

Contents

1	Permutations and Combinations	2
1.1	Basic Counting Principles	2
1.2	Permutations	5
1.2.1	Linear Permutations	5
1.2.2	Circular Permutations	6
1.2.3	Permutations with Repetitions	7
1.3	Combinations	8
1.3.1	Counting Subsets	9
1.3.2	Combinations with Repetitions	10
1.4	Fun Problems Collection	11
1.5	Distribution Problems	13
1.5.1	Stirling Numbers	13
1.5.2	Distinct Objects into Distinct Collections	14
1.6	Binomial and Multinomial Coefficients	15
1.6.1	Binomial Coefficients	15
1.6.2	Multinomial Coefficients	17
2	Pigeonhole Principle	19
2.1	Pigeonhole Principle	19
2.2	Ramsey Types of Problems	19
3	Principle of Inclusion and Exclusion	23
3.1	Principle of Inclusion and Exclusion (PIE)	23
3.2	General Principle of Inclusion and Exclusion (GPIE)	24
3.3	Stirling Numbers of the Second Kind	24
3.4	Derangement	25

Permutations and Combinations

1.1 Basic Counting Principles

An important motivation to study combinatorics is to count the **number of ways** in which an event may occur. Intuitively, we have two approaches to count.

The first approach is to categorise the event into **non-overlapping cases**. This means that we break an event into mutually exclusive sub-events, after which we can count the number of ways for each sub-event to occur. The aggregate of these counts is the total number of ways for the original event to occur.

Those familiar with basic set theory may consider E to be the set containing all distinct ways for an event to occur. By breaking up the event, we essentially establish a **partition** of E , so that the sum of cardinalities of all the elements in that partition equals the cardinality of E .

This motivates us to write the following principle using set notations.

Theorem 1.1.1 ► Addition Principle (AP)

Let $k \in \mathbb{N}^+$ and let A_1, A_2, \dots, A_k be k finite sets which are pairwise disjoint, i.e. $A_i \cap A_j = \emptyset$ whenever $i \neq j$, then

$$\left| \bigcup_{i=1}^k A_i \right| = \sum_{i=1}^k |A_i|.$$

Proof. The case where $k = 1$ is trivial.

Suppose that when $k = n$, we have

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n |A_i|$$

for any n finite sets which are pairwise disjoint. Let A_{n+1} be an arbitrary finite set

which is disjoint with any of the A_i 's from the n sets. So we have:

$$\begin{aligned}
 \left| \bigcup_{i=1}^{n+1} A_i \right| &= \left| \left(\bigcup_{i=1}^n A_i \right) \cup A_{n+1} \right| \\
 &= \left| \bigcup_{i=1}^n A_i \right| + |A_{n+1}| - \left| \left(\bigcup_{i=1}^n A_i \right) \cap A_{n+1} \right| \\
 &= \left(\sum_{i=1}^n |A_i| \right) + |A_{n+1}| - |\emptyset| \\
 &= \sum_{i=1}^{n+1} |A_i|.
 \end{aligned}$$

Therefore, the original statement holds for all $k \in \mathbb{N}^+$. □

Remark. In more casual language, this means that if an event E_k has n_k distinct ways to occur, then there is $\sum_{i=1}^k n_k$ ways for at least one of the events E_1, E_2, \dots, E_k to occur, provided that E_i and E_j can never occur concurrently whenever $i \neq j$.

Given an event E , the other approach to count the number of ways for it to occur is to break E up internally into **non-overlapping stages**.

With set notations, we can write the i -th stage for E to occur as e_i , and so a way for E to occur can be represented by an ordered tuple (e_1, e_2, \dots, e_k) , where k is the total number of stages to undergo for E to occur.

Let E_i denote the set of all distinct ways to undergo the i -th stage of E , then it is easy to see that E is just the **Cartesian product** of all the E_i 's. Hence, we derive the following principle:

Theorem 1.1.2 ► Multiplication Principle (MP)

Let $k \in \mathbb{N}^+$ and let A_1, A_2, \dots, A_k be k pairwise disjoint finite sets, then

$$\left| \prod_{i=1}^k A_i \right| = \prod_{i=1}^k |A_i|.$$

Proof. The case where $k = 1$ is trivial.

Suppose that when $k = n$, we have

$$\left| \prod_{i=1}^n A_i \right| = \prod_{i=1}^n |A_i|$$

for any n finite sets which are pairwise disjoint. Let A_{n+1} be an arbitrary finite set which is disjoint with any of the A_i 's from the n sets. Take $a_i, a_j \in A_{n+1}$. Note that for all $\mathbf{a} \in \prod_{i=1}^n A_i$, $(\mathbf{a}, a_i) \neq (\mathbf{a}, a_j)$ whenever $a_i \neq a_j$. This means that

$$\begin{aligned} \left| \prod_{i=1}^{n+1} A_i \right| &= \left| \prod_{i=1}^n A_i \times A_{n+1} \right| \\ &= \left| \prod_{i=1}^n A_i \right| |A_{n+1}| \\ &= \left(\prod_{i=1}^n |A_i| \right) |A_{n+1}| \\ &= \prod_{i=1}^{n+1} |A_i| \end{aligned}$$

Therefore, the original statement holds for all $k \in \mathbb{N}^+$. □

Remark. In more casual language, this means that if an event E requires k stages to be undergone before it occurs and the i -th stage has n_i ways to complete, then there is $\prod_{i=1}^k n_k$ ways for E to occur, provided that no two different stages complete concurrently.

