전문가 시스템을 위한 우도의 활용

전 영 삼 *

서 언

이른바 "전문가 시스템"(expert system)에 있어서는 흔히 새로 주어진 정보에 의해 관련 가설들을 새로이 평가하는 일이 필요하다.

전문가 시스템이란, 일상적으로 말해, 어떤 분야의 전문 지식을 지식 베이스(knowledge base)로 삼아, 주어진 문제 상황에 대해 전문적인 조언을 주는, 인공 지능(artificial intelligence)의 하나를 말한다. 예컨대 어떤 의학적 지식을 지식 베이스로 하는 전문가 시스템[1]에 있어서는 어느 신체 증상에 관한 기술(記述)을 입력으로 하여 그에 해당하는 병명(病名)과 함께 그에 따른 적절한 치료법을 출력으로 줄 수 있을 것이다.

이 경우 만일 그러한 증상과 병명을 연결해 주는 관계가 확고하게 확립되어 있다면, 일단 시스템 내에 들어가 있는 그러한 관계에 대한 지식은 새로이 평가되거나 수정될 필요가 없을지 모른다.

그러나 해당 영역 내에서의 끊임없는 발전으로 인하여 그와 같은 전문 지식들은 늘 새로운 추가적 정보에 의해 새로운 평가에 놓이는 것이 현실이다. 이러한 의미에서 그와 같은 지식들은 어떤 확고한 지식이라기보다는 차라리 계속해서 테스트에 부쳐져야 할 잠정적인 것이며, 따라서 그러한

---

* 고려대학교 강사

1) 이러한 시스템으로 대표적인 것은 MYCIN을 들 수 있다. 그러나 이 시스템에서는 불확실한 상황을 다루기 위해 전통적인 확률의 방법보다는 이른바 "확신률"(certainty factor)을 이용하고 있다. 이것은 주로, 이하에서 다룰, 베이즈 정리에 있어서의 사전 확률 부여의 난점에 기인한다. (D. W. Rolston, *Principles of Artificial Intelligence and Expert Systems Development*, New York: McGraw-Hill, 1988, p. 99 참조)

지식에 의한 결론 역시 하나의 가설로 봄이 적절하다. 예컨대 위의 의학의 전문가 시스템 예에 있어, 주어진 증상에 관해 그 원인으로 제시된 병명은 하나의 가설로 볼 수 있을 것이다.[2]

바로 이 순간 문제의 가설을 새로운 정보의 증거에 의해 평가함에 있어서는 주어진 가설이 주어진 증거를 내용적으로 넘어서고 있다는 점에 있어 귀납의 문제가 개입되어 있음을 볼 수 있다.

여기서는 일단 문제의 가설과 증거가 이미 주어져 있다는 점에서, 증거로부터 가설을 제시하는 귀납의 방법론(methodology of induction) 문제보다는 그 양자 사이의 지지(支持, support)의 정도를 문제삼는 귀납 논리(inductive logic)의 문제가 개입되게 된다.[3]

이와 같은 지지의 정도를 양적(量的)으로 평가함에 있어서는 일면 확률의 논리가 유용할 수 있고, 특히 변화하는 증거에 따른 가설의 평가 작업에는 이른바 "베이즈의 정리"(Bayes' theorem)가 유용할 수 있다.[4] 이것은 기본적인 확률 공리로부터 순수 연역적으로 주어질 수 있는 하나의 정리로서, 증거 언명(e)을 조건으로 하는 가설 언명(h)의 확률, 즉 P(h/e)는, 가설 언명이 참으로 주어졌을 때의 증거 언명의 확률 P(e/h)와 가설 언명만의 확률 P(h)의 곱에 비례하는 것으로 주어진다. 그러므로 변화하는 증거에 따라 해당 가설을 평가하기 위해서는 P(e/h)와 P(h)를 계산해

---

2) 본 논문에서는, 전문 지식을 표현하는 여러 방법 중 단지 지금과 같이, 'IF~ THEN~' 구문으로 표현되는 "규칙에 근거한 지식"(rule-based knowledge)만에 그 논의를 한정하기로 한다. 'IF~ THEN~' 구문에 있어 IF-절에는 문제의 증상이나 증거 또는 결과들이 표현되어 들어가고, THEN-절에는 그에 대한 원인으로서의 가설이 표현되어 들어간다. (전문가 시스템 내에서 지식을 표현하는 다양한 방법에 관해서는 J.G.Carbonell et al., "An Overview of Machine Learning," in:R.S.Michalski et al. (eds.), *Machine Learning*, Palo Alto: Tioga, 1983(pp. 3-23), pp. 11-3 참조)

3) 귀납에 관한 이와 같은 구분에 관해서는 R.Carnap, *Logical Foundations of Probability*(1950), London:Routledge & Kegan Paul, 1951. sec. 44 참조.

4) 전문가 시스템 내에서 불확실한 상황을 다루는 여러 방법에 관해서는 D.W. Rolston, op. cit., Ch. 6 참조.

낼 필요가 있는데, 이 경우 가장 큰 난점은 흔히 "사전 확률"(事前 確率, prior probability)이라 부르는 P(h)를 어떻게 계산해 낼 것이냐 하는 점이다. 아직 참으로 확정되지 않은 가설에 대해 확률치를 부여함에 있어 자의성(恣意性)을 배제하기 어렵기 때문이다.

그러나 흔히 "우도"(尤度, likelihood)라 부르는 P(e/h)의 경우에는, 만일 전체 모집단(母集團)의 확률적 분포에 관한 일정한 가정만 받아들여진다면, 그러한 자의성을 벗어나 훨씬 용이하게 그 값을 계산해 내는 일이 가능하다.

그러므로 전문가 시스템 설계에 있어서도 역시 이와 같은 우도를 활용해 볼 수 있는 가능성을 고려해 볼 수 있는데, 이와 같은 점에 관해서는 이미 보오간(R. A. Vaughan)에 의해 논의된 바 있다.[5] 그는 앞서 기술한 바와 같은 전문가 시스템에 있어 활용 가능한 베이즈의 방법과 우도의 방법을 차례로 제시하면서, 그 각각의 장단점을 논하고 있다. 그러나 예컨대 어떠한 가설이 그와 같은 평가에 적합하며, 증거 자체의 성격, 베이즈 방법과 우도의 방법의 관계 등에 관해서는 자세한 언급을 하지 않고 있다.

타가드(P. Thagard)의 지적대로[6] 어떠한 가설이든 위와 같은 평가 방식에 적합한 것은 아니며, 이것은 일반적인 확률 개념과 우도의 개념을 좀더 자세히 고찰해 보면 분명히 드러나는 점이다.

그러므로 이제 본 논문의 과제는 바로 이와 같은 가설이나 증거, 그리고 확률 및 우도에 대한 좀더 세밀한 논의를 전개하여 보오간의 상기(上記) 논문을 해명하고, 이를 기반으로 몇몇 점에서 그 내용을 개선시키는 일이다. 특히 후자의 경우 논자는, 추가적 증거에 의한 새로운 우도의 계산시 보오간이 고려하고 있는 표본의 크기 외에 증거의 속성들도 고려의 대상이 될 수 있음을 주장하고, 이를 반영하는 새로운 식을 제시하였다. 또한 전

5) R. A. Vaughan, "Maintaining an Inductive Database," in: J. H. Fetzer (ed.), *Aspects of Artificial Intelligence*, Dordrecht: Kluwer, 1988, pp. 323- 35.

