解答例

# STEP III 2007、第 14 問

確率・統計 ・ 幾何的確率

> **STEP III 2007、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/07-s3-q14

## 鍵となる考え方

すべての穴を囲む、中心を固定した最小の円（または正方形）は**最も遠い**穴によって決まる。したがって求めたいのは最大値（maximum）の分布であり、$P(\max \le r)$ は 1 回投げたときの確率の $n$ 乗にすぎない。円板の場合、1 つの穴の $O$ からの距離の 2 乗 $R^2$ は $[0,1]$ 上の一様分布（uniform distribution）に従うことがわかり、これですべてがすっきりする。正方形の場合、適切な「距離」は $\max(|x|,|y|)$ であり、これは円板上の半径とまったく同じふるまいをする。この一致のおかげで、小問 (iii) もほとんど労せずに片がつく。

## 解答

### 小問 (i)

$n$ 個の穴の $O$ からの距離を $R_1,\dots,R_n$ とする。投げる試行が独立（independent）なので、これらも独立である。1 回の試行と $0 \le r \le 1$ について、穴が $O$ から距離 $r$ 以内にあるのは、それが半径 $r$ の円板に落ちるときであり、かつそのときに限るから、

$$P(R_i \le r) = \frac{\pi r^2}{\pi \cdot 1^2} = r^2.$$

$U_i = R_i^2$ とおく。すると $0 \le u \le 1$ に対して $P(U_i \le u) = P(R_i \le \sqrt{u}) = u$ であるから、各 $U_i$ は$[0,1]$ **上の一様分布に従い**、$U_i$ たちは独立である。

$O$ を中心としすべての穴を囲む最小の円の半径は $M = \max_i R_i$、面積は $\pi M^2 = \pi \max_i U_i$ である。$V = \max_i U_i$ とおく。独立性より

$$P(V \le v) = P(U_1 \le v)\cdots P(U_n \le v) = v^n \qquad (0 \le v \le 1),$$

となるので、$V$ の確率密度関数（density）は $nv^{n-1}$ であり、

$$E(V) = \int_0^1 v \cdot nv^{n-1}\,\mathrm{d}v = \frac{n}{n+1}.$$

よって、$n$ 個の穴をすべて囲む最小の円の面積の期待値（expectation）は

$$\boxed{\frac{n\pi}{n+1}}.$$

$O$ に最も近い $n-1$ 個の穴を囲む円については、半径は **2 番目に大きい**距離であるから、面積は $\pi W$ となる。ここで $W$ は $U_1, \ldots, U_n$ のうち **2** 番目に大きいものである。

**?** **考えてみよう** **2** 番目に大きい値の分布はどうすれば求められるか？

$W \le w$ となるのは、$U_i$ のうち $w$ を超えるものが高々 **1** 個のときであり、かつそのときに限る。超えるものが **1** つもない（確率 $w^n$）か、ちょうど **1** つだけある場合であり、後者はどれが超えるかの選び方が $n$ 通りあるので確率は $n(1-w)w^{n-1}$ となる。よって

$$P(W \le w) = w^n + n(1-w)w^{n-1} = nw^{n-1} - (n-1)w^n.$$

微分する代わりに、$[0,1]$ に値をとる任意の確率変数（random variable）について成り立つ $E(W) = \int_0^1 P(W > w)\,\mathrm{d}w$ を用いることができる：

$$E(W) = \int_0^1 (1 - nw^{n-1} + (n-1)w^n)\,\mathrm{d}w = 1 - 1 + \frac{n-1}{n+1} = \frac{n-1}{n+1}.$$

したがって、最も近い $n-1$ 個の穴を囲む円の面積の期待値は

$$\boxed{\frac{(n-1)\pi}{n+1}}.$$

（これはきれいなパターンに合致している。$n$ 個の独立な一様分布の点は $[0,1]$ を、長さの期待値がどれも $\frac{1}{n+1}$ である $n+1$ 個の区間に分けるので、$k$ 番目に小さい点の平均は $\frac{k}{n+1}$ となる。）

### 小問 (ii)

$Q$ を通り盤の辺に平行な座標軸をとると、盤は $-1 \le x, y \le 1$ となる。囲む正方形は、辺が盤の辺に平行なものとする。$Q$ を中心とし、辺の長さの半分が $s$ である正方形は $\{|x| \le s,\ |y| \le s\}$ であるから、穴 $(x, y)$ を含むのは $\max(|x|, |y|) \le s$ のときであり、かつそのときに限る。

