풀이

STEP I 2017, 문제 13

확률과 통계 • 확률의 점화식

> **STEP I 2017, 13번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/17-s1-q13

핵심 아이디어

각 빵 조각을 보고 그것이 어떻게 쓰이는지 묻는다. 빵 조각은 토스트가 되거나, 샌드위치의 첫 번째 조각이 되거나, 샌드위치의 두 번째 조각이 될 수 있다. 이 중 정확히 하나가 일어나며, 바로 앞 조각이 어떤 경우가 가능한지를 결정한다. 조각 $r-1$이 간식 하나를 끝내면 조각 r은 새 간식을 시작하고, 그렇지 않으면 조각 r은 샌드위치의 나머지 절반이어야 한다. 이로부터 1계 점화식(first-order recurrence)을 얻으며, 이것은 고정점(fixed point)을 빼는 통상적인 방법으로 푼다.

풀이

설정하기

빵 조각에 사용되는 순서대로 $1, 2, \ldots, n$의 번호를 붙인다. 각 조각 r에 대해 다음 중 정확히 하나가 일어난다.

- T_r: 조각 r이 토스트로 쓰인다.
- A_r: 조각 r이 샌드위치를 이루는 두 조각 중 첫 번째이다.
- S_r: 조각 r이 샌드위치를 이루는 두 조각 중 두 번째이다.

따라서 $t_r = \mathrm{P}(T_r)$, $s_r = \mathrm{P}(S_r)$이고 $\mathrm{P}(A_r) = 1 - s_r - t_r$이다. T_r 또는 S_r이 일어날 때 조각 r이 **간식을 끝낸다**고 하자. 이 확률은 $s_r + t_r$이다.

첫 번째 조각은 첫 간식을 시작하므로 샌드위치의 두 번째 조각이 될 수 없다.

$$s_1 = 0.$$

(또한 $n \geq 2$이면 $t_1 = p$이고, $n = 1$이면 $t_1 = 1$이다.)

두 방정식

t_r에 대한 방정식. 조각 r이 새 간식을 시작하는 것은 정확히 조각 $r-1$이 간식을 끝낼 때이며, 그 확률은 $s_{r-1} + t_{r-1}$이다. 조각 r이 간식을 시작한다는 조건 아래에서 조각 $r, r+1, \ldots, n$이 남아 있고, $r \leq n-1$이면 이는 적어도 두 조각이므로, 앞에서 무슨 일이 있었든 나는 확률 p로 토스트를 고른다. 조각 $r-1$이 간식을 끝내지 않으면 조각 r은 샌드위치를 완성하므로 토스트가 아니다. 따라서

$$t_r = (s_{r-1} + t_{r-1})\, p \qquad (2 \leq r \leq n-1).$$

$r = n$일 때는 이 식이 성립하지 않는데, 마지막 조각에서 시작하는 간식은 반드시 토스트이기 때문이다.

s_r에 대한 방정식. 조각 r이 샌드위치의 두 번째 조각인 것은 정확히 조각 $r-1$이 샌드위치의 첫 번째 조각일 때, 즉 조각 $r-1$이 간식을 끝내지 **않을** 때이다. 따라서

$$s_r = 1 - (s_{r-1} + t_{r-1}) \qquad (2 \leq r \leq n).$$

여기서는 남은 조각의 개수에 대한 조건이 필요 없으므로, 이 식은 $r = n$까지 그대로 성립한다.

s_r에 대한 점화식

두 번째 방정식에서 r을 $r+1$로 바꾸어 적용하면 $s_r + t_r = 1 - s_{r+1}$을 얻으므로, 첫 번째 방정식은 다음과 같이 쓸 수 있다.

$$t_r = p\,(1 - s_r) \qquad (2 \leq r \leq n-1).$$

$t_1 = p$이고 $s_1 = 0$이므로, 이 식은 $n \geq 2$일 때 $r = 1$에서도 성립한다. $1 \leq r-1 \leq n-1$일 때 성립하는 $t_{r-1} = p(1 - s_{r-1})$을 s_r에 대한 방정식에 대입하면

$$\begin{aligned} s_r &= 1 - s_{r-1} - p(1 - s_{r-1}) \\ &= (1-p)(1 - s_{r-1}) \\ &= q\,(1 - s_{r-1}), \end{aligned}$$

이 $2 \leq r \leq n$에 대해, 특히 요구된 대로 $2 \leq r \leq n-1$에 대해 성립한다.

점화식 풀기

생각해 보기 $s_r = q - qs_{r-1}$은 어떻게 풀까? 어떤 상수 수열이 이 식을 만족할까?

