Chapter 1: Combinatorial Analysis

A major branch of combinatorial analysis called enumerative combinatorics consists of studying methods for counting the number of ways that certain patterns can be formed from finite sets. We will see how it applies to probability theory.

1 Basic counting principles

Example 1. Birthday: 50 students, 12 months. At least 2 students were born in the same month.

1.1 Pigeonhole principle

Theorem 1.1 (Pigeonhole principle, Dirichlet, 1834). If \( n \) items are to be put into \( m \) containers, with \( n > m \), then at least one container must contain more than one item.

Proof. Let us assume that objects and boxes are labeled respectively \( o_1, \ldots, o_n \) and \( b_1, \ldots, b_m \). Without loss of generality (w.l.o.g.), we can put \( o_1 \) into \( b_1, \ldots, o_m \) into \( b_m \). Therefore, there remain \( m - n > 0 \) objects, namely \( o_{m+1}, \ldots, o_n \) that need to be assigned. So at least one container will contain more than one item. \( \square \)

The pigeonhole principle relates to the concept of injection on finite sets.

Theorem 1.2 (Injections on finite sets). Let \( E \) and \( F \) be finite sets and \( f : E \to F \) be a function from \( E \) to \( F \). If \( |E| > |F| \) then \( f \) is not injective. In other words, if \(|E| > |F|\), there is no injection from \( E \) to \( F \).

Lemma 1.3. If \( n \) items are to be put into \( m \) containers, with \( n > m \), then at least one container must contain at least \( \lceil n/m \rceil \) items.

Reminders:

- Floor function: \( \forall x \in \mathbb{R}, \lfloor x \rfloor = \sup \{ n \in \mathbb{Z} \mid n \leq x \} \)
- Ceiling function: \( \forall x \in \mathbb{R}, \lceil x \rceil = \inf \{ n \in \mathbb{Z} \mid n \geq x \} \)

Before proving the above lemma, let us first show the following result:

Property 1.1. \( \forall x \in \mathbb{R}, \lfloor x \rfloor < x + 1 \) (1)
Proof. Let us use a Reductio ad Absurdum argument. Let us assume there is some \( x_0 \in \mathbb{R} \) such that \( \lceil x_0 \rceil \geq x_0 + 1 \). In that case, \( \lceil x_0 \rceil - 1 \) is an integer smaller than \( \lceil x_0 \rceil \) satisfying \( \lceil x_0 \rceil - 1 \geq x_0 \), which contradicts the definition of \( \lceil x_0 \rceil \). \( \square \)

We are now ready to prove the lemma.

Proof. All \( n \) items are put into the \( m \) containers. We will use a Reductio ad Absurdum argument. Let us assume that all containers contain at most \( \lceil n/m \rceil - 1 \) items. Then, the maximum number of objects in the boxes is

\[
m \left( \lceil n/m \rceil - 1 \right) < m \left( n/m \right) = n
\]

which contradicts the fact that the \( n \) items are all in the boxes. \( \square \)

Example 1 bis. Birthday : 50 students, 12 months. At least \( \lceil 50/12 \rceil = 5 \) students were born in the same month.

1.2 Rule of sum

Proposition 1.1 (Rule of sum). Let us consider \( r \) events. If there are \( n_1 \) possible outcomes for the first event, \( \ldots \), \( n_r \) possible outcomes for the \( r \)th event and if any two distinct events cannot both occur (events are mutually exclusive), then there are \( \sum_{i=1}^{r} n_i \) total possible outcomes for the events.

The rule of sum relates to the following set theory property.

Proposition 1.2. If \( S_1, \ldots, S_n \) are pairwise disjoint sets then

\[
\left| \bigcup_{i=1}^{n} S_i \right| = \sum_{i=1}^{n} |S_i|
\]

Example 2. You plan a trip. You hesitate between different destinations : Europe (3 cities) / Asia (4 cities) / South America (2 cities). But you can only choose one of them due to a restricted budget. How many different destinations are possible ? \( 3 + 4 + 2 = 9 \)

1.3 Rule of product

Proposition 1.3 (Rule of product). If \( r \) experiments are to be performed sequentially and the first experiment can be performed in \( n_1 \) ways, \( \ldots \), the \( r \)th experiment in \( n_r \) ways, then there are \( \prod_{i=1}^{r} n_i \) ways to perform the \( r \) experiments.

