详细解答

STEP I 2006，第 13 题

概率与统计 · 泊松混合分布

STEP I 2006，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/06-s1-q13

关键思路

钻石以固定的平均率独立地散布，这就是一个泊松过程（Poisson process），因此 m 公斤朱古力中的钻石数目服从平均值为 m 的泊松分布（Poisson distribution）。朱古力的份量本身也是随机的，所以我们以它为条件：先就固定的份量求出答案，再利用全概率定理（law of total probability）（以及它在期望值方面的对应版本）对掷骰结果取平均。这样每一部分都归结为级数求和：(i) 是有限级数，(ii) 是几何级数（geometric series）。

解答

设 D 为我找到的钻石数目。由于钻石是以每公斤平均一颗的比率独立地藏在朱古力中，质量为 m kg 的一排朱古力所含的钻石数目服从平均值为 m 的泊松分布。特别地，已知质量为 m kg 时，

$$\mathrm{P}(D=0)=\mathrm{e}^{-m}, \qquad \mathrm{E}(D)=m.$$

以下一律把 $100k$ 克写成 $0.1k$ 公斤。

第 (i) 部分

设 N 为骰子掷出的点数，则对 $n=1,\ldots,6$ 有 $\mathrm{P}(N=n)=\frac{1}{6}$。已知 $N=n$ 时，我买 $0.1n$ kg，所以 $D \mid N=n$ 服从平均值为 $0.1n$ 的泊松分布。由全概率定理，

$$\begin{aligned}\mathrm{P}(D=0) &= \sum_{n=1}^{6} \mathrm{P}(N=n)\,\mathrm{P}(D=0 \mid N=n) = \frac{1}{6}\sum_{n=1}^{6} \mathrm{e}^{-0.1n} \\ &= \frac{1}{6}\left(\mathrm{e}^{-0.1}+\mathrm{e}^{-0.2}+\cdots+\mathrm{e}^{-0.6}\right).\end{aligned}$$

这是一个几何级数，首项为 $\mathrm{e}^{-0.1}$，公比（common ratio）为 $\mathrm{e}^{-0.1}$，共有六项，所以

$$\mathrm{P}(D=0)=\frac{\mathrm{e}^{-0.1}}{6}\left(\frac{1-\mathrm{e}^{-0.6}}{1-\mathrm{e}^{-0.1}}\right),$$

正是所求。

至于期望值（expectation），我们用同样的方法取条件：

$$\mathrm{E}(D)=\sum_{n=1}^{6}\mathrm{P}(N=n)\,\mathrm{E}(D\mid N=n)=\frac{1}{6}\sum_{n=1}^{6}0.1n=\frac{0.1\times 21}{6}=0.35.$$

换言之，$\mathrm{E}(D)=0.1\,\mathrm{E}(N)=0.1\times 3.5$，这正是我们要证明的结果。

第 (ii) 部分

现在设 T 为首次掷出 6 点时的掷骰次数。各次掷骰互相独立，而 $T=t$ 表示先掷出 $t-1$ 次非 6 点，然后掷出一次 6 点，所以

$$\mathrm{P}(T=t)=\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6},\qquad t=1,2,3,\ldots$$

已知 $T=t$ 时，我买 $0.1t$ kg，所以 $\mathrm{P}(D=0\mid T=t)=\mathrm{e}^{-0.1t}$。因此

$$\mathrm{P}(D=0)=\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}\,\mathrm{e}^{-0.1t}=\frac{\mathrm{e}^{-0.1}}{6}\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\mathrm{e}^{-0.1}\right)^{t-1}.$$

? **想一想** 为甚么我们可以把这个级数求和至无穷?

这是一个公比为 $r=\frac{5}{6}\mathrm{e}^{-0.1}$ 的几何级数，而 $0<r<1$，所以它收敛（converge），其和为 $1/(1-r)$。于是

$$\mathrm{P}(D=0)=\frac{\mathrm{e}^{-0.1}}{6}\cdot\frac{1}{1-\frac{5}{6}\mathrm{e}^{-0.1}}=\frac{\mathrm{e}^{-0.1}}{6-5\mathrm{e}^{-0.1}},$$

其中分子和分母同乘以 6。这就是所求的结果。

至于钻石数目的期望值，

$$\mathrm{E}(D)=\sum_{t=1}^{\infty}\mathrm{P}(T=t)\,\mathrm{E}(D\mid T=t)=\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}\times 0.1t=\frac{0.1}{6}\sum_{t=1}^{\infty}t\left(\tfrac{5}{6}\right)^{t-1}.$$

题目的提示在这里派上用场。当 $|x|<1$ 时，由二项展开式（binomial expansion）可得

$$(1-x)^{-2}=1+2x+3x^2+4x^3+\cdots=\sum_{t=1}^{\infty}tx^{t-1}.$$

取 $x=\frac{5}{6}$，这个和就是 $\left(1-\frac{5}{6}\right)^{-2}=36$，所以

$$\mathrm{E}(D)=\frac{0.1}{6}\times 36=0.6.$$

这同样等于 $0.1\,\mathrm{E}(T)$，其中 $\mathrm{E}(T)=6$ 就是大家熟悉的、掷出一次 6 点所需掷骰次数的平均值。

验算

- 按数值计算，(i) 的概率约为 0.7150，(ii) 的概率约为 0.6131。以泊松分布产生钻石数目，模拟 2×10^5 位顾客，所得的结果为 0.7151 和 0.6136，而钻石数目的平均值则为 0.350 和 0.599。
- (ii) 的概率比 (i) 的小。这是合理的：我平均买 600 g，而不是 350 g。
- 两个期望值都是那个随机乘数的平均值的 0.1 倍。钻石数目的期望值等于公斤数的期望值，因为已知质量时，钻石数目的期望值是质量的线性函数。

常见错误

- 以 $\mathrm{e}^{-0.35}$（或 $\mathrm{e}^{-0.6}$）作为没有钻石的概率。这样只用到平均质量，但 e^{-m} 并不是 m 的线性函数，所以必须先取 e^{-m} 再求平均，而不是先求平均。期望值是线性的，所以这条捷径对 $\mathrm{E}(D)$ 确实行得通。
- 忘记把克换算成公斤，结果得到平均值 $100n$ 而不是 $0.1n$。
- 写成 $\mathrm{P}(T = t) = (\frac{5}{6})^t \frac{1}{6}$，这样便多算了一次非 6 点。

STEP 之外

这里泊松分布的参数本身是随机的，这使 D 成为一个混合（mixture）分布。所用的工具，即全概率定理和重期望定律（tower law）$\mathrm{E}(D) = \mathrm{E}\big(\mathrm{E}(D \mid T)\big)$，都是大学一年级概率论的核心内容。概率母函数（probability generating function）可以俐落地处理这类问题：$\mathrm{P}(D = 0)$ 就是 T 的母函数在 $\mathrm{e}^{-0.1}$ 处的值。先定出随机的试验次数，再在每次试验中取一个随机的数目，由此便可进一步引出复合分布（compound distribution）和分支过程（branching process）。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门教材）对以上各项都有详细阐述。

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。