詳細解答

STEP III 1995，第 13 題

概率與統計 • 有噪聲的二元信道

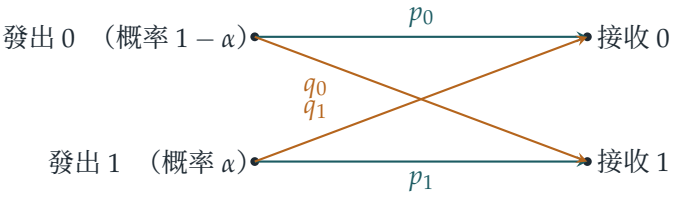
> **STEP III 1995，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/95-s3-q13

關鍵思路

一切都由逐個數字考慮而來。每個數字都經過一條「二元信道」（binary channel）：它以 0 或 1 發出，然後要麼原樣保留，要麼被翻轉。一旦弄清楚單一數字的情況，整段訊息不過是許多獨立的複本，所以概率相乘即可。至於增強站，竅門在於留意到：一個數字經過兩段電纜後正確到達，當且僅當它被翻轉了偶數次。

解答

全篇以 $N = 10^k$ 表示訊息的長度。我們按照題目的原意，假設原訊息的各個數字互相獨立（independent）（每個數字是 1 的概率為 α），並假設電纜對每個數字的處理與其他數字無關。單一數字的情況如下圖所示。



整段訊息正確接收

設 C 為某個指定數字被正確接收的事件。這個數字要麼是保留下來的 0，要麼是保留下來的 1，所以

$$\mathrm{P}(C) = (1-\alpha)p_0 + \alpha p_1.$$

由於 N 個數字互相獨立，

$$\mathrm{P}(\text{整段訊息正確}) = \left[(1-\alpha)p_0 + \alpha p_1\right]^{10^k}.$$

概率 β

一個數字被接收為 1，要麼它原本是 1 而且保留下來，要麼它原本是 0 而被翻轉：

$$\beta = \alpha p_1 + (1-\alpha)q_0.$$

現在

$$\begin{aligned}\beta = \alpha &\iff \alpha p_1 + (1-\alpha)q_0 = \alpha \iff (1-\alpha)q_0 = \alpha(1-p_1) = \alpha q_1 \\ &\iff q_0 = \alpha(q_0+q_1) \iff \alpha = \frac{q_0}{q_0+q_1},\end{aligned}$$

其中最後一步除以 q_0+q_1，而除非電纜完美無誤（$q_0 = q_1 = 0$，此時 $\beta = \alpha$ 顯然成立），否則這個數必為正數。每一步都可以逆推，所以這就是所要求的「當且僅當」(if and only if)。用文字來說：電纜不改變 1 所佔的比例，恰好就是 $0 \to 1$ 的流量（即 $(1-\alpha)q_0$）與 $1 \to 0$ 的流量（即 αq_1）互相抵銷之時。

全是零的訊息

假設上述條件成立，而且接收到的每個數字都是 0。由於各個位置上的（發出數字，接收數字）這些數對互相獨立，整段訊息正確的條件概率（conditional probability）就是各個單一數字的概率 $\mathrm{P}(\text{發出0} \mid \text{接收0})$ 的乘積。

> ? **想一想** 怎樣把「接收到 0」轉化為一句關於發出了甚麼的陳述？

用貝葉斯定理（Bayes' theorem）。一個數字被接收為 0 的概率是 $1-\beta$，而條件說明 $\beta = \alpha$，所以 $1-\beta = 1-\alpha$。因此

$$\mathrm{P}(\text{發出0} \mid \text{接收0}) = \frac{(1-\alpha)p_0}{1-\beta} = \frac{(1-\alpha)p_0}{1-\alpha} = p_0,$$

於是

$$\mathrm{P}(\text{訊息正確} \mid \text{接收到的全是零}) = p_0^{10^k}.$$

一半一半的機會

當 $q_0 = q_1 = q$ 且 $\alpha = \frac{1}{2}$ 時，每個數字正確的概率是 $\frac{1}{2}(1-q) + \frac{1}{2}(1-q) = 1-q$。對於一段有 10^6 個數字的訊息，我們需要

$$(1-q)^{10^6} = \tfrac{1}{2} \quad\Longrightarrow\quad 10^6 \ln(1-q) = -\ln 2.$$