6) P. Thagard, *Computational Philosophy of Science*, Cambridge: The MIT Press, 1988, p. 97.

문가 시스템 내에서의 가설 평가에 있어 베이즈의 방법보다는 우도의 방법이 궁극에 있어 좀더 효과적임을 보여주기로 한다.

1. 가설의 평가와 베이즈의 정리

「전산 과학적 과학 철학」(電算科學的 科學哲學, *Computational Philosophy of Science*)의 저자 타가드는 같은 책에서 과학에 있어 이론의 평가(theory evaluation) 문제를 논하면서 다음과 같이 지적한 바 있다.

> "AI에서의 최근 작업에서는 〔 … 〕 의학적 진단이나 여타 가설·상정론법(假說 想定論法, abduction)과 밀접하게 관련된 것으로 확률을 고려하고 있다."
>
> "많은 의학적 진단을 포함하여, 베이즈의 정리가 유용한 다양한 경우가 존재한다. 그러나 이론의 평가는 그러한 것 중의 하나로 보이지 않는다. 그와 관련된 확률을 발견하는 일은, 불가능하지는 않다 하더라도, 거의 어렵기 때문이다."[7]

일반적인 확률의 기본 공리에 따를 때 증거 언명 e와 가설 언명 h에 대한 베이즈의 정리는 다음과 제시할 수 있다.[8]

$$P(h/e) = \frac{P(h) * P(e/h)}{P(e)} \qquad (1)$$

---

7) Ibid..

8) 이와 같은 베이즈의 정리는 일반적인 통계학 관련 교재에서 흔히 발견할 수 있는 정리이다. 예컨대 R. V. Hogg & E. A. Tanis, *Probability and Statistical Inference*, 2nd ed., New York:Macmillan, 1983, 안철원 · 한상문 공역, 「통계학 원론」, 서울:비봉, 1989, pp. 72-5 참조(이하 본 논문에서 언급되는 일반적인 통계학 관련 내용에 관해서는 본서 참조).

여기서 P(h/e)는 증거 언명 e가 주어져 있을 때 가설 언명 h의 확률을 말하는 것으로, 따라서 만일 우리가 이 확률을 구해 낼 수만 있다면, 불확실한 상황하에서 어떤 증거 언명을 기반으로 어떤 가설 언명이 높은 개연성을 갖는가를 말할 수 있을 것이다. 즉 확률의 관점에서 문제의 가설을 평가하는 것이다.

예컨대 "어떤 사람이 계속 설사를 하며 토(吐)를 한다"라는 언명을 e로 삼고, 그와 같은 증상의 원인으로서 예컨대 장염(腸炎)을 고려하여 "그 사람은 장염을 앓고 있다"라는 언명을 h로 삼았다고 해보자. 그렇다면 어떤 사람이 장염일 때 그가 계속 설사와 토를 하는 확률을 P(e/h)=0.9, 그리고 P(e)=0.2, P(h)=0.1이라 할 경우, 어떤 사람이 계속 설사를 하며 토를 할 때 그가 장염을 앓고 있을 확률 P(h/e)=(0.1 * 0.9) / 0.2 = 0.45가 될 것이다.

그러나 이것은 확률의 기본 공리로부터 주어지는, 관련 확률들간의 상대적인 관계를 나타내 주는 정리일 뿐, 그 자체로 어떤 특정 확률값을 규정해 주는 것은 아니다. 그러므로 사실상 증거 언명 e가 주어져 있을 때 가설 언명 h에 대한 확률 P(h/e)를 구하기 위해서는 위의 식 (1)에 있어 우변의 각 확률을 별도로 계산해 낼 필요가 있다.

그러나 이 각각의 확률을 어떻게 구할 것인가? 이와 같은 경우 타가드는 무엇보다 일반적인 과학 이론의 경우 그러한 확률의 계산이 매우 어려움을 지적하고 있는 셈이다. 예컨대 소리의 파동 이론(wave theory of sound)에 대해 베이즈의 정리를 적용하기 위해 관련 확률들을 계산해 본다 하자. 소리의 파동 이론이란 간략히 말해 우리가 알고 있는 소리란 일종의 파동 형태를 띠고 있음을 주장하는 이론을 말한다. 그러므로 그것은 소리가 반사, 전파, 통과 등의 현상을 보이는 까닭을 그것의 파동성(波動性)에 의해 설명, 예측하려 하고 있다. 이와 같은 경우 만일 반사, 전파, 통과 등의 모든 현상에 관한 언명을 e라 하면, 그에 대한 원인으로서의 파동에 대한 가설 언명은 h라 할 수 있을 것이다. 그러나 사실상 이러한 경우 베이즈의 정리를 이용하기 위한 관련 확률들을 구하기란 매우 어려워 보인다. 무엇

보다 파동 이론과 같은 질적(質的, qualitative)인 이론에 있어서는 해당 현상들이 거의 문제의 가설에 의해 설명되는 것으로 상정되고 있으며, 만일 어떤 반례(反例)의 현상이 발견되는 경우 적어도 논리적으로는 문제의 가설에 대한 반증 사례로 간주되고 있으므로, 이러한 경우 어떤 경험적 의미에 있어 상대 빈도(相對 頻度, relative frequency)와 같은 확률 계산이 어렵기 때문이다.

물론 확률을 이와 같은 경험적 빈도가 아닌, 도박에 있어 합리적 투기율(投機率, betting quotient)에 기반을 둔 주관적 신념도(degree of subjective belief)로 해석하는 경우도 가능하다. 예컨대 어느 두 사람이 각기 a와 b의 판돈을 놓고 어떤 가설 h에 대해 도박을 거는 경우(물론 여타 증거 언명을 기반으로), 전자가 문제의 가설에 대해 부여하는 주관적 확률치는 a/(a+b)로 둘 수 있을 것이다. 그러나 비록 이러한 주관적 확률이라 할지라도 위와 같은 질적 이론에 대해 일종의 도박을 건다는 일은 불합리한 것으로 보인다. 불합리한 도박이 되지 않기 위해 적어도 확률의 기본 공리를 따른다 할 때, 유한한 증거 언명을 기반으로 한 보편 언명의 가설에 대한 확률값은 어느 경우에나 0이 되기 때문이다.[9]

이와는 달리 의학적 진단의 경우에는 최소한도 일정한 가설을 참으로 가정할 때 실제의 증거적 사례의 빈도를 구하는 일이 상대적으로 훨씬 용이하다. 예컨대 앞의 장염의 예의 경우 우리는 실제에 있어 장염 환자 중 몇 %나 계속 설사와 토를 하는가의 빈도를 계산해 낼 수 있기 때문이다. 여기에 있어서는 확실히 주어진 증상과 그 원인에 대한 가설 사이에, 앞서의 파동 이론에서와 같은 어떤 필연적인 설명의 관계가 발견되지 않은 것으로 보인다. 즉 장염이라고 해서 반드시 계속적인 설사와 토의 증상을 나타내는 것도 아니고, 또 장염에 의해서만 계속적인 설사와 토의 증상이 야기되는 것도 아닌 어떤 불확실한 상황(uncertain circumstances)인 것이다. 무엇보다 의학적 진단과 같은 상황하에서 확률을 도입하는 이유는 바

9) J. H. Fetzer, *Scientific Knowledge*, Dordrecht: D. Reidel, 1981, p. 217 참조.

