详细解答

STEP III 1991，第 15 题

概率与统计 • 样本中配对的计数

> **STEP III 1991，第 15 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/91-s3-q15

关键思路

每一组 k 张牌出现的机会均等，所以每个概率（probability）都等于符合条件的手牌数目除以 $\binom{2n}{k}$。要数出具有指定对子结构的手牌，可以分阶段构造。首先选定哪些种类以完整的一对出现，然后选定另外哪些种类只出现一张，最后决定每个单张种类的两张牌中，我们手上拿的是哪一张。这个计数一旦完成，题目的每一部分都只是同一条公式的特殊情况。

解答

把 $2n$ 张牌看成全部可以互相区分（同一种类的两张牌不妨编号为 1 和 2）。这样，不放回地抽出 k 张牌时，$\binom{2n}{k}$ 个大小为 k 的子集各有相同的概率。当某一种类的两张牌都被抽中时，我们说手牌中有该种类的一个对子（pair）。

四张牌，没有对子

没有对子的手牌包含四个不同的种类，每个种类各一张。选种类有 $\binom{n}{4}$ 种方法，每个种类选哪一张牌有 2 种方法，所以共有 $2^4\binom{n}{4}$ 手牌。由于

$$\binom{2n}{4} = \frac{2n(2n-1)(2n-2)(2n-3)}{24} = \frac{4n(n-1)(2n-1)(2n-3)}{24},$$

我们得到

$$\mathrm{P}(\text{没有对子}) = \frac{16 \cdot n(n-1)(n-2)(n-3)/24}{4n(n-1)(2n-1)(2n-3)/24} = \frac{4(n-2)(n-3)}{(2n-1)(2n-3)},$$

正如所求。（这里 $n \geq 4$ 保证了四个不同的种类确实存在。）

四张牌，恰好一个对子

先选出以对子出现的种类（n 种方法）。其余两张牌必须来自余下 $n-1$ 个种类中的两个不同种类，选法有 $\binom{n-1}{2}$ 种，而每个种类选哪一张牌又有 2 种选择。手牌的数目是

$$n \cdot \frac{(n-1)(n-2)}{2} \cdot 4 = 2n(n-1)(n-2),$$

所以

$$\mathrm{P}(\text{恰好一个对子}) = \frac{24 \cdot 2n(n-1)(n-2)}{4n(n-1)(2n-1)(2n-3)} = \frac{12(n-2)}{(2n-1)(2n-3)}.$$

k 张牌，恰好 r 个对子

假设手牌恰好有 r 个对子。它们占去 $2r$ 张牌。余下的 $k-2r$ 张牌全部属于不同的种类，而且这些种类都不是那 r 个成对的种类，否则对子就会更多。于是我们数：

- 成对的种类：$\binom{n}{r}$ 种方法；
- 只出现一张的 $k-2r$ 个种类，从其余 $n-r$ 个种类中选出：$\binom{n-r}{k-2r}$ 种方法；
- 每个单张种类的两张牌中，我们拿的是哪一张：2^{k-2r} 种方法。

每一手恰好有 r 个对子的牌，都恰好由一组这样的选择得出。因此

$$\boxed{\mathrm{P}(\text{恰好}r\text{ 个对子}) = \frac{2^{k-2r}\binom{n}{r}\binom{n-r}{k-2r}}{\binom{2n}{k}}}$$

想一想 这个概率会不会等于零？它与四张牌的答案是否吻合？

单张的种类需要满足 $k-2r \le n-r$，即 $r \ge k-n$。若 $k > n$，任意 k 张牌必定包含至少 $k-n$ 个对子，而 $\binom{n-r}{k-2r} = 0$ 自动处理了这个情况。取 $k=4$，$r=0$ 的情况给出 $2^4\binom{n}{4}/\binom{2n}{4}$，$r=1$ 则给出 $2^2\, n\binom{n-1}{2}/\binom{2n}{4}$，正好就是上面的两个计数。

k 为偶数，所有牌都成对

设 $k = 2m$。手牌由 m 个对子组成，当且仅当它是 m 个完整种类的并集（union）。这样的手牌有 $\binom{n}{m}$ 手（这就是一般公式取 $r = m$ 的情况）。所以

$$\mathrm{P} = \frac{\binom{n}{m}}{\binom{2n}{2m}} = \frac{n!}{m!\,(n-m)!} \cdot \frac{(2m)!\,(2n-2m)!}{(2n)!}.$$

