풀이

# STEP I 1991, 문제 16

확률과 통계 • 조건부 확률

**STEP I 1991, 16번 문제**

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

step.maths.org/questions/91-s1-q16

## 핵심 아이디어

모든 것을 각 파충류가 처음 길을 살핀 뒤 $n = 0, 1, 2, \ldots$ 분에 길을 확인하는 시점(inspection time)에 따라 정리한다. 두꺼비의 대기 시간은 기하분포(geometric distribution)를 따르므로 두꺼비는 반드시 길을 건넌다. 개구리는 다르다. 개구리의 규칙에 따르면 늦어도 $n = 3$에서는 결정을 내려야 하므로, 개구리의 이야기는 네 단계뿐이고 각 단계에서 개구리가 무사히 건널 확률을 나열할 수 있다. 이 목록만 얻으면 두꺼비와의 비교와 조건부확률(conditional probability)은 몇 번의 합 계산에 지나지 않는다.

## 풀이

풀이 전체에서 서로 다른 확인 시점의 길 상태는 독립(independent)이고 각각 확률 0.1로 안전하며, 두꺼비와 개구리는 서로 독립적으로 행동한다고 가정한다. 개구리는 "나중에", 즉 다른 시각에 건너므로 이는 합리적인 가정이다. 건너는 데 걸리는 시간은 무시할 수 있으므로, 확인 시점 $n$에 건너는 파충류는 $n$분에 도착한다.

### 두꺼비

두꺼비는 안전할 때만 건너므로 결코 차에 치이지 않는다. 확인 시점 $0, 1, \ldots, k$가 지난 뒤에도 여전히 기다리고 있을 확률은 $0.9^{k+1}$이고, 이는 0으로 수렴한다. 따라서 두꺼비가 결국 안전하게 건널 확률은

$$\sum_{k=0}^{\infty} 0.9^k(0.1) = \frac{0.1}{1 - 0.9} = 1.$$

두꺼비가 건너는 시각을 $T$라 하자. 그러면 $P(T = k) = 0.9^k(0.1)$이고, 뒤에서 쓰기 위해 적어 두면 $P(T > k) = 0.9^{k+1}$이다.

### 개구리

확인 시점 $n$에 개구리가 아직 기다리고 있다면,

- 확률 0.1로 길이 안전하고 개구리는 안전하게 건넌다.
- 확률 $0.9 \cdot \frac{n}{3}$로 길이 안전하지 않지만 그래도 건너며, 이때 확률 0.2로 무사히 도착하고 확률 0.8로 차에 치인다.
- 확률 $0.9\left(1 - \frac{n}{3}\right)$로 다음 확인 시점까지 기다린다.

$n = 0$에서는 개구리가 결코 위험을 무릅쓰지 않고, $n = 3$에서는 안전하지 않을 때 건널 확률이 1이므로, 그때까지 출발하지 않았다면 $n = 3$에서 반드시 출발한다. 특히 $n > 3$인 경우는 생기지 않는다.

확인 시점 $n$에 개구리가 아직 기다리고 있을 확률은

$$w_0 = 1, \quad w_1 = 0.9, \quad w_2 = 0.9 \cdot 0.9 \cdot \tfrac{2}{3} = 0.54, \quad w_3 = 0.54 \cdot 0.9 \cdot \tfrac{1}{3} = 0.162.$$

확인 시점 $n$에 개구리가 무사히 도착할 확률은 $w_n\left(0.1 + 0.9 \cdot \frac{n}{3} \cdot 0.2\right) = w_n(0.1 + 0.06n)$이다. 이 확률을 $s_n$이라 쓰면

$$s_0 = 1 \times 0.1 = 0.1, \qquad s_1 = 0.9 \times 0.16 = 0.144,$$
$$s_2 = 0.54 \times 0.22 = 0.1188, \qquad s_3 = 0.162 \times 0.28 = 0.04536.$$

따라서

$$P(\text{개구리가 안전하게 건넘}) = 0.1 + 0.144 + 0.1188 + 0.04536 = \boxed{0.40816 = \tfrac{2551}{6250}}.$$

개구리가 차에 치일 확률은 나머지인 $1 - 0.40816 = 0.59184$이다.

