解答例

STEP I 2011、第 13 問

確率・統計 ・ 連続型確率変数

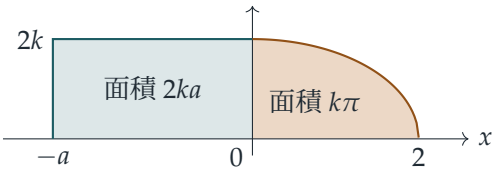
STEP I 2011、問 13

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/11-s1-q13

鍵となる考え方

確率密度（density）のグラフを描いてみる。それは $[-a,0)$ 上の高さ $2k$ の長方形と、$[0,2]$ 上の半径 2 の四分円（縦方向に k 倍したもの）をつなげた形である。グラフの下の面積が確率を表すので、積分をしなくても長方形と四分円でほとんどの計算が済む。全体の面積は 1 でなければならず、これにより k が決まる。その後は、d に関するどの問いも、ある点より右側にどれだけの面積があるかを問うことに帰着する。



解答

k を求める

$[0,2]$ 上で $\sqrt{4-x^2}$ のグラフは円 $x^2+y^2=4$ の四分の一であるから、

$$\int_0^2 \sqrt{4-x^2}\,\mathrm{d}x = \tfrac{1}{4}\pi\cdot 2^2 = \pi.$$

（与えられた原始関数（antiderivative）を用いても一致する：$2\arcsin 1 + 0 - 0 = \pi$。）したがって全確率は $2ka + k\pi = 1$ であり、

$$k = \frac{1}{2a+\pi}.$$

小問 (i)

$\mathrm{E}(X)$ の積分を 0 で分ける：

$$\int_{-a}^{0} 2kx\,\mathrm{d}x = k[x^2]_{-a}^{0} = -ka^2.$$

後半については、$(4-x^2)^{3/2}$ の導関数（derivative）が $-3x(4-x^2)^{1/2}$ であるから、

$$\int_{0}^{2} kx\sqrt{4-x^2}\,\mathrm{d}x = k\Big[-\tfrac{1}{3}(4-x^2)^{3/2}\Big]_{0}^{2} = k\cdot\tfrac{1}{3}\cdot 8 = \tfrac{8}{3}k.$$

二つを足し合わせて、

$$\mathrm{E}(X) = k\Big(\frac{8}{3} - a^2\Big) = \frac{8-3a^2}{3(2a+\pi)}.$$

小問 (ii)

0 より右側の確率は四分円の面積である：

$$\mathrm{P}(X>0) = k\pi = \frac{\pi}{2a+\pi}.$$

考えてみよう $\mathrm{P}(X>0)$ がわかると、なぜ d の符号がわかるのだろうか。

確率密度は $(-a,2)$ 上で正であるから、関数 $t \mapsto \mathrm{P}(X>t)$ は $-a \le t \le 2$ で連続かつ狭義単調減少（strictly decreasing）である。この関数は 1 から 0 まで減少するので、$\mathrm{P}(X>d) = \frac{1}{10}$ を満たす d はちょうど一つ存在する。関数は減少するので、$d<0$ となるのは $\mathrm{P}(X>0) < \frac{1}{10}$ のときであり、かつそのときに限る。言い換えれば、0 より右側にある確率が十分の一未満のとき、d は負になる。

いま $2a > 9\pi$ と仮定する。すると $2a+\pi > 10\pi$ であるから、

$$\mathrm{P}(X>0) = \frac{\pi}{2a+\pi} < \frac{1}{10},$$

よって $d<0$ となり、示すべきことが示された。（この議論は逆向きにもたどれる：$d<0$ ならば $2a>9\pi$ である。）

この場合 $-a \le d < 0$ であるから、d より右側の面積は、長方形のうち d から 0 までの幅 $-d$ の部分と、四分円全体とを合わせたものである：

$$\mathrm{P}(X>d) = 2k(-d) + k\pi = \frac{\pi-2d}{2a+\pi}.$$

これを $\frac{1}{10}$ とおくと $10\pi - 20d = 2a+\pi$ となり、

$$d = \frac{9\pi - 2a}{20}.$$

この値は確かに $[-a,0)$ に属する。$2a>9\pi$ より負であり、また $d \ge -a$ は $9\pi+18a \ge 0$ と同値で、これは成り立つ。

小問 (iii)

今度は $d = \sqrt{2} > 0$ であるから、d は円弧の部分の下にあり、

$$\mathrm{P}(X > \sqrt{2}) = k \int_{\sqrt{2}}^{2} \sqrt{4 - x^2}\,\mathrm{d}x.$$

与えられた結果を用いると、

$$\begin{aligned}\int_{\sqrt{2}}^{2} \sqrt{4 - x^2}\,\mathrm{d}x &= \left[2\arcsin(\tfrac{1}{2}x) + \tfrac{1}{2}x\sqrt{4 - x^2}\right]_{\sqrt{2}}^{2} \\ &= (2 \cdot \tfrac{\pi}{2} + 0) - (2 \cdot \tfrac{\pi}{4} + \tfrac{\sqrt{2}}{2} \cdot \sqrt{2}) \\ &= \frac{\pi}{2} - 1.\end{aligned}$$

（幾何学的には、これは $x = \sqrt{2}$ における弦で切り取られる弓形（circular segment）である。その面積は円板の八分の一 $\frac{\pi}{2}$ から面積 1 の直角三角形を引いたものに等しい。）よって

$$\frac{\frac{\pi}{2} - 1}{2a + \pi} = \frac{1}{10} \quad \Longrightarrow \quad 5\pi - 10 = 2a + \pi \quad \Longrightarrow \quad a = 2\pi - 5.$$

これは正であり（$a \approx 1.283$）、$2a \leq 9\pi$ を満たす。$d > 0$ であるから、これは小問 (ii) の結果と整合する。

検算

- 小問 (i) で $a = 1$ とすると、平均は $\frac{5}{3(2+\pi)} \approx 0.3242$ である。標本数 2×10^5 のモンテカルロ・シミュレーション（Monte Carlo simulation）では 0.3236 が得られた。
- 小問 (ii) の式は境界 $2a = 9\pi$ で $d = 0$ を与え、条件と整合する。$a = 20$ のときは $d = \frac{9\pi - 40}{20} \approx -0.586$ となり、$\mathrm{P}(X > d) = \frac{1}{10}$ を数値的に解いても同じ値が得られる。
- 小問 (iii) で $a = 2\pi - 5$ とすると、シンプソンの公式（Simpson's rule）により $\mathrm{P}(X > \sqrt{2}) = 0.1000$ が得られる。

よくある間違い

- k を未知のまま放置すること。正規化（normalisation）の条件 $2ka + k\pi = 1$ が隠れた最初の一歩であり、すべての答えは a のみで表される。
- 小問 (ii) で四分円を忘れ、$\mathrm{P}(X > d) = -2kd$ だけを書くこと。$[0, 2]$ 上の確率もすべて、負の d より右側にある。
- $\int_{-a}^{0} 2kx\,\mathrm{d}x$ の符号を誤ること。積分区間全体で $x < 0$ であるから、この値は負である。

STEP のその先へ

値 d は X の上側 10% 点、すなわち 0.9 分位点（quantile）である。これを求めることは累積分布関数（cumulative distribution function）F の逆関数を求めることであり、その逆関数こそ、一様分布に従う確率変数から別の確率変数をシミュレートする際に用いられるもの（逆変換法（inverse transform sampling））である。一方では平らで他方では曲がっているというこの確率密度の混合した形は、分位点をしばしば区間ごとに求めなければならない理由を示している。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、分布関数と分位点を丁寧に扱っている。

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。