解答例

STEP III 2014、第 13 問

確率・統計 ・ 確率母関数

> **STEP III 2014、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/14-s3-q13

鍵となる考え方

第 1 ラウンドで何が起こるかで条件付け（conditioning）を行う。このゲームには記憶がない。ゲームが続くどのラウンドの後でも、その後に続くのは同じゲームのまったく新しいコピーであり、したがってその時点から先の**追加の**得点は N 自身と同じ分布をもつ。これにより G をそれ自身を用いて表す方程式が得られ、それを解いて等比級数（geometric series）として展開する。

解答

$p_n = P(N = n)$ とおくと、$G(t) = \sum_{n \geq 0} p_n t^n$ である。第 1 ラウンドの結果がそれぞれ 1、2、3 であるという事象を E_1, E_2, E_3 とすると、$P(E_1) = a$, $P(E_2) = b$, $P(E_3) = c$ である。

小問 (i)

E_1 が起きたとき、ゲームは得点 0 で終了するので、$P(N = 0 \mid E_1) = 1$ であり、$n \geq 1$ に対して $P(N = n \mid E_1) = 0$ である。したがって条件付き母関数（conditional generating function）は

$$\sum_{n \geq 0} P(N = n \mid E_1)\, t^n = 1 \cdot t^0 = 1.$$

である。

小問 (ii)

E_2 が起きたとき、第 1 ラウンドの後も得点は 0 のままで、ゲームは続く。各ラウンドの確率は同じで、それ以前に起きたことに依存しないので、第 2 ラウンド以降のラウンドは得点 0 から始まる新しいゲームをなす。その新しいゲームが生み出す得点を N' とすると、これは N と同じ分布をもつ。最終得点は $0 + N' = N'$ であるから、

$$P(N = n \mid E_2) = P(N' = n) = p_n \quad \text{任意の} n \geq 0,$$

であり、条件付き母関数は $\sum p_n t^n = G(t)$ である。

小問 (iii)

同様に、E_3 が起きたとき、第 1 ラウンドの後の得点は 1 であり、残りは得点 N' の新しいゲームであるから、最終得点は $1 + N'$ である。よって $P(N = 0 \mid E_3) = 0$ であり、$n \geq 1$ に対して $P(N = n \mid E_3) = p_{n-1}$ である。条件付き母関数は $\sum_{n\geq 1} p_{n-1} t^n = tG(t)$ となる。

全確率の法則（law of total probability）により、

$$p_n = a\,P(N = n \mid E_1) + b\,P(N = n \mid E_2) + c\,P(N = n \mid E_3).$$

$G(t)$ における t^n の係数は p_n である。右辺について、$1, G(t), tG(t)$ における t^n の係数は、小問 (i)–(iii) で求めた $P(N = n \mid E_1), P(N = n \mid E_2), P(N = n \mid E_3)$ にちょうど一致する。したがって $G(t)$ と $a + bG(t) + ctG(t)$ における t^n の係数はすべての $n \geq 0$ について一致し、これにより

$$G(t) = a + bG(t) + ctG(t).$$

が示された。

> **考えてみよう** $1 - b - ct$ で割ってよいか。また、等比級数として展開することは正当か。

式を整理すると $G(t)(1 - b - ct) = a$ となる。$1 - b = a + c > c > 0$ であるから、$|t| < (1 - b)/c$ のとき常に $1 - b - ct \neq 0$ であり、$(1 - b)/c = 1 + a/c > 1$ なので、この範囲は $|t| \leq 1$ を含む。このような t に対して、$|ct/(1 - b)| < 1$ であるから

$$G(t) = \frac{a}{1 - b - ct} = \frac{a}{1 - b} \cdot \frac{1}{1 - \dfrac{ct}{1 - b}} = \frac{a}{1 - b} \sum_{n\geq 0} \left(\frac{c}{1 - b}\right)^n t^n,$$

となる。t^n の係数を読み取ると、

$$P(N = n) = \frac{ac^n}{(1 - b)^{n+1}}, \qquad n \geq 0.$$

（同じことであるが、係数比較により $p_0 = a + bp_0$、および $n \geq 1$ に対して $p_n = bp_n + cp_{n-1}$ が得られる。すなわち $p_0 = a/(1-b)$ かつ $p_n = \frac{c}{1-b}p_{n-1}$ であり、これからも同じ公式が導かれる。）

小問 (iv)

平均は $\mu = G'(1)$ である。$G(t) = a(1-b-ct)^{-1}$ を微分すると、$1-b-c = a$ を用いて

$$G'(t) = \frac{ac}{(1-b-ct)^2}, \qquad \text{よって} \qquad \mu = G'(1) = \frac{ac}{(1-b-c)^2} = \frac{ac}{a^2} = \frac{c}{a},$$

となる。すると $1+\mu = \dfrac{a+c}{a} = \dfrac{1-b}{a}$ であり、したがって

$$\frac{\mu^n}{(1+\mu)^{n+1}} = \frac{c^n}{a^n} \cdot \frac{a^{n+1}}{(1-b)^{n+1}} = \frac{ac^n}{(1-b)^{n+1}} = P(N=n),$$

となり、示すべきことが示された。

検算

- 確率の総和は 1 である：$\sum_n p_n = G(1) = a/(1-b-c) = a/a = 1$。
- 直接的な議論：結果が 2 となる「無駄な」ラウンドを無視する。残りの各ラウンドは確率 $a/(a+c)$ で終了、確率 $c/(a+c)$ で 1 点獲得となるので、N は最初の終了までに獲得する点数であり、$P(N=n) = \left(\frac{c}{a+c}\right)^n \frac{a}{a+c}$ となる。$a+c = 1-b$ であるから、これは同じ公式である。
- $a = \frac{1}{5}, b = \frac{3}{10}, c = \frac{1}{2}$ とすると、$\mu = \frac{5}{2}$、$P(N=0) = \frac{2}{7} \approx 0.2857$、$P(N=1) = \frac{10}{49} \approx 0.2041$ となる。2×10^5 回のゲームをコンピュータでシミュレーションしたところ、0.2855、0.2038、平均得点 2.498 が得られた。

よくある間違い

- 場合 3 の条件付き母関数を $tG(t)$ ではなく $G(t)$ と書いてしまうこと。すでに獲得した 1 点によって、t のすべての累乗が 1 つずつ上にずれる。
- 第 1 ラウンド後のゲームが元のゲームのコピーである**理由**を述べ忘れること。これは、各ラウンドが過去とは独立に同じ確率をもつことに依拠している。

STEP のその先へ

第 1 ステップで条件付けることはマルコフ連鎖（Markov chain）の理論における基本的な道具であり、そこでは**第一ステップ解析**（first-step analysis）と呼ばれる。また、ここでの答えは姿を変えた幾何分布（geometric distribution）である。小問 (iv) は、$\{0, 1, 2, \dots\}$ 上の幾何分布が平均だけで決まることを示している。G についての方程式を立てるのと同じ手法は分枝過程（branching process）にも使え、そこでは家系の名字が途絶える確率に関する有名な結果につながる。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、母関数と分枝過程を非常に読みやすく扱っている。

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。