈도주의자도 동의한다. 예를 들어 공정한 주사위를 백 번 던졌는데 1이 한번도 나오지 않았을 수 있다. 이런 일이 매우 개연성이 적은 일이긴 하지만, 불가능한 일은 아니다. 우리가 유한 빈도해석보다는 무한 빈도해석을 택할 때에 가지는 직관은 다음과 같은 것이다: 1이 나올 확률이 1/6임에도 불구하고 유한계열 속에선 그것이 현실화되지 (actualized) 않을 수 있다. 즉 극단적인 경우에 1은 한 번도 나오지 않을 수 있다. 그러나 무한계열 속에선, 1이 나올 가능성이 있음에도 현실화되지 않을 수는 없다. 이같은 생각을 받아들임으로써 유한계열에 대해 던져질 수 있는 위의 물음이 무한계열에 대해서는 성립하지 않는다고 믿는다. 흔히 '충만의 원리(principle of plenitude)'라고 불리는 이 원리는 우리가 꼭 받아들여야 할 원리인지 의심스럽지만, 무한 빈도해석론자들의 편의를 위해 이 원리를 타당한 것으로 받아들이기로 하자. 그러나 이 원리를 받아들이기만 하면 무한 빈도해석 자신이 위협받게 되는 아이러니가 발생한다. 무한 빈도해석이 확률을 상대빈도 $F^n(H)$의 극한치라고 정의할 때, 극한치 $lim F^n(H)$는 다음과 같이 정의된다.

$$\lim_{n \to \infty} F^n(H) = L$$

iff 아무리 작은 양의 ε에 대해서도, 다음을 만족하는 수 N이 존재한다. $n > N$ 이면 $|F^n(H) - L| < \varepsilon$이다.

'iff'의 오른쪽이 성립하는 경우가 있다고 해보자. 이는 모든 ε에 대해 성립하므로 특정한 ε(상수 ε_o이라 하자)에 대해서도 성립할 것이다. '$n > N$이면 $|F^n(H) - L| < \varepsilon_o$'을 만족하는 수 N이 존재한다고 했으므로, 이 수를 No라고 부르기로 하자. 그러면 이 No보다 큰 모든 수 n에 대해 $|F^n(H) - L| < \varepsilon_o$ 가 성립한다. 즉 No이후의 무한 계열에서 $F^n(H)$와 L의 차이가 ε_o보다 크게 벌어지는 일이 단 한번도 발생하지 않는다. 그러나 유한한 상대빈도 $F^n(H)$가 L과 매우 다르게 실현되는 것은 아무리 큰 n에 대해서도 아무리 미세하게나마 항상 가능하다. ($F^n(H)$가 L과 매우 다르게 실현되는 것의 확률은 n이 ∞에 가까워질 수록 0에 가까워진다. 그러나 n이 자연수인 한 그 확률은 0과 같아지지 않는다.)

그리고 충만의 원리에 의하면, 가능성이 있는 일은 무한계열 속에서 실현되는 경우가 반드시 존재한다. 따라서 'iff'의 오른쪽이 성립한다는 가정은 모순을 낳고, 결국 어떤 상대빈도에 대해서도 극한치가 존재하지 않게 된다. 그렇다고 해서 충만의 원리를 받아들이지 않게 되면 무한계열에서의 극한치를 확률과 동일시하는 것은 정당화될 수 없게 된다.

V. 극한에의 추론과 귀납의 정당화 문제

우리 모두가 비록 지금까지 귀납을 사용하고 그것이 유용했지만, 귀납적 추론의 결과를 믿지 않는 것이, 확률 계산의 결과를 따르지 않는 것이 비합리적이라는 의미에서처럼 그렇게 비합리적인 것이라고는 생각되지 않는다. 우리는 귀납이 귀납을 자기반박할 만큼 심하게 제일적(uniform)이지 못한 세계를 상상할 수 있다. 그러나 확률계산에 대해 그것을 받아들이

지 않는 것이 좋을 세계를 상상하는 것은 어렵다. 확률계산은 그것을 어길 경우 더치 북(Dutch Book)에 말려든다는(어떤 경우에건 손해가 되는 내기상황을 공정한 것으로 받아들일 수 있다는) 사실을 통해 정당화될 수 있다. 이렇게 이해된 확률부여는 기존 믿음의 정도가 주어진 한에서, 귀납의 합리성에 의존하지 않으며, 귀납이 정당화될 수 있는지의 여부에 무관하게 정당화되어진다.

