排列组合与概率问题：

by GondorFu

对于排列组合问题，存在以下几种基本的思考方法：

1、求解问题的补集。

2、划分（隔板）的思想。

3、序贯的思想。

4、动态规划的思想

接下来，本文将讨论一类**经典问题——M 个小球放入 N 个盒子**，其中分球相同与不同，盒相同与不同，盒能否为空这八种情况。并在最后对一种特殊地指定盒子中小球数的问题进行讨论。从分析中可以发现很多题目都是这类问题的变式。

一、M 个相同的小球放入 N 个不同的盒子中，盒不能为空

这是经典划分思想的应用。想象将 M 个小球一字排开，一共将形成 M+1 个空隙。而将球放入盒子就相当于从这些空隙中选择 N-1 个。由于盒子不能为空，所以首尾两个空隙不能选择，因此 $N(D)=C_{M-1}^{N-1}$。

这里还存在一种动态规划的思想。首先假定 $f_1(M,N)$ 为该问题的解，那么

$$\begin{aligned} f_1(M,N) &= f_1(M-1,N-1)+\cdots+f_1(N-1,N-1) \\ &= f_1(M-1,N)+f_1(M-1,N-1) \end{aligned}$$

其中 $f_1(M-1,N-1)$ 表示第一个盒子中的球数为 1，$f_1(N-1,N-1)$ 表示第一个盒子中的球数为 $M-N+1$，这也是第一个盒子所能放入最多的小球数量。而根据定义可得 $f_1(M,1)=f_1(i,i)=1,\forall M,i$。根据上述公式可得下表。

4	0	0	0	1	4
3	0	0	1	3	6
2	0	1	2	3	4
1	1	1	1	1	1
N/M	1	2	3	4	5

（python/1Solution.py 给出了仿真程序）

二、M 个相同的小球放入 N 个不同的盒子中，盒能为空

这里可能有人会借用序贯的思想：将 M 个小球一字排开，一共将形成 M+1 个空隙，而放入盒子的动作能够理解为从这 M+1 个空隙中选择 N-1 次，因此 $N(D)=(M+1)^{N-1}$。但这里其实重复计算了很多情况，比如第一次选择了第 4 个空隙，第二次选择了第 8 个空隙，和第一次选择第 8 个空隙，第二次选择第 4 个空隙是相同的情况却被记入为不同的情况。

正确的思想是由于所有球相同，且盒子不能为空，那么可以假象先在每个盒子中放入了一个球，共 N 球，再将 M 个球放入 N 个盒中。通过这两步操作保证了盒子不为空，并没有添加其他限制条件，因此根据之前的公式得到此时的方法数 $N(\tilde{D})=C_{M+N-1}^{N-1}$。同时由于第二步操作其实就是将 M 个球放入 N 个盒子，盒子可以为空的情况，而第一步在 N 个盒子中各放入一个球的方法数为 1，因此得到 M+N 个球放入 N 个盒子，盒子不为空的方法数应等于 M 个球放入 N 个盒子，盒子可以为空的方法数，即 $N(D)=1\times N(\tilde{D})=C_{M+N-1}^{N-1}$。

同样也存在动态规划方法。相比于盒子不能为空的情况，只需要在公式中加上球数少于盒子数导致盒子必然为空的情况即可。

$$\begin{aligned} f_2(M,N) &= f_2(M,N-1)+f_2(M-1,N-1)+\cdots+f_2(0,N-1) \\ &= f_2(M-1,N)+f_2(M,N-1) \end{aligned}$$

其中 $f_2(M,N-1)$ 就表示第一个盒子中的球数为 0，$f_2(0,N-1)$ 就表示所有后面的 $N-1$ 个盒子都为空。根据定义可得 $f_2(M,1)=1,\forall M$，$f_2(0,N)=1,\forall N$。根据上述公式可得下表。

