조건부 확률

劉熙世(고려대 이과대 수학과 교수, 이학박사)

월간 수학세계 1980년 5월호

편집자 주: 이 글은 지금은 은퇴하신 고려대 수학과 유희세(1919–) 교수님께서 월간지 수학세계에 실으셨던 글이다. 이 잡지에서는 매 년 각 학교의 본고사 문제를 분석하는 좌담회를 개최하였다. 1980년 4월호에서는 고려대 본고사 수학 문제를 다루었고, 여기서 조건부 확률을 묻는 문제는 "학생들 중에는 조건부 확률이라 하면 아예 손도 안 대기로 작정이라도 한듯 전연 답안지 난을 비워 둔 학생들이 꽤 있었어요."라는 평이 있을 정도로 성적이 좋지 않았다.

이에 학생들에게 조건부 확률에 대해 설명하기 위해 유희세 교수님께서 다음 호에 글을 실으셨는데, 조건부 확률뿐 아니라 확률 개념 자체를 이해하는 데도 도움이 되는 좋은 글이어서 보여주는 사람마다 한 부씩 복사해 갈 정도였다. 그러나 오래된 책이어서 구하기도 어렵고 편집 상태도 썩 좋지 않아서, 널리 알리고 싶은 마음에 TEX으로 조판하여 PDF 파일을 만들었다. 되도록 원문 그대로 만들었으나, 원문이 국판 2단 편집으로 조판이 빽빽한 편이어서 이 파일에서는 2단 편집은 하지 않았다.

유희세 교수님께 허락을 받지도 못했고 수학세계를 발행했던 성지사의 허락을 받은 글도 아니지만 수학을 공부하는 학생들에게, 또 교사들에게도 도움이 될 수 있도록 널리 양해해 주시기를 부탁 드린다.

조건부 확률에 대한 개념 인식이 크게 부족한 것 같아, 이 기회에 그에 대한 이야기를 하고자 한다. 그러기 위하여 먼저 금년도 고려대 입시문제로 출제되었던 조건부 확률 문제를 가지고 설명을 한 뒤 보충 설명을 하겠지만, 시간이 바쁜 학생, 또는 조건부 확률의 개념에 대하여 자신 있게 배운 독자는 뒷부분은 읽을 필요가 없을 것이다.

문제(지난 4월호 p.15 참조)는 2개의 소문항으로 되어 있는데, 첫 문항을 생각해 보자.[1]

---

[1](편집자) 문제는 다음과 같다.

문제 5. 세 번 중에 한 번의 비율로 모자를 잊어버리는 버릇이 있는 사람이 차례로 A, B, C의 세 집을 방문하고 돌아왔을 때 (15점)

(1) 모자를 잊어버리지 않고 돌아왔을 확률을 구하여라.

(2) 모자를 잊어버리고 돌아왔을 때, 그 모자를 C의 집에서 잊어버렸을 확률을 구하여라.

A, B, C의 세집을 이 순서대로 차례로 방문하고 돌아왔다. 지금 첫째 집 A에서 모자를 잊어버린다는 사건을 A_0, B에서 잊어버린다는 사건을 B_0, C에서 잊어버린다는 사건을 C_0 라고 하자. 그리고 모자를 잊어버리지 않고 집으로 돌아오는 사건을 D_0 라고 하자.

$$
\begin{aligned}
P(A_0) &= \frac{1}{3} \\
P(B_0) &= \frac{2}{3} \times \frac{1}{3} = \frac{2}{9} \\
P(C_0) &= \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3} = \frac{4}{27} \\
P(D_0) &= \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{8}{27}
\end{aligned}
$$

첫 문항은 D_0 의 확률을 요구한 것이니까 위의 마지막 식만 써 놓으면 된다.

위의 네 사건 A_0, B_0, C_0, D_0 는 서로 배반이고, 또 넷 중의 어느 하나는 꼭 일어나므로 D_0 의 확률은 다음과 같이 구할 수도 있다.

$$
\begin{aligned}
P(D_0) &= 1 - P(A_0) \cup B_0 \cup C_0 \\
&= 1 - P(A_0) + P(B_0) + P(C_0) \\
&= 1 - \left(\frac{1}{3} + \frac{2}{9} + \frac{4}{27}\right) \\
&= \frac{8}{27}
\end{aligned}
$$

둘째 문항이 조건부 확률의 문제였다. 분명히 모자를 잊어버리고 돌아왔다면 어느 집에서 잊어버린 것이겠느냐, 즉 잊어버리고 돌아왔다는 조건이 붙어 있다.

