解答例

# STEP III 1991、第 15 問

確率・統計 ・ 標本中の組の数え上げ

> **STEP III 1991、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/91-s3-q15

## 鍵となる考え方

$k$ 枚のカードの組はどれも等しい確率で現れるので、各確率は手札（hand）の数を $\binom{2n}{k}$ で割ったものになる。与えられたペアの構造をもつ手札を数えるには、段階を追って組み立てればよい。まず、どの種類（sort）が完全なペアとして現れるかを選ぶ。次に、それ以外のどの種類が 1 枚だけ現れるかを選ぶ。最後に、1 枚だけの各種類について、2 枚のうちどちらのカードを持っているかを選ぶ。この数え方が確立すれば、問題のすべての小問は 1 つの公式の特別な場合となる。

## 解答

$2n$ 枚のカードをすべて区別できるものと考える（同じ種類の 2 枚に 1 と 2 の番号を付けておけばよい）。すると、$k$ 枚を非復元抽出（without replacement）で引くことにより、大きさ $k$ の $\binom{2n}{k}$ 個の部分集合のそれぞれが同じ確率で得られる。ある種類の 2 枚がともに引かれたとき、手札はその種類の**ペア**（pair）を含むという。

### 4 枚、ペアなし

ペアを含まない手札は、異なる 4 種類のカードを 1 枚ずつ含む。種類の選び方は $\binom{n}{4}$ 通り、各種類のカードの選び方は 2 通りであるから、手札は $2^4\binom{n}{4}$ 通りとなる。

$$\binom{2n}{4} = \frac{2n(2n-1)(2n-2)(2n-3)}{24} = \frac{4n(n-1)(2n-1)(2n-3)}{24},$$

であるから、

$$\mathrm{P}(\text{ペアなし}) = \frac{16\cdot n(n-1)(n-2)(n-3)/24}{4n(n-1)(2n-1)(2n-3)/24} = \frac{4(n-2)(n-3)}{(2n-1)(2n-3)},$$

となり、題意が示された。（ここで $n \geq 4$ により、異なる 4 種類が存在することが保証される。）

### 4 枚、ちょうど 1 組のペア

ペアとして現れる種類を選ぶ（$n$ 通り）。残りの 2 枚は、残りの $n-1$ 種類のうちの異なる 2 種類から来なければならず、その選び方は $\binom{n-1}{2}$ 通りで、それぞれについてカードの選び方が 2 通りある。手札の数は

$$n \cdot \frac{(n-1)(n-2)}{2} \cdot 4 = 2n(n-1)(n-2),$$

であるから、

$$\mathrm{P}(\text{ちょうど 1 組のペア}) = \frac{24 \cdot 2n(n-1)(n-2)}{4n(n-1)(2n-1)(2n-3)} = \frac{12(n-2)}{(2n-1)(2n-3)}.$$

### $k$ 枚、ちょうど $r$ 組のペア

手札がちょうど $r$ 組のペアを含むとする。これらで $2r$ 枚を使う。残りの $k-2r$ 枚はすべて異なる種類に属し、そのどれも $r$ 個のペアの種類ではない。そうでなければペアがさらに増えてしまうからである。したがって、次のように数える。

- ペアになる種類：$\binom{n}{r}$ 通り。
- 1 枚だけ現れる $k-2r$ 種類を、残りの $n-r$ 種類から選ぶ：$\binom{n-r}{k-2r}$ 通り。
- 1 枚だけの各種類について、2 枚のうちどちらを持つか：$2^{k-2r}$ 通り。

ちょうど $r$ 組のペアを含む各手札は、このような選択の列のちょうど 1 つから得られる。よって

$$\boxed{\mathrm{P}(\text{ちょうど } r \text{ 組のペア}) = \frac{2^{k-2r}\binom{n}{r}\binom{n-r}{k-2r}}{\binom{2n}{k}}}$$

**考えてみよう** これが 0 になることはあるか。また、4 枚の場合の答えと一致するか。

1 枚だけの種類については $k-2r \leq n-r$、すなわち $r \geq k-n$ が必要である。$k > n$ ならば、どの $k$ 枚も少なくとも $k-n$ 組のペアを含まなければならず、このことは $\binom{n-r}{k-2r} = 0$ によって自動的に処理される。$k = 4$ のとき、$r = 0$ の場合は $2^4\binom{n}{4}/\binom{2n}{4}$、$r = 1$ の場合は $2^2\, n\binom{n-1}{2}/\binom{2n}{4}$ となり、これらはまさに上の 2 つの数え上げに一致する。

