풀이

STEP III 2001, 문제 13

확률과 통계 • 동전 수열 게임

STEP III 2001, 13번 문제

더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:

step.maths.org/questions/01-s3-q13

이 문제는 *Beyond STEP*의 12.8 절 **페니의 게임**에서 단계별 안내와 함께 자세히 다룬다:
https://beyondstep.keithmcnulty.org/ko/ch12/penney.html

핵심 아이디어

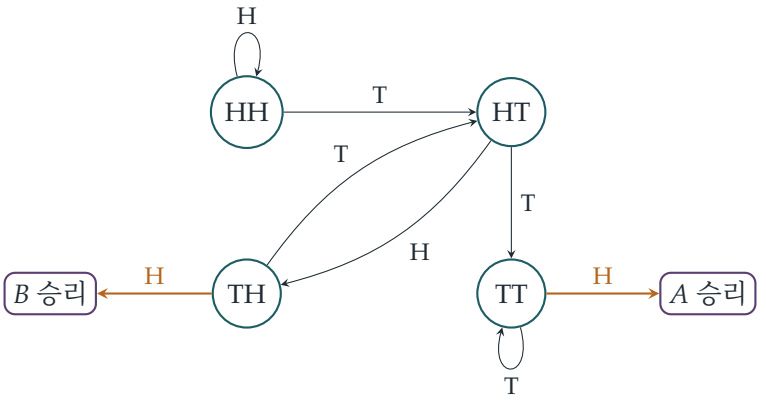
두 목표 수열의 길이가 모두 3이므로 세 번째 던지기 전에는 아무도 이길 수 없고, 그 이후로 앞으로의 진행에 영향을 주는 것은 **마지막 두 번의 결과**뿐이다. 따라서 "상태(state)"는 HH, HT, TH, TT의 네 가지뿐이며, 다음 던지기의 결과에 대해 조건을 걸면(conditioning) 게임은 각 상태에서의 승리 확률에 대한 작은 연립일차방정식(system of linear equations)이 된다. 소문항 (ii)와 (iii)도 같은 방법으로 풀 수 있지만, 네 게임 가운데 두 게임에서는 짧은 직접 논증이 더 빠르다.

풀이

이 풀이 전체에서 각 던지기는 독립이고, 각각 확률 $\frac{1}{2}$로 H 또는 T가 나온다. 적어도 두 번 던졌고 아직 아무도 이기지 않았다면, 다음 결과와 마지막 두 결과만으로 지금 누가 이기는지, 그리고 이기지 않았다면 새로운 마지막 두 결과가 무엇인지가 결정된다. 각 목표 수열의 길이가 3뿐이므로 그 이전의 결과는 관계가 없다. 따라서 어느 시점에서의 승리 확률은 마지막 두 결과에만 의존하며, 그 두 결과가 새 게임의 처음 두 결과였을 때의 승리 확률과 같다.

소문항 (i)

처음 두 결과가 XY일 때 A(TTH)가 B(THH)를 이길 확률을 p_{XY}라 하자. 아래 그림은 다음 결과에 따라 상태가 어떻게 바뀌는지를 보여 준다.



$p_{\mathrm{TT}} = 1$인 이유. 마지막 두 결과가 TT라고 하자. 확률 1로 언젠가 앞면이 나온다(n번 연속 뒷면이 나올 확률은 $2^{-n} \to 0$이다). 그때까지는 모두 뒷면이므로 THH가 나타날 수 없고, 첫 앞면 앞에는 적어도 두 개의 뒷면이 있으므로 그 앞면이 TTH를 완성한다. 따라서 A가 반드시 이기며, $p_{\mathrm{TT}} = 1$이다.

$p_{\mathrm{HT}} = \frac{1}{2}p_{\mathrm{TH}} + \frac{1}{2}p_{\mathrm{TT}}$인 이유. HT 다음에 나오는 HTH와 HTT는 어느 쪽도 목표 수열이 아니므로 세 번째 던지기에서는 아무도 이기지 않는다. 세 번째 결과가 H이면(확률 $\frac{1}{2}$) 마지막 두 결과는 TH가 되고, T이면 TT가 된다. 세 번째 결과에 대해 조건을 걸고, 위에서 말한 대로 마지막 두 결과만 중요하다는 사실을 이용하면

$$p_{\mathrm{HT}} = \tfrac{1}{2}p_{\mathrm{TH}} + \tfrac{1}{2}p_{\mathrm{TT}}.$$

나머지 두 상태. TH 다음에 앞면이 나오면 THH가 되어 B가 이기고, 뒷면이 나오면 HT가 남는다. HH 다음에 앞면이 나오면 HH가 남고, 뒷면이 나오면 HT가 남는다(여기서 HHH, HHT는 목표 수열이 아니다). 따라서

$$p_{\mathrm{TH}} = \tfrac{1}{2}\cdot 0 + \tfrac{1}{2}p_{\mathrm{HT}}, \qquad p_{\mathrm{HH}} = \tfrac{1}{2}p_{\mathrm{HH}} + \tfrac{1}{2}p_{\mathrm{HT}}.$$

