풀이

# STEP II 2025, 문제 11

확률과 통계 • 기댓값과 이항 합

> **STEP II 2025, 11번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 OCR에서 내려받을 수 있는 2025년 STEP 시험지 묶음(zip 파일) 가운데 STEP II 시험지에 있다:
>
> www.ocr.org.uk/Images/750486-2025-past-papers.zip

## 핵심 아이디어

두 게임 모두 처음으로 $2k$가 나오면 끝나므로, 각각은 주사위를 던진 횟수 $r$에 대한 합으로 나타난다. 알리(Ali)의 경우 그 합은 등비급수(geometric series)와 비슷한 꼴이며, 소문항 (i)이 바로 던진 횟수의 기댓값(expected value)을 구하는 도구이다. 젠(Zen)의 경우 이기는 게임은 $2k$ 이전에 어떤 수들이 나왔는지에 따라 결정된다. 그 집합을 알면 증가하는 순서는 저절로 정해지기 때문이다. 이러한 집합을 세면 이항계수(binomial coefficients)가 등장하고, 따라서 그 합은 소문항 (iii)의 합이 된다.

## 풀이

### 소문항 (i)

$|x| < 1$일 때 등비급수에 의해

$$\sum_{r=0}^{\infty} x^r = \frac{1}{1-x}$$

이다. 멱급수(power series)는 수렴 구간(interval of convergence) 안에서 항별로 미분할 수 있으므로, 양변을 $x$에 대하여 미분하면

$$\sum_{r=1}^{\infty} r x^{r-1} = \frac{1}{(1-x)^2}, \qquad |x| < 1,$$

을 얻고, 이것이 증명하려던 것이다.

**소문항 (ii)**

한 번 던질 때 결과는 세 종류이다.

$$\mathrm{P}(\text{눈}2k) = \frac{1}{2k}, \qquad \mathrm{P}(\text{짝수, 제외: } 2k) = \frac{k-1}{2k}, \qquad \mathrm{P}(\text{홀수}) = \frac{1}{2}.$$

던진 횟수를 $R$라 하자. 알리가 $R = r$로 이기는 것은 정확히 $1, \ldots, r-1$번째 결과가 $2k$가 아닌 짝수이고 $r$번째 결과가 $2k$일 때이다. 각 시행은 독립(independent)이므로 $x = \dfrac{k-1}{2k}$라 하면

$$\mathrm{P}(\text{알리가 이기고} R = r) = \frac{1}{2k}\, x^{r-1}$$

이다.

**(a)** 등비급수를 더하면 ($0 \le x < 1$임에 주의한다)

$$\mathrm{P}(\text{알리가 이김}) = \frac{1}{2k} \cdot \frac{1}{1-x} = \frac{1}{2k} \cdot \frac{2k}{k+1} = \frac{1}{k+1}.$$

이를 빠르게 알아보는 방법이 있다. $2k$가 아닌 짝수는 결과에 영향을 주지 않으므로 무시한다. 게임은 홀수와 $2k$ 중 어느 것이 먼저 나오느냐로 결정된다. 홀수 면은 $k$개, $2k$가 적힌 면은 하나이고 모두 같은 확률로 나오므로, $2k$가 먼저 나올 확률은 $\frac{1}{k+1}$이다.

**(b)** 알리는 $R = r$로 이기면 £$r$를 받고 그렇지 않으면 아무것도 받지 못하므로, 소문항 (i)에 의해

$$\mathrm{E}(\text{알리의 수입}) = \sum_{r=1}^{\infty} r \cdot \frac{1}{2k}\, x^{r-1} = \frac{1}{2k} \cdot \frac{1}{(1-x)^2} = \frac{1}{2k}\left(\frac{2k}{k+1}\right)^2 = \frac{2k}{(k+1)^2} \text{ 파운드.}$$

**소문항 (iii)**

주어진 식을 $S$라 하자. 일반항은 $r = 0, 1, \ldots, n$에 대하여 $(r+1)\binom{n}{r}x^r$이다. 인수 $r+1$을 보면 $x^{r+1}$을 미분하는 것을 떠올릴 수 있다.

$$S = \sum_{r=0}^{n} (r+1)\binom{n}{r}x^r = \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{r=0}^{n}\binom{n}{r}x^{r+1}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(x(1+x)^n\Big),$$

여기서 이항정리(binomial theorem)를 사용하였다. 곱의 미분법(product rule)에 의해

$$S = (1+x)^n + nx(1+x)^{n-1} = (1+x)^{n-1}(1+(n+1)x).$$

**소문항 (iv)**

젠이 정확히 $r$번 던져서 이긴다고 하자. 마지막 결과는 $2k$이고, 처음 $r-1$번의 결과는 $\{1, 2, \ldots, 2k-1\}$에서 나온 순증가하는 수들이다. (게임은 처음 $2k$가 나올 때 멈추므로 이 수들은 어차피 모두 $2k$보다 작아야 한다.)

**생각해 보기** 길이가 $r$인 이기는 수열은 몇 개인가?

순증가 수열은 그 안에 어떤 수들이 나타나는지만 알면 결정된다. 서로 다른 수들의 집합을 증가하는 순서로 늘어놓는 방법은 하나뿐이기 때문이다. 따라서 길이가 $r$인 이기는 수열은 $\{1,\dots,2k-1\}$의 원소 $(r-1)$개짜리 부분집합(subset)과 일대일로 대응한다. 그러한 부분집합은 $\binom{2k-1}{r-1}$개이고, $1 \le r \le 2k$이다.

