풀이

STEP II 2018, 문제 13

확률과 통계 • 점화식

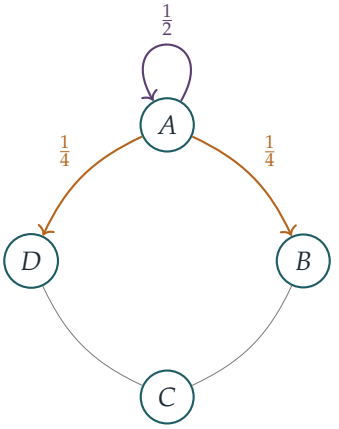
> **STEP II 2018, 13번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/18-s2-q13

핵심 아이디어

소포가 어디에 있든 다음 호루라기 소리에 소포는 정해진 규칙에 따라 움직인다. 따라서 A_{n+1}, B_{n+1}, C_{n+1}, D_{n+1} 각각은 A_n, B_n, C_n, D_n의 고정된 일차결합이다. 서로 얽힌 점화식(recurrence) 네 개는 만만치 않아 보이지만, 알맞은 **결합**(마주 보는 두 아이의 확률의 합과 차)을 잡으면 각각이 한 줄짜리 점화식을 만족하고, 이런 점화식은 곧바로 풀 수 있다.

풀이

원 위에 앉은 아이들에게 이름을 붙이자. 호루라기가 울릴 때마다 소포를 든 아이는 확률(probability) $\frac{1}{2}$로 소포를 그대로 가지고 있고, 양옆의 이웃에게 각각 확률 $\frac{1}{4}$로 소포를 넘긴다. 문제의 설정에 따라 시계 방향으로 넘길 확률과 반시계 방향으로 넘길 확률이 같다는 점에 주의한다.



그림은 A에서 일어나는 이동을 나타낸다. 모든 아이에게서 같은 규칙이 적용된다.

소문항 (i)

A에서 시작하면, 호루라기가 한 번 울린 뒤

$$A_1 = \tfrac{1}{2}, \qquad B_1 = \tfrac{1}{4}, \qquad C_1 = 0, \qquad D_1 = \tfrac{1}{4},$$

이다. C는 A에서 두 칸 떨어져 있으므로 한 번의 이동으로는 도달할 수 없기 때문이다.

두 번째 호루라기 뒤에 소포가 어떤 아이에게 있으려면, 소포가 이미 그 아이에게 있었고 그대로 머물렀거나(확률 $\frac{1}{2}$), 이웃에게 있다가 올바른 방향으로 넘어왔어야 한다(각 이웃에서 확률 $\frac{1}{4}$). 따라서

$$\begin{aligned} A_2 &= \tfrac{1}{2}A_1 + \tfrac{1}{4}B_1 + \tfrac{1}{4}D_1 = \tfrac{1}{4} + \tfrac{1}{16} + \tfrac{1}{16} = \tfrac{3}{8}, \\ B_2 &= \tfrac{1}{2}B_1 + \tfrac{1}{4}A_1 + \tfrac{1}{4}C_1 = \tfrac{1}{8} + \tfrac{1}{8} + 0 = \tfrac{1}{4}, \\ C_2 &= \tfrac{1}{2}C_1 + \tfrac{1}{4}B_1 + \tfrac{1}{4}D_1 = 0 + \tfrac{1}{16} + \tfrac{1}{16} = \tfrac{1}{8}, \\ D_2 &= \tfrac{1}{2}D_1 + \tfrac{1}{4}C_1 + \tfrac{1}{4}A_1 = \tfrac{1}{8} + 0 + \tfrac{1}{8} = \tfrac{1}{4}. \end{aligned}$$

검산하면 $\frac{3}{8} + \frac{1}{4} + \frac{1}{8} + \frac{1}{4} = 1$이다.

소문항 (ii)

$n = 2$일 때 쓴 논증은 전확률 법칙(law of total probability)에 의해 모든 $n \geq 0$에 대해 성립한다(n번째 호루라기 뒤에 누가 소포를 들고 있는지를 조건으로 삼는다).

$$\begin{aligned} A_{n+1} &= \tfrac{1}{2}A_n + \tfrac{1}{4}(B_n + D_n), & B_{n+1} &= \tfrac{1}{2}B_n + \tfrac{1}{4}(A_n + C_n), \\ C_{n+1} &= \tfrac{1}{2}C_n + \tfrac{1}{4}(B_n + D_n), & D_{n+1} &= \tfrac{1}{2}D_n + \tfrac{1}{4}(A_n + C_n), \end{aligned}$$

이고, 초기 조건은 $A_0 = 1$, $B_0 = C_0 = D_0 = 0$이다.

