Mathematical Background

Binomial coefficients

- $[n] \stackrel{\text{def}}{=} \{1, 2, \ldots, n\}$.
- The binomial coefficient $\binom{n}{k}$ equals the number of subsets of $[n]$ that have size k.
- $$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k!}$$
- $$\frac{(n-k)^k}{k!} < \binom{n}{k} \leq \frac{n^k}{k!} < \left(\frac{ne}{k}\right)^k$$
- $$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$
- *Binomial expansion*:
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Inequalities

- For all real x, we have $1 + x \leq e^x$. For small x, we have $e^x \approx 1 + x$.
- For $k \geq 1$ and $x \geq -1$, we have $(1+x)^k \geq 1+kx$. For small x, we have $(1+x)^k \approx 1+kx$.
- The nth harmonic number is $H_n \stackrel{\text{def}}{=} 1 + 1/2 + \cdots + 1/n$. Then $H_n \approx \ln n$; specifically,
$$\ln(n+1) \leq H_n \leq 1 + \ln n.$$
- *Stirling's approximation*: $n! \approx \sqrt{2\pi n}(n/e)^n$. Specifically,
$$(n/e)^n < \sqrt{2\pi n}(n/e)^n < n! < 3\sqrt{n}(n/e)^n.$$
- *Convexity*: A function $f : \mathbb{R} \to \mathbb{R}$ is *convex* if for any real x, y and any $\lambda \in [0, 1]$, we have
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$
If f is twice differentiable, then f is convex iff $f''(x) \geq 0$ for all x.
- *Jensen's Inequality*: For a convex function f, we have
$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

Probability

- Probability and events:

  1. A *probability distribution* on a finite set S is an assignment of probabilities $\Pr[x]$ to each element $x \in S$, where $\sum_{x \in S} \Pr[x] = 1$. The *uniform distribution* is the probability distribution where $\Pr[x] = 1/|S|$ for all $x \in S$.
  2. An *event* T is a subset of S. We have $\Pr[T] = \sum_{x \in T} \Pr[x]$, but often this probability can be computed more directly.
  3. For any events A, B,
  $$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B].$$
  4. *Union bound*: for any events $A_1, A_2, \ldots A_n$,
  $$\Pr[A_1 \cup A_2 \cup \ldots \cup A_n] \leq \Pr[A_1] + \Pr[A_2] + \ldots + \Pr[A_n].$$
  5. For *independent* events $A_1, A_2, \ldots A_n$,
  $$\Pr[A_1 \cap A_2 \cap \ldots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2] \cdot \ldots \cdot \Pr[A_n].$$

- Conditional probability:

  1. The *conditional probability* of A given B, denoted $\Pr[A|B]$, is the probability that A occurs given that B occurs. It satisfies
  $$\Pr[A|B] = \Pr[A \cap B] / \Pr[B].$$
  2. *Bayes' Law*:
  $$\Pr[A|B] = \frac{\Pr[A]\Pr[B|A]}{\Pr[B]}.$$

- Random variables:

  1. A *random variable* is a function on a probability space.
  2. Random variables $X_1, X_2, \ldots, X_n$ are *independent* if and only if for all $x_1, \ldots, x_n$, we have
  $$\Pr[(X_1 = x_1) \wedge (X_2 = x_2) \wedge \ldots \wedge (X_n = x_n)] = \prod_{i=1}^{n} \Pr[X_i = x_i].$$
  3. If $X_1, \ldots, X_n \in \{0, 1\}$ are independent, with $\Pr[X_i = 1] = p$, then
  $$\Pr\left[\sum_{i=1}^{n} X_i = k\right] = \binom{n}{k} p^k (1-p)^{n-k}.$$

- Expectation:

  1. The *expectation* of a random variable X with range S is

  $$\mathbb{E}[X] \stackrel{\text{def}}{=} \sum_{x \in S} x \cdot \Pr[X = x].$$

  2. For $X \in \mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$, we have

  $$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X \geq i].$$

  3. Expectation is linear: for constants a, b and random variables X, Y we have

  $$\mathbb{E}[aX + bY] = a\,\mathbb{E}[X] + b\,\mathbb{E}[Y].$$

  4. Expectation is multiplicative *for independent random variables.* That is, for independent X, Y, we have

  $$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y].$$

- Variation distance

  The variation distance, or statistical distance, between probability distributions P and Q defined on the same space S is

  $$\|P - Q\| \stackrel{\text{def}}{=} \max_{T \subseteq S} |P(T) - Q(T)| = \frac{1}{2} \sum_{s \in S} |P(s) - Q(s)|.$$

Set Theory

- A *k-set* is a set of size k.
- $[n] = \{1, 2, \ldots, n\}$.
- $\binom{[n]}{k}$ is the set of k-subsets of $[n]$.
- The *Hamming distance* between two strings a and b of the same length is

  $$d(a, b) = |\{i : a_i \neq b_i\}|.$$

  Think of: The number of different characters between the two strings.

