详细解答

STEP I 1994，第 14 题

概率与统计 · 指数分布的顺序统计量

STEP I 1994，第 14 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/94-s1-q14

关键思路

每个 T_i 都服从指数分布（exponential distribution），所以第 (i) 部分只是两个标准的分部积分（integration by parts）。至于最后一篇和第一篇文章，窍门在于把关于最大值或最小值的陈述，转化为关于每一个 T_i 的陈述：最后一篇文章在时刻 u 或之前交到，当且仅当所有文章都在那时或之前交到；第一篇文章在时刻 t 仍未交到，当且仅当所有文章在那时都仍未交到。独立性（independence）随即把“所有”变成乘积，再对分布函数（distribution function）求导，便得到密度函数（density）。

解答

第 (i) 部分

以 X 表示任意一个 T_i，其密度函数为 $f(t) = \lambda \mathrm{e}^{-\lambda t}$（$t \geq 0$）。用分部积分，取 $u = t$ 及 $\mathrm{d}v = \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t$（故 $v = -\mathrm{e}^{-\lambda t}$），

$$\mathrm{E}(X) = \int_0^\infty t\,\lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \Big[-t\mathrm{e}^{-\lambda t}\Big]_0^\infty + \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t = 0 + \frac{1}{\lambda} = \frac{1}{\lambda},$$

因为当 $t \to \infty$ 时 $t\mathrm{e}^{-\lambda t} \to 0$（指数函数的增长胜过幂函数）。同样地，

$$\mathrm{E}(X^2) = \int_0^\infty t^2 \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \Big[-t^2\mathrm{e}^{-\lambda t}\Big]_0^\infty + \int_0^\infty 2t\,\mathrm{e}^{-\lambda t}\,\mathrm{d}t = \frac{2}{\lambda}\,\mathrm{E}(X) = \frac{2}{\lambda^2},$$

其中我们认出 $\int_0^\infty t\,\mathrm{e}^{-\lambda t}\,\mathrm{d}t$ 就是 $\frac{1}{\lambda}\mathrm{E}(X)$。因此

$$\mathrm{E}(T_i) = \frac{1}{\lambda}, \qquad \mathrm{Var}(T_i) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

第 (ii) 部分

首先我们需要单独一个 T_i 的分布函数：对于 $u \geq 0$,

$$\mathrm{P}(T_i \leq u) = \int_0^u \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = 1 - \mathrm{e}^{-\lambda u}.$$

现在 $U = \max(T_1, \ldots, T_n)$，而各个 T_i 中最大的一个不超过 u，当且仅当每一个都不超过 u：

$$\{U \leq u\} = \{T_1 \leq u\} \cap \{T_2 \leq u\} \cap \cdots \cap \{T_n \leq u\}.$$

? **想一想** 为甚么这里可以把各个概率相乘?

事件 $\{T_i \leq u\}$ 涉及不同的 T_i，而按假设它们互相独立，所以交集的概率等于各概率的乘积。因此，对于 $u \geq 0$,

$$\mathrm{P}(U \leq u) = \prod_{i=1}^{n} \mathrm{P}(T_i \leq u) = (1 - \mathrm{e}^{-\lambda u})^n,$$

正是题目所求。U 的密度函数是这个分布函数的导数；由链式法则（chain rule），

$$f_U(u) = n(1 - \mathrm{e}^{-\lambda u})^{n-1} \cdot \lambda \mathrm{e}^{-\lambda u} = n\lambda\, \mathrm{e}^{-\lambda u}(1 - \mathrm{e}^{-\lambda u})^{n-1} \qquad (u \geq 0),$$

而当 $u < 0$ 时 $f_U(u) = 0$。

第 (iii) 部分

现在 $T = \min(T_1, \ldots, T_n)$。这里有用的是补事件（complement）：第一篇文章在时刻 t 之后才交到，当且仅当到那时没有一篇交到，即每一个 $T_i > t$。利用 $\mathrm{P}(T_i > t) = \mathrm{e}^{-\lambda t}$，并再次运用独立性，

$$\mathrm{P}(T > t) = \prod_{i=1}^{n} \mathrm{P}(T_i > t) = (\mathrm{e}^{-\lambda t})^n = \mathrm{e}^{-n\lambda t} \qquad (t \geq 0).$$

所以 $\mathrm{P}(T \leq t) = 1 - \mathrm{e}^{-n\lambda t}$，求导得

$$f_T(t) = n\lambda\, \mathrm{e}^{-n\lambda t} \qquad (t \geq 0),$$

而当 $t < 0$ 时 $f_T(t) = 0$。

第 (iv) 部分

第 (iii) 部分的密度函数，正是第 (i) 部分的密度函数把 λ 换成 $n\lambda$：T 服从参数为 $n\lambda$ 的指数分布。所以我们可以由第 (i) 部分直接写出

$$\mathrm{E}(T) = \frac{1}{n\lambda}, \qquad \mathrm{Var}(T) = \frac{1}{n^2\lambda^2}.$$

这合乎情理：有 n 名学生时，第一篇文章交到的时间，应该比任何指定一篇快大约 n 倍。

验算

- 当 $n = 1$ 时只有一篇文章，所以 $T = U = T_1$；两个密度函数的确都简化为 $\lambda e^{-\lambda t}$。
- $P(U \le u)$ 由 $u = 0$ 时的 0 上升至 $u \to \infty$ 时的 1，符合分布函数应有的性质。
- 我们以 $n = 5$、$\lambda = 0.7$ 模拟了 2×10^5 个班别。例如，$U \le 2$ 所占的比例为 0.2425，对比 $(1 - e^{-1.4})^5 \approx 0.2427$；$T$ 的平均值为 0.2860，对比 $1/(5 \times 0.7) \approx 0.2857$。用辛普森法则（Simpson's rule）亦验证了 f_U 的积分为 1。

常见错误

- 处理最小值时直接计算 $P(T \le t)$：“至少有一个 $T_i \le t$”并不是交集，所以不能分解成乘积。应该经由 $P(T > t)$ 来计算。
- 忘记说明乘积之所以成立，是因为独立性。
- 写成 $\operatorname{Var}(T_i) = E(T_i^2) = 2/\lambda^2$，忘记减去 $(E\,T_i)^2$。

STEP 之外

T 和 U 分别是样本中最小和最大的顺序统计量（order statistics）。独立指数随机变量的最小值仍服从指数分布，而其率参数是各率参数之和，这是泊松过程（Poisson process）和排队论（queueing theory）的基础。由无记忆性（memoryless property）还可得出一个漂亮的结果：相继两篇文章之间的时间间隔是互相独立的指数随机变量，率参数依次为 $n\lambda, (n-1)\lambda, \ldots, \lambda$，所以 $E(U) = \frac{1}{\lambda}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$。这两个结果在 Grimmett 与 Welsh 的 *Probability: An Introduction*（概率入门），以及 Blitzstein 与 Hwang 的 *Introduction to Probability*（概率导论）中都有详细讨论。

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。