상수해 $s_r = c$가 되려면 $c = q - qc$이어야 하므로 $c = \dfrac{q}{1+q}$이다. 점화식에서 $c = q - qc$를 빼면

$$s_r - \frac{q}{1+q} = -q\left(s_{r-1} - \frac{q}{1+q}\right),$$

이므로 $s_r - \frac{q}{1+q}$는 한 단계마다 $-q$배가 된다. $s_1 - \frac{q}{1+q} = -\frac{q}{1+q}$에서 출발하면

$$s_r - \frac{q}{1+q} = (-q)^{r-1}\left(-\frac{q}{1+q}\right) = \frac{(-q)^r}{1+q},$$

이고, 따라서

$$s_r = \frac{q + (-q)^r}{1+q} \qquad (1 \le r \le n-1).$$

(점화식이 $r = n$까지 성립하므로, 실제로는 같은 공식이 $r = n$에서도 성립한다.)

t_r에 대해서는 $t_r = p(1 - s_r)$과

$$1 - s_r = \frac{1 + q - q - (-q)^r}{1+q} = \frac{1 - (-q)^r}{1+q},$$

을 이용하면

$$t_r = \frac{p\left(1 - (-q)^r\right)}{1+q} \qquad (1 \le r \le n-1).$$

마지막 조각

앞에서 언급했듯이 s_r에 대한 방정식과 점화식 $s_r = q(1 - s_{r-1})$은 $r = n$에서도 성립하므로, 공식이 그대로 이어진다.

$$s_n = \frac{q + (-q)^n}{1+q}.$$

마지막 조각 뒤에는 아무것도 없으므로 마지막 조각은 샌드위치의 첫 번째 조각이 될 수 없다. 따라서 마지막 조각은 토스트이거나 샌드위치의 두 번째 조각이며, $t_n = 1 - s_n$이다.

$$t_n = \frac{1 - (-q)^n}{1+q}.$$

$r < n$일 때의 t_r과 비교하면 인수 p가 사라졌다. 마지막 조각에서 간식이 시작되면 토스트가 확실하기 때문이다. (이 식들은 $n = 1$일 때도 성립하며, $s_1 = 0$과 $t_1 = 1$을 준다.)

검산

- $n = 3$, $p = q = \frac{1}{2}$로 두자. 가능한 순서는 토스트–토스트–토스트(확률 $\frac{1}{4}$), 토스트 다음 샌드위치(확률 $\frac{1}{4}$), 샌드위치 다음 토스트(확률 $\frac{1}{2}$)이다. 따라서 $s_3 = \frac{1}{4}$이고 $t_3 = \frac{3}{4}$이다. 공식은 $s_3 = \dfrac{\frac{1}{2} - \frac{1}{8}}{\frac{3}{2}} = \frac{1}{4}$, $t_3 = \dfrac{1 + \frac{1}{8}}{\frac{3}{2}} = \frac{3}{4}$를 준다.
- $q = 0$(항상 토스트)이면 모든 r에 대해 $s_r = 0$, $t_r = 1$을 얻으며, 마땅히 그래야 한다.
- r이 커지면 $s_r \to \frac{q}{1+q}$이다. 이는 타당하다. 장기적으로 간식 중 q의 비율이 샌드위치이고 각 샌드위치에는 "두 번째 조각"이 하나씩 있으며, 간식 하나는 평균 $1 + q$개의 조각을 쓰므로, 조각 중 $\frac{q}{1+q}$의 비율이 두 번째 조각이다.
- $n \leq 12$와 여러 p 값에 대해 모든 간식 순서를 정확히 열거하여 컴퓨터로 모든 공식을 확인하였다.

흔한 실수

- $t_n = p(s_{n-1} + t_{n-1})$을 사용하는 것: 마지막 조각에서 간식이 시작되면 그것은 확률 p가 아니라 확률 1로 토스트이다.
- 사건 "조각 $r - 1$이 토스트이다"와 "조각 r이 토스트이다"가 독립(independent)이라고 생각하는 것. 그렇지 않다. 중요한 것은 조각 r이 간식을 시작하는지이며, 이는 이전의 이력에 의존한다.
- 점화식을 풀 때 부호를 놓치는 것: 공비는 $-q$이므로 답은 $\frac{q}{1+q}$를 중심으로 번갈아 진동한다.

STEP 너머

조각 r이 새 간식을 시작하는지 여부는 단순한 **재생 과정**(renewal process)을 이룬다. 간식의 길이는 서로 독립이고 각각 1 또는 2이며, 극한 $\frac{q}{1+q}$는 재생 정리(renewal theorem)의 한 예이다(장기적인 재생 비율은 평균 간격의 역수이다). 동등하게, 각 조각의 유형은 마르코프 연쇄(Markov chain)를 이루며, $(-q)^r$은 이 연쇄가 시작 상태를 얼마나 빨리 잊는지를 보여 준다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 마르코프 연쇄와 이런 종류의 점화식을 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.