The rule of product relates to the concept of cartesian product.
Proposition 1.4. Let $S_1, \ldots, S_n$ be $n$ sets then

$$|S_1 \times \ldots \times S_n| = \prod_{i=1}^{n} |S_i|$$

Example 3. You plan a tour of Western Europe and want to visit London, Paris and Rome. There are 2 recommended roads from London to Paris and 3 from Paris to Rome. $2 \cdot 3 = 6$

Example 4. Let us consider 3 songs consisting of 20 lines. You can compose $3^{20} = 3,486,784,401$ different songs of 20 lines where line 1 comes from line 1 of any of the 3 songs, \ldots, line 20 comes from line 20 of any of the 3 songs. It takes around 6,634 years to listen to all the songs at the pace of 1 song per minute.

Example 5. You go to a restaurant. There you can either choose one starter and one course or one course and one dessert, but you cannot take a starter, a course and a dessert. That day, the restaurant proposes 4 starters, 4 courses and 3 desserts. How many different menus are possible? $4 \cdot 4 + 4 \cdot 3 = 28$

2 Permutations

Example 6. I want to visit 10 people, each of them living in different cities. How many different orders are possible to visit them? $10 \cdot 9 \cdot 8 \ldots 1 = 3,628,800$

Definition 2.1 (Permutation). An ordered ranking of $n \in \mathbb{N}^*$ distinct elements is called a permutation.

Proposition 2.1. There are $n(n-1) \ldots 1 = n!$ permutations of $n \in \mathbb{N}^*$ distinct elements.

Proof. Rule of product

Definition 2.2 (Factorial). $\forall n \in \mathbb{N}^*, \ n! = \prod_{i=1}^{n} i$

By convention $0! = 1$

Example 6 bis. I want to visit 10 people, each of them living in different cities. Among these 10 people, there are 6 relatives of mine. How many different orders are possible if I want to visit my family first? $6!4! = 720 \cdot 24 = 17,280$

Example 7. I have 11 DVDs that I want to put on my shelf. Of these, 4 are action movies, 2 are science-fiction movies, 3 are fantasy movies, and 2 are comedy movies. I want to arrange my DVDs so that all the DVDs dealing with the same subject are together on the shelf. How many different arrangements are possible? $4!4!2!3!2! = 13,824$
3 Partial permutations–Arrangements

Definition 3.1 (Partial Permutation). A partial permutation or a (combinatorial) arrangement is an ordered ranking of \( p \) items among \( n \in \mathbb{N}^* \) elements \((p \leq n)\)

Proposition 3.1. The number of arrangements of \( p \) items among \( n \in \mathbb{N}^* \) elements \((p \leq n)\) is denoted \( A_n^p \) and is equal to

\[
A_n^p = n \cdot (n - 1) \ldots (n - p + 1) = \frac{n!}{(n-p)!}
\]

Proof. Rule of product \( \square \)

Example 8. Formula 1 racing. There are 22 drivers. Only the first 10 drivers crossing the finish line earn championship points. \( A_{22}^{10} = \frac{22!}{(22-10)!} \)

4 Combinations

We are often interested in determining the number of different groups of \( r \) objects that could be formed from a total of \( n \) objects

Definition 4.1 (Combination). A combination is an unordered collection of \( p \) items among \( n \in \mathbb{N}^* \) elements \((p \leq n)\)

Proposition 4.1. The number of combinations of \( p \) items among \( n \in \mathbb{N}^* \) elements \((p \leq n)\) is denoted \( \binom{n}{p} \) (read ‘\( n \) choose \( p \)’) and is equal to

\[
\binom{n}{p} = \frac{n!}{p!(n-p)!}
\]

Proof. If we perform all the permutations of each combination, we obtain all the arrangements of \( p \) items among \( n : A_n^p = p!(\binom{n}{p}) \).

