풀이

STEP I 1998, 문제 12

확률과 통계 • 경우의 수와 독립

> **STEP I 1998, 12번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/98-s1-q12

핵심 아이디어

이 문제는 모두 세기(counting)이다. 해가 모두 몇 개인지, 그리고 그중 X가 주어진 값을 갖는 해가 몇 개인지를 알면, 각 확률은 두 개수의 비가 된다. 독립(independence)인지 확인하는 가장 빠른 방법은 극단적인 경우를 보는 것이다. 즉 결합확률(joint probability)이 주변확률(marginal probability)의 곱과 결코 같을 수 없는 값의 쌍 하나를 찾으면 된다. 마지막 부분에서 "XYZ가 5로 나누어떨어진다"는 것은 "X, Y, Z 중 적어도 하나가 5로 나누어떨어진다"는 뜻이므로, 포함–배제의 원리(inclusion–exclusion)를 쓰면 된다.

풀이

해의 개수 세기

$X = x$를 고정하면($0 \le x \le 20$), Y는 $0, 1, \ldots, 20 - x$ 중 어느 값이든 될 수 있고, 그러면 $Z = 20 - x - Y$는 결정되며 음이 아니다. 따라서 $X = x$인 해는 정확히 $21 - x$개이고, 해의 총 개수는

$$N = \sum_{x=0}^{20} (21 - x) = 21 + 20 + \cdots + 1 = \frac{21 \cdot 22}{2} = 231.$$

이다. (이는 "별과 막대(stars and bars)" 방법으로 센 $\binom{22}{2}$와 같다. 별 20개와 막대 2개를 한 줄로 배열하는 것이다.) 각 해의 확률은 $1/231$이고,

$$\mathrm{P}(X = x) = \frac{21 - x}{231}, \qquad x = 0, 1, \ldots, 20.$$

이다. 대칭성에 의해 Y(그리고 Z)도 같은 분포를 따른다.

X와 Y는 독립인가?

아니다. $X = 20$, $Y = 20$을 생각하자. 각 사건은 따로 보면 일어날 수 있다.

$$\mathrm{P}(X = 20) = \mathrm{P}(Y = 20) = \frac{1}{231} > 0,$$

그러나 $X + Y \leq 20$이므로 두 사건은 동시에 일어날 수 없다. 따라서

$$\mathrm{P}(X = 20,\ Y = 20) = 0 \neq \frac{1}{231^2} = \mathrm{P}(X = 20)\,\mathrm{P}(Y = 20),$$

이므로 X와 Y는 독립이 아니다. 직관적으로 말하면, X가 크면 Y가 가질 수 있는 여지가 줄어든다.

> **생각해 보기** "$X + Y \leq 20$이므로 당연히 종속이다"라고 말하는 것으로 충분하지 않을까?

그런 제약 조건에 주목하는 것은 올바른 직감이지만, 독립은 확률에 관한 정확한 명제이므로, 근거를 대려면 $\mathrm{P}(X = x, Y = y) \neq \mathrm{P}(X = x)\mathrm{P}(Y = y)$인 값의 쌍 하나를 제시해야 한다. 반례(counterexample) 하나면 충분하며, 극단적인 쌍 $(20, 20)$을 택하면 계산이 아주 간단해진다.

X가 5로 나누어떨어지는 경우

범위 안에 있는 5의 배수는 $X = 0, 5, 10, 15, 20$이고, 각각에 해당하는 해의 개수는 $21, 16, 11, 6, 1$이다. 따라서

$$\mathrm{P}(5 \mid X) = \frac{21 + 16 + 11 + 6 + 1}{231} = \frac{55}{231} = \frac{5}{21},$$

이고, 이것이 구하려던 결과이다.

XYZ가 5로 나누어떨어지는 경우

5는 소수(prime)이므로, $5 \mid XYZ$일 필요충분조건은 5가 X, Y, Z 중 적어도 하나를 나누는 것이다. (0도 5로 나누어떨어진다는 것을 잊지 말자.) A, B, C를 각각 X, Y, Z가 5로 나누어떨어지는 해들의 집합이라 하자. 앞 부분에서 $|A| = |B| = |C| = 55$이다.

