詳細解答

STEP I 1989，第 15 題

概率與統計 • 分佈的比較

> **STEP I 1989，第 15 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/89-s1-q15

關鍵思路

以 $12\,\mathrm{km\,h^{-1}}$ 的速率行駛，每公里剛好需時 5 分鐘，所以每條路線的行程時間都是一個固定的分鐘數，再加上某個隨機變量（random variable）的簡單倍數：這個變量分別服從均勻分佈（uniform distribution）、泊松分佈（Poisson distribution）或二項分佈（binomial distribution）。於是，每一個關於時間的問題，都變成關於這一個變量的問題。中間的一問則是貝葉斯定理（Bayes' theorem）的應用，其中三條路線被選中的可能性相等。

解答

行程時間

設 T_A、T_B、T_D 為三條路線的行程時間，以分鐘計。由於 $12\,\mathrm{km\,h^{-1}}$ 即每 5 分鐘行 1 km：

- Angle Avenue：$T_A = 25 + A$，其中 $A \sim \mathrm{U}(0, 10)$。
- Bend Boulevard：$T_B = 20 + 3B$，其中 $B \sim \mathrm{Po}(4)$。
- Detour Drive：設 X 為繞道的次數。假設五處道路工程互相獨立（independent），每處導致繞道的概率均為 $\frac{4}{5}$，則 $X \sim \mathrm{B}(5, \frac{4}{5})$，路程為 $2 + X$ km，而 $T_D = 10 + 5X$。

至少 27 分鐘

Angle Avenue。$T_A \geq 27 \iff A \geq 2$，所以

$$\mathrm{P}(T_A \geq 27) = \frac{10 - 2}{10} = \frac{4}{5}.$$

Bend Boulevard。$T_B \geq 27 \iff 3B \geq 7 \iff B \geq 3$，因為 B 是整數。所以

$$\mathrm{P}(T_B \geq 27) = 1 - \mathrm{e}^{-4}\left(1 + 4 + \frac{4^2}{2!}\right) = 1 - 13\mathrm{e}^{-4} \approx 0.7619.$$

Detour Drive。$T_D \geq 27 \iff 5X \geq 17 \iff X \geq 4$。所以

$$\mathrm{P}(T_D \geq 27) = 5\left(\frac{4}{5}\right)^4\frac{1}{5} + \left(\frac{4}{5}\right)^5 = \left(\frac{4}{5}\right)^4\left(1 + \frac{4}{5}\right) = \frac{256}{625}\cdot\frac{9}{5} = \frac{2304}{3125} = 0.73728.$$

已知行程較長，走的是哪條路線？

設 R 為所選的路線，其中 $\mathrm{P}(R = A) = \mathrm{P}(R = B) = \mathrm{P}(R = D) = \frac{1}{3}$，並設 L 為行程時間至少 27 分鐘這一事件。根據貝葉斯定理，各個 $\frac{1}{3}$ 因子互相約去：

$$\begin{aligned}\mathrm{P}(R = B \mid L) &= \frac{\mathrm{P}(L \mid B)}{\mathrm{P}(L \mid A) + \mathrm{P}(L \mid B) + \mathrm{P}(L \mid D)} = \frac{1 - 13\mathrm{e}^{-4}}{\frac{4}{5} + 1 - 13\mathrm{e}^{-4} + \frac{2304}{3125}} \\ &= \frac{1 - 13\mathrm{e}^{-4}}{\frac{7929}{3125} - 13\mathrm{e}^{-4}} \approx \frac{0.7619}{2.2992} \approx 0.331.\end{aligned}$$

這個數值只比 $\frac{1}{3}$ 略高，十分合理：三條路線令行程達 27 分鐘或以上的可能性幾乎相同。

第 (i) 部分

利用 $\mathrm{E}(A) = 5$、$\mathrm{E}(B) = 4$ 和 $\mathrm{E}(X) = 5 \times \frac{4}{5} = 4$，以及期望值的線性性質（linearity of expectation），

$$\mathrm{E}(T_A) = 25 + 5 = 30, \qquad \mathrm{E}(T_B) = 20 + 3 \times 4 = 32, \qquad \mathrm{E}(T_D) = 10 + 5 \times 4 = 30.$$

? **想一想** 有兩條路線打成平手。這是計錯了嗎？

不是。Angle Avenue 和 Detour Drive 的平均時間都剛好是 30 分鐘，而 Bend Boulevard 的平均時間是 32 分鐘。所以要令平均時間盡量短，我應該選用 Angle Avenue 或 Detour Drive 其中任何一條：兩者同樣好，而且都勝過 Bend Boulevard。

第 (ii) 部分

現在比較各條路線的 $\mathrm{P}(T < 32)$。

- $T_A < 32 \iff A < 7$，所以 $\mathrm{P}(T_A < 32) = \frac{7}{10} = 0.7$。
- $T_B < 32 \iff 3B < 12 \iff B \leq 3$，所以

$$\mathrm{P}(T_B < 32) = \mathrm{e}^{-4}\left(1 + 4 + 8 + \frac{32}{3}\right) = \frac{71}{3}\mathrm{e}^{-4} \approx 0.4335.$$

- $T_D < 32 \iff 5X < 22 \iff X \leq 4$，所以

$$\mathrm{P}(T_D < 32) = 1 - \left(\frac{4}{5}\right)^5 = 1 - \frac{1024}{3125} = \frac{2101}{3125} \approx 0.6723.$$

由於 $0.7 > 0.6723 > 0.4335$，我應該每次都選用 **Angle Avenue**，這樣有 70% 的日子可以在 32 分鐘內到達。

驗算

- 模擬 400 000 次行程，得出 $\mathrm{P}(T \geq 27) \approx 0.801,\ 0.761,\ 0.740$，平均時間 $\approx 30.0,\ 32.0,\ 30.0$ 分鐘，$\mathrm{P}(T < 32) \approx 0.700,\ 0.435,\ 0.670$，以及 $\mathrm{P}(R = B \mid L) \approx 0.331$。全部與精確值吻合。
- 檢查邊界情況是否合理：在 Detour Drive 上，繞道 3 次需時 25 分鐘，繞道 4 次則需時 30 分鐘，所以「至少 27 分鐘」確實等同於「至少繞道 4 次」。

常見錯誤

- 處理整數條件時不夠嚴謹：$3B \geq 7$ 表示 $B \geq 3$，而不是 $B \geq 2$；$3B < 12$ 表示 $B \leq 3$，而不是 $B \leq 4$。
- 忘記 Detour Drive 本身有 2 km 的基本路程，或者忘記其他路線上每公里五分鐘的騎車時間。
- 在 (i) 中聲稱只有一條路線勝出。其實兩個平均值完全相等。

STEP 之外

第 (i) 和第 (ii) 部分說明，「最好」取決於我們要優化的是甚麼：平均值只是概括一個分佈的其中一個數字，兩條平均值相同的路線，其離散程度和尾部概率可以相差很遠。在決策理論（decision theory）中，這就是如何選擇*損失函數*（loss function）的問題。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教科書）涵蓋了本題用到的各種分佈和貝葉斯定理。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）則收錄了許多這一類的題目。

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。