佔有問題

劉涵初

緒　言

在組合學（ combinatorics ）與機率論的發展歷史中，將 n 個球放在 k 個箱中是個很重要的模式，這個問題稱爲佔有問題（ occupancy problem ），佔有問題的應用很廣，我們將探討這些問題。首先看看以下的例子：

(1)車禍：n 件車禍發生於週一到週日的可能情形，相當於置 n 個球於 $k=7$ 個箱中的分佈狀況。

(2)抽樣：n 個人按職業分類，則人相當於球，類別相當於箱子。

(3)生物放射實驗：眼睛的網膜細胞曝光時，光線的粒子相當於球，細胞則相當於箱子。同理，在研究放射對遺傳的影響時，α 粒子相當於球，而染色體相當於箱子。

(4)宇宙射線實驗：射入蓋氏計數器的粒子相當於球，而計數器相當於箱子。

(5)骰子：擲一個骰子 n 次所出現點數的分配，相當於置 n 個球於 $k=6$ 個箱子的分佈。

(6)隨機亂數：n 個數字的各種排列，相當於置 n 個球於 $k=10$ 個箱中，10 個箱子標示爲 0，1，2，………，9 。

(7) n 個人的性別分配：相當於置 n 個人於 $k=2$ 個箱中。

(8)彩券收集：所收集的彩券相當於球，彩券種類相當於箱子。

(9)基因分配：每一個體（人、植物或動物）都從上一代遺傳下來某些基因，設某一基因可能有 A_1，A_2，……，A_k k 種型態出現，則其後代可依基因形態而分類。故後代相當於球而遺傳型態 A_1，A_2，……，A_k 相當於 k 個箱子。

(10)化學：長鏈聚合體與氧分子發生反應，每一鏈可與 0，1，2，……個氧分子發生作用，則反應的氧分子相當於球，而聚合鏈相當於箱子。

(11)感光乳劑理論：感光板塗上對光量子敏感的藥劑粒子，若一藥劑粒子被 r 個光量子擊中，則發生反應，我們想知道有幾個粒子可能被 r 個光量子擊中，這裡光量子相當於球，粒子相當於箱子。

(12)錯印：一本書 k 頁中有 n 個印錯的分配情形，相當於 n 個球置於 k 個箱中的分配情形。

分配型態

將 n 個球放在 k 個箱中，根據球是否可區分，箱子是否可區分，及是否允許有空箱，這個問題的模式共有 8 種如表 1，稍後我們再一一探討各種不同情況下，分配方法的個數。

	球是否可區分	箱子是否可區分	是否允許有空箱	n個球放在k個箱中的方法數
1a	是	是	是	k^n
1b	是	是	否	$k!\ S_{n,k}$
2a	否	是	是	$C(n+k-1, k-1)$
2b	否	是	否	$C(n-1, k-1)$
3a	是	否	是	$S_{n,1}+\cdots\cdots+S_{n,k}$
3b	是	否	否	$S_{n,k}$
4a	否	否	是	n分割爲k個或更少正整數和的方法數
4b	否	否	否	n分割爲k個正整數和的方法數

表 1

考慮$n=3$，$k=2$的例子

(1a) 球可區分標識爲a、b、c，箱子可區分標識爲 1、2，允許有空箱，共有 8 種方法如表 2：

分配方法

	1	2	3	4	5	6	7	8
箱 1	abc	ab	ac	bc	a	b	c	
子 2		c	b	a	bc	ac	ab	abc

表 2

(1b) 同(1a)但不允許有空箱，共有 6 種方法，即表 2 的方法 2-7 。

(2a) 球不可區分表爲a、a、a，箱子可區分標識爲 1、2，允許有空箱共有 4 種方法如表 3：

分配方法

	1	2	3	4
箱 1	aaa	aa	a	
子 2		a	aa	aaa

表 3

(2b) 同(2a)但不允許有空箱，共有 2 種方法，即表 3 的方法 2-3 。

(3a) 球可區分表爲a、b、c，但箱子不可區分，允許有空箱，共有 4 種方法，即表 2 的方法 1-4（因箱子不可區分，故表 2 中，方法 1 與 8 同，2 與 7 同，3 與 6 同，4 與 5 同）。

