

## From counting to computing the probabilities

Consider a scenario of putting  $n$  objects into  $k$  buckets. We are interested in computing the probability of a particular arrangement of the objects into buckets. Note that multiple objects can be assigned to a single bucket. Furthermore, the ordering of the objects does not matter, meaning that once the objects are in the buckets we only care about the arrangement of the objects not the labeling of the objects in each bucket. Thus, this setting corresponds to “sampling with replacement without ordering”, as explained in the textbook.

For the ease of discussion, let us suppose that  $n = 2k$  and the arrangement of interest is “having exactly 2 objects in each bucket”. If we were to count the total number of ways that  $n$  objects can be put into  $k$  buckets, having in mind the “arrangement of  $n$  balls and  $k - 1$  dash symbols (separators)” interpretation, there are

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

ways to do so. And if the labeling of the objects is not important, there is only one possible “arrangement” where you have exactly two objects in each bucket. But, can we divide these numbers and jump into the conclusion that the probability of this arrangement is the relative frequency  $\frac{1}{\binom{n+k-1}{n}}$ ?

The answer is no and the reason lies in the fundamental fact that “the probability of an event is equal to the relative frequency of the event only when all different arrangements are equally likely”.

So, is it really the case in the argument above? Consider two different arrangements:

- 1) All the objects are put in the first bucket
- 2) Every bucket contains two objects

Giving it some thought, one realizes that, even if the objects are labeled there is only one way for the first arrangement to occur. But for the second arrangement, all different ways of putting a pair of labeled objects in buckets (How many of them are there?  $\underbrace{(2, 2, \dots, 2)}_k$ )

correspond to the same arrangement.

So how can you make sure that you are computing the right relative frequency as the probability? Simply to make sure that all arrangements are equally likely, one can assume that the objects are labeled and account for all possible ways of putting labeled objects in the buckets that correspond to the same “arrangement”.

In the light of this, as computed earlier, there are  $\binom{n}{2, 2, \dots, 2}$  ways to put  $k$  pairs of objects in

$k$  buckets, and there are in total  $k^n$  (each object can be assigned to  $k$  possible buckets) ways of putting  $n$  objects into  $k$  buckets, which are all equally likely. Thus the probability of the event of interest is

$$\frac{\binom{n}{2,2,\dots,2}}{k^n}$$