详细解答

STEP I 1999，第 14 题

概率与统计 • 条件概率

STEP I 1999，第 14 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/99-s1-q14

关键思路

这个连续随机变量（continuous random variable）只是用来引出一个离散（discrete）随机变量。只要把密度函数（density）化成得 4、3、2、1 分的四个概率，第 (i) 部分就只是一行的期望值（expectation）计算。第 (ii) 部分是条件概率（conditional probability）；由于总分必须刚好等于 3，获胜的方式寥寥可数，我们可以把它们全部列出。

解答

第 (i) 部分

对于 $0 \le a \le b \le 1$，

$$\mathrm{P}(a \le X < b) = \int_a^b 2x \,\mathrm{d}x = b^2 - a^2.$$

（对连续随机变量而言，包括或不包括端点并无分别。）设 S 为掷一支飞镖所得的分数，则

$$\mathrm{P}(S=4) = \left(\tfrac{1}{4}\right)^2 - 0 = \tfrac{1}{16}, \qquad \mathrm{P}(S=3) = \left(\tfrac{1}{2}\right)^2 - \left(\tfrac{1}{4}\right)^2 = \tfrac{3}{16},$$

$$\mathrm{P}(S=2) = \left(\tfrac{3}{4}\right)^2 - \left(\tfrac{1}{2}\right)^2 = \tfrac{5}{16}, \qquad \mathrm{P}(S=1) = 1 - \left(\tfrac{3}{4}\right)^2 = \tfrac{7}{16}.$$

这四个概率之和为 1，正如所料。留意外围的环较容易掷中：密度 $2x$ 随 x 增大，因为离中心愈远，面积愈大。

所以

$$\mathrm{E}(S) = 4 \cdot \tfrac{1}{16} + 3 \cdot \tfrac{3}{16} + 2 \cdot \tfrac{5}{16} + 1 \cdot \tfrac{7}{16} = \frac{4+9+10+7}{16} = \frac{30}{16} = \frac{15}{8},$$

即为所求。

第 (ii) 部分

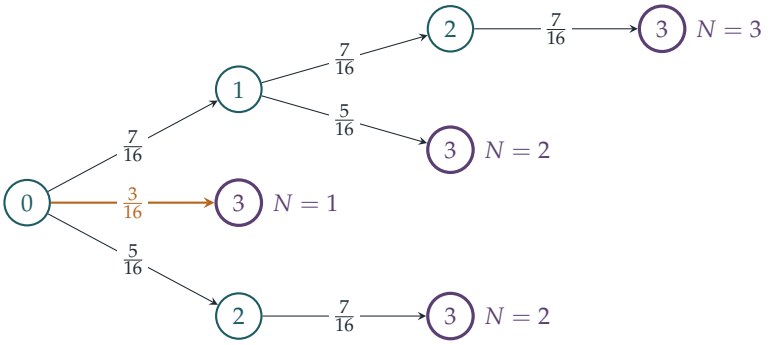
假设各次投掷互相独立（independent），每次的分数都依从上述 S 的分布。设 W 为我获胜的事件，N 为这局游戏的投掷次数。

> **想一想** 哪些得分序列会导致获胜?

我获胜，当且仅当累积总分达到 3 而之前没有超过它。每次得分至少为 1，所以总分每次至少增加 1，游戏最多只会掷三次。一局获胜的游戏，就是一个由正数得分组成、总和刚好为 3 的序列。（只有当总分低于 3 时我才会再掷，所以每个这样的序列确实是一局完整的游戏。）这样的序列只有四个:

$$(3), \qquad (1,2), \qquad (2,1), \qquad (1,1,1).$$

得分 4 永远不会出现，因为它会令总分立即超过 3。



树形图显示每条以获胜告终的路线上的累积总分（获胜状态以紫色圆圈标示），箭头上标示每次投掷的概率。

根据独立性，每个序列的概率等于其中各个得分的概率之积:

$$\begin{aligned}
\mathrm{P}(W \cap N = 1) &= \mathrm{P}(3) = \tfrac{3}{16} = \tfrac{768}{4096}, \\
\mathrm{P}(W \cap N = 2) &= \mathrm{P}(1,2) + \mathrm{P}(2,1) = 2 \cdot \tfrac{7}{16} \cdot \tfrac{5}{16} = \tfrac{70}{256} = \tfrac{1120}{4096}, \\
\mathrm{P}(W \cap N = 3) &= \mathrm{P}(1,1,1) = \left(\tfrac{7}{16}\right)^3 = \tfrac{343}{4096}.
\end{aligned}$$

这三个事件互斥（disjoint），合起来就是 W，所以

$$\mathrm{P}(W) = \frac{768 + 1120 + 343}{4096} = \frac{2231}{4096}.$$

因此

$$\mathrm{P}(N = 3 \mid W) = \frac{\mathrm{P}(W \cap N = 3)}{\mathrm{P}(W)} = \frac{343/4096}{2231/4096} = \frac{343}{2231},$$

即为所求。

验算

- 四个得分概率 $\frac{1}{16}, \frac{3}{16}, \frac{5}{16}, \frac{7}{16}$ 之和为 1。它们都是奇数除以 16，因为 $(k+1)^2 - k^2 = 2k+1$。
- 目标答案中的 $343 = 7^3$ 一看便知，掷三次而获胜的情况必定是 $(1,1,1)$，这印证了我们的设置。
- 以电脑模拟 10^6 局游戏，得出 $\mathrm{P}(W) \approx 0.5452$（准确值为 0.5447）及 $\mathrm{P}(N=3 \mid W) \approx 0.1535$（准确值为 $343/2231 \approx 0.1537$）。

常见错误

- 把各个得分区当作等可能。概率来自对密度函数求积分，而不是来自区间的阔度。
- 计算了 $(1,2)$ 却忘记 $(2,1)$，结果得出 $\mathrm{P}(W) = \frac{1671}{4096}$。
- 答案写成 $\mathrm{P}(N=3 \text{ 及} W) = \frac{343}{4096}$，而没有除以 $\mathrm{P}(W)$。

STEP 之外

第 (ii) 部分其实是一条小型的马可夫链（Markov chain）：状态就是累积总分，而 3 和“爆分”是吸收态（absorbing states）。目标分数较大时，逐一列出路线便不再可行。这时我们改为以 w_t 表示由总分 t 出发而获胜的概率，从而得到递推关系（recurrence）$w_t = \sum_s \mathrm{P}(S=s)\, w_{t+s}$。这种首步分析（first-step analysis）是大学一年级概率论的重要工具。Grimmett 与 Welsh 所著的 *Probability: An Introduction*（概率论入门教科书）在讨论随机漫步（random walks）和马可夫链的部分对此有所讲解。

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。