로 이 때문이다.

그러므로 질적인 이론과 의학적 진단과 같은 상황에 대한 확률 부여에 관해 이상의 난점을 인정하는 한, 최소한도 전문가 시스템 내에 있어 확률을 이용한 가설 평가에 있어 전자(前者)에 대한 평가는 배제되어야 하리라 본다. 이 점은 질적 이론 역시 하나의 가설로 본다 할 때, 일견 의학적 진단의 상황과 동일하게 그것 역시 전문가 시스템 내에서 확률적 평가에 놓일 수 있다는, 어쩌면 있을 수 있는 그 적용 범위에의 확대에 대한 하나의 제한이다.

이와 관련하여 새먼(W. C. Salmon)은 일찌기, 전문가 시스템과는 무관하게, 질적인 이론 역시 베이즈의 정리에 의해 확률적 평가가 가능하리라 그 가능성을 제시한 바 있다. 그러나 그는 과연 관련 확률들을 어떻게 계산해 낼 것인가에 관한 받아들일 만한 구체적인 방법을 제시하지는 못하였다.[10] 이제 전문가 시스템을 논함에 있어서는 베이즈 정리를 적용함에 있어 관련된 확률치들을 구체적으로 계산해 내는 일이 필요하다. 그러나 소리의 파동 이론에 관한 위의 논의를 통해 볼 때, 전문가 시스템 내에 있어 그와 같은 확률 계산은 어려운 것으로 보인다.

이로써 전문가 시스템 내에서 확률을 이용해 가설을 평가함에 있어서는 적어도 주어진 현상이나 증상 및 그에 대한 원인으로서의 가설 사이에 불확실한 상황하에서의 통계적 분포(statistical distribution)가 가능한 경우로 그 범위가 제한될 필요가 있다.

2. 가설의 사전 확률과 우도

앞절을 통해 지적한 대로 비록 그 가설을 해당 증거와의 사이에 일정한 통계적 분포가 형성 가능한 경우로 제한을 한다 할지라도, 전문가 시스템

10) W. C. Salmon, *The Foundations of Scientific Inference*, Pittsburgh: Univ. of Pittsburgh Press, 1967, p. 115ff., 특히 p. 128 참조.

내에서 베이즈의 정리를 이용함에 있어서는 아직 해결해야 할 또 다른 문제가 남아 있다. 과연 증거 언명 e나 가설 언명 h 자체에 대한 확률 P(e)나 P(h)를 어떻게 계산해 낼 것이냐 하는 문제이다.

이 문제를 적절히 다루기 위해서는 먼저 통계학에 있어 베이즈의 정리를 적용하는 일반적 용례(用例)를 검토하는 일이 도움이 될 수 있다.

이제 A, B 두 개의 상자가 있다라고 해보자. 상자 A에는 같은 크기, 같은 재질의 붉은색 공(R)이 1개, 흰색 공(W)이 2개, 상자 B에는 붉은색 공이 5개, 흰색 공이 2개 들어 있다라고 해보자. 이러한 상황하에서 만일 어느 상자인지는 모르나 둘 중 어느 하나에서 붉은색 공 하나를 끄집어내게 되었다라고 해보자. 그렇다면 이미 주어진 붉은색 공 하나를 기반으로 우리는 그것이 어느 상자에서 추출된 것인가를 알아 내고자 한다. 그러나 우리에게 주어진 그 공이 특별히 상자 A나 B에 속했던 공이었음을 보여주는 특정의 징표가 없는 한, 그것이 상자 A나 B 어느 쪽에 속한다 확실히 말할 수 있는 상황은 아니다. 즉 그 결과가 우리에게 알려져 있지 않은 불확실한 상황인 셈이다. 그럼에도 불구하고 이러한 상황하에서라도 그 붉은색 공이 어느 상자에서 나왔는가를 추정해 볼 수 있는 확률적 계산은 가능하다.

이제 그 문제의 붉은색 공이 상자 A로부터 나왔을 것이라는 가설을 h_A로, 상자 B로부터 나왔을 것이라는 가설을 h_B로 나타내 보기로 하자. 문제의 붉은색 공이 나왔다는 경험적 사실에 관한 증거 언명을 e라 할 때, 증거 언명 e를 기반으로 가설 h_A와 h_B의 확률은 각기 다음과 같이 계산할 수 있다.[11]

11) 지금의 예와 같은 경우 일반적인 통계학 문헌에서는 흔히 문제의 사상(事象, event), 예컨대 붉은색 공 하나를 꺼내는 사상에 대해 확률 부여를 하고 있으나, 이와 같은 확률값은 대응하는 언명에 대해서도 그대로 부여 가능하다.

$$P(h_A/e)=\frac{P(h_A\cdot e)}{P(e)}=\frac{P(h_A\cdot e)}{P(h_A\cdot e)+P(h_B\cdot e)}=\frac{P(h_A)\cdot P(e/h_A)}{P(h_A)\cdot P(e/h_A)+P(h_B)\cdot P(e/h_B)} \quad (1)$$

$$P(h_B/e)=\frac{P(h_B)\cdot P(e/h_B)}{P(h_A)\cdot P(e/h_A)+P(h_B)\cdot P(e/h_B)} \quad \text{(위와 마찬가지 방식으로)} \quad (2)$$

그러나 위의 두 식이 일정한 수치를 갖기 위해서는 각각의 우변에 있어 각 확률의 값이 계산될 필요가 있다. 여기에 있어 우선 $P(e/h_A)$나 $P(e/h_B)$는 위에 제시된 가정만으로 쉽사리 계산 가능하다. 즉 전자의 경우에는 상자 A에서 붉은색 공 하나를 꺼내는 경우의 확률이므로 즉시 1/3의 값을 얻을 수 있고, 후자의 경우에는 즉시 5/7의 값을 얻을 수 있다. 그러나 $P(h_A)$나 $P(h_B)$의 경우에는 어떻게 할 것인가 ? 이것이야말로 어떤 증거가 주어지기 이전의 가설 h_A나 h_B의 사전 확률로서, 상자 A나 B 중 하나를 선택하는 확률을 말한다. 그러나 어떤 확률로써 그것을 선택할 것인가 ?

이러한 선택의 난점은 무엇보다 그것이 경험적 증거와 같은 객관적 기준 없이 이루어져야 한다는 데 기인한다. 이 경우 만일 어떠한 기준 없이 주어진 가설을 선택해야 한다는 입장에서, 각 가설에 대해 동일한 확률값을 부여하는 이른바 "무차별의 원리"나 "불충분 이유율" (principle of indifferncе or insufficient reason)에 따라 그 각각의 확률을 1/2로 삼아야 한다고 볼지 모른다. 그러나 이것의 자의성은 일견 명백한데, 만일 지금과 같이 있을 수 있는 가설을 단지 두 가지로 제한하는 대신 그 이상의 여러 가지로 제시를 한다면, 이 경우 단지 임의로 각 가설의 사전 확률값이 변화할 수 있기 때문이다.

그러므로 우리가 만일 이와 같은 사전 확률 계산의 난점을 인정한다면, 그러한 사전 확률을 배제한 채, 상대적으로 훨씬 그 계산이 용이한 확률

$P(e/h_A)$나 $P(e/h_B)$만으로 확률 $P(h_A/e)$나 $P(h_B/e)$를 비교해 볼 수 있는 가능성을 모색할 필요가 있다.

