详细解答

STEP I 2001，第 13 题

概率与统计 • 条件泊松分布

STEP I 2001，第 13 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/01-s1-q13

关键思路

以打破碟子的总数作为条件，便可把一道泊松（Poisson）问题化为二项（binomial）问题。假如四名学生各自独立地以相同的比率打破碟子，那么一旦知道总共打破了多少只碟子，每一只破碟由四人中任何一人打破的可能性都相等，而且与其他破碟互相独立。因此，数学家在破碟中所占的数目服从二项分布，而未知的 λ 亦随之消失。

解答

设置

设 M 为数学家打破的碟子数目，R 为其余三名学生打破的碟子总数。题目指出每名学生打破的碟子数目均服从 $\mathrm{Po}(\lambda)$，且彼此独立。所以 $M \sim \mathrm{Po}(\lambda)$，而 R 是三个独立的 $\mathrm{Po}(\lambda)$ 变量之和，故服从 $\mathrm{Po}(3\lambda)$。此外，M 与 R 独立（independent），因为 R 只取决于其余三名学生。总数为 $T = M + R \sim \mathrm{Po}(4\lambda)$。

? **想一想** 为甚么独立泊松变量之和仍然服从泊松分布?

对于独立的 $X \sim \mathrm{Po}(\alpha)$ 和 $Y \sim \mathrm{Po}(\beta)$，就 n 的各种拆分方式求和：

$$\mathrm{P}(X+Y=n) = \sum_{k=0}^{n} \frac{\mathrm{e}^{-\alpha}\alpha^k}{k!} \cdot \frac{\mathrm{e}^{-\beta}\beta^{n-k}}{(n-k)!} = \frac{\mathrm{e}^{-(\alpha+\beta)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \alpha^k \beta^{n-k} = \frac{\mathrm{e}^{-(\alpha+\beta)}(\alpha+\beta)^n}{n!},$$

最后一步用了二项式定理（binomial theorem）。所以 $X + Y \sim \mathrm{Po}(\alpha + \beta)$。应用这个结果两次，得 $R \sim \mathrm{Po}(3\lambda)$；再应用一次，得 $T \sim \mathrm{Po}(4\lambda)$。

M 的条件分布

对于 $m = 0, 1, \ldots, 5$，事件 $\{M = m,\ T = 5\}$ 与 $\{M = m,\ R = 5 - m\}$ 相同。由于 M 与 R 独立，

$$\begin{aligned}\mathrm{P}(M = m \mid T = 5) &= \frac{\mathrm{P}(M = m)\,\mathrm{P}(R = 5 - m)}{\mathrm{P}(T = 5)} \\ &= \frac{\dfrac{\mathrm{e}^{-\lambda}\lambda^m}{m!} \cdot \dfrac{\mathrm{e}^{-3\lambda}(3\lambda)^{5-m}}{(5-m)!}}{\dfrac{\mathrm{e}^{-4\lambda}(4\lambda)^5}{5!}} \\ &= \frac{5!}{m!\,(5-m)!} \cdot \frac{3^{5-m}}{4^5} = \binom{5}{m}\left(\tfrac{1}{4}\right)^m \left(\tfrac{3}{4}\right)^{5-m}.\end{aligned}$$

指数项互相约去，λ 的所有幂次也一样（分子和分母各有 λ^5）。所以，在 $T = 5$ 的条件下，不论 λ 取甚么值，都有 $M \sim \mathrm{B}(5, \frac{1}{4})$。

答案

把 $m = 3, 4, 5$ 的二项概率相加：

$$\begin{aligned}\mathrm{P}(M \geq 3 \mid T = 5) &= \binom{5}{3}\frac{3^2}{4^5} + \binom{5}{4}\frac{3}{4^5} + \binom{5}{5}\frac{1}{4^5} \\ &= \frac{10 \times 9 + 5 \times 3 + 1}{1024} = \frac{106}{1024} = \frac{53}{512}.\end{aligned}$$

所以，五只破碟之中有三只或以上由数学家打破的概率是 $\dfrac{53}{512} \approx 0.104$。

验算

- 六个条件概率（conditional probability）$\binom{5}{m}3^{5-m}/4^5$ 之和为 $(1+3)^5/4^5 = 1$，正如所料。
- 对称性：按同样的论证，每名学生在五只破碟中所占数目的期望值（expected value）为 $5 \times \frac{1}{4} = \frac{5}{4}$，而四人所占数目之和为 5。
- 我们直接由四个泊松分布计算 $\mathrm{P}(M \geq 3 \mid T = 5)$（就五只破碟在四名学生之间的每一种分配方式求和），取 $\lambda = 0.3, 1, 2.7$ 和 5，每次都得到 $0.103515625 = 53/512$。另外模拟了 4×10^5 轮洗碗，在恰好打破五只碟子的各轮之中，所得比例约为 0.105。

常见错误

- 答案中仍留有 λ。它必定会约去；如果约不去，就表示取条件的步骤出了错。

- 除以 $\mathrm{P}(M+R=5)$ 时，把 R 当作 $\mathrm{Po}(\lambda)$ 而不是 $\mathrm{Po}(3\lambda)$，这等于只有两名学生。
- 漏掉二项式系数，或把事件当作“恰好三只”而不是“三只或以上”。

STEP 之外

这是一个一般事实的小例子：若 $X_1, \ldots, X_k$ 是独立的泊松变量，平均值分别为 $\lambda_1, \ldots, \lambda_k$，则在它们的和为 n 的条件下，向量 $(X_1, \ldots, X_k)$ 服从多项分布（multinomial distribution），其概率为 $\lambda_i / \sum \lambda_j$。反过来看，这个事实说明：若事件的数目服从 $\mathrm{Po}(\mu)$，而每个事件各自独立地标记为 k 种类型之一，则各类型的数目是互相独立的泊松变量。这种“泊松分拆”（Poisson splitting）是研究泊松过程（Poisson process）的基本工具。Blitzstein 与 Hwang 合著的 *Introduction to Probability*（概率论入门教材）以大量例子讨论了这两个方向。

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。