조건부 확률에 관한 정리

$$P(A \cap B) = P(A) \cdot P(B|A)$$

를 쓰면, 첫째 집에서 잊어버리고 왔을 조건부 확률 $P(A_0|D_0^C)$ 는

$$P(D_0^C \cap A_0) = P(D_0^C) \cdot P(A_0|D_0^C)$$

그런데, 사건 $D_0^C \cap A_0$ 는 사건 A_0 와 같으니까(그 이유는 모르는 독자를 위하여 나중에 설명하겠음), 윗식은

$$P(A_0) = P(D_0^C) \cdot P(A_0|D_0^C)$$

$$
\begin{aligned}
\therefore \mathrm{P}(\mathrm{A}_0|\mathrm{D}_0^{\mathrm{C}}) &= \mathrm{P}(\mathrm{A}_0)/\mathrm{P}(\mathrm{D}_0^{\mathrm{C}}) \\
&= \mathrm{P}(\mathrm{A}_0)/\{1-\mathrm{P}(\mathrm{D}_0)\} \\
&= \frac{1}{3}\Big/\left(1-\frac{8}{27}\right) \\
&= \frac{9}{19}
\end{aligned}
$$

마찬가지로 생각하여 둘째 집에서 잊어버리고 왔을 조건부 확률은

$$\mathrm{P}(\mathrm{B}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{2}{9}\Big/\left(1-\frac{8}{27}\right) = \frac{6}{19}$$

세째 집에서 잊어버리고 왔을 확률은

$$\mathrm{P}(\mathrm{C}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{4}{27}\Big/\left(1-\frac{8}{27}\right) = \frac{4}{19}$$

가 됨을 알 수 있다.

그러면, 앞에서 보류해 두었던 설명을 하자. 즉, $\mathrm{D}_0^{\mathrm{C}} \cap \mathrm{A}_0 = \mathrm{A}_0$ 의 설명이다.

D_0 가 모자를 잊어버리지 않고 집으로 돌아온다는 사건이었으니까, $\mathrm{D}_0^{\mathrm{C}}$ 는 결국 모자를 잊어버리고 집으로 돌아온다는 사건이다.

그런데, 모자를 잊어버린다는 사건은 $\mathrm{A}_0, \mathrm{B}_0, \mathrm{C}_0$ 의 셋뿐이고, 이 셋은 서로 배반이니까

$$\mathrm{D}_0^{\mathrm{C}} = \mathrm{A}_0 \cup \mathrm{B}_0 \cup \mathrm{C}_0$$

따라서

$$\mathrm{A}_0 \subset \mathrm{D}_0^{\mathrm{C}}, \quad \text{즉 } \mathrm{D}_0^{\mathrm{C}} \cap \mathrm{A}_0 = \mathrm{A}_0$$

가 성립한다.

조건부 확률에 관한 공식, 즉 곱셈정리는 교과서마다 자세히 설명하고 있으니까 우리는 여기서 그 증명을 반복하지 않겠다. 다만 조건부 확률과 조건을 안 붙인 확률과를 비교해 보자.

$$
\begin{aligned}
&\mathrm{P}(\mathrm{A}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{9}{19}, \quad \mathrm{P}(\mathrm{A}_0) = \frac{1}{3} \\
&\mathrm{P}(\mathrm{B}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{6}{19}, \quad \mathrm{P}(\mathrm{B}_0) = \frac{2}{9} \\
&\mathrm{P}(\mathrm{C}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{4}{19}, \quad \mathrm{P}(\mathrm{C}_0) = \frac{4}{27} \\
&\mathrm{P}(\mathrm{A}_0|\mathrm{D}_0^{\mathrm{C}}) \neq \mathrm{P}(\mathrm{A}_0), \quad \mathrm{P}(\mathrm{B}_0|\mathrm{D}_0^{\mathrm{C}}) \neq \mathrm{P}(\mathrm{B}_0) \\
&\mathrm{P}(\mathrm{A}_0|\mathrm{D}_0^{\mathrm{C}}) \neq \mathrm{P}(\mathrm{A}_0)
\end{aligned}
$$

여기서 좌변이 우변보다 크지만 반드시 조건부 확률이 조건을 안 붙인 확률보다 커진다는 이유는 없다. 오히려 작게 되는 수도 물론 있다.

