# 재생 스코어링 함수의 구간확장

(An Extension Of Reproducing-Scoring-Function For Intervals Of Probabilities)

정 민 교, 김 진 형

한국 과학 기술원 전산학과

요약

이성적이고 건전한 사고 방식을 갖고있는 사람으로 하여금 자기의 의견을 진솔하게 나타내도록 하는 한 방법으로 재생 스코어링 함수(Reproducing Scoring Function)가 있다. 이때 평가를 받는 사람은 자기의 믿음의 정도를 이산확률값(Discrete Point Estimate)으로 나타낸다 반면, 믿음의 정도를 확률구간(Intervals Of Probabilities)으로 표시하면 이산확률값으로 나타내는 것보다 훨씬 자연스럽고 융통성이 있다 본 논문에서는 재생 스코어링 함수가 사람의 정직성을 유도하면서 확률구간에도 적용될 수 있도록, 확률구간에 임의의 확률밀도함수가 주어진다고 가정하고 새로운 상금보상공식(New Reward Sum)을 제안했다 또한, 상금보상체계에서 만족되어야 할 두 가지 기본성질이 소개되고, "음고양저 확률분포"와 "음저양고 확률분포"가 이 두 성질을 모두 만족함을 보여준다.

## I 서론

흔히 인생은 불확실한 사건들의 연속이라고 한다. 이때마다 우리는 미래의 상황과 조건을 나름대로 판단하여 자기 주관적인 결정을 내리게 된다. 미래의 사건은 불확실하며, 그 불확실 정도를 어떤 정량화된 수로 표시해야 한다. 이런 불확실 정도를 나타내는 데는 확률이 가장 적합한 표현방법이 된다 [6]. 따라서, 우리는 각 개인의 의사결정을 주관적 확률(Subjective Judgmemtal Probability) 관점에서 보고, 이를 의사결정 대안들(Decision Alternatives)의 확률을 정하는데 사용한다.

Bruno de Finetti (1937) 와 Savage (1954)에 의해 개발된 주관확률론 (Subjective Probability Theory)은, 의사결정 환경에서 전문가(또는 예보관, 측정자, 평가자)가 예측한 확률측정치(Personal Probability Assessments)와 전문가가 실제로 생각하고 있는 판단(Judgments)은 일치해야 한다는 이론적 배경이된다 그러나 전문가의 판단은 마음속에 있기 때문에 전문가의 마음속에 들어가지 않고는 정확히 감지해낼 수 가 없다 [2, 8]. 이런 문제를 해결하기 위해 여러가지 종류의 방법이 개발되었고 그 중 하나가 재생 스코어링

함수(Reproducing Scoring Function)이다. 스코어링 함수의 스코어는 측정확률(Assessed Probability)과 판단확률(Judgmental Probability)의 근접도에 따라 정해지기 때문에, 이 함수는 전문가의 확률측정치를 내면판단과 일치시켜 주는 기능을 한다. 또한, 스코어링 함수는 전문가의 능력을 평가하는 데에도 쓰일 수 있다 [5]. 스코어링 함수의 또다른 적용 분야는 [7]에 잘 언급되어 있다.

기상예보관은 일기예보를 확률을 사용해 표현한다. 그러나, 이 예보들이 얼마나 정확한지 측정할 방법이 없었다. 이런 저런 이유로해서 기상학에서 제일 먼저 예보관들을 평가 하는 방법들이 개발되기 시작했다 [1]. 그래서 스코어링 함수인 "Probability Score"(PS)와 "Ranked Probability Score"(RPS)는 일기예보와 관계가 있다 [3, 4]

그런데 스코어링 함수는 이산확률값(Discrete Point Estimate)에는 잘 적용되지만, 확률구간(Intervals Of Probabilities)에는 적용되지 않는다. 반면 믿음의 정도를 확률구간으로 표시하면 이산확률값으로 나타내는 것보다 훨씬 자연스럽고 융통성이 있다. 본 논문에서는 새로운 상금보상공식(New Reward Sum)을 제안하여, 이 보상공식이 확률구간에도 적용됨을 보여준다.

## II 재생 스코어링 함수(Reproducing Scoring Function)

우리는 전문가가 사실을 곡해하지 말고, 그의 진실된 생각을 그대로 전해주었으면 하고 바란다. 그러나, 여러가지 이유로 해서 전문가는 사실을 숨기고 왜곡하여 말한다. 이런 문제를 해결하여, 전문가의 정직한 의견을 이끌어내는 한 방법이 재생 스코어링 함수이다. 그러면 재생 스코어링 함수가 어떻게 이런 문제를 해결하는데 사용되는지 알아보자.

