详细解答

STEP I 2013，第 12 题

概率与统计 • 不放回抽样

STEP I 2013，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/13-s1-q12

关键思路

“随机”选取 k 粒药片，意思是容器中剩下的药片里，每一组 k 粒药片被选中的机会均等，所以每天的概率（probability）都是两个计数之比：有利的选法数目除以 $\binom{\text{剩下的药片}}{k}$。至于最后一天的问题，窍门在于对称性（symmetry）：留到最后一天的药片，与第一天选出的药片同样随机。在 (iii) 中，各天之间环环相扣，十分顺畅，因为在“好”的一天之后，容器的状态恰好就像一个较短的假期。

解答

在整个解答中，把药片视为可区分的（例如编上号码），尽管它们看起来一模一样。这样做不会改变任何概率，而且令从 N 粒中选 k 粒的全部 $\binom{N}{k}$ 种选法机会均等。

第 (i) 部分

当 $k = 3$ 及 $n = 3$ 时，共有 9 粒药片，每种 3 粒。从中选 3 粒的方法数为 $\binom{9}{3} = 84$。要每种各得一粒，就要从 3 粒第 1 种药片中选一粒、从 3 粒第 2 种中选一粒、从 3 粒第 3 种中选一粒，所以共有 $3^3 = 27$ 种方法。因此

$$\mathrm{P}(\text{第 1 天每种各一粒}) = \frac{27}{84} = \frac{9}{28},$$

正如所求。

想一想 为甚么求最后一天的概率时，我不需要逐一分析第 1 天和第 2 天发生了甚么？

在整个假期中，九粒药片被随机分成三组，每组三粒：第 1 天的一组、第 2 天的一组和第

3 天的一组（即剩下的）。无论按 1、2、3 的次序还是按 3、2、1 的次序选出各组，所得的分布（distribution）都相同：把药片分成三个有标记的组，每一种分法的机会都均等。特别地，留给第 3 天的那组药片是一组均匀随机（uniformly random）的三粒药片，与第 1 天选出的那组完全一样。所以最后一天剩下每种各一粒药片的概率同样是

$$\frac{9}{28}.$$

第 (ii) 部分

现在共有 $3n$ 粒药片，每种 n 粒。共有 $\binom{3n}{3}$ 种机会均等的选法，其中 n^3 种包含每种各一粒。所以

$$\begin{aligned}\mathrm{P}(\text{第 1 天每种各一粒}) &= \frac{n^3}{\binom{3n}{3}} = \frac{6n^3}{3n(3n-1)(3n-2)} \\ &= \frac{2n^2}{(3n-1)(3n-2)}.\end{aligned}$$

当 $n=3$ 时，此值为 $\frac{18}{8\cdot 7} = \frac{9}{28}$，与 (i) 吻合。

第 (iii) 部分

现在 $k=2$，共有 $2n$ 粒药片，两种药片各 n 粒。对 $m = n, n-1, \ldots, 1$，设 G_m 为以下事件：在某天开始时容器中每种药片各有 m 粒，而我在当天选到每种各一粒（所以第 1 天时 $m=n$）。

若容器中每种各有 m 粒，则选两粒药片共有 $\binom{2m}{2} = m(2m-1)$ 种选法，其中 $m\cdot m$ 种是每种各一粒。所以

$$\mathrm{P}(\text{每种各一粒} \mid m \text{ 为每种剩余粒数}) = \frac{m^2}{m(2m-1)} = \frac{m}{2m-1}.$$

若到目前为止每一天都是“好”的，那么每天都恰好取走了每种各一粒，所以下一天开始时每种各有 $m-1$ 粒。把每天的条件概率（conditional probability）逐一相乘，

$$\mathrm{P}(\text{每天都是每种各一粒}) = \prod_{m=1}^{n} \frac{m}{2m-1} = \frac{n!}{1\times 3\times 5\times \cdots \times (2n-1)}.$$

