풀이

STEP III 2005, 문제 13

확률과 통계 • 기하분포

> **STEP III 2005, 13번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/05-s3-q13

핵심 아이디어

$w+1$부터 n까지의 카드는 아무 역할도 하지 않는다. 돈을 주지도 않고 게임을 끝내지도 않는다. 따라서 (i)에서는 이 카드들을 그냥 무시하고 $0, 1, \ldots, w$가 나오는 뽑기만 지켜보면 된다. 그런 뽑기 하나하나는 동등하게 가능한 $w+1$개의 수 가운데 하나를 공정하게 고르는 것이고, 그러면 총상금은 기하 확률변수(geometric random variable)가 된다. 뽑기들이 더 이상 독립(independent)이 아닌 (ii)에서는 총상금을 돈을 주는 카드마다 하나씩 대응하는 지시변수(indicator)들의 합으로 쪼개고, 기댓값의 선형성(linearity of expectation)을 이용한다.

풀이

뽑은 카드가 $0, 1, \ldots, w$ 중 하나이면 그 뽑기를 **유효한** 뽑기, $w+1, \ldots, n$ 중 하나이면 **무관한** 뽑기라고 부르자.

소문항 (i)

두 사건의 확률이 같은 이유. 무관한 뽑기는 아무것도 바꾸지 않는다. 참가자는 아무것도 받지 않고 게임은 계속된다. 따라서 참가자의 총상금은 유효한 뽑기들의 순서에만 의존한다. 0이 아닌 유효한 뽑기마다 £1을 받고, 0이 나오는 첫 번째 유효한 뽑기에서 게임이 끝난다. 그러므로 참가자가 정확히 £3을 받는 것은 유효한 뽑기들 가운데 처음 세 번은 0이 아닌 수가 나오고 네 번째에 0이 나오는 경우와 정확히 일치한다. 두 사건은 같은 사건이므로 확률도 같다.

(각 뽑기는 서로 독립적으로 확률 $\frac{w+1}{n+1} > 0$로 유효하므로, 확률 1로 유효한 뽑기가 무한히 많이 일어나고, 특히 첫 번째, 두 번째, 세 번째, 네 번째 유효한 뽑기가 존재한다. 게임은 확

률 1로 끝난다.)

> **생각해 보기** 앞에서 무엇이 나왔든 관계없이, 각 유효한 뽑기에서 $0, 1, \ldots, w$ 중 어느 것이 나올 확률이 모두 같은 이유는 무엇인가?

매번 카드를 되돌려 놓으므로(복원 추출, replacement) 뽑기들은 서로 독립이고, 각각 $\{0, 1, \ldots, n\}$ 위의 균등분포(uniform distribution)를 따른다. 어떤 뽑기가 유효하다는 조건 아래에서 그 뽑기는 $\{0, 1, \ldots, w\}$ 위에서 균등하다. 이 $w+1$개의 결과는 처음부터 동등하게 가능했기 때문이다. 뽑기들이 독립이므로 연속된 유효한 뽑기들의 값도 서로 독립이다. 따라서 각 유효한 뽑기는 다른 뽑기와 독립적으로 확률 $\frac{1}{w+1}$로 0이고 확률 $\frac{w}{w+1}$로 0이 아니다.

그러므로

$$P(\text{상금이 정확히 £3}) = \left(\frac{w}{w+1}\right)^3 \cdot \frac{1}{w+1} = \frac{w^3}{(w+1)^4},$$

이고, 이것이 보이려던 것이다.

일반적인 r. 3 대신 r을 넣고 똑같이 논증하면(처음 r번의 유효한 뽑기는 0이 아니고, $(r+1)$번째는 0)

$$P(\text{상금이 정확히 £}r) = \frac{w^r}{(w+1)^{r+1}}, \qquad r = 0, 1, 2, \ldots$$

를 얻는다. $q = \frac{w}{w+1}$로 놓으면 이것은 $(1-q)q^r$, 즉 기하분포이다. 이 확률들의 합은 당연히 1이다.

총상금의 기댓값. 총상금(파운드 단위)을 X라 하자. $|q| < 1$일 때 $\sum_{r\geq 0} q^r = \frac{1}{1-q}$를 미분하고 q를 곱하면 $\sum_{r\geq 0} rq^r = \frac{q}{(1-q)^2}$를 얻는다. 따라서

$$E(X) = \sum_{r=0}^{\infty} r(1-q)q^r = (1-q) \cdot \frac{q}{(1-q)^2} = \frac{q}{1-q} = \frac{w/(w+1)}{1/(w+1)} = w.$$

총상금의 기댓값은 £w이다.

