详细解答

STEP III 1992，第 15 题

概率与统计 · 几何概率

> **STEP III 1992，第 15 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/92-s3-q15

关键思路

由于山羊的位置在田地上均匀分布（uniformly distributed），每一个概率都只是某块面积除以 a^2。所以整道题归根究底，就是要求出正方形内与角 O 的距离小于 r 的那部分的面积。当 r 超过 a 之后，四分之一圆会溢出正方形以外，而要求出仍留在正方形内的面积，最利落的做法是把它切成两个直角三角形和一个扇形（sector）。之后把分布函数微分便得到密度，而系绳的山羊则是同一个想法，只是换了一个区域。

解答

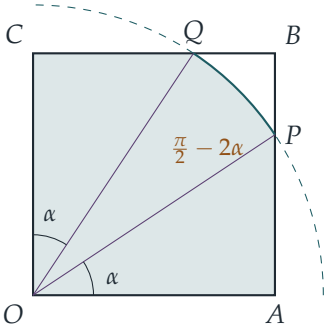
把 O 放在原点，OA 沿 x 轴，OC 沿 y 轴，于是田地就是 $0 \le x \le a$、$0 \le y \le a$。记 $F(r) = \mathrm{P}(R < r)$，它等于田地中满足 $x^2 + y^2 < r^2$ 的那部分的面积除以 a^2。

R 的分布

当 $0 < r \le a$ 时。以 O 为圆心、半径为 r 的四分之一圆盘完全落在田地之内，所以

$$F(r) = \frac{\frac{1}{4}\pi r^2}{a^2} = \frac{\pi r^2}{4a^2}.$$

当 $a \le r \le a\sqrt{2}$ 时。这时圆弧与边 AB（即直线 $x = a$）相交于 $P = (a, \sqrt{r^2 - a^2})$，与边 CB（即直线 $y = a$）相交于 $Q = (\sqrt{r^2 - a^2}, a)$。设 $\alpha = \cos^{-1}(a/r)$，它就是角 AOP（因为 $\cos \angle AOP = a/r$）。根据对称性，角 COQ 也等于 α。



阴影区域（即田地中与 O 的距离小于 r 的部分）可分成三块：

- 三角形 OAP，直角在 A，两条直角边分别为 a 和 $\sqrt{r^2 - a^2}$，所以面积为 $\frac{1}{2}a\sqrt{r^2 - a^2}$；
- 三角形 OCQ，与它全等（congruent），面积相同；
- 扇形 OPQ，半径为 r，圆心角为 $\frac{\pi}{2} - 2\alpha$，面积为 $\frac{1}{2}r^2\left(\frac{\pi}{2} - 2\alpha\right) = \frac{\pi r^2}{4} - r^2\alpha$。

把三者相加再除以 a^2，

$$\begin{aligned} F(r) &= \frac{a\sqrt{r^2 - a^2}}{a^2} + \frac{\pi r^2}{4a^2} - \frac{r^2}{a^2}\,\alpha \\ &= \left(\frac{r^2}{a^2} - 1\right)^{\frac{1}{2}} + \frac{\pi r^2}{4a^2} - \frac{r^2}{a^2}\cos^{-1}\left(\frac{a}{r}\right), \end{aligned}$$

这正是所求的结果。（当 $r \geq a\sqrt{2}$ 时，整块田地与 O 的距离都小于 r，所以 $F(r) = 1$。在 $r = a\sqrt{2}$ 处，公式给出 $1 + \frac{\pi}{2} - 2 \cdot \frac{\pi}{4} = 1$，与预期相符。）

中位数

在 $r = a$ 处有 $F(a) = \pi/4 \approx 0.785 > \frac{1}{2}$，所以中位数（median）$m$ 落在 $0 < m < a$ 的范围内，这里适用的是较简单的那条公式：

$$\frac{\pi m^2}{4a^2} = \frac{1}{2} \quad \Longrightarrow \quad m = a\sqrt{\frac{2}{\pi}} \approx 0.798a.$$

概率密度函数

概率密度函数（probability density function）为 $f(r) = F'(r)$。当 $0 < r < a$ 时，$f(r) = \dfrac{\pi r}{2a^2}$。

当 $a < r < a\sqrt{2}$ 时，我们需要 $\cos^{-1}(a/r)$ 的导数：

$$\frac{\mathrm{d}}{\mathrm{d}r}\cos^{-1}\left(\frac{a}{r}\right) = -\frac{1}{\sqrt{1 - a^2/r^2}} \cdot \left(-\frac{a}{r^2}\right) = \frac{a}{r\sqrt{r^2 - a^2}}.$$

然后对 F 逐项微分（最后一项要用积法则（product rule）），

$$\begin{aligned} f(r) &= \frac{r}{a\sqrt{r^2-a^2}} + \frac{\pi r}{2a^2} - \frac{2r}{a^2}\cos^{-1}\left(\frac{a}{r}\right) - \frac{r^2}{a^2}\cdot\frac{a}{r\sqrt{r^2-a^2}} \\ &= \frac{r}{a^2}\left(\frac{\pi}{2} - 2\cos^{-1}\left(\frac{a}{r}\right)\right), \end{aligned}$$

