详细解答

STEP III 2025，第 12 题

概率与统计 • 递归随机变量

STEP III 2025，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2025 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/750486-2025-past-papers.zip

关键思路

第 (i) 部分其实只是把一个二重和（double sum）用两种方式写出来：两边都是把 $\mathrm{f}(r)\,\mathrm{g}(s)$ 在同一组数对 (r,s) 上加起来，一边逐行计算，另一边逐列计算。在第 (ii) 部分，以前一个值 X_{n-1} 为条件（conditioning），可把每个概率 $\mathrm{P}(X_n = r)$ 写成恰好是 (i) 左边那种形式的和。交换求和次序后，$\mathrm{E}(X_n)$ 就变成一个能以 $\mathrm{E}(X_{n-1})$ 表示的式子。

解答

第 (i) 部分

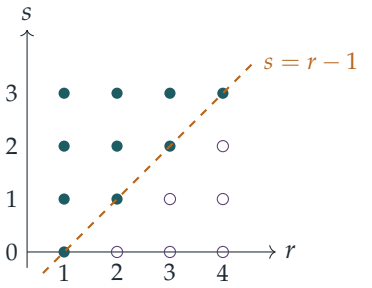
把括号展开，左边是

$$\sum_{r=1}^{m+1}\sum_{s=r-1}^{m}\mathrm{f}(r)\,\mathrm{g}(s),$$

即 $\mathrm{f}(r)\,\mathrm{g}(s)$ 在所有满足

$$1\le r\le m+1 \quad 及 \quad r-1\le s\le m.$$

的整数对 (r,s) 上的和。我们先固定 s 来描述同一组数对。这些条件迫使 $0\le s\le m$（因为 $s\ge r-1\ge 0$），而对于这样的某个 s，容许的 r 是满足 $1\le r$ 及 $r\le s+1$ 的那些；注意 $s+1\le m+1$，所以界限 $r\le m+1$ 自动成立。反过来，任何满足 $0\le s\le m$ 及 $1\le r\le s+1$ 的数对都符合原来的条件。下图显示 $m=3$ 时的三角形数对区域。



按列求和（固定 r）得到左边；按行求和（固定 s）则得到

$$\sum_{s=0}^{m}\sum_{r=1}^{s+1} \mathrm{f}(r)\,\mathrm{g}(s) = \sum_{s=0}^{m}\Big(\mathrm{g}(s)\sum_{r=1}^{s+1}\mathrm{f}(r)\Big),$$

这正是右边，得证。由于这些都是有限和，重新排列求和次序总是容许的。

第 (ii)(a) 部分

X_1 取 0 或 1 的机会均等，所以 $\mathrm{E}(X_1) = \frac{1}{2}$。

至于 X_2，以 X_1 为条件。若 $X_1 = 0$，则 X_2 在 $\{0,1\}$ 上均匀分布（uniform）；若 $X_1 = 1$，则 X_2 在 $\{0,1,2\}$ 上均匀分布。由全概率定律（law of total probability），

$$\begin{aligned}
\mathrm{P}(X_2 = 0) &= \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{3} = \tfrac{5}{12},\\
\mathrm{P}(X_2 = 1) &= \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{3} = \tfrac{5}{12},\\
\mathrm{P}(X_2 = 2) &= \tfrac{1}{2}\cdot 0 + \tfrac{1}{2}\cdot\tfrac{1}{3} = \tfrac{1}{6},
\end{aligned}$$

如题所求（三者之和为 1，这是个有用的验算）。因此

$$\mathrm{E}(X_2) = 0\cdot\tfrac{5}{12} + 1\cdot\tfrac{5}{12} + 2\cdot\tfrac{1}{6} = \tfrac{3}{4}.$$

第 (ii)(b) 部分

由归纳法（induction），X_{n-1} 的取值在 $\{0,1,\ldots,n-1\}$ 之中（每一步把可能的最大值提高一）。已知 $X_{n-1} = s$ 时，X_n 在 $s+2$ 个值 $0,1,\ldots,s+1$ 上均匀分布，所以

$$\mathrm{P}(X_n = r \mid X_{n-1} = s) = \begin{cases} \dfrac{1}{s+2}, & 0 \le r \le s+1,\\[2mm] 0, & r > s+1. \end{cases}$$

值 $r = 0$ 总是可能出现的，所以全概率定律给出

$$\mathrm{P}(X_n = 0) = \sum_{s=0}^{n-1} \frac{\mathrm{P}(X_{n-1} = s)}{s+2}.$$

对于 $1 \le r \le n$，只有满足 $s+1 \ge r$，即 $s \ge r-1$ 的那些 s 才有贡献：

$$\mathrm{P}(X_n = r) = \sum_{s=r-1}^{n-1} \frac{\mathrm{P}(X_{n-1}=s)}{s+2}, \qquad r = 1, 2, \ldots, n.$$