이를테면 $P(B_0|A_0)$를 생각하자.

$$P(A_0 \cap B_0) = P(A_0) \cdot P(B_0|A_0)$$

에서 A_0와 B_0는 배반사건이었으니까

$$P(A_0 \cap B_0) = 0$$

또 $P(A_0) = \frac{1}{3}$ 따라서

$$0 = \frac{1}{3} \cdot P(B_0|A_0) \quad \text{즉}, P(B_0|A_0) = 0$$

이것은 $P(B_0) = \frac{2}{9}$보다 더 작아졌다.

한편, 조건부 확률과 조건을 안 붙인 확률이 같게 될 때, 즉

$$P(B|A) = P(B)$$

일 때 사건 A와 B가 서로 독립이라고 한다.

사건 A와 사건 B가 서로 독립이라는 개념은 실로 희한한 개념이다. 사건의 관계에 관한 중요한 개념 중에 독립이란 개념이라든가 배반이라는 개념 같은 것이 있는데, 이 두 개념은 본질적으로 다른 개념이다. 즉, 배반인 사건이라 할 때, 집합의 개념으로 말한다면 서로 공통부분이 없는 집합이라는 뜻에 불과하지만, 독립인 사건이라 할 때에는 집합의 개념만으로는 도저히 옮겨 말할 수가 없고, 꼭 확률의 개념이 들어 있고서야 그 사실을 분간할 수가 있는 것이다. 즉, 사건의 독립의 개념이야 말로 확률 개념이 가지고 있는 고유의 특이한 개념인 것이다.

이 정도까지 읽고 이해한 독자는 처음에 약속한 대로 이 다음은 읽지 않아도 무방하다고 생각하는데, 나는 여기서 우리나라 고등학교에서의 확률과 통계의 교재의 개관을 하고, 프랑스의 현재의 고등학교에서의 확률과 통계의 교재에 대하여 다소의 비교를 해 보고자 한다. 아울러 프랑스의 수학교육에서 가르치는 확률의 입장에 서서 이야기해 보겠다.

확률과 통계의 개념과 방법은 현대 사회에 있어서 매우 중요한 위치를 차지하고 있고, 자연과학에 있어서 뿐만 아니라 사회과학에 있어서도 응용의 폭이 넓은 긴요한 수학의 한 분야가 되어 있다. 따라서 수학교육의 중요한 내용으로 등장하게 되는 것인데, 이번의 입시문제에서와 같은 기초적인 문제에 대해서조차 전연 손을 안 대는 학생이 많았다는 것은 확률이란 것이 가지고 있는 고유한 성질을 파악하지 못한 학생이 많기 때문인 것 같다.

확률이란, 불확실성의 정도를 수량화하려고 하는 의도에서 생긴 개념이라고 볼 수 있다. 이 확률의 개념을 수학적으로 정의하는 방법이 셋이 있는데, 그것은

① 수학적 확률(선험적 확률)

② 통계적 확률(경험적 확률)

③ 공리적 확률

이다. 현대수학의 한 분야로서의 확률의 이론 연구에 있어서는 ③에 의한 정의만이 완전한 정의라고 말할 수 있고, 사실 프랑스 고교에서의 확률의 개념은 ③에 의한 정의만을 사용하고 있다. 그런데, 우리나라 고교 수학교육에 있어서 는 ③에 의한 정의는 전혀 소개하지 않고 있으며, ①과 ②만을 사용하고 있다는 데에 큰 문제가 있다고 생각된다.

①에 의한 정의에서는, 생길 수 있는 모든 경우가 n개일 때, 그리고 어느 두 개도 중복되어 생기는 일이 없을 때, 그리고 어느 경우가 생기는 것도 같은 정도로 확실성이 있을 때, 이 n개의 경우 중에서 어떤 사건 A가 생길 경우의 수가 r이면 A가 생길 확률을 $\frac{r}{n}$이라고 정의하는 것이다.

그런데, 이 ①의 정의에서 문제가 되는 것은 '같은 정도로 확실성이 있다' 라는 말의 뜻이다. 그리고 ①에서 또 하나의 문제점은, 생길 수 있는 경우의 수가 유한개인 n개일 때에만 확률을 생각해야 한다는 점이다. 이와 같은 애로가 있기 때문에 ①은 확률의 개념의 엄밀한 수학적 정의가 될 수는 없다.