因为第一项与最后一项互相抵消。所以

$$f(r) = \begin{cases} \dfrac{\pi r}{2a^2}, & 0 < r \le a, \\ \dfrac{r}{a^2}\left(\dfrac{\pi}{2} - 2\cos^{-1}\left(\dfrac{a}{r}\right)\right), & a < r \le a\sqrt{2}, \\ 0, & \text{其他情况。} \end{cases}$$

这在几何上也说得通：$f(r)\,\mathrm{d}r$ 等于（半径为 r 的圆在田地内的弧长）$\times\,\mathrm{d}r/a^2$，而该段弧所对的圆心角在 $r \le a$ 时为 $\frac{\pi}{2}$，在 $r > a$ 时为 $\frac{\pi}{2} - 2\alpha$。两段公式在 $r = a$ 处吻合（都给出 $\pi/(2a)$），而 f 在 $r = a\sqrt{2}$ 处降至 0，那里 $\alpha = \pi/4$。

系绳的山羊

山羊被系住之后，只能到达田地中 $R \le a$ 的点，即以 O 为圆心、半径为 a 的四分之一圆盘，其面积为 $\frac{1}{4}\pi a^2$。牠的位置仍然是均匀分布，只是现在分布在这个四分之一圆盘上。这恰好就是原来的模型以事件 $R \le a$ 为条件之后的情况。

> ? **想一想** 哪一个坐标量度的是与围栏 OC 的距离?

由于 OC 沿 y 轴，与 OC 的距离就是 x。所以我们需要求出四分之一圆盘中 $x > a/2$ 的那部分的面积：

$$\begin{aligned} \int_{a/2}^{a} \sqrt{a^2-x^2}\,\mathrm{d}x &= \left[\frac{x\sqrt{a^2-x^2}}{2} + \frac{a^2}{2}\sin^{-1}\frac{x}{a}\right]_{a/2}^{a} \\ &= \frac{\pi a^2}{4} - \left(\frac{\sqrt{3}\,a^2}{8} + \frac{\pi a^2}{12}\right) = \frac{\pi a^2}{6} - \frac{\sqrt{3}\,a^2}{8}. \end{aligned}$$

（从几何上看，这是一个圆心角为 $\pi/3$ 的扇形减去一个边长为 $a/2$ 和 $\sqrt{3}a/2$ 的三角形，因为直线 $x = a/2$ 与圆弧的交点和 OA 所成的角为 $\pi/3$。）除以面积 $\frac{1}{4}\pi a^2$，

$$\mathrm{P}\left(\text{距}OC > \tfrac{a}{2} \,\middle|\, R \le a\right) = \frac{2}{3} - \frac{\sqrt{3}}{2\pi} \approx 0.391.$$

验算

- 连续性：F 的两条公式在 $r = a$ 处都给出 $\pi/4$，而且 $F(a\sqrt{2}) = 1$。

- 以单位正方形内 4×10^5 个均匀分布的点进行蒙地卡罗模拟（Monte Carlo simulation），结果在 $r = 0.3, 0.7, 1.1, 1.25, 1.4$ 处与 $F(r)$ 吻合至大约小数点后三位。模拟得出的样本中位数为 0.7986，而 $\sqrt{2/\pi} = 0.7979$；条件比例为 0.3914，而 $\frac{2}{3} - \frac{\sqrt{3}}{2\pi} = 0.3910$。
- 密度在 $[0, a\sqrt{2}]$ 上的数值积分为 1，并且在若干个点上与 F 的数值导数相符。
- 系绳情况的答案小于 $\frac{1}{2}$，这是应有的结果：四分之一圆盘在远离 OC 的地方较窄。

常见错误

- 求中位数时忘记了 F 在 $r > a$ 时要用另一条公式。值得先检查 $F(a) > \frac{1}{2}$，由此可知中位数落在容易处理的范围内。
- 把扇形的圆心角写成 $\frac{\pi}{2} - \alpha$ 而不是 $\frac{\pi}{2} - 2\alpha$。被切去的角位有两块，两边各一。
- 在系绳的部分除以 a^2。条件概率（conditional probability）应除以四分之一圆盘的面积 $\frac{1}{4}\pi a^2$。

STEP 之外

这里的 R 是在某个区域内均匀选取的一点的函数，而它的分布是通过量度 $R < r$ 的区域而得出的。这正是求随机变量（random variable）的函数的分布的一般方法，而密度公式 $f(r) = (\text{弧长})/a^2$ 则让人初尝“余面积”（coarea）的想法，即把一个区域切成一片片水平集（level sets）。同样的几何推理也是布丰投针（Buffon's needle）以及正方形内两个随机点之间的距离等问题的基础。Grimmett 与 Welsh 的 *Probability: An Introduction*，以及 Blitzstein 与 Hwang 的 *Introduction to Probability*，都有阐述这种处理连续随机变量的几何方法。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。