详细解答

STEP III 2021，第 11 题

概率与统计 · 指数分布

STEP III 2021，第 11 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 2021 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696588-2021-past-papers.zip

本题在 *Beyond STEP* 第 12.18 节无记忆性中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch12/memoryless.html

关键思路

所有关于 Y 和 Z 的问题，其实都是在问 X 落在哪一个区间（interval）。事件 $Y = n$ 就是 $n \le X < n+1$，而事件 $Z < z$ 可拆成一系列互斥（disjoint）的部分 $n \le X < n+z$，每个 n 对应一部分。指数分布（exponential distribution）的尾概率（tail probability）$\mathrm{P}(X \ge x) = \mathrm{e}^{-\lambda x}$ 会把每一部分变成一个几何级数（geometric series）中的项。

解答

全文记 $q = \mathrm{e}^{-\lambda}$；由于 $\lambda > 0$，故 $0 < q < 1$。对 $x \ge 0$，

$$\mathrm{P}(X \ge x) = \int_x^\infty \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \mathrm{e}^{-\lambda x},$$

又因 X 是连续（continuous）随机变量，对 $0 \le a < b$ 有 $\mathrm{P}(a \le X < b) = \mathrm{P}(a < X < b) = \mathrm{e}^{-\lambda a} - \mathrm{e}^{-\lambda b}$。

第 (i) 部分

Y 是 X 的整数部分（integer part），所以 $Y = n$ 当且仅当 $n \le X < n+1$。因此

$$\mathrm{P}(Y = n) = \mathrm{e}^{-n\lambda} - \mathrm{e}^{-(n+1)\lambda} = (1 - \mathrm{e}^{-\lambda})\mathrm{e}^{-n\lambda},$$

即为所求。（所以 Y 服从由 0 开始、公比为 q 的几何分布（geometric distribution）。）

第 (ii) 部分

$Z = X - Y$ 是 X 的小数部分（fractional part），所以 $0 \le Z < 1$。取 $0 \le z \le 1$。事件 $Z < z$ 发生，当且仅当对整数 $n = Y$ 有 $n \le X < n + z$。对 $n = 0, 1, 2, \ldots$，这些事件两两互斥，所以

$$\begin{aligned} \mathrm{P}(Z < z) &= \sum_{n=0}^{\infty} \mathrm{P}(n \le X < n + z) = \sum_{n=0}^{\infty} \left(\mathrm{e}^{-n\lambda} - \mathrm{e}^{-n\lambda - \lambda z}\right) \\ &= (1 - \mathrm{e}^{-\lambda z}) \sum_{n=0}^{\infty} q^n = \frac{1 - \mathrm{e}^{-\lambda z}}{1 - \mathrm{e}^{-\lambda}}, \end{aligned}$$

这里把几何级数求和，而因 $0 < q < 1$，该级数收敛（converges）。这就是所求的结果。

第 (iii) 部分

把 (ii) 的分布函数（distribution function）求导，得 Z 的概率密度（probability density）为

$$g(z) = \frac{\lambda \mathrm{e}^{-\lambda z}}{1 - \mathrm{e}^{-\lambda}} \qquad (0 \le z \le 1),$$

其余情况 $g(z) = 0$。取 $u = z$ 及 $\mathrm{d}v = \lambda \mathrm{e}^{-\lambda z}\,\mathrm{d}z$ 进行分部积分（integration by parts），

$$\int_0^1 z\lambda \mathrm{e}^{-\lambda z}\,\mathrm{d}z = \left[-z\mathrm{e}^{-\lambda z}\right]_0^1 + \int_0^1 \mathrm{e}^{-\lambda z}\,\mathrm{d}z = -\mathrm{e}^{-\lambda} + \frac{1 - \mathrm{e}^{-\lambda}}{\lambda}.$$

除以 $1 - \mathrm{e}^{-\lambda}$，

$$\mathrm{E}(Z) = \frac{1}{\lambda} - \frac{\mathrm{e}^{-\lambda}}{1 - \mathrm{e}^{-\lambda}} = \frac{1}{\lambda} - \frac{1}{\mathrm{e}^{\lambda} - 1}.$$

