Instructions and Cheat Sheet
AIDAMS Mathematical examination

Antonin Della Noce

Contents

1 Instructions **2**

2 Notations for sets and set operations **3**

3 Combinatorics and counting **4**

4 Vectors, lines, and planes in space **6**

5 Dot products, orthogonality, and distances **7**

6 Parametric lines and Cartesian planes **9**

7 Sequences **10**

8 Limits of functions **12**

9 Differentiation and convexity **13**

10 Continuity and solving equations **14**

11 Exponential and natural logarithm **14**

12 Sine and cosine **15**

13 Antiderivatives and differential equations **16**

14 Integral calculus **17**

15 Probability **18**

1 Instructions

This mathematics and logic test is organized into fourteen sections covering the mathematical prerequisites to join AIDAMS. You will have 90 minutes to answer 30 questions. Each question must be answered in order. You are recommended to skip a question if you cannot solve it, rather than picking a random answer (skipping questions leads to zero gain, but giving a wrong answer leads to negative points).

Test Details:

- Total duration: 90 minutes
- Unless otherwise specified, each question has only one correct answer.
- Answer choices will be displayed randomly.
- The order of the questions will be random.
- Calculator is useless for this test: when numerical computations are required, they can be done either mentally, either the results of the computation will be given.
- You can use a printed version of this document during the test, and also a printed version of the mock exam, that is available on AIDAMS website.
- You may use blank sheets or scrap paper during the test.

Instructions for Fraud Detection and Surveillance:

> Live monitoring and video surveillance will be conducted during the test. Ensure that your webcam and microphone are working correctly before the test begins. You must prepare a valid identification document (passport or national ID) and show it to your camera at the beginning of the exam.

Please Note:

- The only items allowed on your desk are a printed version of this document, a printed version of the mock exam, blank scrap paper, and a pen.
- Phones, tablets, and other devices outside of your computer should not be near your testing environment.
- Internet searches, using other browser tabs or software, or using phones, headphones, or earphones is not allowed.
- Leaving the screen during the test is prohibited and will be automatically detected as fraud.
- Communication with others during the test is not allowed.
- You must be alone for the entire duration of the test.

2 Notations for sets and set operations

Sets. Let A and B be sets, and let x denote an arbitrary element of some set. The notation $x \in A$ means that x **belongs** to A; $x \notin A$ means that it does not.
The **empty set** is denoted by $\varnothing$. It is the unique set containing no elements, or $x \notin \varnothing$ is always true for any x.
The notation $A \subseteq B$ reads "A is **a subset** of B" and means that every element of A belongs to B.

We use the notation $\{x \in A : P(x)\}$ to describe the set of elements x in A that satisfy the property $P(x)$. For example, $\{x \in \mathbb{N} : x \text{ is even}\}$ is the set of even natural numbers.

The **intersection** of two sets A and B is the set of elements that belong to both A and B:

$$A \cap B = \{x \in A : x \in B\}$$

The **set difference** of two sets A and B is the set of elements that belong to A but not to B:

$$A \setminus B = \{x \in A : x \notin B\}$$

This notation also applies to other set operations, such as union and power set; in these cases, the domain of the dummy variable is not specified.

For instance, the **union** of two sets A and B can be written as

$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

The **power set** of a set A is the set of all subsets of A:

$$\mathcal{P}(A) = \{B : B \subseteq A\}.$$

The **Cartesian product** of two sets A and B is the set of ordered pairs (a, b) with $a \in A$ and $b \in B$:

$$A \times B = \{(a, b) : a \in A,\ b \in B\}.$$

More generally, the Cartesian product of k sets $A_1, \ldots, A_k$ is

$$A_1 \times \cdots \times A_k = \{(a_1, \ldots, a_k) : a_i \in A_i \text{ for } i \in \{1, \ldots, k\}\}.$$

When $A_1 = \cdots = A_k = A$, we have $A^k = \underbrace{A \times \cdots \times A}_{k \text{ factors}}$.

Functions. A function f from a set A to a set B is a rule that assigns each element $x \in A$ exactly one element $f(x) \in B$. We write $f : A \to B$ and call A the **domain** and B the **codomain** of f. We may also write:

$$f : x \in A \mapsto f(x) \in B.$$

This notation specifies the domain, the codomain, and the rule defining the function.

Quantifiers. For a property $P(x)$ and a set A,

$$\forall x \in A,\ P(x) \quad \text{means "for every } x \text{ in } A,\ P(x) \text{ holds"},$$

$$\exists x \in A,\ P(x) \quad \text{means "there is an } x \text{ in } A \text{ for which } P(x) \text{ holds".}$$

When $A = \varnothing$, the universal statement is true and the existential statement is false.

Number sets. We use the following conventions:

$$\begin{aligned}
&\mathbb{N} = \{0, 1, 2, \ldots\},\\
&\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\},\\
&\mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z},\ q \in \mathbb{Z} \setminus \{0\}\right\},\\
&\mathbb{R} = \text{the set of real numbers},\\
&\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}, \qquad\qquad i^2 = -1.
\end{aligned}$$

Thus, $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.

We write $\mathbb{R}_+ = [0, +\infty[$.

Whenever a number set excludes zero, we use a superscript asterisk; for example, $\mathbb{N}^* = \{1, 2, 3, \ldots\}$, or $\mathbb{R}^* = \mathbb{R} \setminus \{0\} =]-\infty, 0[\cup]0, +\infty[$.