$A \cap B$에 대해서는 $a, b \geq 0$, $a + b \leq 4$인 a, b에 대하여 $X = 5a$, $Y = 5b$로 쓴다. 이러한 쌍의 개수는 $5 + 4 + 3 + 2 + 1 = 15$이다. 이들 모두에서 $Z = 20 - 5a - 5b$ 역시 5의 배수이므로

$$|A \cap B| = |B \cap C| = |C \cap A| = |A \cap B \cap C| = 15.$$

이다. 포함–배제의 원리에 의해

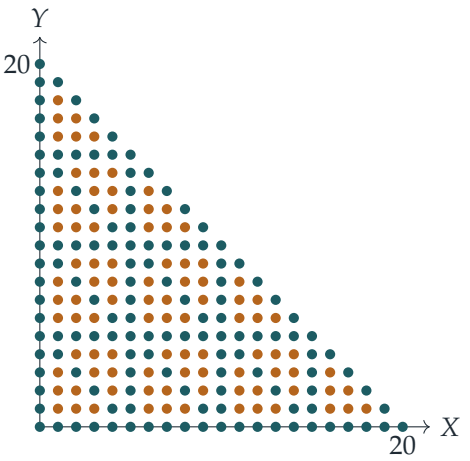
$$\begin{aligned} |A \cup B \cup C| &= 3 \cdot 55 - 3 \cdot 15 + 15 \\ &= 135, \end{aligned}$$

이고, 따라서

$$\mathrm{P}(5 \mid XYZ) = \frac{135}{231} = \frac{45}{77}.$$

이다.

아래 그림은 해들을 $X + Y \leq 20$인 격자점(lattice point) (X, Y)로 나타낸 것이다. 청록색 점은 직선 $X = 5k$, $Y = 5k$, $X + Y = 20 - 5k$(즉 $Z = 5k$) 중 하나 위에 있는 점이고, 호박색 점은 X, Y, Z 중 어느 것도 5의 배수가 아닌 96개의 해이다.



검산

- 여사건에는 X, Y, Z 중 어느 것도 5의 배수가 아닌 해가 $231 - 135 = 96$개 있어야 한다. 그림의 호박색 점들이 바로 그것이다. 컴퓨터로 231개의 해를 모두 나열하고 정확한 분수로 계산하면 96, $|A \cap B| = 15$, $5/21$, $135/231 = 45/77 \approx 0.584$가 확인된다.
- $5/21$에 대한 타당성 점검: X가 5로 나누어떨어질 확률이 $1/5$보다 조금 큰 것은, 작은 X 값일수록 확률이 높고 그중에서도 $X = 0$의 확률이 가장 높기 때문이다.

흔한 실수

- 0이 5로 나누어떨어진다는 것을 잊는 실수. $X = 0$을 빠뜨리면 $5/21$이 아니라 $34/231$이 나온다.
- $3 \times \frac{5}{21}$을 더하는 실수. 이렇게 하면 겹치는 부분을 여러 번 세게 된다. 여기서는 겹치는 부분이 크다. 두 변수가 5의 배수이면 나머지 하나도 반드시 5의 배수이기 때문이다.
- 곱셈 법칙이 성립하지 않는 구체적인 값의 쌍을 제시하지 않고 종속이라고 주장하는 실수.

STEP 너머

해를 균등하게 무작위로 고르는 것은 20을 세 부분으로 나누는 **합성**(*composition*)을 무작위로 고르는 것과 같다. 20을 합계 n으로 바꾸고 n을 키우면, $(X/n, Y/n)$은 삼각형 위의 균등분포(uniform distribution)를 따르게 되는데, 이는 디리클레 분포(Dirichlet distribution)의 가장 간단한 경우이다. 포함–배제의 원리는 "적어도 하나" 사건을 다루는 표준 도구로, 교란순열(derangement)과 오일러(Euler)의 토티언트 함수(totient function)에서도 다시 등장한다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문서)은 세기와 독립성에 관한 아이디어를 모두 다룬다.

Beyond STEP 풀이: https://beyondstep.keithmcnulty.org/ko/. 풀이 글은 CC0 1.0이다. 문제 자체의 저작권은 © Cambridge University Press & Assessment에 있으며 여기에 싣지 않는다. 영어 원문을 기계 번역하였으며 아직 검토되지 않았다.