详细解答

STEP I 2015，第 12 题

概率与统计 · 泊松分布的稀疏化

> **STEP I 2015，第 12 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/15-s1-q12

关键思路

以到达人数作条件（conditioning），需要做手术的人数便成为一个二项（binomial）计数。然后对所有可能的到达人数求和（即全概率定律（law of total probability）），便得出需要做手术人数的分布，而这个和式会化简为一个指数级数（exponential series）。至于最后一部分，关键事实是：若两个平均值相同的独立（independent）泊松（Poisson）计数以其总和作条件，则两者之间的分配服从二项分布。

解答

设 X 为某一天到达的伤者人数，Y 为其中需要做手术的人数。

第 (i) 部分

已知 $X = n$，这 n 名伤者各自以概率 $\frac{1}{4}$ 需要做手术，且彼此独立。所以在此条件下 $Y \sim \mathrm{B}(n, \frac{1}{4})$，而对于 $0 \le r \le n$，

$$\mathrm{P}(Y = r \mid X = n) = \binom{n}{r}\left(\frac{1}{4}\right)^r\left(\frac{3}{4}\right)^{n-r} = \binom{n}{r}\frac{3^{n-r}}{4^n}.$$

（若 $r > n$，则概率为 0。）

第 (ii) 部分

固定 $r \geq 0$。由于 $Y = r$ 只可能在 $X \geq r$ 时发生，全概率定律给出

$$\begin{aligned} \mathrm{P}(Y = r) &= \sum_{n=r}^{\infty} \mathrm{P}(X = n)\,\mathrm{P}(Y = r \mid X = n) \\ &= \sum_{n=r}^{\infty} \frac{\mathrm{e}^{-8} 8^n}{n!} \cdot \frac{n!}{r!\,(n-r)!} \left(\frac{1}{4}\right)^r \left(\frac{3}{4}\right)^{n-r}. \end{aligned}$$

$n!$ 互相抵消。接着把 $8^n = 8^r \cdot 8^{n-r}$ 拆开，使得 $8^r (\frac{1}{4})^r = 2^r$ 及 $8^{n-r}(\frac{3}{4})^{n-r} = 6^{n-r}$：

$$\begin{aligned} \mathrm{P}(Y = r) &= \frac{\mathrm{e}^{-8}\, 2^r}{r!} \sum_{n=r}^{\infty} \frac{6^{n-r}}{(n-r)!} \\ &= \frac{\mathrm{e}^{-8}\, 2^r}{r!} \sum_{k=0}^{\infty} \frac{6^k}{k!} \qquad (k = n - r) \\ &= \frac{\mathrm{e}^{-8}\, 2^r}{r!}\, \mathrm{e}^{6} = \frac{\mathrm{e}^{-2}\, 2^r}{r!}. \end{aligned}$$

此式对每个 $r = 0, 1, 2, \ldots$ 都成立，所以每天需要做手术的人数服从平均值为 $8 \times \frac{1}{4} = 2$ 的泊松分布（Poisson distribution）。

> ? **想一想** e^6 从何而来?

把所有只与 r 有关的部分移到和式之外后，剩下的就是 $\sum 6^k/k!$，即 e^6 的指数级数。把 8^n 拆成 $8^r \cdot 8^{n-r}$，正正是令余下的和式只依赖于 $n - r$ 的关键。

第 (iii) 部分

设 M 和 T 分别为星期一和星期二需要做手术的人数。由第 (ii) 部分，两者均为 $\mathrm{Po}(2)$。由于不同日子的到达人数互相独立，而伤者是否需要做手术亦彼此独立，所以 M 与 T 独立。于是

$$\mathrm{P}(M = 8 \mid M + T = 12) = \frac{\mathrm{P}(M = 8,\ T = 4)}{\mathrm{P}(M + T = 12)}.$$

