解答例

STEP II 1994、第 12 問

確率・統計 ・ 条件付き確率

STEP II 1994、問 12

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/94-s2-q12

鍵となる考え方

最初の投げで「ポイント」s が定まると、それ以降の投げはすべて同じ独立（independent）な試行であり、結果は三通りである：s が出る（勝ち）、7 が出る（負け）、それ以外が出る（続行）。したがって、ポイントを固定すれば、ゲームが終わる投げの番号は幾何分布（geometric distribution）に従う。問題全体は、ポイントによって場合分けし、全確率の公式（law of total probability）で足し合わせることに尽きる。最後の小問は、条件となる事象を変えて同じ計算をするだけである。

解答

S を公平な二つのさいころの目の和とする。同様に確からしい 36 通りの結果を数えると、

$$\mathrm{P}(S=s)=\frac{6-|s-7|}{36},\qquad s=2,3,\ldots,12,$$

であり、和 $2,3,4,5,6,7,8,9,10,11,12$ の確率は $\frac{1}{36}\times(1,2,3,4,5,6,5,4,3,2,1)$ である。

最初の投げで勝つ確率

最初の和が 7 または 11 ならば、彼女はただちに勝つ：

$$\mathrm{P}(\text{勝つ投げの番号が1})=\frac{6}{36}+\frac{2}{36}=\frac{2}{9}.$$

複数回投げたという条件のもとで第 n 投で勝つ確率

彼女が複数回投げるのは、最初の和がポイント $4,5,6,8,9,10$ のいずれかであるときに限り、その確率は

$$\mathrm{P}(M) = \frac{3+4+5+5+4+3}{36} = \frac{2}{3},$$

である。ここで M は事象「複数回投げる」を表す。

ポイント s を固定し、$p_s = \mathrm{P}(S = s)$ とおく。以降の各投げはそれ以前の投げと独立であり、各投げについて

$$\mathrm{P}(\text{和が}s) = p_s, \qquad \mathrm{P}(\text{和が}7) = \tfrac{1}{6}, \qquad \mathrm{P}(\text{どちらでもない}) = r_s = 1 - p_s - \tfrac{1}{6}.$$

が成り立つ（$s \neq 7$ なので、これら二つの事象は互いに排反である）。彼女が第 $n \geq 2$ 投で勝つのは、第 $2,\ldots,n-1$ 投がすべて「どちらでもない」であり、かつ第 n 投で s が出るときに限る。よって独立性より

$$\mathrm{P}(\text{勝つ投げの番号が}n \mid \text{ポイント}s) = r_s^{n-2}\, p_s.$$

ポイントは確率の等しい組に分かれる（4 と 10、5 と 9、6 と 8）。$p_s = k/36$（$k = 3,4,5$）と書くと $r_s = (30-k)/36$ であり、これはそれぞれ $\frac{3}{4}$、$\frac{13}{18}$、$\frac{25}{36}$ となる。全確率の公式より、$n \geq 2$ のとき

$$\mathrm{P}(\text{勝つ投げの番号が}n) = \sum_s p_s \cdot r_s^{n-2} p_s = 2\sum_{k=3}^{5} \left(\frac{k}{36}\right)^2 \left(\frac{30-k}{36}\right)^{n-2}.$$

第 $n \geq 2$ 投で勝つならば M が起こっているので、$\mathrm{P}(M) = \frac{2}{3}$ で割ると各係数は $\frac{3}{2}$ 倍になる：

$$2 \cdot \frac{k^2}{1296} \cdot \frac{3}{2} = \frac{k^2}{432} \quad \text{より} \quad \frac{9}{432} = \frac{1}{48}, \quad \frac{16}{432} = \frac{1}{27}, \quad \frac{25}{432}.$$

したがって

$$\mathrm{P}(\text{勝つ投げの番号が}n \mid M) = \frac{1}{48}\left(\frac{3}{4}\right)^{n-2} + \frac{1}{27}\left(\frac{13}{18}\right)^{n-2} + \frac{25}{432}\left(\frac{25}{36}\right)^{n-2},$$

