详细解答

# STEP II 1996，第 12 题

概率与统计 · 最大值与最小值的分布

**STEP II 1996，第 12 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/96-s2-q12

## 关键思路

多个独立（independent）变量的最大值很难直接处理，但“最大值很小”这个事件却很简单：它发生，当且仅当每一个变量都很小，而独立性让我们可以把概率相乘。所以我们先求累积分布函数（cumulative distribution function），再求导得出密度。第 (ii) 部分正好是它的镜像：最小值很大，当且仅当每一个变量都很大。

## 解答

### 第 (i) 部分

每个 $X_j$ 都在 $[0,1]$ 上均匀分布（uniform），所以当 $0 \le y \le 1$ 时，$\mathrm{P}(X_j < y) = y$。对于这个范围内的 $y$，$Y < y$ 发生当且仅当对每个 $j$ 都有 $X_j < y$，而由于各 $X_j$ 互相独立，

$$F(y) = \mathrm{P}(Y < y) = \prod_{j=1}^{n} \mathrm{P}(X_j < y) = y^n, \qquad 0 \le y \le 1.$$

当然，当 $y < 0$ 时 $F(y) = 0$，当 $y > 1$ 时 $F(y) = 1$。求导后，$Y$ 的概率密度函数（probability density function）为

$$f(y) = \begin{cases} n y^{n-1}, & 0 \le y \le 1, \\ 0, & \text{其他情况。} \end{cases}$$

（作为简单的检查，$\int_0^1 n y^{n-1}\,\mathrm{d}y = 1$。）

接着计算矩（moments）：

$$\mathrm{E}(Y) = \int_0^1 y \cdot n y^{n-1}\,\mathrm{d}y = \frac{n}{n+1},$$
$$\mathrm{E}(Y^2) = \int_0^1 y^2 \cdot n y^{n-1}\,\mathrm{d}y = \frac{n}{n+2}.$$

因此

$$\mathrm{Var}(Y) = \frac{n}{n+2} - \frac{n^2}{(n+1)^2} = \frac{n\left[(n+1)^2 - n(n+2)\right]}{(n+2)(n+1)^2} = \frac{n}{(n+2)(n+1)^2},$$

因为 $(n+1)^2 - n(n+2) = n^2 + 2n + 1 - n^2 - 2n = 1$。这正是所要求的结果。

**想一想** 为甚么要处理 $\mathrm{P}(Y < y)$，而不是一开始就尝试写出 $\mathrm{P}(Y = y)$ 或 $Y$ 的密度?

对连续变量而言，密度并不是概率，而“最大的那个等于 $y$”这个事件，把“哪一个 $X_j$ 最大”和“其余各个的情况”混在一起。事件 $\{Y < y\}$ 是 $n$ 个独立事件的交集，所以它的概率只是一个简单的乘积。先求分布函数再求导，是处理最大值和最小值的标准方法。

### 第 (ii) 部分

设 $T_j$ 为第 $j$ 盏灯的失效时间，则 $\mathrm{P}(T_j \le t) = 1 - \mathrm{e}^{-t/\lambda}$，且

$$\mathrm{P}(T_j > t) = \mathrm{e}^{-t/\lambda}, \qquad t > 0.$$

各盏灯互相独立，所以各 $T_j$ 互相独立。设 $T = \min(T_1, \ldots, T_n)$ 为首次失效的时间。这一次方便处理的事件是 $T > t$，它发生当且仅当每盏灯在时间 $t$ 仍然运作：

$$\mathrm{P}(T > t) = \prod_{j=1}^{n} \mathrm{P}(T_j > t) = \left(\mathrm{e}^{-t/\lambda}\right)^n = \mathrm{e}^{-nt/\lambda}.$$

所以 $\mathrm{P}(T \le t) = 1 - \mathrm{e}^{-nt/\lambda}$，求导便得密度

$$g(t) = \frac{n}{\lambda}\,\mathrm{e}^{-nt/\lambda}, \qquad t > 0.$$

换句话说，$T$ 的分布与单独一盏灯的分布属同一类型，只是 $\lambda$ 换成了 $\lambda/n$。取 $u = t$ 作分部积分（integration by parts），

$$\begin{aligned}\mathrm{E}(T) &= \int_0^\infty t\,\frac{n}{\lambda}\,\mathrm{e}^{-nt/\lambda}\,\mathrm{d}t = \Big[-t\,\mathrm{e}^{-nt/\lambda}\Big]_0^\infty + \int_0^\infty \mathrm{e}^{-nt/\lambda}\,\mathrm{d}t \\ &= 0 + \Big[-\frac{\lambda}{n}\,\mathrm{e}^{-nt/\lambda}\Big]_0^\infty = \frac{\lambda}{n},\end{aligned}$$

其中边界项为零，因为当 $t \to \infty$ 时 $t\,\mathrm{e}^{-nt/\lambda} \to 0$。所以直至首次失效的期望时间是 $\lambda/n$，正如所求。

## 验算

- 当 $n = 1$ 时，$Y$ 就是 $[0,1]$ 上的均匀分布，公式给出 $\frac{1}{3\cdot4} = \frac{1}{12}$，正是我们熟悉的均匀变量的方差（variance）。
- 当 $n$ 增大时，方差的表现近似 $1/n^2$：许多个均匀变量的最大值会被挤向 1，一如所料。
- 在第 (ii) 部分取 $n = 1$，一盏灯的期望寿命是 $\lambda$，这正是参数为 $1/\lambda$ 的指数分布（exponential distribution）的平均值。
- 模拟 200 000 次五个均匀变量的最大值，得出方差为 0.01993，而理论值是 $\frac{5}{7\cdot36} \approx 0.01984$；以 $\lambda = 3$ 模拟四盏灯，得出首次失效时间的平均值为 0.749，而理论值是 $\frac{3}{4}$。

## 常见错误

- 对 $\mathrm{P}(T > t)$ 求导时忘记负号，结果得出一个负的“密度”。
- 写成 $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)$，而没有减去 $(\mathrm{E}(Y))^2$。
- 以为 $n$ 盏灯各自平均在时间 $\lambda$ 失效，所以第一盏也在 $\lambda$ 失效。其实 $n$ 个变量的最小值远小于其中任何一个。

## STEP 之外

最大值和最小值是*顺序统计量*（order statistics）的极端情况：把一个样本由小至大排序，$n$ 个均匀变量中第 $k$ 小的那个服从贝塔分布（beta distribution），而这里的 $Y$ 就是 $k = n$ 的情况。第 (ii) 部分显示了一个关键事实：独立指数变量的最小值仍然服从指数分布，而且各个率参数相加。这个事实是泊松过程（Poisson processes）和连续时间马可夫链（continuous-time Markov chains）理论的基础。Grimmett 与 Welsh 的 *Probability: An Introduction*（概率论入门）以及 Blitzstein 与 Hwang 的 *Introduction to Probability*（概率论导论），都对这些概念有详细阐述。

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。