풀이

STEP II 2015, 문제 12

확률과 통계 • 동전 수열 게임

> **STEP II 2015, 12번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/15-s2-q12

핵심 아이디어

여기서 다루는 수열은 모두 길이가 3이므로, 어떤 던지기에서 게임이 끝나는지는 그 던지기와 바로 앞의 두 던지기에만 달려 있다. 따라서 **마지막 두 번의 던지기**를 추적하는 것이 자연스럽다. 가능한 상태(state)는 HH, HT, TH, TT의 네 가지뿐이고, 각 상태에서 다음 던지기는 게임을 끝내거나 다른 상태로 옮겨 간다. 어떤 상태에서는 승자가 곧바로 정해진다. 나머지 상태에 대해서는 다음 던지기에 대한 조건부 확률(conditioning)을 이용해 간단한 연립방정식을 세운다.

풀이

전체에 걸쳐 각 던지기는 독립(independent)이고, 각각 확률 $\frac{1}{2}$로 H 또는 T가 나온다. 처음 두 번의 던지기만으로 완성되는 수열은 없으므로, 게임은 항상 두 번 던진 뒤의 네 상태 중 하나에 도달하며 각 상태의 확률은 $\frac{1}{4}$이다. 또한 게임이 확률 1로 끝난다는 점에 주목하자. 어떤 위치에서든 다음 세 번의 던지기가 주어진 수열을 완성할 확률은 $\frac{1}{8}$이므로, 게임이 앞으로 $3n$번보다 더 오래 계속될 확률은 많아야 $(\frac{7}{8})^n \to 0$이다.

소문항 (i)

여기서는 HHT(A의 수열)와 THH(B의 수열)만 고려한다.

처음 두 번이 HH인 경우: 동전은 한동안 H를 계속 보이고, 처음 나오는 T는 앞에 적어도 두 개의 H가 있으므로 HHT를 완성한다. 그 T 이전에는 뒷면이 한 번도 나오지 않았으므로 THH는 나올 수 없다. 따라서 A가 이긴다.

처음 두 번 중에 T가 있는 경우: 그 T 이후 처음으로 앞면이 두 번 연속 나오는 때를 생각하

자. 이 HH 바로 앞의 던지기는 H일 수 없다(그렇다면 T 이후에 더 이른 HH가 있었거나, 이 HH가 처음 두 번의 던지기여야 하는데, 이는 이미 배제하였다). 따라서 그것은 T이고 THH가 완성된다. HHT는 HH가 있어야 하는데 이것이 첫 번째 HH이므로, HHT가 그 전에 나왔을 수 없다. 따라서 B가 이긴다.

그러므로 A는 처음 두 번이 모두 앞면일 때, 그리고 그때에만 이기며,

$$\mathrm{P}(A \text{ 승리}) = \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{4},$$

이다. 이로써 요구된 결과를 얻는다.

소문항 (ii)

이제 네 수열 HHT, THH, TTH, HTT가 모두 유효하다. 처음 두 번의 던지기에 대해 조건을 걸어 생각한다.

- **HH.** 소문항 (i)에서처럼 던지기는 처음 T가 나올 때까지 H, H, ...로 이어지고, 그 T가 HHT를 완성한다. 다른 수열 중 어느 것도 HH 뒤에 H가 오는 형태를 포함하지 않으며, THH, TTH, HTT를 시작할 뒷면도 아직 나오지 않았다. 따라서 A가 이긴다.
- **TT.** H와 T의 역할을 바꾸어 같은 논증을 하면, 처음 나오는 H가 TTH를 완성하므로 C가 이긴다.
- **HT.** 다음 던지기가 T이면 HTT가 되어 D가 이긴다. H이면 마지막 두 번은 TH가 된다(그리고 HTH는 아무 수열도 완성하지 않는다).
- **TH.** 다음 던지기가 H이면 THH가 되어 B가 이긴다. T이면 마지막 두 번은 HT가 된다(그리고 THT는 아무 수열도 완성하지 않는다).

