詳細解答

STEP I 2003，第 14 題

概率與統計 • 泊松分佈

STEP I 2003，第 14 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/03-s1-q14

關鍵思路

Jane 見到的朋友數目並不是來訪的人數，而是這個人數以 2 為上限後的值。一個設有上限的變量只需考慮三種情況：無人來訪、恰好一人來訪、兩人或以上來訪。因此我們根本不必把整個泊松（Poisson）級數加起來，只需求出 $P(0)$ 和 $P(1)$，而「兩人或以上」的概率就是餘下的部分。至於新朋友，我們按其是否來訪作條件分析（conditioning）。

解答

原有的朋友

設 N 為某一天來訪的朋友數目，則 $N \sim \text{Po}(2)$，且

$$P(N=k) = \frac{\mathrm{e}^{-2}2^k}{k!}, \qquad k = 0,1,2,\ldots$$

Jane 見到 $S = \min(N,2)$ 位朋友。所以當 $N=0$ 時 $S=0$，當 $N=1$ 時 $S=1$，當 $N \geq 2$ 時 $S=2$。現在

$$P(N=0) = \mathrm{e}^{-2}, \qquad P(N=1) = 2\mathrm{e}^{-2}, \qquad P(N\geq 2) = 1-3\mathrm{e}^{-2},$$

因此

$$\mathrm{E}(S) = 0\cdot\mathrm{e}^{-2} + 1\cdot 2\mathrm{e}^{-2} + 2\left(1-3\mathrm{e}^{-2}\right) = 2-4\mathrm{e}^{-2},$$

這正是所要證明的結果。

同一個計算還有一個更簡潔的寫法，下文會再次用到。對任何取值於 $\{0,1,2,\ldots\}$ 的隨機變量（random variable）X，

$$\mathrm{E}(\min(X,2)) = 2 - 2P(X=0) - P(X=1),$$

因為 $\min(X,2)$ 與 2 的差距，在 $X=0$ 時是 2，在 $X=1$ 時是 1，其餘情況則為零。取 $X=N$，便再次得到 $2-2\mathrm{e}^{-2}-2\mathrm{e}^{-2}$。

加入新朋友之後

若新朋友來訪，設 $B=1$，否則設 $B=0$，於是 $P(B=1)=p$。題目指出 B 與 N 獨立（independent）。現在來訪的總人數是 $M=N+B$，而 Jane 見到 $\min(M,2)$ 位朋友。

> **想一想** 我們需要 M 的完整分佈嗎？

不需要。根據上面的恆等式，我們只需要 $P(M=0)$ 和 $P(M=1)$。利用獨立性把概率相乘：

$$\begin{aligned}P(M=0)&=P(N=0)\,P(B=0)=(1-p)\mathrm{e}^{-2},\\P(M=1)&=P(N=1)\,P(B=0)+P(N=0)\,P(B=1)=2(1-p)\mathrm{e}^{-2}+p\,\mathrm{e}^{-2}.\end{aligned}$$

因此

$$\begin{aligned}\mathrm{E}(\min(M,2))&=2-2(1-p)\mathrm{e}^{-2}-2(1-p)\mathrm{e}^{-2}-p\,\mathrm{e}^{-2}\\&=2-(4-3p)\mathrm{e}^{-2}.\end{aligned}$$

另一個做法是以 B 為條件。若新朋友沒有來訪（概率為 $1-p$），平均值一如之前，是 $2-4\mathrm{e}^{-2}$。若新朋友來訪（概率為 p），Jane 見到 $\min(N+1,2)$ 位朋友，這個數在 $N=0$ 時是 1，其餘情況是 2，其平均值為 $\mathrm{e}^{-2}+2(1-\mathrm{e}^{-2})=2-\mathrm{e}^{-2}$。所以平均值是

$$(1-p)(2-4\mathrm{e}^{-2})+p(2-\mathrm{e}^{-2})=2-(4-3p)\mathrm{e}^{-2}.$$

所以 Jane 現在平均每天見到 $2-(4-3p)\mathrm{e}^{-2}$ 位朋友，比之前多了 $3p\,\mathrm{e}^{-2}$。

驗算

- 當 $p=0$ 時，新朋友永不來訪，我們重新得到 $2-4\mathrm{e}^{-2}\approx 1.4587$。
- 當 $p=1$ 時，我們得到 $2-\mathrm{e}^{-2}\approx 1.8647$，這正是 $\min(N+1,2)$ 的平均值，與預期相符。
- 答案隨 p 遞增，而且恆小於 2，這是合理的。
- 按精確分佈對 $\min(k+b,2)$ 求和，以及取 $p=0.3$ 模擬 4×10^5 天，兩者都得出約 1.580，與 $2-3.1\mathrm{e}^{-2}$ 吻合。

常見錯誤

- 忘記了最多只見兩位朋友的上限，答成 $\mathrm{E}(N) = 2$ 或 $2 + p$。
- 把 $N + B$ 當作泊松變量。伯努利（Bernoulli）變量並非泊松變量，所以兩者之和也不是泊松變量。
- 恰好一人來訪有兩種情況（一位舊朋友來訪而新朋友沒來，或沒有舊朋友來訪而新朋友來了），卻漏掉其中一種。

STEP 之外

恆等式 $\mathrm{E}(\min(X,2)) = P(X \geq 1) + P(X \geq 2)$ 是尾和公式（tail-sum formula）$\mathrm{E}(X) = \sum_{k\geq 1} P(X \geq k)$ 的特殊情況，該公式適用於取非負整數值的隨機變量。它是對生存函數（survival function）積分的離散版本。這類設有上限的平均值常見於排隊模型和存貨模型，在容量有限的情況下，它們給出獲得服務的顧客數目的期望值（expected value）。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門教材）涵蓋尾和公式和條件期望（conditional expectation），正是本題所用的兩件工具。

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。