Contents

1 How to Count

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 agregate 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 $A_1, A_2, \cdots, A_k$ be k finite sets which are pairwise disjoint, i.e. for all i, j such that $1 \leq i, j \leq k$, $A_i \cap A_j = \varnothing$ 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:

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

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

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 $E_1, E_2, \cdots, E_k$ 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 $(e_1, e_2, \cdots, e_k)$, 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 $A_1, A_2, \cdots, A_k$ 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 $\boldsymbol{a} \in \prod_{i=1}^{n} A_i$, $(\boldsymbol{a}, a_i) \neq (\boldsymbol{a}, a_j)$ 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 $k \in \mathbb{N}^+$. □

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.

1.2 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

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

1.2.1 Permutations with Idential Objects

Theorem 1.2.2 ▶ Generalised Formula for Permutations

Let $k \in \mathbb{N}^+$ and let $A_1, A_2, \cdots, A_k$ be k distinct objects, where A_i occurs $n_i > 0$ times for $i = 1, 2, \cdots, k$, then the number of permutations for these k objects are given by

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

1.3 Combinations

Definition 1.3.1 ▶ Combinations

Let A be a finite set such that $|A| = n$, an r-combination of A is a way to choose r elements from A regardless of the order of selection, denoted as C_r^n and 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. $C_r^n = C_{n-r}^n$.

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

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

1.4 Binomial and Multinomial Coefficients

Consider the expansion of $(x+y)^n$ where $n \in \mathbb{N}$. Note that this expansion is a linear combination of terms in the form of $x^k y^{n-k}$ where $k = 0, 1, 2, \cdots, n$.

Thus, fix any k, to determine how many copies of $x^k y^{n-k}$ there are, it suffices to compute C_k^n. Therefore, in the expanded form of $(x+y)^n$, the coefficient is exactly C_r^n.

Theorem 1.4.1 ▸ Binomial Expansion

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

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

We can extend the idea of binomial coefficients onto multinomial expansions, i.e. expressions in the form of $\left(\sum_{i=1}^{r} x_i\right)^n$.

Note that the binomial coefficient C_r^n is essentially equivalent to dividing n distinct elements into two groups with r and $(n-r)$ members respectively. Now we consider dividing n distinct elements into r groups with $n_1, n_2, \cdots, n_r$ members respectively for each group.

Notice that we can simply permute the n distinct elements and assign them sequentially into the r groups, i.e. the first n_1 elements will go into the first group and so on.

Since the order of elements within each group does not matter, we need to remove repeated selections by dividing by $\prod_{i=1}^{r}(n_i!)$. So we have the following definition:

Definition 1.4.2 ▸ Multinomial Coefficients

The **multinomial coefficient** is defined by

$$\binom{n}{n_1, n_2, \cdots, n_k} = \frac{n!}{\prod_{i=1}^{k}(n_i!)}$$

Theorem 1.4.3 ▸ Multinomial Expansion

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

$$\left(\sum_{i=1}^{r} x_i\right)^n = \sum_{\substack{n_1, n_2, \cdots, n_r \in \mathbb{N} \\ \sum_{j=1}^{r} n_j = n}} \left[\binom{n}{n_1, n_2, \cdots, n_r} \prod_{i=1}^{r} x_i^{n_i} \right]$$

2

Axioms of Probability

2.1 Sample Space and Events

Definition 2.1.1 ▸ Sample Space

Consider an experiment whose outcome is **not** predictable, then the set of all possible outcomes of the experiment is called the **sample space** of the experiment, denoted by S.

Remark. Note that $S \neq \varnothing$.

Definition 2.1.2 ▸ Events

Let S be a sample space, a set $E \subseteq S$ is known as an **event**.

Remark. S itself is known as the **sure event** and $\varnothing$ is known as the **null event**.

Note that since sample spaces and events are sets, we can apply operations onto events precisely in the same way for sets.

By convention, the intersection of two events E and F is preferably written as EF. Two events which are disjoint are called *mutually exclusive*.

Definition 2.1.3 ▸ Probability

Let E be any event of an experiment and let $n(E)$ be the number of occurrences of E in the first n repetitions of the experiment, then the **probability** of E is

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n},$$

if the limit exists.

Remark. Note that the following properties are satisfied:

1. $0 \leq P(E) \leq 1$.
2. Let S be a sample space, then $P(S) = 1$.
3. If E and F are mutually exclusive, then $P(E \cup F) = P(E) + P(F)$.

With induction, one can easily show that if $E_1, E_2, \cdots$ to be any sequence of events in a sample space S, then

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Theorem 2.1.4 ▶ The Null Event

Consider the null event $\varnothing$, we have

$$P(\varnothing) = 0.$$

Proof. Let S be a sample space and let $E_1, E_2, \cdots$ be a countably infinite sequence of events such that $E_i = \varnothing$ for all $i \in \mathbb{N}^+$. We can write

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Note that the countable union of empty sets is empty, so the above is equivalent to

$$P\left(\bigcup_{i=1}^{\infty} \varnothing\right) = P(\varnothing) = \sum_{i=1}^{\infty} P(\varnothing).$$

This means that $P(\varnothing)$ equals the sum of a countably infinite sequence of itself, so

$$P(\varnothing) = 0.$$

□

Theorem 2.1.5 ▶ Monotonity of Probability

Let E and F be events such that $E \subseteq F$, then

$$P(F) \geq P(E).$$

Proof. Note that E and $F - E$ are mutually exclusive, so

$$P(F) = P(E \cup (F - E)) = P(E) + P(F - E).$$

Note that $P(F - E) \geq 0$, so $P(E) + P(F - E) \geq P(E)$, which means

$$P(F) \geq P(E).$$

□

Theorem 2.1.6 ▸ Inclusion-Exclusion Principle

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

Theorem 2.1.7 ▸ Boole's Inequality

Let $E_1, E_2, \cdots, E_n, \cdots$ be a countable sequence of events, then

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} P(E_i).$$

In particular, equality is achieved if and only if the E_i's are mutually exclusive.

Theorem 2.1.8 ▸ Probability in a Finite Sample Space

et S be a sample space which is finite and let $E \subseteq S$ be an event, then

$$P(E) = \frac{|E|}{|S|}.$$