그에 반해, 확률을 무한계열에서의 상대빈도로 정의할 때에 확률부여는 불가피하게 귀납의 정당함을 요구한다. 왜냐하면 유한한 증거로부터 상대빈도의 극한치를 추론하기 위해서는 항상 귀납이 요구되기 때문이다.

실제로 라이헨바하와 새먼은 나름대로 귀납을 정당화하려는 시도를 하고 있다.[9] 라이헨바하는, 최소한 다른 어떤 방법이건 미래에 대한 예측을 할 수 있다면 귀납도 그런 예측을 할 수 있음을 보일 수 있다고 말한다. 자연은 제일적이거나 제일적이지 않은데, 자연이 제일적이라면 귀납은 예측력을 가질 것이고 제일적이지 않다면 다른 어떤 방법도 예측력을 가지지 않을 것이다. 또한 만약 다른 방법이 예측력을 가진다면 귀납적 방법은 그 방법이 예측력을 가진다는 사실을 찾아낼 수 있고 따라서 그 방법을 통해 귀납도 미래를 예측할 수 있다. 그러나, 그들이 행하는 귀납에 대한 실용적 정당화(pragmatic justification)는 확률해석 문제에 적용되었을 때에는 원래의 귀납 자체에 대한 정당화만큼 그럴듯하지는 못하다. 극한치를 추론하는 문제에 있어서도, 극한치가 존재하지 않을 경우 어떠한 방법도 극한치를 찾을 수 없을 것이라는 점에서 귀납을 택할만한 이유가 있는 것처럼 보인다. 그러나 이 경우 극한치가 존재하지 않는다면, 어떠한 방법도 극한치를 찾을 수 없을 것이라는 말을 할 수 있어도 어떠한 해석도 확률을 부여할 수 없을 것이라는 말은 할 수가 없다. 만약 어떤 경우에 극한치가 존재하지 않는다면, 그것은 단지 (모든 계열에 대해 확률값을 부여해야 할) 빈도해석이 그 경우에 성공적이지 못하다는 것을 의미할 뿐이다.

상대 빈도의 유한한 계열로부터 극한치를 추론하기 위해 라이헨바흐와 새먼이 제안한 '매거에 의한 귀납 규칙'은 n항까지의 상대빈도 $F^n(B/A)$

9) Reichenbach (1933) p. 473, Salmon (1966) pp. 52-3

의 값을 극한 $lim F^n(B/A)$의 값으로 추정(posit)하는 것이었다. 라이헨바흐는 특별히 이 귀납규칙에 대한 실용적 정당화를 시도한다. 만약 상대빈도의 계열에 극한치가 존재한다면, 매거에 의한 귀납 규칙의 반복 적용은 언젠가는 극한치를 발견할 것이다. 만약 극한치가 존재하지 않는다면, (당연한 이야기지만) 어떠한 방법도 극한치를 발견할 수 없을 것이다. 이 규칙의 반복적용이 극한치를 '언젠가는 발견한다'는 말의 의미가 더 분명하게 되어져야 한다. 그것은 아무리 작은 δ에 대해서도, 이 규칙으로 추정한 극한치가 실제 극한치에 δ만큼보다 더 차이나지 않게 가까와지는 항 n이 존재한다는 것이다. 이것은 다소 사소하게 성립한다. 추정치들 $E^n(B/A)$의 계열은 상대빈도의 계열과 일치한다. 그러므로 추정치들의 계열의 극한치는 곧 상대빈도 계열의 극한치와 같다. 즉 추정치들의 계열은 상대빈도 계열의 극한치에로 수렴한다.

이렇게 보았을 때, 빈도 계열의 실제 극한치에로 계속 가까와져가는 속성은 특정한 추정치가 가지는 속성이 아니라 추정치들의 무한 계열이 가지는 속성이라는 점이 부각되어 드러난다. 그리하여 문제는 매우 당황스러운 것으로 남는다. 이 귀납규칙에 의해 추정되는 특정한 값들은 우리에게 접근 가능하지만 결국 추정되는 값들의 극한치는 상대빈도 계열의 극한치와 마찬가지로 알려지지 않기 때문이다.