Intervals of real numbers. Let $a, b \in \mathbb{R}$. We consider the following notations for intervals of real numbers

$$\begin{aligned}
[a, b] &= \{x \in \mathbb{R} : a \le x \le b\},\\
[a, b[&= \{x \in \mathbb{R} : a \le x < b\},\\
]a, b] &= \{x \in \mathbb{R} : a < x \le b\},\\
]a, b[&= \{x \in \mathbb{R} : a < x < b\}.
\end{aligned}$$

The symbols $+\infty$ and $-\infty$ are not real numbers, but are elements of the extended real line used to describe unbounded intervals. In particular, every real number x satisfies $-\infty < x < +\infty$. Let $a \in \mathbb{R}$. Unbounded intervals are denoted by

$$\begin{aligned}
]-\infty, a] &= \{x \in \mathbb{R} : x \le a\},\\
]-\infty, a[&= \{x \in \mathbb{R} : x < a\},\\
[a, +\infty[&= \{x \in \mathbb{R} : x \ge a\},\\
]a, +\infty[&= \{x \in \mathbb{R} : x > a\}.
\end{aligned}$$

Intervals of integers. Similar notation is used for intervals of integers. Let $m, n \in \mathbb{Z}$. We define

$$\begin{aligned}
[\![m, n]\!] &= \{k \in \mathbb{Z} : m \le k \le n\},\\
[\![m, n[\![&= \{k \in \mathbb{Z} : m \le k < n\},\\
]\!]m, n]\!] &= \{k \in \mathbb{Z} : m < k \le n\},\\
]\!]m, n[\![&= \{k \in \mathbb{Z} : m < k < n\}.
\end{aligned}$$

3 Combinatorics and counting

Cardinality of finite sets For finite sets, $|A|$ denotes the number of elements of A.

Cardinality of a disjoint union Let $A_1, \ldots, A_r$ be pairwise disjoint sets; that is, $A_i \cap A_j = \varnothing$ whenever $i \neq j$. Then

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

Cardinality of a Cartesian product Let $A_1, \ldots, A_k$ be finite sets. Then

$$|A_1 \times \cdots \times A_k| = \prod_{i=1}^{k} |A_i|.$$

In particular, if $|A| = n$, then $|A^k| = n^k$ for $k \geq 1$. By convention, $A^0 = \{\varnothing\}$, so $|A^0| = 1$.

Cardinality of a union of two sets . A direct consequence of equation (1) is that, for any sets A and B,

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

Ordered choices, factorials, and permutations. For $0 \leq k \leq n$, the number of ordered k-tuples of distinct elements of an n-element set is

$$n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}, \qquad n! = 1 \cdot 2 \cdots n, \qquad 0! = 1.$$

The empty product is 1. A permutation orders every element exactly once, and there are $n!$ permutations. If repetition is allowed, there are n^k ordered choices.

Subsets and combinations. A combination of k elements is a k-element subset: its order does not matter and no element is repeated. Each subset has $k!$ orderings, hence

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}, \qquad 0 \leq k \leq n.$$

In particular,

$$\binom{n}{0} = \binom{n}{n} = 1, \qquad \binom{n}{1} = n, \qquad \binom{n}{2} = \frac{n(n-1)}{2}, \qquad \binom{n}{k} = \binom{n}{n-k}.$$

The last identity pairs each subset with its complement. Binary words of length n with exactly k ones or subsets of size k of an n-element set are all counted by $\binom{n}{k}$.

All subsets and the binomial sum (proof). A subset of $\{a_1, \ldots, a_n\}$ corresponds to its membership word $(\varepsilon_1, \ldots, \varepsilon_n) \in \{0,1\}^n$, where $\varepsilon_i = 1$ exactly when a_i is selected. Thus

$$|\mathcal{P}(A)| = 2^n, \qquad \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

For the second equality, partition all subsets by their size. The same 2^n counts binary words, complete paths in a binary tree of depth n, and possible records of n Bernoulli trials.

Pascal's identity and triangle (two proofs). For $n \geq 1$ and $1 \leq k \leq n$,

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

By counting, a k-subset either contains a fixed distinguished element (choose its other $k-1$ elements) or omits it (choose all k elements). Algebraically,

$$\frac{n!}{(k-1)!(n-k+1)!}+\frac{n!}{k!(n-k)!}=\frac{n!\,[k+(n-k+1)]}{k!(n-k+1)!}=\binom{n+1}{k}.$$

Pascal's triangle starts with row (1); every row begins and ends with 1, and each interior entry is the sum of the two entries above it.

Counting method. Specify the objects precisely: does order matter, is repetition allowed, and are cases disjoint? Use a tree for successive choices, a table for pairs, and sets or diagrams for unions. For example, q^n counts words of length n over q symbols; $\binom{n}{k}$ counts teams of size k; $n!$ counts rankings of n distinct players. Restrictions are handled by partitioning into cases or counting the complement.

4 Vectors, lines, and planes in space

Vectors, translations, and linear combinations. A vector describes a translation, independently of its starting point. The translation of M by $\mathbf{u}$ is M' with $\overrightarrow{MM'}=\mathbf{u}$. For any points A,B,C,

$$\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC},\qquad \overrightarrow{BA}=-\overrightarrow{AB},\qquad \overrightarrow{AA}=\mathbf{0}.$$

A linear combination is an expression of the form $a\mathbf{u}+b\mathbf{v}+c\mathbf{w}$, with real coefficients. Two vectors are **collinear** when one is a scalar multiple of the other, including when one vector is zero.

Lines and planes. A line through A with nonzero direction vector $\mathbf{u}$ is

$$\ell=\{M:\overrightarrow{AM}=t\mathbf{u},\ t\in\mathbb{R}\}.$$

Thus, assuming $A\neq B$, the points A,B,C are collinear exactly when $\overrightarrow{AC}=t\overrightarrow{AB}$ for some t. A plane through A with noncollinear vectors $\mathbf{u},\mathbf{v}$ is

$$P=\{M:\overrightarrow{AM}=s\mathbf{u}+t\mathbf{v},\ (s,t)\in\mathbb{R}^2\}.$$

Its direction consists of all $s\mathbf{u}+t\mathbf{v}$ for all real numbers s and t. For noncollinear points A,B,C, a point D lies in their plane exactly when $\overrightarrow{AD}=s\overrightarrow{AB}+t\overrightarrow{AC}$ for some $s,t\in\mathbb{R}$.