(3b) 同(3a)但不允許有空箱，共有 3 種方法，即表 2 的方法 2-4 。

(4a) 球不可區分表爲a、a、a，箱子亦不可區分，允許有空箱，共有 2 種方法，即表 3 的方法 1-2（因箱子不可區分，故表 3 中，方法 1 與 4 同，2 與 3 同）。

(4b) 同(4a)但不允許有空箱，共有 1 種方法，即表 3 的方法 2 。

現在我們來求表 1 中各種不同情況下，分配方法的個數：

(1a) n個可區分的球，放在k個可區分的箱中，允許有空箱：因每一個球放入箱中有k種選擇，根據乘法原理，共有 k^n 種方法。

(2a) n個不可區分的球，放在k個可區分的箱中，允許有空箱：

n個不可區分的球表爲$aa\cdots\cdots a$，我們在這n個a中插入$k-1$條橫線，例如$n=5$，$k=3$，則$a|aaa|a$爲一種分配方法，即 1 號箱 1 個球，2 號箱 3 個球，3 號箱 1 個球。又如$|aaaa|a$亦爲一種分配方法，即 1 號箱 0 個球，2 號箱 4 個球，3 號箱 1 個球。故n個

a 與 $k-1$ 條橫線排列的方法共有

$$\frac{(n+k-1)!}{n!(k-1)!}=C(n+k-1,k-1)$$

(2b) 同(2a)但不允許有空箱：故每箱先放一個球，還剩 $n-k$ 個球，如同(2a)的組合推理，$n-k$ 個 a 與 $k-1$ 條橫線排列的方法共有

$$\frac{(n-k+k-1)!}{(n-k)!(k-1)!}$$
$$=C(n-1,k-1)。$$

(4a) n 個不可區分的球放在 k 個不可區分的箱中，允許有空箱：定義一個正整數 n 的分割爲一群正整數其和爲 n，例如 5 的分割共有 7 種如下：

$$\begin{aligned}&1+1+1+1+1\\=\,&1+1+1+2\\=\,&1+2+2\\=\,&1+1+3\\=\,&2+3=1+4=5\end{aligned}$$

故將 n 個不可區分的球放入 k 個不可區分的箱中，且允許有空箱的分配方法，正如將 n 分割爲 k 個或更少正整數和的方法，例如，$n=5$，$k=3$，則共有 $1+2+2=1+1+3=2+3=1+4=5$，這五種方法，如其中 $1+4$ 即表一個箱子含 1 個球，1 個箱子含 4 個球，另一個箱子爲空箱的分配方法。

(4b) 如同(4a)但不允許有空箱：同理，分配方法正如將 n 分割成 k 個正整數和的方法，例如 $n=5$，$k=3$，則共有 $1+2+2=1+1+3$，這 2 種方法。

(3b) n 個可區分的球放在 k 個不可區分的箱中，不允許有空箱：

定義這種情況下分配方法的個數爲 $S_{n,k}$，這個數稱爲第二型史特林數 (Stirling number of the second kind)，我們將利用包含排除原理 (Principle of inclusion and exclusion) 證明下式：

$$S_{n,k}=\frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$$

例如，$n=4$，$k=2$，則

$$S_{4,2}=\frac{1}{2!}\sum_{i=0}^{2}(-1)^i\binom{2}{i}(2-i)^4$$
$$=\frac{1}{2}(16-2+0)=7$$

若 4 個球表爲 a、b、c、d，分成兩堆 (放入兩個不可區分的箱中)，則有以下 7 種方法：

$$\begin{gathered}\{a,b\},\{c,d\}\\\{a,c\},\{b,d\}\\\{a,d\},\{b,c\}\\\{a\},\{b,c,d\}\\\{b\},\{a,c,d\}\\\{c\},\{a,b,d\}\\\{d\},\{a,b,c\}\end{gathered}$$

