OCCUPANCY PROBLEMS

Consider the combinatorial problem of allocating r items (objects,balls) to n boxes (cells): for example, can we enumerate

the number of ways of allocating 6 items to 4 boxes

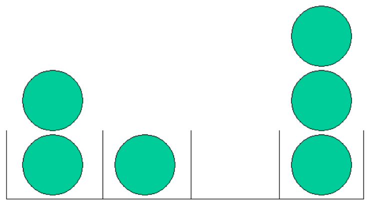

Fig1. Allocation of $n = 6$ balls to $r = 4$ cells

To count the possible number of allocations, we consider the cases of *DISTINGUISHABLE* and *INDISTINGUISHABLE* items separately.

DISTINGUISHABLE ITEMS

If the items are distinguishable, that is, labelled $1, 2, ...r$, we consider *ordered* arrangements. The following two allocations are different:

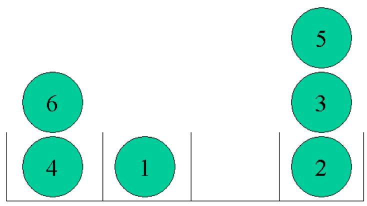

Fig 2a. Allocation 1

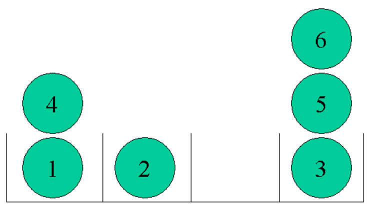

Fig 2b. Allocation 2

and we can record the two allocations as follows

	ITEM	1	2	3	4	5	6
ALLOCATION 1	BOX LABEL SEQUENCE	2	4	4	1	4	1
ALLOCATION 2	BOX LABEL SEQUENCE	1	2	4	1	4	4

and, essentially, we have selected r box labels from n with replacement, where the box labels are ordered. Hence the number of possible allocations is n^r.

If we require

$$\begin{array}{ll} r_1 & \text{items in BOX 1} \\ r_2 & \text{items in BOX 2} \\ \vdots & \quad\vdots \\ r_n & \text{items in BOX } n \end{array}$$

then we must **partition** the box label sequence to contain r_1 1s, r_2 2s,...,r_n ns. Hence the number of possible allocations is given by the partition formula

$$\frac{r!}{r_1!r_2!...r_n!} \qquad \text{where}$$

INDISTINGUISHABLE ITEMS

If the items are indistinguishable, that is, here completely identical, and we wish to consider **distinct** allocation patterns, we must consider *unordered* arrangements; the two allocations in Fig 2a. and Fig 2b are regarded as identical, and identical to the allocation in Fig 1., as the items are not labelled.

For example, consider forming the allocation pattern by dropping the items into the boxes in sequence:

ITEM	1	2	3	4	5	6
SEQUENCE (1)	2	1	2	4	1	4
SEQUENCE (2)	3	2	4	1	3	3
SEQUENCE (3)	2	4	4	1	1	2

Then, we have that the patterns obtained by sequences (1) and (3) are both of the form

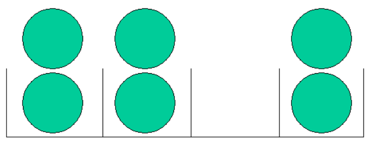

Fig 3a. Allocation patterns for sequences (1) and (3)

and thus we do not count (1) and (3) as distinct patterns, whereas sequence (2) produces an allocation pattern of

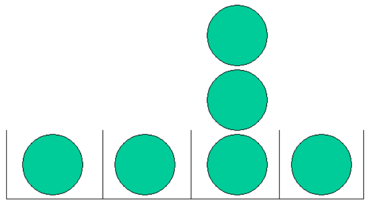

Fig 3b. Allocation patterns for sequence (2)

which is distinct from the pattern for (1) and (3)