- The *incidence vector* of a set $S \subseteq U$ (universe U) is a vector v, often denoted 1_S, with entries labeled with the members of U: $v_u = 1 \Leftrightarrow u \in S$, otherwise $v_u = 0$.

  Think of: The idea is similar to the adjacency matrix idea. Just put a 1 in all the columns labeled with elements of S.

- The *symmetric set difference* between sets S and T is

$$S \triangle T = (S \cup T) - (S \cap T)$$

  Think of: The elements that are in S or T, but not both. Kind of like set XOR.

Graph Theory

- A *graph* $G = (V, E)$ contains a vertex set V and edge set E. Vertices are also called nodes. In an *undirected graph*, each edge is a 2-subset of V. In a *directed graph*, each edge is an ordered pair of elements of V.

  Common Notation: $|V| = n$, $|E| = m$. i.e., n vertices, m edges.

- A graph $G' = (V', E')$ is a *subgraph* of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. G' is a *spanning subgraph* if $V' = V$.

- The *degree* of a vertex is the number of edges incident to the vertex.

- A graph is *d-regular* if all of its vertices have degree d.

- A graph $G = (V, E)$ is *bipartite* if its vertices can be partitioned into A and B, i.e., $V = A \cup B$ with $A \cap B = \emptyset$, such that $\{a, b\} \in E \implies a \in A, b \in B \vee a \in B, b \in A$.

  Think of: A and B are each on a different side of a river and all of the edges in E are bridges over the river.

- An undirected graph is *connected* if there is a path between every two nodes.

- The *adjacency matrix* of a graph $G = (V, E)$ is a square matrix $A = (a_{i,j})$ where the rows and columns are labeled with V, and $a_{i,j} = 1 \Leftrightarrow (i, j) \in E$ else $a_{i,j} = 0$.

  Think of: In the row for vertex a, if there is an edge between a and b, put a 1 in b's column.

- The *distance* $d(x, y)$ between $x, y \in V$ is the length of the shortest path between x and y.

- The *diameter* of a graph $G = (V, E)$ is $\max_{x,y \in V} d(x, y)$, the largest distance between two nodes.

- A *cycle* is a sequence of vertices $v_1, \dots, v_k$ s.t. $\{v_i, v_{i+1}\} \in E$ for $i = 1, \dots, k-1$ and $\{v_k, v_1\} \in E$.

  Think of: A "circle" of vertices.

- The *girth* of a graph is the length of its smallest cycle.

- A *tree* is a connected graph without cycles.

  Note: Equivalently: A connected graph s.t. removing any edge disconnects it.

  Note: Equivalently: A cycle-free graph s.t. introducing a new edge makes a cycle.

- A *clique* is a subset $K \subseteq V$ s.t. $\forall a, b \in K \implies \{a, b\} \in E$.

  Think of: All pairs of vertices in K have an edge between them.

  Common Notation: The size of a largest clique in a graph G is $\omega(G)$.

- The *complete graph* K_n on n nodes is a clique of n vertices.

- An *independent set* is a subset $I \subseteq V$ s.t. $\forall a, b \in I \implies \{a, b\} \notin E$.

  Think of: The opposite of *clique*. Each pair of vertices in I must not have an edge between them.

  Common Notation: The size of a largest independent set in a graph G is $\alpha(G)$.

- A *vertex cover* is a subset $A \subseteq V$ s.t. $\{a, b\} \in E \implies a \in A$ or $b \in A$.

  Think of: Pick a set of vertices such that each edge has an end in the set.

  Common Notation: $\tau(G)$ is the minimum size vertex cover.

- A *matching* is a subset $M \subseteq E$ s.t. $\{a, b\}, \{c, b\} \in M \implies a = c$.

  Think of: A set of edges so that no two elements in the set share an endpoint.

  Common Notation: $\nu(G)$ is the maximum size matching.

- A *perfect matching* is a matching of size $n/2$.

  Think of: All vertices are matched.

- A function $c : V \to [r]$ is a *proper coloring* of V with r colors if $\{a, b\} \in E \implies c(a) \neq c(b)$. The *chromatic number* of G is the smallest r such that there is a proper coloring with r colors.

  Common Notation: The chromatic number is $\chi(G)$.

  Think of: Color the vertices such that the ends of each edge are different.

Number Theory

- $\mathbb{Z}$ denotes the set of integers, and $\mathbb{Z}^+$ denotes the set of positive integers.
- For $d \in \mathbb{Z}^+$ and $a, b \in \mathbb{Z}$:

  1. $d|a$ means there exists an integer c such that $a = dc$.
  2. $d|a$ and $d|b$ implies $d|(a+b)$ and $d|(a-b)$.
  3. $d|a$ implies $d|ab$.
  4. The common divisors of a and b are all positive integers that divide both a and b. $\gcd(a, b)$ is the greatest (largest) common divisor of a and b.

