詳細解答

STEP II 1999，第 12 題

概率與統計 • 貝葉斯定理

STEP II 1999，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/99-s2-q12

關鍵思路

這是一道保留了參數的貝葉斯定理（Bayes' theorem）題目。只要用全概率公式（law of total probability）求出 $\mathrm{P}(S)$，兩個條件概率（conditional probability）都是 p 的簡單有理函數（rational function），而且各自在容許的區間上都是單調（monotonic）的。因此，每位檢查員的做法中所謂的「最大值」，都落在區間 $0 \le p \le \frac{1}{3}$ 的端點上；而真正有趣的地方，在於這對單憑一次觀察作估計有甚麼啟示。

解答

條件概率

設 A、B、C 分別為隨機抽出的一塊晶片由 A、B、C 製造的事件，並設 S 為該晶片完好的事件。這三個事件構成樣本空間（sample space）的一個分割（partition），其中

$$\mathrm{P}(A) = 2p, \quad \mathrm{P}(B) = p, \quad \mathrm{P}(C) = 1 - 3p,$$
$$\mathrm{P}(S \mid A) = \tfrac{7}{10}, \quad \mathrm{P}(S \mid B) = \tfrac{8}{10}, \quad \mathrm{P}(S \mid C) = \tfrac{9}{10}.$$

根據全概率公式，

$$\begin{aligned}\mathrm{P}(S) &= 2p \cdot \tfrac{7}{10} + p \cdot \tfrac{8}{10} + (1 - 3p) \cdot \tfrac{9}{10} \\ &= \tfrac{1}{10}(14p + 8p + 9 - 27p) = \tfrac{1}{10}(9 - 5p).\end{aligned}$$

留意在整個 $0 \le p \le \frac{1}{3}$ 的範圍內都有 $9 - 5p > 0$，所以可以用它作除數。根據貝葉斯定理，

$$\mathrm{P}(A \mid S) = \frac{\mathrm{P}(S \mid A)\,\mathrm{P}(A)}{\mathrm{P}(S)} = \frac{\frac{14p}{10}}{\frac{9-5p}{10}} = \frac{14p}{9-5p},$$

同理

$$\mathrm{P}(C \mid S) = \frac{\frac{9}{10}(1-3p)}{\frac{1}{10}(9-5p)} = \frac{9(1-3p)}{9-5p} = \frac{9-27p}{9-5p}.$$

第一位檢查員的估計

當 p 由 0 增加至 $\frac{1}{3}$ 時，分子 $14p$ 遞增，而正的分母 $9-5p$ 遞減，所以 $\mathrm{P}(A \mid S)$ 嚴格遞增（strictly increasing）。（等價地說，它的導數是 $126/(9-5p)^2 > 0$。）因此它的最大值出現在右端點 $p = \frac{1}{3}$，其值為

$$\frac{14/3}{9-5/3} = \frac{14/3}{22/3} = \frac{7}{11}.$$

所以第一位檢查員的估計是 $\hat{p} = \frac{1}{3}$。

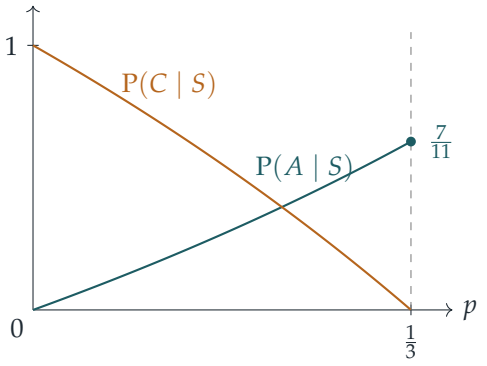
第二位檢查員的估計

想一想 $\mathrm{P}(C \mid S)$ 是遞增還是遞減？隨着 p 增大，它的分子和分母都在減小。

當分子和分母朝同一方向變化時，我們不能一眼看出答案，所以要用商法則（quotient rule）求導：

$$\frac{\mathrm{d}}{\mathrm{d}p}\frac{9-27p}{9-5p} = \frac{-27(9-5p)+5(9-27p)}{(9-5p)^2} = \frac{-243+135p+45-135p}{(9-5p)^2} = \frac{-198}{(9-5p)^2} < 0.$$

所以 $\mathrm{P}(C \mid S)$ 在 $[0, \frac{1}{3}]$ 上嚴格遞減，它的最大值出現在 $p = 0$，其值為 1。第二位檢查員的估計是 $\hat{p} = 0$。



評論

兩位檢查員用的是同一個方法，各自依據的單次觀察也同樣有效，卻得出容許範圍內兩個相反的極端：$\hat{p} = \frac{1}{3}$（完全沒有晶片來自 C）和 $\hat{p} = 0$（所有晶片都來自 C）。每個估

計其實只是「令我這塊晶片的製造商盡可能常見的 p 值」，而只有一個樣本時，必然會得出這樣的結果。第二個估計甚至與第一位檢查員手上的晶片互相矛盾，因為 $p = 0$ 表示 A 根本不製造晶片。兩個估計都不可信：一次觀察所含的信息實在太少，不足以估計 p，我們需要更大的樣本（連同來自每個製造商的晶片數目）。

還有一點值得留意：晶片是否完好，幾乎不起任何作用。「完好且來自 A」的似然（likelihood）是 $\frac{7}{10} \cdot 2p$，它取得最大值的位置與 $2p$ 本身相同；因子 $\frac{7}{10}$ 與 p 無關。

驗算

- 當 $p = 0$ 時，所有晶片都來自 C，所以 $\mathrm{P}(C \mid S)$ 應為 1，而 $\mathrm{P}(A \mid S)$ 應為 0；兩條公式都與此吻合。
- 同理可得 $\mathrm{P}(B \mid S) = 8p/(9 - 5p)$，而三個條件概率之和為

$$\frac{14p + 8p + 9 - 27p}{9 - 5p} = 1.$$

- 當 $p = 0.2$ 時，公式給出 $\mathrm{P}(A \mid S) = 0.35$ 和 $\mathrm{P}(C \mid S) = 0.45$；以 400 000 塊晶片進行模擬，得出 0.3505 和 0.4505。

常見錯誤

- 令導數等於零，然後尋找轉向點（turning point）。這裏根本沒有轉向點：兩個最大值都在 $[0, \frac{1}{3}]$ 的端點上。
- 忘記 $\mathrm{P}(C \mid S)$ 需要仔細論證：分子和分母都在減小，所以單說「分子變小了」並不足夠。

STEP 之外

兩位檢查員所做的，實際上就是最大似然估計（maximum likelihood estimation）：選出令他們所見的結果概率最大的參數值。只有一次觀察時，最大似然估計值往往落在參數空間（parameter space）的邊界上，這提醒我們：它需要大量數據才會表現良好。貝葉斯學派（Bayesian）的做法則是先為 p 設定一個先驗分佈（prior distribution），然後每觀察一塊晶片就更新一次，這樣便不會單憑一次觀察而得出如此極端的結論。Blitzstein 與 Hwang 所著的 *Introduction to Probability*（概率論入門教材）對這兩種觀點都有討論。

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。