

# Combinatorics

## Contents

# 1 Counting and Enumeration

## 1.1 Strings and sets

An  $X$ -string of length  $n$  is written  $x_1x_2x_3\cdots x_n$  for  $x_i \in X$ , where  $X$  is an alphabet its elements are characters.

$$s : [n] \mapsto X, \quad [n] = \{1, 2, 3 \cdots, n\}$$

## 1.2 Exponential

An enumeration such that:

1. Order of items matters.
2. Counts include duplication or removals of items.

If  $X$  is a finite set, the number of all  $X$ -strings of length  $n$  will be  $|X|^n$ .

## 1.3 Permutations

If  $X$  is an  $n$  element set, the number of  $X$ -strings of length  $m$  is,

$$P(n, m) = \frac{n!}{(n - m)!}$$

1. Order of items matters.
2. Counts do not include duplication or removals of items.
3. Collection of counts could be stored in *arrays*.

## 1.4 Combinations (Binomial Coefficient)

The number of combinations of  $n$  objects taken  $k$  at a time. Given  $n, k \in \mathbb{Z}, 0 \leq k \leq n$  then the binomial coefficient is,

$$C(n, k) = \binom{n}{k} = \frac{P(n, k)}{k!} = \frac{n!}{k!(n - k)!} = \binom{n}{n - k}$$

1. Order of items doesn't matter.
2. Counts do not include duplication or removals of items.
3. Collection of counts could be stored in *sets*.

**Identity 1.1.**

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

## 1.5 n-th Partial Sums

**Identity 1.2.**

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

**Identity 1.3.**

$$1 + 3 + 5 + \cdots + 2n - 1 = n^2$$

## 1.6 Pascal's Triangle

A triangular array made up of the binomial coefficients in which the entry in the  $n$ th row and  $k$ th column is  $\binom{n}{k}$ .

**Identity 1.4.**

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

**Identity 1.5.**

$$\sum_{i=0}^n \binom{n}{i} = 2^n$$

## 1.7 Binomial Theorem

Let  $x, y \in \mathbb{R}, x + y \neq 0$ . Then  $\forall n \in \mathbb{Z}^+$ ,

$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^{n-i} y^i$$

## 1.8 Multinomial Coefficients

The number of combinations in which  $k$  elements are chosen from subsets of  $X$  with  $|X| = n$  will be,

$$\binom{n}{k_1, k_2, \dots, k_n} = \frac{n!}{k_1! k_2! \cdots k_n!}$$

## 1.9 Multinomial Theorem

Let  $x_1, \dots, x_n \in \mathbb{R}^+$ . Then  $\forall n \in \mathbb{N}^+$ ,

$$(x_1 + \cdots + x_n)^n = \sum_{k_1 + \dots + k_r = n} \binom{n}{k_1, \dots, k_r} x_1^{k_1} x_2^{k_2} \cdots x_r^{k_r}$$

## 1.10 Lattice Paths

A sequence of ordered pairs  $(m_1, n_1), (m_2, n_2), \dots, (m_t, n_t)$  such that either:

1.  $m_{i+1} = m_i + 1$  and  $n_{i+1} = n_i$
2.  $m_{i+1} = m_i$  and  $n_{i+1} = n_i + 1$ .

The construction of lattice paths forms a bijection with  $X$ -strings where  $X = \{H, V\}$  with  $H, V$  encoding horizontal or vertical moves. The number of lattice paths from  $(m_1, n_1)$  to  $(m_2, n_2)$  is,

$$\binom{m_2 - m_1 + n_2 - n_1}{m_2 - m_1}.$$

## 1.11 Catalan Number

The number of lattice paths from  $(0, 0)$  to  $(n, n)$  that do not go above the diagonal line  $y = x$ , also known as Dyck paths, are equal to the  $n$ th Catalan number,

$$C(n) = \frac{1}{n+1} \binom{2n}{n}$$

# 2 Number Theory

## 2.1 Well-Ordered Property

$\forall A \subset \mathbb{N}^+, A \neq \emptyset, A$  must have a minimum element. This statement is equivalent to mathematical induction.