由於 q 極小，$\ln(1-q) \approx -q$，得出

$$q \approx \frac{\ln 2}{10^6} \approx 6.93 \times 10^{-7}.$$

加設增強站

長度減半會使 q 除以 $2^2 = 4$，因為 q 與長度的平方成正比。所以每半段電纜翻轉一個數字的概率是 $q/4$，而且與另一半段無關（增強站只是把收到的東西原樣再發出）。

> **想一想** 「兩半段都正確傳送」是數字正確到達的唯一途徑嗎？

不是：一個在兩半段都被翻轉的數字會被翻回原狀。所以

$$\mathrm{P}(\text{數字正確}) = \left(1-\frac{q}{4}\right)^2 + \left(\frac{q}{4}\right)^2 = 1 - \frac{q}{2} + \frac{q^2}{8} \approx 1 - \frac{q}{2},$$

這裏略去了 q^2 階的項。（在這裏，雙重翻轉正是這類項之一，所以它對答案沒有影響，但明白箇中原因仍然值得。）整段訊息正確的概率現在是

$$\left(1-\frac{q}{2}\right)^{10^6}.$$

在同一階的近似下，$(1-\frac{q}{2})^2 \approx 1-q$，所以

$$\left(1-\frac{q}{2}\right)^{10^6} \approx (1-q)^{10^6/2} = \left[(1-q)^{10^6}\right]^{1/2} = \left(\tfrac{1}{2}\right)^{1/2} = \frac{1}{\sqrt{2}} \approx 0.707.$$

所以增強站把訊息完全正確的機會由 $\frac{1}{2}$ 提高到約 0.71。留意我們在這裏根本用不着 q 的數值：原來的條件 $(1-q)^{10^6} = \frac{1}{2}$ 已經包辦了一切。同樣的想法配合 $e^{-10^6 q} = \frac{1}{2}$，可得 $e^{-10^6 q/2} = 1/\sqrt{2}$。

驗算

- 把長度為 4 的所有發出訊息和接收訊息逐一準確列舉（取 $p_0 = 0.9$，$p_1 = 0.8$），可以重現 $[(1-\alpha)p_0 + \alpha p_1]^4$，以及在 α 滿足該條件時的條件概率 $p_0^4 = 0.6561$。
- 準確解出 $(1-q)^{10^6} = \frac{1}{2}$ 可得 $q = 6.9315 \times 10^{-7}$，與 $\ln 2/10^6$ 吻合至五位有效數字。
- 取這個 q 值，增強站情況下的準確概率 $\left[(1-\frac{q}{4})^2 + (\frac{q}{4})^2\right]^{10^6}$ 是 0.707107，與 $1/\sqrt{2}$ 相符。
- 合理性檢查：若 $q_0 = q_1$，條件給出 $\alpha = \frac{1}{2}$，正如對稱性所要求。

常見錯誤

- 把每半段的出錯概率當作 $q/2$。出錯概率與長度的平方成正比，所以應是 $q/4$。

- 在「全是零」的部分只答 $(1-\alpha)^N$ 或單單一個 p_0：我們需要的是每個數字的條件概率，再取 10^k 次方。

STEP 之外

這就是信息論（information theory）中的二元信道。當 $q_0 = q_1$ 時，它就是二元對稱信道（binary symmetric channel），是香農（Shannon）理論的起點——這套理論研究有噪聲的信道能夠傳送多少信息；它也是糾錯碼（error-correcting codes）的起點，糾錯碼加入冗餘，使偶爾出現的翻轉可以被偵測和修復，而不僅僅是變得較為罕見。增強站的計算也是一條兩步的馬可夫鏈（Markov chain）：翻轉概率的複合，就是把轉移矩陣（transition matrices）相乘。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）以恰到好處的程度講解條件概率和馬可夫鏈。

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。