详细解答

STEP III 2025，第 11 题

概率与统计 · 最大值的分布

STEP III 2025，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 2025 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/750486-2025-past-papers.zip

本题在 *Beyond STEP* 第 12.25 节众多之中的最大值中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch12/extremes.html

关键思路

当且仅当每一个 X_i 都不超过 y 时，X_i 之中最大者才不超过 y；而独立性（independence）把这件事化为一个乘积。此后，整道题就是从 G 读出分位数（quantile），再看看当 n 很大时它们会怎样变化。真正起作用的只有一个近似：$\alpha^{1/n} = \mathrm{e}^{(\ln\alpha)/n} \approx 1 + \frac{\ln\alpha}{n}$，因为当 n 很大时，$(\ln\alpha)/n$ 很小。

解答

第 (i) 部分

对于 $x \geq 0$，

$$F(x) = \int_0^x \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = 1 - \mathrm{e}^{-\lambda x}.$$

由于 $Y = \max(X_1, \ldots, X_n)$，事件 $\{Y \leq y\}$ 与“对每一个 i 都有 $X_i \leq y$”这个事件相同。各 X_i 互相独立，所以对于 $y \geq 0$，

$$G(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(X_1 \leq y) \cdots \mathrm{P}(X_n \leq y) = F(y)^n = \left(1 - \mathrm{e}^{-\lambda y}\right)^n,$$

即为所求。

第 (ii) 部分

Y 是连续随机变量（continuous random variable），所以 $\mathrm{P}(Y < y) = \mathrm{P}(Y \le y) = G(y)$。因此 $L(\alpha)$ 须满足的条件是 $G(L(\alpha)) = \alpha$：

$$\begin{aligned}\left(1 - \mathrm{e}^{-\lambda L}\right)^n = \alpha &\iff 1 - \mathrm{e}^{-\lambda L} = \alpha^{1/n} \qquad \text{（取正的 } n \text{ 次方根）} \\ &\iff \mathrm{e}^{-\lambda L} = 1 - \alpha^{1/n} \\ &\iff L(\alpha) = -\frac{1}{\lambda}\ln\left(1 - \alpha^{1/n}\right).\end{aligned}$$

这个对数（logarithm）有定义，因为由 $0 < \alpha < 1$ 可得 $0 < \alpha^{1/n} < 1$。

至于 $U(\alpha)$，条件 $\mathrm{P}(Y > U) = \alpha$ 即 $G(U) = 1 - \alpha$，所以把 α 换成 $1 - \alpha$ 再重复同样的步骤，便得

$$U(\alpha) = -\frac{1}{\lambda}\ln\left(1 - (1-\alpha)^{1/n}\right).$$

由于 $\alpha \le \frac{1}{2} \le 1 - \alpha$，而 G 是递增函数，所以 $L(\alpha) \le U(\alpha)$，且只有当 $\alpha = \frac{1}{2}$ 时才取等号。

第 (iii) 部分

写 $\alpha^{1/n} = \mathrm{e}^{(\ln\alpha)/n}$。当 α 固定而 n 很大时，$t = (\ln\alpha)/n$ 很小，所以由 $\mathrm{e}^t \approx 1 + t$ 得

$$1 - \alpha^{1/n} \approx -\frac{\ln\alpha}{n} = \frac{1}{n}\ln\left(\frac{1}{\alpha}\right).$$

留意这是正数，因为 $\alpha < 1$。代入第 (ii) 部分的结果，

$$\lambda L(\alpha) = -\ln\left(1 - \alpha^{1/n}\right) \approx -\ln\left(\frac{1}{n}\ln\frac{1}{\alpha}\right) = \ln(n) - \ln\left(\ln\frac{1}{\alpha}\right).$$

> **想一想** 对一个近似值取对数，是否安全？$1 - \alpha^{1/n}$ 本身非常小，所以 $1 - \alpha^{1/n}$ 的误差相对于它本身可能很大。

关键在于相对误差（relative error）。设 $t = (\ln\alpha)/n$，则 $1 - \mathrm{e}^t = -t\,(1 + \frac{t}{2} + \cdots)$，所以准确值与近似值之比是 $1 + O(1/n)$。取对数后，$\lambda L(\alpha)$ 的误差是 $\ln(1 + O(1/n))$，它趋于 0。所以这个近似不仅在 n 很大时很好；两边之差确实趋于零。