- For $a, b, c, d, m \in \mathbb{Z}$, $m \geq 2$:

  1. $a \equiv b \pmod{m}$ means $m|(a-b)$.
  2. $a \bmod m$ is the unique $b \in \{0, 1, \ldots, m-1\}$ such that $a \equiv b \pmod{m}$.
  3. $a \equiv c \pmod{m}$ and $b \equiv d \pmod{m}$ imply both

  $$\begin{aligned} a + b &\equiv c + d \pmod{m} \\ a \cdot b &\equiv c \cdot d \pmod{m}. \end{aligned}$$

  Therefore $((a \bmod m)(b \bmod m)) \bmod m = (ab) \bmod m$.

- For $m \in \mathbb{Z}$, $m \geq 2$:

  1. $\mathbb{Z}_m = \{0, 1, \ldots, m-1\}$ where the operations $+$, $-$, and $\cdot$ are performed mod m.
  2. $\mathbb{Z}_m^* = \{x \in \mathbb{Z}_m : \gcd(x, m) = 1\}$.

- For $a, b, c, m \in \mathbb{Z}$, $m \geq 2$:

  1. If $\gcd(a, m) = 1$, then $ab \equiv ac \pmod{m}$ implies $b \equiv c \pmod{m}$.
  2. If $\gcd(a, m) = 1$, then there is a unique solution $x \in \mathbb{Z}_m^*$ to $ax \equiv b \pmod{m}$.
  3. For $a \in \mathbb{Z}_m^*$, the *multiplicative inverse of* a, denoted a^{-1}, is the unique element in $\mathbb{Z}_m^*$ such that $a \cdot a^{-1} \equiv 1 \pmod{m}$. Division b/a in $\mathbb{Z}_m$ means $b \cdot a^{-1}$.

Abstract Algebra

A *field* is a set $\mathbb{F}$ and operations $+, *$ s.t. $\forall a, b, c \in \mathbb{F}$

- $a + b \in \mathbb{F}$, $ab \in \mathbb{F}$
- $a + b = b + a$, $ab = ba$
- $(a + b) + c = a + (b + c)$, $(ab)c = a(bc)$
- $a(b + c) = ab + ac$
- $\exists 0 \in \mathbb{F}$ s.t. $\forall a, a + 0 = a$
- $\exists 1 \in \mathbb{F}$ s.t. $\forall a, 1 * a = a$
- $\forall a, \exists - a \in \mathbb{F}$ s.t. $a + -a = 0$
- $\forall a \neq 0, \exists a^{-1} \in \mathbb{F}$ s.t. $a * a^{-1} = 1$

Think of: A set and operations that behave like the real numbers. Pretty much all theorems that hold for real numbers hold in fields, a notable exception being theorems involving limits.

Note: Finite fields may only have sizes that are powers of primes, i.e., size p^k for some prime p and positive integer k. A finite field with q elements is denoted $\mathbb{F}_q$.

Note: The integers mod a prime p are a finite field, denoted as $\mathbb{Z}_p = \mathbb{F}_p$.

Polynomials

The following holds for polynomials over any field.

- A *monomial* is a product of non-negative powers of variables, such as x^2y^3z.
- A *polynomial* is a linear combination of monomials.
- A univariate polynomial is *monic* if it has leading coefficient 1.
- For any $k + 1$ distinct points, there is a unique ($\exists!$) univariate polynomial of degree $\leq k$ passing through them.
- For a univariate polynomial p, $p(a) = 0$ iff $(x - a)|p$, i.e., $p(x) = (x - a)q(x)$ for a polynomial q.

Linear Algebra

- A set of vectors $v_1, \dots, v_k$ over a field $\mathbb{F}$ are *linear independent* if the only solution to

  $$\lambda_1 v_1 + \lambda_2 v_2 + \dots \lambda_k v_k = 0,$$

  where $\lambda_i \in \mathbb{F}$, is $\lambda_i = 0, \forall i$.

- $\lambda \in \mathbb{C}$ is an *eigenvalue* of a complex-valued matrix M if:

  $$Mx = \lambda x$$

  for some $x \neq 0$.

  Note: The above x is called an *eigenvector*.

- A *Vandermonde matrix* is a square matrix of the form

  $$V = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \dots & a_3^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{bmatrix}$$

  Note: Vandermonde matrices are very useful when trying to find a polynomial going through a set of points.

  Note: If the a_i's are distinct, V is invertible.

- The *determinant* of a matrix $M = (M_{ij})$ is

  $$\begin{vmatrix} M_{11} & M_{12} & \dots & M_{1n} \\ M_{21} & M_{22} & \dots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \dots & M_{nn} \end{vmatrix} = \sum_{\pi \in S_n} \text{sgn}(\pi) M_{1\pi_1} M_{2\pi_2} \cdots M_{n\pi_n}$$

  where S_n is the group of all possible permutations of n elements.

  Think of: One of the permutations in the sum is the identity permutation, where $\pi(i) = i$.

  Note: The determinant may also be computed with the cofactor expansion.

- The *sign* of a permutation π is

  $$\text{sgn}(\pi) = (-1)^k$$

  where k is the number of inversions in the permutation. An inversion is a pair x, y s.t. $x < y$ but $\pi(x) > \pi(y)$.

Last updated December 30, 2025.