4	1	4	10	20	35	56
3	1	3	6	10	15	21
2	1	2	3	4	5	6
1	1	1	1	1	1	1
N/M	0	1	2	3	4	5

（python/2Solution.py 给出了仿真程序）

对于小球和盒子都相同的情况，可以按小球的数量对盒子进行排序。假设 N 个盒子中的小球数量分别为 $a,b,\cdots,z$，则

$$\underbrace{a+b+\cdots+z}_{N}=M,a\le b\le\cdots\le z$$

这是典型的整数分拆问题。

三、M 个相同的小球放入 N 个相同的盒子中，盒子不为空

由于球和盒都相同，且盒子不为空，可以首先选出 N 个球平均分配到 N 个盒子中，这样就满足了盒子不为空的条件，且方法数为 1。此时就可以将剩下的 M-N 个球无限制的放入 N 个盒中。这里可以对剩下的 M-N 个球进行如下分类：只有一个盒子被放入了球，有两个盒子被放入了球，一直到所有 N 个盒子都被放入了球。如果定义 $f_3(M,N)$ 为所求问题的答案，根据以上可得

$$f_3(M,N)=\underbrace{f_3(M-N,1)+f_3(M-N,2)+\cdots+f_3(M-N,N)}_{N}$$

其中 $f_3(M,1)=f_3(i,i)=1,\forall M,i$。需要注意如果 $M-N<N$，那么最后几项一定不满足要求可以直接置零，即 $f_3(M,N)=0,M<N$。

例 1，将 7 个相同的球放入 4 个相同的盒子中，盒子不为空，有几种分法？（1114,1123,1222）

$$f_3(7,4)=f_3(7-4,1)+f_3(7-4,2)+f_3(7-4,3)+f_3(7-4,4)$$

其中

$$\begin{aligned}
&f_3(7-4,1)=f_3(3,1)=1\\
&f_3(7-4,2)=f_3(3,2)=f_3(3-2,1)+f_3(3-2,2)=1+0=1\\
&f_3(7-4,3)=f_3(3,3)=1\\
&f_3(7-4,4)=f_3(3,4)=0
\end{aligned}$$

因此

$$f_3(7,4)=1+1+1+0=3$$

（python/3Solution.py 给出了仿真程序）

四、M 个相同的小球放入 N 个相同的盒子中，盒子能为空

根据二中求解盒子为空的思想，同理可得 M+N 个小球放入 N 个盒子，盒子不为空的的方法数应等于 M 个小球放入 N 个盒子,盒子可以为空的方法数相同，即 $f_4\left(M,N\right)=f_3\left(M+N,N\right)$

（python/4Solution.py 给出了仿真程序）

五、M 个不同的小球放入 N 个相同的盒子中，盒子不为空

对于该问题，可以针对某一个小球，比如 A，将问题拆分为它被单独放置在一个盒中和它放置的盒中还有其他球的两种情况。如果定义 $f_5(M,N)$ 为所求结果，那么对于第一种情况（小球 A 被单独放置），首先由于盒子相同，放入的方法数为 1，接下来需要将剩下的 M-1 个小球不空放置到剩下 N-1 个盒中，即 $f_5(M-1,N-1)$。对于第二种情况，首先考虑将剩下的 M-1 个小球不空地放入 N 个盒中，得 $f_5(M-1,N)$。接着考虑将 A 放入，由于盒中都有小球，造成与放入 N 个不同的盒子中的效果等价，因此此步的方法数为 N。综合以上得

$$f_5(M,N)=\underbrace{f_5(M-1,N-1)}_{1}+\underbrace{N\times f_5(M-1,N)}_{2}$$

这是第二类斯特林数，其中 $f_5(M,1)=f_5(i,i)=1,\forall M,i$， $f_5(M,N)=0,M<N$

（python/5Solution.py 给出了仿真程序）

六、M 个不同的小球放入 N 个相同的盒子中，盒子能为空

对于本问题可以理解为第五个问题的推广，将该问题拆分为有一个盒子为空，有两个盒子为空，一直到有 N-1 个盒子为空的情况，因此本题的结果

$$\begin{aligned}f_6(M,N)&=\underbrace{f_5(M,N)+f_5(M,N-1)+\cdots f_5(M,1)}_{N}\\&=f_6(M,N-1)+f_5(M,N)\end{aligned}$$