이러한 가능성은 위의 식 (1)과 (2)에 있어 그 상대적인 비(比)를 고려함으로써 열릴 수 있다. 이를 위해 이제 식 (1)과 (2)를 변변 나누어 보기로 하자.

$$\frac{P(h_A/e)}{P(h_B/e)} = \frac{P(e/h_A)}{P(e/h_B)} \cdot \frac{P(h_A)}{P(h_B)} \qquad (3)$$

이 경우 위의 식 (3)에서는 일단 $P(e)$의 확률 부분은 사라짐을 알 수 있다. 물론 앞서 문제가 되었던 $P(h_A)$나 $P(h_B)$는 그대로 남아 있기는 하나, 그럼에도 불구하고 이것은 상대적인 비로 나타나고 있다. 이러한 사실은 지금의 경우 매우 중요한데, 이렇게 상대적인 비를 고려함으로써, 앞서 문제되었던 무차별의 원리를 고려한다 할지라도 그것은 위의 식 (3)의 결과에 아무런 영향도 미치지 못하기 때문이다. 예컨대 가설의 수가 임의적으로 늘어나 2, 3, 4, ... 등이 된다 할지라도, 언제나 $P(h_A)/P(h_B) = 1$일 뿐이기 때문이다. 물론 이와 같이 무차별의 원리를 고려치 않고, 가설 h_A나 h_B 중 어느 것을 선호(選好)할 여타의 합리적 기준이 존재하여, 이에 따라 h_A나 h_B에 대한 나름의 확률치 부여가 가능하다면, 그에 따라 $P(h_A)/P(h_B)$의 구체적 계산도 가능하고,[12] 이것은 그대로 베이즈의 정리를 이용한 위의 식 (1)과 (2)에 대해서도 적용 가능하다.

그러므로 적어도 베이즈의 정리에 있어 해당 가설에 대한 사전 확률값을 구하기 위해 무차별의 원리를 적용함에 문제가 있다면, 위의 식 (1)과

12) 어쩌면 이것은 주관적 확률에 의해 계산될 수도 있고(예컨대 M. R. Genesereth & N. J. Nilsson, *Logical Foundations of Artificial Intelligence*, Los Altos:Morgan Kaufmann, 1987, 177-86 참조), 언어나 개념 체계에 기반을 둔 논리적 확률에 의해 계산될 수도 있다(예컨대 R. Caranp, op. cit. ; "A Basic

(2)의 방식보다는 위의 식 (3)의 방식을 이용하는 것이 훨씬 유리할 수 있다. 식 (3)에서는 무차별의 원리를 적용하되, 어떤 불합리한 결과에 빠지지 않기 때문이다. 물론 가설 선택에 관해 근본적으로 무지(無知)한 상태에서는 해당 가설에 대해 아무런 확률치도 부여할 수 없다는 반론도 있을 수 있으나,[13] 해당 가설에 대해 아무런 배경 지식이나 사전(事前) 지식이 없는 상태에서 그에 대해 동일한 확률값을 부여하여, 위의 식 (3)에서와 같이 그것의 효과를 무력화(無力化)하는 일은 그러한 반론에도 적절한 답이 될 수 있다.

만일 이와 같은 식으로 위의 식 (3)에 있어 $P(h_A)/P(h_B)$의 요소를 무시한다면, 이제 식(3)에 있어 문제가 되는 것은 $P(e/h_A)$와 $P(e/h_B)$이다. 그러나 이것들은 이미 앞 1절에서 살펴본 바대로 쉽사리 계산 가능한 것들이다.

사실상 지금의 이 확률값들은 통계학에 있어 이른바 "우도"와 관련된 것들이다. 일상적 용어로 말해 통계학에서의 우도란, 이미 주어진 표본적 증거에 비추어 보았을 때 모집단에 관해 어떤한 통계적 가설이 "그럴 법"(likely)한가를 말해 주는 정도를 말한다. 이러한 정도는 일단 일정한 통계적 가설을 전제하였을 때 그러한 전제하에서 우리에게 주어져 있는 문제의 증거가 나타날 수 있는 정도를 말하는 것으로, 그 정의상 이러한 정도는

System of Inductive Logic, Part, I" in:R. Carnap & R. C. Jeffrey (eds.), *Studies in Inductive Logic and Probability*, Vol. I, Los Angeles: Univ. of California Press, 1971, pp. 33-165, Part II, in:R. C. Jeffrey (ed.), ibid, Vol. II, 1980, pp. 7-155). 또한 어렵기는 하지만, 상대적 빈도로서의 빈도적 확률로 계산될 가능성이 없는 것도 아니다(예컨대 W. C. Salmon, op. cit., pp. 123-6 참조). 그러나 단순히 계산될 수 있다는 것과 그것에 자의성이 없다는 것은 다른 것이다. 만일 경험만을 합리적 기준으로 삼는다면, 어떤 경험적 증거가 주어지기 이전에 부여된 이와 같은 확률들에 있어 자의성은 배제되기 어려운 듯 보인다(W. C. Salmon, op. cit., p. 128 참조). 이에 대한 본격적인 논의는 다른 기회로 돌리기로 한다.

13) A. W. F. Edwards, *Likelihood*, Cambridge:Cambridge Univ. Press, 1972, p. 59.

문제의 가정된 가설(h)을 기반으로 증거 언명(e)의 확률에 비례하는 것으로 주어질 수 있다.[14] 즉 다음과 같은 식이 가능하다.

$$L(h/e) = k\,P(e/h)\ ,\ \text{단 } k\text{는 비례상수} \qquad (4)$$

이제 우도에 관한 이와 같은 정의를 채택하고, 가설 h_A와 h_B에 대한 사전 확률의 비(比)의 효과를 무시한다면, 위의 식 (3)은 다음과 같이 바꿔 쓸 수 있다.

$$\frac{P(h_A/e)}{P(h_B/e)} = \frac{P(e/h_A)}{P(e/h_B)} = \frac{L(h_A/e)}{L(h_B/e)} \qquad (5)$$

그렇다면 결국 지금까지 논의된 상황하에서 통계적 가설의 평가 문제는 각 가설에 대한 우도비(尤度比, likelihood ratio)의 확정 문제로 대치될 수 있다. 그리고 이것은 일단 해당 가설에 대한 사전 확률의 확정에 비해 문제가 없는 것으로 보인다.

그러나 이와 같은 우도비를 이용한다 할지라도 이것을 전문가 시스템 내에서 사용하기 위해서는 또 다른 고려 사항이 필요하다.

3. 가설의 지지도와 추가적 증거

우도 함수 L을 앞절의 식 (4)에서와 같이 정의하는 한, 그것이 그 자체로는 확률의 기본 공리를 벗어나고 있음은 분명하다. 그렇다면 우도 L(h/e)이 확률 P(h/e)와 달리 의미하는 바는 무엇인가 ?

---

14) R. A. Fisher, *Statistical Methods and Scientific Inference*, New York: Hafner, 1956, p. 68.