$r$번 던진 결과의 특정한 수열 하나의 확률은 $\left(\frac{1}{2k}\right)^r$이다. 따라서

$$\mathrm{E}(\text{젠의 수입}) = \sum_{r=1}^{2k} r\binom{2k-1}{r-1}\left(\frac{1}{2k}\right)^r = \frac{1}{2k}\sum_{j=0}^{2k-1}(j+1)\binom{2k-1}{j}\left(\frac{1}{2k}\right)^j,$$

이고, 여기서 $j = r-1$이다. 이것은 소문항 (iii)의 합에서 $n = 2k-1$, $x = \frac{1}{2k}$로 둔 것이며, 이때 $1+(n+1)x = 1+2k\cdot\frac{1}{2k} = 2$이다. 그러므로

$$\mathrm{E}(\text{젠의 수입}) = \frac{1}{2k}\left(1+\frac{1}{2k}\right)^{2k-2}\cdot 2 = \frac{1}{k}\left(1+\frac{1}{2k}\right)^{2k-2} \text{ 파운드.}$$

**소문항 (v)**

젠의 수입의 기댓값을 알리의 것으로 나누면

$$\rho = \frac{1}{k}\left(1+\frac{1}{2k}\right)^{2k-2}\cdot\frac{(k+1)^2}{2k} = \frac{(k+1)^2}{2k^2}\cdot\frac{\left(1+\frac{1}{2k}\right)^{2k}}{\left(1+\frac{1}{2k}\right)^2}.$$

이제 $k$가 크다고 하자. 그러면

- $\dfrac{(k+1)^2}{k^2} = \left(1+\dfrac{1}{k}\right)^2 \approx 1$;
- $\left(1+\dfrac{1}{2k}\right)^2 \approx 1$;
- 주어진 근사식에서 $n = 2k$로 두면 $\left(1+\dfrac{1}{2k}\right)^{2k} \approx \mathrm{e}$이다.

따라서

$$\rho \approx \frac{\mathrm{e}}{2} \approx \frac{2.71828}{2} = 1.359\dots,$$

이고, $k$가 클 때 젠의 수입의 기댓값은 알리의 약 1.359배, 즉 35%보다 조금 더 많으며, 이것이 보이려던 것이다.

## 검산

- $k = 1$이면 주사위는 면이 1, 2인 동전이다. 알리는 첫 결과가 2일 때만 이기며, 그 확률은 $\frac{1}{2} = \frac{1}{k+1}$이고 이때 £1을 받으므로 수입의 기댓값은 $\frac{1}{2} = \frac{2k}{(k+1)^2}$이다. 젠은 수열 2 (확률 $\frac{1}{2}$, 수입 £1)와 수열 1, 2 (확률 $\frac{1}{4}$, 수입 £2)로 이기므로 수입의 기댓값은 $\frac{1}{2} + \frac{1}{2} = 1$이다. 공식도 일치한다. $k = 1$이면 지수 $2k - 2$가 0이므로 공식은 $\frac{1}{1} \cdot 1 = 1$을 준다.
- $k = 3$이면 알리의 수입의 기댓값은 $\frac{3}{8}$이고 젠의 것은 $\frac{1}{3}\left(\frac{7}{6}\right)^4 = \frac{2401}{3888} \approx 0.618$이다. 두 값 모두 이기는 수열을 직접 모두 나열한 결과, 그리고 게임의 컴퓨터 시뮬레이션 결과와 일치한다.
- 비 $\rho$는 $k = 10$에서 약 1.456, $k = 100$에서 약 1.369, $k = 10^4$에서 약 1.3592이다. 이 값은 $\mathrm{e}/2 = 1.35914\ldots$를 향해 감소한다.

## 흔한 실수

- 소문항 (iv)에서 부분집합 대신 계승을 써서 순서 있는 수열을 세는 것. 각 집합을 증가하도록 배열하는 방법은 하나뿐이다.
- 젠의 게임에서 앞선 결과들에 $2k$가 포함될 수 없다는 것을 잊는 것. 이 때문에 범위는 $\{1, \ldots, 2k-1\}$이 되고, 따라서 $n$은 $2k$가 아니라 $n = 2k - 1$이다.
- 소문항 (ii)에서 게임이 계속될 확률로 $\frac{k-1}{2k}$ 대신 $\frac{k}{2k}$를 쓰는 것. $2k$도 짝수이지만 이 결과가 나오면 게임이 끝난다.

## STEP 너머

소문항 (iii)은 **생성함수**(generating functions)의 작은 예이다. 수열을 멱급수의 계수로 부호화한 다음 그 급수에 미적분을 적용하는 것이다. 확률생성함수(probability generating functions)는 이를 한 걸음 더 나아가게 한다. $G(s) = \mathrm{E}(s^X)$이면 $\mathrm{E}(X) = G'(1)$이며, 이것이 바로 여기서 두 번 사용한 미분 기법이다. 극한 $\left(1 + \frac{1}{n}\right)^n \to \mathrm{e}$는 확률론 곳곳에서 등장하는데, 예를 들어 이항분포의 푸아송 근사(Poisson approximation)에서 나타난다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)에는 생성함수에 관한 좋은 장이 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.