Often times, it is not straight-forward to count directly due to the presence of restrictions. We shall consider the following:

Let E be the set of all possible ways for an event to occur. Let p be some predicate representing some restriction and let $E(p)$ denote the set of all possible ways for the event to occur while p holds. Note that:

$$E(p) \cup E(\neg p) = E \quad \text{and} \quad E(p) \cap E(\neg p) = \emptyset,$$

i.e. $\{E(p), E(\neg p)\}$ is a partition of E . Therefore, to count the number of ways for the event to occur while p holds, it suffices to compute $E(\neg p)$, i.e. find the number of ways for the event to occur while p does not hold.

Theorem 1.1.3 ► Principle of Complementation

Let U be a set and let $E \subseteq U$, then

$$|E| = |U| - |U - E|.$$

Remark. It may also help to think the Principle of Complementation as an inverse of the Addition Principle, where

$$E = \bigcup_{i=1}^n E_i$$

is the total number of ways for an event to occur and E_p is the number of ways for the event to occur with restriction p .

In some cases, it is difficult to count the objects directly. However, note that the events for which we are counting are just sets, so we can treat “number of occurrences” of an event as the **cardinality** of a set. Therefore, we have the following principles:

Theorem 1.1.4 ► Injection Principle

Let A, B be finite sets. If there exists an injection

$$f : A \hookrightarrow B,$$

then $|A| \leq |B|$.

Theorem 1.1.5 ► Bijection Principle

Let A, B be finite sets. If there exists a bijection

$$f : A \rightarrow B,$$

then $|A| = |B|$.

1.2 Permutations

1.2.1 Linear Permutations

A fundamental problem in combinatorics is described as follows: given a set S , how many ways are there to arrange r elements in S , i.e. how many **distinct sequences** can be formed using the elements in S without repetition? The process of selecting elements from S and arranging them as a sequence is known as **permutation**.

Note that forming a sequence using r elements from a set S is an event consisting of r stages, as we need to select an element for each of the r terms of the sequence. Suppose S has n elements. For the first term of the sequence, we can choose any of the elements in S , so there is n ways to do it. For the second term, since we cannot repeat the elements, we are left with $n - 1$ choices.

Continue choosing elements in this way, we realise that if we choose the terms sequentially, when we reach the k -th term we will be left with $n - k + 1$ options as the previous $(k - 1)$ terms have taken away $(k - 1)$ elements. By Theorem 1.1.2, we know that the number of sequences which can be formed is given by $\prod_{i=1}^r (n - r + i)$.

Definition 1.2.1 ► Permutations

Let A be a finite set such that $|A| = n$, an r -permutation of A is a way to arrange r elements of A , denoted as P_r^n and given by

$$P_r^n = \prod_{i=1}^r (n - r + i) = \frac{n!}{(n - r)!}.$$

With some algebraic manipulations, it is easy to derive the following formula, which we, however, will prove in a combinatorial manner.

Theorem 1.2.2

Let $n, r \in \mathbb{N}$ with $r \leq n$, then $P_r^{n+1} = P_r^n + rP_{r-1}^n$.

Proof. Let $S = \{x \in \mathbb{N}^+ : x \leq n + 1\}$ represent $(n + 1)$ distinct objects. Consider a permutation of S :

If $n + 1$ is not inside the permutation, this is equivalent to an r -permutation of $S - \{n + 1\}$, so there are P_r^n such permutations.

If $n + 1$ is inside the permutation, it means we need to first find an $(r - 1)$ -permutation of $S - \{n + 1\}$, which has P_{r-1}^n ways to do. After that, we need to insert $n + 1$ into each of these $(r - 1)$ -permutations. Note that for each of such permutations, there are r positions into which we can place $n + 1$. Therefore, the total number of r -permutations of S derived in this manner is rP_{r-1}^n .

Therefore, there are $P_r^n + rP_{r-1}^n$ r -permutations of S , i.e. $P_r^{n+1} = P_r^n + rP_{r-1}^n$. \square

1.2.2 Circular Permutations

Consider arranging n distinct objects around a circle. If the slots around the circle are uniquely labelled, this is exactly the same as permutations along a straight line.

However, if the slots are identical, i.e. we are arranging n distinct objects around a circle with identical slots, only the **relative positions** of the objects matter.

Let \mathbf{x}_i be an arbitrary straight-line permutations of the n objects and let \mathbf{y}_i be the corre-

sponding circular permutation of the n objects.

Note that if we translate every element in x_i by k positions, this will result in a different straight-line permutation x_j but does not change the corresponding circular permutation because the relative positions of the objects remain unchanged.

Notice that k can take the values $0, 1, 2, \dots, n-1$, so for the same set of n distinct objects, every circular permutation is mapped to n straight-line permutations.

Definition 1.2.3 ► Circular Permutations

Let A be a finite set such that $|A| = n$, a circular r permutation of A is a way to arrange r elements of A around a circular locus, denoted as Q_r^n and given by

$$Q_r^n = \frac{P_r^n}{r} = \frac{n!}{r(n-r)!}.$$

1.2.3 Permutations with Repetitions

Consider the following scenario: suppose T_1, T_2, \dots, T_n are distinct labels and a_1, a_2, \dots, a_r are r distinct objects. If we wish to associate each object with one label, then each of the ways to do so is a permutation, the terms of which are taken from the T_i 's. Notice that under this setting, the T_i 's are allowed to be repeated in a permutation.