상호배타적(Mutually Exclusive)이면서, 포괄적(Exhaustive)인 사건들의 집합 $\{E_1, E_2, E_3, \cdots\cdots E_n\}$ 을 가정하자. 이 집합에서는 꼭 한 사건만 다음 실험(Trial)에서 일어난다. 사건 $E_i$ 가 진실로 일어날 확률을 $t_i$ , 전문가가 이 사건이 일어난다고 예측한 확률을 $r_i$ , 이 사건이 실제로 일어났을때 전문가가 받는 점수(상금)를 나타내는 함수를 $R_i$ 라하자 그러면, 전문가가 실제로 믿는 확률을 $t_1, t_2, t_3, \cdots, t_n$ , 예측한 확률을 $r_1, r_2, r_3, \cdots\cdots, r_n$ 이라 했을때, 전문가가 받는 상금기대치 S는 다음과 같

다.

$$S = \sum_{i=1}^{n} t_i . R_i(r_1, r_2, \ldots, r_n) \quad \ldots\ldots\ldots\ldots (a)$$

이때 상금함수 $R_i$ 는 $r_1 = t_1$, $r_2 = t_2$, $r_3 = t_3$, ....., $r_n = t_n$ 일때, 상금기대치 S가 최대값을 갖게하는 함수이어야 한다 즉, $R_i$ 는 전문가가 외부로 표현한 사실이 자기자신의 진실된 생각일때 S가 최대가 되게 해주는 함수이어야 한다 상금함수 $R_i$ 는 전문가가 마음속에 갖고 있는 진실된 생각을 그대로 끄집어 재생시키기 때문에 재생 스코어링 함수(Reproducing Scoring Function)라고 부른다. 그러나, (a)식에 관련규칙(Relevance Principle)과 동등규칙(Invariance or Symmetric Principle)을 적용하면 표준화된 식(b)를 얻을 수 있고, 본 논문 에서는 상금기대치공식으로 (b)식을 사용해 이론을 전개한다

$$S = \sum_{i=1}^{n} t_i . R(r_i) \quad \ldots\ldots\ldots\ldots (b)$$

라그랑지 배수(Lagrangian Multipliers) 방법을 사용하면, 어떤 함수가 재생 스코어링 함수가 되기위한 조건은 다음과 같다.

(1) R은 절대오목(Strictly Concave), 단조증가(Monotone Increasing), 이차연속(Twice Continuous)이다.

(2) $0 < x < 1$ 일때, $x.R'(x) = c$ . 여기서 c는 상수이다.

그리고, 위 성질을 모두 만족하는 전형적인 재생 스코어링 함수의 예로는 이차재생함수(Quadratic Scoring Function) 인 $R(p) = 1 - (1 - p)(1 - p)$ 와 로그재생함수(Logarithmic Scoring Function) 인 $R(p) = a \log p + b$ 가 있다. III 장 에서는 위에서 설명한 상금기대치공식 S 와 재생 스코어링 함수에 기초해서, 새로운 상금보상공식을 제안한다.

## III. 새로운 상금보상공식(New Reward Sum)

여기에서 소개하는 새로운 상금보상공식은, 사건이 상호 배타적(Mutually Exclusive)이며 발생 가능한 사건의 수가 둘인 경우에만 확률구간에 적용된다. 두 사건 $E_1$, $E_2$을 가정하자. 첫번째 사건 $E_1$ 의 확률구간은 [a,b], 두번째 사건 $E_2$ 의 확률구간은 [c,d]로 하자. 구간 [a,b]상의 임의의 점을 $r_1$ 이라 하면, $r_2$ 는 $r_1+r_2=1$ 을 만족하는 구간 [c,d]상의 임의의 점이 된다. 따라서 구간 끝점은 a+d=1, b+c=1 와 같은 관계가 성립한다. 첫번째 사건의 확률 밀도함수 $f_1(x)$ 가 주어진다고 가정하면, 새로운 상금보상공식은 다음과 같이 정의된다.

$$NS=\int_a^b f_1(x)\,S(x)\,dx$$

where

$$S(x)=t_1\,R(x)+t_2\,R(1-x) \quad (\,t_1=1-t_2\,).$$

여기서, NS는 새로운 상금보상공식, R(x)는 재생 스코어링 함수, $t_i$ 는 i 번째 사건이 진실로 일어날 확률을 각기 나타낸다.

첫번째, 두번째 사건의 확률밀도함수를 각각 $f_1(x)$ , $f_2(y)$ 라하면 위식은 다시 다음과 같이 표현된다

$$NS=t_1\int_a^b f_1(x)\,R(x)\,dx+t_2\int_c^d f_2(y)\,R(y)\,dy$$

이 새로운 상금보상공식은 기존의 스코어링 함수에 의해 표현된 상금기대치 공식 S이 만족하는 성질을 그대로 만족하면서, 스코어링 함수를 확률구간에서도 적용될 수 있도록 해준다. 새로운 상금보상공식은 아래의 두 가지 기본성질 을 만족한다.

(가). 두 전문가의 확률밀도함수는 같지만 확률밀도함수가 놓이는 확률구간이 다를 경우, 진실점(True Belief Point)에 가깝게 확률구간을 놓은 전문가에게 보다 많은 점수(상금)가 주어진다. 증명은 [9]을 참조하시오.

