解答例

STEP III 2023、第 12 問

確率・統計 ・ 組合せ論的確率

STEP III 2023、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 2023 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/707102-2023-past-papers.zip

鍵となる考え方

ペアになっている 2 枚の靴下は見た目こそ同じであるが、選び方は $2n$ 枚の**物理的な**靴下から $2k$ 枚を一様ランダム（uniformly random）に選んだ集合である。そこで、すべての靴下に番号を付け、集合を数えることにする。あとはすべて、選び方を段階的に組み立てる問題になる。まず、どのペアが完全にそろって現れるかを選び、次に、ほかのどのペアから 1 枚だけが選ばれるかを選び、最後に、そうした各ペアのどちらの靴下が選ばれるかを選ぶ。小問 (iii) はおなじみの技法 $r\binom{n}{r} = n\binom{n-1}{r-1}$ であり、これによって重み付きの和が確率の和に変わる。

解答

$2n$ 枚の靴下がすべて区別できるように番号を付ける。選び方は $\binom{2n}{2k}$ 通りあり、どれも同様に確からしい。以下の確率はすべて（条件を満たす選び方の数）$/\binom{2n}{2k}$ である。

小問 (i)

完全なペアを 1 組も含まない選び方では、各ペアから高々 1 枚しか選ばれないので、$2k$ 枚の靴下は $2k$ 個の異なるペアから来ている。このような選び方を作るには、靴下が選ばれる $2k$ 個のペアを $\binom{n}{2k}$ 通りで選び（$2k \le n$ なので可能である）、次にそれぞれのペアから 2 枚のうち 1 枚を 2^{2k} 通りで選べばよい。このような選択ごとに異なる集合が得られ、ペアを含まない集合はすべてこの方法で得られるから、

$$\mathrm{P}(\text{ペアなし}) = \frac{\binom{n}{2k}\,2^{2k}}{\binom{2n}{2k}}.$$

小問 (ii)

$0 \le r \le k$ を満たす r を固定する。完全なペアをちょうど r 組含む選び方は、次のものからなる。

- 丸ごと選ばれた r 組のペア（$2r$ 枚の靴下）
- それ以外の $2k - 2r = 2(k-r)$ 枚の靴下。これらはすべて異なるペアから来ており（そうでなければペアが r 組より多くなる）、すでに使った r 組のペアからは 1 枚も来ていない。

したがって、完全なペア r 組を $\binom{n}{r}$ 通りで選び、残りの $n-r$ 組のペアから 1 枚ずつ靴下を出す $2(k-r)$ 組のペアを $\binom{n-r}{2(k-r)}$ 通りで選び、さらにそれぞれから靴下を 1 枚ずつ $2^{2(k-r)}$ 通りで選ぶ。この場合も、ペアをちょうど r 組含む集合はそれぞれちょうど 1 回ずつ得られるので、

$$\mathrm{P}(X_{n,k} = r) = \frac{\binom{n}{r}\binom{n-r}{2(k-r)}\, 2^{2(k-r)}}{\binom{2n}{2k}},$$

となり、示すべき式が得られた。（$r = 0$ のとき、これは再び小問 (i) の結果である。）

小問 (iii)

$1 \le r \le k$ とする。

$$r\binom{n}{r} = \frac{r \cdot n!}{r!\,(n-r)!} = n \cdot \frac{(n-1)!}{(r-1)!\,(n-r)!} = n\binom{n-1}{r-1},$$

であるから、小問 (ii) より

$$r\,\mathrm{P}(X_{n,k} = r) = \frac{n\binom{n-1}{r-1}\binom{n-r}{2(k-r)}\, 2^{2(k-r)}}{\binom{2n}{2k}}.$$

ここで、ペアが $n-1$ 組、取り出す靴下が $2(k-1)$ 枚、見つかるペアが $r-1$ 組の場合に小問 (ii) を適用する。$2k \le n$ から $2(k-1) \le n-1$ が従い、また $0 \le r-1 \le k-1$ であるから、これは許される。$(n-1)-(r-1) = n-r$ かつ $(k-1)-(r-1) = k-r$ であるから、

$$\mathrm{P}(X_{n-1,k-1} = r-1) = \frac{\binom{n-1}{r-1}\binom{n-r}{2(k-r)}\, 2^{2(k-r)}}{\binom{2n-2}{2k-2}}.$$

