풀이

STEP II 2004, 문제 13

확률과 통계 • 최적 정지

> **STEP II 2004, 13번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/04-s2-q13

핵심 아이디어

한 번에 한 번씩 뽑는 것을 생각한다. 그녀가 흰 공 n개를 무사히 모았다면, 한 번 더 뽑을 때 £1이 더해지거나 지금 가진 £n이 모두 사라지므로, 그 기대 효과는 쉽게 적을 수 있다. 목표를 공 N개로 할 때의 전체 기대 보상은 추가로 한 번 더 뽑는 것의 기대 이득(expected gain)이 양수일 때에만 정확히 증가하므로, 최선의 목표는 그 이득이 아직 음수가 아닌 마지막 목표이다.

풀이

처음부터 끝까지 $1 \le r < b$라고 가정한다. 따라서 빨간 공이 적어도 하나, 흰 공이 적어도 하나 있다. ($r = 0$이면 그녀는 그냥 모든 공을 꺼내면 된다.)

다음 뽑기의 기대 이득

처음 n번의 뽑기에서 모두 흰 공이 나왔다고 하자. 이제 주머니에는 공이 $b - n$개 있다. 빨간 공이 r개, 흰 공이 $b - n - r$개이며, 각 공이 다음에 나올 확률(probability)은 모두 같다(흰 공이 남아 있으려면 $n \le b - r - 1$이어야 한다. 어떤 경우든 $b - n \ge r \ge 1$이다).

- 확률 $\dfrac{b-n-r}{b-n}$로 흰 공이 나오고, 그녀의 보상은 £n에서 £$(n+1)$이 된다. 즉 이득은 $+1$이다.
- 확률 $\dfrac{r}{b-n}$로 빨간 공이 나오고, 그녀는 원래 가졌을 £n을 잃는다. 즉 이득은 $-n$이다.

따라서 다음 뽑기에서 얻는 그녀의 기대 이득은

$$G_n = \frac{(b-n-r)\cdot 1 + r\cdot(-n)}{b-n} = \frac{b-r-n(r+1)}{b-n}.$$

분모가 양수이므로 $G_n = 0$인 것은 정확히 $b-r = n(r+1)$일 때, 즉

$$n = \frac{b-r}{r+1},$$

일 때이며, 이는 문제에서 요구한 결과이다. 더 유용하게는, $n < \frac{b-r}{r+1}$이면 $G_n > 0$이고 $n > \frac{b-r}{r+1}$이면 $G_n < 0$이다.

최선의 공의 개수

처음 N번의 뽑기가 모두 흰 공일 확률을 P_N이라 하고, 그녀가 공 N개를 꺼내는 것을 목표로 할 때의 전체 기대 보상을 E_N이라 하자. 그 N개의 공이 모두 흰 공이면 그녀는 £N을 받고, 그렇지 않으면 아무것도 받지 못하므로

$$E_N = NP_N.$$

> ? **생각해 보기** 한 번의 뽑기에서의 기대 이득은 전체 기대 보상 E_N과 어떤 관계가 있는가?

N을 목표로 하는 것과 $N+1$을 목표로 하는 것을 비교하자. 두 계획은 처음 N개의 공이 모두 흰 공일 때에만 달라지며(그렇지 않으면 두 경우 모두 이미 진 것이다), 이 일은 확률 P_N로 일어난다. 그리고 그 경우 추가 뽑기는 그녀의 보상을 기댓값(expected value)이 G_N인 양만큼 바꾼다. 따라서

$$E_{N+1} - E_N = P_N G_N = P_N \frac{b-r-N(r+1)}{b-N}.$$

(이는 $P_{N+1} = P_N \cdot \frac{b-r-N}{b-N}$로부터 직접 확인할 수도 있다.)

$N \le b-r$이면 $P_N > 0$이므로, $E_{N+1} - E_N$의 부호는 $b-r-N(r+1)$의 부호와 같다. $k = N+1$이라 쓰면, 이는

$$(b+1) - k(r+1).$$

의 부호이다. $m = \frac{b+1}{r+1}$이라 두자. 그러면 $k < m$일 때 $E_k > E_{k-1}$, $k = m$일 때 $E_k = E_{k-1}$, $k > m$일 때 $E_k < E_{k-1}$이다. 따라서 수열(sequence) $E_0, E_1, E_2, \ldots$는 $E_{\lfloor m \rfloor}$까지 순증가하고(m이 정수이면 $E_{m-1} = E_m$으로 같은 값이 생긴다) 그다음에는 순감소한다. 적어도 $k-1 \le b-r$인 동안은 그렇다. 그 너머에서는 $P_N = 0$이고 $E_N = 0$이므로 더 나을 것이 없다.

$\lfloor m \rfloor$이 타당한 목표인지, 즉 $\lfloor m \rfloor \le b - r$인지 확인해야 한다. $b = r + 1$이면 $m = \frac{r+2}{r+1}$이므로 $\lfloor m \rfloor = 1 = b - r$이다. $b \ge r + 2$이면

$$(b - r)(r + 1) - (b + 1) = r(b - r - 1) - 1 \ge r - 1 \ge 0,$$

이므로 $m \le b - r$이다. 어느 경우든 E_N의 최댓값은

$$n = \left\lfloor \frac{b+1}{r+1} \right\rfloor,$$

즉 $\frac{b+1}{r+1}$의 정수 부분(integer part)에서 얻어지며, 이는 문제에서 요구한 결과이다.

