解答例

STEP I 2002、第 14 問

確率・統計 ・ 確率における等比級数

STEP I 2002、問 14

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/02-s1-q14

鍵となる考え方

入力する番号の総数で条件付け（conditioning）を行う。その数を固定すれば、警報が鳴ってから私がお金を手にするまでの時間も定まるので、警察が先に到着する確率はその時間 t に対して単に $1-\mathrm{e}^{-\lambda t}$ となる。これを試行回数について平均すれば答えが得られる。どちらの方法でも、この平均は $q=\mathrm{e}^{-\lambda/r}$ を公比とする等比級数（geometric series）になるので、早い段階で q に名前をつけておくとよい。

解答

順番に入力する場合

リスト $1,2,\ldots,n$ における私の番号の位置を K とする。このとき K は私が行う試行の回数であり、$k=1,\ldots,n$ に対して $P(K=k)=\frac{1}{n}$ である。

$K=k$ と仮定する。私は 1 分あたり r 個の割合で番号を入力するので、続けて入力される番号の間隔は $1/r$ 分である。時刻は最初の番号を入力した時点から測る。

- $k=1$ ならば最初の番号が正しいので、誤った番号は一度も入力されず、警察が呼ばれることはない。警察が先に到着する確率は $0=1-\mathrm{e}^{0}$ である。
- $k\geq 2$ ならば最初の番号は誤りなので、警報は時刻 0 に鳴り、正しい番号は k 番目で、時刻 $(k-1)/r$ に入力される。警察が $t=(k-1)/r$ 分以内に到着すれば、私がお金を手にする前に警察が着くことになり、その確率は $1-\mathrm{e}^{-\lambda(k-1)/r}$ である。

したがって式 $1-\mathrm{e}^{-\lambda(k-1)/r}$ は $k=1$ を含むすべての k に対して成り立つ。私がお金を手にする前に警察が到着するという事象を A とする。全確率の法則（law of total

probability）により

$$P(A) = \sum_{k=1}^{n} P(K=k)\,P(A \mid K=k) = \sum_{k=1}^{n} \frac{1-\mathrm{e}^{-\lambda(k-1)/r}}{n},$$

となり、求める式が得られた。

これを簡単にするため、$q = \mathrm{e}^{-\lambda/r}$ とおく。$\lambda, r > 0$ より $0 < q < 1$ であり、$\mathrm{e}^{-\lambda(k-1)/r} = q^{k-1}$ である。よって

$$\begin{aligned} P(A) &= 1 - \frac{1}{n}\sum_{k=1}^{n} q^{k-1} = 1 - \frac{1-q^n}{n(1-q)} \\ &= 1 - \frac{1-\mathrm{e}^{-\lambda n/r}}{n\left(1-\mathrm{e}^{-\lambda/r}\right)}, \end{aligned}$$

となる。ここで等比級数の和を用いたが、$q \neq 1$ であるからこれは正当である。

無作為に入力する場合

今度は各試行で、それまでの試行とは独立（independently）に、$1, \ldots, n$ から一様（uniformly）に選んだ新しい番号を入力する（私は何を試したかを覚えていない）。したがって各試行は、互いに独立に、確率 $\frac{1}{n}$ で正しい番号となる。正しい番号を入力する試行までを含めた試行回数を N とすると、N は幾何分布（geometric distribution）に従う：

$$P(N=k) = \frac{1}{n}\left(\frac{n-1}{n}\right)^{k-1}, \qquad k = 1, 2, 3, \ldots$$

? **考えてみよう** 番号が重複しうる場合にも、時間に関する議論はそのまま通用するだろうか？

通用する。警察が呼ばれるのは最初に誤った番号を入力したときであり、$N = k \geq 2$ ならばそれはまさに最初の試行、すなわち時刻 0 である。お金が出てくるのは k 回目の試行、すなわち時刻 $(k-1)/r$ である。重複によって変わるのは N の分布だけで、時間は変わらない。よって前と全く同様に $P(A \mid N=k) = 1 - q^{k-1}$ であり、これは $k = 1$ のときも正しい。

$p = \frac{n-1}{n}$ とおく。すると

$$\begin{aligned} P(A) &= \sum_{k=1}^{\infty} \frac{1}{n} p^{k-1}\left(1 - q^{k-1}\right) = \frac{1}{n}\sum_{k=1}^{\infty} p^{k-1} - \frac{1}{n}\sum_{k=1}^{\infty} (pq)^{k-1} \\ &= \frac{1}{n}\cdot\frac{1}{1-p} - \frac{1}{n}\cdot\frac{1}{1-pq}. \end{aligned}$$

$0 \leq p < 1$ かつ $0 < pq < 1$ よりどちらの級数も収束（converge）するので、和を分けてよい。第 1 項は $\frac{1}{n}\cdot n = 1$ であり、これは N が確実に有限であることをそのまま述べ

ている。第 2 項については

$$n(1-pq) = n-(n-1)q,$$

であるから

$$P(A) = 1 - \frac{1}{n-(n-1)\mathrm{e}^{-\lambda/r}},$$

となり、求める式が得られた。

検算

- $n = 1$ のときは誤った番号が入力されることはなく、どちらの式も 0 を与える。前者は $1-q^1 = 1-q$ となるから、後者は分母が 1 となるからである。
- $\lambda \to \infty$（警察が即座に到着する）のとき $q \to 0$ となり、どちらの答えも $1-\frac{1}{n}$ に近づく。これは最初の試行が誤りである確率であり、当然そうなるべき値である。
- 無作為に入力するほうが不利なはずである。$n = 5$、$\lambda = 0.7$、$r = 3$ とすると、順番に入力した場合は 0.3382、無作為に入力した場合は 0.4543 となる。それぞれ 2×10^5 回の試行でシミュレーションを行ったところ、0.339 と 0.456 が得られた。

よくある間違い

- 時刻 $(k-1)/r$ の代わりに k/r を使ってしまう。時間は最初の（誤った）番号から測り始めるので、k 番目の番号はそこから $k-1$ 個の間隔の後に来る。
- 無作為の場合に、各 k に確率 $\frac{1}{n}$ を割り当ててしまう、あるいは和を n で打ち切ってしまう。重複が許されるので、N は幾何分布に従い、上限はない。

STEP のその先へ

ここでの警察の到着時刻は指数分布（exponential distribution）に従う確率変数（random variable）であり、この計算は確率変数 X に対する $P(T < X)$ を、X で条件付けして求めたものである。X が幾何分布に従うとき、$E\left(q^X\right)$ は X の確率母関数（probability generating function）の q における値であるため、答えはきれいな閉じた形になる。実際、どちらの答えも $1-E\left(q^{N-1}\right)$ である。母関数と指数分布は大学 1 年次の確率論の中心的な題材である。Grimmett と Welsh による *Probability: An Introduction*（確率論入門）は、その両方をよく扱っている。

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。