詳細解答

STEP II 1996，第 12 題

概率與統計 • 最大值與最小值的分佈

> **STEP II 1996，第 12 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/96-s2-q12

關鍵思路

多個獨立（independent）變量的最大值很難直接處理，但「最大值很小」這個事件卻很簡單：它發生，當且僅當每一個變量都很小，而獨立性讓我們可以把概率相乘。所以我們先求累積分佈函數（cumulative distribution function），再求導得出密度。第 (ii) 部分正好是它的鏡像：最小值很大，當且僅當每一個變量都很大。

解答

第 (i) 部分

每個 X_j 都在 $[0,1]$ 上均勻分佈（uniform），所以當 $0 \le y \le 1$ 時，$\mathrm{P}(X_j < y) = y$。對於這個範圍內的 y，$Y < y$ 發生當且僅當對每個 j 都有 $X_j < y$，而由於各 X_j 互相獨立，

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

當然，當 $y < 0$ 時 $F(y) = 0$，當 $y > 1$ 時 $F(y) = 1$。求導後，Y 的概率密度函數（probability density function）為

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

（作為簡單的檢查，$\int_0^1 n y^{n-1}\,\mathrm{d}y = 1$。）

接着計算矩（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

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

因為 $(n+1)^2 - n(n+2) = n^2 + 2n + 1 - n^2 - 2n = 1$。這正是所要求的結果。

? **想一想** 為甚麼要處理 $\mathrm{P}(Y < y)$，而不是一開始就嘗試寫出 $\mathrm{P}(Y = y)$ 或 Y 的密度？

對連續變量而言，密度並不是概率，而「最大的那個等於 y」這個事件，把「哪一個 X_j 最大」和「其餘各個的情況」混在一起。事件 $\{Y < y\}$ 是 n 個獨立事件的交集，所以它的概率只是一個簡單的乘積。先求分佈函數再求導，是處理最大值和最小值的標準方法。

第 (ii) 部分

設 T_j 為第 j 盞燈的失效時間，則 $\mathrm{P}(T_j \le t) = 1 - \mathrm{e}^{-t/\lambda}$，且

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

各盞燈互相獨立，所以各 T_j 互相獨立。設 $T = \min(T_1, \ldots, T_n)$ 為首次失效的時間。這一次方便處理的事件是 $T > t$，它發生當且僅當每盞燈在時間 t 仍然運作：

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

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

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

換句話說，T 的分佈與單獨一盞燈的分佈屬同一類型，只是 λ 換成了 λ/n。取 $u = t$ 作分部積分（integration by parts），

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

其中邊界項為零，因為當 $t \to \infty$ 時 $t\,\mathrm{e}^{-nt/\lambda} \to 0$。所以直至首次失效的期望時間是 λ/n，正如所求。

驗算

- 當 $n = 1$ 時，Y 就是 $[0,1]$ 上的均勻分佈，公式給出 $\frac{1}{3\cdot 4} = \frac{1}{12}$，正是我們熟悉的均勻變量的方差（variance）。
- 當 n 增大時，方差的表現近似 $1/n^2$：許多個均勻變量的最大值會被擠向 1，一如所料。
- 在第 (ii) 部分取 $n = 1$，一盞燈的期望壽命是 λ，這正是參數為 $1/\lambda$ 的指數分佈（exponential distribution）的平均值。
- 模擬 200 000 次五個均勻變量的最大值，得出方差為 0.01993，而理論值是 $\frac{5}{7\cdot 36} \approx 0.01984$；以 $\lambda = 3$ 模擬四盞燈，得出首次失效時間的平均值為 0.749，而理論值是 $\frac{3}{4}$。

常見錯誤

- 對 $\mathrm{P}(T > t)$ 求導時忘記負號，結果得出一個負的「密度」。
- 寫成 $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)$，而沒有減去 $(\mathrm{E}(Y))^2$。
- 以為 n 盞燈各自平均在時間 λ 失效，所以第一盞也在 λ 失效。其實 n 個變量的最小值遠小於其中任何一個。

STEP 之外

最大值和最小值是*順序統計量*（order statistics）的極端情況：把一個樣本由小至大排序，n 個均勻變量中第 k 小的那個服從貝塔分佈（beta distribution），而這裏的 Y 就是 $k = n$ 的情況。第 (ii) 部分顯示了一個關鍵事實：獨立指數變量的最小值仍然服從指數分佈，而且各個率參數相加。這個事實是泊松過程（Poisson processes）和連續時間馬可夫鏈（continuous-time Markov chains）理論的基礎。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論導論），都對這些概念有詳細闡述。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。