Bases and coordinate frames. Two noncollinear vectors form a basis of a plane's direction. Three vectors $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ form a basis of space when they are linearly independent, equivalently noncoplanar:

$$\text{for all } a,b,c\in\mathbb{R},\ \text{if } a\mathbf{e}_1+b\mathbf{e}_2+c\mathbf{e}_3=\mathbf{0}\quad\text{then}\quad a=b=c=0.$$

Every vector then has a unique decomposition in this basis:

$$\mathbf{u}=x\mathbf{e}_1+y\mathbf{e}_2+z\mathbf{e}_3.$$

A frame $(O;\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ assigns coordinates $M=(x,y,z)$ through $\overrightarrow{OM}=x\mathbf{e}_1+y\mathbf{e}_2+z\mathbf{e}_3$. Thus, for $A=(x_A,y_A,z_A)$ and $B=(x_B,y_B,z_B)$,

$$\overrightarrow{AB}=(x_B-x_A,y_B-y_A,z_B-z_A).$$

In a parallelepiped, let B, D, E be the vertices adjacent to A. The three edge vectors form a basis, and the opposite vertex G satisfies $\overrightarrow{AG} = \overrightarrow{AB} + \overrightarrow{AD} + \overrightarrow{AE}$. In a tetrahedron $ABCD$, the vectors $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$ form a basis.

Relative positions. Two distinct lines are parallel when their direction vectors are collinear, intersecting when they share one point, and skew when they are neither parallel nor intersecting (so lie in no common plane). For a line and a plane, there are three cases: the line is contained in the plane, is parallel and disjoint, or intersects at one point. Two planes are equal, parallel and disjoint, or intersect in a line.

5 Dot products, orthogonality, and distances

Dot product and norm. For nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ of norms $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ with angle $\theta \in [0, \pi]$, the dot product of the two vectors is

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta, \qquad \mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2.$$

The dot product is symmetric and bilinear:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}, \qquad (a\mathbf{u} + b\mathbf{v}) \cdot \mathbf{w} = a(\mathbf{u} \cdot \mathbf{w}) + b(\mathbf{v} \cdot \mathbf{w}).$$

Orthogonality is characterized by $\mathbf{u} \perp \mathbf{v} \iff \mathbf{u} \cdot \mathbf{v} = 0$. The zero vector is orthogonal to every vector.

Orthonormal coordinates. A basis is orthonormal if its vectors have norm 1 and are pairwise orthogonal. In an orthonormal frame only, for $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$,

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3, \qquad \|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2},$$

$$AB = \|\overrightarrow{AB}\| = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}.$$

Norm expansions and polarization.

$$\|\mathbf{u} \pm \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \pm 2\mathbf{u} \cdot \mathbf{v},$$

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= \frac{\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2}{2} \\ &= \frac{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2}{2} \\ &= \frac{\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2}{4}. \end{aligned}$$

In particular,

$$\overrightarrow{AB} \cdot \overrightarrow{AC} = (AB^2 + AC^2 - BC^2)/2.$$

This formula can be used to compute the angle at A in triangle ABC from the lengths of its three sides.

Orthogonal lines, normal vectors, and planes. Two lines are orthogonal when their direction vectors have dot product 0, even if the lines are skew. They are perpendicular when they also intersect. A nonzero vector $\mathbf{n}$ is normal to a plane directed by $\mathbf{u}, \mathbf{v}$ exactly when

$$\mathbf{n} \cdot \mathbf{u} = \mathbf{n} \cdot \mathbf{v} = 0.$$

A line is perpendicular to a plane exactly when its direction vector is normal to the plane. Two planes are perpendicular exactly when their normal vectors are orthogonal. Given A and $\mathbf{n} \neq \mathbf{0}$, the plane through A normal to $\mathbf{n}$ is $\{M : \mathbf{n} \cdot \overrightarrow{AM} = 0\}$.

Orthogonal projections and nearest points. The orthogonal projection H of M onto a line or plane belongs to that line or plane and $\overrightarrow{HM}$ is perpendicular to its direction. For every other point X of the line or plane, Pythagoras gives

$$MX^2 = MH^2 + HX^2 \geq MH^2,$$

with equality only for $X = H$. Hence, H is the unique nearest point and the distance is MH. For a line $\ell = A + \mathbb{R}\mathbf{u}$, write $\mathbf{w} = \overrightarrow{AM}$. Then

$$\overrightarrow{AH} = \frac{\mathbf{w} \cdot \mathbf{u}}{\|\mathbf{u}\|^2}\mathbf{u}, \qquad d(M, \ell)^2 = \|\mathbf{w}\|^2 - \frac{(\mathbf{w} \cdot \mathbf{u})^2}{\|\mathbf{u}\|^2}.$$

For a plane P through A with normal vector $\mathbf{n}$,

$$\overrightarrow{MH} = -\frac{\overrightarrow{AM} \cdot \mathbf{n}}{\|\mathbf{n}\|^2}\mathbf{n}, \qquad d(M, P) = \frac{|\overrightarrow{AM} \cdot \mathbf{n}|}{\|\mathbf{n}\|}.$$

Angles, areas, and volumes. For $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$,

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

The acute or right angle α between two lines uses the absolute value of this quotient. The angle $\beta \in [0, \pi/2]$ between a line with direction $\mathbf{u}$ and a plane with normal $\mathbf{n}$ satisfies

$$\sin\beta = \frac{|\mathbf{u} \cdot \mathbf{n}|}{\|\mathbf{u}\|\|\mathbf{n}\|}.$$

For two planes, the acute or right angle is given by the absolute value of the dot-product quotient of their normal vectors. Useful measurements are

$$\begin{aligned} \operatorname{area}(ABC) &= \tfrac{1}{2} AB\, AC \sin\widehat{BAC} \\ &= \tfrac{1}{2}\sqrt{\|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2}, \\ V_{\text{prism}} &= Bh, \qquad V_{\text{pyramid}} = \tfrac{1}{3}Bh, \\ V_{ABCD} &= \tfrac{1}{3}\operatorname{area}(ABC)\, d(D, (ABC)), \end{aligned}$$

where $\mathbf{u} = \overrightarrow{AB}$, $\mathbf{v} = \overrightarrow{AC}$, B is the base area, and h is the perpendicular height.

A simple geometric locus. For distinct A, B with midpoint I, expansion gives

$$MA = MB \iff \overrightarrow{IM} \cdot \overrightarrow{AB} = 0.$$

Thus points equidistant from A and B form the plane through I normal to $\overrightarrow{AB}$, the perpendicular bisector plane.