其中 $f_6(M,1)=1,\forall M$。

（python/6Solution.py 给出了仿真程序）

七、M 个不同的小球放入 N 个不同的盒子中，盒子能为空

由于小球不同，盒子也不同，因此不同小球之间的放置互不影响，可以对每个小球的可能数量累乘得到最后结果。每个小球总有 N 中不同方法（放入 N 个不同的盒子中的任意一个），重复操作 M 次，因此 $N(D)=\underbrace{N\times N\times\cdots\times N\times N}_{m}=N^M$ 种方法。

八、M 个不同的小球放入 N 个不同的盒子中，盒子不为空

首先介绍几种错误的想法。

i）如果考虑先选出 N 个球平均放入 N 个盒子中，从而满足盒子不为空，然后在对剩下的 M-N 个球只有序贯放入，得 $N(D)=A_M^N N^{M-N}$。那么存在如下重复，以 M=3（A，B，C），N=2 为例，一种情况首先选取了 A，C 分别放入两个盒子，而 B 放入了第一个盒子，而另一种情况选取了 B，C 分别放入两个盒子，而 A 进入了第一个盒子，这样 AB|C 和 BA|C 被认为了是两种不同的情况。其同样需要除以盒内球数的全排列，而盒内球数很难确定。

ii）那能不能不选球直接对 N 个球进行全排列放入 N 个盒子中，然后序贯放入剩下的小球，得 $N(D)=A_N^N N^{M-N}$。这样又少算了开始选择的 N 个小球可能两两或全部出现在同一盒子中的情况。

正确的思路在第五个问题的基础上考虑唯一的区别就是盒子由相同变为不同，且由于球不同盒不为空，因此相对于第五个问题的每一种方法，都多出了盒子数量的全排列种方法。因此

$$f_8(M,N)=A_N^N f_5(M,N)$$

以上就是对所有八种情况的讨论。现总结如下：

1、小球相同，盒子相同：整数分拆。

2、小球相同，盒子不同：用划分。

3、小球相同，盒子能否为空，用 N 小球先填。

4、小球不同盒不同，无限制，序贯处理

5、小球不同，斯特林数。

==

华丽的分割线

==

下面将针对一类特殊问题——在放盒子时指定了球被划分的数量，即 $\underbrace{a,b,\cdots,z}_{N}$，并满足 $a+b+\cdots+z=M$。

1、可以发现对于球相同盒相同的情况，方法只有一种。

2、而对于球相同盒不同的情况，由于小球相同且指定了数量，小球的所有可能划分只有一种，因此只需要对盒子进行全排列从而将对应数量的小球放入即可，得 $N\left(\tilde{D}\right)=A_N^N$。但是不难发现由于如果存在指定的数 $c==d$ 的情况，那么交换这两个盒子将可得相同的方法。因此对于一般情况，结果为

$$N\left(D\right)=\frac{A_N^N}{A_m^m(i)\times\cdots\times A_n^n(j)}$$

其中 $A_m^m(i)$ 表示存在 m 个盒子被指定了必须具有 i 个小球，$A_n^n(j)$ 表示存在 n 个盒子被指定了必须具有 j 个小球。

同时还有另一种思考方法，先从 N 个盒子中选择一个装 a，再选择一个来装 b，直到最后一个盒子装 z，得到 $N\left(\tilde{D}\right)=C_N^1C_{N-1}^1\cdots C_1^1$。此时如果存在 $c==d$，将同样导致重复。原因是如果一种方法选择了 x 来装 c，y 来装 d；另一种方法选择了 y 来装 c，x 来装 d，而它们应该是相同的方法。可以发现这使得每种方法重复了 $A_m^m(i)\times\cdots\times A_n^n(j)$ 次，定义如上。最后得到

$$N\left(D\right)=\frac{C_N^1C_{N-1}^1\cdots C_1^1}{A_m^m(i)\times\cdots\times A_n^n(j)}$$