앞절에서 언급한 대로 우도란 일단 주어진 표본적 증거에 비추어 모집단에 관해 어떠한 통계적 가설이 그럴 법한가를 말해 주는 정도를 말한다. 즉 문제의 증거는 해당 가설을 전제로 할 때 어느 정도 나올 법한가를 말해 주는 것이다. 그렇다면 이제 이것을 증거의 입장에서 본다면, 문제의 증거는 해당 가설을 바로 그 정도만큼 "지지"해 준다고 말할 수 있다. 이것은 "지지"에 대한 우리의 직관적 개념에 기인한 것으로, 만일 동일한 증거 e를 기반으로 가설 h와 h'에 대한 각각의 우도에 있어 L(h/e)가 L(h'/e)보다 크다면, 증거 e는 가설 h'보다는 가설 h를 더 높게 지지한다고 말할 수 있다는 것이다.[15] 그러므로 이러한 의미에서 어떤 가설에 대한 우도는 주어진 증거를 기반으로 그 가설에 대한 지지도(支持度, measure of support)로서 간주할 수 있다. 이러한 우도가 확률 P(h/e)와 상이한 주요한 점 중 하나는, 확률의 경우 만일 확률 P(h/e)가 높다면, 가설 h에 대한 대립 가설 ~h (not-h)에 대한 확률은 낮을 수 밖에 없는 반면, 우도의 경우에는 경우에 따라 그 각각의 우도 L(h/e)와 L(~h/e)가 동시에 높을 수 있다는 점이다. 동일한 증거에 대해 그러한 증거가 나타날 수 있는 높은 확률을 갖는 가설들을 얼마든지 상정할 수 있기 때문이다.

그러나 어떤 통계적 가설 평가에 있어 그러한 가설에 대한 사전 확률 부여의 문제점을 피해 이와 같은 우도를 도입하게 되면, 전문가 시스템 내에서 풀어야 할 또 다른 문제에 부딪치게 된다. 일종의 지지의 정도로서 우도를 고려하게 되면, 새로운 증거가 시스템 내로 유입될 때마다 관련 가설에 대한 우도를 독립적으로 계산해 내야 하고, 이러한 독립성으로 인해 이전에 계산된 우도는 전혀 고려하지 못하는 상황에 빠지게 되는 것이다. 즉 시스템 내로 새로운 증거가 새로운 데이타로서 유입되는 경우, 시스템 내에서는 그러한 증거와 관련된 모든 가설의 우도를 계산하여 그 중 최대의 우도를 지닌 가설을 찾아내야 할 것이고, 이렇게 추가된 증거를 기반으로

15) 우도 개념에 대한 이와 같은 해석에 관해서는 I. Hacking, *Logic of Statistical Inference*, Cambridge: Cambridge Univ. Press, 1965, Ch. V 참조.

한 우도는 문제의 추가 증거가 있기 이전의 우도와는 아무 관련이 없는 것이다. 이 점은 사전 확률에 의한 베이즈의 정리를 이용할 경우와는 대비되는 점인데, 그 경우 첫 단계에서 계산된 사후 확률(事後 確率, posterior probability) P(h/e)는 새로운 추가 증거 e'가 유입되는 다음 단계에서 그 자체 다시 사전 확률로서 사용될 수 있기 때문이다.

그러므로 가설 평가에 있어 우도를 이용시 이러한 문제를 해결하기 위해서는 또 다른 새로운 전략이 요구된다. 이러한 전략으로서 보오간은 다음과 같은 두 가지 가능성을 제시한 바 있다. (i)각 가설과 관련된 증거(데이타)들을 통합하고, 그것을 커다란 하나의 증거 집합으로 간주하는 방법, (ii)데이타 베이스(data base) 내에서 각각의 관련 증거 집합에 관해 개별적으로 계산된 우도를 통합하는 방법.[16]

이러한 가능성 중 전자(前者)는 전문가 시스템 내에서 현실적 이유 때문에 난점에 부딪치게 된다. 이것은 각 가설과 관련된 모든 증거들을 하나로 통합해야 하므로, 각 가설과 관련된 증거들을 모두 데이타 베이스 내에 저장하거나, 그 어느 증거가 어느 가설과 관련되는가를 지시해 주는 포인터(pointer)의 리스트를 저장해야 하는 커다란 부담을 안게 되는 것이다. 더군다나 이 경우에는 새로운 추가 증거가 유입될 때마다 각 가설과 관련된 모든 증거들을 다시금 전체적으로 계산해 내야 할 필요가 있는데, 이것은 시간과 비용에 있어 매우 비효율적인 것이다.

이와는 달리 후자(後者)의 경우에는 기존의 증거들에 비해 새로이 추가된 증거에 의한 지지도의 상대적 가중(加重)의 정도를 알 필요가 있다. 이미 계산된 우도를 기반으로, 새로이 추가된 증거에 의한 우도를 통합할 필요가 있기 때문이다. 그러나 무엇에 의해 그와 같은 가중의 정도를 정할 것인가는 또 다른 문제가 될 수 있다. 이 점에 관해 보오간은 주어진 증거의 표본 크기(sample size)에 의해 그것을 정할 수 있으리라 제안한 바 있다. 예컨대 과거에 통합된 표본의 크기가 900이고, 새로운 증거의 표본의

16) Op. cit., p. 329.

크기가 100이라면, 새로 통합된 표본의 크기는 900+100=1,000이고, 이에 의하면 과거의 우도에 대한 가중치가 900일 때, 새로운 증거에 의한 우도의 가중치는 100이 될 것이다. 따라서 이 경우 만일 지금까지 통합된 우도가 0.8이고, 새로운 증거에 의한 우도가 0.7이라면, 새로이 통합된 우도는 다음과 같이 계산된다(여기서 e는 기존의 통합된 증거, e'는 추가적 증거, W는 기존의 통합된 표본 크기에 의한 가중치, W'는 추가적 증거의 표본 크기에 의한 가중치를 나타낸다).

$$L(h/e \cdot e') = \frac{L(h/e) \cdot W + L(h/e') \cdot W'}{W + W'} \qquad (1)$$

$$= \frac{(0.8 * 900) + (0.7 * 100)}{900 + 100} = 0.77$$

지금의 방법은 새로이 추가된 증거에 의한 우도에 가중치를 고려함으로써 확실히, 첫번째 방법에서와는 달리, 각 가설과 관련된 모든 증거를 하나로 통합할 필요도 없을 뿐더러, 전체 증거에 의해 다시금 우도를 계산해 내야 하는 번거로움도 회피하고 있다. 그러나 이 경우에는 문제의 가중치를 얼마나 타당하게 설정할 것이냐에 관한 새로운 문제가 대두된다. 위에 지적한 대로 보오간은 주어진 증거의 표본 크기에 의해 이 문제를 해결하고 있으나, 과연 주어진 증거들의 중요성이 그 표본 크기에 의해서만 주어지는가는 의문이다. 만일 거의 동일한 종류의 추가적 증거들이라면, 이 때 그것들의 중요성은 단지 그와 같은 표본 크기에 의해서만 평가되도 좋을지 모른다. 그러나 좀더 중요한 추가적 증거 사례는 이전의 증거 중에는 나타나 있지 않던 새로운 속성, 예컨대 앞서 장염의 예에서라면 고열(高熱)과 같은 새로운 증상이 장염과 결부되어 나타나는 경우이다. 그러므로 만일 단순히 거의 동일한 증거들만의 추가가 아닌, 진정으로 새로운 증거가 추

가되는 경우에는 그 표본의 크기뿐 아니라, 그와 같은 새로움을 고려하는 새로운 가중치가 필요하다. 이를 위해 논자는 어떤 증거가 주어져 있을 때, 그것의 표본 크기뿐 아니라, 그 증거에 있어 문제가 되고 있는 특징적 속성들의 개수까지를 고려할 것을 제안한다. 만일 이전의 증거들에 있어 그러한 속성들의 총개수를 P, 새로운 증거에 있어 문제의 속성들의 개수를 P′ 라 하면, 그 각각에 있어 표본의 크기를 S 및 S′ 라 할 때, 그 각각의 가중치는 다음과 같이 수정할 수 있다(여기서 P와 P′ 에 각기 1을 더한 이유는 P나 P′ 가 1인 경우에는 그것을 곱하는 효과가 사라지기 때문이다).

$$W = S * (P + 1)$$
$$W' = S' * (P' + 1) \qquad (2)$$

이제 이와 같은 새로운 가중치를 사용한다면, 만일 새로운 증거가 이전의 증거들과 거의 동일하여, P = P′ 인 경우에는, W = S, W′ = S′ 인 원래의 식 (1)로 환원이 될 것이다. 그러므로 식 (2)와 같은 새로운 가중치를 사용한 통합된 우도식은 위의 식 (1)의 확장이라 할 수 있다.