6 Parametric lines and Cartesian planes

In all the section, unless otherwise stated, we work in dimension 3, i.e., points are elements of $\mathbb{R}^3$, and vectors are elements of $\mathbb{R}^3$ as well.

Parametric representation of a line. In an orthonormal frame, a line through $A = (x_A, y_A, z_A)$ with direction $\mathbf{u} = (u_1, u_2, u_3) \neq \mathbf{0}$ has representation

$$\ell : \quad \begin{cases} x = x_A + tu_1, \\ y = y_A + tu_2, \\ z = z_A + tu_3, \end{cases} \qquad t \in \mathbb{R}.$$

All three coordinates use the same parameter. To test membership, solve the three equations for one common t; a zero direction coordinate means that the corresponding point coordinate is constant. For distinct points A and B, take $\mathbf{u} = \overrightarrow{AB}$.

Cartesian equation of a plane (derivation). Let $A = (x_A, y_A, z_A)$ and $\mathbf{n} = (a, b, c) \neq \mathbf{0}$. Then

$$\begin{aligned} M \in P &\iff \mathbf{n} \cdot \overrightarrow{AM} = 0 \\ &\iff a(x - x_A) + b(y - y_A) + c(z - z_A) = 0 \qquad d = -ax_A - by_A - cz_A. \\ &\iff ax + by + cz + d = 0, \end{aligned}$$

Conversely, any equation $ax + by + cz + d = 0$ with $(a, b, c) \neq (0, 0, 0)$ defines a plane with normal (a, b, c). Multiplying every coefficient by the same nonzero scalar gives the same plane. For three noncollinear points A, B, C, find a nonzero $\mathbf{n}$ solving $\mathbf{n} \cdot \overrightarrow{AB} = \mathbf{n} \cdot \overrightarrow{AC} = 0$, then use the equation above.

Projection coordinates. For $M = (x_M, y_M, z_M)$ and $P : ax + by + cz + d = 0$, set

$$q = \frac{ax_M + by_M + cz_M + d}{a^2 + b^2 + c^2}.$$

Then

$$H = (x_M - qa,\ y_M - qb,\ z_M - qc), \qquad d(M, P) = \frac{|ax_M + by_M + cz_M + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

For $\ell = A + \mathbb{R}\mathbf{u}$, the projection is

$$H = A + t_0\mathbf{u}, \qquad t_0 = \frac{\overrightarrow{AM} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}.$$

Alternatively, solve simultaneously $H \in P$ and $\overrightarrow{MH} \parallel \mathbf{n}$, or $H \in \ell$ and $\overrightarrow{MH} \cdot \mathbf{u} = 0$.

Linear systems for bases and configurations. For $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, $\mathbf{w} = (w_1, w_2, w_3)$, a decomposition $\mathbf{r} = a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$ means

$$\begin{cases} au_1 + bv_1 + cw_1 = r_1, \\ au_2 + bv_2 + cw_2 = r_2, \\ au_3 + bv_3 + cw_3 = r_3. \end{cases}$$

Use substitution or elimination. The three vectors form a basis exactly when the corresponding homogeneous system ($\mathbf{r} = \mathbf{0}$) has only $a = b = c = 0$. In a basis the solution

gives the unique coordinates of $\mathbf{r}$. To test collinearity for $\mathbf{u} \neq \mathbf{0}$, solve $\mathbf{v} = t\mathbf{u}$; to test whether three vectors lie in the same plane, solve $\mathbf{w} = s\mathbf{u} + t\mathbf{v}$ for noncollinear $\mathbf{u}, \mathbf{v}$.

Line–line intersection and parallelism. For $\ell = A + \mathbb{R}\mathbf{u}$ and $m = B + \mathbb{R}\mathbf{v}$, solve

$$A + t\mathbf{u} = B + s\mathbf{v}$$

as three scalar equations in t, s. Collinear directions give parallel lines: they coincide exactly when $\overrightarrow{AB}$ is collinear with $\mathbf{u}$, and otherwise are disjoint. Noncollinear directions give intersecting lines if the system has a solution, and skew lines if it has no solution. Orthogonality is tested separately by $\mathbf{u} \cdot \mathbf{v} = 0$.

Line–plane position. For $\ell = A + \mathbb{R}\mathbf{u}$ and $P : \mathbf{n} \cdot \mathbf{x} + d = 0$, substitution gives the single equation

$$(\mathbf{n} \cdot \mathbf{u})t + \mathbf{n} \cdot \mathbf{a} + d = 0,$$

where $\mathbf{a} = \overrightarrow{OA}$. If $\mathbf{n} \cdot \mathbf{u} \neq 0$, there is exactly one intersection. If $\mathbf{n} \cdot \mathbf{u} = 0$, the line is contained in P when $\mathbf{n} \cdot \mathbf{a} + d = 0$, and parallel and disjoint otherwise. The line is perpendicular to P exactly when $\mathbf{u}$ is collinear with $\mathbf{n}$.

Plane–plane position. For $P : \mathbf{n} \cdot \mathbf{x} + d = 0$ and $Q : \mathbf{m} \cdot \mathbf{x} + e = 0$, collinear normal vectors give parallel planes. If $\mathbf{m} = \lambda\mathbf{n}$, the planes coincide exactly when $e = \lambda d$; otherwise they are disjoint. Noncollinear normal vectors give a line of intersection: solve the two equations with one coordinate as a free parameter. The planes are perpendicular exactly when $\mathbf{n} \cdot \mathbf{m} = 0$.

7 Sequences

Limits and bounds. A real sequence is a function $u : \mathbb{N} \to \mathbb{R}$, written $(u_n)_{n\in\mathbb{N}}$. It is increasing if $u_{n+1} \geq u_n$ for every $n \in \mathbb{N}$, bounded above if there is an $M \in \mathbb{R}$ such that $u_n \leq M$ for every $n \in \mathbb{N}$, and bounded below if there is an $m \in \mathbb{R}$ such that $u_n \geq m$ for every $n \in \mathbb{N}$.

$$\begin{aligned} u_n \to \ell \in \mathbb{R} &\iff \forall \varepsilon > 0,\ \exists N,\ \forall n \geq N, \quad |u_n - \ell| < \varepsilon, \\ u_n \to +\infty &\iff \forall A \in \mathbb{R},\ \exists N,\ \forall n \geq N, \quad u_n > A. \end{aligned}$$

The meaning of $u_n \to -\infty$ is $-u_n \to +\infty$. A convergent sequence has a unique finite limit and is bounded; a bounded sequence need not converge, as $((-1)^n)_{n\in\mathbb{N}}$ shows.