②에 의한 정의를 생각해 보자. n번의 시행에서 사건 A가 r번 생겼다고 하자. 그 상대빈도 $\frac{r}{n}$을 생각하여, 여기서 n을 충분히 크게 할 때 상대빈도가 일정한 값 p로 접근한다면 그 값 p를 사건 A가 생길 확률이라고 하자는 것이다. 그런데, 여기서는 '상대빈도(상대돗수)가 일정한 값에 접근한다' 는 말의 뜻이 문제가 된다. 그 뜻에 대하여는 Von Mises 씨에 의한 엄격한 설명이 있기는 하지만, 모든 순수수학자들을 납득시킬 수는 없었다. 결국 ③에 의한 정의가 현대 수학으로서 확률론을 발전시킬 수 있는 유일한 정의인 것이다. 이 정의가 나온 것이 1933년이다. 소련의 수학자 Kolmogorov의 업적이다.

우리 나라 고교생들이 확률의 개념에 대하여 선명한 파악을 못하고 있는 까닭의 하나는, 위에서 말한 바와 같이 엄밀하지 못한 정의 ①과 ②를 배우고 그 상호간의 관계가 선명하게 파악되지 못하고 있기 때문일지도 모른다.

위에서 말한 상호간의 관계에 관하여 고등학교의 교과서에는 '1큰 수의 법칙(大數의 법칙)' 이란 것이 있다. 어느 교과서에도 그 증명이 엄밀하게 나와 있지는 않을 뿐 아니라 아마 그 법칙이 있음으로써 확률의 개념이 보다 선명하게 되어야 할 터인데, 실제로는 더 모호하게 되어 버리는 것이 아닐까 하는 염려가 있다. 사실은 다음과 같다.

수학적 확률과 통계적 확률과의 관계를 생각해 볼 때, 통계적 확률을 생각하는 사건에 대하여 반드시 그 수학적 확률도 생각할 수 있는 것은 아니지만, 역으로 수학적 확률을 생각하는 사건에 대하여 그것을 통계적 확률의 입장에서 생각해 본다면 그 값이 일치한다고 하는 것은 증명할 수가 있다. 이 성질을 뜻하는 것이 곧 큰 수의 법칙이다.

어떤 독립시행에서 사건 A가 생길 확률(수학적 확률)이 p일 때, n번의 시행에 의한 A의 상대돗수 $f_n(\mathrm{A})$는 다음을 만족한다.

i) $\epsilon > 0$이 아무리 작더라도 $\lim_{n\to\infty} \mathrm{P}(|f_n(\mathrm{A}) - p| < \epsilon) = 1$

ii) $\mathrm{P}(\lim_{n\to\infty} f_n(\mathrm{A}) = p) = 1$

여기서 i)을 큰 수의 약법칙, ii)를 큰 수의 강법칙이라고 한다. i)의 경우 $f_n(\mathrm{A})$가 p에 확률수렴한다고 하고, ii)의 경우 $f_n(\mathrm{A})$는 p에 확률 1로써의 수렴(또는 거의 수렴한다) 한다고 말한다. 성질 i)은 이미 18세기 초에 발견되었지만, 성질 ii)는 20세기에 들어와서야 발견된 것이다. 고교 교과서에 나와 있는 것은 성질 i)이다.

그러면, 여기서 프랑스 고교 교과서에 의한 공리적 확률의 개념을 간단히 알아보기로 하자.

이 개념에서는 확률공간(또는 표본공간)의 개념부터 출발한다. 확률공간(S라고 하자)이란 하나의 집합인데, 그 집합의 부분집합은 '사건'이라고 한다. 사건 A에 대하여 다음 세 조건을 만족하는 실수값 P(A)를 대응시킬 때, P(A)를 A의 확률이라고 한다.

i) 임의의 사건 A_i에 대하여 $\mathrm{P}(\mathrm{A}_i) \geq 0$

ii) 전사건 S에 대하여 $\mathrm{P}(S) = 1$

iii) $\mathrm{A}_1, \mathrm{A}_2, \ldots$가 서로 배반사건일 때 $\mathrm{P}(\mathrm{A}_1 \cup \mathrm{A}_2 \cup \cdots) = \mathrm{P}(\mathrm{A}_1) + \mathrm{P}(\mathrm{A}_2) + \cdots$

확률공간의 개념은 확률문제를 풀 때 여러 관계를 엄밀하게 처리하는 데 긴요한 작용을 한다.