为了整理奇数之积，把分子和分母同乘以偶数之积 $2\times 4\times\cdots\times(2n) = 2^n\, n!$：

$$1\times 3\times\cdots\times(2n-1) = \frac{(2n)!}{2^n\, n!}.$$

因此

$$\boxed{\mathrm{P}(\text{每天都是每种各一粒}) = \frac{2^n\,(n!)^2}{(2n)!}}$$

（验算一下：$n=1$ 时得 1，这是正确的：每种只有一粒时，我必定两粒都取。）

现在对 $n!$ 及 $(2n)!$ 使用斯特林（Stirling）近似（Stirling's approximation）：

$$(n!)^2 \approx 2n\pi\left(\frac{n}{\mathrm{e}}\right)^{2n},$$
$$(2n)! \approx \sqrt{4n\pi}\left(\frac{2n}{\mathrm{e}}\right)^{2n} = 2\sqrt{n\pi}\,2^{2n}\left(\frac{n}{\mathrm{e}}\right)^{2n}.$$

因子 $(n/\mathrm{e})^{2n}$ 互相抵消，剩下

$$\frac{2^n\,(n!)^2}{(2n)!} \approx \frac{2^n\cdot 2n\pi}{2\sqrt{n\pi}\,2^{2n}} = 2^{-n}\,\frac{n\pi}{\sqrt{n\pi}} = 2^{-n}\sqrt{n\pi},$$

这正是所求的近似式。

验算

- 用电脑穷举所有抽取药片的方式，证实了 (i) 中第 1 天和第 3 天的概率都是 $\frac{9}{28}$，(ii) 的公式在 $n\le 7$ 时成立，以及 (iii) 的公式在 $n\le 4$ 时成立。
- 当 $n=10$ 时，(iii) 的准确概率为 $\frac{2^{10}(10!)^2}{20!}\approx 0.005542$，而 $2^{-10}\sqrt{10\pi}\approx 0.005474$：偏低约 1.3%。随着 n 增大，两者之比趋向 1（$n=50$ 时约为 1.0025），这与斯特林近似的预期一致。
- (iii) 的另一种计数方法：把 $2n$ 粒药片分成 n 对并排成有序列，共有 $(2n)!/2^n$ 种机会均等的分法；好的分法是每天把第 1 种药片与第 2 种药片配对，共有 $n!\times n!$ 种。两者之比同样是 $2^n(n!)^2/(2n)!$。

常见错误

- 每次抽一粒药片，写出 $\frac{3}{9}\times\frac{3}{8}\times\frac{3}{7}$，但这是按此次序抽到第 1、2、3 种的机会。三种药片出现的次序共有 3! 种，所以应得 $6\times\frac{27}{504}=\frac{9}{28}$。（又或者：第一粒可以是任何一种，所以是 $1\times\frac{6}{8}\times\frac{3}{7}=\frac{9}{28}$。）
- 求最后一天的概率时，尝试对第 1 天发生的所有情况作条件分析；这方法可行，但十分冗长，而对称性能立即给出答案。
- 对 $(2n)!$ 用斯特林近似时，要把所有 n 都换成 $2n$，包括平方根之内：$\sqrt{2(2n)\pi}=2\sqrt{n\pi}$，而不是 $\sqrt{2n\pi}$。

STEP 之外

(iii) 的答案是 $2^n/\binom{2n}{n}$，而估计式 $\binom{2n}{n} \approx 4^n/\sqrt{n\pi}$ 在概率论中随处可见：它给出简单随机游走（simple random walk）在 $2n$ 步后回到起点的机会，并且是证明一维或二维的这类游走以概率 1 回到原点的关键。(i) 中的对称性论证，就是可交换性（exchangeability）的概念：随机抽取的次序不会改变它们的联合分布（joint distribution）。这两个概念在 Grimmett 与 Welsh 的 *Probability: An Introduction*（《概率导论》）以及 Blitzstein 与 Hwang 的 *Introduction to Probability*（《概率导论》）中都有深入发展。

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。