Operations and comparison. If $u_n \to u$, $v_n \to v$ are finite, then

$$u_n + v_n \to u + v, \qquad u_n v_n \to uv, \qquad \frac{u_n}{v_n} \to \frac{u}{v} \quad (v \neq 0).$$

For infinite limits, the following rules apply:

$$\begin{aligned} u_n \to +\infty,\ v_n \to v \in \mathbb{R} &\implies u_n + v_n \to +\infty, \quad v_n/u_n \to 0, \\ u_n \to 0^+ &\implies 1/u_n \to +\infty, \\ u_n \to 0^- &\implies 1/u_n \to -\infty. \end{aligned}$$

If $u_n \to +\infty$ and $v_n \to v \neq 0$, then $u_n v_n$ tends to $+\infty$ for $v > 0$ and to $-\infty$ for $v < 0$. Products of two infinite limits follow the same sign rule; sums of two infinite limits with the

same sign retain that sign. Taking the opposite of a sequence gives the corresponding rules for $-\infty$. If $u_n \le v_n \le w_n$ for all $n \ge N$ and $u_n, w_n \to \ell$, then $v_n \to \ell$ (squeeze theorem). If $u_n \ge v_n \to +\infty$, then $u_n \to +\infty$: for any A, eventually $v_n > A$, hence $u_n > A$. The corresponding result holds with the inequalities reversed for $-\infty$. Inequalities are preserved by finite limits, but strict inequalities can become equalities.

Convergence of monotone sequences. An increasing sequence bounded above converges; a decreasing sequence bounds below converges. An increasing sequence that is unbounded above tends to $+\infty$: choose N with $u_N > A$, then $u_n \ge u_N > A$ for $n \ge N$. Similarly, a decreasing sequence that is unbounded below tends to $-\infty$. Monotonicity alone does not imply a finite limit.

Arithmetic and geometric models.

$$\begin{gathered}
\forall n \in \mathbb{N},\ u_{n+1} = u_n + d \implies \forall n \in \mathbb{N},\ u_n = u_0 + nd, \\
\forall n \in \mathbb{N},\ u_{n+1} = qu_n \implies \forall n \in \mathbb{N},\ u_n = u_0 q^n, \\
\forall n \in \mathbb{N},\ \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \quad (q \neq 1), \\
\forall n \in \mathbb{N},\ \sum_{k=0}^{n} (a+kd) = (n+1)a + \frac{dn(n+1)}{2}.
\end{gathered}$$

For $q = 1$ the geometric sum is $n + 1$. All real ratios are covered by

q	behavior of q^n as $n \to \infty$
$-1 < q < 1$	0
$q = 1$	1
$q > 1$	$+\infty$
$q = -1$	alternates between 1 and -1
$q < -1$	$\lvert q^n \rvert \to +\infty$ with alternating signs

For $q = 0$, take $n \ge 1$. For $q \le -1$, q^n has no finite or infinite limit.

Bernoulli's inequality and geometric limits. For $t \ge -1$ and integers $n \ge 1$,

$$(1+t)^n \ge 1 + nt.$$

The base case $n = 1$ is equality. The induction step uses multiplication by $1 + t \ge 0$ and $(1 + nt)(1 + t) = 1 + (n + 1)t + nt^2 \ge 1 + (n + 1)t$. For $q > 1$, set $t = q - 1 > 0$ to obtain $q^n \to +\infty$. For $0 < |q| < 1$, apply this result to $1/|q|$, then take reciprocals; $|q^n| \to 0$ implies $q^n \to 0$.

Recurrences and evolution. For $u_{n+1} = au_n + b$, $a \neq 1$, the equilibrium $L = b/(1 - a)$ gives

$$u_n = L + a^n(u_0 - L).$$

When $|a| < 1$, $u_n \to L$; when $a = 1$, use the arithmetic formula. For a general $u_{n+1} = f(u_n)$, prove an invariant interval and bounds by induction, study $u_{n+1} - u_n$, then use monotone convergence. If the limit exists and f is continuous there, it satisfies $L = f(L)$; solving this equation alone does not prove convergence.

Thresholds and numerical constants. Starting from $(n, u) = (0, u_0)$, repeat $(n, u) \leftarrow (n+1, f(u))$ while $u < T$ to find the first index with $u_n \ge T$, when such an index exists. A

proved error bound gives a reliable stopping condition. Typical convergent approximations are

$$e = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!}, \qquad \pi = \lim_{n\to\infty} 4 \sum_{k=0}^{n} \frac{(-1)^k}{2k+1},$$
$$h_{n+1} = \frac{1}{2}\left(h_n + \frac{2}{h_n}\right) \to \sqrt{2} \quad (h_0 > 0),$$
$$v_{n+1} = 1 + \frac{1}{v_n} \to \frac{1+\sqrt{5}}{2} \qquad (v_0 = 1).$$

The alternating-series approximation to π has error at most $4/(2n+3)$.

8 Limits of functions

Accumulation points and limits. Let A be a subset of $\mathbb{R}$ and $a \in \mathbb{R}$. Then a is an accumulation point of A if every neighborhood of a contains at least one point of A different from a. Equivalently,

$$\forall \delta > 0,\ \exists x \in A,\ 0 < |x - a| < \delta.$$

For example, 1 is an accumulation point of $[0, 1[$.

Finite, infinite and one-sided limits. For $f : D \to \mathbb{R}$ and an accumulation point a of D,

$$\lim_{x\to a} f(x) = \ell \iff \forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in D,$$
$$0 < |x - a| < \delta \implies |f(x) - \ell| < \varepsilon.$$

In a one-sided limit $x \to a^+$ or a^-, restrict to $x > a$ or $x < a$. At $+\infty$, replace the condition by $x > R$; at $-\infty$, use $x < -R$. A limit $+\infty$ means that $f(x) > M$ eventually for every $M > 0$; a limit $-\infty$ means $f(x) < -M$ eventually. A two-sided limit exists precisely when both approachable one-sided limits exist and agree. The value $f(a)$ does not determine the limit.