分子は因子 n を除いて一致するので、

$$\begin{aligned}\frac{r\,\mathrm{P}(X_{n,k} = r)}{\mathrm{P}(X_{n-1,k-1} = r-1)} &= \frac{n\binom{2n-2}{2k-2}}{\binom{2n}{2k}} = n \cdot \frac{(2n-2)!}{(2k-2)!\,(2n-2k)!} \cdot \frac{(2k)!\,(2n-2k)!}{(2n)!} \\ &= n \cdot \frac{2k(2k-1)}{2n(2n-1)} = \frac{k(2k-1)}{2n-1}.\end{aligned}$$

（割り算をしても問題はない。$n-r \geq 2k-r \geq 2(k-r)$ であるから、ここに現れる二項係数（binomial coefficient）はすべて正である。）分母を払えば、示すべき結果が得られる。

? **考えてみよう** この等式は $\mathrm{E}(X_{n,k})$ を求めるのにどう役立つか。

期待値（expectation）はちょうど左辺の和であり、右辺を足し合わせると一つの確率分布（probability distribution）全体の和になる。期待値の $r=0$ の項は 0 であるから、

$$\begin{aligned}\mathrm{E}(X_{n,k}) = \sum_{r=1}^{k} r\,\mathrm{P}(X_{n,k}=r) &= \frac{k(2k-1)}{2n-1}\sum_{r=1}^{k}\mathrm{P}(X_{n-1,k-1}=r-1)\\ &= \frac{k(2k-1)}{2n-1}\sum_{s=0}^{k-1}\mathrm{P}(X_{n-1,k-1}=s).\end{aligned}$$

確率変数 $X_{n-1,k-1}$ は $2(k-1)$ 枚の靴下の中のペアの数を数えるので、値 $0,1,\ldots,k-1$ しかとらず、最後の和は 1 である。よって

$$\mathrm{E}(X_{n,k}) = \frac{k(2k-1)}{2n-1}.$$

検算

- $k=1$（靴下 2 枚）のとき、答えは $\frac{1}{2n-1}$ である。最初にどの靴下が出ても、残りの $2n-1$ 枚のうちちょうど 1 枚がその相方である。
- 指示変数（indicator variable）を使えば平均が直接求まる。ペア i がそろう確率は $\binom{2n-2}{2k-2}/\binom{2n}{2k} = \frac{2k(2k-1)}{2n(2n-1)}$ であり、これにペアの数 n を掛けると、やはり $\frac{k(2k-1)}{2n-1}$ が得られる。
- $n=6, k=3$ のとき、小問 (i) から $\frac{64}{924} = \frac{16}{231}$ が得られ、平均は $\frac{15}{11}$ である。$n \leq 8$ のすべての場合について、コンピュータですべての選び方を列挙し、(ii) の分布、(iii) の等式、および平均が正しいことを確認した。

よくある間違い

- ペアの靴下を本当に区別できないものとみなし、「色のパターン」を数えてしまう。それらは同様に確からしくないので、番号を付けた靴下について数えなければならない。
- 1 枚だけ選ばれた各ペアについて、どちらの靴下が選ばれたかを表す因子 $2^{2(k-r)}$ を忘れる。
- (iii) で $\sum_s \mathrm{P}(X_{n-1,k-1}=s)=1$ に気づかず、和を力ずくで計算しようとする。

STEP のその先へ

2 つ目の検算は、指示変数を用いた期待値の線形性（linearity of expectation）であり、分布をまったく知らなくても平均が求まる。事象「ペア i がそろう」どうしが独立でなくても、この方法は成り立つ。(iii) の等式はサイズバイアス（size-biasing）の関係式である。分布を r で重み付けすると、より小さな問題の分布をずらしたものが得られる。この手法を使えば、二項分布（binomial distribution）や超幾何分布（hypergeometric distribution）の平均も求められる。Blitzstein and Hwang, *Introduction to Probability*（確率論入門）では、このような指示変数やストーリー証明が全編にわたって用いられている。

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。