두 번째 식에서 $p_{\mathrm{HH}} = p_{\mathrm{HT}}$를 얻는다. p_{HT}에 대한 식에 $p_{\mathrm{TH}} = \frac{1}{2}p_{\mathrm{HT}}$와 $p_{\mathrm{TT}} = 1$을 대입하면

$$p_{\mathrm{HT}} = \tfrac{1}{4}p_{\mathrm{HT}} + \tfrac{1}{2} \quad\Longrightarrow\quad \tfrac{3}{4}p_{\mathrm{HT}} = \tfrac{1}{2} \quad\Longrightarrow\quad p_{\mathrm{HT}} = \tfrac{2}{3}.$$

따라서 $p_{\mathrm{HH}} = p_{\mathrm{HT}} = \frac{2}{3}$이고 $p_{\mathrm{TH}} = \frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3}$이므로, 요구된 결과가 성립한다.

생각해 보기 예를 들어 TH가 실제로는 5번째와 6번째 던지기에서 나왔는데도, 마치 TH가 처음 두 결과인 것처럼 TH에서 게임을 "다시 시작"해도 되는 이유는 무엇일까?

게임의 앞으로의 진행은 앞으로의 던지기와 마지막 두 결과에만 의존하기 때문이다. 아직 완성되지 않은 TTH나 THH는 이미 나온 결과 가운데 많아야 마지막 두 개만 사용할 수 있

다. 그 이전의 기록은 어느 수열에도 도움이 되지 않으며, 앞으로의 던지기는 그것과 독립이다.

전체 확률. 처음 두 번의 던지기에서는 아무도 이길 수 없고, 가능한 네 쌍은 각각 확률 $\frac{1}{4}$를 가진다. 전확률 법칙(law of total probability)에 의해

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

소문항 (ii)

이제 B는 THH, C는 HHT를 가진다. 여기서는 직접 논증이 가장 깔끔하다.

처음 두 결과가 HH이면 C가 이긴다. 그 뒤에 처음 나오는 뒷면(확률 1로 나온다) 앞에는 HH가 있으므로 HHT가 완성되고, 그 전에는 모두 앞면이므로 THH가 나타났을 수 없다.

그렇지 않으면 처음 두 결과 가운데 뒷면이 있다. 확률 1로 패턴 HH가 언젠가 나타난다. 그 첫 번째 출현을 생각하자. 그것은 1번째 던지기에서 시작하지 않으므로 바로 앞의 결과가 존재하고, 그 결과는 H가 아니므로(그렇다면 더 이른 HH가 있었을 것이다) T이다. 따라서 첫 번째 HH는 THH를 완성한다. 한편 HHT가 그 전에 나타나려면 더 이른 HH가 필요하므로 HHT는 그 전에 나타났을 수 없다. 따라서 B가 이긴다.

그러므로 처음 두 결과가 HH가 아닌 한 B가 이긴다.

$$\mathrm{P}(B \text{ 승리}) = 1 - \tfrac{1}{4} = \tfrac{3}{4}.$$

소문항 (iii)

$C(HHT)$ 대 $D(HTT)$. 마지막 두 결과가 XY일 때 C가 이길 확률을 q_{XY}라 하자. HH에서는 (ii)와 같은 이유로 C가 반드시 이긴다. 처음 나오는 뒷면이 HHT를 완성하고, HTT는 HT 뒤에 뒷면이 와야 하므로 먼저 나올 수 없다. 따라서 $q_{\mathrm{HH}} = 1$이다. 나머지 각 상태에서 다음 결과에 대해 조건을 걸면

$$\begin{aligned}
&\text{HT:} && \text{T이면 HTT } (D \text{ 승리}),\ \text{H이면 TH:} && q_{\mathrm{HT}} = \tfrac{1}{2} q_{\mathrm{TH}},\\
&\text{TH:} && \text{H이면 HH, T이면 HT:} && q_{\mathrm{TH}} = \tfrac{1}{2} + \tfrac{1}{2} q_{\mathrm{HT}},\\
&\text{TT:} && \text{H이면 TH, T이면 TT:} && q_{\mathrm{TT}} = \tfrac{1}{2} q_{\mathrm{TH}} + \tfrac{1}{2} q_{\mathrm{TT}}.
\end{aligned}$$

마지막 식에서 $q_{\mathrm{TT}} = q_{\mathrm{TH}}$를 얻는다. 첫 번째 식을 두 번째 식에 대입하면 $q_{\mathrm{TH}} = \frac{1}{2} + \frac{1}{4} q_{\mathrm{TH}}$이므로 $q_{\mathrm{TH}} = \frac{2}{3}$, $q_{\mathrm{TT}} = \frac{2}{3}$, $q_{\mathrm{HT}} = \frac{1}{3}$이다. 따라서

$$\mathrm{P}(C \text{ 승리}) = \tfrac{1}{4}\left(1 + \tfrac{1}{3} + \tfrac{2}{3} + \tfrac{2}{3}\right) = \tfrac{1}{4} \cdot \tfrac{8}{3} = \tfrac{2}{3} > \tfrac{1}{2},$$

이므로 C가 D보다 이길 가능성이 크다.