由于 $A_N^N=C_N^1C_{N-1}^1\cdots C_1^1$，因此上述两种方法是等价的。

从后面的讨论中可以发现对于此类问题都存在以上两种思路，这里分别称为全排列（交换造成重复）和序贯选择（交叉选择造成重复）。

例 2，将 5 个相同的小球放到四个不同的盒子，分别为 s，x，y，z 中，且指定盒子中小球的数量为 1，1，2，2 一共有多少种方法？

$$N\left(D\right)=\frac{A_4^4}{A_2^2A_2^2}=6$$

	s	x	y	z
第 1 种	1	1	2	2
第 2 种	1	2	1	2
第 3 种	1	2	2	1
第 4 种	2	1	1	2
第 5 种	2	1	2	1
第 6 种	2	2	1	1

3、对于球不同盒相同的情况，同样从全排列和序贯选择来求解。

对于全排列方法，由于此时球不同，因此首先对球进行全排列，得 $N\left(\tilde{D}\right)=A_M^M$。再依据当前的顺序按要求将球放入盒中，由于盒都相同，因此方法数为 1。但不难发现重复，对于一个盒中的球交换顺序并不会得到不同的结果，而重复数为盒子大小的全排列，同时如果存在指定数目相等的情况，如 $c==d$，由于盒子相同，交换它们的球也将造成重复，因此最后的结果为

$$N\left(D\right)=\frac{A_M^M}{\left(A_a^a\times\cdots\times A_z^z\right)\left(A_m^m(i)\times\cdots\times A_n^n(j)\right)}$$

同理，对于序贯选择方法，首先选择 a 个球，之后选择 b 个球，直到最后剩下 z 个球，得 $N\left(\tilde{D}\right)=C_M^aC_{M-a}^b\cdots C_z^z$。此时需要将球放入盒中，由于盒子相同，因此方法数为 1。至此，如果存在 $c==d$ 的情况，也存在重复。由于字面比较难解释，现举例说明如下。

例 3，将 4 个不同小球分别为 A，B，C，D 放入二个相同盒子中，且指定小球数目为 2，2，一共有几种方法？

如果使用 $N\left(\tilde{D}\right)=C_4^2C_2^2=6$ 得到的所有情况如下

	x	x
1	AB	CD
2	AC	BD
3	AD	BC
4	BC	AD
5	BD	AC
6	CD	AB

由于两个盒子相同，可以发现 1 与 6，2 与 5，3 与 4 重复，同时重复发生在多次选择中，不同方法选择的球只在选择次序上出现了不同。可以总结出每种情况重复的次数为指定具有相同小球数的盒子数的全排列。因此**例 3** 最后的结果为

$$N\left(D\right)=\frac{C_4^2C_2^2}{A_2^2}=3$$

对于一般情况，最后的结果为

$$N\left(D\right)=\frac{C_M^aC_{M-a}^b\cdots C_z^z}{A_m^m(i)\times\cdots\times A_n^n(j)}$$

通过推导

$$\begin{aligned}&C_M^aC_{M-a-b}^b\times\cdots\times C_z^z\\&=\frac{M!}{a!\left(M-T\right)!}\times\frac{\left(M-a\right)!}{b!\left(M-a-b\right)!}\times\cdots\times\frac{z!}{z!\left(0\right)!}\\&=\frac{M!}{a!\times b!\times\cdots\times z!}\\&=\frac{A_M^M}{A_a^a\times\cdots\times A_z^z}\end{aligned}$$

