풀이

# STEP II 견본 문제지, 문제 16

확률과 통계 • 토너먼트의 확률

**STEP II 견본 문제지, 16번 문제**

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

step.maths.org/questions/spec-s2-q16

이 문제는 *Beyond STEP*의 12.7 절 **토너먼트에서 누가 누구와 만나는가**에서 단계별 안내와 함께 자세히 다룬다: https://beyondstep.keithmcnulty.org/ko/ch12/knockout.html

## 핵심 아이디어

선택한 두 선수 $A$와 $B$를 라운드별로 따라간다. 두 선수가 모두 아직 남아 있고 아직 서로 만나지 않은 라운드에서는 두 가지만 중요하다. $A$가 $B$와 대진되는가, 그리고 그렇지 않다면 두 선수가 모두 이기는가? 각각은 간단한 확률(probability)이고, 이때 나타나는 곱들은 아름답게 연쇄적으로 소거된다(telescoping). 마지막 문항에는 깔끔한 대칭성(symmetry) 논법도 있다. 바로 토너먼트 전체의 경기 수를 세는 것이다.

## 풀이

$N = 2^n$이라 하자. 선택한 두 선수를 $A$와 $B$라 부른다. $r-1$개의 라운드가 끝나면 $2^{n-r+1}$명의 선수가 남으므로, 라운드 $r$에는 $M_r = 2^{n-r+1}$명의 선수가 있고, 라운드 $n$($M_n = 2$)이 결승전이다.

반복해서 사용할 사실은 다음과 같다. $A$와 $B$가 모두 무작위로 짝지어지는 $M$명의 선수 안에 있다면, $A$의 상대는 나머지 $M-1$명 중 누구든 같은 가능성으로 정해지므로

$$\mathrm{P}(A\text{가 }B\text{와 대진됨}) = \frac{1}{M-1}.$$

(결승전에서는 대진 추첨이 없지만, $M = 2$이면 이 공식은 1을 주므로 올바르다.)

**소문항 (i)**

첫 라운드에서 $M_1 = N$이므로

$$\mathrm{P}(A\text{와 } B\text{가 첫 라운드에서 만남}) = \frac{1}{2^n - 1}.$$

**소문항 (ii)**

$S_r$를 라운드 $r$가 시작될 때 $A$와 $B$가 모두 아직 토너먼트에 남아 있는 사건이라 하자(경기에서 진 선수는 탈락하므로, 이 경우 두 선수는 아직 만나지 않았어야 한다). $S_r$가 주어졌을 때, 두 선수가 함께 라운드 $r+1$에 진출하는 것은 라운드 $r$에서 서로 대진되지 **않고**, 각자의 (서로 다른) 경기에서 모두 이길 때이며, 그때뿐이다. 두 경기 결과는 독립(independent)이고 각각의 확률은 $\frac{1}{2}$이므로

$$\mathrm{P}(S_{r+1} \mid S_r) = \frac{M_r - 2}{M_r - 1} \cdot \frac{1}{4}.$$

? **생각해 보기** 이것을 실제로 계산할 수 있는 곱으로 어떻게 바꿀 수 있을까?

$M = 2^k$라 쓰고, $M - 2 = 2(2^{k-1} - 1)$임에 주목하면

$$\frac{M-2}{M-1} = \frac{2\,(2^{k-1}-1)}{2^k - 1}.$$

한 라운드의 분자는 (인수 2를 제외하면) 다음 라운드의 분모이다. $S_1$은 확실한 사건이므로, $k = n, n-1, \ldots, n-r+2$인 라운드들에 대해 곱하면 연쇄적으로 소거되는 곱을 얻는다.

$$\mathrm{P}(S_r) = \frac{1}{4^{r-1}} \prod_{k=n-r+2}^{n} \frac{2\,(2^{k-1}-1)}{2^k - 1} = \frac{2^{r-1}}{4^{r-1}} \cdot \frac{2^{n-r+1} - 1}{2^n - 1} = \frac{2^{n-r+1} - 1}{2^{r-1}\,(2^n - 1)}.$$

$S_r$가 주어졌을 때, 두 선수가 라운드 $r$에서 만날 확률은 $1/(M_r - 1) = 1/(2^{n-r+1} - 1)$이므로

$$\mathrm{P}(A\text{와 } B\text{가 라운드 } r\text{에서 만남}) = \frac{1}{2^{r-1}\,(2^n - 1)}.$$

이는 기분 좋을 만큼 간단한 공식이다. 라운드가 하나 지날 때마다 그 라운드에서 만날 확률이 절반으로 줄어든다. $r = 1$로 두면 소문항 (i)의 결과가 다시 나온다. 결승전은 $r = n$이므로

$$\mathrm{P}(A\text{와 } B\text{가 결승전에서 만남}) = \frac{1}{2^{n-1}\,(2^n - 1)}.$$

대칭성으로 이를 검증해 보자. 규칙은 모든 선수를 똑같이 다루므로, 결승 진출자 쌍은 $\binom{N}{2} = 2^{n-1}(2^n - 1)$개의 선수 쌍 중 어느 것이든 같은 가능성으로 정해지고, 이로부터 같은 답을 얻는다.

