解答例

STEP III 2020、第 12 問

確率・統計 ・ 幾何分布と独立性

STEP III 2020、問 12

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2020 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：

www.ocr.org.uk/Images/696589-2020-past-papers.zip

鍵となる考え方

すべては独立（independent）な幾何分布（geometric distribution）に従う確率変数の組 (X, Y) の関数であり、$\mathrm{P}(X = x,\, Y = y) = p^2 q^{x+y-2}$ である。和と差については、この同時確率（joint probability）を (x, y) 格子上の直線に沿って足し合わせるのがよく、最大値と最小値については $\mathrm{P}(T \le t)$ のような累積確率（cumulative probability）を通して扱うのが最も簡単である。独立性については、一方の変数のある値が他方の値を完全に決めてしまう場合を探すとよい。

解答

A は最初に表が出るまで硬貨を投げるので、X が値 $k \ge 1$ をとるのは、最初の $k-1$ 回がすべて裏で、k 回目が表のときである。

$$\mathrm{P}(X = k) = pq^{k-1}, \qquad k = 1, 2, 3, \ldots,$$

Y についても同様である。2 人は別々に硬貨を投げるので、X と Y は独立であり、

$$\mathrm{P}(X = x,\, Y = y) = p^2 q^{x+y-2} \qquad (x, y \ge 1).$$

繰り返し用いる事実が 2 つある。$\mathrm{P}(X > k) = q^k$（最初の k 回がすべて裏）であり、したがって $\mathrm{P}(X \le k) = 1 - q^k$ である。また $1 - q^2 = (1-q)(1+q) = p(1+q)$ である。

小問 (i)

和。S は値 $s \geq 2$ をとる。$S = s$ となるには、$1 \leq x \leq s-1$ として $X = x$ かつ $Y = s - x$ であればよい。これら $s-1$ 通りの結果はそれぞれ確率 p^2q^{s-2} をもつので、

$$\mathrm{P}(S = s) = (s-1)p^2q^{s-2}, \qquad s = 2, 3, 4, \ldots$$

最大値。$T \leq t$ となるのは $X \leq t$ かつ $Y \leq t$ のときに限るので、独立性より $\mathrm{P}(T \leq t) = (1-q^t)^2$ である。よって $t \geq 1$ に対して

$$\begin{aligned}\mathrm{P}(T = t) &= \mathrm{P}(T \leq t) - \mathrm{P}(T \leq t-1) = (1-q^t)^2 - (1-q^{t-1})^2 \\ &= [(1-q^t) - (1-q^{t-1})][(1-q^t) + (1-q^{t-1})] \\ &= (q^{t-1} - q^t)(2 - q^{t-1} - q^t) \\ &= pq^{t-1}(2 - q^{t-1} - q^t),\end{aligned}$$

ここで $q^{t-1} - q^t = q^{t-1}(1-q) = pq^{t-1}$ を用いた。これが求める結果である。($t = 1$ のとき、この式は正しく $p(1-q) = p^2$ を与える。)

小問 (ii)

差。$U = 0$ となるには、ある $k \geq 1$ について $X = Y = k$ であればよい。

$$\mathrm{P}(U = 0) = \sum_{k=1}^{\infty} p^2q^{2k-2} = \frac{p^2}{1-q^2} = \frac{p^2}{p(1+q)} = \frac{p}{1+q}.$$

$u \geq 1$ のときは、$Y = X + u$ または $X = Y + u$ である。これらは互いに排反であり、対称性から確率は等しいので、

$$\mathrm{P}(U = u) = 2\sum_{k=1}^{\infty} \mathrm{P}(X = k, Y = k+u) = 2\sum_{k=1}^{\infty} p^2q^{2k+u-2} = \frac{2p^2q^u}{1-q^2} = \frac{2pq^u}{1+q}.$$

最小値。$W > w$ となるのは $X > w$ かつ $Y > w$ のときに限るので、$\mathrm{P}(W > w) = q^{2w}$ であり、$w \geq 1$ に対して

$$\mathrm{P}(W = w) = \mathrm{P}(W > w-1) - \mathrm{P}(W > w) = q^{2w-2} - q^{2w} = (1-q^2)\,q^{2(w-1)}.$$

したがって W 自身も「成功」確率 $1-q^2$ の幾何分布に従う。A と B が 1 回ずつ投げる各回で、少なくとも 1 つ表が出る確率が $1-q^2$ だからである。

小問 (iii)

$S = 2$ ならば、$X, Y \geq 1$ より $X = Y = 1$ でなければならず、このとき $T = 1$ である。よって $S = 2$ と $T = 3$ が同時に起こることはない。

$$\mathrm{P}(S = 2 \text{ かつ} T = 3) = 0.$$

一方、$\mathrm{P}(S=2)=p^2>0$ であり、$0<q<1$ より $q^2+q^3<2$ となるので $\mathrm{P}(T=3)=pq^2(2-q^2-q^3)>0$ である。よって $\mathrm{P}(S=2)\times\mathrm{P}(T=3)>0$ であり、両辺は一致しない。

小問 (iv)

U と W は独立である。 すべての $u\geq 0$ と $w\geq 1$ に対して $\mathrm{P}(U=u,\,W=w)=\mathrm{P}(U=u)\,\mathrm{P}(W=w)$ を示せばよい。U と W の両方がわかれば、小さい方の値が w、大きい方の値が $w+u$ であることがわかる。