소문항 (ii)

이제 카드를 되돌려 놓지 않으므로, 이 게임은 카드 묶음을 섞어 균등하게 무작위인 순서로 만든 다음 0이 나올 때까지 카드를 한 장씩 뒤집는 것과 같다. 참가자는 이 순서에서 카드 0보다 앞에 오는 카드 $1, \ldots, w$ 각각에 대해 £1을 받는다(0보다 앞의 카드는 모두 뽑히고, 뒤의 카드는 하나도 뽑히지 않는다).

$k = 1, \ldots, w$에 대해 카드 k가 카드 0보다 앞에 오면 $I_k = 1$, 그렇지 않으면 $I_k = 0$이라 하자. 총상금은 $Y = I_1 + I_2 + \cdots + I_w$이다.

생각해 보기 I_k들은 독립이 아니다. 그것이 문제가 되는가?

그렇지 않다. 합의 기댓값은 종속 관계와 상관없이 기댓값들의 합이다. 따라서 $P(I_k = 1)$만 구하면 된다. 균등하게 무작위인 순서에서 카드 k와 0의 위치를 맞바꾸는 것은 k가 0보다 앞에 오는 순서들과 0이 k보다 앞에 오는 순서들 사이의 전단사(bijection)이므로, 두 경우는 동등하게 가능하고 $P(I_k = 1) = \frac{1}{2}$이다. 그러므로

$$E(Y) = \sum_{k=1}^{w} E(I_k) = \sum_{k=1}^{w} \frac{1}{2} = \frac{w}{2},$$

이고, 이는 (i)의 기대 상금 w의 절반이므로 보이려던 것이 성립한다.

검산

- $w = n$(0이 아닌 모든 카드가 돈을 줌)이고 비복원일 때, 카드 0의 위치는 $1, \ldots, n+1$ 위에서 균등하므로 그 앞에 있는 카드 수의 평균은 $\frac{n}{2}$이며, $\frac{w}{2}$와 일치한다.
- 복원이고 $w = 1$일 때, 각 유효한 뽑기는 0과 1 사이의 공정한 동전 던지기이므로 $P(X = r) = 2^{-(r+1)}$이고 $E(X) = 1$이며, $w^r/(w+1)^{r+1}$ 및 $E(X) = w$와 일치한다.
- $n = 5$, $w = 2$로 2×10^5번의 게임을 시뮬레이션한 결과, $\frac{8}{81} \approx 0.0988$에 대해 $P(X = 3) \approx 0.0995$, 2에 대해 평균 상금 1.995를 얻었다. 작은 카드 묶음(예를 들어 $n = 5$, $w = 3$)에 대해 모든 순서를 정확히 열거하면 비복원일 때 평균은 $\frac{3}{2} = \frac{w}{2}$이다.

흔한 실수

- 한 번의 뽑기가 0일 확률로 $\frac{1}{n+1}$을 쓰는 것. 그것은 뽑기 한 번당 확률이지만, 그 사이에 무관한 뽑기가 몇 번이든 일어날 수 있다. 유효한 뽑기에 대해 조건부로 생각해야 깔끔한 $\frac{1}{w+1}$이 나온다.
- $r = 0$을 빠뜨리는 것. 첫 번째 유효한 뽑기가 0일 수 있으며, 그러면 아무것도 받지 못할 확률은 $\frac{1}{w+1}$이다.
- (ii)에서 상금의 분포 전체를 구하려 하는 것. 그럴 필요가 없다. 지시변수와 선형성이면 두 줄로 해결된다.

STEP 너머

무관한 뽑기를 무시하는 것은 **솎아내기**(thinning), 즉 흥미로운 일이 일어나는 시점에서만 과정을 살펴보는 방법의 한 예이다. 이 생각은 마르코프 연쇄(Markov chains)("도약 연

쇄(jump chain)”)와 푸아송 과정(Poisson processes)에서 다시 등장한다. 개수를 지시변수들의 합으로 쓰고, 지시변수들이 종속이더라도 기댓값의 선형성을 이용하는 것은 대학 1학년 확률론과 조합론에서 가장 유용한 기법 가운데 하나이다. 두 아이디어 모두 Blitzstein과 Hwang의 *Introduction to Probability*(확률론 입문)와 Grimmett과 Welsh의 *Probability: An Introduction*(확률: 입문)에서 다루고 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.