풀이

STEP I 2009, 문제 13

확률과 통계 • 배열과 연

> **STEP I 2009, 13번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/09-s1-q13

핵심 아이디어

중요한 것은 여학생들의 **위치**뿐이고, 어느 여학생 또는 어느 남학생이 누구인지는 중요하지 않다. 따라서 $(n+3)!$가지 순서를 일일이 다루는 대신, 여학생들이 차지하는 세 자리의 집합을 고른다. $\binom{n+3}{3}$개의 집합은 모두 똑같이 일어날 가능성이 있다. 그러면 유리한 집합을 세는 일은 n명의 남학생을 먼저 줄 세운 다음, 여학생들이 그 사이와 양 끝의 틈 가운데 어디로 들어가는지를 묻는 방식이 가장 쉽다.

풀이

모든 $(n+3)!$가지 순서가 똑같은 확률로 일어나고, 여학생들이 차지하는 세 자리의 각 집합은 같은 개수, 즉 $3!\,n!$개의 순서에서 나오므로, 여학생들의 위치 집합은 가능한

$$\binom{n+3}{3} = \frac{(n+3)(n+2)(n+1)}{6}$$

개의 집합 위에서 균등분포(uniform distribution)를 따른다. 전체에 걸쳐 K는 1, 2, 3 가운데 하나의 값만 가질 수 있다.

한 줄로 선 n명의 남학생은 $n+1$개의 틈(gap)을 만든다(첫 번째 남학생 앞에 하나, 마지막 남학생 뒤에 하나, 이웃한 남학생 사이에 $n-1$개):



여학생들을 틈에 넣으면(한 틈에 몇 명이든 넣을 수 있다) 하나의 줄이 만들어지고, 모든 줄은 정확히 한 가지 방법으로 만들어진다. 같은 틈에 들어간 여학생들은 서로 이웃해 앉고,

서로 다른 틈에 들어간 여학생들은 적어도 한 명의 남학생으로 떨어져 있다.

소문항 (i)

$K = 3$은 세 여학생이 모두 한 덩어리를 이룬다는 뜻, 즉 모두 같은 틈에 들어간다는 뜻이다. 틈을 고르는 방법은 $n+1$가지이므로

$$\mathrm{P}(K=3) = \frac{n+1}{\binom{n+3}{3}} = \frac{6(n+1)}{(n+3)(n+2)(n+1)} = \frac{6}{(n+2)(n+3)}.$$

(같은 말로, 연속한 세 자리로 된 덩어리는 $1, 2, \ldots, n+1$ 가운데 어느 위치에서든 시작할 수 있다.)

소문항 (ii)

$K = 1$은 어떤 두 여학생도 이웃하지 않는다는 뜻, 즉 여학생들이 세 개의 **서로 다른** 틈에 들어간다는 뜻이다. $n+1$개의 틈 가운데 3개를 $\binom{n+1}{3}$가지 방법으로 고르고, 각 틈에 여학생을 한 명씩 넣는다. 따라서

$$\mathrm{P}(K=1) = \frac{\binom{n+1}{3}}{\binom{n+3}{3}} = \frac{(n+1)n(n-1)}{(n+3)(n+2)(n+1)} = \frac{n(n-1)}{(n+2)(n+3)},$$

이고, 이것이 증명하려던 식이다. ($n = 0$ 또는 1이면 틈이 너무 적은데, 이때 공식도 올바르게 0을 준다.)

? **생각해 보기** 왜 경우의 수가 $\binom{n+1}{3} \times 3!$가 아니라 $\binom{n+1}{3}$인가?

우리는 이름이 붙은 아이들의 순서가 아니라 **위치**의 집합을 세고 있다. 어느 세 틈을 쓰는지 고르면 여학생들의 위치가 완전히 정해진다. 어느 여학생이 어디에 앉는지는 모든 위치 집합이 똑같은 확률로 일어난다는 주장 안에 이미 반영되어 있다.

소문항 (iii)

먼저 $\mathrm{P}(K=2)$가 필요하다. $K \in \{1, 2, 3\}$이므로

$$\begin{aligned}\mathrm{P}(K=2) &= 1 - \frac{6}{(n+2)(n+3)} - \frac{n(n-1)}{(n+2)(n+3)} \\ &= \frac{(n^2+5n+6) - 6 - (n^2-n)}{(n+2)(n+3)} = \frac{6n}{(n+2)(n+3)}.\end{aligned}$$

검산으로 직접 세어 보자. $K = 2$는 두 여학생이 한 틈에 있고 나머지 한 여학생은 다른 틈에 있다는 뜻이다. 두 명이 들어갈 틈을 고르고 나서 한 명이 들어갈 다른 틈을 고르면 $(n+1)n$개의 집합이 나오고, $(n+1)n/\binom{n+3}{3} = 6n/((n+2)(n+3))$이므로 앞의 결과와 같다.

이제 기댓값(expectation)은

$$\begin{aligned}\mathrm{E}(K) &= 1\cdot\frac{n(n-1)}{(n+2)(n+3)} + 2\cdot\frac{6n}{(n+2)(n+3)} + 3\cdot\frac{6}{(n+2)(n+3)}\\ &= \frac{n^2+11n+18}{(n+2)(n+3)} = \frac{(n+2)(n+9)}{(n+2)(n+3)}\\ &= \frac{n+9}{n+3}.\end{aligned}$$

검산

- $n=0$: 여학생만 있으므로 반드시 $K=3$이고, $\frac{n+9}{n+3}=3$이다.
- $n=1$: 줄은 GGGB, GGBG, GBGG, BGGG이고 똑같은 확률로 일어나므로, $K=3$ 또는 2가 각각 확률 $\frac{1}{2}$로 일어나고 $\mathrm{E}(K)=\frac{5}{2}=\frac{10}{4}$이다.
- $n\to\infty$일 때 $\mathrm{E}(K)\to 1$이다. 남학생이 많으면 여학생들은 거의 이웃하지 않는다.
- $n=0,\ldots,9$에 대해 모든 위치 집합을 정확한 분수로 열거했고(예를 들어 $n=3$이면 $\mathrm{P}(K=1,2,3)=\frac{1}{5},\frac{3}{5},\frac{1}{5}$이고 $\mathrm{E}(K)=2$이다), $n\le 4$에 대해 이름 붙은 아이들의 모든 $(n+3)!$가지 순서도 열거했다. 모든 결과가 일치했다.

흔한 실수

- 같은 분수 안에서 순서 있는 배열의 개수와 위치 집합의 개수를 섞어 쓰는 실수. 표본공간(sample space)을 하나 정하고 끝까지 그것만 사용하라.
- 양 끝의 두 틈을 빠뜨려서 $n+1$개 대신 $n-1$개의 틈을 쓰는 실수.

STEP 너머

"남학생이 틈을 만들고, 여학생이 틈에 들어간다"는 기법은 별과 막대(stars and bars) 아이디어이며, 무작위 배열에서 가장 긴 연속 구간(longest run)에 관한 일반적인 질문은 고전적인 문제이다. 긴 무작위 수열에서 가장 긴 연속 구간의 길이는 수열 길이의 로그처럼 증가한다. 대학 학부의 확률 강좌에서는 지시변수(indicator variable)와 생성함수(generating function)를 통해 이런 내용을 접한다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 이런 세기 논법이 전개되는 모습을 살펴보기에 좋은 책이다.

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.