### $k$ が偶数で、すべてのカードがペアになる場合

$k = 2m$ とおく。手札が $m$ 組のペアからなるのは、それが $m$ 個の完全な種類の和集合であるときに限る。そのような手札は $\binom{n}{m}$ 通りある（これは一般の公式で $r = m$ とした場合である）。よって

$$\mathrm{P} = \frac{\binom{n}{m}}{\binom{2n}{2m}} = \frac{n!}{m!\,(n-m)!} \cdot \frac{(2m)!\,(2n-2m)!}{(2n)!}.$$

これを簡単にするために、階乗の偶数の因数と奇数の因数を分けてまとめる。

$$(2j)! = \big(2 \cdot 4 \cdots 2j\big)\big(1 \cdot 3 \cdots (2j-1)\big) = 2^j\, j!\,\big(1 \cdot 3 \cdots (2j-1)\big).$$

$O_j = 1 \cdot 3 \cdots (2j-1)$ とおき、$O_0 = 1$ とする。すると $(2j)!/j! = 2^j O_j$ であるから、

$$\mathrm{P} = \frac{2^m O_m \cdot 2^{n-m} O_{n-m}}{2^n O_n} = \frac{O_m\, O_{n-m}}{O_n}.$$

$O_n$ を $O_{n-m}$ で割ると、$2n - 2m + 1 = 2n - k + 1$ から $2n - 1$ までの奇数が残るので、

$$\mathrm{P}(\tfrac{1}{2}k \text{ 組のペア}) = \frac{1 \times 3 \times 5 \times \cdots \times (k-1)}{(2n-1)(2n-3)\cdots(2n-k+1)},$$

となり、題意が示された。

直接的な鮮やかな議論もある。山札全体をよく切り、上から $k$ 枚を手札とする。一番上のカードに注目する。その相方のカードは、他の $2n-1$ 個の位置のどれにも等しい確率で存在し、そのうち $k-1$ 個は手札の中にあるので、相方が手札にある確率は $\frac{k-1}{2n-1}$ である。そうであれば、そのペアを取り除く。残るのは $2n-2$ 枚のカードのランダムな並びで、手札はその上から $k-2$ 枚である。同じ議論を繰り返すと $\frac{k-3}{2n-3}$ が得られ、以下同様にして $\frac{1}{2n-k+1}$ まで続く。これらの条件付き確率（conditional probability）を掛け合わせると、同じ積が得られる。

## 検算

- 4 枚の場合、ペアの数は 0、1、2 の 3 通りであり、2 組のペアとなる確率は $\binom{n}{2}/\binom{2n}{4} = \frac{3}{(2n-1)(2n-3)}$ である。実際

  $$4(n-2)(n-3) + 12(n-2) + 3 = 4n^2 - 8n + 3 = (2n-1)(2n-3),$$

  であるから、3 つの確率の和は 1 になる。

- $n = 5$ で 4 枚の場合、確率は $\frac{8}{21}$、$\frac{4}{7}$、$\frac{1}{21}$ である。$4 \le n \le 8$ および 3 から $2n-1$ までのすべての $k$ について、コンピュータですべての手札を列挙したところ、一般の公式ともすべてペアになる場合の公式とも一致した。

- $k = 2$ のとき、すべてペアになる場合の公式は $\frac{1}{2n-1}$ を与える。これは 2 枚目のカードが 1 枚目と一致する確率である。

## よくある間違い

- 1 枚だけのカードを、残りの $2n-2r$ 枚すべてから $\binom{2n-2r}{k-2r}$ として選んでしまうこと。これでは、そのうち 2 枚が同じ種類になることを許してしまい、余分なペアが生じる。
- 因数 $2^{k-2r}$ を忘れること。1 枚だけ現れる各種類について、そのカードは 2 枚のうちどちらでもありうる。

## STEP のその先へ

ちょうど $r$ 組のペアを含む手札の数は、「構造を選んでから中身を埋める」という典型的な数え上げであり、積 $1\cdot 3\cdots(2m-1)$（二重階乗（double factorial）$(2m-1)!!$）は $2m$ 個の対象をペアに分ける方法の数を数えている。二重階乗は、標準正規分布（standard normal distribution）の偶数次のモーメント（moment）としても再び現れる。ランダムな標本の中にペアがいくつ残るかという同じ組合せの問題は、古典的な「なくした靴下」の問題や誕生日型の問題にも現れる。Grimmett と Welsh による *Probability: An Introduction*（確率論入門）は、この種の数え上げの良い基礎を与えてくれる。

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。