## 2.2 Principle of Mathematical Induction

Let  $S_n$  be an open statement involving a positive integer  $n$ . If  $S_1$  is true, and if for each positive integer  $k$ , assuming that the statement  $S_k$  is true implies that the statement  $S_{k+1}$  is true, then  $S_n$  is true for every positive integer  $n$ .

## 2.3 Pigeonhole Principle

Given  $m$  items and  $n$  containers, there exists a container with at least  $\lceil \frac{m}{n} \rceil$  items.

Expressed in terms of functions and their domains and ranges: if  $f : X \mapsto Y$  and  $|X| > |Y|$  then there exists a  $y \in Y$  and  $x, x' \in X, x \neq x'$  so that  $f(x) = f(x') = y$

## 2.4 Erdos–Szekeres theorem

If  $m, n \in \mathbb{Z}^+$ , then any sequence of  $mn + 1$  distinct real numbers either has an increasing sub-sequence of  $m + 1$  items or a decreasing sub-sequence of  $n + 1$  terms.

## 2.5 Fibonacci Sequence

$$\phi_n = \phi_{n-1} + \phi_{n-2}; \phi_1 = 1, \phi_2 = 2$$

## 2.6 Euclidean Algorithm

The Euclidean algorithm is an efficient method for computing the greatest common divisor (GCD) of two integers.

Given  $m = q_i n_i + r_i$  where  $m_i, n_i \in \mathbb{Z}^+$  with  $0 \leq r_i < n_i$  we see that  $r_i$  represents the remainder and  $q_i$  represents the quotient. We use the fact that  $\gcd(m, n_i) = \gcd(n_i, r_i)$  to iteratively calculate the gcd on the decomposition of the  $n_i$  term and the remainder  $r_i$ .

## 3 Inclusion-Exclusion Principle

The inclusion-exclusion principle is a counting technique that generalizes the method of obtaining the number of elements in the union of two finite sets without double counting elements,

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

The principle is particularly useful in the case of three sets, i.e. for the sets  $A$ ,  $B$  and  $C$ , the count of the union is given by,

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

### 3.1 The Euler Totient Function

For a positive integer  $n \geq 2$ , Euler's Totient function is defined by

$$\phi(n) = |\{m \in \mathbb{N} : m \leq n, \gcd(m, n) = 1\}|.$$

## 4 Graphs Theory

Let  $G = (V, E)$  be a graph with  $V$  vertices and  $E$  edges.

### 4.1 Subgraph

$H = (W, F)$  is a subgraph of  $G$  if  $W \subseteq V$  and  $F \subseteq E$ . That is, a subgraph does not contain any additional vertices or edges but may contain fewer.

### 4.2 Induced Subgraph

$H = (W, F)$  is an induced subgraph if  $W \subseteq V$  and  $F = \{xy \in E : xy \in W\}$ . That is, an induced subgraph is formed from a subset of the vertices of the graph with all of the edges connecting pairs of vertices in the subset being kept from the original graph.

### 4.3 Spanning

$H = (W, F)$  is spanning when  $W = V$ .

### 4.4 Isomorphic

$G$  is isomorphic to  $H$ , denoted by  $G \cong H$ , if there exists an edge-preserving bijective mapping,  $f : V \mapsto W$ .

### 4.5 Homeomorphic

The subdivision of an edge  $e$  with endpoints  $\{u, v\}$  yields a graph containing one new vertex  $w$  with an edge set replacing  $e$  by two new edges,  $\{u, w\}$  and  $\{w, v\}$ .

Two graphs  $G$  and  $H$  are homeomorphic if there is a graph isomorphism from some subdivision of  $G$  to some subdivision of  $H$ .

### 4.6 Contains

$G$  contains  $H$  if there exists a subgraph of  $G$  that is isomorphic to  $H$ .