To reason such problems, we first introduce the notion of a *multi-set*.

Definition 1.2.4 ► Multi-set

The **multi-set** is defined to be

$$M := \{r_1 \cdot a_1, r_2 \cdot a_2, \dots, r_n \cdot a_n\},$$

where a_i has r_i identical copies.

Observe that in a permutation of such a multi-set M , the object a_i is repeated r_i times. This brings us to the following formula:

Proposition 1.2.5 ► Generalised Formula for Permutations

Let $k \in \mathbb{N}^+$ and let A_1, A_2, \dots, A_k be k distinct objects, where A_i occurs $n_i > 0$ times for $i = 1, 2, \dots, k$, then each permutation for these objects, corresponds to a unique permutation of the multi-set

$$M = \{n_1 \cdot A_1, n_2 \cdot A_2, \dots, n_k \cdot A_k\},$$

and the total number of permutations is given by

$$\frac{\left(\sum_{i=1}^k n_i\right)!}{\prod_{i=1}^k (n_i)!}.$$

We may also consider the special case where we want to form an r -permutation with k objects, where each object can repeat an arbitrary number of times. Notice that this is equivalent to finding the permutations of the multi-set

$$M = \{\infty \cdot a_1, \infty \cdot a_2, \dots, \infty \cdot a_k\}.$$

In this case, we can simply iterate through each of the slots in the r -permutation and choose the object to be placed there. Trivially, there are always k choices for each slot.

Proposition 1.2.6 ▶ A Special Case

Consider the multi-set

$$M = \{\infty \cdot a_1, \infty \cdot a_2, \dots, \infty \cdot a_n\}.$$

The number of r -permutations formed using the elements from M is given by n^r .

1.3 Combinations

Beside permutations, there are also occasions where we only care about which elements from a particular set are selected instead of the order of selection.

Note that if we want to find a selection of r elements from a set A where the order of selected elements does not matter, it is equivalent to finding a subset of A containing r elements. This motivates us to give the following definition:

Definition 1.3.1 ▶ Combinations

Let A be a finite set such that $|A| = n$, an r -combination of A is a set $B \subseteq A$ with $|B| = r$. The number of combinations of A is given by

$$C_r^n = \frac{P_r^n}{P_r^r} = \frac{n!}{r!(n-r)!} = \binom{n}{r}.$$

Remark. Two obvious results:

1. If $r > n$ or $r < 0$, $C_r^n = 0$;
2. $C_r^n = C_{n-r}^n$ (By Theorem 1.1.3).

Similar to permutations, we have the following important identity:

Theorem 1.3.2 ► Pascal's Triangle

Let n be an integer with $n \geq 2$ and let r be an integer with $0 \leq r \leq n$, then

$$C_r^{n+1} = C_{r-1}^n + C_r^n.$$

Proof. Let $S = \{x \in \mathbb{N}^+ : x \leq n+1\}$ represent $(n+1)$ distinct objects. Consider an r -combination T of S :

If $n+1 \notin T$, this is equivalent to an r -combination of $S - \{n+1\}$, so there are C_r^n such permutations.

If $n+1 \in T$, it suffices to find an $(r-1)$ -combination of $S - \{n+1\}$, which has C_{r-1}^n ways to do.

Therefore, there are $C_r^n + C_{r-1}^n$ r -combinations of S , i.e. $C_r^{n+1} = C_r^n + C_{r-1}^n$. □

1.3.1 Counting Subsets

A useful application of combinations, derived directly from the definition, is to count the number of subsets for a given set which is finite. In other words, given a set A with $|A| = n \in \mathbb{N}$, we wish to find a general formula for $|\mathcal{P}(A)|$.

Let A_i be the set of all subsets of A whose cardinality is i , then clearly

$$|\mathcal{P}(A)| = \sum_{i=0}^n |A_i| = \sum_{i=0}^n C_i^n.$$

We can expand the above expression algebraically and realise that it simplifies to 2^n . However, in a combinatorial perspective, we are able to prove this result in a more succinct manner:

Theorem 1.3.3 ▶ General Formula for $|\mathcal{P}(A)|$

Let A be a finite set. If $|A| = n$, then $|\mathcal{P}(A)| = 2^n$.

Proof. Let S be an arbitrary subset of A . Consider an arbitrary element $a \in A$, then either $a \in S$ or $a \notin S$.

Let $a_i \in A$ for $i = 1, 2, \dots, n$. For all $S \in \mathcal{P}(A)$, We replace a_i by 1 if $a_i \in S$, and by 0 otherwise. Let B be the set of all binary sequences of length n . It is clear that there exists a bijection between $\mathcal{P}(A)$ and B , and so $|\mathcal{P}(A)| = |B|$.

For each binary sequence of length n , each of its digits is either 0 or 1. By Theorem 1.1.2, this means that there are in total 2^n such binary sequences. Therefore,

$$|\mathcal{P}(A)| = |B| = 2^n.$$

□

1.3.2 Combinations with Repetitions

Given a set of n distinct objects, we wish to find the number of combinations taken from the set where any element is allowed to be selected for multiple times. To reason about such problems, we introduce the following notion of a *multi-subset*.