Example 9. Tennis. ATP World Tour Finals. There are 100 players. Only the first 8 in the ATP ranking at the end of the year get access to the finals. \( \binom{100}{8} \)

Property 4.1.

\[
\forall n, p \in \mathbb{N}, \binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}, \; p \leq n \tag{2}
\]

Proof. OK for an analytic proof.

Let us think of a combinatorial argument. Consider a set \( S = \{e_1, \ldots, e_{n+1}\} \) of \( n+1 \) distinct elements and focus on a particular element, say \( e_1 \) w.l.o.g. Combinations of \( p+1 \) elements among \( n+1 \) can be divided as follows: combinations containing \( e_1 \), there are \( \binom{n}{p} \) such combinations and there are \( \binom{n}{p+1} \) combinations without \( e_1 \). \( \square \)
Another useful combinatorial identity is

**Property 4.2.**
\[ \forall n, p \in \mathbb{N}, \binom{n}{p} = \binom{n}{n-p}, \quad p \leq n \quad (3) \]

**Proof.** Think of a combinatorial argument.

**Theorem 4.1 (Binomial theorem).** For all \( x, y \in \mathbb{R} \) and \( n \in \mathbb{N}^* \),
\[ (x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \quad (4) \]

**Proof.** By induction. Let \( x, y \) be 2 real numbers. For a given \( n \in \mathbb{N}^* \), denote by \( P_n \) the proposition : \( (x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \).

First, let us prove \( P_1 \). We have:
\[ (x + y)^1 = x + y = (\binom{1}{0}) x^0 y^{1-0} + (\binom{1}{1}) x^1 y^{1-1} \]

Assuming \( P_n \) is true for some \( n \in \mathbb{N}^* \), we have:
\[
(x + y)^{n+1} = (x + y)(x + y)^n \\
= (x + y) \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \\
= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
= \sum_{i=1}^{n+1} \binom{n}{i-1} x^i y^{n+1-i} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
= x^{n+1} + \sum_{i=1}^{n} \left\{ \left( \binom{n}{i-1} \right) + \binom{n}{i} \right\} x^i y^{n+1-i} + y^{n+1} \\
= \binom{n+1}{n+1} x^{n+1} y^{n+1-(n+1)} + \sum_{i=1}^{n} \binom{n+1}{i} x^i y^{n+1-i} + \binom{n+1}{0} y^{n+1-0} \\
= \sum_{i=0}^{n+1} \binom{n+1}{i} x^i y^{n+1-i} 
\]

\[ \square \]
5 Multinomial Coefficients

We shall now determine the number of ways to divide \( n \) items into \( r \) distinct groups of respective sizes \( n_1, \ldots, n_r \) such that \( \sum_{i=1}^{r} n_i = n \). W.l.o.g., there are \( \binom{n}{n_1} \) choices for the first group; for each choice of the first group, there are \( \binom{n-n_1}{n_2} \) for the second group; and so on. From the rule of product there are:

\[
\binom{n}{n_1} \binom{n-n_1}{n_2} \cdots \binom{n-n_1-\cdots-n_{r-1}}{n_r} = \frac{n!}{\prod_{i=1}^{r} n_r!}
\]

possible divisions.

**Definition 5.1.** Let \( n, n_1, \ldots, n_r \in \mathbb{N}^* \) such that \( \sum_{i=1}^{r} n_i = n \), we define the multinomial coefficient by

\[
\binom{n}{n_1, \ldots, n_r} = \frac{n!}{\prod_{i=1}^{r} n_r!}
\]

Thus, \( \binom{n}{n_1, \ldots, n_r} \) represents the number of possible divisions of \( n \) distinct objects into \( r \) distinct groups of respective sizes \( n_1, \ldots, n_r \).

It also relates to the problem of finding the number of permutations of a set of \( n \) objects when certain of the objects are indistinguishable from each other.