### 4.7 Complete Graphs

A complete graph is a graph in which each pair of graph vertices is connected by an edge, i.e.  $x, y \in E$  for all distinct  $x, y \in V$ .

The complete graph with  $n$  graph vertices is denoted as  $K_n$ .  $K_n$  has  $\binom{n}{2}$  edges.

### 4.8 Independent Graphs

An independent graph, denoted  $I_n$ , is a graph with  $xy \notin E$  for all distinct pairs of vertices  $x, y \in V$ .

### 4.9 Paths and Cycles

$P_n = \{x_1, x_2, \dots, x_n\}$  is a path of length  $n - 1$ .

$C_n = \{x_1, x_2, \dots, x_n, x_1\}$  is a cycle with length  $n$ .

### 4.10 Connected

A graph  $G$  is connected if every vertex can be reached by any other vertex in the graph, i.e. a path exists connecting  $v_1$  to  $v_2$  for all  $v_1, v_2 \in V$ .

### 4.11 Component

A maximal connected subgraph of a disconnected graph.

### 4.12 Acyclic / Forrests

A graph that does not contain any cycles.

The number of components in a forest is  $|V| - |E|$

### 4.13 Tree

A connected acyclic graph.

### 4.14 Spanning Tree

If  $G$  is a connected graph and  $H$  is a spanning subgraph, then we can call  $H$  a spanning tree.

### 4.15 Degree Sum Formula

The degree sum formula states that, given a graph  $G = (V, E)$ ,

**Identity 4.1.**

$$\sum_{v \in V} \deg(v) = 2|E|.$$

The formula implies that in any undirected graph, the number of vertices with odd degree is even, which is known as the Handshaking lemma. The name comes from a popular problem: in any group of people the number of people who have shaken hands with an odd number of other people from the group must be even.

### 4.16 Eulerian Cycles and Graphs

A Eulerian cycle is a cycle that contains every edge in a given graph exactly once. A Eulerian graph is a graph that contains a Eulerian cycle.

**Theorem 4.2.** *A graph admits a Eulerian cycle if and only if all vertices have even degree*

### 4.17 Hamiltonian Cycles and Graphs

A Hamiltonian cycle is a cycle that traverses each vertex exactly once. A Hamiltonian graph is a graph that contains a Hamiltonian cycle.

**Theorem 4.3.** *If a graph has  $n$  vertices and  $\forall v \in V, \deg(v) \geq \lceil \frac{n}{2} \rceil$ , then it must be Hamiltonian.*

**Theorem 4.4.** *(Ore's Theorem) If for all non-adjacent vertices  $v, w \in V$ ,  $\deg(v) + \deg(w) \geq n$ , then  $G$  must be Hamiltonian.*

## 4.18 Bipartite Graphs

A graph is bipartite if there exists a partition of  $V$  so that the resulting induced subgraphs are independent.

A graph is complete bipartite, denoted  $K_{m,n}$ , if it is bipartite with  $V = V_1 \cup V_2$ ,  $|V_1| = m$ ,  $|V_2| = n$  and has edges  $xy$  for all  $x \in V_1, y \in V_2$ .

## 4.19 Graph Coloring

A graph can be colored by assigning a color or label to each of its vertices. A proper coloring exists when no adjacent vertices have the same color.

## 4.20 Chromatic Number

The chromatic number of a graph, denoted  $\chi(G)$ , is the smallest number of  $k$  colors so that  $G$  is  $k$ -colorable.

**Proposition 4.4.1.** *Recall  $C_n$  is a cycle of  $n$  vertices. Then,  $\chi(C_n) = \begin{cases} 2, & \text{if } n \text{ is even} \\ 3, & \text{if } n \text{ is odd} \end{cases}$*

## 4.21 Clique Number

The clique number of a graph, denoted  $\omega(G)$ , is the maximum number  $n$  so that  $G$  has a subgraph isomorphic to the complete graph  $K_n$ .

**Proposition 4.4.2.** *For any graph  $G$ ,  $\chi(G) \geq \omega(G)$ .*

If  $\chi(G) = \omega(G)$ , then  $G$  is called perfect or reduced.