Definition 1.3.4 ▶ Multi-subset

Let $M = \{\infty \cdot a_1, \infty \cdot a_2, \dots, \infty \cdot a_n\}$ be a multi-set where $n \in \mathbb{N}$. An r -element **multi-subset** of M is the set

$$\{m_1 \cdot a_1, m_2 \cdot a_2, \dots, m_n \cdot a_n\}$$

where each of the m_i 's is a non-negative integer such that $\sum_{i=1}^n m_i = r$. We denote the number of r -element multi-subsets of M by H_r^n .

Intuitively, we can view each of the a_i 's as a “box” which holds m_i balls. Therefore, if we can find the ways to distribute a total of r balls into the n “boxes”, each of the distribution will then correspond to a multi-subset of M !

Proposition 1.3.5 ▶ A Formula for H_r^n

Let $M = \{\infty \cdot a_1, \infty \cdot a_2, \dots, \infty \cdot a_n\}$, then the number of r -element multi-subsets of M is given by

$$H_r^n = C_r^{r+n-1}.$$

Proof. Consider a binary string formed using r zeros. Note that we can insert a number of ones into the binary string to partition the zeros into different sections. Now, suppose we insert $(n - 1)$ ones into the binary string, it will result in a binary string containing r zeros and $(n - 1)$ ones, such that the zeros are partitioned into n sections.

We can then number the sections using $1, 2, 3, \dots, n$. For the i -th section, the number of zeros is recorded as m_i . Notice that in this case, each of the binary strings will correspond to a multi-subset in the form of

$$\{m_1 \cdot a_1, m_2 \cdot a_2, \dots, m_n \cdot a_n\}.$$

One can check that this will establish a bijection between the set of all multi-subsets containing n distinct types of objects and the set of all binary strings with r zeros and $(n - 1)$ ones.

Therefore, the above has proven that $H_r^n = C_r^{r+n-1}$. □

1.4 Fun Problems Collection

No Consecutive

Let $X = \{x \in \mathbb{N}^+ : x \leq n\}$. Find the number of r -combinations of X such that $|x_i - x_j| \geq 2$ whenever $i \neq j$.

Solution. Note that for each of the r -combinations of X , we can re-write it into a strictly increasing sequence $\{x_i\}_{i=1}^r$ with $x_i \in X$ and $x_{i+1} - x_i \geq 2$. Let the set of all such sequences be S_X . Consider the function

$$f : S_X \rightarrow Y$$

where $f(\{x_i\}_{i=1}^r) = \{x_i - i + 1\}_{i=1}^r$ and Y is the set of all strictly increasing sequences whose terms are integers bounded between 1 and $n - r + 1$ inclusive.

Take two different sequences $s_1, s_2 \in S_X$. Let the k -th terms of s_1 and s_2 be k_1 and k_2 respectively such that $k_1 \neq k_2$. Let $t_1 = f(s_1)$ and $t_2 = f(s_2)$, whose k -th terms are h_1 and h_2 respectively. Note that $h_1 = k_1 - k + 1 \neq k_2 - k + 1 = h_2$, so $t_1 \neq t_2$, which means f is injective.

Take an arbitrary $\{y_i\}_{i=1}^r \in Y$ and consider the sequence $\{y_i + i - 1\}_{i=1}^r$. Note that $1 \leq y_i \leq n - r + 1$, so $1 \leq y_i + i - 1 \leq n$ for $i = 1, 2, \dots, r$, which means $\{y_i + i - 1\}_{i=1}^r \in S_X$. This means that for all $y \in Y$, there exists some $s \in S_X$ such that $f(s) = y$, so f is surjective.

Therefore, f is a bijection and so

$$|S_X| = |Y| = C_r^{n-r+1}.$$

Top Secret

Six scientists are working on a secret project. They wish to lock up the documents in a cabinet so that the cabinet can be opened when and only when three or more of the scientists are present. Suppose that each lock has one and only one key with the locks being pairwise distinct.

1. What is the smallest number of locks needed?
2. What is the smallest number of keys that each scientist needs to carry?

Solution. Let $S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$ be the set of scientists and L be the set of locks. Define T to be the set of all subsets of S with 2 elements, then for any $P \in T$, P is a pair of distinct scientists.

Let $L_P \subseteq L$ be the set of locks which cannot be opened by the pair of scientists in P . Note that for all $P \in T$, $L_P \neq \emptyset$.

Consider $P, Q \in T$ with $P \neq Q$, then $|P \cup Q| \geq 3$. Thus, $L_P \cap L_Q = \emptyset$. Now, define the function

$$f: T \rightarrow L$$

such that $f(P) = \ell_P$ where ℓ_P is an arbitrary element from L_P (note that we can do this without using Choice as each of the L_P 's is finite).

