詳細解答

STEP I 1996，第 13 題

概率與統計 • 離散分佈與期望值

STEP I 1996，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/96-s1-q13

本題在 *Beyond STEP* 第 12.11 節郵票、杯子與 e 中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/stamps-and-e.html

關鍵思路

遊戲在第一次猜錯時結束，所以 $N = k$ 的意思是「用 $2, 3, \ldots, k-1$ 個杯時都猜中，然後用 k 個杯時猜錯」。把這些概率（probability）相乘，分母便會出現階乘（factorial）；而求和式中出現階乘，正強烈暗示指數級數（exponential series）$\mathrm{e}^x = \sum x^n/n!$ 即將登場。計算期望值的竅門，是把 k 的因子與 $k!$ 約去，使每個和都變成 $\sum 1/n!$ 平移後的版本。

解答

N 的分佈

把 c 個杯打亂之後，Jim 實際上是隨機亂猜，所以他猜中的概率是 $1/c$，猜錯的概率是 $1 - 1/c$。每一回合都重新打亂一次，所以各回合互相獨立（independent）。

由於杯的數目至少有兩個，故 $N \geq 2$。對於 $k \geq 2$，事件 $N = k$ 發生，當且僅當 Jim 用 $2, 3, \ldots, k-1$ 個杯時都猜中，然後用 k 個杯時猜錯。由獨立性，

$$\begin{aligned} \mathrm{P}(N = k) &= \frac{1}{2} \cdot \frac{1}{3} \cdots \frac{1}{k-1} \cdot \left(1 - \frac{1}{k}\right) = \frac{1}{(k-1)!} \cdot \frac{k-1}{k} \\ &= \frac{k-1}{k!}, \qquad k = 2, 3, 4, \ldots \end{aligned}$$

（當 $k = 2$ 時，猜中部分的乘積是空積，我們得到 $\frac{1}{2}$，正如所料。）

值得留意的是，這個結果有一個簡潔的等價形式：$N > k$ 表示 Jim 用 $2, \ldots, k$ 個杯時都

猜中，所以

$$\mathrm{P}(N > k) = \frac{1}{k!} \qquad (k \geq 1).$$

事實上 $\frac{1}{(k-1)!} - \frac{1}{k!} = \frac{k-1}{k!}$，與上面的公式吻合。當 $k \to \infty$ 時，這個概率趨向 0，所以遊戲以概率 1 結束，而各概率之和為 1。

期望值

把 k 與 $k!$ 約去，

$$\begin{aligned}\mathrm{E}(N) &= \sum_{k=2}^{\infty} k \cdot \frac{k-1}{k!} = \sum_{k=2}^{\infty} \frac{k(k-1)}{k!} = \sum_{k=2}^{\infty} \frac{1}{(k-2)!} \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} = \mathrm{e},\end{aligned}$$

其中設 $j = k-2$，並在 $x = 1$ 處使用 $\mathrm{e}^x = \sum_{j\geq 0} x^j/j!$。這就是所要求的結果。

方差

> **想一想** 這裏哪一個二階矩（second moment）最容易求和：$\mathrm{E}(N^2)$ 還是 $\mathrm{E}\big(N(N-1)\big)$？

兩者皆可，因為 $k^2\,\mathrm{P}(N = k) = k \cdot \dfrac{k(k-1)}{k!}$ 中的因子 $k(k-1)$ 本身已能乾淨利落地約去。我們求得

$$\begin{aligned}\mathrm{E}(N^2) &= \sum_{k=2}^{\infty} \frac{k^2(k-1)}{k!} = \sum_{k=2}^{\infty} \frac{k}{(k-2)!} = \sum_{j=0}^{\infty} \frac{j+2}{j!} \\ &= \sum_{j=1}^{\infty} \frac{1}{(j-1)!} + 2\sum_{j=0}^{\infty} \frac{1}{j!} = \mathrm{e} + 2\mathrm{e} = 3\mathrm{e},\end{aligned}$$

其中在中間一步，$\sum j/j!$ 的 $j = 0$ 項為零，而當 $j \geq 1$ 時 $j/j! = 1/(j-1)!$。因此

$$\mathrm{Var}(N) = \mathrm{E}(N^2) - \big(\mathrm{E}(N)\big)^2 = 3\mathrm{e} - \mathrm{e}^2 = \mathrm{e}(3-\mathrm{e}) \approx 0.766.$$

驗算

- 各概率 $\frac{1}{2}, \frac{1}{3}, \frac{1}{8}, \frac{1}{30}, \ldots$ 之和為 1，這可由尾部形式 $\mathrm{P}(N > k) = 1/k!$ 看出。
- 利用尾部公式，$\mathrm{E}(N) = \sum_{k\geq 0} \mathrm{P}(N > k) = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots = \mathrm{e}$，這是得出同一答案的另一條獨立途徑。
- 方差（variance）為正數，這是必然的，因為 $\mathrm{e} < 3$。精確求和得 $\mathrm{E}(N) = 2.71828\ldots$ 及 $\mathrm{Var}(N) = 0.76579\ldots$，而模擬 400 000 局遊戲所得的平均值為 2.720，方差為 0.767。

常見錯誤

- 把分佈由 $k = 1$ 開始：遊戲一開始已有兩個杯，所以 $N \geq 2$。
- 漏掉代表猜錯的最後一個因子 $1 - \frac{1}{k}$，結果得出 $1/(k-1)!$，而這些「概率」之和是 $\mathrm{e} - 1$，而不是 1。
- 寫成 $\mathrm{Var}(N) = \mathrm{E}(N^2) - \mathrm{E}(N)$，而沒有減去 $\big(\mathrm{E}(N)\big)^2$。

STEP 之外

尾和公式（tail-sum formula）$\mathrm{E}(N) = \sum_{k \geq 0} \mathrm{P}(N > k)$ 對任何取值於 $\{0, 1, 2, \dots\}$ 的隨機變量（random variable）都成立，是大學一年級的標準工具，往往能把繁複的期望值計算化為一行算式，正如本題所見。要把所有矩一併整理起來，更簡潔的方法是使用概率生成函數（probability generating function）$G(s) = \mathrm{E}(s^N)$。在本題中可算出 $G(s) = 1 + (s-1)\mathrm{e}^s$，由此 $G'(1) = \mathrm{e}$ 及 $G''(1) = 2\mathrm{e}$ 立即給出平均值和方差。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門）正是以這種精神闡述生成函數。

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。