現在介紹包含排除原理，由圖(1a)，我們可輕易看出

$$|A_1\cup A_2|=|A_1|+|A_2|-|A_1\cap A_2|$$

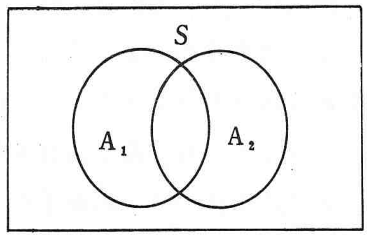
其中 A 爲集合，$|A|$ 表示集合元素之個數。

$$|A_1'\cap A_2'|=|S|-(|A_1|+|A_2|)+|A_1\cap A_2|$$

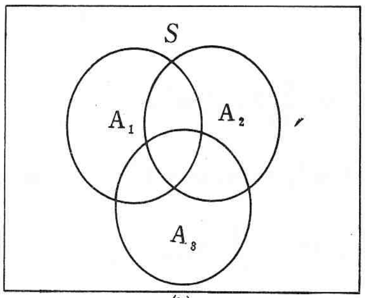
由圖(1b)亦不難得知

$$|A_1\cup A_2\cup A_3|=|A_1|+|A_2|+|A_3|-|A_1\cap A_2|-|A_1\cap A_3|-|A_2\cap A_3|+|A_1\cap A_2\cap A_3|$$

$$|A_1'\cap A_2'\cap A_3'|=|S|-(|A_1|+|A_2|+|A_3|)+(|A_1\cap A_2|+|A_1\cap A_3|+|A_2\cap A_3|)-|A_1\cap A_2\cap A_3|$$



(a)



(b)

圖1

將上式推廣，以數學歸納法可證得以下定理:

定理1：包含排除原理

設A_1，……，A_n 爲集合S的子集，則

$$|A_1 \cup A_2 \cup \cdots\cdots \cup A_n|$$
$$= \sum_{1\le i\le n}|A_i| - \sum_{1\le i<j\le n}|A_i \cap A_j|$$
$$+ \sum_{1\le i<j<k\le n}|A_i \cap A_j \cap A_k| - \cdots$$
$$+(-1)^{n-1}|A_1 \cap A_2 \cap \cdots \cap A_n|$$
$$|A_1' \cap A_2' \cap \cdots\cdots \cap A_n'|$$
$$= |S| - \sum_{1\le i\le n}|A_i| + \sum_{1\le i<j\le n}|A_i$$
$$\cap A_j| - \sum_{1\le i<j<k\le n}|A_i \cap A_j \cap A_k|$$
$$+\cdots+(-1)^n|A_1 \cap A_2 \cap \cdots \cap A_n|$$

定理2：$S_{n,k} = \dfrac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$

證明：考慮將n個可區分的球放在k個可區分的箱中，且不允許有空箱。令A_i表集合含i號箱爲空的方法，故問題爲求$|A_1' \cap A_2' \cap \cdots\cdots A_k'|$，因$|S| = k^n$，若$i$號箱爲空，則每一個球放入箱中只有$k-1$種選擇，故$|A_i| = (k-1)^n$，同理，$|A_i \cap A_j| = (k-2)^n$，$|A_i \cap A_j \cap A_k| = (k-3)^n$，……，根據定理1可得

$$|A_1' \cap A_2' \cap \cdots\cdots \cap A_k'|$$
$$= k^n - \binom{k}{1}(k-1)^n + \binom{k}{2}(k-2)^n$$
$$+\cdots\cdots+(-1)^k\binom{k}{k}(k-k)^n$$
$$= \sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n$$