4. 우도와 증거의 신뢰도

이상에서 논한 바와 같이 우도를 이용해 가설을 평가하는 경우에는 확실히 최대의 우도를 갖는 가설을 선택해야 할 것이다. 물론 이 때라도 그 최대 우도를 받아들일 것이냐 아니냐의 여부에 관해서는 또 다른 약속된 기준이 필요할지 모른다. 그러나 만일 지금의 단계에서 그러한 기준이 설정되고, 따라서 최대 우도치에 의해 어떤 가설을 선택한다 할지라도, 여기에는 우도의 성격상 또 다른 문제가 남아 있다.

이미 페처(J. H. Fetzer)가 지적한 대로,[17] 어떤 증거와 관련한 최대 우

17) Op. cit., p. 231.

도를 지닌 가설은 원리상 얼마든지 가능하다. 모집단의 다양한 분포에도 불구하고 문제의 증거가 나타날 수 있는 가능성이 모두 동일할 수 있기 때문이다. 예컨대 앞의 3절에서 지적한 대로 어떤 가설 h와 그 대립 가설 ~h에 있어 그 각각의 우도 L(h/e)와 L(~h/e)가 동시에 높을 수 있는 것이다. 따라서 만일 이와 같은 가능성을 인정한다면, 단순히 최대 우도를 지닌 가설을 선택하는 대신에 이제 또 다른 선택의 기준이 필요하다.

이와 관련하여 페처는 그 자신 하나의 해결책을 제시하였는데, 그것은 주어진 증거에 주목하는 일이었다. 즉 같은 우도를 갖는 가설이라 할지라도 그것이 얼마나 신뢰할 만한 증거에 의한 우도인가에 주목하는 일이다. 포퍼(K. R. Popper)의 지적대로,[18] 주어진 증거에 대해 동일하게 적용될 수 있는 가설일지라도, 만일 그 증거가 문제의 가설에 대해 테스트를 위한 진지한 시도(sincere attempt)의 결과가 아니라면, 그 가설들의 동일한 확인의 정도(degree of corroboration)란 우리를 기만(欺瞞)할 수 있다.

이와 같은 "진지한 시도"를 실현하기 위한 실제적 방안은 무엇보다 다양한 상황으로부터 문제의 증거를 수집하는 일이다. 그러나 전문가 시스템 내에서라면 그와 같은 수집 행위 자체는 불가능하므로, 그렇게 수집된 증거의 분포 상황을 검토하는 일이 필요하다. 통계학적으로 볼 때 이와 같은 증거들은 가능한 한 무작위적(無作爲的, random)으로 분포되어 있는 것이 바람직하며, 가설의 테스트를 위해서는 이와 같은 증거가 신뢰할 만하다 할 수 있다.

자연 상태에서 무작위적 표본 추출이 이루는 통계적 분포 중 가장 흔한 것은 정규 분포(normal distribution)이다. 더군다나 이른바 "중심 극한 정리"(central limit theorem)에 의하면,[19] 분산 σ^2과 평균 θ를 지닌 어

18) K. R. Popper, *The Logic of Scientific Discovery*, New York:Harper & Row, 1968, p. 418.

19) 이 정리에 대한 증명을 위해서는 R. V. Hogg & E. A. Tanis, 안철원 · 한상문 공역, op. cit., pp. 355-9 참조.

떠한 분포로부터 추출한 무작위 표본의 표본 평균이라도 근사적으로는 정규 분포에 가까이 다가가게 된다. 그러므로 만일 이와 같은 사실을 원용(援用)한다면, 어떤 증거가 얼마나 신뢰할 만한가 하는 것은 그것이 이루는 분포가 얼마나 정규 분포로부터 멀리 떨어져 있는가를 측정함으로써 측정 가능하다.

페처는 이와 같은 이탈의 정도(degree of divergence)를 측정하기 위해 다음과 같은 방식을 취하고 있다.[20] 예컨대 우리의 증거가 이루는 분포의 분포 함수를 F(x), 정규 분포의 함수를 N(x)라 하면, 모든 x와 어떤 작은 값 ε에 대해 다음과 같은 식이 가능하고,

$$N(x-\varepsilon)-\varepsilon \leq F(x) \leq N(x+\varepsilon)+\varepsilon \quad (1)$$

이 때 직관적인 특정의 몇몇 조건[21]을 만족시키는 가장 작은 ε를 그 증거의, 정규 분포로부터의 이탈의 정도로 삼는 것이다.

그러므로 만일 분포 함수 F(x)에 대한 사전 신뢰도(事前 信賴度, prior degree of confidence)를 1이라 한다면, 위의 이탈의 정도를 측정한 후의 사후(事後, posterior) 신뢰도는 $1-\varepsilon$ 로 삼을 수 있을 것이다.

이제 만일 이와 같은 식으로 증거의 신뢰도를 새로 도입한다면, 앞 3절의 식 (1)을 통한 통합된 우도의 식 역시 수정할 필요가 있다. 물론 식 (1)에 있어 수정할 부분은 추가적 증거 e'에 의한 가설 h의 우도 L(h/e')이다. 그러나 증거의 신뢰도 $1-\varepsilon$ 와 이것을 어떻게 관련시킬 것인가?

보오간은 이 문제를 해결하기 위해 직접 신뢰도 $1-\varepsilon$ 를 고려하는 대신에 증거의 이탈 정도 를 이용해 다음과 같은 관계를 고려하고 있다.[22] 즉 이탈의 정도가 낮으면 신뢰도가 높고, 전자가 높으면 후자가 낮으므로, 증

20) Op. cit., pp. 250-3.
21) Ibid., p. 251 참조.
22) Op. cit., pp. 330-1.

거의 이탈 정도와 우도 역시 역(逆)의 관계에 있다는 것이다. 물론 이와 같은 역의 관계를 수식화하기 위해서는 해당 우도를 이탈의 정도 ε 로 나누는 것이 적절할 수 있다. 다만 $\varepsilon = 0$인 경우, 즉 증거가 완전히 정규 분포를 이루는 경우에는 분모가 0이 될 수 있으므로, 보오간은 분모에 1을 더할 것을 제안하고 있다. 이로써 이탈의 정도 ε 를 고려한 새로운 우도 $L'(h/e')$는 다음과 같이 구할 수 있다.

$$L'(h/e') = \frac{L(h/e')}{1+\varepsilon} \qquad (2)$$

이 경우 만일 $\varepsilon = 0$인 경우 식 (2)의 좌변과 우변은 동일하게 되어 우도값에 있어 아무런 변화도 없을 것이다.