Suppose that there exist $P, Q \in T$ with $P \neq Q$ such that $f(P) = f(Q) = \ell_0$. Notice that this would mean that $\ell_0 \in L_P \cap L_Q$, which is a contradiction. Therefore, f must be injective. Hence,

$$|L| \geq |T| = C_2^6 = 15.$$

Now, choose an arbitrary element from S , say s_k , and consider the set $S' = S - \{s_k\}$. Let T' be the set of all subsets of S' with 2 elements, then for any $P \in T'$, $L_{P \cup \{s_k\}} = \emptyset$, i.e., s_k must have the keys to open all locks which the pair P cannot open with their own keys. Let this set of keys be K_P . Note that $L_P \cap L_Q = \emptyset$ whenever $P \neq Q$, so $K_P \cap K_Q = \emptyset$ whenever $P \neq Q$. Let the set of all keys held by s_k be H_k , then

$$|H_k| \geq |S'| = C_2^5 = 10.$$

1.5 Distribution Problems

Another problem in which we are interested is to **distribute** the objects from a collection into a finite number of sub-collections. Specifically, we consider the number of ways to

1. distribute r distinct objects into n identical collections;
2. distribute r distinct objects into n distinct collections;
3. distribute r identical objects into n identical collections;
4. distribute r identical objects into n distinct collections.

These 4 basic models of distribution problems give rise to many variations. One important tool to solve these problems is the *Stirling Numbers*.

1.5.1 Stirling Numbers

Definition 1.5.1 ► Stirling Numbers of the First Kind

Let $r, n \in \mathbb{N}$ such that $0 \leq n \leq r$, then the **Stirling Number of the First Kind**, $s(r, n)$, is the number of ways to arrange r **distinct** objects around n **identical** circles such that no circle is empty.

Remark. Some obvious results:

1. $s(r, 0) = 0$ if $r \geq 1$.
2. $s(r, r) = 1$ if $r \geq 0$.
3. $s(r, 1) = (r - 1)!$ if $r \geq 2$.
4. $s(r, r - 1) = C_2^r$ if $r \geq 2$.

One thing to take note of is that there is no general algebraic formula for $s(r, n)$. Instead, all Stirling Numbers of the First Kind follow a recurrence relation in 2 parameters.

Theorem 1.5.2 ▶ A Recurrence Relation for $s(r, n)$

For $r, n \in \mathbb{N}^+$ and $r \leq n$, we have

$$s(r, n) = s(r - 1, n - 1) + (r - 1)s(r - 1, n).$$

Proof. Consider the set

$$\{x_1, x_2, \dots, x_r\}$$

to be the r distinct objects. We consider two cases.

If x_r is distributed to a circle such that it is the only objects around that circle, it suffices to find the number of ways to arrange the rest $(r - 1)$ distinct objects around the rest $(n - 1)$ identical circles, which there are $s(r - 1, n - 1)$ ways to do.

If x_r is adjacent to some other object, we can first arrange the rest $(r - 1)$ distinct objects around the n identical circles, which there are $s(r - 1, n)$ ways to do. After that, we can choose any one of the $(r - 1)$ spaces between two adjacent objects to slot in x_r . Therefore, there are $(r - 1)s(r - 1, n)$ ways to distribute the objects.

By Theorem 1.1.1, the total number of distributions is

$$s(r, n) = s(r - 1, n - 1) + (r - 1)s(r - 1, n).$$

□

1.5.2 Distinct Objects into Distinct Collections

The easiest case for this distribution problem is that each collection contains at most one object, i.e., each collection either contains an object, or contains no object at all.

Proposition 1.5.3

The number of ways to distribute r distinct objects into n distinct collections such that each collection contains at most 1 object is given by P_r^n .

We then consider the case where each collection can contain any number of objects.

Proposition 1.5.4

The number of ways to distribute r distinct objects into n distinct collections such that

each collection can contain any number of objects is given by

$$\frac{(n+r-1)!}{(n-1)!}.$$

1.6 Binomial and Multinomial Coefficients

1.6.1 Binomial Coefficients

In previous sections on combinations, we have already introduced the ways to choose r items from a collection of n distinct items, the number of which is given by

$$C_r^n = \binom{n}{r}.$$

In this section, we will see that there is a clear relation between the number of r -combinations to the *binomial expansion*.

Theorem 1.6.1 ► Binomial Expansion

Let $n \in \mathbb{N}$, then

$$(x+y)^n = \sum_{i=0}^n \left[\binom{n}{i} x^{n-i} y^i \right].$$

Proof. Note that

$$(x+y)^n = \prod_{i=1}^n (x+y).$$

Let the coefficient of $x^{n-r}y^r$ be k , where r is an integer with $0 \leq r \leq n$. Notice that k is exactly the number of ways to choose $(n-r)$ copies of x 's from the distinct n terms of $(x+y)$, which is given by

$$\binom{n}{n-r} = \binom{n}{r}.$$

Therefore, summing up the terms, we have

$$(x+y)^n = \sum_{i=0}^n \left[\binom{n}{i} x^{n-i} y^i \right].$$

□

There are many nice algebraic properties about the binomial coefficient $\binom{n}{r}$. Beside Theorem 1.3.2, the following can be proven with a few simple algebraic manipulations:

1. $\binom{n}{r} = \frac{n}{r} \binom{n-1}{r-1}$.
2. $\binom{n}{r} = \frac{n-r+1}{r} \binom{n}{r-1}$.
3. $\binom{n}{m} \binom{m}{r} = \binom{n}{r} \binom{n-r}{m-r}$.

