詳細解答

STEP II 2021，第 11 題

概率與統計 • 條件概率的遞推關係

STEP II 2021，第 11 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2021 年 STEP 試卷（zip 檔案）內的 STEP II 試卷：

www.ocr.org.uk/Images/696588-2021-past-papers.zip

本題在 *Beyond STEP* 第 12.5 節最後一位乘客的座位中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/airline-seat.html

關鍵思路

以 T_1 坐在哪裏為條件（condition on）。若 T_1 坐了自己的座位 S_1，其他人都不受影響；若 T_1 坐了 S_k，那麼 $T_2, \ldots, T_{k-1}$ 都能安然入座，而 T_k 就成了新的「隨機選座者」。關鍵的觀察是：此時 T_k 的處境與當初的 T_1 完全相同，只是列車變小了，而空着的座位 S_1 則扮演 T_k 自己座位的角色。這樣問題便轉化為一條遞推關係（recurrence）。

解答

全文中，「T_r 被擠走」指 T_r 到達時發現 S_r 已被人坐了。只有被擠走的乘客（或 T_1）才會隨機選座。

第 (i) 部分

當 $n = 2$ 時，T_1 各以概率（probability）$\frac{1}{2}$ 坐 S_1 或 S_2，而 T_2 能坐到 S_2 恰好當 T_1 坐了 S_1。所以

$$P_2 = \frac{1}{2}.$$

當 $n = 3$ 時，以 T_1 的座位為條件，每種情況的概率均為 $\frac{1}{3}$：

- S_1：T_2 和 T_3 都發現自己的座位空着，所以 T_3 坐到 S_3。
- S_3：T_3 必定坐不到 S_3。

► S_2：T_2 被擠走，在 S_1 和 S_3 之間隨機選擇；T_3 坐到 S_3 恰好當 T_2 選了 S_1，概率為 $\frac{1}{2}$。

因此

$$P_3 = \tfrac{1}{3}\cdot 1 + \tfrac{1}{3}\cdot 0 + \tfrac{1}{3}\cdot\tfrac{1}{2} = \tfrac{1}{2}.$$

第 (ii) 部分

假設 T_1 坐在 S_k，其中 $2 \le k \le n-1$。那麼 $T_2, \ldots, T_{k-1}$ 都發現自己的座位空着並坐下。當 T_k 到達時，空座位是

$$S_1,\ S_{k+1},\ S_{k+2},\ \ldots,\ S_n,$$

共 $n-k+1$ 個座位，而 T_k 在其中均勻隨機地選擇。尚未到達的乘客是 $T_{k+1}, \ldots, T_n$，各人都想坐自己的座位，而尚未到達的人之中沒有人被分配到 S_1。

想一想 為甚麼這恰好就是有 $n-k+1$ 個座位的原問題？

重新標記：把 T_k 稱為「第一位乘客」，並當作 S_1 是 T_k 獲分配的座位。在有 m 個座位的原問題中，第一位乘客從全部 m 個座位中均勻地選擇，之後每位乘客若自己的座位空着便坐下，否則從空座位中均勻地選擇，而第一位乘客獲分配的座位沒有其他人想要。這裏取 $m = n-k+1$，乘客 $T_k, T_{k+1}, \ldots, T_n$ 在座位 $S_1, S_{k+1}, \ldots, S_n$ 上正好遵循這些規則，而 S_1 沒有其他人想要。這些規則除了用來判斷乘客自己的座位是否空着之外，從不用到座位的標籤，所以兩個過程的概率完全相同。在較小的問題中，T_n 仍是最後一位乘客，所以

$$\mathrm{P}(T_n \text{ 坐在} S_n \mid T_1 \text{ 坐在} S_k) = P_{n-k+1}.$$

現在，對 $n \ge 3$，以 T_1 的座位為條件，$S_1, \ldots, S_n$ 每個的概率均為 $\frac{1}{n}$。若是 S_1，T_n 必定坐到 S_n；若是 S_n，T_n 必定坐不到。所以由全概率公式（law of total probability），

$$P_n = \frac{1}{n}\cdot 1 + \frac{1}{n}\sum_{k=2}^{n-1} P_{n-k+1} + \frac{1}{n}\cdot 0 = \frac{1}{n}\left(1 + \sum_{r=2}^{n-1} P_r\right),$$

其中我們代入了 $r = n-k+1$，當 k 由 2 變到 $n-1$ 時，它由 $n-1$ 遞減至 2。這就是所求的結果。

第 (iii) 部分

我們斷言對每個 $n \ge 2$ 都有 $P_n = \frac{1}{2}$。我們使用強歸納法（strong induction）。

*基本情況。*由第 (i) 部分得 $P_2 = \frac{1}{2}$。

*歸納步驟。*設 $n \ge 3$，並假設對 $r = 2, 3, \ldots, n-1$ 有 $P_r = \frac{1}{2}$。第 (ii) 部分中的和有 $n-2$ 項，每項都等於 $\frac{1}{2}$，所以

$$P_n = \frac{1}{n}\left(1 + \frac{n-2}{2}\right) = \frac{1}{n}\cdot\frac{n}{2} = \frac{1}{2}.$$