이것은 물론 하나의 좋은 해결책이다. 그러나 위와 같이 이탈의 정도 ε 만을 사용해 식 (2)에 이를 필요는 없다. 이미 신뢰도를 $1-\varepsilon$ 로 계산해 낼 수 있는 가능성이 있으므로, 이를 이용해도 역시 유사한 효과를 얻을 수 있다. 이제 이와 같은 신뢰도를 이용할 경우에는 그 자체 새로운 우도 L'와 비례할 수 있으므로 다음과 같은 새로운 식이 가능하다.

$$L'(h/e') = L(h/e') * (1-\varepsilon) \qquad (3)$$

이 경우에도 역시 $\varepsilon = 0$인 경우에는 우도값에 아무런 변화가 없을 것이다.

그러므로 위의 식 (2)나 (3), 그리고 앞 3절에서의 식 (2)를 적용하여 같은 절의 식 (1)을 최종적으로 정리하여 제시하면 다음과 같다.

$$L(h/e \cdot e') = \frac{L(h/e) * (P+1) + L'(h/e') * (P'+1)}{S * (P+1) + S' * (P'+1)} \qquad (4)$$

5. 베이지언 조건화와 우도

앞 2절에서의 논의대로 가설의 평가에 있어 우도의 이점(利點)을 살펴보았음에도 불구하고, 그러나 보오간은 베이즈의 방법이 우도의 방법에 비해 반드시 불리한 것은 아니라 보고 있다.

무엇보다 베이즈의 정리를 이용시 최대의 약점인, 해당 가설에 대한 사전 확률의 주관성 문제는 베이즈의 정리를 거듭 적용함으로써, 즉 이른바 "베이지언 조건화" (Bayesian conditionalization)를 거듭 행함으로써 그 문제성이 감소될 수 있다는 것이다.[23]

그러므로 만일 그와 같은 조건화의 횟수를 계산해 주는 어떤 카운터(counter)를 각 가설과 결부시키면, 그 사후 확률뿐 아니라 그러한 카운터가 일정한 기준을 넘어설 때 문제의 가설을 채택할 수 있다는 것이다.

그러나 이러한 조건화의 횟수를 증가시킴으로써 사전 확률의 문제성을 감소시킬 수 있다 하더라도 문제의 가설에 대해 그 최초의 사전 확률을 부여하는 일은 여전히 자의적일 수 밖에 없다. 과연 높은 사전 확률을 부여할 것인가 낮은 사전 확률을 부여할 것인가 ?

이 문제에 관해 보오간은, 새먼의 제안에 따라,[24] 낮은 사전 확률값을 부여할 것을 권고하고 있다.[25] 베이즈의 정리에 있어 그 우변에 나타나는 분수(分數)의 성격상, 만일 해당 가설 h에 대한 우도 $L(h/e)=P(e/h)$〔이 경우 상수 k=1〕가 높고, 그와 대립되는 가설인 $\sim h$에 대한 우도 $L(\sim h/e)=P(e/\sim h)$가 매우 낮아 거의 0에 가까운 경우에는, 아주 낮은 사전 확률값으로부터도 높은 사후 확률값이 나올 수 있기 때문이다. 지금의 경우에는 그 분모값이 매우 작아지기 때문이다. 반면 우도 $L(h/e)$가 낮고 우도 $L(\sim h/e)$가 매우 높아 거의 1에 가까운 경우에는 높은 사전 확률값으로부

23) Ibid., p. 327 ; 또한 W. C. Salmon, op. cit., pp. 128-9 참조.
24) W. C. Salmon, ibid..
25) Op. cit., pp. 325-7.

터 언제나 낮은 확률값이 나온다 할 수는 없다. 이 경우에는 그 분모값이 중간이나 높은 값을 취할 수 있기 때문이다. 위의 두 경우를 비교해 볼 때 (이 때 어느 경우에건 그 극단적 값인 0이나 1의 값은 배제된다. 이 값들은 베이지언 조건화의 값을 언제나 0 또는 1로 만들어 줄 뿐이기 때문이다), "낮은 사전 확률을 높은 사후 확률로 올릴 때보다는 높은 사전 확률로부터 낮은 사후 확률로 내릴 때 훨씬 더 극단적인 값이 필요하므로, (결국) 낮은 사전 확률값이 바람직하다" [26]는 것이다.

그러나 이와 같은 이유에서 낮은 사전 확률값을 취한다 할지라도 구체적으로 어느 정도의 낮은 확률값을 취해야 하는가는 또 다시 남아 있는 문제이다. 이 문제에 관해 보오간은 그 최초의 사전 확률값은 문제의 가설을 테스트하게 될 최초의 증거 집합을 이용해 계산해 낼 수 있으리라 제안하고 있다. 즉 그렇게 주어진 최초의 증거 집합에 의한 가설의 우도를 바로 최초의 조건화를 위한 최초의 사전 확률값으로 볼 수 있다는 것이다.

만일 이렇게 된다면, 결국 베이즈의 조건화는 최초의 우도값으로부터 출발하여 그 조건화를 진행시키는 셈이다. 그러나 이 경우 만일 그 최초의 우도값이 매우 낮은 경우라면 위의 새먼이나 보오간의 제안대로 그러한 값이 적합할 수 있으나, 만일 그 우도값이 매우 높은 경우라면 바로 그들의 지적대로 적절한 조건화의 수행은 어렵게 된다.

더군다나 최초의 높은 우도값 중 극단적으로 그 값이 1인 경우에는 그 사전 확률값이 1이 되어, 앞서 지적한 대로, 금지된 사전 확률값에 이르게 된다. 그러므로, 보오간의 지적대로,[27] 이러한 경우를 해결하기 위한 단순한 한 가지 해결책은 그러한 극단적 값을 회피하기 위한 어떤 한계값을 설정하는 일이다. 그러나 어떠한 기준에 의해 그 한계값을 설정할 것인가? 여기에 있어 그 자의성은 그대로 남아 있는 셈이다.

그러나 비록 이와 같은 자의성을 인정한 채, 이제 만일 문제의 가설에

---

26) Ibid., p. 327.
27) Ibid..

대해 일정하게 낮은 사전 확률값을 부여했다 할지라도, 그 자체로 곧 높은 사후 확률값이 보장되는 것은 아니다. 앞서 지적했듯, 이러한 결과가 나오기 위해서는 해당 가설에 대한 높은 우도와 대립 가설에 대한 낮은 우도가 있지 않으면 안 된다. 그러므로 이 점에서 본다면, 결국 어떤 가설에 대해 높은 사후 확률값에 의해 그것을 채택했다면, 그것은 곧 해당 가설에 대한 높은 우도와 그에 대립하는 가설에 대한 낮은 우도에 기인하는 셈이다. 그렇다면 결국 직접 우도를 비교하는 대신에 베이즈의 정리라는 하나의 우회로를 택할 이유는 전혀 없다.