となり、示すべき式が得られた。

考えてみよう 一つの「平均的な」ポイントと一つの幾何分布で考えるのは、なぜ誤りなのだろうか。

続行の確率 r_s がポイントによって異なるからである。三つの幾何分布の混合分布（mixture）はそれ自体は幾何分布ではない。答えに一つではなく三つの別々のべきが現れるのはそのためである。

m 回より多く投げたという条件のもとで

M_m を事象「m 回より多く投げる」とする。ただし $m > 1$ である。これが起こるのは、最初の投げでポイント s が出て、第 $2, \ldots, m$ 投がすべて「どちらでもない」であるときに限るので、

$$\mathrm{P}(M_m) = \sum_s p_s\, r_s^{m-1} = \frac{2}{36} \sum_{k=3}^{5} k\Big(\frac{30-k}{36}\Big)^{m-1}.$$

$n \le m$ のとき、第 n 投で勝つことは M_m と両立しないので、条件付き確率（conditional probability）は 0 である。$n > m$ のとき、第 n 投で勝つならば M_m が起こっているので、

$$\mathrm{P}(\text{勝つ投げの番号が}n \mid M_m) = \frac{\mathrm{P}(\text{勝つ投げの番号が}n)}{\mathrm{P}(M_m)} = \frac{9(\frac{3}{4})^{n-2} + 16(\frac{13}{18})^{n-2} + 25(\frac{25}{36})^{n-2}}{36\Big[3(\frac{3}{4})^{m-1} + 4(\frac{13}{18})^{m-1} + 5(\frac{25}{36})^{m-1}\Big]},$$

となる。ここで因子 $\frac{2}{1296}$ と $\frac{2}{36}$ は約分してある。同じことだが、分子と分母に $\frac{1}{432}$ を掛けると

$$\frac{\frac{1}{48}(\frac{3}{4})^{n-2} + \frac{1}{27}(\frac{13}{18})^{n-2} + \frac{25}{432}(\frac{25}{36})^{n-2}}{\frac{1}{4}(\frac{3}{4})^{m-1} + \frac{1}{3}(\frac{13}{18})^{m-1} + \frac{5}{12}(\frac{25}{36})^{m-1}}, \qquad n > m.$$

分母が $\mathrm{P}(M_m \mid M)$ になっていることに注意しよう。これは当然である：分子は先ほどの M を条件とする答えであり、$M_m \subseteq M$ だからである。

検算

- $m = 1$ とすると分母は $\frac{1}{4} + \frac{1}{3} + \frac{5}{12} = 1$ となり、前の小問の結果に戻る。
- $n = 2$ のとき、条件付きの式は $\frac{1}{48} + \frac{1}{27} + \frac{25}{432} = \frac{50}{432} = \frac{25}{216} \approx 0.1157$ を与える。直接計算しても $\frac{3}{2} \cdot \frac{2(9+16+25)}{1296}$ で同じ値になる。
- $m = 2$、$n = 3$ のとき、答えは $\frac{107}{930} \approx 0.1151$ である。すべての式を、分数を用いてゲームを一手ずつ厳密に追跡した計算と照合し、さらに 4×10^5 回のゲームのシミュレーションも標本誤差の範囲内で一致した。

よくある間違い

- $\mathrm{P}(M) = \frac{2}{3}$ で割るのを忘れること。そうすると係数は $\frac{1}{72}$、$\frac{2}{81}$、$\frac{25}{648}$ になってしまう。
- 単一の公比で $\mathrm{P}(M_m) = \big(\text{何か}\big)^{m-1}$ とすること。条件となる事象もまた、ポイントについての混合なのである。
- 「どちらでもない」投げの回数を $n-2$ ではなく $n-1$ と数えること。第 1 投でポイントが決まり、第 n 投が勝ちの投げである。

STEP のその先へ

このゲームは「勝ち」と「負け」を吸収状態（absorbing state）とする小さなマルコフ連鎖（Markov chain）であり、ここでの吸収までの時間は幾何分布の混合分布である。m 回を超えて続いたという条件をつけると、混合の重みは最も長く続くポイント（ここでは再び出る確率が最も低い 4 と 10）の方へ移る。これは、無記憶性（memoryless）をもつ分布の混合がそれ自体は無記憶でないことを示す簡単な例である。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、このレベルで条件付き確率とマルコフ連鎖を扱っている。

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。