第 (iv) 部分

设 $0 \le z_1 < z_2 \le 1$。事件 $\{Y = n \text{ 及} z_1 < Z < z_2\}$ 恰好就是 $n + z_1 < X < n + z_2$，所以

$$\mathrm{P}(Y = n \text{ 及} z_1 < Z < z_2) = \mathrm{e}^{-\lambda(n+z_1)} - \mathrm{e}^{-\lambda(n+z_2)} = \mathrm{e}^{-n\lambda}\left(\mathrm{e}^{-\lambda z_1} - \mathrm{e}^{-\lambda z_2}\right).$$

? **想一想** 要判断是否独立，我们需要拿这个结果跟甚么比较?

跟乘积 $\mathrm{P}(Y = n)\,\mathrm{P}(z_1 < Z < z_2)$ 比较。由于 Z 是连续的，$\mathrm{P}(Z = z_1) = 0$，故由 (ii)

$$\mathrm{P}(z_1 < Z < z_2) = \mathrm{P}(Z < z_2) - \mathrm{P}(Z < z_1) = \frac{\mathrm{e}^{-\lambda z_1} - \mathrm{e}^{-\lambda z_2}}{1 - \mathrm{e}^{-\lambda}}.$$

乘以 (i) 中的 $\mathrm{P}(Y = n) = (1 - \mathrm{e}^{-\lambda})\mathrm{e}^{-n\lambda}$，因子 $1 - \mathrm{e}^{-\lambda}$ 互相抵消，正好得到上面的联合概率（joint probability）。所以

$$\mathrm{P}(Y = n \text{ 及} z_1 < Z < z_2) = \mathrm{P}(Y = n)\,\mathrm{P}(z_1 < Z < z_2)$$

对每个非负整数 n 及每个区间 $(z_1, z_2) \subseteq [0,1]$ 都成立。由于 Y 只取非负整数值，而 Z 只取 $[0,1)$ 中的值，这类事件决定了整个联合分布。因此 Y 和 Z 是独立（independent）的：知道 X 落在哪一个单位区间，并不会告诉我们它落在该区间内的哪个位置。

验算

- 在 (ii) 中代入 $z = 1$，得 $\mathrm{P}(Z < 1) = 1$，理应如此；代入 $z = 0$ 则得 0。
- 当 $\lambda \to 0$ 时，X 的密度在每个单位区间上几乎是平的，所以 Z 应接近 $[0,1]$ 上的均匀分布（uniform distribution）。事实上，由 $\mathrm{e}^{\lambda} - 1 = \lambda + \frac{1}{2}\lambda^2 + \cdots$ 得 $\frac{1}{\mathrm{e}^{\lambda}-1} = \frac{1}{\lambda} - \frac{1}{2} + O(\lambda)$，所以 $\mathrm{E}(Z) \to \frac{1}{2}$。当 $\lambda \to \infty$ 时，X 几乎总是非常小，而 $\mathrm{E}(Z) \to 0$。
- 当 $\lambda = 1$ 时，$\mathrm{E}(Z) = 1 - \frac{1}{\mathrm{e}-1} \approx 0.4180$，这与以电脑模拟 200 000 个 X 值的结果吻合；(iv) 中模拟所得的联合频率亦与边际分布（marginal distributions）的乘积吻合。

常见错误

- 写成 $\mathrm{P}(Z < z) = \mathrm{P}(X < z)$，忘记了 $Z < z$ 可以在每一个单位区间发生，而不只是第一个。
- 只凭一组 n、z_1、z_2 的取值就断定独立：因式分解必须对所有取值都成立。

STEP 之外

(iv) 中的独立性，其实是指数分布无记忆性（memoryless property）的一个影子：已知 $X \geq n$，无论 n 是多少，超出部分 $X - n$ 仍然服从相同参数的指数分布。这正是指数分布用来描述泊松过程（Poisson process）中等候时间的原因，也是它的整数部分服从几何分布——即离散的无记忆分布——的原因。Grimmett 与 Welsh 所著的 *Probability: An Introduction*（《概率导论》）对这两种分布及泊松过程都有详细阐述。

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。