$D(HTT)$ 대 $A(TTH)$. 이것은 소문항 (ii)에서 H와 T를 바꾼 것과 같다. 처음 두 결과가 TT이면 A가 이긴다. 처음 나오는 앞면이 TTH를 완성하고, HTT는 앞면 뒤에 뒷면 두 개가 와야

하므로 그 첫 앞면 전에는 나타날 수 없다. 그렇지 않으면 처음 두 결과 가운데 앞면이 있다. 확률 1로 패턴 TT가 언젠가 나타난다. 그 첫 번째 출현은 1번째 던지기에서 시작하지 않으며, 바로 앞의 결과는 T일 수 없으므로(그렇다면 더 이른 TT가 있었을 것이다) H이고, 따라서 첫 번째 TT는 HTT를 완성한다. TTH는 더 이른 TT가 필요하므로 그 전에 나타났을 수 없다. 따라서 D가 이기며,

$$\mathrm{P}(D \text{ 승리}) = 1 - \tfrac{1}{4} = \tfrac{3}{4} > \tfrac{1}{2}.$$

그러므로 D가 A보다 이길 가능성이 크다.

> **생각해 보기** A는 B를, B는 C를, C는 D를, D는 A를 이긴다. 어떻게 이럴 수 있을까?

"이길 가능성이 더 크다"는 추이적(transitive) 관계가 아니므로 모순이 없다. 네 사람은 순환을 이루며, 각자의 게임에서 A, B, C, D가 이길 확률은 각각 $\frac{2}{3}$, $\frac{3}{4}$, $\frac{2}{3}$, $\frac{3}{4}$이다. 두 번째로 고르는 사람은 언제나 첫 번째 사람의 수열을 이기는 수열을 고를 수 있다.

검산

- 소문항 (i)에서 다시 대입해 보면 $\frac{1}{2}\cdot\frac{1}{3} + \frac{1}{2}\cdot 1 = \frac{2}{3} = p_{\mathrm{HT}}$이고, $\frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3} = p_{\mathrm{TH}}$이다.
- 상태를 이용하는 방법은 소문항 (ii)와 D–A 게임에도 통한다. B 대 C에서 HH, HT, TH, TT로부터 B가 이길 확률은 $0, 1, 1, 1$이고, 그 평균은 $\frac{3}{4}$이다.
- 각 대결마다 200 000번의 게임을 시뮬레이션한 결과, 승리 빈도는 0.664, 0.752, 0.666, 0.749로 정확한 값 $\frac{2}{3}$, $\frac{3}{4}$, $\frac{2}{3}$, $\frac{3}{4}$과 잘 맞았다.

흔한 실수

- 세 번째 던지기에서 게임이 끝날 수 있다는 것을 잊는 실수. 예를 들어 TH 다음에 H가 나오면 B가 곧바로 이긴다는 것을 알아차리지 못하고 $p_{\mathrm{TH}} = \frac{1}{2}p_{\mathrm{HH}} + \frac{1}{2}p_{\mathrm{HT}}$라고 쓰는 경우이다.
- 처음 두 결과의 각 쌍이 확률 $\frac{1}{4}$를 가지며 세 번째 던지기 전에는 아무도 이길 수 없다는 것을 밝히지 않고 네 상태의 확률을 평균 내는 실수.

STEP 너머

이것은 페니 게임(Penney's game)으로, 비추이적(non-transitive) 확률의 유명한 예이다. 상태를 이용하는 방법은 흡수 마르코프 연쇄(absorbing Markov chain)의 이론이며, 여기서 흡수 확률은 이 풀이에서 푼 것과 정확히 같은 종류의 일차방정식을 만족한다. 임의의 두

수열 사이의 승산을 주는 닫힌 공식도 있는데, 바로 콘웨이(Conway)의 "선행 수(leading numbers)" 알고리즘이며, 이는 마팅게일(martingale)을 이용해 증명할 수 있다. 마르코프 연쇄와 첫 단계 분석(first-step analysis)에 대해서는 Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)이나 Blitzstein과 Hwang의 *Introduction to Probability*(확률 입문)를 참고하라.

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.