详细解答

STEP III 1988，第 15 题

概率与统计 • 期望值的递推关系

> **STEP III 1988，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/88-s3-q15

关键思路

对当天的第一本书取条件（conditioning）。这本书一旦上了架，当天余下的过程就与书架较短的全新一天完全一样，所以长度为 n 的书架的期望（expected）本数，与长度为 $n-1$ 和 $n-2$ 时的期望本数有关。由此得到一条线性差分方程（difference equation），解法与二阶递推关系（recurrence）一样：齐次（homogeneous）部分用辅助二次方程（auxiliary equation）处理，再为右边的常数 1 找一个特解（particular solution）。

解答

以下一律设 N_n 为书架长度为 n 的一天内放上书架的书本数目，因此 $M(n) = \mathrm{E}(N_n)$。各本书是厚是薄互相独立（independent），每本为厚书的概率是 p。题目说明书本数量总是多于书架所能容纳的，所以每天最终必定有某本书放不下。

第 (i) 部分

当 $n = 0$ 时，书架完全没有空位。第一本书不论厚薄都放不下，所以它和其后每一本书都要放到手推车上。因此 $N_0 = 0$ 恒成立，而 $M(0) = 0$。

第 (ii) 部分

当 $n = 1$ 时，考虑第一本书。

- 若它是薄书（概率为 q），它放得下，并刚好放满书架。下一本书至少需要一个单位，所以放不下，自此所有书都放到手推车上。故 $N_1 = 1$。

- 若它是厚书（概率为 p），它需要两个单位，放不下，故 $N_1 = 0$。

因此 $M(1) = q \cdot 1 + p \cdot 0 = q$。

第 (iii) 部分

设 $n \geq 2$, 同样对第一本书取条件。不论它是哪一种书都放得下, 因为它最多只需要 $2 \leq n$ 个单位。

- 若它是薄书（概率为 q），它上架后余下 $n-1$ 个单位。
- 若它是厚书（概率为 p），它上架后余下 $n-2$ 个单位。

> ? **想一想** 为甚么在第一种情况下，其后上架书本数目的期望值恰好是 $M(n-1)$?

第一本书之后，其余的书依次到来，它们是厚是薄与第一本书独立，概率亦相同，而且按同一规则摆放：一直上架，直到有一本放不下为止。第一本书唯一改变了的是剩余空间，现在是 $n-1$ 个单位。所以当天余下的过程，就是书架长度为 $n-1$ 的完整一天的翻版，其后上架书本数目的期望值是 $M(n-1)$。同理，若第一本是厚书，这个期望值就是 $M(n-2)$。（根据第 (i) 部分，当 $n-2=0$ 时这也成立。）

把第一本书也计算在内，由全期望定律（law of total expectation）得

$$\begin{aligned} M(n) &= q\big(1 + M(n-1)\big) + p\big(1 + M(n-2)\big) \\ &= (p+q) + qM(n-1) + pM(n-2), \end{aligned}$$

而由于 $p+q=1$,

$$M(n) - qM(n-1) - pM(n-2) = 1 \qquad (n \geq 2),$$

这正是所要证明的。

闭合公式

先考虑齐次方程 $M(n) - qM(n-1) - pM(n-2) = 0$。代入 $M(n) = \lambda^n$ 试解，得辅助方程

$$\lambda^2 - q\lambda - p = 0.$$

由于 $q = 1-p$，$\lambda = 1$ 是一个根（$1 - q - p = 0$），而两根之积为 $-p$，所以

$$\lambda^2 - q\lambda - p = (\lambda - 1)(\lambda + p).$$

因此对任意 A 和 B，$A(-p)^n$ 和常数 B 都满足齐次方程。

因为 1 是一个根，常数无法产生右边的 1，所以我们改试 Cn。代入得

$$\begin{aligned} Cn - qC(n-1) - pC(n-2) &= C\big[(1-q-p)n + q + 2p\big] \\ &= C(q+2p) = C(1+p). \end{aligned}$$

当 $C = \dfrac{1}{1+p}$ 时，上式等于 1；由于 $1+p > 0$，这个取值没有问题。于是由线性性质（linearity），不论 A 和 B 取何值，$M(n) = A(-p)^n + B + Cn$ 对所有 $n \geq 2$ 都满足 (iii) 中的方程。

现在代入初始值。由第 (i) 部分得

$$M(0) = A + B = 0, \qquad \text{所以} B = -A,$$

而由第 (ii) 部分得

$$M(1) = -Ap + B + C = -A(1+p) + \frac{1}{1+p} = 1 - p.$$

因此

$$A(1+p) = \frac{1}{1+p} - (1-p) = \frac{1-(1-p^2)}{1+p} = \frac{p^2}{1+p},$$

所以

$$A = \frac{p^2}{(1+p)^2}, \qquad B = -\frac{p^2}{(1+p)^2}, \qquad C = \frac{1}{1+p}.$$

即

$$M(n) = \frac{n}{1+p} - \frac{p^2}{(1+p)^2}\big(1 - (-p)^n\big).$$

由于 (i)、(ii) 和 (iii) 确定了 $M(0)$、$M(1)$，继而逐一确定其后每一个值，所以这不仅是一个可能的解，更是唯一的解。

验算

- $p = 0$（全是薄书）：$M(n) = n$，理应如此。
- $p = 1$（全是厚书）：$M(n) = \frac{n}{2} - \frac{1}{4}\big(1 - (-1)^n\big)$，当 n 为偶数时是 $\frac{n}{2}$，当 n 为奇数时是 $\frac{n-1}{2}$，即 $\lfloor n/2 \rfloor$。正确。
- 取 $p = q = \frac{1}{2}$，直接由递推关系计算：$M(2) = 1 + \frac{1}{4} = \frac{5}{4}$，而公式给出 $\frac{2}{3} - \frac{1}{9} \cdot \frac{3}{4} = \frac{5}{4}$。
- 当 n 很大时，每本书平均占用 $1+p$ 个单位，所以我们预期大约有 $n/(1+p)$ 本书；公式与此吻合，另加一个有界的修正项，来自书架末端浪费掉的空间。

我们已把这个闭合公式与上架过程的精确枚举结果核对（$n \leq 12$，并取多个 p 值），亦以模拟作了核对；例如当 $p = \frac{1}{2}$ 及 $n = 5$ 时，两者都给出 $M(5) = \frac{103}{32} \approx 3.219$。

常见错误

- 以为厚书放不下之后，薄书仍然可以上架。题目说明其后所有书都放到手推车上，这正是 $M(1) = q$ 而不是更大的数值的原因。
- 以常数作为特解来试：因为 $\lambda = 1$ 是辅助方程的解，常数会被齐次部分吸收，所以需要用 Cn。

STEP 之外

对第一步取条件，并留意到余下的部分是同一问题的较小翻版，这种方法称为*首步分析*（first-step analysis）。它是在马可夫链（Markov chain）中求期望首达时间（hitting time）和吸收概率（absorption probability）的标准方法。我们的答案亦说明了*更新定理*（renewal theorem）：若把大小随机的物件装进长度 n 之内，能装下的数目的增长速度与 n 除以平均大小相若。Grimmett 与 Welsh 的 *Probability: An Introduction*（概率入门教科书）既涵盖随机漫步（random walk）的差分方程，亦有马可夫链的入门介绍。

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。