解答例

STEP III 1993、第 15 問

確率・統計 ・ 待ち時間の分布

> **STEP III 1993、問 15**
>
> 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/93-s3-q15

鍵となる考え方

最後の 1 回の投げに注目する。実験が第 n 回でちょうど終わるのは、その回の結果が、まだ足りなかった面の 2 回目を与えるときである。したがって最初の $n-1$ 回には、その面がちょうど 1 回、残りはすべてもう一方の面が出ていなければならない。これにより、分布（distribution）は 2 つの簡単な項の和になる。期待値（expectation）は $\sum n(n-1)x^{n-2}$ の形の級数（series）2 つで表され、これは等比級数（geometric series）を 2 回微分することで求められる。

解答

$0 < p < 1$ と仮定し（そうでなければ実験は終わらない）、各回の投げは独立（independent）であるとする。

N の分布

表が少なくとも 2 回、裏が少なくとも 2 回必要なので、$N \geq 4$ である。$n \geq 4$ を固定する。事象 $N = n$ のもとで、第 n 回は表か裏のいずれかであり、この 2 つの場合は互いに排反（disjoint）である。

第 n 回が表の場合。その直前までは目標に達しておらず、第 n 回で表が加わって目標に達した。よって最初の $n-1$ 回の時点で裏はすでに少なくとも 2 回出ていたが、表は 1 回しか出ていなかった。すなわち、表がちょうど 1 回、裏が $n-2$ 回である（$n \geq 4$ より $n-2 \geq 2$ である）。逆に、そのような列の後に表が出れば、ちょうど第 n 回で終

わる。1 回の表の位置は $n-1$ 通りのいずれでもよいので、この場合の確率は

$$(n-1)\,p\,q^{n-2}\cdot p = (n-1)p^2q^{n-2}.$$

である。

第 n 回が裏の場合。 表と裏を入れ替えて同じ議論をすれば、確率は $(n-1)q^2p^{n-2}$ である。

これらを足して

$$\boxed{\mathrm{P}(N=n) = (n-1)(p^2q^{n-2}+q^2p^{n-2}),\qquad n=4,5,6,\ldots}$$

を得る。また $n\le 3$ に対しては $\mathrm{P}(N=n)=0$ である。

> **考えてみよう** 最後の投げが表のとき、最初の $n-1$ 回に表が 2 回含まれることがないのはなぜか。

もし含まれていたとすると、そこには表が 2 回と（終了の条件から）裏が 2 回すでに含まれていることになり、実験は第 n 回より前に終わっていたはずである。これにより、回数がちょうど 1 回に定まるのである。

期待値

$|x|<1$ のとき、$\sum_{n\ge 0} x^n = (1-x)^{-1}$ を 2 回微分すると

$$\sum_{n=2}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}.$$

を得る。ここで扱う和は $n=4$ から始まるので、$n=2$ と $n=3$ の項、すなわち 2 と $6x$ を取り除く：

$$\sum_{n=4}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3} - 2 - 6x.$$

さて

$$\mathrm{E}(N) = \sum_{n=4}^{\infty} n\,\mathrm{P}(N=n) = p^2\sum_{n=4}^{\infty} n(n-1)q^{n-2} + q^2\sum_{n=4}^{\infty} n(n-1)p^{n-2}.$$

である。この級数を $x=q$（このとき $1-x=p$）として、次に $x=p$（このとき $1-x=q$）として用いる。$0<p,q<1$ であるから、どちらも有効である：

$$\begin{aligned}\mathrm{E}(N) &= p^2\Big(\frac{2}{p^3}-2-6q\Big) + q^2\Big(\frac{2}{q^3}-2-6p\Big)\\ &= \frac{2}{p}+\frac{2}{q}-2(p^2+q^2)-6pq(p+q).\end{aligned}$$

$p+q=1$ であるから、$\frac{1}{p}+\frac{1}{q}=\frac{p+q}{pq}=\frac{1}{pq}$ かつ $p^2+q^2=(p+q)^2-2pq=1-2pq$ である。したがって

$$\begin{aligned}\mathrm{E}(N) &= \frac{2}{pq}-2(1-2pq)-6pq \\ &= \frac{2}{pq}-2-2pq = 2\Big(\frac{1}{pq}-1-pq\Big),\end{aligned}$$

となり、示すべき式が得られた。

検算

- 確率の総和は 1 になる。同じ級数で因子を 1 つ減らしたものから $\sum_{n\geq 4}(n-1)p^2q^{n-2}=1-p^2-2p^2q$ が得られ、表裏を入れ替えたものを足すと、$p^2+q^2=1-2pq$ を用いて $1-(p^2+q^2)-2pq(p+q)=1-1+2pq-2pq=1$ が残る。
- 公平な硬貨では $pq=\frac{1}{4}$ であり、$\mathrm{E}(N)=2(4-1-\frac{1}{4})=\frac{11}{2}$ となる。直接計算すると $\mathrm{P}(N=4)=\frac{3}{8}$、$\mathrm{P}(N=5)=\frac{1}{4}$、…であり、$n\,\mathrm{P}(N=n)$ の部分和は 5.5 に近づく。
- $p\to 0$ のとき $\mathrm{E}(N)\approx 2/p$ である。表が 2 回出るのを待つことになり、1 回あたりおよそ $1/p$ 回の投げを要するからである。
- 長さ 10 までのすべての列を数え上げると分布が正確に再現され、$p=0.3$ でのシミュレーションでは平均 7.10 が得られた。公式による値は 7.1038 である。

よくある間違い

- $n-1$ の代わりに $\binom{n}{2}$ と書いてしまう。最後の投げの位置は固定されているので、自由に動けるのは最初の $n-1$ 回の中にただ 1 回ある表（または裏）だけである。
- 標準的な級数を用いる際に、$n=2$ と $n=3$ の項を引き忘れる。

STEP のその先へ

N は、ずらした 2 つの**負の二項分布**（negative binomial）に従う待ち時間の混合であり、等比級数を微分するという技法は**確率母関数**（probability generating function）の手作り版である。$G(s)=\mathrm{E}(s^N)$ とおけば、平均は $G'(1)$ として、分散は $G''(1)$ から、同じくらい手早く求められる。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）では、母関数や、このような待ち時間の問題が詳しく扱われている。

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。