B_n과 D_n 구하기. B_{n+1}과 D_{n+1}에 대한 식을 더하면

$$B_{n+1} + D_{n+1} = \tfrac{1}{2}(B_n + D_n) + \tfrac{1}{2}(A_n + C_n) = \tfrac{1}{2}(A_n + B_n + C_n + D_n) = \tfrac{1}{2},$$

이다. 언제나 누군가는 소포를 들고 있으므로 네 확률의 합은 1이기 때문이다.

? **생각해 보기** 이제 합 $B + D$는 알지만, 이것을 B와 D에 어떻게 나누어야 할까?

더하는 대신 빼 보자. $A_n + C_n$ 항이 소거되어

$$B_{n+1} - D_{n+1} = \tfrac{1}{2}(B_n - D_n).$$

을 얻는다. $B_0 - D_0 = 0$이므로 모든 n에 대해 $B_n - D_n = 0$이다. (이것은 이 놀이의 좌우 대칭성(symmetry)이다. 시계 방향과 반시계 방향으로 넘길 확률이 같고, B와 D는 A의 양쪽에 대칭으로 앉아 있다.) 이를 $B_n + D_n = \frac{1}{2}$와 함께 쓰면

$$B_n = D_n = \tfrac{1}{4} \qquad \text{모든} n \geq 1,$$

에 대해 성립하고, 물론 $B_0 = D_0 = 0$이다. 이 결과는 소문항 (i)과 일치한다.

A_n과 C_n 구하기. 이제 $n \geq 1$에 대해 $A_n + C_n = 1 - (B_n + D_n) = \frac{1}{2}$임을 안다. A_{n+1}과 C_{n+1}에 대한 식을 빼면 $B_n + D_n$ 항이 소거되어

$$A_{n+1} - C_{n+1} = \tfrac{1}{2}(A_n - C_n),$$

이고, 이 식은 모든 $n \geq 0$에 대해 성립한다. $A_0 - C_0 = 1$이므로 모든 $n \geq 0$에 대해 $A_n - C_n = \left(\frac{1}{2}\right)^n$이다. 이를 $A_n + C_n = \frac{1}{2}$와 연립하여 풀면

$$\boxed{A_n = \frac{1}{4} + \frac{1}{2^{n+1}}, \qquad C_n = \frac{1}{4} - \frac{1}{2^{n+1}} \qquad (n \geq 1).}$$

이 공식은 $n = 0$에서는 성립하지 않는다($A_0 = \frac{3}{4}$가 되어 버린다). 관계식 $A_n + C_n = \frac{1}{2}$는 첫 번째 호루라기 뒤부터 성립하기 때문이다.

검산

- $n = 1$: $A_1 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, $C_1 = 0$이다. $n = 2$: $A_2 = \frac{3}{8}$, $C_2 = \frac{1}{8}$이다. 둘 다 소문항 (i)과 일치한다.
- $n = 3$: 공식에 따르면 $(A_3, B_3, C_3, D_3) = (\frac{5}{16}, \frac{1}{4}, \frac{3}{16}, \frac{1}{4})$이고, 합은 1이며 점화식을 직접 한 단계씩 계산한 결과와 일치한다. 또한 놀이 200 000회를 컴퓨터로 모의실험한 결과 $A_3 \approx 0.313$, $C_3 \approx 0.188$을 얻었다.
- $n \to \infty$일 때 모든 확률은 $\frac{1}{4}$로 수렴(converge)한다. 호루라기가 여러 번 울리고 나면 대칭성에서 예상되듯 소포는 어느 아이에게나 같은 확률로 있게 된다.

흔한 실수

- B_{n+1}을 B**로** 들어오는 확률이 아니라 B**에서** 나가는 확률로 나타내는 것. "바로 직전에 소포가 어디에 있었을 수 있는가?"를 물어보라.
- 모든 $n \geq 0$에 대해 $A_n = \frac{1}{4} + 2^{-(n+1)}$이라고 쓰는 것. 이 식은 $n = 0$에서 틀리다.
- 아무 근거도 없이 "대칭성에 의해" $B_n = D_n$이라고 가정하는 것. 차에 대한 점화식이 이를 엄밀하게 만들어 준다.

STEP 너머

이것은 **마르코프 연쇄**(Markov chain)이다. 즉 네 상태로 이루어진 순환 위에서 전이 행렬(transition matrix) P를 갖는 무작위 행보(random walk)이며, $(A_n, B_n, C_n, D_n) = (1,0,0,0)P^n$이다. 결합 $B_n - D_n$, $A_n - C_n$, $A_n + C_n - B_n - D_n$은 P의 고유벡터(eigenvector)에서 나오고(각각의 고윳값(eigenvalue)은 $\frac{1}{2}$, $\frac{1}{2}$, 0이다), 균등 분포(uniform distribution)로 2^{-n}의 속도로 수렴하는 현상은 두 번째로 큰 고윳값이 결정한다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 마르코프 연쇄와 그 장기적 거동을 친절하게 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.