

明天还会出太阳吗?¹

钟开莱 / 文 胡佳 / 译

如果已经连续出了 n 天太阳, 那么下一天继续出太阳的几率是多少呢? 拉普拉斯 (Pierre-Simon Laplace) 给出了答案: $\frac{n+1}{n+2}$ 。对 $n=1$, 答案就是 $\frac{2}{3}$ 。既然他假定每天出太阳或不出太阳的几率相等, 那么答案为什么不是 $\frac{1}{2}$ 呢?

考虑箱子模型, 箱子中有两个球, 颜色或黑或白。从箱子里摸出第一个球并发现这个球是白色的, 那么另一个球也是白色的几率是多大呢?

箱子里可能包含了 i 个黑球, $i=0, 1, 2$ 。拉普拉斯称这些可能性为“条件”, 以后将用 C_i 来表示它们 (记号具有不可思议的重要性!), 用 E 来表示摸到的第一个球是白球的“事件”, 用 F 来表示第二个球也是白球的事件。那么我们就可以得到条件概率

$$P(E|C_i) = \frac{2-i}{2}, \quad i=0, 1, 2.$$

拉普拉斯认为“逆概率”应该正比于上述条件概率。这被称为贝叶斯法则 (Bayes's Rule)。令

$$S = \sum_i P(E|C_i). \quad (1)$$

则

$$P(C_i|E) = P(E|C_i) / S. \quad (2)$$

读者现在可以计算所需的结果

$$P(F|E) = \sum_i P(C_i|E)P(F|C_i \cap E). \quad (3)$$

并对拉普拉斯的答案感到满意了。

1935年, 我在中国上海立达学园上高三。刘炳震那时念高二。我们在一堂代数课上学习了拉普拉斯的日出定理。概率是代数课程中排列组合的后继内容, 例如范恩 (Henry Burchard Fine, 普林斯顿大学那栋古老而庄严的范恩大楼即以他的名字命名) 那本不可思议的著作。² 那时这本书被中国的很多高中作为教材来使用, 取代了早先一些更难的英国教材。

我们对老师关于这一结果的证明并不满意, 所以决定自己想出一个证明。经过一番冥思苦想, 我们得到了答案, 事实上我们在王星拱³ 的一本论科学方法的书里找到了一个一般形式。这里有一个中间结果, 是问题的核心。

¹ 原文标题是“Will the Sun Rise Again?”, 发表于新加坡数学会普及刊物 *Mathematical Medley*, Volume 29, No. 2, December 2002, pp. 67-73.

² H. B Fine: *A College Algebra*, Ginn & Co., 1901. 范恩的书有中译本, 即《范氏大代数》。

³ 王星拱 (1888-1949), 著名教育家、化学家、哲学家。著作多部论科学方法通俗著作。

一个箱子里包含了若干个黑色和白色的球。连续取出 n 个球（不再放回）并发现有 $n-r$ 个白球， $0 \leq r \leq n$ 。那么下一个取出来的球（当然，假设箱子不空）是白色的几率是多大？答案是 $\frac{n-r+1}{n+2}$ 。对 $r=0$ ，就是先前的情形。

令箱子里的球的总数为 $n+m$ ，这里 $m \geq 1$ 。用 C_i 来表示箱子里的黑球总数为 $r+i$ 的拉普拉斯条件，用 E 来表示事件取出来的前 n 个球里恰好包含了 r 个黑球，用 F 表示事件第 $n+1$ 个取出的球是白球。正如之前的那个最简单的情形一样，我们可以计算出各种情况下的条件概率。现在我们需要学会计算排列组合，可参阅前面提到的《范氏大代数》或其他初等教材，如⁴。

结果记录如下，以供读者验证。 S 如 (1) 定义， i 具有新的取值范围 $0 \leq i \leq m$ 。

$$P(E|C_i) = \binom{n}{r} \binom{m}{i} \binom{n+m}{r+1}^{-1};$$

$$P(E|F \cap C_i) = \binom{m-1}{n} \binom{m}{i}^{-1}.$$

经过一些显然的操作，我们可以得到

$$S = \binom{n+m}{n}^{-1} \sum_{i=0}^m \binom{r+i}{r} \binom{n+m-r-i}{n-r}. \quad (4)$$

这个求和的计算有赖于下面这个**主要等式**：对于非负整数 x, y, z 和 i ，有：

$$\sum_{i=0}^z \binom{x+i}{x} \binom{y+z-i}{y} = \binom{x+y+z+1}{x+y+1}. \quad (5)$$