因此對 $r = 2, \ldots, n$ 有 $P_r = \frac{1}{2}$，由歸納法，對所有 $n \geq 2$ 都有 $P_n = \frac{1}{2}$。（第 (i) 部分的 $P_3 = \frac{1}{2}$ 與此吻合。）

第 (iv) 部分

當 $n = 2$ 時，$T_{n-1} = T_1$，他以概率 $\frac{1}{2}$ 坐在 S_1。所以 $Q_2 = \frac{1}{2}$。

現在設 $n \geq 3$，於是 T_{n-1} 不是 T_1，再次以 T_1 的座位 S_k 為條件：

- $k = 1$：每個人都坐到自己的座位，所以 T_{n-1} 坐到 S_{n-1}。
- $k = n$：$T_2, \ldots, T_{n-1}$ 都發現自己的座位空着，所以 T_{n-1} 同樣坐到 S_{n-1}。
- $k = n-1$：T_{n-1} 必定坐不到 S_{n-1}。
- $2 \leq k \leq n-2$：與第 (ii) 部分完全一樣，餘下的旅程就是以 T_k 為首、有 $m = n - k + 1$ 個座位的原問題。在這列較小的列車中，T_{n-1} 是倒數第二位乘客，所以概率為 Q_{n-k+1}。這裏 m 取遍 $3, 4, \ldots, n-1$。

因此，對 $n \geq 3$，

$$Q_n = \frac{1}{n}\left(2 + \sum_{m=3}^{n-1} Q_m\right).$$

當 $n = 3$ 時和為空，故 $Q_3 = \frac{2}{3}$。若對 $3 \leq m \leq n-1$ 有 $Q_m = \frac{2}{3}$，則

$$Q_n = \frac{1}{n}\left(2 + \frac{2(n-3)}{3}\right) = \frac{1}{n} \cdot \frac{2n}{3} = \frac{2}{3},$$

所以由強歸納法，對每個 $n \geq 3$ 都有 $Q_n = \frac{2}{3}$。總結：

$$Q_2 = \tfrac{1}{2}, \qquad Q_n = \tfrac{2}{3} \quad (n \geq 3).$$

> **想一想** 有沒有方法不用任何遞推關係就看出 $\frac{1}{2}$ 和 $\frac{2}{3}$？

有的。每當有人隨機選座時，S_1 與每個後來乘客仍然空着的座位都受到同等對待，所以在過程結束之前，S_1 和 S_n 兩者中先被坐的那個是哪一個，可能性相等；T_n 成功恰好當 S_1 先被坐，得 $\frac{1}{2}$。對於 T_{n-1}（其中 $n \geq 3$），考慮 S_1, S_{n-1}, S_n 之中哪一個最先被隨機選座者坐了；每個的可能性相等，而 T_{n-1} 只有在那個是 S_{n-1} 時才會被擠走，得 $1 - \frac{1}{3} = \frac{2}{3}$。

驗算

- $n = 3$ 的情況可以用手列舉：(T_1, T_2, T_3) 的結果是 (S_1, S_2, S_3)，概率為 $\frac{1}{3}$；(S_3, S_2, S_1)，概率為 $\frac{1}{3}$；以及 (S_2, S_1, S_3)、(S_2, S_3, S_1)，概率各為 $\frac{1}{6}$。所以 T_3 坐到 S_3 的概率為 $\frac{1}{2}$，T_2 坐到 S_2 的概率為 $\frac{2}{3}$，與所求得的一致。

- 用電腦對 $n = 2, \ldots, 8$ 精確列舉每一種可能的旅程，得出 $P_n = \frac{1}{2}$ 全部成立、對 $n \geq 3$ 有 $Q_n = \frac{2}{3}$，並證實第 (ii) 部分的條件概率（conditional probability）對每個 k 都成立。以 $n = 20$ 模擬 10^5 次旅程，得到 0.500 和 0.668。

常見錯誤

- 忘記 T_1 本身可以選 S_1，這正是遞推關係中「$1+$」的來源。
- 在第 (iv) 部分漏掉 T_1 坐了 S_n 的情況：這對 T_{n-1} 是好消息，所以常數是 2 而不是 1。
- 在 $n = 2$ 時套用 Q 的遞推關係，但此時 T_{n-1} 就是 T_1，答案是 $\frac{1}{2}$ 而不是 $\frac{2}{3}$。

STEP 之外

這裏的遞推關係是一種首步分析（first-step analysis）：以第一次隨機選擇為條件，並認出同一問題的較小版本。同樣的技巧可以求出馬爾可夫鏈（Markov chain）的命中概率（hitting probability）和期望命中時間（expected hitting time），這是大學一年級概率論的核心課題。上面的對稱性論證讓人初嘗一個選得好的「哪件事先發生？」問題如何取代冗長的計算。Grimmett 與 Welsh 的 *Probability: An Introduction*（《概率導論》）對這兩種思路都有闡述。

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。