### 둘 다 안전하고 개구리가 더 빠른 경우

개구리가 무사히 도착하는 시각을 $F$라 하자. 둘 다 무사히 도착하고 $F < T$이어야 한다. 두꺼비는 반드시 안전하므로 독립성에 의해

$$P(\text{개구리가 시점}n\text{에서 안전하고}T > n) = s_n\, P(T > n) = s_n(0.9)^{n+1}.$$

**?** **생각해 보기** 왜 $T \geq n$이 아니라 $T > n$일까?

둘이 같은 확인 시점 $n$에 건너면 걸린 시간이 같으므로, 개구리가 엄밀하게 더 빠른 것이 아니다. 따라서 두꺼비는 시각 $n$의 확인이 끝난 뒤에도 여전히 기다리고 있어야 하고, 그 확률은 $0.9^{n+1}$이다. 개구리의 가능한 네 시각에 대해 더하면

$$\begin{aligned} P &= 0.1(0.9) + 0.144(0.81) + 0.1188(0.729) + 0.04536(0.6561) \\ &= 0.09 + 0.11664 + 0.0866052 + 0.029760696 \\ &= 0.323005896 \approx \boxed{0.323}. \end{aligned}$$

정확한 분수로 나타내면 $\frac{40375737}{125000000}$이다.

### 마지막 문항

개구리가 첫 확인 후 2분이 지났을 때 아직 건너편에 도착하지 못한 사건을 $A$, 개구리가 차에 치이는 사건을 $R$이라 하자. 건너는 것은 순간적으로 이루어지므로, $n = 2$의 확인 시점에 무사히 건넌 개구리는 2분까지 도착한 것이다. 따라서 $A$는 "$n = 0, 1$ 또는 2에서 안전하게 건넘"의 여사건(complement)이다.

$$P(A) = 1 - (s_0 + s_1 + s_2) = 1 - 0.3628 = 0.6372.$$

차에 치인 개구리는 어느 시각에도 도착하지 못하므로 $R \subseteq A$이고 $P(R \cap A) = P(R) = 0.59184$이다. 그러므로

$$P(R \mid A) = \frac{0.59184}{0.6372} = \frac{59184}{63720} = \boxed{\frac{274}{295} \approx 0.929}.$$

여기에는 이미 ($n = 1$ 또는 $n = 2$에서) 차에 치인 개구리와, 아직 기다리다가 $n = 3$에서 차에 치이는 개구리가 모두 포함되며, 이는 "어느 단계에서든"이라는 조건이 요구하는 바이다.

(어떤 독자는 이 조건을 $n = 0$과 $n = 1$의 확인 뒤에, 즉 $n = 2$에서 살피기 **전에** 도착하지 못했다는 뜻으로 받아들인다. 이렇게 해석하면 $0.59184/(1 - 0.244) = \frac{137}{175} \approx 0.783$을 얻는다. 정확히 2분에 도착한 것도 출발한 지 "2분 후"에 도착한 것이므로, 우리는 위의 해석을 택한다.)

## 검산

- 개구리의 결과들을 모두 합하면 전체가 되어야 한다. $n = 1, 2, 3$에서 차에 치일 확률은 $0.9(0.24)$, $0.54(0.48)$, $0.162(0.72)$, 즉 $0.216 + 0.2592 + 0.11664 = 0.59184$이고, 이는 정확히 $1 - 0.40816$이다.
- 개구리가 안전할 확률은 두꺼비의 확실한 1보다 훨씬 작은데, 참을성 없는 쪽이니 당연히 그래야 한다.
- 분수를 이용한 정확한 열거와 파충류 $10^6$쌍에 대한 몬테카를로 시뮬레이션(Monte Carlo simulation)으로 0.4087, 0.3234, 0.9286을 얻었으며, 이는 0.40816, 0.32301, 0.92881과 일치한다.

## 흔한 실수

- 개구리가 $n = 0$에서 위험을 무릅쓰게 하는 것. 그 시점에서는 규칙에 따라 $n/3 = 0$이다.

- 개구리의 과정이 $n = 3$에서 끝난다는 것을 잊고, 두꺼비처럼 개구리에 대해서도 기하급수를 더하는 것.
- 마지막 문항에서 $P$(안전하지 않음, $n \leq 2$)로 나누면서 분자에는 $n = 2$**까지** 차에 치인 경우만 넣는 것. 문제는 $n = 3$을 포함하여 어느 단계에서든 차에 치일 확률을 묻는다.

## STEP 너머

개구리의 여정은 흡수 상태(absorbing state)("안전"과 "치임")를 가진 작은 **마르코프 연쇄**(Markov chain)이고, 두꺼비의 대기 시간은 무기억성(memoryless)을 가진 지수분포(exponential distribution)의 이산 버전인 기하분포이다. 마르코프 연쇄의 흡수 확률(absorption probability)과 도달 시간(hitting time)은 대학 첫 확률론 강의의 핵심 주제이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 이 두 가지를 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.