**Proposition 4.4.3.** *For any  $t \geq 3$ , there exists a graph  $G$  so that  $\omega(G) = 2$  and  $\chi(G) = t$ .*

This means that graphs can have an arbitrarily large chromatic number. If the original graph is triangle-free, a Mycielskian graph may be constructed to increase the graph's chromatic number.

## 4.22 Planar Graph

A graph is planar if it has a drawing or embedding so that its edges only intersect one another at one of the graph's vertices.

## 4.23 (Euler's formula)

**Theorem 4.5.** *Given a finite, connected, planar graph where  $v$  is the number of vertices,  $e$  is the number of edges and  $f$  is the number of faces,*

$$v - e + f = 2.$$

**Theorem 4.6.** *A planar graph on  $n$  vertices has at most  $3n - 6$  edges, for  $n \geq 3$ .*



## 4.24 Kuratowski's Theorem

**Theorem 4.7.** *A graph is planar if and only if it does not contain a subgraph homeomorphic to  $K_5$  or  $K_{3,3}$ .*

## 4.25 Four Color Theorem

**Theorem 4.8.** *Every planar graph has chromatic number of at most 4.*

## 4.26 Petersen Graph

A Petersen graph has the following properties:  $|V| = 10$ ,  $|E| = 15$ , there are no cycles of length 3 or 4, and  $\deg(v) = 3, \forall v \in V$ . A Petersen graph is never Hamiltonian.

## 4.27 Dense Graph

A dense graph is a graph in which the number of edges is close to the maximal number of edges. For undirected simple graphs, the graph density is:

$$D = \frac{|E|}{\binom{|V|}{2}} = \frac{2|E|}{|V|(|V| - 1)}$$

## 4.28 Sparse Graph

Conversely, a sparse graph is a graph in which the number of edges is close to the minimal number of edges. A sparse graph can be a disconnected graph.

# 5 Generating Functions

Given a sequence  $\sigma = \{a_n : n \geq 0\}$  of real numbers, we define the ordinary generating function  $F(x)$  of the sequence  $\sigma$  as,

$$F(x) = \sum_{n=0}^{\infty} a_n x^n.$$

This differs from a formal power series in that we are not concerned with the series convergence or divergence and are generally not interested in substituting a value of  $x$  and obtaining a specific value for  $F(x)$ . Instead we may use  $F(x)$  to encode and generate the sequence  $\sigma$ .

**Identity 5.1.** *The geometric series:*

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots$$

## 5.1 Newton's Binomial Theorem

**Theorem 5.2.** For all real  $p$  with  $p \neq 0$ ,

$$(1+x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n.$$

## 5.2 Partitions of an Integer

A partition  $P$  of an integer is a collection of (not necessarily distinct) positive integers such that,

$$\sum_{i \in P} i = n.$$

We can find the number of partitions of a given integer using a generator function

$$P(x) = \left( \sum_{m=0}^{\infty} x^m \right) \left( \sum_{m=0}^{\infty} x^{2m} \right) \left( \sum_{m=0}^{\infty} x^{3m} \right) \cdots \left( \sum_{m=0}^{\infty} x^{km} \right) \cdots = \prod_{m=1}^{\infty} \frac{1}{1-x^m}.$$

**Theorem 5.3.** For each  $n \geq 1$ , the number of partitions of  $n$  into distinct parts is equal to the number of partitions of  $n$  into odd parts.

## 5.3 Exponential Generating Function

The exponential generating function for the sequence  $\{a_n : n \geq 0\}$  is  $\sum_n \frac{a_n x^n}{n!}$ .

**Identity 5.4.** The exponential function formula:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

## 6 Recurrence Equations

A recurrence tells us how to compute the  $n$ th term of a sequence given some number of previous values. A recursion is linear when it has the following form,

$$c_0 a_{n+k} + c_1 a_{n+k-1} + \cdots + c_k a_n = g(n),$$

where  $k \geq 1$  is an integer,  $c_0, c_1, \dots, c_k$  are constants with  $c_0, c_k \neq 0$ , and  $g : \mathbb{Z} \rightarrow \mathbb{R}$  is a function.