We will focus on two important identities about the binomial coefficients, one of which is as follows:

Theorem 1.6.2 ▶ Vandermonde's Identity

Let $m, n, r \in \mathbb{N}$, then

$$\sum_{i=0}^r \left[\binom{m}{i} \binom{n}{r-i} \right] = \binom{m+n}{r}.$$

Proof. Consider the set

$$X = \{a_1, a_2, \dots, a_m, b_1, b_2, \dots, b_n\}.$$

Clearly, $|X| = m+n$. Let $A \subseteq X$ such that $|A| = r$, we consider the number of such A 's.

Let i be an integer with $0 \leq i \leq r$. If A contains exactly i elements from the a_j 's, then it will contain exactly $(r-i)$ elements from the b_j 's. Therefore, the number of A 's for each $i = 0, 1, 2, \dots, r$ is given by

$$\binom{m}{i} \binom{n}{r-i}.$$

However, we know that the total number of such A 's is just C_r^{m+n} , so we have

$$\sum_{i=0}^r \left[\binom{m}{i} \binom{n}{r-i} \right] = \binom{m+n}{r}.$$

□

Recall that in Theorem 1.3.2, we essentially established a recurrence relation for the bino-

mial coefficients. Thus in theory, it is possible to generate any binomial coefficient recursively. The following theorem illustrates this.

Theorem 1.6.3 ► Chu Shih-Chieh (CSC) Identity

Let $r, n, k \in \mathbb{N}$ with $n \geq r$, then

$$\sum_{i=0}^{n-r} \binom{r+i}{r} = \binom{n+1}{r+1}, \quad \sum_{i=0}^k \binom{r+i}{i} = \binom{r+k+1}{k}.$$

Proof. Consider the set

$$X = \{x \in \mathbb{N}^+ : x \leq n+1\}.$$

We will count the number of $(r+1)$ -element subsets of X in the following way:

Fix k to be the smallest element in some $Y \subseteq X$, then it suffices to choose r elements from $X - \{x \in \mathbb{N}^+ : x \leq k\}$. Note that the maximum for k is $n - r + 1$, and so there are $n - r + 1$ disjoint cases. Therefore, it is easy to see that the total number of subsets is given by

$$\sum_{i=0}^{n-r} \binom{r+i}{r}.$$

However, we know that this is equivalent to choosing $(r+1)$ elements directly from X , so

$$\sum_{i=0}^{n-r} \binom{r+i}{r} = \binom{n+1}{r+1}.$$

Note that $C_r^{r+i} = C_i^{r+i}$ and $C_{r+1}^{n+1} = C_{n-r}^{n+1}$. Therefore, by setting $k = n - r$, it is obvious by symmetry that

$$\sum_{i=0}^k \binom{r+i}{i} = \binom{r+k+1}{k}.$$

□

1.6.2 Multinomial Coefficients

Definition 1.6.4 ► Multinomial Coefficient

Let $m, n \in \mathbb{N}$, with $m \geq 1$. The **multinomial coefficient**

$$\binom{n}{n_1, n_2, \dots, n_m} = \frac{n!}{\prod_{i=1}^m n_i!}$$

is the number of ways to distribute n distinct objects into m distinct collections such that there are exactly n_i objects in the i -th collection.

Theorem 1.6.5 ► Multinomial Expansion

For $m, n \in \mathbb{N}+$, we have

$$\left(\sum_{i=1}^m x_i \right)^n = \sum_{\sum_{j=1}^m n_j = n} \binom{n}{n_1, n_2, \dots, n_m} \prod_{i=1}^m x_i^{n_i}.$$

Pigeonhole Principle

2.1 Pigeonhole Principle

Theorem 2.1.1 ► Pigeonhole Principle

Let $k, n \in \mathbb{N}^+$. If at least $kn + 1$ distinct objects are distributed into n distinct sets, then there exists a set K such that $|K| \geq k + 1$.

Proof. Let $a_1, a_2, \dots, a_{kn+1}$ be the $kn + 1$ distinct objects to be distributed into the sets K_1, K_2, \dots, K_n . Suppose there is no $i \in \mathbb{N}^+$ such that $|K_i| \geq k + 1$, i.e. $|K_i| \leq k$ for all $i = 1, 2, \dots, n$.

This means that $\sum_{i=1}^n |K_i| \leq kn < kn + 1$, which is a contradiction. \square

Note that Pigeonhole Principle only guarantees existence rather than uniqueness. We may also consider the following generalised version:

Theorem 2.1.2 ► Generalised Pigeonhole Principle

Let $n, k_1, k_2, \dots, k_n \in \mathbb{N}^+$. If at least $\left[\sum_{i=1}^n k_i + (n - 1)\right]$ distinct objects are distributed into n distinct sets S_1, S_2, \dots, S_n , then there exists at least one S_i for $i = 1, 2, \dots, n$ such that $|S_i| \geq k_i$.

Proof. Suppose there is no $i \in \mathbb{N}^+$ such that $|S_i| \geq k_i$, i.e., $|S_i| \leq k_i - 1$ for all $i = 1, 2, \dots, n$.

This means that $\sum_{i=1}^n |S_i| \leq \sum_{i=1}^n (k_i - 1) = \sum_{i=1}^n k_i - n$, which is a contradiction. \square

Note that by setting $k_1 = k_2 = \dots = k_n = 2$ we get from Theorem 2.1.2 to Theorem 2.1.1.

The Pigeonhole Principle is closely linked to the *Ramsey Types of Problems*.

