풀이

STEP I 2015, 문제 13

확률과 통계 • 조건부 확률

> **STEP I 2015, 13번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/15-s1-q13

핵심 아이디어

첫 번째 6이 나오는 시점을 조건(conditioning)으로 삼는다. 첫 번째 6이 n번째 던지기에서 나온다는 것을 알면, 그 앞의 $n-1$번의 던지기는 각각 독립적으로 $1,\dots,5$ 중 어느 것이든 같은 확률로 나오므로, 첫 번째 6 이전의 4와 5에 관한 질문은 평범한 이항분포(binomial distribution) 문제가 된다. 이를 n에 대해 더하면 등비급수(geometric series)와 그 도함수가 나오는데, 바로 이 지점에서 힌트에 주어진 이항전개(binomial expansion)가 쓰인다.

풀이

각 던지기는 독립이고, 각 면이 나올 확률은 $\frac{1}{6}$이다. 첫 번째 6이 n번째 던지기에서 나오는 사건을 A라 하고 $p_n = \mathrm{P}(A)$라 쓴다.

소문항 (i)

첫 번째 6이 n번째 던지기에서 나오는 것은 $1,\dots,n-1$번째 던지기가 모두 6이 아니고 n번째 던지기가 6일 때와 정확히 같다. 독립성(independence)에 의해

$$\mathrm{P}(A) = p_n = \left(\tfrac{5}{6}\right)^{n-1}\tfrac{1}{6}, \qquad n = 1,2,3,\dots$$

조건부에 관한 사실. A가 주어지면, $1,\dots,n-1$번째 던지기 각각에 대해 알 수 있는 것은 "6이 아니다"라는 것뿐이다. 주사위가 공정하고 던지기가 독립이므로, 이 $n-1$번의 던지기는 서로 독립이고 각각 $\{1,2,3,4,5\}$ 위에서 균등분포(uniform distribution)를 따른다. 이후의 모든 소문항에서 이 사실과 함께 전확률 법칙(law of total probability) $\mathrm{P}(X) = \sum_n \mathrm{P}(X \mid A)\, p_n$을 사용한다.

소문항 (ii)

첫 번째 6 이전에 5가 한 번도 나오지 않을 확률 $\mathrm{P}(B')$를 구하는 것이 더 쉽다. A가 주어지면 앞의 $n-1$번의 던지기 각각이 5를 피할 확률은 $\frac{4}{5}$이므로

$$\mathrm{P}(B') = \sum_{n=1}^{\infty} \left(\tfrac{4}{5}\right)^{n-1} \left(\tfrac{5}{6}\right)^{n-1} \tfrac{1}{6} = \tfrac{1}{6} \sum_{n=1}^{\infty} \left(\tfrac{2}{3}\right)^{n-1} = \tfrac{1}{6} \cdot \frac{1}{1-\frac{2}{3}} = \tfrac{1}{2}.$$

따라서 $\mathrm{P}(B) = \frac{1}{2}$이다.

소문항 (iii)

포함배제의 원리(inclusion–exclusion principle)를 여사건 꼴로 쓰면

$$\mathrm{P}(B \cap C) = 1 - \mathrm{P}(B') - \mathrm{P}(C') + \mathrm{P}(B' \cap C').$$

면 4와 5 사이의 대칭성(symmetry)에 의해 $\mathrm{P}(C') = \mathrm{P}(B') = \frac{1}{2}$이다. $B' \cap C'$의 경우, A가 주어지면 앞의 각 던지기는 4와 5를 모두 피해야 하고 그 확률은 $\frac{3}{5}$이다.

$$\mathrm{P}(B' \cap C') = \sum_{n=1}^{\infty} \left(\tfrac{3}{5}\right)^{n-1} \left(\tfrac{5}{6}\right)^{n-1} \tfrac{1}{6} = \tfrac{1}{6} \cdot \frac{1}{1-\frac{1}{2}} = \tfrac{1}{3}.$$

그러므로 $\mathrm{P}(B \cap C) = 1 - \frac{1}{2} - \frac{1}{2} + \frac{1}{3} = \frac{1}{3}$이다.

소문항 (iv)

A가 주어지면 앞의 $n-1$번의 던지기 중 5가 나오는 횟수는 $\mathrm{B}(n-1, \frac{1}{5})$를 따르므로

$$\mathrm{P}(D \mid A) = (n-1)\tfrac{1}{5} \left(\tfrac{4}{5}\right)^{n-2} \qquad (n \geq 2),$$

이고 $n = 1$일 때는 0이다. 따라서

$$\mathrm{P}(D) = \sum_{n=2}^{\infty} (n-1)\tfrac{1}{5} \left(\tfrac{4}{5}\right)^{n-2} \left(\tfrac{5}{6}\right)^{n-1} \tfrac{1}{6}.$$

$\frac{5}{6}$ 하나를 밖으로 꺼내고 나머지를 짝지으면 $\frac{1}{5} \cdot \frac{5}{6} \cdot \frac{1}{6} = \frac{1}{36}$이고 $\frac{4}{5} \cdot \frac{5}{6} = \frac{2}{3}$이다. $m = n-1$로 놓으면

$$\mathrm{P}(D) = \tfrac{1}{36} \sum_{m=1}^{\infty} m \left(\tfrac{2}{3}\right)^{m-1} = \tfrac{1}{36} \left(1 - \tfrac{2}{3}\right)^{-2} = \tfrac{9}{36} = \tfrac{1}{4},$$

이며, 여기서 지수가 -2인 주어진 전개 $(1-x)^{-2} = 1 + 2x + 3x^2 + \cdots$를 사용하였다.

