详细解答

STEP II 2001，第 12 题

概率与统计 · 泊松近似

> **STEP II 2001，第 12 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/01-s2-q12

关键思路

合伙集团（syndicate）必定恰好持有一张中奖彩票，所以一切取决于有多少其他投注者选中了中奖号码。这个数目服从二项分布（binomial distribution），试验次数极多而成功概率极小，正是泊松分布（Poisson distribution）接手的情况。于是所得奖金的期望值（expectation）就是 W 乘以 $1/(1+X)$ 的期望值，而这个和可以把 $k+1$ 并入阶乘（factorial），从而求得闭合形式（closed form）。

解答

建立模型

合伙集团买了 N 张彩票，所以其他投注者买了 $3.8N - N = 2.8N$ 张。设 X 为这些非集团彩票中印有中奖号码的张数。

每名非集团投注者都是均匀随机地选一个号码，与其他人的选择及开奖结果独立，所以无论中奖号码是甚么，他们每张彩票与之相符的概率都是 $1/N$，而且各张彩票互相独立。因此

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

由于 N 非常大，$2.8N$ 很大而 $1/N$ 很小，同时平均值 $2.8N \cdot \frac{1}{N} = 2.8$ 保持不变，所以 X 近似服从平均值为 $\lambda = 2.8$ 的泊松分布：

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

合伙集团每个号码都恰好持有一张彩票，所以它总是恰好持有一张中奖彩票，而中奖彩票的总数是 $T = 1 + X$。

最可能出现的中奖彩票数目

记 $p_k = \mathrm{P}(X = k)$。比较相邻的概率，

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

当 $k \le 2$ 时此比值大于 1，当 $k \ge 3$ 时小于 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。（对精确的二项分布，结论相同：此时 $p_k/p_{k-1} = \frac{(2.8N-k+1)}{k(N-1)}$，它大于 1 当且仅当 $k < 2.8 + \frac{1}{N}$。）

因此最可能出现的中奖彩票数目是

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

净损失的期望值

头奖奖金由 $T = 1 + X$ 张中奖彩票平分，而合伙集团持有其中一张，所以它得到 $\dfrac{W}{1+X}$。它付出了 $£N$，所以净损失是

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

> ? **想一想** 怎样求 $\sum \frac{1}{k+1} \cdot \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$ 这个和?

窍门在于 $(k+1)\cdot k! = (k+1)!$，所以那个碍事的因子 $\frac{1}{k+1}$ 可以并入阶乘。同时乘以和除以 λ，这个和就变成缺了一项的泊松级数：

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

其中我们令 $j = k+1$，并利用 $\sum_{j\ge 0} \lambda^j/j! = \mathrm{e}^{\lambda}$，再减去 $j = 0$ 的一项，即 1。

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

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

正是所求的结果。

验算

- 同一窍门也适用于精确的二项分布 $X \sim \mathrm{B}(n,p)$：$\mathrm{E}\left(\frac{1}{1+X}\right) = \dfrac{1-(1-p)^{n+1}}{(n+1)p}$，当 $n \to \infty$ 而 $np = \lambda$ 时，它趋向 $(1 - \mathrm{e}^{-\lambda})/\lambda$。取 $N = 1000$ 时，精确值是 0.33536，而泊松近

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

- 以 $N = 200$ 模拟 10^5 次开奖，所得奖金的平均值为 $0.335W$，而最常出现的结果是 3 张中奖彩票。
- 极限情况：如果没有其他人投注（$\lambda \to 0$），则 $(1-\mathrm{e}^{-\lambda})/\lambda \to 1$，合伙集团独得全部头奖奖金，与预期相符。

常见错误

- 把其他投注者的彩票数目当作 $3.8N$ 而不是 $2.8N$：合伙集团自己的彩票并非随机，它们恰好贡献一张中奖彩票。
- 写成 $\mathrm{E}\left(\frac{W}{1+X}\right) = \frac{W}{1+\mathrm{E}(X)} = \frac{W}{3.8}$。期望值不能直接穿过非线性函数；事实上 $\frac{1}{3.8} \approx 0.263$ 明显小于 0.335（延森不等式（Jensen's inequality）指出它必定不会较大，因为 $1/(1+x)$ 是凸函数（convex function））。

STEP 之外

以 $\mathrm{Po}(\lambda)$ 逼近 $\mathrm{B}(n, \lambda/n)$，就是所谓的“稀有事件定律”（law of rare events）；大学一年级的课程会证明各个概率收敛，或者证明概率生成函数（probability generating function）收敛，从而证明这个定律。计算 $\mathrm{E}\big(1/(1+X)\big)$ 其实就是对生成函数求积分 $\int_0^1 G_X(s)\,\mathrm{d}s$，这是生成函数在计算平均值和方差（variance）以外的一个巧妙用途。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门）涵盖了这两个概念。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。