12.11 우표, 컵, 그리고 e

학습지

문제

STEP I 1996, 13번 문제

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

step.maths.org/questions/96-s1-q13

힌트

먼저 혼자 힘으로 문제를 풀어 보자. 막히면 힌트를 **하나씩** 읽고, 힌트를 하나 읽을 때마다 다시 문제로 돌아가자. **잠깐, 직접 해 보기**라고 표시된 힌트가 핵심 단계이다. 각 힌트 뒤에는 풀이를 적을 공간이 있다.

? **생각해 보기** $N = 4$가 되려면 무슨 일이 일어나야 하는가? 그 확률은 얼마인가?

1단계 분포

다루는 내용: N의 분포(distribution)

? **생각해 보기** $N = k$이려면 Jim은 컵이 $2, 3, \ldots, k-1$개일 때는 맞게 추측하고 컵이 k개일 때는 틀리게 추측해야 한다. 이 확률들을 곱한다.

2단계 기댓값

다루는 내용: 기댓값(expectation)

잠깐, 직접 해 보기 $k \cdot \frac{k-1}{k!}$을 간단히 정리해 보자. 어떤 유명한 급수(series)가 나오는가?

3단계 분산

다루는 내용: 분산(variance)

> **생각해 보기** $k^2 = k(k-1) + k$로 써서 $\mathrm{E}(N^2)$을 구하거나, $\mathrm{E}\big(N(N-1)\big)$을 먼저 구한다.

숨은 수학

모습만 바꾼 같은 분포

난수 $U_1, U_2, \ldots$를 $[0,1]$에서 균등하게(uniformly) 뽑고, 그 합이 1을 넘는 데 필요한 개수를 M이라 하자. 처음 k개의 수를 더한 값이 1보다 작을 확률은 영역 $\{u_1 + \cdots + u_k < 1, u_i \geq 0\}$의 부피이며, 이 값은 $\frac{1}{k!}$이다(넓이가 $\frac{1}{2}$인 삼각형, 부피가 $\frac{1}{6}$인 사면체(tetrahedron) 등). 따라서

$$\mathrm{P}(M > k) = \frac{1}{k!}, \qquad \mathrm{P}(M = k) = \frac{1}{(k-1)!} - \frac{1}{k!} = \frac{k-1}{k!},$$

이고, 이는 N의 분포와 정확히 같다. 두 문제 모두에서 $\mathrm{P}(N > k) = \frac{1}{k!}$이다. 컵 문제에서 이것은 $k-1$번 연속으로 맞힐 확률이다. 공식 $\mathrm{E}(N) = \sum_{k \geq 0} \mathrm{P}(N > k)$(Section 9.9에서 나온 공식)를 쓰면 한 줄 만에 $\mathrm{E}(N) = \sum \frac{1}{k!} = e$를 얻는다.

e가 자꾸 나타나는 이유

k개의 대상이 특정한 순서로 놓이는 경우이든, 확률이 $\frac{1}{2}, \frac{1}{3}, \ldots$인 사건 k개가 연달아 일어나는 경우이든, 단체(simplex)의 부피이든, 확률에 $\frac{1}{k!}$이 들어 있으면 k에 대한 합은 지수급수가 된다. 교란순열(derangement, Section 9.8)에서는 $\sum \frac{(-1)^k}{k!} = e^{-1}$이 나왔고, 여기서는 $\sum \frac{1}{k!} = e$가 나왔다. 그리고 일어날 가능성이 낮은 독립 사건이 많이 있을 때 생기는 $\left(1 - \frac{1}{n}\right)^n \to e^{-1}$ 꼴의 극한은 비서 문제와 포아송 분포에 e가 나타나는 이유를 설명해 준다.

몬테카를로 방법으로 구하는 e

"합이 1을 넘는" 문제를 이용하면 실험으로 e를 추정할 수 있다. 난수를 만들어 1을 넘는 데 몇 개가 필요한지 세고, 이를 여러 번 반복하여 평균을 내면 된다. 이는 π를 추정하는 뷔퐁(Buffon)의 바늘(Section 12.20)과 사촌 격이다. 카지노의 이름을 딴 이러한 **몬테카를로**(Monte Carlo) 방법은 직접 계산하기 어려운 양을 무작위 표본추출로 추정한다. 오늘날 이 방법은 입자물리학에서 금융에 이르기까지 과학 전반에서 쓰이며, 주로 고차원 적분에 사용된다.

더 도전하기

1 **생성함수.** $\mathrm{E}(t^N) = 1 + (t-1)e^t$임을 보이고, 이를 이용하여 $\mathrm{E}(N)$과 $\mathrm{E}(N(N-1))$을 구하여라.

2 **꼬리 공식.** 음이 아닌 정수 값을 갖는 확률변수(random variable)에 대하여 $\mathrm{E}(N) = \sum_{k\geq 0} \mathrm{P}(N > k)$임을 증명하고, 이를 이용하여 한 줄 만에 $\mathrm{E}(N)$을 구하여라.

3 **컵 하나로 시작하기.** 게임이 컵 하나(짐이 틀릴 수 없는 경우)로 시작한다면 분포는 어떻게 달라지는가? 가격의 기댓값은 얼마인가?

4 **단체.** $\{u_1 + u_2 + u_3 < 1, u_i \geq 0\}$의 부피가 $\frac{1}{6}$임을 보이고, 수학적 귀납법(induction)을 이용하여 k차원에서는 그 부피가 $\frac{1}{k!}$임을 보여라.

더 읽을거리

- **Eli Maor, *e: The Story of a Number*** (Princeton University Press). 확률에서의 등장을 포함한 e의 역사를 다룬다.
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction*** (Oxford University Press). 생성함수를 다룬다.
- **Paul Nahin, *Digital Dice*** (Princeton University Press). 몬테카를로 시뮬레이션으로 푸는 확률 퍼즐 모음이다.

안내에 따른 전체 풀이는 제 12.11절에 있다:

https://beyondstep.keithmcnulty.org/ko/ch12/stamps-and-e.html

매주 STEP 문제 하나를 끝까지 풀어 본다. The STEP Up은 이메일로 무료로 받아 볼 수 있다: keithmcnulty.substack.com