Operations, composition and comparison. The sequence limit rules apply to functions. If $f(x) \to b$ and g is continuous at b, then $g(f(x)) \to g(b)$. More generally one may compose limits when the arguments approach within the required domain; if g is undefined at b, check that $f(x) \neq b$ eventually. Squeeze and divergence comparisons apply near the point being studied.

Reference limits and growth. Polynomials are continuous; rational functions are continuous where their denominators are nonzero. For integers $m \geq 1$,

$$x^m \to +\infty\ (x \to +\infty), \qquad x^{-m} \to 0\ (x \to \pm\infty),$$
$$\sqrt{x} \to 0\ (x \to 0^+), \qquad \sqrt{x} \to +\infty\ (x \to +\infty),$$
$$e^x \to +\infty\ (x \to +\infty), \qquad e^x \to 0\ (x \to -\infty),$$
$$\frac{x^m}{e^x} \to 0\ (x \to +\infty), \qquad x^m e^x \to 0\ (x \to -\infty).$$

At $-\infty$, integer powers have the sign given by their parity. For $c > 0$, $x^m/e^{cx} \to 0$ at $+\infty$ by setting $t = cx$. At 0^+, $1/x^m \to +\infty$; at 0^- its sign depends on parity.

Proof ideas for exponential growth. Since $e^x \geq 1 + x$, $e^x \to +\infty$ as $x \to +\infty$; $e^{-x} = 1/e^x$ gives the negative-infinity limit. For $x \geq 0$, repeatedly integrate $e^t \geq 1$ to obtain $e^x \geq x^{m+1}/(m+1)!$. Hence, $0 \leq x^m/e^x \leq (m+1)!/x \to 0$.

Asymptotes and dominant terms. If $f(x) \to L$ at $+\infty$ or $-\infty$, then $y = L$ is a horizontal asymptote on that side. If $f(x) \to \pm\infty$ as $x \to a^+$ or a^-, then $x = a$ is a vertical asymptote. At infinity, factor the highest power in a polynomial or rational function:

$$\frac{a_m x^m + \cdots + a_0}{b_n x^n + \cdots + b_0} = x^{m-n} \frac{a_m + \cdots + a_0/x^m}{b_n + \cdots + b_0/x^n}, \qquad a_m b_n \neq 0.$$

Compare the resulting power and the signs on the side considered.

9 Differentiation and convexity

Interior point of an interval. A real number a is an interior point of an interval I if there exists $\delta > 0$ such that $(a - \delta, a + \delta) \subseteq I$.

Derivative, tangent and calculation rules. Let $f : I \to \mathbb{R}$ be a function defined on an interval I and let a be an interior point of I. We say that f is differentiable at a if the following limit exists:

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

The equation of the tangent line to the graph of f at a is

$$T_a : x \in \mathbb{R} \mapsto f(a) + f'(a)(x - a) \in \mathbb{R}.$$

Where the expressions are defined and differentiable,

$$(u+v)' = u' + v', \qquad (cu)' = cu', \qquad (uv)' = u'v + uv',$$
$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \quad (v \neq 0), \quad (v \circ u)' = (v' \circ u)u'.$$

Useful derivatives, with their natural domain restrictions, are

$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
x^n	nx^{n-1} $(n \in \mathbb{Z} \setminus \{0\})$	$\sqrt{x}$	$1/(2\sqrt{x})$ $(x > 0)$
e^x	e^x	$\ln x$	$1/x$ $(x > 0)$
$\sin x$	$\cos x$	$\cos x$	$-\sin x$

A constant has derivative zero; for negative powers exclude $x = 0$. In particular, $(e^u)' = u'e^u$, $(\ln u)' = u'/u$ for $u > 0$, $(\sin u)' = u' \cos u$, $(\cos u)' = -u' \sin u$. The second derivative is $f'' = (f')'$.

Variations, maxima, minima and function studies. On an interval, $f' \geq 0$ implies f is increasing and $f' \leq 0$ implies it is decreasing. A change of sign of f' from $+$ to $-$ gives a local maximum, from $-$ to $+$ a local minimum. At an interior differentiable local maximum or minimum, $f' = 0$; this condition alone is insufficient. For an absolute maximum or minimum on a closed interval, compare endpoint values and all interior critical points.

Convexity and concavity. Let I be an interval. A function $f : I \to \mathbb{R}$ is convex on I if, for $x, y \in I$ and $t \in [0, 1]$,

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

Its graph lies below every chord joining two of its points. For twice differentiable f on an interval, the following are equivalent:

$$\begin{aligned} f \text{ is convex} &\iff f' \text{ is increasing} \iff f'' \geq 0 \\ &\iff \forall a, x \in I, \quad f(x) \geq f(a) + f'(a)(x-a). \end{aligned}$$

Concavity reverses these inequalities and makes f' decreasing. An inflection point is a point of the graph where convexity changes. If f'' is continuous, a zero of f'' with a sign change gives an inflection point; $f''(a) = 0$ alone does not (example: $f(x) = x^4$ at 0). Read convexity from the increase of f' or the sign of f'', and variation of f from the sign of f'.

Tangent inequalities: proof and examples. Fix a and set $g(x) = f(x) - f(a) - f'(a)(x-a)$. If $f'' \geq 0$, then $g' = f' - f'(a) \leq 0$ to the left of a and $g' \geq 0$ to its right. Thus, $g(x) \geq g(a) = 0$. Applied to e^x at 0, this gives $e^x \geq 1 + x$. Applied to the concave logarithm at 1, it gives $\ln x \leq x - 1$.

