詳細解答

STEP I 2010，第 12 題

概率與統計 • 以尾和求期望值

> **STEP I 2010，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/10-s1-q12

關鍵思路

尾和公式（tail-sum formula）的來源，是把 $n\,\mathrm{P}(X=n)$ 寫成 $\mathrm{P}(X=n)$ 相加 n 次，然後把各項按另一種次序重新歸併。就企鵝這個問題而言，事件 $X \geq n$ 比 $X = n$ 容易描述得多：開了 $n-1$ 個盒之後仍須再開，當且僅當這 $n-1$ 個盒裏全是同一種企鵝。因此尾概率（tail probability）十分簡單，而剛證明的公式就把 $\mathrm{E}(X)$ 化成兩個等比級數（geometric series）。

解答

定義與尾和公式

由於 X 只取 $1, 2, 3, \ldots$ 這些值，

$$\mathrm{E}(X) = \sum_{m=1}^{\infty} m\,\mathrm{P}(X=m),$$

前提是這個級數收斂（converge）。

對每個 $n \geq 1$，事件 $X \geq n$ 是各事件 $X = m$（$m \geq n$）的不相交併集（disjoint union），所以 $\mathrm{P}(X \geq n) = \sum_{m \geq n} \mathrm{P}(X=m)$。因此

$$\begin{aligned}\sum_{n=1}^{\infty} \mathrm{P}(X \geq n) &= \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} \mathrm{P}(X=m) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{m} \mathrm{P}(X=m) \\ &= \sum_{m=1}^{\infty} m\,\mathrm{P}(X=m) = \mathrm{E}(X).\end{aligned}$$

想一想 為甚麼可以對調兩個求和的次序？

這個二重和是把 $\mathrm{P}(X=m)$ 對所有滿足 $1 \le n \le m$ 的數對 (n,m) 相加。第一行是固定 n，對 m 求和；第二行是固定 m，對 $n=1,\ldots,m$ 這 m 個值求和。每一項都是非負的，而對於非負項級數，總和與求和的次序無關（兩邊都可能是無窮大，若然則兩邊同為無窮大）。不妨把各項想像成排成一個三角形，第 m 行有 m 個 $\mathrm{P}(X=m)$：逐行相加得 $\mathrm{E}(X)$，逐列相加得尾和。公式得證。

尾概率

按題目原意，假設不同盒子的內容互相獨立（independent）。設 D_k 為第 k 個盒內是企鵝爸爸的事件，則 $\mathrm{P}(D_k)=p$，$\mathrm{P}(D_k')=q$。

取 $n \ge 2$。$X \ge n$ 成立，當且僅當首 $n-1$ 個盒尚未湊齊兩種企鵝，即它們全是爸爸或全是媽媽。這兩種情況互不相交，故由獨立性得

$$\mathrm{P}(X \ge n) = p^{n-1} + q^{n-1} \qquad (n \ge 2).$$

特別地，取 $n=4$，

$$\mathrm{P}(X \ge 4) = p^3 + q^3,$$

正是所求。此外，$X \ge 2$ 必然成立，因為一個盒不可能同時裝有兩種企鵝，所以 $\mathrm{P}(X \ge 1) = \mathrm{P}(X \ge 2) = 1$。（上面的公式在 $n=2$ 時給出 $p+q=1$，與此相符；但在 $n=1$ 時卻給出 2，所以 $n=1$ 這一項必須另行處理。）

期望值

由尾和公式，

$$\begin{aligned}\mathrm{E}(X) &= \mathrm{P}(X \ge 1) + \sum_{n=2}^{\infty} (p^{n-1} + q^{n-1}) \\ &= 1 + \sum_{k=1}^{\infty} p^k + \sum_{k=1}^{\infty} q^k \\ &= 1 + \frac{p}{1-p} + \frac{q}{1-q},\end{aligned}$$

其中兩個等比級數都收斂，因為 $0<p<1$ 及 $0<q<1$（這裏正是用到 $p \ne 0$、$q \ne 0$ 的地方）。由於 $1-p=q$ 及 $1-q=p$，

$$\mathrm{E}(X) = 1 + \frac{p}{q} + \frac{q}{p} = 1 + \frac{p^2+q^2}{pq}.$$

而 $p^2+q^2 = (p+q)^2 - 2pq = 1-2pq$，所以

$$\mathrm{E}(X) = 1 + \frac{1-2pq}{pq} = 1 + \frac{1}{pq} - 2 = \frac{1}{pq} - 1,$$

正是所求。

最後一部分

對正數 p 和 q 應用算術平均—幾何平均不等式（AM–GM inequality），

$$\sqrt{pq} \le \frac{p+q}{2} = \frac{1}{2}, \qquad \text{所以} \qquad 0 < pq \le \frac{1}{4}.$$

想一想 取倒數之後，不等號朝哪個方向？

由於 $pq > 0$，取倒數會令不等號反向：$\dfrac{1}{pq} \ge 4$。因此

$$\mathrm{E}(X) = \frac{1}{pq} - 1 \ge 4 - 1 = 3,$$

等號成立當且僅當 $p = q = \frac{1}{2}$。（等價地，$pq = p(1-p) = \frac{1}{4} - (p - \frac{1}{2})^2 \le \frac{1}{4}$。）

驗算

- 取 $p = q = \frac{1}{2}$：$X - 1$ 是等到第一個與第 1 盒不同的盒子所需的次數，是一個平均值為 2 的幾何變量（geometric variable），所以 $\mathrm{E}(X) = 3$，與 $\frac{1}{1/4} - 1$ 吻合。
- 當 $p \to 0$ 時，$\mathrm{E}(X) \to \infty$：企鵝爸爸十分罕有，所以要等很久。這與 $\frac{1}{pq} - 1$ 趨向無窮大相符。
- 直接計算，對 $n \ge 2$ 有 $\mathrm{P}(X = n) = p^{n-1}q + q^{n-1}p$。取 $p = \frac{1}{3}$，以數值方法把 $n\,\mathrm{P}(X = n)$ 相加得 3.5；另外模擬 200 000 名收集者（取 $p = 0.3$），所得平均值為 3.761，而準確值是 $\frac{1}{0.21} - 1 \approx 3.762$。

常見錯誤

- 在 $n = 1$ 時也使用 $p^{n-1} + q^{n-1}$，結果得出一個等於 2 的「概率」，期望值亦因而多了 1。
- 寫成 $\mathrm{P}(X \ge 4) = p^4 + q^4$：只要首三個盒相同，就需要開第四個盒。
- 在最後一部分忘記了正數之間的不等式取倒數後會反向。

STEP 之外

尾和公式有一個連續版本的孿生公式：對於非負隨機變量（random variable），$\mathrm{E}(X) = \int_0^\infty \mathrm{P}(X > x)\,\mathrm{d}x$，證明方法同樣是對調次序（在一般情況下，這一步由托

內利定理（Tonelli's theorem）保證成立）。企鵝問題是贈券收集問題（coupon collector's problem）中只有兩種類型的情況；若有 N 種機會均等的類型，所需盒數的期望值是 $N(1+\frac{1}{2}+\cdots+\frac{1}{N})$，約為 $N\ln N$。這兩個想法都見於 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門），以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）。

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。