詳細解答

STEP II 2003，第 12 題

概率與統計 • 分支過程

STEP II 2003，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/03-s2-q12

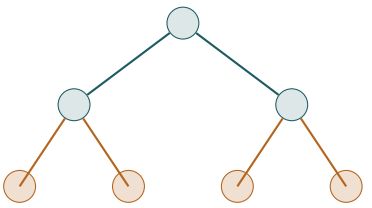
本題在 *Beyond STEP* 第 12.15 節滅絕還是爆發中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/branching.html

關鍵思路

這是一個分支過程（branching process）：每一代都由上一代產生，而每顆粒子各自獨立地分裂。整道題只靠一個事實：若 m 顆粒子各自獨立地產生後代，而每顆的後代數目服從 $\mathrm{B}(2, p)$，則後代總數服從 $\mathrm{B}(2m, p)$。因此，無論是求概率還是求期望值（expectation），我們都以上一代的大小作為條件。

解答

我們按照模型的原意，假設不同粒子的衰變彼此獨立，亦與過去的情況獨立。記 $X_0 = 1$，代表在 $0 < t < T$ 期間存活的那一顆粒子。在 $t = nT$ 時，存活的 X_{n-1} 顆粒子各自被自己的後代取代，所以 X_n 就是第 $n-1$ 代的後代總數。



$X_0 = 1$

$X_1 = 2$

$X_2 \leq 4$

圖中所示為 $X_1 = 2$ 的情況：這兩顆粒子合共最多有四個後代，而四個可能的後代各自

以概率 p 獨立地誕生。

聯合概率

由於 $X_1 \sim \mathrm{B}(2,p)$，我們有 $\mathrm{P}(X_1=2)=p^2$。已知 $X_1=2$ 時，兩顆粒子各自獨立地產生後代，每顆的後代數目服從 $\mathrm{B}(2,p)$，所以 X_2 就是 4 次獨立試驗中成功的次數，而每次成功的概率為 p；即 $X_2 \mid X_1=2 \sim \mathrm{B}(4,p)$。因此

$$\mathrm{P}(X_1=2 \text{ 及} X_2=2)=\mathrm{P}(X_1=2)\,\mathrm{P}(X_2=2 \mid X_1=2)=p^2\cdot\binom{4}{2}p^2q^2=6p^4q^2,$$

正如題目所求。

條件概率

要運用條件概率（conditional probability）的定義，我們需要 $\mathrm{P}(X_2=2)$，這可以透過以 X_1 為條件求得：

- 若 $X_1=0$，整個羣體已經滅絕，所以 $X_2=0$；
- 若 $X_1=1$（概率為 $2pq$），則 $X_2\sim\mathrm{B}(2,p)$，而 $\mathrm{P}(X_2=2 \mid X_1=1)=p^2$；
- 若 $X_1=2$，其貢獻就是剛才求得的 $6p^4q^2$。

所以

$$\mathrm{P}(X_2=2)=2pq\cdot p^2+6p^4q^2=2p^3q\,(1+3pq),$$

因此，當 $0<p<1$ 時，

$$\mathrm{P}(X_1=2 \mid X_2=2)=\frac{6p^4q^2}{2p^3q(1+3pq)}=\frac{3pq}{1+3pq}.$$

令上式等於 $\frac{9}{25}$，得 $75pq=9+27pq$，所以 $48pq=9$，即 $pq=\frac{3}{16}$。代入 $q=1-p$，這就是

$$p^2-p+\tfrac{3}{16}=0 \quad\Longleftrightarrow\quad (p-\tfrac{1}{4})(p-\tfrac{3}{4})=0,$$

所以 $p=\frac{1}{4}$ 或 $p=\frac{3}{4}$。（$p=0$ 和 $p=1$ 這兩種情況須排除：當 $p=0$ 時，事件 $X_2=2$ 不可能發生；當 $p=1$ 時，條件概率為 0。）

> **想一想** 為甚麼 p 的兩個值會給出相同的答案？

條件概率只透過 $pq=p(1-p)$ 依賴於 p，而把 p 換成 $1-p$ 時，這個量保持不變。所以方程的根必定成對出現，即 p 與 $1-p$，在這裏就是 $\frac{1}{4}$ 和 $\frac{3}{4}$。

期望值

設 $X_{n-1} = m$。這 m 顆粒子每顆平均產生 $\mathrm{E}(X) = 2p$ 個後代，所以下一代大小的期望值為 $2pm$：

$$\mathrm{E}(X_n \mid X_{n-1} = m) = 2pm.$$

對 m 的所有可能取值求平均（全期望值定律，law of total expectation），

$$\mathrm{E}(X_n) = \sum_m \mathrm{P}(X_{n-1} = m)\, 2pm = 2p\,\mathrm{E}(X_{n-1}).$$

用文字來說：在某一段時間內存活的每顆粒子，平均為下一段時間貢獻 $2p$ 顆粒子。由於 $\mathrm{E}(X_1) = 2p$（或 $\mathrm{E}(X_0) = 1$），這個遞推關係(recurrence)給出一個等比數列(geometric sequence)：

$$\mathrm{E}(X_n) = (2p)^n.$$

極限

當 $p = \frac{1}{4}$ 時，我們得到 $\mathrm{E}(X_n) = \left(\frac{1}{2}\right)^n \to 0$（當 $n \to \infty$），因為 $0 < \frac{1}{2} < 1$。

當 $p = \frac{3}{4}$ 時，我們得到 $\mathrm{E}(X_n) = \left(\frac{3}{2}\right)^n \to +\infty$，因為 $\frac{3}{2} > 1$。

由此可見，p 的這兩個值在條件概率中看似毫無分別，長遠表現卻截然不同：平均而言，當 $p = \frac{1}{4}$ 時羣體逐漸消亡，而當 $p = \frac{3}{4}$ 時羣體則急劇膨脹。

驗算

- 取 $p = \frac{3}{4}$：$6p^4q^2 = \frac{243}{2048} \approx 0.1187$，而模擬 10^6 個兩代羣體所得的結果是 0.1186。
- 當 $p = \frac{1}{4}$ 時，$pq = \frac{3}{16}$，所以 $3pq = \frac{9}{16}$，而 $\dfrac{9/16}{25/16} = \frac{9}{25}$，與預期相符。
- 對若干個 p 值計算 $X_1, \ldots, X_5$ 的精確分佈（已知上一代時，每一代服從 $\mathrm{B}(2m, p)$），所得的平均值恰好是 $(2p)^n$。

常見錯誤

- 在求 $\mathrm{P}(X_2 = 2)$ 時，遺漏了 $X_1 = 1$、$X_2 = 2$ 這條路徑；至於經過 $X_1 = 0$ 的路徑，其貢獻確實為零。
- 把 $X_2 \mid X_1 = 2$ 當作 $2 \times \mathrm{B}(2, p)$（即一顆粒子後代數目的兩倍），而不是兩個獨立副本之和，後者才是 $\mathrm{B}(4, p)$。

STEP 之外

這是一個高爾頓–沃森（Galton–Watson）分支過程，最初是為了解釋姓氏為何會消失而提出的。後代數目的平均值 $\mu = 2p$ 決定了一切：若 $\mu \leq 1$，羣體以概率 1 滅絕；若 $\mu > 1$，羣體則有正概率永遠存活下去。即使 $p = \frac{3}{4}$，羣體仍然有 $\frac{1}{9}$ 的概率滅絕，這個數值是 $s = (q + ps)^2$ 的最小根，而這條方程來自概率母函數（probability generating function）。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）以母函數處理分支過程，其講法正好可以直接承接本題。

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。