因$S_{n,k}$表n個可區分的球放入k個不可區分的箱中，且不允許有空箱的方法數，而前面的結果爲k個可區分的箱子，故除以$k!$即得$S_{n,k}$的式子。

(1b) n個可區分的球，放在k個可區分的箱中，且不允許有空箱：

根據$S_{n,k}$的定義，顯然共有$k!\,S_{n,k}$種方法。

(3a) n個可區分的球，放在k個不可區分的箱中，允許有空箱：

情況如同(3b)但允許有空箱，在(3b)中不允許有空箱，共有$S_{n,k}$種方法，現在允許有空箱，可有以下不同情況：沒有空箱，恰有1個空箱，恰有2個空箱，……，恰有$k-1$個空箱。因此分配方法共有

$$S_{n,k} + S_{n,k-1} + \cdots\cdots + S_{n,1}$$

第二型史特林數

我們亦可利用生成函數(generating function)的技巧求得定理2中第二型史特林數的公式，參閱〔4〕。我們現在將利用組合推理求得第二型史特林數$S_{n,k}$的遞廻關係(recurrence relation)。

定理3：$S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$

$S_{n,1} = 1$，$S_{n,n} = 1$

證明：根據$S_{n,k}$的定義，$S_{n,k}$表n個可區分的球，放在k個不可區分的箱中，且不允許有空箱的方法數，顯然$S_{n,1} = 1$，$S_{n,n} = 1$，這是遞廻關係的起始條件。

我們先考慮$n-1$個球放在k個箱中，有以下兩種情況：(i)有一個空箱：故有$S_{n-1,k-1}$種方法，剩下的一個球必得放入此空箱。(ii)沒有空箱：有

$S_{n-1,k}$ 種方法，而剩下的一個球可放在任意一個箱中，因此共有$kS_{n-1,k}$種方法。得證。

例如，$n=4$，$k=2$，則

$$S_{4,2}=S_{3,1}+2S_{3,2}$$
$$S_{3,2}=S_{2,1}+2S_{2,2}$$
$$=1+2=3$$

故$S_{4,2}=1+2\cdot 3=7$

$S_{n,k}=S_{n-1,k-1}+kS_{n-1,k}$與$C(n,k)=C(n-1,k-1)+C(n-1,k-1)$這個二項關係很相似，故正如巴斯卡三角形(Pascal's triangle)，可利用表4的史特林三角形(Stirling's triangle)求連續的$S_{n,k}$ 值。

$S_{1,1}$

$S_{2,1}$ $S_{2,2}$

$S_{3,1}$ $S_{3,2}$ $S_{3,3}$

$S_{4,1}$ $S_{4,2}$ $S_{4,3}$ $S_{4,4}$

$S_{5,1}$ $S_{5,2}$ $S_{5,3}$ $S_{5,4}$ $S_{5,5}$

1

1 1

1 3 1

1 7 6 1

1 15 25 10 1

表4

將n個可區分的球放在m個可區分的箱中，共有 m^n 種方法，若限制不能有空箱，則有$m!\,S_{n,m}$種方法；若限制只能有一個空箱，則有$\binom{m}{m-1}(m-1)!\,S_{n,m-1}$種方法，若限制只能有兩個空箱，則有$\binom{m}{m-2}(m-2)!\,S_{n,m-2}$ 種方法，………，故

$$m^n=\sum_{k=0}^{m}\binom{m}{k}k!\,S_{n,k}\ (\text{若 } n<m\text{，則 } S_{n,k}=0\text{，}k>n)$$

$$=\sum_{k=0}^{n}\binom{m}{k}k!\,S_{n,k}\ (\text{若 } n>m\text{，則 } \binom{m}{k}=0\text{，}k>m)$$

$$=\sum_{k=0}^{n}S_{n,k}[m]_k$$

定義$[m]_k=m(m-1)\cdots\cdots(m-k+1)$

故得，$x^n=\sum_{k=0}^{n}S_{n,k}[x]_k$

因$S_{n,0}=0$，故

$$x^n=\sum_{k=1}^{n}S_{n,k}[x]_k=S_{n,1}x+S_{n,2}x(x-1)+\cdots\cdots+S_{n,n}x(x-1)\cdots(x-n+1)$$