따라서 게임이 일단 상태 HT나 TH에 들어가면, D나 B가 이길 때까지 이 두 상태 사이를 오간다. HT에서 시작할 때와 TH에서 시작할 때 D가 이길 확률을 각각 d_1, d_2라 하자. 다음 던지기에 대해 조건을 걸면

$$d_1 = \tfrac{1}{2} + \tfrac{1}{2}d_2, \qquad d_2 = \tfrac{1}{2}d_1.$$

대입하면 $d_1 = \frac{1}{2} + \frac{1}{4}d_1$이므로 $d_1 = \frac{2}{3}$, $d_2 = \frac{1}{3}$이다. 이 상태들에서는 B나 D만 이길 수 있으므로, B는 HT에서 확률 $\frac{1}{3}$, TH에서 확률 $\frac{2}{3}$로 이긴다. 그러므로

$$\mathrm{P}(A) = \tfrac{1}{4}, \qquad \mathrm{P}(C) = \tfrac{1}{4},$$
$$\mathrm{P}(D) = \tfrac{1}{4} \cdot \tfrac{2}{3} + \tfrac{1}{4} \cdot \tfrac{1}{3} = \tfrac{1}{4}, \qquad \mathrm{P}(B) = \tfrac{1}{4} \cdot \tfrac{1}{3} + \tfrac{1}{4} \cdot \tfrac{2}{3} = \tfrac{1}{4}.$$

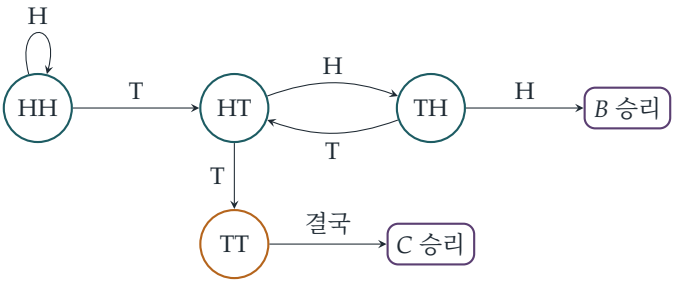
각 참가자가 이길 확률은 모두 $\frac{1}{4}$이다.

소문항 (iii)

이제 THH(B의 수열)와 TTH(C의 수열)만 고려한다.

처음 두 번이 TT인 경우. 그 뒤 던지기는 처음 H가 나올 때까지 T, T, ...로 이어지고, 그 H가 TTH를 완성한다. THH가 나오려면 TH 뒤에 H가 와야 하는데 아직 H가 한 번도 나오지 않았으므로, B가 먼저 이겼을 수는 없다. 따라서 C가 확률 1로 이긴다.

다른 시작의 경우에는 게임이 끝나지 않았다는 조건에서 상태(마지막 두 번의 던지기)가 어떻게 바뀌는지 보면 도움이 된다.



(TT에서 다음에 T가 나오면 TT에 머물고, H가 나오면 TTH가 완성된다. 그래서 TT를 C의 승리로 취급할 수 있다.)

HT, TH, HH에서 C가 이길 확률을 각각 p, q, r라 하자.

HT에서. T가 나오면 TT가 되고, 여기서 C는 확률 1로 이긴다. H가 나오면 HTH가 되는데, 이는 어느 수열도 완성하지 않고 상태 TH에 놓이게 한다. 따라서

$$p = \tfrac{1}{2} \cdot 1 + \tfrac{1}{2} \cdot q = \tfrac{1}{2} + \tfrac{1}{2}q,$$

이고, 이로써 요구된 결과를 얻는다.

TH에서. H가 나오면 THH가 완성되어 B가 이긴다. T가 나오면 THT가 되어 아무것도 완성되지 않고 상태 HT에 놓인다. 따라서 $q = \tfrac{1}{2} \cdot 0 + \tfrac{1}{2}p = \tfrac{1}{2}p$이다.

> ? **생각해 보기** 마지막 두 번의 던지기를 알고 나면 왜 그 이전의 기록은 중요하지 않을까?