2.2 Ramsey Types of Problems

The Ramsey Types of Problems are motivated by the following question statement:

Example 2.2.1

Let $A_1, A_2, A_3, A_4, A_5, A_6$ be six arbitrary points in a plane. We join these points pairwise with either a **red** line segment or a **blue** line segment. Prove that after each point has been joined with any other point, there exists at least one **blue** triangle or at least one **red** triangle whose vertices are three of the six points.

Proof. Consider the line segments joining A_h with the other five points. Note that this is equivalent to colouring 5 line segments using either red or blue. By Theorem 2.1.1, this means that at least 3 of these line segments have the same colour.

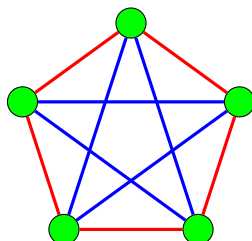
Without loss of generality, suppose the line segments A_hA_i , A_hA_j and A_hA_k are red. Consider the line segments A_iA_j , A_jA_k and A_iA_k . If they are all blue, then we will have a blue triangle. Otherwise, if any one of them is red, then those two vertices together with A_h will form a red triangle. \square

We can extend the above question a bit. Suppose we have 17 arbitrary points X_1, X_2, \dots, X_{17} in a plane and we join these points in the same way as Example 2.2.1 but with three distinct colours: **red**, **blue** and **orange**. Can we still guarantee the existence of at least one triangle whose sides are all of the same colour?

It turns out that we can! Again we can consider all line segments joining X_1 with the other 16 vertices. By Theorem 2.1.1, we must have at least 6 of them having the same colour. Without loss of generality, assume that these 6 line segments are all red.

Now, examine these 6 vertices joined to X_1 with the same-coloured line segments. If any line segment between these 6 vertices is red, then we have a red triangle. Otherwise, the line segments joining these 6 vertices are coloured with only two colours, and by Example 2.2.1 we are guaranteed to have a triangle with all three sides having the same colour as well.

Going back to Example 2.2.1, one might notice that 6 is the smallest possible number of vertices for us to guarantee the existence of such a triangle. In the case of 5 vertices, consider:



This motivates the following definitions:

Definition 2.2.2 ▶ Clique

A **clique** is defined to be a finite set of vertices such that every pair of vertices are adjacent to each other (i.e., there exists a line segment joining any two vertices). A clique with k vertices is called a k -clique.

Definition 2.2.3 ▶ Ramsey Number

Given $p, q \in \mathbb{N}^+$. The **Ramsey Number** $R(p, q)$ is the smallest natural number n such that for any colouring of the edges of an n -clique by 2 colours, there exists either a monochromatic p -clique or a monochromatic q -clique.

Remark. Some obvious results:

- $R(p, q) = R(q, p)$.
- $R(1, q) = 1$.
- $R(2, q) = q$.

These develop into what we now know as *Ramsey Theory* which probes into the question:

How many elements of a particular structure S must there exist to guarantee that a certain property p will always hold for S ?

It turns out that calculating the Ramsey Numbers can be extremely difficult. However, we do have a way to establish an upper bound for each of the Ramsey Numbers.

Theorem 2.2.4 ▶ Recurrence Relation of Ramsey Numbers

For all integers $p, q \geq 2$,

$$R(p, q) \leq R(p-1, q) + R(p, q-1).$$

In particular, if both $R(p-1, q)$ and $R(p, q-1)$ are even, then

$$R(p, q) \leq R(p-1, q) + R(p, q-1) - 1.$$

Proof. Let $n = R(p-1, q) + R(p, q-1)$. To prove that $R(p, q) \leq n$, it suffices to prove that for any n -clique, any colouring of its edges using blue and red will result in either a blue p -clique or a red q clique.

Consider an arbitrary vertex V of any n -clique. Note that V is incident with $(R(p-1, q) + R(p, q-1) - 1)$ vertices and edges. By Theorem 2.1.2 this means that either at least $R(p-1, q)$ edges are blue or at least $R(p, q-1)$ edges are red.

Without loss of generality, consider the $R(p - 1, q)$ -clique whose vertices are joined to V by blue edges. By Definition 2.2.3, this clique contains either a blue $(p - 1)$ -clique or a red p -clique. If it contains a red p -clique, then we are done. If it contains a blue $(p - 1)$ -clique, this clique together with V will form a blue p -clique. Either way, it is guaranteed that

$$R(p, q) \leq n = R(p - 1, q) + R(p, q - 1).$$

Now we proceed to proving the case where $R(p - 1, q)$ and $R(p, q - 1)$ are both even.

Let $m = R(p - 1, q) + R(p, q - 1) - 1$. Fix an arbitrary vertex W , then it suffices to prove that among the $(m - 1)$ edges incident to W , either at least $R(p - 1, q)$ are blue or at least $R(p, q - 1)$ are red.

Suppose on contrary that there are exactly $(R(p - 1, q) - 1)$ blues edges and exactly $(R(p, q - 1) - 1)$ red edges for all vertices in the m -clique. This means that the total number of blue edges in the m -clique is

$$N = \frac{m(R(p - 1, q) - 1)}{2}.$$

However, m and $R(p - 1, q) - 1$ are both odd, this means that N is not an integer, which is impossible. □

The above recurrence allows us to prove the following theorem:

Theorem 2.2.5 ► Ramsey's Theorem

For all $p, q \in \mathbb{N}^+$, the Ramsey Number $R(p, q)$ always exists.

