解答例

STEP III 2007、第 13 問

確率・統計 ・ ランダムウォークの期待値

STEP III 2007、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

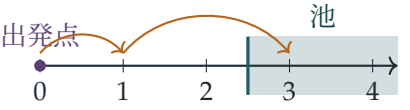
step.maths.org/questions/07-s3-q13

鍵となる考え方

0.5 m という半端な距離は、「ちょうど縁に着地する」場合についての曖昧さを取り除くためのものである。$(n-\frac{1}{2})$ m 離れた地点から出発すると、カエルが池に入るのは、跳んだ距離の合計が n **以上**になったときであり、そのときに限る。したがって問題はすべて、1 と 2 の累積和が初めて n に達する時点に帰着する。n が小さいときは可能性を書き出せばよい。一般の公式については、u_n の三つの値から三つの定数についての三つの方程式が得られる。

解答

最初の j 回の跳躍で進む距離の合計（メートル単位）を S_j とし、池に達するまでに必要な跳躍の回数を T とする。$(n-\frac{1}{2})$ m 離れた地点から出発するので、j 回跳んだ後にカエルが池の中にいるのは $S_j > n-\frac{1}{2}$ のとき、そしてそのときに限る。S_j は整数であるから、これは $S_j \geq n$ と同値である。よって $p_n(j) = \mathrm{P}(T=j)$ であり、T は $S_j \geq n$ を満たす最小の j である。



図は $n=3$（縁までの距離は 2.5 m）の場合で、「短、長」という跳躍の列を示している。この列では 2 回の跳躍で池に達する。

小問 (i)

$n = 2$ のとき、縁までの距離は 1.5 m である。最初の跳躍が長ければ（確率 q）カエルはそのまま池に入るので、$p_2(1) = q$ である。そうでなければ最初の跳躍は短く（確率 p）、カエルは縁から 0.5 m の位置にいるので、2 回目の跳躍はどちらであっても池に入る。各跳躍は独立（independent）であるから

$$p_2(2) = p \times 1 = p,$$

となり、示すべきことが得られた。（また $j \geq 3$ に対して $p_2(j) = 0$ であり、$p + q = 1$ により確率分布（probability distribution）が尽くされていることが確かめられる。）

小問 (ii)

$\frac{1}{2}$ m 離れた地点からはどの跳躍でも池に着地するので、$u_1 = 1$ である。

小問 (i) を用いると

$$u_2 = 1 \cdot p_2(1) + 2 \cdot p_2(2) = q + 2p = 2 - q.$$

$n = 3$ のときは $S_j \geq 3$ となる必要がある。

- $S_1 \leq 2$ であるから、1 回の跳躍では決して足りない。よって $p_3(1) = 0$ である。
- 2 回の跳躍で足りないのは、両方とも短い場合（$S_2 = 2$）に限られる。それ以外の組ではすべて $S_2 \geq 3$ となる。よって $p_3(2) = 1 - p^2$ である。
- 短い跳躍が 2 回続いた後、カエルは 0.5 m 離れた位置にいるので、3 回目の跳躍で必ず池に着地する。よって $p_3(3) = p^2$ である。

したがって

$$\begin{aligned} u_3 &= 2(1 - p^2) + 3p^2 = 2 + p^2 \\ &= 2 + (1 - q)^2 = 3 - 2q + q^2, \end{aligned}$$

となり、示すべきことが得られた。

考えてみよう これらの値をもっと手早く求める方法はあるだろうか。また、それはすべての n に拡張できるだろうか。

ある。最初の跳躍で条件付けすればよい。最初の跳躍の後、カエルは n が 1 または 2 だけ減った同じ状況に置かれるので、$n \geq 2$ に対して

$$u_n = 1 + p\,u_{n-1} + q\,u_{n-2}, \qquad u_0 = 0,$$

が成り立つ。ここで $u_0 = 0$ は、1.5 m の地点から最初に長い跳躍をする場合を表している。この漸化式（recurrence relation）から $u_2 = 1 + p$、$u_3 = 1 + p(1 + p) + q = 2 + p^2$ が再び得られる。小問 (iii) の形が正しい理由もここにある。特性方程式（auxiliary equation）$\lambda^2 - p\lambda - q = 0$ は $(\lambda - 1)(\lambda + q) = 0$ と因数分解でき、これが $(-q)^{n-1}$ の項を与える。そして（「+1」から生じる）重解（repeated root）1 が $B + Cn$ を説明する。

小問 (iii)