第 (iv) 部分

Y 的中位数（median）m 满足 $G(m) = \frac{1}{2}$，所以 $m = L(\frac{1}{2})$（它同时也是 $U(\frac{1}{2})$）。由第 (iii) 部分，

$$\lambda m \approx \ln n - \ln(\ln 2).$$

由于 $\ln(\ln 2)$ 是一个固定的数（约为 -0.37），而 $\ln n \to \infty$，所以当 n 增大时，中位数趋于无限大。

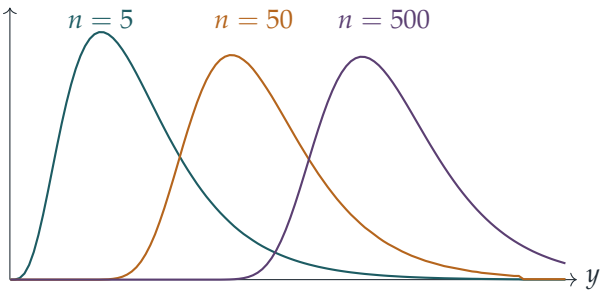
至于 $U(\alpha)$，在第 (iii) 部分中把 α 换成 $1-\alpha$：

$$\lambda U(\alpha) \approx \ln n - \ln\left(\ln\frac{1}{1-\alpha}\right).$$

两式相减，$\ln n$ 项互相抵消：

$$U(\alpha) - L(\alpha) \approx \frac{1}{\lambda}\left[\ln\left(\ln\frac{1}{\alpha}\right) - \ln\left(\ln\frac{1}{1-\alpha}\right)\right].$$

由于两个近似的误差都趋于零（正如上面方框所述），区间宽度趋于这个值，它取决于 α 和 λ，但与 n 无关。所以当 n 增大时，Y 的分布像 $\frac{1}{\lambda}\ln n$ 那样往无限大漂移，但并不会扩散开来。



上图显示 $\lambda = 1$ 时 Y 的概率密度（density）：n 每增大十倍，曲线便向右移约 $\ln 10 \approx 2.3$，而曲线的形状则趋于稳定。

第 (v) 部分

取 $\alpha = 0.05$。那么

$$\mathrm{P}\big(L(\alpha) \le Y \le U(\alpha)\big) = 1 - \mathrm{P}(Y < L) - \mathrm{P}(Y > U) = 1 - 2(0.05) = 0.9,$$

所以无论 n 是多少，Y 落在区间 $[L(0.05), U(0.05)]$ 内的概率恰好是 0.9。当 n 很大时，由第 (iv) 部分，这个区间的宽度约为

$$\frac{1}{\lambda}\left[\ln(\ln 20) - \ln\left(\ln\frac{1}{0.95}\right)\right].$$

现在运用题目给出的两个事实：

- $\mathrm{e}^3 \approx 20$，所以 $\ln 20 \approx 3$，而 $\ln(\ln 20) \approx \ln 3 \approx 1.1$；
- 利用 $\ln(1+t) \approx t$（取 $t = -0.05$），得 $\ln\frac{1}{0.95} = -\ln(1-0.05) \approx 0.05 = \frac{1}{20}$；所以 $\ln\left(\ln\frac{1}{0.95}\right) \approx \ln\frac{1}{20} = -\ln 20 \approx -3$。

因此宽度约为

$$\frac{1}{\lambda}(\ln 3 + 3) \approx \frac{4.1}{\lambda} \approx 4\lambda^{-1},$$

即为所求。

验算

- 在第 (ii) 部分中，把 $L(\alpha)$ 代回 G，得 $\left(1-(1-\alpha^{1/n})\right)^n = \alpha$，完全吻合。
- 当 $\alpha = 0.05$ 而 $n = 10^4$ 时，$\lambda L(\alpha)$ 的准确值是 8.11330，而近似式 $\ln n - \ln\ln 20$ 给出 8.11315；当 $n = 10$ 时，两者分别是 1.35 和 1.21，可见“足够大”这个条件确实重要。
- 第 (v) 部分中宽度的准确极限值是 $4.067/\lambda$，与 $4/\lambda$ 相当接近。以电脑模拟五个指数分布随机变量的最大值，结果亦与 $G(y)$ 吻合至三位小数。

常见错误

- 写成 $G(y) = 1-(\mathrm{e}^{-\lambda y})^n$：这是最小值的累积分布函数（cumulative distribution function），而不是最大值的。
- 把 $\mathrm{P}(Y > U) = \alpha$ 写成 $G(U) = \alpha$，而不是 $G(U) = 1-\alpha$。
- 在第 (v) 部分中选取 $\alpha = 0.1$：区间两端的尾部各应留下 0.05，而不是 0.1。

STEP 之外

这是初窥极值理论（extreme value theory）。第 (iii) 部分其实是在说明 $\lambda Y - \ln n$ 有一个极限分布（limiting distribution），且 $\mathrm{P}(\lambda Y - \ln n \le x) \to \exp(-\mathrm{e}^{-x})$：这就是甘贝尔分布（Gumbel distribution），可用来为破纪录的洪水、一年中最强的阵风以及最大额的保险索偿建立模型。一条了不起的定理指出，经适当缩放的最大值只可能收敛到三个分布族之一。Blitzstein 和 Hwang 的 *Introduction to Probability*（《概率导论》）讨论了顺序统计量（order statistics），是入门的好起点。

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。