详细解答

STEP III 2024，第 11 题

概率与统计 · 二项分布的期望值

STEP III 2024，第 11 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2024 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

第 (i) 部分的恒等式（identity）把和式下半部分的每一项 $r\binom{2n}{r}$ 与上半部分的一项配对起来。这种对称性（symmetry）正是处理 X 所需要的：正面与反面次数中较大者，不是正面的次数，就是反面的次数；把正面和反面的角色互换，便能把期望值（expectation）折叠到和式的上半部分。之后，第 (iii) 部分只是一个比值检验（ratio test），而第 (iv) 部分只需一行。

解答

第 (i) 部分

当 $1 \le r \le 2n$ 时，把 r 与 $r!$ 约去，得

$$r\binom{2n}{r} = \frac{r\,(2n)!}{r!\,(2n-r)!} = \frac{(2n)!}{(r-1)!\,(2n-r)!}.$$

另一边，$2n-(2n+1-r) = r-1$，所以

$$(2n+1-r)\binom{2n}{2n+1-r} = \frac{(2n+1-r)\,(2n)!}{(2n+1-r)!\,(r-1)!} = \frac{(2n)!}{(2n-r)!\,(r-1)!},$$

同样是把分子的因子与阶乘（factorial）约去。（我们需要 $1 \le r \le 2n$，使 $2n+1-r$ 介乎 1 与 $2n$ 之间，这样两个阶乘才有意义。）两个表达式相同，故恒等式成立。

至于和式，$r = 0$ 的项为零。把其余各项分成 $1 \le r \le n$ 和 $n+1 \le r \le 2n$ 两组。在较低的一组中，设 $s = 2n+1-r$。当 r 由 1 走到 n，s 便由 $2n$ 递减至 $n+1$，而根据恒等式，

$$\sum_{r=1}^{n} r\binom{2n}{r} = \sum_{r=1}^{n}(2n+1-r)\binom{2n}{2n+1-r} = \sum_{s=n+1}^{2n} s\binom{2n}{s}.$$

因此较低一组等于较高一组，从而

$$\sum_{r=0}^{2n} r\binom{2n}{r} = 2\sum_{r=n+1}^{2n} r\binom{2n}{r},$$

得证。

第 (ii) 部分

设 H 为正面的次数，则 $H \sim \mathrm{B}(2n, \frac{1}{2})$，且 $\mathrm{P}(H=r) = \binom{2n}{r}/2^{2n}$。反面的次数为 $2n-H$，所以 $X = \max(H, 2n-H)$。即是说，当 $H \geq n+1$ 时 $X = H$；当 $H = n$ 时 $X = n$；当 $H \leq n-1$ 时 $X = 2n-H$。因此

$$2^{2n}\,\mathrm{E}(X) = \sum_{r=n+1}^{2n} r\binom{2n}{r} + n\binom{2n}{n} + \sum_{r=0}^{n-1}(2n-r)\binom{2n}{r}.$$

在最后一个和式中设 $s = 2n-r$，它由 $2n$ 递减至 $n+1$。由于 $\binom{2n}{r} = \binom{2n}{2n-r}$，该和式变成 $\sum_{s=n+1}^{2n} s\binom{2n}{s}$，与第一个和式相同。（这就是正反面的对称性：得到 s 次反面的机会与得到 s 次正面的机会相同。）

> **想一想** $\sum_{r=n+1}^{2n} r\binom{2n}{r}$ 只涵盖一半的范围，该如何求它的值?

这正是第 (i) 部分的用途。题目已给出全范围的版本：以 $2n$ 代替 n，得 $\sum_{r=0}^{2n} r\binom{2n}{r} = 2n\cdot 2^{2n-1}$。所以由第 (i) 部分，

$$2\sum_{r=n+1}^{2n} r\binom{2n}{r} = 2n\cdot 2^{2n-1} = n\,2^{2n}.$$

因此

$$2^{2n}\,\mathrm{E}(X) = n\,2^{2n} + n\binom{2n}{n}, \qquad \text{所以} \qquad \mathrm{E}(X) = n\left(1 + \frac{1}{2^{2n}}\binom{2n}{n}\right),$$

得证。

第 (iii) 部分

设 $a_n = \binom{2n}{n}/2^{2n}$，它是正数。于是

$$\frac{a_{n+1}}{a_n} = \frac{(2n+2)!}{(n+1)!\,(n+1)!}\cdot\frac{n!\,n!}{(2n)!}\cdot\frac{1}{4} = \frac{(2n+2)(2n+1)}{4(n+1)^2} = \frac{2n+1}{2n+2},$$

这里用了 $2n+2 = 2(n+1)$。这个比值小于 1，而 $a_n > 0$，所以对每个 $n \geq 1$ 都有 $a_{n+1} < a_n$。此数列（sequence）严格递减（strictly decreasing）。

第 (iv) 部分

你赢得的金额正是第 (ii) 部分的 X，而你要付 £n，所以每付一英镑的期望赢得金额为

$$\frac{\mathrm{E}(X)}{n} = 1 + \frac{1}{2^{2n}}\binom{2n}{n} = 1 + a_n.$$

由第 (iii) 部分，n 最小时 a_n 最大，所以你应选 $n = 1$。这时每付一英镑可取回 $1 + \frac{1}{2} = \frac{3}{2}$ 英镑，即每英镑的期望利润为 50 便士（50p）。（若“赢得金额”指的是利润，则为 $\mathrm{E}(X)/n - 1 = a_n$，结论不变。）

验算

- 当 $n = 1$（掷两次）时，HH 或 TT 对应 $X = 2$，HT 或 TH 对应 $X = 1$，所以 $\mathrm{E}(X) = \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 1 = \frac{3}{2}$。公式给出 $1\cdot(1 + \frac{2}{4}) = \frac{3}{2}$。
- 当 $n = 2$ 时，公式给出 $2(1 + \frac{6}{16}) = \frac{11}{4}$，所以每英镑的回报为 $\frac{11}{8} = 1.375 < 1.5$。当 $n = 3$ 时，回报为 $\frac{21}{16} = 1.3125$，正如第 (iii) 部分所预测，仍在递减。
- 对于 $n \le 5$，已用电脑逐一检查全部 2^{2n} 个掷币序列，确认 $\mathrm{E}(X)$ 的公式正确。

常见错误

- 把 $H = n$ 的项计算两次，或完全漏掉。这一项不属于分割的任何一边，只贡献一次 $n\binom{2n}{n}$。
- 在第 (iv) 部分中，去最大化会无限增大的 $\mathrm{E}(X)$ 本身，而不是每付一英镑的 $\mathrm{E}(X)$。

STEP 之外

超出的部分 $\mathrm{E}(X) - n = na_n$ 是简单随机漫步（simple random walk）与起点之期望距离的一半。由史特灵公式（Stirling's formula），$a_n \approx 1/\sqrt{\pi n}$，所以这个超出量的增长与 $\sqrt{n/\pi}$ 相若：正面与反面次数的典型差距，按掷币次数的平方根增长。与此同时，每英镑的超出量 a_n 趋向 0，这正是大数定律（law of large numbers）在起作用。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（《概率导论》）讲述了随机漫步及这些极限。

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。