10 Continuity and solving equations

Continuity and convergent sequences. A function is continuous at a if $\lim_{x \to a} f(x) = f(a)$, with one-sided limits at an endpoint; it is continuous on an interval if it is continuous at every point. Differentiability implies continuity; the converse is false, as $|x|$ at 0 shows. If $u_n \to L$ and f is continuous at L, then $f(u_n) \to f(L)$. For $u_{n+1} = f(u_n)$ on an invariant interval, combine this property with bounds and monotonicity to identify the limit.

Intermediate value theorem (IVT). If f is continuous on $[a, b]$, it takes every value between $f(a)$ and $f(b)$. In particular, $f(a)f(b) < 0$ guarantees a root in $]a, b[$. If f is also strictly monotone, each such value is taken exactly once: f is a bijection from $[a, b]$ onto the interval with endpoints $f(a)$ and $f(b)$. For $f(x) = k$, apply the theorem to $g = f - k$. Endpoint limits can replace endpoint values on an open interval.

Bisection. For continuous function $g : [a, b] \to \mathbb{R}$ with opposite endpoint signs $g(a)g(b) < 0$, take $m = (a+b)/2$. If $g(m) = 0$, stop; otherwise retain the half-interval whose endpoint signs are opposite. After n steps the width is $(b-a)/2^n$; its midpoint is within $(b-a)/2^{n+1}$ of a root in that interval. Strict monotonicity ensures that this is the unique root.

Newton and secant iterations.

$$\begin{aligned} x_{n+1} &= x_n - \frac{g(x_n)}{g'(x_n)} && \text{(Newton, } g'(x_n) \neq 0\text{)}, \\ x_{n+1} &= x_n - g(x_n)\frac{x_n - x_{n-1}}{g(x_n) - g(x_{n-1})} && \text{(secant, nonzero denominator)}. \end{aligned}$$

Newton intersects the tangent with the horizontal axis; the secant method uses the line through two successive graph points. Convergence requires suitable hypotheses and starting values; small successive changes alone do not certify the error.

11 Exponential and natural logarithm

Inverse functions and algebra. The exponential function $\exp : x \in \mathbb{R} \mapsto \exp(x) = e^x \in \mathbb{R}_+^*$ is a strictly increasing bijection. Its inverse is the logarithm $\ln : \mathbb{R}_+^* \to \mathbb{R}$.

We have the following identities for all $x, y \in \mathbb{R}$ and $a, b > 0$, $r \in \mathbb{R}$:

$$\ln(e^x) = x, \qquad e^{\ln x} = x \ (x > 0), \qquad \ln 1 = 0, \quad \ln e = 1,$$
$$e^{x+y} = e^x e^y, \qquad e^{-x} = \frac{1}{e^x},$$
$$\ln(ab) = \ln a + \ln b, \quad \ln(a/b) = \ln a - \ln b, \quad \ln(a^r) = r \ln a.$$

The graphs of the logarithm and exponential are reflections across the line $y = x$.
The exponential function is increasing and convex since for all $x \in \mathbb{R}$

$$\exp'(x) = e^x > 0, \qquad \exp''(x) = e^x > 0.$$

The logarithm is increasing and concave since, for all $x > 0$,

$$\ln'(x) = \frac{1}{x} > 0, \qquad \ln''(x) = -\frac{1}{x^2} < 0.$$

To obtain the above derivative, differentiate the relation $e^{\ln x} = x$ for $x > 0$: $e^{\ln x} \ln'(x) = 1$.
Limits and comparison. For every integer $m \geq 1$,

$$\lim_{x \to 0^+} \ln x = -\infty, \qquad \lim_{x \to +\infty} \ln x = +\infty,$$
$$\lim_{x \to +\infty} \frac{\ln x}{x^m} = 0, \qquad \lim_{x \to 0^+} x^m \ln x = 0, \qquad \lim_{h \to 0} \frac{\ln(1+h)}{h} = 1.$$

For $x \to 0^+$ set $t = -\ln x \to +\infty$: $x \ln x = -t/e^t \to 0$ by exponential dominance. To prove the limit as $x \to +\infty$, set $t = \ln x$.

12 Sine and cosine

Radians, identities and curves. Angles are in radians: a full turn is 2π.
For all $x, y \in \mathbb{R}$ we have the following fundamental identities:

$$\sin^2 x + \cos^2 x = 1, \qquad |\sin x| \leq 1, |\cos x| \leq 1,$$
$$\sin(-x) = -\sin x, \qquad \cos(-x) = \cos x,$$
$$\sin(x + 2\pi) = \sin x, \qquad \cos(x + 2\pi) = \cos x,$$
$$\sin(x + y) = \sin x \cos y + \cos x \sin y,$$
$$\cos(x + y) = \cos x \cos y - \sin x \sin y.$$

Here are some basic values of sine and cosine:

x	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
$\sin x$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\cos x$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0

The derivatives of sine and cosine are as follows:

$$\forall x \in \mathbb{R}, \quad \sin'(x) = \cos x, \quad \cos'(x) = -\sin x.$$

The sine function increases on $[-\pi/2, \pi/2]$ and decreases on $[\pi/2, 3\pi/2]$; cosine increases on $[-\pi, 0]$ and decreases on $[0, \pi]$. These variations repeat periodically.

Trigonometric equations and inequalities. For $a \in [-1,1]$, let $\theta \in [0,\pi]$ be the unique angle with $\cos\theta = a$ (also denoted $\arccos a$). Then, for any $x \in \mathbb{R}$,

$$\cos x = a \iff \exists k \in \mathbb{Z},\ x = 2k\pi \pm \theta,$$
$$\{x \in [-\pi,\pi] : \cos x \le a\} = [-\pi,-\theta] \cup [\theta,\pi],$$
$$\{x \in [-\pi,\pi] : \cos x \ge a\} = [-\theta,\theta].$$

If $a < -1$, $\cos x \le a$ has no solution; if $a > 1$ it holds for every x. The equation $\cos x = a$ has no solution when $|a| > 1$. To solve a sine equation, use the identity $\sin x = \cos(\pi/2 - x)$.

13 Antiderivatives and differential equations

Antiderivatives. On an interval I, a function F is an antiderivative of a function f if F is differentiable and $F'(x) = f(x)$ for all $x \in I$. Every continuous function $f : I \to \mathbb{R}$ has an antiderivative. If two functions F and G satisfy $F' = G' = f$, then $(F-G)' = 0$, hence $F - G$ is constant on I; all functions y satisfying $y' = f$ are of the form $y = F + C$, with $C \in \mathbb{R}$.