可以得到两个公式能够得到相同的结果。最后给出一个实际的例子。

例 4，将 54 张扑克牌，平均分成三堆，其中大小王在同一堆的概率是多少？

题目可以理解为 54 个不同的小球平均放到三个相同的盒子里。但由于其中两张大小王不同与其他的牌，所以其实是指定分配为（16,18,18）的情况。因此答案为

$$P\left(D\right)=\frac{N\left(D\right)}{N\left(S\right)}=\frac{C_{52}^{16}C_{36}^{18}C_{18}^{18}\big/A_2^2}{C_{54}^{18}C_{36}^{18}C_{18}^{18}\big/A_3^3}=\frac{17}{53}。$$

（python/7Solution.py 给出了仿真程序）

4、最后针对球不同盒也不同的情况，进行讨论。

对于全排列方法，由于球不同对球进行全排列，盒也不同再对盒进行全排列，得 $N\left(\tilde{D}\right)=A_M^MA_N^N$。此时同样存在盒内球之间的全排列被作为了不同的方法，同时盒子具有相同数量小球的情况由于在球和盒下都进行了全排，因此也存在重复。

最终得到

$$N\left(D\right)=\frac{A_M^M A_N^N}{\left(A_a^a\times\cdots\times A_z^z\right)\left(A_m^m(i)\times\cdots\times A_n^n(j)\right)}$$

同理，对于序贯选择方法，首先对球进行选择，得 $N\left(\tilde{D}\right)=C_M^a C_{M-a}^b\cdots C_z^z$。同时，由于盒子不同，对盒子进行全排列。由于同样存在交叉选择的重复方法，因此最后的结果为

$$N\left(D\right)=\frac{C_M^a C_{M-a}^b\cdots C_z^z A_N^N}{A_m^m(i)\times\cdots\times A_n^n(j)}$$

根据之前的推导，两式也是等价的。

例 5，将四个不同的小球，分别为 A，B，C，D，放入三个不同的盒子，分别为 x，y，z 中，其中指定盒中小球数目为 1，1，2，一共有多少种分法？

$$N\left(D\right)=\frac{C_4^1 C_3^1 C_2^2 A_3^3}{A_2^2}=36$$

通过对上述问题的分析，可以发现只有对于不同类型的情况才能使用 C_M^i 的操作来进行选择。

对于指定数量的问题，现总结如下：

1、对于所有情况，都存在全排列和序贯选择两种方法。

2、全排列的重复发生在同类交换中，序贯选择的重复发生在交叉选择中。

3、综合比较下来，序贯选择方法更简单：对不同类型的东西进行选择，之后去除交叉重复。

==

华丽的分割线

==

最后给出一道经典例题。

例 6，求解任意 30 人中至少有两个人生日（简化为 365 天）相同的概率？

如果直接求解问题十分复杂，但如果求解没有任何人的生日相同就变得比较简单。如此，定义至少两个人生日相同的事件为 D（所要求解的问题），那么没有任何人生日相同的事件（D 的补集）为 $\bar{D}$，而全集为 S。

第一种方法：借助序贯思想，全集 S 的数量 $N\left(S\right)=365^{30}$；而补集就是从 365 天中选出 30 天分配给 30 个人，因此 $N\left(\bar{D}\right)=A_{365}^{30}$。最终，可得

$$P\left(D\right)=1-P\left(\bar{D}\right)=1-\frac{A_{365}^{30}}{365^{30}}=1-\frac{365\times364\times\cdots\times336}{\underbrace{365\times\cdots\times365}_{30}}=70.63\%$$

第二种方法,对每个人的情况进行分析,第一个人可以从 365 天中随意选取，第二个人在除去第一个人选取剩下的 364 天中随机选取，依此类推可得

$$P\left(\bar{D}\right)=\frac{365}{365}\times\frac{364}{365}\times\cdots\times\frac{336}{365}=29.37\%$$

（python/8birthday.py 中给出了仿真程序）