A linear equation is homogeneous if the function  $g(n)$  is the zero function is non-homogeneous otherwise. The process of solving recurrence equations is similar to solving ordinary differential equations as both functions express future states in relation to differences to their previous states.

## 6.1 Advancement Operator

The advancement operator is a linear map that shifts or advances a function's input forward by one, i.e.  $A : \mathbb{R} \mapsto \mathbb{R}$ , where  $(Af)(n) = f(n+1)$ .

It is commutative and has the properties that  $A^k = \underbrace{A \circ \cdots \circ A}_{k \text{ times}}$  and  $(A^k f)(n) = f(n+k)$ .

## 6.2 Finding Closed-Form Expressions

**Theorem 6.1.** *The set of all solutions to the homogeneous linear equation of degree  $k$  is a  $k$ -dimensional subspace of  $V$ .*

This means that we may solve a linear recurrence stated in terms of advancement operations by finding a basis for the vector space of solutions.

**Lemma 6.2.** *Let  $r \neq 0$ , and let  $f$  be a solution to the operator equation  $(A - r)f = 0$ . If  $c = f(0)$ , then  $f(n) = cr^n$  for every  $n \in \mathbb{Z}$ .*

This means that the roots of the advancement operator equation will give us solutions to its corresponding recurrence of the form  $f(n) = cr^n$ .

**Lemma 6.3.** *Consider a non-homogeneous operator equation and let  $W$  be the subspace of  $V$  consisting of all solutions to its corresponding homogeneous equation.*

*If  $f_0$  is a solution to the homogeneous equation, then every solution  $f$  to has the form  $f = f_0 + f_1$  where  $f_1 \in W$  is a particular solution.*

If the polynomial in the advancement operator has distinct roots we may use the following theorem,

**Theorem 6.4.** *Consider the following advancement operator equation  $p(A)f = (A - r_1)(A - r_2) \cdots (A - r_k)f = 0$ . with  $r_1, r_2, \dots, r_k$  distinct non-zero constants. Then every solution has the form  $f(n) = c_1 r_1^n + c_2 r_2^n + c_3 r_3^n + \cdots + c_k r_k^n$ .*

If there are repeated roots we use the following theorem,

**Theorem 6.5.** *Let  $k \geq 1$  and consider the equation  $(A - r)^k f = 0$ . Then the general solution has the form  $f(n) = c_1 r^n + c_2 n r^n + c_3 n^2 r^n + c_4 n^3 r^n + \cdots + c_k n^{k-1} r^n$ .*

## 6.3 Using Generating Functions

Alternatively, we may use generating function to solve recursive equations. This approach does not require determining an appropriate form for the particular solution but instead the method of generating functions often requires that the resulting generating function be expanded using partial fractions.

Generating functions can be used in solving certain nonlinear recurrence equation, in particular, we can use it to determine the number of rooted, unlabeled, binary, ordered trees. A tree is rooted if we have designated a special vertex called its root. An unlabeled tree is one in which we do not make distinctions based upon names given to the vertices. A binary tree is one in which each vertex has 0 or 2 children, and an ordered tree is one in which the children of a vertex have some ordering (first, second, third, etc.).

**Theorem 6.6.** *The generating function for the number  $c_n$  of rooted, unlabeled, binary, ordered trees with  $n$  leaves is*

$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2} = \sum_{n=1}^{\infty} \frac{1}{n} \binom{2n-2}{n-1} x^n.$$

Notice that  $c_n$  is a Catalan number, which also counts the number of lattice paths that do not cross the diagonal line  $y = x$ .

## References

- [1] Keller, M.T. and Trotter, W.T., Applied Combinatorics, Open Textbook Library, ISBN9781534878655. <http://www.rellek.net/book/app-comb.html>