詳細解答

STEP II 2001，第 12 題

概率與統計 • 泊松近似

> **STEP II 2001，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/01-s2-q12

關鍵思路

合夥集團（syndicate）必定恰好持有一張中獎彩票，所以一切取決於有多少其他投注者選中了中獎號碼。這個數目服從二項分佈（binomial distribution），試驗次數極多而成功概率極小，正是泊松分佈（Poisson distribution）接手的情況。於是所得獎金的期望值（expectation）就是 W 乘以 $1/(1+X)$ 的期望值，而這個和可以把 $k+1$ 併入階乘（factorial），從而求得閉合形式（closed form）。

解答

建立模型

合夥集團買了 N 張彩票，所以其他投注者買了 $3.8N - N = 2.8N$ 張。設 X 為這些非集團彩票中印有中獎號碼的張數。

每名非集團投注者都是均勻隨機地選一個號碼，與其他人的選擇及開獎結果獨立，所以無論中獎號碼是甚麼，他們每張彩票與之相符的概率都是 $1/N$，而且各張彩票互相獨立。因此

$$X \sim \mathrm{B}\left(2.8N,\ \tfrac{1}{N}\right).$$

由於 N 非常大，$2.8N$ 很大而 $1/N$ 很小，同時平均值 $2.8N \cdot \frac{1}{N} = 2.8$ 保持不變，所以 X 近似服從平均值為 $\lambda = 2.8$ 的泊松分佈：

$$\mathrm{P}(X = k) \approx \frac{\mathrm{e}^{-2.8}\,(2.8)^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

合夥集團每個號碼都恰好持有一張彩票，所以它總是恰好持有一張中獎彩票，而中獎彩票的總數是 $T = 1 + X$。

最可能出現的中獎彩票數目

記 $p_k = \mathrm{P}(X = k)$。比較相鄰的概率，

$$\frac{p_k}{p_{k-1}} = \frac{\lambda}{k} = \frac{2.8}{k},$$

當 $k \le 2$ 時此比值大於 1，當 $k \ge 3$ 時小於 1。所以 $p_0 < p_1 < p_2$ 而 $p_2 > p_3 > p_4 > \cdots$，X 最可能取的值是 2。（對精確的二項分佈，結論相同：此時 $p_k/p_{k-1} = \frac{(2.8N-k+1)}{k(N-1)}$，它大於 1 當且僅當 $k < 2.8 + \frac{1}{N}$。）

因此最可能出現的中獎彩票數目是

$$T = 1 + 2 = 3.$$

淨損失的期望值

頭獎獎金由 $T = 1 + X$ 張中獎彩票平分，而合夥集團持有其中一張，所以它得到 $\dfrac{W}{1+X}$。它付出了 £N，所以淨損失是

$$L = N - \frac{W}{1+X}, \qquad \text{而} \qquad \mathrm{E}(L) = N - W\,\mathrm{E}\left(\frac{1}{1+X}\right).$$

> ? **想一想** 怎樣求 $\sum \frac{1}{k+1} \cdot \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$ 這個和？

竅門在於 $(k+1)\cdot k! = (k+1)!$，所以那個礙事的因子 $\frac{1}{k+1}$ 可以併入階乘。同時乘以和除以 λ，這個和就變成缺了一項的泊松級數：

$$\begin{aligned}\mathrm{E}\left(\frac{1}{1+X}\right) &\approx \sum_{k=0}^{\infty} \frac{1}{k+1} \cdot \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!} = \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} \\ &= \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{j=1}^{\infty} \frac{\lambda^j}{j!} = \frac{\mathrm{e}^{-\lambda}}{\lambda}\left(\mathrm{e}^{\lambda} - 1\right) = \frac{1 - \mathrm{e}^{-\lambda}}{\lambda},\end{aligned}$$

其中我們令 $j = k+1$，並利用 $\sum_{j\ge 0} \lambda^j/j! = \mathrm{e}^{\lambda}$，再減去 $j = 0$ 的一項，即 1。

代入 $\lambda = 2.8 = \frac{14}{5}$，上式等於 $\frac{5}{14}\left(1 - \mathrm{e}^{-2.8}\right)$，於是

$$\mathrm{E}(L) \approx N - \frac{5\left(1 - \mathrm{e}^{-2.8}\right)}{14} W,$$

正是所求的結果。

驗算

- 同一竅門也適用於精確的二項分佈 $X \sim \mathrm{B}(n, p)$：$\mathrm{E}\left(\frac{1}{1+X}\right) = \dfrac{1 - (1-p)^{n+1}}{(n+1)p}$，當 $n \to \infty$ 而 $np = \lambda$ 時，它趨向 $(1 - \mathrm{e}^{-\lambda})/\lambda$。取 $N = 1000$ 時，精確值是 0.33536，而泊松

近似值是 $\frac{5}{14}(1-\mathrm{e}^{-2.8}) \approx 0.33542$。

- 以 $N = 200$ 模擬 10^5 次開獎，所得獎金的平均值為 $0.335W$，而最常出現的結果是 3 張中獎彩票。
- 極限情況：如果沒有其他人投注（$\lambda \to 0$），則 $(1-\mathrm{e}^{-\lambda})/\lambda \to 1$，合夥集團獨得全部頭獎獎金，與預期相符。

常見錯誤

- 把其他投注者的彩票數目當作 $3.8N$ 而不是 $2.8N$：合夥集團自己的彩票並非隨機，它們恰好貢獻一張中獎彩票。
- 寫成 $\mathrm{E}\left(\frac{W}{1+X}\right) = \frac{W}{1+\mathrm{E}(X)} = \frac{W}{3.8}$。期望值不能直接穿過非線性函數；事實上 $\frac{1}{3.8} \approx 0.263$ 明顯小於 0.335（延森不等式（Jensen's inequality）指出它必定不會較大，因為 $1/(1+x)$ 是凸函數（convex function））。

STEP 之外

以 $\mathrm{Po}(\lambda)$ 逼近 $\mathrm{B}(n, \lambda/n)$，就是所謂的「稀有事件定律」（law of rare events）；大學一年級的課程會證明各個概率收斂，或者證明概率生成函數（probability generating function）收斂，從而證明這個定律。計算 $\mathrm{E}\big(1/(1+X)\big)$ 其實就是對生成函數求積分 $\int_0^1 G_X(s)\,\mathrm{d}s$，這是生成函數在計算平均值和方差（variance）以外的一個巧妙用途。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）涵蓋了這兩個概念。

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。