詳細解答

STEP I 1987，第 15 題

概率與統計 • 連續隨機變量

STEP I 1987，第 15 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/87-s1-q15

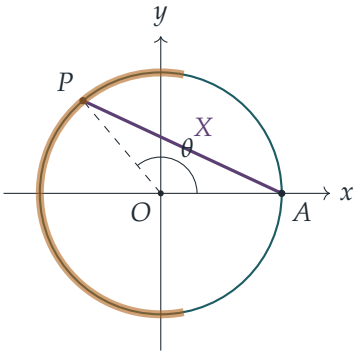
關鍵思路

「在圓上均勻分佈（uniformly distributed）」的意思是 P 在圓心所成的角是均勻分佈的，所以自然的做法是設 $P = (\cos\theta, \sin\theta)$，其中 θ 在 $[0, 2\pi)$ 上均勻分佈。這樣，與 $(1,0)$ 的距離就是單位圓的一條弦（chord），而半角公式（half-angle formula）可把它化成單一個正弦。之後，平均值（mean）只需一個積分，方差（variance）幾乎完全不用積分，而所求概率則變成一個關於角度區間的問題。

解答

設定

設 θ 為由正 x 軸到 OP 的角，其中 O 為原點。在圓上均勻地選取 P，即是均勻地選取弧長（arc length），所以 θ 在 $[0, 2\pi)$ 上均勻分佈，概率密度（density）為 $\frac{1}{2\pi}$。記 $A = (1,0)$ 及 $P = (\cos\theta, \sin\theta)$。



於是

$$X^2 = (\cos\theta - 1)^2 + \sin^2\theta = 2 - 2\cos\theta = 4\sin^2\tfrac{\theta}{2},$$

這裏用了 $1 - \cos\theta = 2\sin^2\frac{\theta}{2}$。由於 $0 \le \frac{\theta}{2} < \pi$，我們有 $\sin\frac{\theta}{2} \ge 0$，所以取正平方根是穩妥的：

$$X = 2\sin\tfrac{\theta}{2}.$$

平均值

$$\mathrm{E}(X) = \frac{1}{2\pi}\int_0^{2\pi} 2\sin\tfrac{\theta}{2}\,\mathrm{d}\theta = \frac{1}{2\pi}\Big[-4\cos\tfrac{\theta}{2}\Big]_0^{2\pi} = \frac{1}{2\pi}(4+4) = \frac{4}{\pi}.$$

所以 $\mathrm{E}(X) = \dfrac{4}{\pi} \approx 1.273$。

方差

想一想 要求 $\mathrm{E}(X^2)$，我們需要對 $4\sin^2\frac{\theta}{2}$ 積分嗎？

不需要：未化成平方的形式更容易處理。$X^2 = 2 - 2\cos\theta$，而 $\cos\theta$ 在整整一圈上的平均值為零，因為 $\int_0^{2\pi}\cos\theta\,\mathrm{d}\theta = 0$。因此 $\mathrm{E}(X^2) = 2$，而

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = 2 - \frac{16}{\pi^2} \approx 0.379.$$

X 大於其平均值的概率

我們需要 $2\sin\frac{\theta}{2} > \frac{4}{\pi}$，即 $\sin\frac{\theta}{2} > \frac{2}{\pi}$。留意 $0 < \frac{2}{\pi} < 1$，所以此不等式有解。設 $\alpha = \arcsin\frac{2}{\pi}$，它位於 $(0, \frac{\pi}{2})$ 內。當 $\frac{\theta}{2}$ 取遍 $[0,\pi)$ 時，正弦值由 0 升至 1 再回落，而它大於 $\sin\alpha$ 當且僅當

$$\alpha < \tfrac{\theta}{2} < \pi - \alpha, \qquad \text{即} \qquad 2\alpha < \theta < 2\pi - 2\alpha.$$

這就是圖中加粗的圓弧：圓上位於 A 對面一側的點。由於 θ 均勻分佈，所求概率就是此區間的長度除以 2π：

$$\mathrm{P}(X > \mathrm{E}(X)) = \frac{2\pi - 4\alpha}{2\pi} = \boxed{1 - \frac{2}{\pi}\arcsin\frac{2}{\pi}} \approx 0.561.$$

驗算

- X 的取值範圍是 $[0,2]$，而平均值 $4/\pi \approx 1.27$ 位於中點 1 之上。這與圖像吻合：在圓上頗大的一部分，弦長都接近其最大值（當 P 經過對面一側時，X 變化緩慢），所以分佈偏向較大的值。

- 基於同樣原因，大於平均值的概率應略大於 $\frac{1}{2}$，而 0.561 正是如此。
- 在數值上，以中點法則（midpoint rule）對 2×10^5 個等距角度取平均，得到平均值 1.273240、方差 0.378861 及 $\mathrm{P} = 0.56066$；而一個 10^6 點的模擬（simulation）在抽樣誤差範圍內亦與此一致。

常見錯誤

- 取 $\theta \in (-\pi, \pi]$ 而仍寫 $X = 2\sin\frac{\theta}{2}$。此時 $\sin\frac{\theta}{2}$ 可以是負數，你需要寫 $X = 2\left|\sin\frac{\theta}{2}\right|$；使用 $[0, 2\pi)$ 就可避免絕對值（modulus）。
- 把 P 的 x 坐標當作在 $[-1, 1]$ 上均勻分佈。在圓上均勻分佈是指在角度（弧長）上均勻，而不是在 x 上均勻。

STEP 之外

這是隨機變量變換（transformation of a random variable）的第一個例子：X 的概率密度函數是 $[0, 2]$ 上的 $f(x) = \dfrac{2}{\pi\sqrt{4 - x^2}}$，屬於反正弦（arcsine）類型的分佈，而上面每一個答案都可以由它重新推導出來。這道題亦暗示了「隨機」一詞可以多麼難以捉摸：以不同的自然方式隨機選取圓的一條弦，會得出不同的答案（貝特朗悖論（Bertrand's paradox））。這兩個概念在 Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）中都有進一步發展。

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。