由独立性，分子为

$$\mathrm{P}(M = 8)\,\mathrm{P}(T = 4) = \frac{\mathrm{e}^{-2} 2^8}{8!} \cdot \frac{\mathrm{e}^{-2} 2^4}{4!} = \frac{\mathrm{e}^{-4}\, 2^{12}}{8!\,4!}.$$

至于分母，独立泊松随机变量之和仍服从泊松分布，其平均值相加，所以 $M + T \sim \mathrm{Po}(4)$。若不想直接引用这个结果，亦可直接推导：

$$\mathrm{P}(M + T = 12) = \sum_{m=0}^{12} \frac{\mathrm{e}^{-2} 2^m}{m!} \cdot \frac{\mathrm{e}^{-2} 2^{12-m}}{(12-m)!} = \frac{\mathrm{e}^{-4} 2^{12}}{12!} \sum_{m=0}^{12} \binom{12}{m} = \frac{\mathrm{e}^{-4}\, 4^{12}}{12!},$$

当中用了 $\sum_m \binom{12}{m} = 2^{12}$。两者相除，

$$\begin{aligned}\mathrm{P}(M = 8 \mid M + T = 12) &= \frac{\mathrm{e}^{-4}\,2^{12}}{8!\,4!} \cdot \frac{12!}{\mathrm{e}^{-4}\,4^{12}} \\ &= \binom{12}{8}\left(\frac{1}{2}\right)^{12} = \frac{495}{4096}.\end{aligned}$$

由于 $495 = 3^2 \cdot 5 \cdot 11$ 是奇数，而 $4096 = 2^{12}$，此分数已是最简分数（lowest terms）。所以所求概率为 $\dfrac{495}{4096}$。

想一想 为甚么答案看起来像一个二项概率?

已知总数为 12，由于两天的平均值相同，这 12 宗手术个案中的每一宗来自星期一或星期二的可能性均等。所以星期一的人数服从 $\mathrm{B}(12, \frac{1}{2})$。留意原本的平均值 8 已完全消失。

验算

- 第 (ii) 部分的平均值与 $\mathrm{E}(Y) = \mathrm{E}\big(\mathrm{E}(Y \mid X)\big) = \mathrm{E}(X/4) = 2$ 一致。
- 对 $r = 0, \ldots, 14$ 以数值方法计算第 (ii) 部分的级数,结果在舍入误差范围内与 $\mathrm{e}^{-2}2^r/r!$ 相同。例如 $\mathrm{P}(Y = 0) = \sum_n \mathrm{e}^{-8}8^n(\frac{3}{4})^n/n! = \mathrm{e}^{-8}\mathrm{e}^{6}$。
- 对 $m = 0, \ldots, 12$，条件概率 $\binom{12}{m}/2^{12}$ 之和为 1，而 $495/4096 \approx 0.121$。以电脑模拟整个模型（泊松到达，然后每人各以概率 $\frac{1}{4}$ 需要做手术），所得数值与此相符。

常见错误

- 在第 (ii) 部分中由 $n = 0$ 而非 $n = r$ 开始求和：若到达人数少于 r，就不可能有 r 名伤者需要做手术。
- 在第 (iii) 部分中每天都用 $\mathrm{Po}(8)$。条件是针对手术个案，所以相关的平均值是 2。（在这里平均值无论如何都会抵消，但推导过程仍应正确。）
- 题目要求最简分数时，却把答案写成 $\binom{12}{8}/2^{12}$ 或 $1980/16384$。

STEP 之外

第 (ii) 部分称为*稀疏化*（thinning）：若泊松过程（Poisson process）中的每一点各自独立地以概率 p 保留，则被保留的点构成一个速率乘以 p 的泊松过程，而且，或许令人意外的是，它与被舍弃的点互相独立。第 (iii) 部分则是其逆向：以总和作条件时，独立的泊松计数按多项分布（multinomial distribution）分配。这两个事实是泊松过程理论的基础，

你会在大学第一个概率课程中学到。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）对两者均有论述。

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。