上式即第二型史特林數的原始定義，例如，

$$x^4=x+7x(x-1)+6x(x-1)\cdot(x-2)+x(x-1)(x-2)\cdot(x-3)$$

故$S_{4,1}=1$，$S_{4,2}=7$，$S_{4,3}=6$，$S_{4,4}=1$

而第一型史特林數的定義$s_{n,k}$則爲

$$[x]_n=\sum_{k=0}^{n}s_{n,k}x^k\text{，}s_{n,0}=0\text{，}n>0$$

例如，$[x]_3=x(x-1)(x-2)$
$=x^3-3x^2+2x$

故$s_{3,3}=1$，$s_{3,2}=-3$，$_{3,1}=2$

第二型史特林數，亦可表示如下，證明參閱〔2〕。

$$S_{n,k}=\frac{k^n}{k!}\frac{\Gamma(n+1)}{\Gamma(n+1-k)}\int_0^{\frac{1}{k}}\cdots\int_0^{\frac{1}{k}}(1-\sum_{i=1}^{k}x_i)^{n-k}\prod_{i=1}^{k}dx_i$$

例如，$n=4$，$k=2$，

$$S_{4,2}=\frac{16}{2}\frac{\Gamma(5)}{\Gamma(3)}\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}}(1-x_1-x_2)^2\,dx_1dx_2=7$$

最後，我們再考慮第二型史特林數的一般化，定義$S_{n,k}^{(r)}$爲n個可區分的球，放在k個不可區分的箱中，且每箱至少r個球的方法數，所以$S_{n,k}^{(1)}=S_{n,k}$，以下定理4爲$S_{n,k}^{(r)}$的遞迴關係。

定理 4：$S^{(r)}_{n,k} = kS^{(r)}_{n-1,k} + \binom{n-1}{r-1}S^{(r)}_{n-r,k-1}$

證明：先考慮 $n-1$ 個球放在 k 個箱中，有以下兩種情況：(i)每箱至少 r 個球：因此另外一個球可放在任意一個箱中，共有 $kS^{(r)}_{n-1,k}$ 種方法。(ii) $k-1$ 個箱中至少有 r 個球，另一個箱子恰有 $r-1$ 個球：因此另外一個球需放在只有 $r-1$ 個球的箱中，故有 $\binom{n-1}{r-1}$ 種方法選這 $r-1$ 個球，剩下 $(n-1)-(r-1)=n-r$ 個球放在 $k-1$ 個箱中且每箱至少 r 個球，則有 $S^{(r)}_{n-r,k-1}$ 種方法，共有 $\binom{n-1}{r-1}S^{(r)}_{n-r,k-1}$ 種方法。得證。

顯然可知，若 $n < kr$，則 $S^{(r)}_{n,k}=0$。若 $n \geq r$，則 $S^{(r)}_{n,1}=1$。例如，

$$S^{(2)}_{4,2} = 2S^{(2)}_{3,2} + \binom{3}{1}S^{(2)}_{2,1}$$

$$= 2\cdot 0 + 3\cdot 1 = 3$$

若球表爲 a、b、c、d，則 3 種方法如下

$\{a, b\}, \{c, d\}$

$\{a, c\}, \{b, d\}$

$\{a, d\}, \{b, c\}$

又如，$S^{(2)}_{5,2} = 2S^{(2)}_{4,2} + \binom{4}{1}S^{(2)}_{3,1}$

$$= 2\cdot 3 + 4\cdot 1 = 10$$

表 5 爲第二型史特林數的一些表值。

$S_{n,k}$ (n \ k)	1	2	3	4	5	6	7	8	9	10
1	1	0	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
3	1	3	1	0	0	0	0	0	0	0
4	1	7	6	1	0	0	0	0	0	0
5	1	15	25	10	1	0	0	0	0	0
6	1	31	90	65	15	1	0	0	0	0
7	1	63	301	350	140	21	1	0	0	0
8	1	127	966	1701	1050	266	28	1	0	0
9	1	255	3025	7770	6951	2646	462	36	1	0
10	1	511	9330	34105	42525	22827	5880	750	45	1