将 (5) 代入 (4) 可得

$$\binom{n+m}{n}^{-1} \binom{n+m+1}{n+1} = \frac{m+n+1}{n+1}. \quad (6)$$

现在我们令

$$f(n, m) = \binom{n}{r}^{-1} S.$$

那么“显而易见”（正如拉普拉斯必然会说的）有

$$P(F|E) = \frac{f(n+1, m-1)}{f(n, m)}.$$

在“奇妙”的抵消之后，化简即得我们想要的结果： $\frac{n-r+1}{n+2}$ 。

堪称奇迹的是，引入的不定的数字 m 在最终的结果里消失了！我们可以先验证这一点吗？现在只需要证明 (5)。我们从这个简洁的恒等式开始

$$\binom{x+i}{x} = \binom{x-1+i}{x-1} + \binom{x-1+i}{x} \quad (7)$$

⁴I. Todhunter: A History of the Mathematical Theory of Probability, MacMillan Co. 1865.

这在范恩的书里第 404 页有提到，同时提到的还有一个整齐简洁的计数论证（不要用愚蠢的计算来破坏它！）。通过重复（归纳）可得

$$\binom{x+i}{x} = \sum_{j=0}^i \binom{x-1+j}{x-1}. \quad (8)$$

现在我们将（8）代入（5）从而将这个简单的求和转换成二重求和或累次求和，然后颠倒求和的次序，再次利用（8）将这个颠倒后的累次求和化简为一个简单的求和。从而（5）左边的和式等于

$$\begin{aligned} & \sum_{i=0}^z \left\{ \sum_{j=0}^i \binom{x-1+j}{x-1} \right\} \binom{y+z-i}{y} \\ &= \sum_{i=0}^z \sum_{j=0}^i \binom{x-1+j}{x-1} \binom{y+z-i}{y} \\ &= \sum_{j=0}^z \sum_{i=j}^z \binom{x-1+j}{x-1} \binom{y+z-i}{y} \\ &= \sum_{j=0}^z \binom{x-1+j}{x-1} \left\{ \sum_{i=j}^z \binom{y+z-i}{y} \right\} \\ &= \sum_{j=0}^z \binom{x-1+j}{x-1} \binom{y+1+z-j}{y+1}. \end{aligned}$$

最后所得的求和不过是将（5）中第一个求和中的 (x, y) 用 $(x-1, y+1)$ 代替。继续这个过程，原始的 (x, y) 就成了 $(0, y+x)$ ，相应的求和就是

$$\sum_{i=0}^z \binom{0+i}{0} \binom{y+x+z-i}{y+x} = \sum_{i=0}^z \binom{y+x+z-i}{y+x}.$$

利用（8），该式就与（5）的右边相等了。豁然开朗了！我们为自己的方法感到特别高兴，并将此命名为层蜕法。

我们将我们的结果告诉给毛路真⁵老师，他当时是在我的家乡杭州的浙江大学任教的一个讲师。他发现在托德亨特（Isaac Todhunter）的历史⁴（第 454 页）中，普雷沃斯特（Pierre Prévost）和吕利耶（Simon A. J. L'Huilier）已经利用箱子模型证明了拉普拉斯的日出定理。拉普拉斯没有这样做，而是采取了一个简单的方法——假设作为密度 $0 \leq p \leq 1$ 的先验概率的连续性，然后计算一个积分。这在⁶（第 123 页）中有相关阐述，文献⁷给出了与此有关的变体。

1936 年，刘和我都考进了北平（北京的旧称）清华大学。他是破格录取，因为他

⁵ 毛路真（1904-1961），数学家，又名信桂，浙江奉化人。1927 年毕业于武昌中山大学（武汉大学）数学系，后任浙江大学教授、数学力学系主任。与陈建功合编过《高中代数学》。

⁶ K. L. Chung: Elementary Probability Theory with Stochastic Processes, Third Edition, Springer Verlag, 1979. 钟开莱的这本书是斯普林格出版的本科生数学丛书（UTM）第 4 号丛书，有中译本，《初等概率论附随机过程》，魏宗舒等译，人民教育出版社，1979 年。

⁷ W. Feller: An Introduction to Probability Theory and Its Applications, Third Edition, Wiley & Sons, 1968. 有中译本，威廉·费勒：《概率论及其应用》，胡迪鹤译，人民邮电出版社，2014 年。