详细解答

# STEP III 2002，第 12 题

概率与统计 • 以条件化求期望值

> **STEP III 2002，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/02-s3-q12

## 关键思路

按游戏如何开始来取条件（conditioning）。每一局游戏的开头不外乎三种：H、TH 或 TT。在前两种情况下，并没有取得任何有用的进展（连续反面已被打断），所以游戏等于从头再来，而余下部分的期望长度仍然是 $E$。这种“更新”（renewal）给出一条关于 $E$ 的线性方程，解之即可。同样的想法——按第一个正面何时出现来分类——也能处理连续 $r$ 次反面的情况。

## 解答

全文中，$p$ 是掷出正面的概率，$q = 1 - p$ 是掷出反面的概率，各次投掷互相独立（independent）。我们假设 $0 < p < 1$，这样游戏以概率 1 结束，而且下文所有期望值（expectation）都是有限的。

### 解释关于 $E$ 的方程

考虑最初的一两次投掷。以下三种情况恰好有一种发生：

- **先掷出 H**（概率为 $p$）。已用去一次投掷，而且没有正在累积的连续反面。由于各次投掷互相独立，游戏余下的部分就是整个游戏的一个全新复本，期望长度为 $E$。期望总长度：$1 + E$。
- **先 T 后 H**（概率为 $qp$）。已用去两次投掷，而正面再次把连续反面清零，所以游戏重新开始。期望总长度：$2 + E$。
- **先 T 后 T**（概率为 $q^2$）。游戏恰好在 2 次投掷后结束。

这三种情况互斥（exclusive）而且穷尽（exhaustive）所有可能（它们的概率之和为 $p + qp + q^2 = p + q = 1$）。根据全期望定律（law of total expectation），

$$E = p(1 + E) + (1 - p)p(2 + E) + 2(1 - p)^2,$$

这正是题目给出的方程。

**求 $E$**

把含 $E$ 的项合并：

$$E\,(1 - p - qp) = p + 2qp + 2q^2.$$

左边，$1 - p - qp = q - qp = q(1 - p) = q^2$。右边，$p + 2q(p + q) = p + 2q = 1 + q$。由于 $q \neq 0$，

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

**期望奖金**

设 $W$ 为正面次数的期望值，亦即以英镑计的期望奖金。沿用同样的三种情况。掷出 H 之后，我们已赢得 £1，游戏重新开始；掷出 TH 之后，我们同样已赢得 £1，游戏重新开始；掷出 TT 之后，我们一无所获，游戏结束。所以

$$W = p(1 + W) + qp(1 + W) + q^2 \cdot 0.$$

? **想一想** 在 TH 的情况中，为甚么是 $1 + W$ 而不是 $W$?

在这种情况下，第二次投掷*确实*是正面，所以在游戏重新开始之前已经赚得 £1。这一题最容易出错的地方，就是漏掉了它。

合并各项，得 $W(1 - p - qp) = p(1 + q)$，所以 $Wq^2 = p(1 + q)$，而

$$W = \frac{p(1 + q)}{q^2} = \frac{p(2 - p)}{(1 - p)^2} \text{ 英镑}.$$

留意 $W = pE$：平均而言，所有投掷中有 $p$ 这个比例是正面，即使游戏停止的时刻取决于投掷的结果。

**连续 $r$ 次反面的游戏**

现在设 $E$ 为游戏在连续出现 $r$ 次反面时停止的情况下，投掷次数的期望值。按第一个正面来取条件。要么最初 $r$ 次投掷全是反面，要么第一个正面出现在第 $j + 1$ 次投掷，其中 $j$ 满足 $0 \leq j \leq r - 1$：

- $j$ **次反面后一次正面**（概率为 $q^j p$）。这用去了 $j+1$ 次投掷，连续反面已被打断，游戏重新开始。期望总长度：$j+1+E$。
- $r$ **次反面**（概率为 $q^r$）。游戏恰好在 $r$ 次投掷后结束。

所以

$$E = \sum_{j=0}^{r-1} q^j p\,(j+1+E) + rq^r.$$

右边 $E$ 的系数是 $p\sum_{j=0}^{r-1} q^j = p\cdot\dfrac{1-q^r}{1-q} = 1-q^r$，因为 $1-q=p$。把它移到左边，

$$q^r E = p\sum_{j=0}^{r-1}(j+1)q^j + rq^r. \qquad (*)$$

**想一想** 没有公式表在手，怎样求 $\sum(j+1)q^j$ 的和?

记 $S=\sum_{j=0}^{r-1}(j+1)q^j = 1+2q+3q^2+\cdots+rq^{r-1}$，然后减去 $qS = q+2q^2+\cdots+(r-1)q^{r-1}+rq^r$：

$$(1-q)S = 1+q+q^2+\cdots+q^{r-1}-rq^r = \frac{1-q^r}{p} - rq^r.$$

因此 $pS = \dfrac{1-q^r}{p} - rq^r$，代入 $(*)$ 后，$rq^r$ 两项互相抵消：

$$q^r E = \frac{1-q^r}{p} - rq^r + rq^r = \frac{1-q^r}{p}.$$

于是

$$E = \frac{1-q^r}{pq^r},$$

这就是所要证明的结果。当 $r=2$ 时，上式为 $\dfrac{(1-q)(1+q)}{pq^2} = \dfrac{1+q}{q^2}$，与第一个游戏的结果一致。

## 验算

- 若硬币是公平的，$p=q=\frac{1}{2}$：第一个游戏平均持续 $E=\frac{3/2}{1/4}=6$ 次投掷，赢得 $W=£3$，正好是 $E$ 的一半，合乎预期。
- $r=1$：游戏在第一次出现反面时停止，所以长度服从几何分布（geometric distribution），平均值为 $1/q$；公式给出 $\dfrac{1-q}{pq} = \dfrac{1}{q}$。
- 当 $q\to 0$（几乎每次投掷都是正面）时，$E$ 和 $W$ 都趋向无穷大，这是理所当然的。

- 我们以分数作精确计算，对 $r = 1, \ldots, 7$ 及若干个 $p$ 值解出“目前已连续出现 $k$ 次反面”这一串状态的方程，确认了这些公式。另外，以 $p = 0.4$ 模拟 $2 \times 10^5$ 局游戏，在 $r = 2$ 时得出平均 4.438 次投掷（理论值 4.444）及奖金 £1.772（理论值 £1.778），在 $r = 3$ 时得出 9.076 次投掷（理论值 9.074）。

## 常见错误

- 在奖金的方程中，把 $qp(1 + W)$ 写成 $qpW$，结果得出错误答案 $p/q^2$。
- 只按第一次投掷来取条件：掷出一次 T 之后，游戏*并不会*重新开始，因为连续反面中的一次已经到手。这正是题目所给方程要往前看两次投掷的原因。

## STEP 之外

“一旦出现正面，未来的情形就与开局时完全一样”这个论证，是马可夫链（Markov chain）的*更新法*或*首步分析法*（first-step method），一般情况下的期望到达时间（expected hitting time）正是这样计算的。$W = pE$ 这个简洁的事实是沃尔德恒等式（Wald's identity）的一个特例：把随机个数 $N$ 个独立同分布（independent, identically distributed）的项相加，其和的期望值等于 $\mathbb{E}[N]$ 乘以单独一项的平均值，前提是 $N$ 为停时（stopping time）。这两个课题都可以在 Grimmett 与 Welsh 的 *Probability: An Introduction*（概率论入门），或 Blitzstein 与 Hwang 的 *Introduction to Probability*（概率论导论）中读到。

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。