考えてみよう $U=u$ かつ $W=w$ となる組 (X,Y) はどれか？

$u=0$ ならば、そのような組は (w,w) だけである。$u\geq 1$ ならば、ちょうど 2 つ、$(w,\,w+u)$ と $(w+u,\,w)$ があり、それぞれの確率は p^2q^{2w+u-2} である。

$u=0$ のとき、$\mathrm{P}(U=0,\,W=w)=p^2q^{2w-2}$ であり、一方

$$\mathrm{P}(U=0)\,\mathrm{P}(W=w)=\frac{p}{1+q}\cdot p(1+q)\,q^{2w-2}=p^2q^{2w-2}.$$

$u\geq 1$ のとき、$\mathrm{P}(U=u,\,W=w)=2p^2q^{2w+u-2}$ であり、一方

$$\mathrm{P}(U=u)\,\mathrm{P}(W=w)=\frac{2pq^u}{1+q}\cdot p(1+q)\,q^{2w-2}=2p^2q^{2w+u-2}.$$

いずれの場合も両者は一致するので、U と W は独立である。

他の組はどれも独立ではない。 残りの組は (S,T)、(S,U)、(S,W)、(T,U)、(T,W) である。2 つの変数が独立でないことを示すには、$\mathrm{P}(A=a,\,B=b)\neq\mathrm{P}(A=a)\,\mathrm{P}(B=b)$ となる値の組を 1 つ挙げれば十分である。

考えてみよう S または T のどの 1 つの値が、他のすべてを決めてしまうか？

事象 $S=2$ と事象 $T=1$ は同じ事象 $\{X=Y=1\}$ であり、その確率は $p^2>0$ である。この事象のもとでは $T=1$、$U=0$、$W=1$ である。$0<p,q<1$ であるから、以下に挙げる T、U、W の値はいずれも、小問 (i) と (ii) の式により正の確率をもつ。よって

- (S,T)：小問 (iii) より $\mathrm{P}(S=2,\,T=3)=0\neq\mathrm{P}(S=2)\,\mathrm{P}(T=3)$；
- (S,U)：$\mathrm{P}(S=2,\,U=1)=0$ であるが、$\mathrm{P}(S=2)\,\mathrm{P}(U=1)>0$；
- (S,W)：$\mathrm{P}(S=2,\,W=2)=0$ であるが、$\mathrm{P}(S=2)\,\mathrm{P}(W=2)>0$；
- (T,U)：$\mathrm{P}(T=1,\,U=1)=0$ であるが、$\mathrm{P}(T=1)\,\mathrm{P}(U=1)>0$；
- (T,W)：$\mathrm{P}(T=1,\,W=2)=0$ であるが、$\mathrm{P}(T=1)\,\mathrm{P}(W=2)>0$。

したがって、S、T、U、W の中で独立な組は U と W だけである。

検算

- 各分布の確率の総和は 1 になる。例えば $\sum_{t\geq 1}\mathrm{P}(T=t)$ は次々に打ち消し合って（telescoping）$\lim(1-q^t)^2=1$ となり、また $\frac{p}{1+q}+\sum_{u\geq 1}\frac{2pq^u}{1+q}=\frac{p}{1+q}\left(1+\frac{2q}{p}\right)=\frac{p+2q}{1+q}=1$ である。
- $p=q=\frac{1}{2}$ のとき、$\mathrm{P}(T=2)=\frac{1}{4}\left(2-\frac{1}{2}-\frac{1}{4}\right)=\frac{5}{16}$ となり、これは 3 つの結果 $(1,2),(2,1),(2,2)$ を書き出して確率 $\frac{1}{8}+\frac{1}{8}+\frac{1}{16}$ を足したものと一致する。
- $p=q=\frac{1}{2}$ のとき、$\mathrm{P}(U=1,W=1)=\mathrm{P}\{(1,2),(2,1)\}=\frac{1}{4}$ であり、$\mathrm{P}(U=1)\,\mathrm{P}(W=1)=\frac{1}{3}\times\frac{3}{4}=\frac{1}{4}$ となって、独立性が要求するとおりである。
- すべての式と U と W の独立性は、$p=0.3$、0.55、0.8 について同時分布を足し合わせることで、コンピュータでも確認した。

よくある間違い

- $u\geq 1$ のときの $\mathrm{P}(U=u)$ で因子 2 を忘れること、あるいは $X=Y$ が 1 回しか数えられない $u=0$ のときに誤ってそれを含めること。
- 場合を直接足し合わせて $\mathrm{P}(T=t)$ を求めようとし、$X=Y=t$ を二重に数えてしまうこと。累積確率を用いる方法ならこれを避けられる。
- 値の組を 1 つ確かめれば独立性が証明できると考えること。独立性には **すべての** u と w についての確認が必要であるが、従属性を示すには 1 つの反例で十分である。

STEP のその先へ

U と W の独立性は、無記憶性（memoryless property）の離散版である。最初の表が時刻 W に出たとき、もう一方のプレイヤーの残りの待ち時間は、それまでにどれだけ時間がかかったかに関係なく、新たな幾何分布に従う確率変数となる。独立な指数分布（exponential distribution）に従う確率変数についての連続版では、最小値と、最大値までの差が独立になる。これはポアソン過程（Poisson process）や待ち行列理論（queueing theory）の基礎となっている。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）がこれらの考え方を展開している。

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。