라이헨바흐는 규칙의 반복되는 적용에 의해 추정치들을 계속 나열하면, 언젠가는 (상대빈도 계열의) 극한치와 매우 작은 차이만 나는 추정치가 나타나게 될 것이라는 점에서, 이 규칙이 유용하고 받아들일만 하다고 본다. 그러나 추정치들의 계열이 결국에는 우리가 구하고자 하는 빈도의 극한치로 수렴할 것이라고 말하는 것은 별로 소용이 없는 것 같다. 우리가 필요로 하는 것은 특정한 추정치가 빈도의 극한치와 일치하거나 그와 비슷하다고 믿을 만한 이유를 제공받는 것이다. 그리고 이는 극한치가 존재한다고 하더라도 어떠한 추정치에 대해서도 보장받지 못한다.

또한, 계열의 유한한 부분의 규칙성을 계열 전체에도 투사하는 방법은 수없이 많으며 라이헨바흐의 '매거에 의한 귀납 규칙'은 그 중 오직 한 귀납 방법일 뿐이다. 이 규칙을 따르는 것이 어떤 경우에나 가장 자연스러

운 추정인 것도 아니다. 예를 들어 어떤 상대빈도 계열이 (8항까지) 다음처럼 주어졌다고 하자.

1/1 1/2 1/3 1/4 1/5 1/6 1/7 1/8 -----

이 계열의 주어진 부분으로부터, (전체 계열에 걸쳐) n항의 일반식이 1/n이라고 추정하는 것은 자연스럽다. 그리고 그렇게 추정했다면 당연히 이 계열의 극한값은 0이라고 말해야 할 것이다. 그러나, 소위 '매거에 의한' 규칙에 의해 귀납추론을 할 경우 주어진 부분의 맨 마지막 항인 $F^8(B/A)$의 값, 즉 1/8을 전체 계열의 극한값으로 추정해야 한다. 물론 이 상대빈도 계열이 '지나치게' 규칙적이어서 확률을 다루는 사람에게 진정으로 관심거리가 될 만한 계열이 되지는 않다는 것을 언급해 두어도 좋을 것이다. 그런데 문제는, 극한값 추정이 특별히 문제가 되지 않는 (그래서 확률 이론에는 별로 관심거리가 안되는) 계열에 대해서조차 오히려 이 규칙이 자연스럽지 않은 추정을 한다는 것이다. 수학적 계열에 대해서까지도 그 상대빈도의 극한 또는 통계적 확률을 구하는 것은 (비록 흥미롭지는 못하더라도) 의미가 있다. 관찰되는 상대빈도들이 규칙적인지 그렇지 않은지, 규칙적이라면 어떤 규칙성을 지니는지는 미리 알려지지 않는다. 그러나 적절한 귀납 규칙이라면 두드러진 규칙성이 있을 경우 그 규칙성을 존중해야 한다. 그런데 '매거에 의한' 귀납은 상대 빈도 계열의 주어진 부분에 어떠한 규칙성이 있다고 하더라도 그 규칙성을 무시하고 그 부분의 맨 마지막 항만을 추정의 자료로 삼는다.

이상의 모든 상황에서 무한 빈도해석은 어려움에 직면하고 있다. 빈도를 세어 확률값을 구하는 것이 가장 객관적인 확률부여라는 생각이 흔히 빈도해석을 택하게 하지만, 상대빈도의 극한치 개념은 여러 어려움들을 낳고 있고 급기야 확률을 부여하는 일을 객관적이게 보다는 어느 정도는 임의적이게끔 만들어 버리고 있다.

VI. 준거 계열 구성의 문제

라이헨바흐나 새먼 같은 빈도주의자들은, 고전적 확률 해석에서 전제되는 무차별의 원리(Principle of indifference, 한 경우를 다른 경우보다 선호할 이유가 없다면 두 경우는 똑같은 확률을 가진다는 원리)가 논리적 모순을 낳는다는 것, 즉 동일한 경우에 대해 서로 다른 두 확률을 부여한다는 것을 지적하는 논변을 폈었다. 그런데 이 논변을 변형함으로써 빈도 해석에서도 같은 구조의 문제가 생긴다는 것을 보일 수 있다.