要化简这个式子，可把阶乘（factorial）中的偶数因子和奇数因子分开归组：

$$(2j)! = \big(2 \cdot 4 \cdots 2j\big)\big(1 \cdot 3 \cdots (2j-1)\big) = 2^j\, j!\,\big(1 \cdot 3 \cdots (2j-1)\big).$$

记 $O_j = 1\cdot 3\cdots(2j-1)$，并设 $O_0 = 1$。那么 $(2j)!/j! = 2^j O_j$，所以

$$\mathrm{P} = \frac{2^m O_m \cdot 2^{n-m} O_{n-m}}{2^n O_n} = \frac{O_m\, O_{n-m}}{O_n}.$$

把 O_n 除以 O_{n-m}，剩下的是由 $2n-2m+1 = 2n-k+1$ 到 $2n-1$ 的奇数，所以

$$\mathrm{P}(\tfrac{1}{2}k \text{ 个对子}) = \frac{1\times 3\times 5\times\cdots\times(k-1)}{(2n-1)(2n-3)\cdots(2n-k+1)},$$

正如所求。

另外还有一个简洁的直接论证。把整副牌洗匀，以最上面的 k 张牌作为手牌。看看最顶的那张牌。它的拍档出现在其余 $2n-1$ 个位置中任何一个的机会均等，而其中 $k-1$ 个位置在手牌之内，所以拍档在手牌中的概率是 $\frac{k-1}{2n-1}$。如果它确实在手牌中，就把这一对拿走。剩下的是 $2n-2$ 张牌的一个随机排列，而手牌就是其中最上面的 $k-2$ 张。重复这个论证得到 $\frac{k-3}{2n-3}$，如此类推，直到 $\frac{1}{2n-k+1}$。把这些条件概率（conditional probability）相乘，便得到同一个乘积。

验算

- 四张牌时共有三种情况：0、1 和 2 个对子，而 2 个对子的概率是 $\binom{n}{2}/\binom{2n}{4} = \frac{3}{(2n-1)(2n-3)}$。果然

  $$4(n-2)(n-3) + 12(n-2) + 3 = 4n^2 - 8n + 3 = (2n-1)(2n-3),$$

  所以三个概率之和为 1。
- 取 $n = 5$ 和四张牌，三个概率分别是 $\frac{8}{21}$、$\frac{4}{7}$ 和 $\frac{1}{21}$。用电脑对 $4 \le n \le 8$ 以及由 3 到 $2n-1$ 的每个 k 穷举所有手牌，结果与一般公式及全部成对的公式均相符。
- 取 $k = 2$，全部成对的公式给出 $\frac{1}{2n-1}$，即第二张牌与第一张牌配成一对的机会。

常见错误

- 用 $\binom{2n-2r}{k-2r}$ 从余下全部 $2n-2r$ 张牌中选出单张的牌。这样会容许其中两张属于同一种类，从而产生额外的对子。
- 遗漏因子 2^{k-2r}。每个只出现一张的种类，拿到的可以是它两张牌中的任何一张。

STEP 之外

恰好有 r 个对子的手牌数目，是典型的“先选结构，再填内容”式计数，而乘积 $1\cdot 3\cdots(2m-1)$（即双阶乘（double factorial）$(2m-1)!!$）数的是把 $2m$ 件物件分成若干对的方法数

目。双阶乘还会以标准正态分布（standard normal distribution）的偶数阶矩（moment）的身分再次出现。同样的配对问题，即随机样本中有多少个对子得以保存，也见于经典的“失袜”问题和生日类型的问题。Grimmett 与 Welsh 的 *Probability: An Introduction*（概率论入门）为这类计数打下良好的基础。

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。