이 목표는 어떤 정지 규칙(stopping rule)을 쓰더라도 그녀가 얻을 수 있는 최선이기도 하다. 계속할지 정하기 전에 그녀가 알게 되는 것은 지금까지의 모든 공이 흰 공이었다는 사실뿐이므로(그렇지 않다면 게임은 끝난 것이다), 모든 규칙은 결국 수 N을 미리 정하는 것과 같다.

$r = 1$인 경우

빨간 공이 하나일 때, 처음 N번의 뽑기에서 빨간 공을 피하는 것은 정확히 빨간 공이 마지막 $b - N$개의 자리 중 하나에 있을 때이며, 모든 자리는 똑같이 가능하다. 따라서

$$P_N = \frac{b - N}{b}, \qquad E_N = \frac{N(b - N)}{b}.$$

최선의 목표는 $n = \lfloor \frac{b+1}{2} \rfloor$이다.

b가 짝수인 경우. 이때 $\frac{b+1}{2} = \frac{b}{2} + \frac{1}{2}$이므로 $n = \frac{b}{2}$이고

$$E_n = \frac{\frac{b}{2} \cdot \frac{b}{2}}{b} = \frac{b}{4}.$$

그녀의 전체 기대 보상은 £$\frac{1}{4}b$이며, 이는 문제에서 요구한 결과이다.

b가 홀수인 경우. 이때 $n = \frac{b+1}{2}$이므로 $b - n = \frac{b-1}{2}$이고

$$E_n = \frac{\frac{b+1}{2} \cdot \frac{b-1}{2}}{b} = \frac{b^2 - 1}{4b}.$$

그녀의 전체 기대 보상은 £$\frac{b^2-1}{4b}$이다. (여기서 $m = \frac{b+1}{2}$는 정수이므로, 대칭인 식 $N(b - N)/b$가 보여 주듯 공 $\frac{b-1}{2}$개를 목표로 해도 같은 값을 얻는다.)

검산

- $r = 1$일 때 $E_N = N(b - N)/b$는 N에 대해 아래로 볼록한(위로 열린 것이 아닌, 아래로 향한) 포물선(parabola)이고 꼭짓점(vertex)은 $N = b/2$이다. 이는 두 답과 모두 맞는다. b가 짝수이면 꼭짓점에서 정확히 $b/4$이고, b가 홀수이면 그보다 조금 작은 $\frac{b}{4} - \frac{1}{4b}$이다.

- $b = 7, r = 2$로 두자. 공식은 $n = \lfloor 8/3 \rfloor = 2$를 주고, 직접 계산하면 $E_0, \ldots, E_5 = 0, \frac{5}{7}, \frac{20}{21}, \frac{6}{7}, \frac{4}{7}, \frac{5}{21}$로 $N = 2$에서 최대가 된다. 공을 늘어놓는 7!가지 순서를 모두 나열하여 이를 확인했고, 모든 $b \leq 24$에 대해 n의 공식을 완전한 최적 정지 계산과 비교하여 확인했다.
- $b = 9, r = 1$이면 답은 $\frac{80}{36} = \frac{20}{9} \approx 2.22$로, $\frac{9}{4}$보다 조금 작다.

흔한 실수

- 빨간 공이 나오면 이미 모은 £n을 잃는다는 것을 잊고, 기대 이득을 그냥 $\frac{b-n-r}{b-n}$로 쓰는 것.
- "이미 뽑은 공 n개"와 "목표인 공 n개"를 혼동하는 것. 이득 G_n은 $n = \frac{b-r}{r+1}$에서 0이 되며, 그래서 최선의 **목표**는 그보다 하나 많은 $\frac{b-r}{r+1} + 1 = \frac{b+1}{r+1}$이다.
- $\frac{b+1}{r+1}$이 정수일 때 값이 같아지는 경우를 놓치는 것. 이때는 n과 $n-1$이 모두 최적이다.

STEP 너머

이것은 작은 **최적 정지**(*optimal stopping*) 문제이며, 위의 논증은 "한 단계 앞보기(one-step look-ahead)" 규칙이다. 즉 한 단계를 더 가는 것의 기대 이득이 음수가 되는 순간 멈춘다. 이 규칙은 문제가 **단조**(*monotone*)일 때, 즉 계속하는 것이 일단 불리해 보이면 그 뒤로도 계속 불리할 때 언제나 최적이다. 여기서는 G_n이 n에 대해 감소하므로 그러하다. 같은 아이디어는 대학 1, 2학년 확률론에서 비서 문제(secretary problem)와 마팅게일(martingale) 방법으로 이어진다. Grimmett and Welsh의 *Probability: An Introduction*(확률론 입문)은 그 바탕이 되는 조건부 기댓값(conditional expectation) 도구를 익히기에 좋은 책이다.

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.