$u_n = A(-q)^{n-1} + B + Cn$ に $n = 1, 2, 3$ を代入し、小問 (ii) を用いると

$$A + B + C = 1, \tag{1}$$
$$-qA + B + 2C = 2 - q, \tag{2}$$
$$q^2A + B + 3C = 3 - 2q + q^2. \tag{3}$$

引き算して B を消去すると

$$(2) - (1): \quad -(1 + q)A + C = 1 - q, \tag{4}$$
$$(3) - (2): \quad q(1 + q)A + C = 1 - q + q^2. \tag{5}$$

次に $q \times (4) + (5)$ により A を消去すると

$$(1 + q)C = q - q^2 + 1 - q + q^2 = 1, \qquad \text{よって} \qquad C = \frac{1}{1 + q},$$

となり、示すべきことが得られた。さらに (4) より

$$(1 + q)A = \frac{1}{1 + q} - (1 - q) = \frac{1 - (1 - q^2)}{1 + q} = \frac{q^2}{1 + q}, \qquad \text{よって} \qquad A = \frac{q^2}{(1 + q)^2},$$

であり、(1) より

$$B = 1 - \frac{q^2}{(1 + q)^2} - \frac{1}{1 + q} = \frac{(1 + q)^2 - q^2 - (1 + q)}{(1 + q)^2} = \frac{q}{(1 + q)^2}.$$

以上より

$$u_n = \frac{q^2(-q)^{n-1} + q + (1 + q)n}{(1 + q)^2}.$$

n が大きい場合。 $0 \leq q \leq 1$ であるから、項 $A(-q)^{n-1}$ と B は有界である（いずれも大きさは高々 $\frac{1}{4}$）。一方 Cn は限りなく大きくなる。したがって n が大きいとき

$$u_n \approx Cn = \frac{n}{1 + q} = \frac{n}{p + 2q},$$

である。ここで $1 = p + q$ より $1 + q = p + 2q$ を用いた。

この結果が予想どおりである理由。 1 回の跳躍の長さの平均は $1 \cdot p + 2 \cdot q = p + 2q$ メートルである。カエルはおよそ n メートル進む必要があり、大数の法則（law of large numbers）により、N が大きいとき N 回の跳躍の長さの合計は $N(p + 2q)$ に近い。よって必要な跳躍の回数は $n/(p + 2q)$ に近い。この近似がよいのは、出発点の 0.5 m のずれと、最後の行き過ぎ（カエルが着地するのは縁から高々 1.5 m 先まで）が有界だからである。これらが回数に与える変化は有界な量であり、n に比べて無視できる。この有界な補正こそが $A(-q)^{n-1} + B$ の部分である。

検算

- 公式に $n = 0$ を代入すると $\left(q^2 \cdot (-1/q) + q\right)/(1 + q)^2 = 0 = u_0$ となり、あるべき値と一致する。また、この公式は上の漸化式を満たす。
- $q = 0$（すべての跳躍が短い）の場合：$u_n = n$ となる。カエルはちょうど n 回の 1 m の跳躍を必要とするので、これは正しい。
- すべての跳躍の列を（分数を用いて）厳密に数え上げた結果は、$n = 1, \ldots, 12$ および p のいくつかの値について公式と一致した。$q = 0.7$、$n = 200$ でのシミュレーションでは跳躍回数の平均は 117.90 となり、公式による値 117.89、および $n/(p + 2q) \approx 117.65$ と比較できる。

よくある間違い

- 「池に着地する」をちょうど $S_j = n$ となることと解釈してしまう。0.5 m のずれがあるためカエルは行き過ぎることがあり、$S_j \geq n$ となる列はすべて数に入る。
- $p_3(2) = 2pq + q^2$ と書くべきところで、長い跳躍が先に来る場合を忘れてしまう。実際には $pq + qp + q^2 = 1 - p^2$ が両方の順序を含んでいる。

STEP のその先へ

u_n についての漸化式は再生方程式（renewal equation）である。μ を 1 歩の平均の長さとするとき $u_n \sim n/\mu$ という結果は、初等再生定理（elementary renewal theorem）と呼ばれる。厳密な形はワルド（Wald）の恒等式 $\mathrm{E}(S_T) = \mu\,\mathrm{E}(T)$ から導かれる。ここでは S_T は n または $n + 1$ であるから、$(1 + q)u_n = n + \mathrm{P}(\text{行き過ぎる})$ となる。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、ランダムウォーク（random walk）と母関数（generating function）をこうした考え方で扱っている。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。