詳細解答

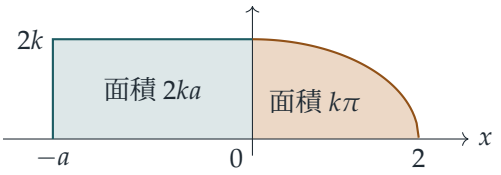
STEP I 2011，第 13 題

概率與統計 • 連續隨機變量

> **STEP I 2011，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/11-s1-q13

關鍵思路

先畫出密度函數（density function）的圖像。它由兩部分拼成：在 $[-a, 0)$ 上是一個高為 $2k$ 的長方形，在 $[0, 2]$ 上則是一個半徑為 2 的四分之一圓（垂直方向按比例 k 縮放）。圖像下方的面積就是概率，因此單靠長方形和四分之一圓，大部分工作已經完成，毋須做任何積分。總面積必須等於 1，由此便可定出 k。之後所有關於 d 的問題，都歸結為同一件事：某一點右方有多少面積？



解答

求 k

在 $[0, 2]$ 上，$\sqrt{4 - x^2}$ 的圖像是圓 $x^2 + y^2 = 4$ 的四分之一，所以

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

（用題目給出的原函數（antiderivative）計算，結果一致：$2\arcsin 1 + 0 - 0 = \pi$。）因此總概率為 $2ka + k\pi = 1$，即

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

第 (i) 部分

把 $\mathrm{E}(X)$ 的積分在 0 處分拆：

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

至於第二部分，$(4-x^2)^{3/2}$ 的導數是 $-3x(4-x^2)^{1/2}$，所以

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

兩部分相加，得

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

第 (ii) 部分

0 右方的概率就是四分之一圓的面積：

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

? **想一想** 知道了 $\mathrm{P}(X>0)$，為何就能判斷 d 的正負？

由於密度函數在 $(-a,2)$ 上為正，函數 $t \mapsto \mathrm{P}(X>t)$ 在 $-a \le t \le 2$ 上連續（continuous）且嚴格遞減（strictly decreasing）。它由 1 降至 0，所以恰好有一個 d 滿足 $\mathrm{P}(X>d) = \frac{1}{10}$。由於函數遞減，$d<0$ 當且僅當 $\mathrm{P}(X>0) < \frac{1}{10}$。換句話說，當 0 右方的概率不足十分之一時，$d$ 便是負數。

現在假設 $2a > 9\pi$，則 $2a+\pi > 10\pi$，所以

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

因此 $d<0$，得證。（這個論證反過來亦成立：只有當 $2a>9\pi$ 時才有 $d<0$。）

在這情況下 $-a \le d < 0$，所以 d 右方的面積，就是長方形由 d 到 0、寬度為 $-d$ 的那一部分，再加上整個四分之一圓：

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

令它等於 $\frac{1}{10}$，得 $10\pi - 20d = 2a+\pi$，所以

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

這個值確實位於 $[-a,0)$ 之內：因為 $2a>9\pi$，所以它是負數；而 $d \ge -a$ 等價於 $9\pi + 18a \ge 0$，這顯然成立。

第 (iii) 部分

現在 $d = \sqrt{2} > 0$，所以 d 落在圓弧部分之下，而

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

利用題目給出的結果，

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

（從幾何角度看，這是被 $x = \sqrt{2}$ 處的弦（chord）割出的弓形（circular segment），它等於圓面積的八分之一，即 $\frac{\pi}{2}$，減去一個面積為 1 的直角三角形。）所以

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

這是正數（$a \approx 1.283$），並且滿足 $2a \leq 9\pi$；由於 $d > 0$，這與第 (ii) 部分的結論吻合。

驗算

- 在第 (i) 部分，取 $a = 1$，均值為 $\frac{5}{3(2+\pi)} \approx 0.3242$。以 2×10^5 個樣本進行蒙地卡羅模擬（Monte Carlo simulation），得出 0.3236。
- 在臨界情況 $2a = 9\pi$ 下，第 (ii) 部分的公式給出 $d = 0$，與條件一致。取 $a = 20$，公式給出 $d = \frac{9\pi - 40}{20} \approx -0.586$，而以數值方法解 $\mathrm{P}(X > d) = \frac{1}{10}$ 亦得出相同的值。
- 在第 (iii) 部分，取 $a = 2\pi - 5$，以辛普森法則（Simpson's rule）計算得 $\mathrm{P}(X > \sqrt{2}) = 0.1000$。

常見錯誤

- 沒有求出 k。歸一化（normalisation）條件 $2ka + k\pi = 1$ 是隱藏的第一步，而所有答案都只應以 a 表示。
- 在第 (ii) 部分漏掉四分之一圓，只寫 $\mathrm{P}(X > d) = -2kd$。其實 $[0, 2]$ 上的全部概率，同樣位於負數 d 的右方。
- 把 $\int_{-a}^{0} 2kx\,\mathrm{d}x$ 的正負號弄錯。由於整個區間上 $x < 0$，這個積分是負數。

STEP 之外

d 這個值就是 X 的上 10% 點，即 0.9 分位數（quantile）。要求出它，就要求累積分佈函數（cumulative distribution function）F 的反函數；而這個反函數，正正是用均勻分佈的隨機變量去模擬另一個隨機變量時所用的工具（逆變換抽樣（inverse transform sampling））。這個密度函數形狀混合，一邊平坦、一邊彎曲，說明了為何分位數往往要分段求出。Grimmett 和 Welsh 所著的 *Probability: An Introduction*（《概率導論》）對分佈函數和分位數有仔細的論述。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。