그럼에도 불구하고 보오간은 그의 논문의 결론적 부분에서 우도를 이용한 방법이 베이즈의 정리를 이용한 방법에 비해 완전한 이점을 갖는 것은 아니라 치적하고 있다. 우도를 이용한 방법은 사전 확률의 부여와 같은 자의성은 갖고 있지 않으나, 대신 그 방법을 적용함에 있어 주어진 증거들이 정규적으로 분포해야 한다는 가정이 깔려 있다는 것이다.[28] 그러나 비록 이러한 전제가 충족되지 않는다 할지라도, 그것이 곧 우도 계산에 있어 문제가 되는 것은 아니다. 앞서 제4절에서 살펴본 바대로, 합리적 기준에 의해, 주어진 증거들이 정규 분포로부터 얼마나 이탈되어 있는가를 계산해 내는 일이 가능하고, 이로써 문제의 우도값을 조정해 주면 될 따름이기 때문이다.

보오간은 언급하지 않았지만, 이 경우 오히려 문제삼아야 할 점은 증거들이 정규 분포에서 아주 멀리 이탈되어 있는 경우 모집단의 분포 형태를 무엇으로 보아야 하겠느냐 하는 점이다. 물론 중심 극한 정리에 의해, 표본의 개수를 한없이 증가시키면 그 분포는 정규 분포에 다가갈지 모른다. 그러나 현실적으로 그러한 극한에 한계가 지어질 경우에는 심한 이탈의 정도를 보이는 증거 자체에 주목하여, 다음의 두 경우를 고려할 필요가 있다. 즉 증거들이 정규 분포로부터 아주 멀리 이탈되어 나타나는 경우에는

28) Ibid., p. 334.

다음의 두 가지 이유에 기인할 수 있다. (i)의도적으로 편향된 증거를 수집한 경우, (ii)다양한 상황하에서 다양한 원천으로부터 수집된 증거일지라도 뚜렷이 정규 분포와는 다른 형태의 분포를 보이는 경우.

전자의 경우에는 물론 시스템 자체 내에서 시정할 수 있는 문제가 아니다. 오직 해당 우도를 낮추는 기능만을 행할 수 있을 뿐이다. 그러나 후자의 경우에는 과연 그러한 증거의 분포 형태가 모집단에 관해 무엇을 말해주는가를 고려해야만 한다. 예를 들어 균일 분포(uniform distribution) 형태를 취하고 있는 모집단으로부터 무작위로 표본 추출을 행했을 때 그러한 표본적 증거가 유한한 상태에서 정규 분포를 이루리라 보기는 어려울 것이다. 그러므로 위의 (ii)의 경우 그 증거들이 보이는 특이한 분포 형태에 관해서는 가설이 제기되고 있는 모집단과의 관계에 관한 새로운 고찰이 필요하다.

물론 이러한 관계 설정이 결정적일 필요는 없다. 다만 그러하리라는 가정만으로 충분하고, 이러한 가정은 계속적인 증거의 수집으로 그 자체 테스트 가능하다. 단지, 잠정적으로나마 그와 같은 관계 설정이 어떻게 이루어지는 것이 좋을 것인가에 관해서는 일종의 배경 지식(background knowledge)으로서 시스템 내에 별도로 저장해 두면 좋을 것이다.

마지막으로 보오간은 모든 상황 속에서 베이즈의 방법과 우도의 방법 중 그 어느 하나가 나머지에 대해 완전히 유리한 것은 아니라 지적하고, 전자는 새로운 증거의 유입이 많아 많은 조건화가 가능한 곳에서 효과적이고, 후자는 그러한 증거가 적은 경우에 효과적이라 말하고 있다.[29] 그러나 여러 이유로(예컨대 비용, 시간, 기계상의 문제 등등으로) 많은 증거의 수집이 어려운 경우에는 어떻게 될 것인가 ? 이 경우 후자의 방법은 적절히 이용될 수 있는 반면, 상대적으로 전자의 방법은 많은 문제를 안게 된다. 또한 비록 많은 증거의 유입으로 많은 조건화가 가능한 경우 전자의 방법

29) Ibid..

이 이용 가능한 곳에서는 언제나 후자의 방법 역시 이용 가능하다. 따라서 오직 제한적인 상황하에서만 그 자의성이 감소될 뿐인 베이즈의 방법을 이용하는 것보다는 그러한 제한을 넘어서는 우도의 방법이 역시 좀더 효과적이라 생각한다.

결 어

본 논문을 통해 논자는 보오간의 주장을 기반으로 전문가 시스템 내에서 우도의 활용 방법을 해명하면서, 특별히 다음과 같은 새로운 사실들을 밝힌 셈이다. (i)전문가 시스템 내에 있어 확률이나 우도를 이용해 평가될 수 있는 가설은 통계적 가설에 제한될 뿐이다, (ii)우도를 이용한 가설의 평가에 있어서는, 보오간이 지적한 증거의 크기뿐 아니라, 그러한 증거 중에 나타난 속성의 개수까지를 고려할 필요가 있다, (iii)가설의 평가에 있어 베이즈의 방법보다는 우도의 방법이 궁극에 있어 좀더 효과적이다.

이 경우 위의 (ii)와 (iii)의 사실은, 전문가 시스템 내에서 우도를 이용해 가설을 평가함에 있어 단지 해당 가설에 대해서뿐 아니라, 동시에 우리에게 주어진 증거에 대해서도 주의를 기울여야 함을 보여주고 있다. 적어도 우도가 이미 주어진 증거를 중시하는 하나의 측도(測度, measure)라면, 우도의 계산시 그와 같은 증거 자체에 대해 좀더 심각한 고려가 있어야 함은 당연하다.

〈참 고 문 헌〉

Carbonell, J. G. et al., "An Overview of *Machine Learning*," in: R. S. Michalski et al. (eds.), *Machine Learning*, Palo Alto: Tioga, 1983, pp. 3-23.

Carnap, R., *Logical Foundations of Probability* (1950), London: Routledge & Kegan Paul, 1951.

———, "A Basic System of Inductive Logic, Part I," in: R. Carnap & R. C. Jeffrey (eds.), *Studies in Inductive Logic and Probability*, Vol. I, Los Angeles: Univ. of California Press, 1971, pp. 33-165 ; Part II, in: R. C. Jeffrey (ed.), ibid., Vol. II, 1980, pp. 7-155.

Edwards, A. W. F., *Likelihood*, Cambridge: Cambridge Univ. Press, 1972.

Fetzer, J. H., *Scientific Knowledge*, Dordrecht: D. Reidel, 1981.

Fisher, R. A., *Statistical Methods and Scientific Inference*, New York: Hafner, 1956.

Genesereth, M. R. & Nilsson, N. J., *Logical Foundations of Artificial Intelligence*, Los Altos: Morgan Kaufmann, 1987.

Hacking, I., *Logic of Statistical Inference*, Cambridge: Cambridge Univ. Press, 1965.

Hogg, R. V. & Tanis, E. A., *Probability and Statistical Inference*, 2nd ed., New York: Macmillan, 1983, 안철원 · 한상문 공역, 「통계학 원론」, 서울: 비봉, 1989.

Popper, K. R., *The Logic of Scientific Discovery*, New York: Harper & Row, 1968.

Rolston, D. W., *Principles of Artificial Intelligence and Expert Systems Development*, New York: McGraw-Hill, 1988.

Salmon, W. C., *The Foundations of Scientific Inference*, Pittsburgh: Univ. of Pittsburgh Press, 1967.

Thagard, P., *Computational Philosophy of Science*, Cambridge: The MIT Press, 1988.

Vaughan, R. A., "Maintaining an Inductive Database," in: J. H. Fetzer (ed.), *Aspects of Artificial Intelligence*, Dordrecht: Kluwer, 1988, pp. 323-35.