예를 들면, 앞에서 생각한 금년도 고려대 입시문제로 되돌아가 보자. 여기서

$$\{\mathrm{A}_0, \mathrm{B}_0, \mathrm{C}_0, \mathrm{D}_0\} = S$$

라고 놓으면 S가 이 문제에 있어서의 확률공간이다. 이 S는 확률 P와 더불어 위의 공리 i), ii), iii)을 전부 만족하고 있다. 조건부 확률의 경우는 확률공간

$$\{\mathrm{A}_0, \mathrm{B}_0, \mathrm{C}_0, \mathrm{D}_0\} = S$$

에 대하여 확률 $\mathrm{P}(\mathrm{X}|\mathrm{D}_0^{\mathrm{C}})$가 다음과 같이 정의 되었다고 생각할 수가 있다. 즉

$$\mathrm{P}(\mathrm{A}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{9}{19}, \quad \mathrm{P}(\mathrm{B}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{6}{19},$$
$$\mathrm{P}(\mathrm{C}_0|\mathrm{D}_0^{\mathrm{C}}) = \frac{4}{19}, \quad \mathrm{P}(\mathrm{D}_0|\mathrm{D}_0^{\mathrm{C}}) = 0$$

이로부터 공리 i), ii), iii)이 만족된다는 것을 증명할 수가 있다.

앞에서 조건부 확률의 이야기를 하기 전에 $\mathrm{A}_0, \mathrm{B}_0, \mathrm{C}_0, \mathrm{D}_0$와 같은 사건의 확률을 계산하였는데, 사실은 거기서 이미 조건부 확률의 개념을 사용한 것이 었다고 볼 수 있다. 즉 좀더 자세히 설명하여 보면 다음과 같다.

지금 집 A에서 나올 때 모자를 쓰고 나온다는 사건을 A, 잊어버리고 나온다 는 사건을 $\overline{\mathrm{A}}$라고 적기로 하고, 집 B, 집 C에서는 모자를 쓰고 나온다는 사건을 각각 B, C로, 모자를

잊어버리고 나온다는 사건을 각각 $\overline{\mathrm{B}}, \overline{\mathrm{C}}$로 적기로 하면

$$
\begin{aligned}
\mathrm{A}_0 &= \overline{\mathrm{A}} \\
\mathrm{B}_0 &= A \cap \overline{\mathrm{B}} \\
\mathrm{C}_0 &= A \cap B \cap \overline{\mathrm{C}} \\
\mathrm{D}_0 &= A \cap B \cap C
\end{aligned}
$$

와 같다. 그런데

$$
\begin{aligned}
\mathrm{P}(\mathrm{A}_0) &= \mathrm{P}(\overline{\mathrm{A}}) = \frac{1}{3} \\
\mathrm{P}(\mathrm{B}_0) &= \mathrm{P}(\mathrm{A} \cap \overline{\mathrm{B}}) \\
&= \mathrm{P}(\mathrm{A}) \cdot \mathrm{P}(\overline{\mathrm{B}}|\mathrm{A}) = \frac{2}{3} \times \frac{1}{3}
\end{aligned}
$$

여기서 $\mathrm{P}(\overline{\mathrm{B}}|\mathrm{A})$는 A가 일어났다는 조건 하에서 $\overline{\mathrm{B}}$가 일어난다는 조건부 확률로서 $\mathrm{P}(\overline{\mathrm{B}}|\mathrm{A}) = \frac{1}{3}$이다. A가 일어났다는 말은 집 A에서 모자를 쓰고 나왔다는 뜻이요, 즉 집 B에 모자를 쓰고 들어갔다는 말이니까 집 B에 모자를 쓰고 들어갔다가 나올 때 잊어버리고 나올 확률이 사실은 $\frac{1}{3}$이라고 가정되어 있기 때문이다.

다음도 마찬가지로 생각된다.

$$
\begin{aligned}
\mathrm{P}(\mathrm{C}_0) &= \mathrm{P}(\mathrm{A} \cap \mathrm{B} \cap \overline{\mathrm{C}}) \\
&= \mathrm{P}(\mathrm{A} \cap \mathrm{B}) \cdot \mathrm{P}(\overline{\mathrm{C}}|\mathrm{A} \cap \mathrm{B}) \\
&= \mathrm{P}(\mathrm{A}) \cdot \mathrm{P}(\mathrm{B}|\mathrm{A}) \cdot \mathrm{P}(\overline{\mathrm{C}}|\mathrm{A} \cap \mathrm{B}) \\
&= \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3}
\end{aligned}
$$

이상으로 조건부 확률에 관한 이야기를 끝맺기로 하거니와, 이 글로써 독자들의 조건부 확률에 관한 이해에 도움이 되었으면 하는 마음 간절하다.