Principle of Inclusion and Exclusion

3.1 Principle of Inclusion and Exclusion (PIE)

Theorem 3.1.1 ► Principle of Inclusion and Exclusion (a special case)

Let A and B be two finite sets, then

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

In particular, if A and B are disjoint, then $|A \cup B| = |A| + |B|$.

Theorem 3.1.2 ► Principle of Inclusion and Exclusion (general)

Let E_1, E_2, \dots, E_n be a sequence of finite sets. In general, we have

$$\left| \bigcup_{i=1}^n E_i \right| = \sum_{j=1}^n \left[(-1)^{j+1} \left(\sum_{1 \leq k_1 \leq k_2 \leq \dots \leq k_j \leq n} \left| \bigcap_{h=1}^j E_{k_h} \right| \right) \right].$$

Proof. Define a function $f_S : S \rightarrow \{0, 1\}$ by

$$f_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases}.$$

Let $E = \bigcup_{i=1}^n E_i$. Consider

$$\prod_{i=1}^n (f_E(x) - f_{E_i}(x)). \quad (*)$$

For any x , if $x \in E$, then $x \in E_k$ for some $k \in \{x \in \mathbb{N} : x \leq n\}$, which means that $f_E(x) - f_{E_k}(x) = 0$; if $x \notin E$, then $f_E(x) = f_{E_i}(x) = 0$ for all $i \in \{x \in \mathbb{N} : x \leq n\}$. In either case, $(*)$ is identically 0.

Note that for any sets A, B , $f_{A \cap B}(x) = f_A(x)f_B(x)$. □

3.2 General Principle of Inclusion and Exclusion (GPIE)

Theorem 3.2.1 ► General Principle of Inclusion and Exclusion

Let S be an n -element set and let $\{P_1, P_2, \dots, P_q\}$ be a set of properties. Suppose that $E(m)$ is the number of elements in S that satisfy exactly m of the q properties and that $\omega(P_{i_1}P_{i_2}\dots P_{i_m})$ is the number of elements in S that satisfy exactly $P_{i_1}, P_{i_2}, \dots, P_{i_m}$, such that

$$\omega(m) = \sum \omega(P_{i_1}P_{i_2}\dots P_{i_m}),$$

then

$$E(m) = \sum_{k=m}^q (-1)^{k-m} \binom{k}{m} \omega(k).$$

3.3 Stirling Numbers of the Second Kind

Definition 3.3.1 ► Stirling Number of the Second Kind

The **Stirling Number of the Second Kind**, denoted as $S(r, n)$, is the number of ways to distribute r distinct objects into n indential collections such that no collection is empty.

Remark. Some trivial results:

- $S(0, 0) = 1$.
- $S(r, 0) = S(0, n) = 0$ for all $r, n \in \mathbb{N}^+$.
- $S(r, n) > 0$ for all $r, n \in \mathbb{N}$ with $r \geq n \geq 1$.
- $S(r, n) = 0$ if $n > r \geq 1$.
- $S(r, 1) = S(r, r) = 1$ for all $r \in \mathbb{N}^+$.

Theorem 3.3.2 ► Recurrence Relation of $S(r, n)$

For all $r, n \in \mathbb{N}$ with $r \geq n$,

$$S(r, n) = S(r-1, n-1) + nS(r-1, n).$$

Definition 3.3.3 ► Partition

Let A be a set, an n -partition of A is a set $S \subset \mathcal{P}(A)$ such that for all $S_1, S_2 \in S$ with $S_1 \neq S_2$, $S_1 \cap S_2 = \emptyset$ and $\bigcup S = A$.

Note that the number of all partitions of A with $|A| = n$ is just

$$\sum_{i=1}^n S(n, i).$$

Theorem 3.3.4 ► Number of Surjective Mappings

Let $F(r, n)$ where $r, n \in \mathbb{N}^+$ denote the number of surjective mappings

$$f : [r] - \{0\} \rightarrow [n] - \{0\},$$

then

$$F(r, n) = \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^r.$$

Some corollaries:

$$\begin{aligned} \frac{1}{n!} \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^r &= S(r, n) \\ \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^r &= 0 \quad \text{if } n > r \geq 1 \end{aligned}$$

3.4 Derangement

Definition 3.4.1 ► Derangement

Let $N_n = [n] - \{0\}$. A **derangement** of N_n , denoted by D_n , is the permutation $a_1 a_2 \cdots a_n$ of N_n such that $a_i \neq i$ for all $i = 1, 2, \dots, n$.

In general, we also say that an r -permutation of N_n has a **fixed point** at i if $a_i = i$. The number of r -permutations of N_n with k fixed points is $D(n, r, k)$.

Theorem 3.4.2 ► A Formula for $D(n, r, k)$

For $n \geq r \geq k \geq 0$ and $r \geq 1$,

$$D(n, r, k) = \frac{C_k^r}{(n-r)!} \sum_{i=0}^{r-k} (-1)^i \binom{r-k}{i} (n-k-i)!.$$

Note that the number of derangement

$$D_n = D(n, n, 0) = n! \sum_{i=0}^n \frac{(-1)^i}{i!}.$$

And so

$$\lim_{n \rightarrow \infty} P(\text{Having a derangement}) = \lim_{n \rightarrow \infty} \frac{D_n}{n!} = e^{-1}.$$