To enumerate the number of possible allocation patterns, we utilize a binary sequence representation. We code an allocation pattern by reading from left to right, and writing a 1 for a box edge, and 0 for an item, so that the pattern in Fig3a. is coded

$$1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 1$$

and the pattern in Fig3b is coded

$$1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1$$

The number of possible allocation patterns is equal to the number of binary sequences that correspond to them, and these sequences are composed as follows; they contain $n+1$ 1s (for the box edges) and r 0s (for the items), but also they begin with a 1, and end with a 1. The number of sequences like this is therefore equal to the number of ways of arranging a sequence of $(n+1)+r-2 = n+r-1$ binary digits containing precisely $n-1$ 1s and r 0s. This number is

$$\binom{n+r-1}{r}$$

from combination/binomial coefficient definition, and this is therefore the total number of distinct allocation patterns.

EXAMPLE

If r identical dice are rolled, with $n = 6$ possible scores for each die, the total number of distinct score patterns is

$$\binom{n+r-1}{r} = \binom{6+r-1}{r} = \binom{5+r}{r}$$

for example, with $r = 4$, we could have

DICE NUMBER				SCORE PATTERN					
1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**5**	**6**
6	1	6	2	1	1	0	0	0	2
3	2	3	4	0	1	2	1	0	0
6	2	1	6	1	1	0	0	0	2

and the number of distinct patterns is

$$\binom{9}{4} = 126$$

OCCUPANCY PROBLEMS : EXAMPLES

EXAMPLE 1 Allocate n items to n boxes. Evaluate the probability of event E that no box is empty.

SOLUTION: Probability is

$$\mathrm{P}(E) = \frac{n_E}{n_\Omega} = \frac{n}{n} \times \frac{n-1}{n} \times ... \times \frac{1}{n} = \frac{n!}{n^n}$$

as we allocate by sampling n boxes **with** and **without** replacement for numerator and denominator respectively.

EXAMPLE 2 Allocate r items to n boxes. Evaluate the probability of event E that no box is contains more than one item.

SOLUTION: Probability is

$$\mathrm{P}(E) = \frac{n_E}{n_\Omega} = \frac{n}{n} \times \frac{n-1}{n} \times ... \times \frac{n-r+1}{n} = \frac{n!}{(n-r)!\, n^r} = \frac{(n)_r}{n^r}$$

as we allocate by sampling r boxes **with** and **without** replacement for numerator and denominator respectively.

NOTE: we can re-write $\mathrm{P}(E)$ using a conditional probability/chain rule argument corresponding to a sequence of selection probabilities:

$$\mathrm{P}(E) = 1 \times \left(1 - \frac{1}{n}\right) \times \left(1 - \frac{2}{n}\right) \times ... \left(1 - \frac{r-1}{n}\right)$$

where each term is the conditional probability of choosing a currently empty box, given the allocations at that instant.

SPECIAL CASE : THE BIRTHDAY PROBLEM

In a group of r people, what is the probability that no two people have the same birthday ? Assuming that all of the $n = 365$ days in the year are equally likely to be a birthday, then we identify in EXAMPLE 2 the "boxes" as days, and "items" as people, and evaluate the probability as

$$\frac{(n)_r}{n^r} = \frac{(365)_r}{365^r}$$

which we can evaluate numerically

r	5	10	20	22	23	50
Probability	0.973	0.883	0.589	0.524	0.493	0.030

EXAMPLE 3 Allocate r items to n boxes. Evaluate the probability of event E that box 1 contains precisely k items.

SOLUTION: For $0 \leq k \leq r$, probability is

$$\mathrm{P}(E) = \frac{n_E}{n_\Omega} = \frac{\binom{r}{k}(n-1)^{r-k}}{n^r} = \binom{r}{k}\left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{r-k}$$

For the numerator, we first select the k **items** from r (**without** replacement) to place in box 1, and then select $(r-k)$ **boxes** from the remaining $(n-1)$ (**with** replacement) to house the remaining $(r-k)$ items. For the denominator, we merely select r boxes from n with replacement.