（当 $r=1$ 时，这与 $r=0$ 的和相同，所以对每个 $n \ge 1$ 都有 $\mathrm{P}(X_n=0) = \mathrm{P}(X_n=1)$，与 (a) 中两个 $\frac{5}{12}$ 吻合。）

第 (ii)(c) 部分

记 $p_s = \mathrm{P}(X_{n-1}=s)$。$r=0$ 的项对期望值（expectation）没有贡献，所以由 (b)，

$$\mathrm{E}(X_n) = \sum_{r=1}^{n} r\,\mathrm{P}(X_n=r) = \sum_{r=1}^{n}\Big(r \sum_{s=r-1}^{n-1} \frac{p_s}{s+2}\Big).$$

? **想一想** 第 (i) 部分在这里如何应用?

这就是 (i) 的左边，其中 $m = n-1$（因 $n \ge 1$，这是容许的）、$\mathrm{f}(r) = r$ 及 $\mathrm{g}(s) = p_s/(s+2)$。交换求和次序，

$$\begin{aligned}
\mathrm{E}(X_n) &= \sum_{s=0}^{n-1}\Big(\frac{p_s}{s+2}\sum_{r=1}^{s+1} r\Big) = \sum_{s=0}^{n-1} \frac{p_s}{s+2}\cdot\frac{(s+1)(s+2)}{2} \\
&= \frac{1}{2}\sum_{s=0}^{n-1}(s+1)\,p_s = \frac{1}{2}\Big(\sum_{s=0}^{n-1} p_s + \sum_{s=0}^{n-1} s\,p_s\Big) \\
&= \tfrac{1}{2}\big(1 + \mathrm{E}(X_{n-1})\big),
\end{aligned}$$

这里用了 $\sum p_s = 1$。这就是所求的结果。

直观而言：已知 $X_{n-1}=s$ 时，在 $\{0,\ldots,s+1\}$ 上均匀分布的变量的平均值是 $\frac{1}{2}(s+1)$，所以这其实是乔装了的塔性质（tower law）。

要解这个递推关系（recurrence relation），记 $e_n = \mathrm{E}(X_n)$。两边同时减去 1，

$$e_n - 1 = \tfrac{1}{2}\,(e_{n-1} - 1),$$

所以 $e_n - 1$ 每一步减半。由于 $e_0 = 0$，得 $e_n - 1 = -\big(\frac{1}{2}\big)^n$，即

$$\mathrm{E}(X_n) = 1 - \frac{1}{2^n}, \qquad n = 1, 2, \ldots.$$

验算

此公式给出 $\mathrm{E}(X_1) = \frac{1}{2}$ 及 $\mathrm{E}(X_2) = \frac{3}{4}$，与 (a) 一致。逐步计算 X_n 的精确分布直至 $n=15$，证实了 (b) 中的表达式及 $\mathrm{E}(X_n) = 1 - 2^{-n}$；例如 $\mathrm{E}(X_5) = \frac{31}{32} \approx 0.969$，而以电脑模拟 200 000 次得到 0.967。当 $n \to \infty$ 时，平均值趋向 1，这是合理的：每一步都把平均值向 1 拉近一半距离。

常见错误

- 把 $\mathrm{P}(X_n = r)$ 的和从 $s = r$ 而非 $s = r - 1$ 开始：由于 X_n 最大可达 $X_{n-1} + 1$，值 r 可以从 $s = r - 1$ 到达。
- 在 (c) 中交换次序时弄错上下限：交换后，r 由 1 取到 $s + 1$，得 $\frac{1}{2}(s+1)(s+2)$，而因子 $s + 2$ 会整齐地约去。
- 以 e_1 作为起始值解递推关系，却写成 $1 - 2^{-(n-1)}$；应以 $e_1 = \frac{1}{2}$ 核对。

STEP 之外

这里的过程是一条马尔可夫链（Markov chain）：X_n 的分布只取决于 X_{n-1}，而 (b) 就是把概率向量（probability vector）乘以转移矩阵（transition matrix）的一步。(c) 中的捷径是全期望定律（law of total expectation），$\mathrm{E}(X_n) = \mathrm{E}\big(\mathrm{E}(X_n \mid X_{n-1})\big)$，它能把许多这类问题化为线性递推关系。这两个概念在 Grimmett 与 Welsh 所著的 *Probability: An Introduction*（《概率导论》）中都有进一步发展。

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。