$S^{(2)}_{n,k}$ (n \ k)	2	3	4	5
4	3	0	0	0
5	10	0	0	0
6	25	15	0	0
7	56	105	0	0
8	119	490	105	0
9	246	1918	1260	0
10	501	6825	9450	945

$S^{(3)}_{n,k}$ (n \ k)	3	4	5
9	280	0	0
10	2100	0	0
11	10395	0	0
12	42735	15400	0
13	158301	204200	0
14	549549	1611610	0
15	1827826	10335325	1401400

表 5

應 用

最後我們看一些佔有問題應用的例子，其中球與箱子是否可區分，係根據我們如何解釋來判斷。

例題 1：某醫院九月份共接生 80 個嬰兒，記錄下每個嬰兒的出生日期，則這個事件的發生共有幾種方法？

解：嬰兒看成是球，日期看成是箱子。若不區分嬰兒，只區分日期，則為 $2a$ 的情況，$n=80$，$k=30$，故有 $C(80+30-1, 30-1)=C(109, 29)$ 種方法。若每天至少接生一個嬰兒，則為 $2b$ 的情況，共有 $C(79, 29)$ 種方法。若不區分嬰兒，亦不區分日期，我們只想知道有幾天接生兩個嬰兒，有幾天接生三個嬰兒，………，則為 $4a$ 的情況，需考慮 80 分割成 30 個或更少的正整數和。

例題 2：某打卡員打卡 100 張，共有 30 個錯誤，這事件的發生共有幾種方法？

解：錯誤看成是球，卡片看成是箱子，錯誤的型態不予考慮，則球不可區分。也許我們想知道是否因工作的枯燥而導至錯誤增加的趨勢。故考慮卡片可區分此為 $2a$ 的情況，有 $C(30+100-1, 100-1)=C(129, 99)$ 種方法。

例題 3：假設我們按畢業先後次序，記錄某校資訊系1000位畢業生的性別，則這事件的發生共有幾種方法？

解：1000 人看成 1000 個球，性別男與女則看成兩個可區分的箱子。若不區分此1000人，亦即只想知道男女分配的情形，則此為 $2a$ 的情況，共有

$$C(1000+2-1, 2-1)=C(1001, 1)=1001 \text{ 種}$$

例題 4：電梯上有 8 位乘客，電梯將在 6 層不同的樓停下這事件的發生共有幾種方法？

解：乘客看成是球，每層樓看成是箱子。若我們想知道那些乘客在同一層樓出去，則球可區分，箱子不可區分，此為 $3a$ 的情況，共有 $S(8, 1)+S(8, 2)+\cdots\cdots+S(8, 6)$ 種方法。

例題 5：統計力學（Statistical Mechanics）

在統計力學中，我們想知道 n 個粒子（中子、質子、電子、……），在 k 個不同能階（energy level）上分佈的情形，共有幾種？(i)假設粒子可區分，則為Maxwell-Boltzmann 模式，(ii)假設粒子不可區分，則為Bose-Einstein模式。(iii)假設粒子不可區分，且服從Pauli排除原理（Pauli exclusion principle），亦即不能有兩個粒子在同一能階，則為Fermi-Dirac 模式。

解：(i)此為 $1a$ 的情況，共有 k^n 種。

(ii)此為 $2a$ 的情況，共有 $C(n+k-1, k-1)$ 種。

(iii)共有 $C(k, n)$ 種。

參考文獻

〔1〕 Feller, W. (1950). "*An Introduction to Probability Theory and its Applications*," 2d.ed., John Wiley & Sons.

〔2〕 Olkin, I. & Sobel, M. (1965). "*Integral expressions for tail probabilities of the multinomial and negative multinomial distributions*," Biometrika, 52, p.167-179.

〔3〕 Roberts, F. (1984). "*Applied Combinatorics*," Prentice-Hall.

〔4〕 劉涵初（1987）. "離散與組合數學"，華泰書局。

—本文作者任教於逢甲大學統計系—