(나) 두 전문가의 확률밀도함수 형(Type)은 같고, 두 확률구간의 중심점도 일치한다 그러나 한 확률구간이 다른 구간을 포함할때 넓은 확률구간을 배정한 전문가는 좁은 확률구간을 배정한 전문가보다 점수(상금)가 적게 주어진다

이 두 성질을 적절히 결합해서 사용하면 전문가의 진실한 생각(또는, 정직)을 유도해 낼 수 있다. 왜냐하면, 성질 (가)때문에 전문가는 진실점에 가깝도록 확률구간을 결정하려고 할 것이고, 성질 (나)때문에 확률구간의 폭을 점점 줄이려고 할 것이기 때문이다. 성질 (나)를 DISC(Different Intervals With Same Center-Point)성질이라고 부르겠다. 그러나, 이 성질은 모든 확률밀도함수에 대해 성립하는 것은 아니다. 다음 장에서는 "음고양저 확률분포"와 "음저양고 확률분포"가 정의되고, 이들 확률밀도함수들이 DISC성질을 만족한다는 것을 보여주겠다.

## IV. DISC 성질을 만족하는 확률밀도함수

이 장에서는 DISC성질을 만족하는 확률밀도함수를 소개한다. 여기에는 평균치정리(Mean Value Theorem)와 다음에 기술하는 두 기본 정리가 사용된다. 이 정리들의 증명은 [9]를 참조하시오.

[정리 1] 상금기대치 공식 S 의 도함수 S'(x)는 스코어링 함수에 관계없이 구간 [0,1]에서 단조감소(Monotone Decreasing) 한다.

[정리 2] $S(x+c) - S(x\alpha+c) > S(-x\alpha+c) - S(-x+c)$ (x 는 음수) ; 여기에서 'c'는 x축의 양의 방향으로의 이동계수를, '$\alpha$'는 x축 방향으로의 스케일 인자(Scale Factor)를 각기 나타낸다. 단, $c > 0$, $\alpha > 1$ 이다.

먼저 이곳에서 쓰이는 새로운 확률밀도함수를 정의하자. 구간 [-a,a]에서 정의된 확률밀도함수 'f'가 구간 [-a,0]에 속하는 모든 'y'에 대해 $f(y) >= f(-y)$ 가 성립하면, 'f'를 "음고양저 확률분포" (Negative-Higher-Than-Positive Distribution)라 하고, $f(y) <= f(-y)$ 가 성립하면, 'f'를 "음저양고 확률분포"(Negative-Lower-Than-Positive Distribution)라 부르자. 그러면, 확률구간의 중심점이 진실점(True Belief Point)의 왼쪽에 위치할때는 "음고양저 확률분포"가, 오른쪽에 위치할때는 "음저양고 확률분포"가 DISC성질을 만족하는 확률밀도함수가 된다. 중심점과 진실점이 일치하면 어떤 확률밀도함수라도 DISC성질을 만족한나. 이들 증명은 [9]를 참조하시오.

## V. 결론

전문가의 측정확률이 이산값(Point Estimate)으로 주어지면, 상금기대치 공식 S(x) 만으로도 충분히 전문가의 정직성을 유도할 수 있다. 그러나, 측정확률이 확률구간으로 주어지면 S(x)는 이제 더 이상 적용될 수 없다. 본 논문에서는 새로운 상금보상 공식을 만들어 이 문제를 해결했다. 그리고 상금보상체계에서 만족되어야할 기본성질(즉, DISC성질)이 소개되고, "음고양저 분포"와 "음저양고 분포"가 DISC 성질을 만족함을 보여 주었다.

## 참고문헌


[1] Winkler, R.L., and Murphy, A.H., " 'Good' probability assessors", Journal of Applied Meteorology 7 (October 1968), pp 751-758.

[2] Lindley, D.V., Tversky, A., and Brown,R.V., "On the reconciliation of probability assesments", J.R.Statist.Soc.A (1979), 142, Part 2, pp 146-180.

[3] Murphy, A.H., "On the 'Ranked Probability Score' ", Journal of Applied Meteorology 8 (1969), pp 988-989.

[4] Murphy, A.H., "On expected-utility measures in cost-loss ratio decision situations", Journal of Applied Meteorology 8 (1969), pp 989-991.

[5] Winkler, R.L., "Scoring rules and the evaluation of probability assessors", Journal of the American Statistical Association 64 (Sep. 1969), pp 1073-1078.

[6] Lindley, D.V., "Scoring rules and the inevitability of probability", International Statistics Review (1982), Vol 50, pp 1-26.

[7] Pearl, J., "A note on the management of probability assessors", IEEE Transactions on Systems,Man,and Cybernetics(May 1977), pp 402-403.

[8] Savage, L.J., "Elicitation of personal probabilities and expectations", in *Studies in Bayesian Econometrics and Statistics, New York,John Wiley and Sons,1974*, pp 111-156.

[9] Chung, M.G., "An Extension Of Reproducing-Scoring-Function For Intervals Of Probabilities", (forthcoming master thesis), 1987.