빈도해석은 준거집합(A)과 속성집합(B)의 관계로서의 확률을 정의한다. 즉, P(B/A) = r처럼, 확률함수는 준거집합이 주어진 한에서 속성집합에 확률값을 부여한다. 우리의 문제는 다음과 같다: 준거집합이 주어지면 그 집합에서의 속성의 상대빈도의 극한치는 단일하게 결정되는가? 우리가 어차피 무한계열을 구성해 그 극한치를 구한다면, 무한계열 속에서 준거집합의 모든 원소들이 나열될 것이라고 안심해서는 안된다. 준거집합은 나열할 수 없이 불가번적으로 무한(indenumerably infinite) 할 수도 있다. 한편 계열이 아무리 무한하게 계속되어도 그것은 나열의 한 방식에 불과하다. (또한 설사 모든 원소들이 나열되더라도 나열의 방식에 따라 다른 상대빈도가 얻어질 수 있다.)

어떤 물체가 등가속도(가속도 = a)로 정지상태로부터 좌표 x_0로부터 양의 x축 방향으로 낙하운동을 했다. (운동을 시작한 시각을 t_0라하고 지면에 떨어진 시각을 t_n이라 하자) 그 물체는 t_n에 이르기까지 시간에 따라 일정하게 속도가 증가했을 것이다.(도표 1 참조)

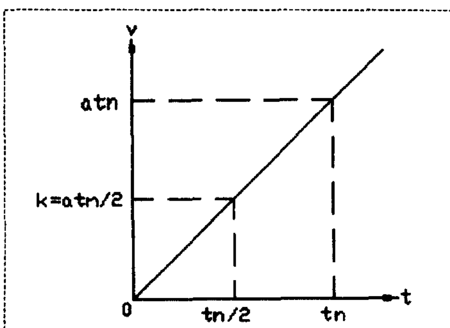
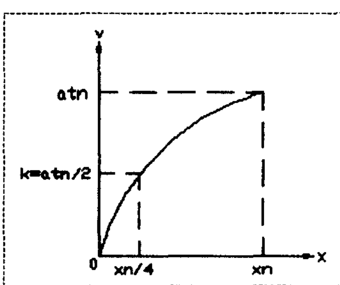


도표 1

이 경우, t_0로부터 t_n까지 중의 특정시점에서 그 물체의 속도가 k보다 빠를 확률은 1/2이다. '특정 시점 t에서의 그 물체의 운동(the movement of the body at t)' 형식은 하나의 사건을 지칭하기에 문제가 없는 기술어구이다. 우리는 이런 형식의 사건기술어구(event description)가 지칭하는 사건들이 임의로 선택되었을 때 우리는 그 사건이 'k보다 빠른 속도를 가진다'라는 속성을 가지는 경우들의 상대빈도를 셀 수가 있다. 그 사건이 $t_n/2$ 이전의 운동이라면 문제의 속성을 가지지 않을 것이고 $t_n/2$ 이후의 운동이라면 문제의 속성을 가질 것이다. 만약 이런 사건들이 충분히 임의적으로 선택된다면, 빈도해석에 의해서도 1/2의 확률값을 얻으리라는 것이 인정될 수 있을 것이다 .

지면에 이르렀을 때의 위치(x_n)는 $at_n^2/2$이다. 이 위치에 이르기까지 물체의 속도는 v= 2ax를 만족하는 방식으로 증가한다. (도표2 참조)

도표 2



이번에는, x_0에서 x_n사이의 특정 위치에서 물체의 속도가 k보다 빠를 확률을 구하기 위해, 그 구간의 임의의 위치에서 그 물체의 속도를 검사한다고 해 보자. 준거집합 속의 사건들을 충분히 임의적으로 선택해 계열을 구성한다면, x_0에서 $x_n/4$ 사이의 위치에서의 물체의 순간 운동들은 $x_n/4$에서 x_n사이의 위치에서의 물체의 순간운동들의 1/3의 비율로 계열 속에 등장할 것이다. 따라서 이 계열에서의 사건들이 'k보다 빠른 속도를 가진다'라는 속성을 가지는 경우들의 상대빈도는 3/4일 것이다.