**소문항 (iii)**

$r = 1, \ldots, n$에 대하여 사건 "라운드 $r$에서 만남"들은 서로 배반(mutually exclusive)이다(진 선수는 떠나므로 두 선수는 많아야 한 번 만날 수 있다). 따라서 등비급수(geometric series)를 이용하여 더하면

$$\begin{aligned}\mathrm{P}(A\text{와 } B\text{가 만남}) &= \frac{1}{2^n-1}\sum_{r=1}^{n}\frac{1}{2^{r-1}} = \frac{1}{2^n-1}\cdot 2\left(1-\frac{1}{2^n}\right)\\ &= \frac{2}{2^n}\cdot\frac{2^n-1}{2^n-1} = \frac{1}{2^{n-1}}.\end{aligned}$$

**생각해 보기** 합을 전혀 계산하지 않고 이 답을 알아낼 방법이 있을까?

있다. 모든 경기는 정확히 한 명의 선수를 탈락시키고, 우승자를 제외한 모든 선수가 탈락하므로, 경기 수는 언제나 정확히 $2^n - 1$이다. 각 경기는 한 쌍의 선수가 치르며, 대칭성에 의해 $\binom{N}{2}$개의 쌍 각각이 경기를 치를 확률 $p$는 모두 같다. 경기 수의 기댓값은 모든 쌍에 대한 이 확률들의 합이므로

$$\binom{2^n}{2}p = 2^n - 1 \quad\Longrightarrow\quad p = \frac{2(2^n-1)}{2^n(2^n-1)} = \frac{2}{2^n} = \frac{1}{2^{n-1}},$$

이고, 이는 합으로 구한 결과와 일치한다.

## 검산

- $n = 1$: 선수가 두 명이고, 이들은 반드시 첫 라운드에서 만나며 그것이 곧 결승전이다. 세 공식 모두 1을 준다.
- $n = 2$: 선수가 네 명이다. 소문항 (i)은 $\frac{1}{3}$, 결승전은 $\frac{1}{6}$을 주고, 합은 소문항 (iii)의 말대로 $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$이다.
- $n = 3$: 모든 대진과 모든 결과를 정확히 나열하면 $\frac{1}{7}$, $\frac{1}{28}$, $\frac{1}{4}$를 얻는다. 선수 16명으로 토너먼트 $2 \times 10^5$회를 모의실험한 결과는 0.0664, 0.0085, 0.1247이었고, 정확한 값은 $\frac{1}{15} \approx 0.0667$, $\frac{1}{120} \approx 0.0083$, $\frac{1}{8}$이다.

## 흔한 실수

- 나중 라운드에서 만나려면 두 선수가 먼저 그 이전의 모든 라운드에서 서로를 **피해야** 한다는 점을 잊는 것. 인수 $\frac{M-2}{M-1}$를 빠뜨리기 쉽다.
- 소문항 (ii)의 답을 $\left(\frac{1}{2^{n-1}}\right)^2$로 쓰는 것. 이는 **두 선수가 독립이라면** 둘 다 결승전에 진출할 확률이다. 그러나 두 선수는 독립이 아니다. 일찍 서로 대진되면 많아야 한 명만 올라갈 수 있다.

## STEP 너머

소문항 (iii)의 논법은 지시확률변수(indicator variable)를 이용한 **기댓값의 선형성**(*linearity of expectation*)이다. 경기 수의 기댓값은 사건들이 독립이든 아니든, 모든 쌍에 대하여 그 쌍이 경기를 치를 확률을 더한 것과 같다. 이는 확률론과 조합론에서 가장 강력한 도구 중 하나이며, 주어진 성질을 갖는 대상이 존재함을 증명하는 "확률론적 방법(probabilistic method)"의 바탕이 된다. Blitzstein과 Hwang의 *Introduction to Probability*(확률론 입문)에는 지시확률변수 논법의 멋진 예들을 담은 장이 통째로 하나 있다.

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.