? **생각해 보기** 하나 또는 둘 다 일어나는 모든 경우를 나열하지 않고 $\mathrm{P}(D \cup E)$를 어떻게 구할 수 있을까?

소문항 (v)

$\mathrm{P}(D\cup E)=\mathrm{P}(D)+\mathrm{P}(E)-\mathrm{P}(D\cap E)$를 이용하고, 4–5 대칭성에 의해 $\mathrm{P}(E)=\mathrm{P}(D)=\frac{1}{4}$이다. 따라서 $\mathrm{P}(D\cap E)$만 구하면 된다.

A가 주어졌을 때, 앞의 $n-1$번의 던지기 중 4가 정확히 한 번, 5가 정확히 한 번 나온다는 것은 어느 던지기가 5이고 어느 던지기가 4인지를 고르고($(n-1)(n-2)$가지의 순서 있는 선택), 나머지 $n-3$번의 던지기가 모두 $\{1,2,3\}$에 속한다는 뜻이다.

$$\mathrm{P}(D\cap E\mid A)=(n-1)(n-2)\left(\tfrac{1}{5}\right)^2\left(\tfrac{3}{5}\right)^{n-3}\qquad(n\geq 3).$$

그러면

$$\mathrm{P}(D\cap E)=\sum_{n=3}^{\infty}(n-1)(n-2)\tfrac{1}{25}\left(\tfrac{3}{5}\right)^{n-3}\left(\tfrac{5}{6}\right)^{n-1}\tfrac{1}{6}.$$

$\left(\frac{5}{6}\right)^2$을 밖으로 꺼내면 상수 $\frac{1}{25}\cdot\frac{25}{36}\cdot\frac{1}{6}=\frac{1}{216}$이 나오고, $\frac{3}{5}\cdot\frac{5}{6}=\frac{1}{2}$이다. $k=n-3$으로 놓으면

$$\begin{aligned}\mathrm{P}(D\cap E)&=\tfrac{1}{216}\sum_{k=0}^{\infty}(k+2)(k+1)\left(\tfrac{1}{2}\right)^k=\tfrac{2}{216}\sum_{k=0}^{\infty}\frac{(k+2)!}{k!\,2!}\left(\tfrac{1}{2}\right)^k\\&=\tfrac{1}{108}\left(1-\tfrac{1}{2}\right)^{-3}=\tfrac{8}{108}=\tfrac{2}{27},\end{aligned}$$

이며, 여기서 지수가 -3인 주어진 전개를 사용하였고 그 x^k의 계수는 $\frac{(k+2)!}{k!\,2!}$이다. 그러므로

$$\mathrm{P}(D\cup E)=\tfrac{1}{4}+\tfrac{1}{4}-\tfrac{2}{27}=\tfrac{27-4}{54}=\tfrac{23}{54}.$$

검산

모든 답을 더 매끄럽게 볼 수 있는 방법이 있으며, 이는 훌륭한 검산이 된다. 1, 2, 3이 나온 던지기(그리고 (ii)와 (iv)에서는 4가 나온 던지기도)는 모두 무시한다. 이들은 사건에 전혀 영향을 주지 않는다. 남은 관련 있는 던지기들은 각 관련 면이 같은 확률로 나오는 수열을 이룬다.

- (ii): 5 또는 6이 나오는 첫 번째 던지기는 둘 중 어느 것이든 같은 확률이므로 $\mathrm{P}(B)=\frac{1}{2}$이다.
- (iii): 4, 5, 6 가운데 세 면 각각이 처음 나타나는 순서에서 마지막이 될 확률은 모두 같고, $B\cap C$는 6이 마지막이라는 뜻이므로 $\frac{1}{3}$이다.
- (iv): D는 5-또는-6 수열이 5, 6으로 시작한다는 뜻이므로 확률은 $\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$이다.
- (v): $D\cap E$는 4-5-6 수열이 4, 5, 6 또는 5, 4, 6으로 시작한다는 뜻이므로 확률은 $2\left(\frac{1}{3}\right)^3=\frac{2}{27}$이다.

다섯 개의 답 모두 3×10^5개의 던지기 수열에 대한 컴퓨터 모의실험(simulation)으로도 확인하였으며, 그 결과 $\mathrm{P}(D \cup E) \approx 0.426$을 얻어 $\frac{23}{54} \approx 0.4259$와 일치하였다.

흔한 실수

- 앞의 던지기들을 여섯 면 전체 위의 균등분포로 취급하는 것: A가 주어지면 이들은 6일 수 없으므로, 5가 나올 확률은 각각 $\frac{1}{6}$이 아니라 $\frac{1}{5}$이다.
- (v)에서 인수 $(n-1)(n-2)$를 빠뜨리거나, 두 던지기 중 어느 것이 4인지를 구별하지 않는 $\binom{n-1}{2}$를 사용하는 것.
- (v)에서 겹치는 부분을 빼지 않고 $\mathrm{P}(D) + \mathrm{P}(E)$를 그대로 더하는 것.

STEP 너머

여기서처럼 무작위 정지 시각(stopping time)을 조건으로 삼고 더하는 것은 전확률 법칙의 가장 유용한 형태이며, 여기서 만난 급수들은 음이항분포(negative binomial distribution)의 생성함수(generating function)이다. "관련 없는 던지기는 무시한다"는 검산은 강력한 아이디어의 한 예이다. 즉, 어떤 과정을 관련 있는 일이 일어나는 순간에만 관찰하면 원래 과정보다 훨씬 단순해지는 경우가 많다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)에서 두 아이디어를 모두 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.