두 목표 수열은 모두 길이가 3이므로, 다음 던지기가 게임을 끝내는지는 그 던지기와 직전 두 번의 던지기에만 달려 있다. 그보다 앞선 던지기는 무관하며, 앞으로의 던지기는 과거와 독립이다. 따라서 주어진 상태에서 C가 이길 확률은 그 상태에 어떻게 도달했는지와 상관없이 같고, 바로 이 때문에 p에 대한 식의 우변에 q가 다시 나타날 수 있다.

HH에서. H가 나오면 HH에 머물고, T가 나오면 HT로 옮겨 간다. 따라서 $r = \tfrac{1}{2}r + \tfrac{1}{2}p$이고, $r = p$를 얻는다.

풀면 $p = \tfrac{1}{2} + \tfrac{1}{4}p$이므로

$$p = \tfrac{2}{3}, \qquad q = \tfrac{1}{3}, \qquad r = \tfrac{2}{3}.$$

마지막으로, 각각 확률 $\frac{1}{4}$인 처음 두 번의 던지기에 대해 조건을 걸면

$$\mathrm{P}(C \text{ 승리}) = \tfrac{1}{4}(1 + p + q + r) = \tfrac{1}{4}\left(1 + \tfrac{2}{3} + \tfrac{1}{3} + \tfrac{2}{3}\right) = \tfrac{1}{4} \cdot \tfrac{8}{3} = \tfrac{2}{3}.$$

검산

- 소문항 (ii)에서 네 확률의 합은 1이고, 답은 대칭적이다. H와 T를 바꾸면 $A \leftrightarrow C$, $B \leftrightarrow D$가 서로 바뀌므로, $\mathrm{P}(A) = \mathrm{P}(C)$와 $\mathrm{P}(B) = \mathrm{P}(D)$는 당연히 성립해야 한다.
- 소문항 (iii)에서 $q = \frac{1}{3} < p$인 것은 이치에 맞는다. TH에서는 B가 한 번만 더 던지면 이길 수 있기 때문이다.
- 각 유형의 게임을 200,000번씩 컴퓨터로 모의실험(simulation)한 결과, 소문항 (i)에서 $\mathrm{P}(A) \approx 0.250$, 소문항 (ii)에서 네 참가자 모두 0.25에 가까운 값, 소문항 (iii)에서 $\mathrm{P}(C) \approx 0.667$을 얻었으며, 이는 마르코프 연쇄(Markov chain)를 따라 정확히 계산한 값과 일치한다.

흔한 실수

- 길이가 같은 두 수열은 먼저 나올 가능성도 같아야 한다고 가정하는 것. HHT와 THH는 각각 따로 보면 기대 대기 시간(expected waiting time)이 8번의 던지기로 같지만, THH는 네 번 중 세 번꼴로 HHT를 이긴다.
- 이기지 못한 던지기도 여전히 쓸모 있는 상태로 이끌 수 있다는 점을 잊는 것. HT에서 H가 나오는 것은 단지 "실패"가 아니라 TH로 옮겨 가는 것이다.
- HTH가 왜 TH에서 시작하는 것과 같은 위치에 놓이게 하는지 설명하지 않고 $p = \frac{1}{2} + \frac{1}{2}q$라고 쓰는 것.

STEP 너머

여기서 쓴 방법은 **마르코프 연쇄**로 들어가는 첫걸음이다. 마르코프 연쇄는 미래가 현재 상태에만 의존하는 과정으로, 흡수 확률(absorption probability)에 대한 일차방정식을 풀어 분석한다. 이 게임은 페니 게임(Penney's game)의 한 형태이다. 페니 게임에서는 상대가 세 번의 던지기로 이루어진 어떤 수열을 고르든, 항상 적어도 $\frac{2}{3}$의 확률로 그것을 이기는 수열을 고를 수 있다. 따라서 "어느 수열이 먼저 나오는가"는 추이적(transitive) 관계가 아니다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 마르코프 연쇄와 흡수 확률을 바로 이 수준에서 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.