그런데, t_0로부터 t_n까지 중의 특정시점에서의 물체의 운동들의 집합은 x_0에서 x_n사이의 특정 위치에서의 물체의 운동들의 집합과 동일한 집합이다. 물체가 시점 α에 위치 α'에 있다고 할 때, '시점 α의 물체의 운동'은 '위치 α'에서의 물체의 운동'과 동일한 하나의 사건을 지칭한다. 그런데,

계열을 어떻게 선택하느냐에 따라 빈도해석은 동일한 한 사건에, 혹은 동일한 한 집합에 대해 서로 다른 (동일한 속성을 가질) 확률을 부여한다.

이런 상황은 사실 다음과 같은 보다 단순한 수학적 예에서도 유사하게 일어난다.(그리고 이 경우, 준거집합은 가번무한임에 주의하자) 자연수의 집합에서 짝수 속성의 상대빈도는 어떻게 되는가? 이 물음에 쉽게 1/2이라고 대답할 수도 있지만, 이것은 자연수의 집합을

1,2,3,4,5,6,7,...

와 같은 계열에 의해 순서지울 때에 한해 올바른 대답이다. 그러나 동일한 집합을

1,2,4,3,6,8,5,10,12,...

와 같은 계열에 의해서도 모두 망라할 수 있다. 그리고 이 계열에 있어 짝수 속성의 상대빈도는 2/3이다.

빈도해석론자는 이런 어려움을 피하기 위해 확률을 준거집합과 속성집합의 관계로 정의하지 않고 준거계열(reference sequence)과 속성집합의 관계로 정의하려 할지도 모른다. 그러나 이런 확률정의는 빈도해석을 우리의 직관적 확률개념으로부터 크게 멀어지게 할뿐이다. 준거집합을 써서 정의되는 빈도해석은 어떤 사건 유형(event type)이 어떤 속성을 가질 확률을 부여하는 것으로 이해되어질 수 있어도, 준거계열을 써서 정의되는 빈도해석은 우리 확률언어 속의 어떤 문장을 해석하는 것인지 이해할 수가 없다.

VII. 맺음말

빈도 해석은 고전적 확률 해석이 부딪힌 논리적 난점에 대한 대안으로

등장했다. 고전적 해석은 무차별의 원리(principle of indifference)를 전제함으로써 주관의 임의적(arbitrary)선택에 의존한다는 것이 빈도주의자의 비판이었다. 그러나 우리가 이 논문에서 본 것은 빈도주의자 역시 그런 선택에 의존하지 않을 수 없다는 것이었다. 첫째, 개별적 사건은 어떻게 기술되는가에 따라 다른 확률을 부여받게 되었다. 둘째, 관찰된 유한 빈도로부터 (무한 빈도로서의) 확률을 추론하는 것도 수없이 가능한 여러 귀납 추론 중의 한 방식에 불과했다. 셋째, 준거 집합이 주어지더라도 어떻게 계열을 구성하는가에 따라 어떤 속성을 가질 확률이 일의적으로 결정되지 않았다.

우리의 교훈은 초기 확률(initial probability) 부여에 있어서의 임의성을 인정하고, 확률 진술의 객관성을 선행 확률(prior probability)이 주어진 한에서의 후행 확률(posterior probability) 부여에서 찾는 해석이 더 고려될 만하다는 것이다. 이런 해석은 확률을 합리적 믿음의 정도와 동일시하는 해석으로서 이미 발전되었다.[10] 이 해석을 적극적으로 옹호하는 것은 이 논문의 범위를 넘어서는 일이다.

---

10) 이 해석은 '베이즈적 해석(Bayesean interpretation)' 혹은 '개인주의적 해석(personalist interpretaion)' 등으로 불린다. 이 해석에 대해서는 Savage (1954), Cohen (1989) 참조.

참고 문헌

Cohen, J.(1989), *The Philosophy of Induction and Probability,* Oxford: Oxford University Press

Kneale, W.(1949), *Probability and Induction,* Oxford: Oxford University Press

Reichenbach, H.(1949), *The Theory of Probability,* Berkeley: University of California Press

Reichenbach, H.(1932), "The Logical Foundation of the Concept of Probability" in *Reading in the Philosophy of Science* (ed. Feigl and Brodbeck)

Salmon, W.(1966), *The Foundations of Scientific Inference,* Pittsburgh: University of Pittsburgh Press

Salmon, W.(1971), *Statistical Explanation and Statistical Relevance*

Savage, L.(1954), *The Foundations of Statistics,* Wiley & Sons