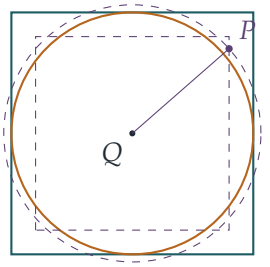
$i$ 番目の穴について $S_i = \max(|x_i|, |y_i|)$ とおく。$0 \le s \le 1$ に対して

$$P(S_i \le s) = \frac{(2s)^2}{2^2} = s^2,$$

であり、これは小問 (i) の $P(R_i \le r) = r^2$ と同じである。囲む最小の正方形の辺の長さの半分は $\max_i S_i$、面積は $(2\max_i S_i)^2 = 4\max_i S_i^2$ である。小問 (i) とまったく同様に、$S_i^2$ たちは独立で $[0,1]$ 上の一様分布に従うから、その最大値の期待値は $\frac{n}{n+1}$ である。囲む最小の正方形の面積の期待値は

$$\boxed{\frac{4n}{n+1}}.$$

**小問 (iii)**




今度は盤が正方形であり、$Q$ を中心としすべての穴を囲む最小の円を求めたい。穴 $P = (x, y)$ の $Q$ からの距離は $\sqrt{x^2+y^2}$ であり、

$$\sqrt{x^2+y^2} \ \ge \ \max(|x|, |y|) = S,$$

が成り立ち、等号は $P$ が座標軸上にあるときに限る。(図において、$P$ を通る破線の円は、$P$ を通る破線の正方形の内接円より小さくなることはない。) したがって、$D$ を穴の $Q$ からの距離の最大値とすると $D \ge \max_i S_i$ であり、囲む円の面積は

$$\pi D^2 \ \ge \ \pi \max_i S_i^2.$$

を満たす。小問 (ii) より、$\max_i S_i^2$ は円形の盤における $\max_i R_i^2$ とまったく同じ分布をもつから、

$$E(\pi D^2) \ \ge \ \pi\, E(\max_i S_i^2) = \frac{n\pi}{n+1},$$

となり、右辺は小問 (i) の円形の盤での答えである。この不等式は等号が成り立たない。正の確率でどの穴も座標軸上になく、そのとき $D > \max_i S_i$ となるからである。よって、正方形の盤における円の面積の期待値のほうが**大きい**。

直観的には、正方形の盤は円形の盤を含み (円板はその内接円である)、ダーツは隅のほうまで散らばるので、最も遠い穴はより遠くにある傾向がある、ということである。

## 検算

- $n = 1$ のとき、公式は $\frac{\pi}{2}$ と 2 を与える。一様分布に従う $R^2$ の平均は $\frac{1}{2}$ なので、これは正しい。$n \to \infty$ のとき面積は $\pi$ と 4、すなわち盤全体に近づき、そうなるべきとおりである。
- $n = 2, 3, 5$ について $2 \times 10^5$ 回の試行によるモンテカルロ・シミュレーション（Monte Carlo simulation）を行ったところ、$\frac{n\pi}{n+1}$、$\frac{(n-1)\pi}{n+1}$、$\frac{4n}{n+1}$ と有効数字（significant figures）約 3 桁で一致した。たとえば $n = 3$ では $\frac{3\pi}{4} \approx 2.3562$ に対して 2.3561 が得られた。
- 小問 (iii) については、厳密な分布の数値積分（quadrature）により、$n = 3$ のとき正方形の盤での面積の期待値は約 3.257、円板では 2.356 となった。主張どおり、前者のほうが大きい。

## よくある間違い

- $E(M)$ を求めてから 2 乗してしまうこと。面積の期待値は $\pi E(M^2)$ であって、$\pi\big(E(M)\big)^2$ ではない。
- 中心からの距離が $[0, 1]$ 上の一様分布に従うとしてしまうこと。面積に比例して当たるので、一様分布に従うのは $R$ ではなく $R^2$ である。外側の輪のほうが大きいのである。
- 2 番目に大きい値について、どの穴が円の外にあるかを表す因子 $n$ を忘れること。

## STEP のその先へ

標本の最大値や 2 番目に大きい値は**順序統計量**（order statistics）である。$[0, 1]$ 上の一様分布に従う $n$ 個の独立な確率変数について、$k$ 番目に小さいものはベータ分布（Beta distribution）に従い、その平均は $\frac{k}{n+1}$ である。小問 (iii) は**確率優越**（stochastic dominance）（カップリング（coupling））の一例である。ある確率変数がつねに別の確率変数以上であれば、その期待値も同様であり、この比較には詳しい計算を必要としない。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）、および Blitzstein と Hwang の *Introduction to Probability*（確率論入門）は、いずれも順序統計量と確率変数の変換を扱っている。

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。