$f(x)$	one antiderivative	conditions
x^n	$x^{n+1}/(n+1)$	$n \in \mathbb{Z},\ n \neq -1$
$1/x$	$\ln\lvert x\rvert$	$x \in \mathbb{R}_-^*$ or $x \in \mathbb{R}_+^*$
$1/\sqrt{x}$	$2\sqrt{x}$	$x > 0$
e^x	e^x	$x \in \mathbb{R}$
$\sin x$	$-\cos x$	$x \in \mathbb{R}$
$\cos x$	$\sin x$	$x \in \mathbb{R}$

For negative powers, zero must be excluded from the domain.

Let $f : I \to \mathbb{R}$ be continuous on an interval I, and let $a \in I$. We use the notation

$$x \in I \mapsto \int_a^x f(t)\,dt \in \mathbb{R}$$

to denote the unique antiderivative of f that takes the value 0 at $x = a$.

Let $u : I \to \mathbb{R}$ and v be differentiable functions, and let $a \in I$. Reversing the chain rule gives the following integration formulas for all $x \in I$:

$$\int_a^x u'(t)\,v'(u(t))\,dt = v(u(x)) - v(u(a)),$$
$$\int_a^x u'(t)e^{u(t)}\,dt = e^{u(x)} - e^{u(a)},$$
$$\int_a^x u'(t)u(t)^r\,dt = \frac{u(x)^{r+1}}{r+1} - \frac{u(a)^{r+1}}{r+1} \quad (r \neq -1).$$

If $u : I \to \mathbb{R}$ is such that $u(t) \neq 0$ for all $t \in I$, then

$$\int_a^x \frac{u'(t)}{u(t)}\,dt = \ln|u(x)| - \ln|u(a)|.$$

If $u : I \to \mathbb{R}_+^*$ and $r \in \mathbb{R} \setminus \{-1\}$, then

$$\int_a^x u'(t)u(t)^r\,dt = \frac{u(x)^{r+1}}{r+1} - \frac{u(a)^{r+1}}{r+1}.$$

Linear differential equations with constant coefficients. Let $a \in \mathbb{R}$. We want to characterize the differentiable functions $y : \mathbb{R} \to \mathbb{R}$ that satisfy the linear differential equation

$$y' = ay.$$

Suppose y is a solution of the differential equation above. For all $x \in \mathbb{R}$, we have

$$\frac{d}{dx}\left(e^{-ax}y(x)\right) = e^{-ax}(y'(x) - ay(x)) = 0.$$

Hence, there exists a constant $C \in \mathbb{R}$ such that for all $x \in \mathbb{R}$,

$$y(x) = Ce^{ax}.$$

Conversely, if $y : x \in \mathbb{R} \mapsto Ce^{ax}$ for some constant C, then, for all $x \in \mathbb{R}$, $y'(x) = aCe^{ax} = ay(x)$, so y satisfies the differential equation.

Similarly, for $b \in \mathbb{R}$, if y satisfies the differential equation

$$y' = ay + b,$$

then there exists a constant $C \in \mathbb{R}$ such that

$$\begin{aligned} y(x) &= -\frac{b}{a} + Ce^{ax}, \quad \text{if } a \neq 0, \\ y(x) &= bx + C, \quad \text{if } a = 0. \end{aligned}$$

An initial condition $y(x_0) = y_0$, for some $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}$, determines C uniquely.

14 Integral calculus

Area and the fundamental theorem. Let $a < b$ be real numbers, and let $f : [a, b] \to \mathbb{R}_+$ be a continuous, non-negative function. The integral $\int_a^b f(x)\,dx$ is the area under the graph of f. For continuous $f : [a, b] \to \mathbb{R}$ of either sign,

$$A(x) = \int_a^x f(t)\,dt, \qquad A(a) = 0, \qquad A' = f.$$

Thus, A is the unique antiderivative that vanishes at a, and for any antiderivative F,

$$\int_a^b f(x)\,dx = F(b) - F(a) = [F(x)]_a^b.$$

For signed f, this formula defines the oriented integral. If f is positive and increasing, then for $h > 0$,

$$f(x) \leq \frac{A(x+h) - A(x)}{h} \leq f(x+h).$$

The analogous bound for $h < 0$, together with continuity, yields $A'(x) = f(x)$ by the squeeze theorem. Since two antiderivatives differ by a constant, the evaluation formula $F(b) - F(a)$ follows.

Integral rules and bounds. For continuous functions $f, g : [a, b] \to \mathbb{R}$, real constants $\alpha, \beta \in \mathbb{R}$, and $c \in [a, b]$, we have

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g,$$
$$\int_a^c f + \int_c^b f = \int_a^b f, \qquad \int_b^a f = -\int_a^b f, \qquad \int_a^a f = 0.$$

For $a \le b$, $f \le g$ implies $\int_a^b f \le \int_a^b g$. In particular,

$$m \le f \le M \implies m(b-a) \le \int_a^b f \le M(b-a), \qquad \left|\int_a^b f\right| \le \int_a^b |f|.$$

Integration by parts. For continuously differentiable functions u and v,

$$\int_a^b u(x)v'(x)\,dx = \int_a^b \frac{d}{dx}[u(x)v(x)]\,dx - \int_a^b u'(x)v(x)\,dx.$$

Integrating $(uv)' = u'v + uv'$ proves the formula. Choose u so that differentiating it simplifies the integral. For example, for all $x \in \mathbb{R}$, $\int_0^x te^t\,dt = (x-1)e^x + 1$.

15 Probability

In this document, the sample space Ω is a nonempty finite or countably infinite set (so its outcomes can be listed as $\omega_1, \omega_2, \ldots$). Its elements are possible outcomes. All subsets of Ω are *events*; their collection is the *power set* $\mathcal{P}(\Omega) = \{A : A \subseteq \Omega\}$, which includes $\varnothing$ and Ω. A discrete probability space is a triple $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, where $\mathbb{P}$ is a probability measure on $\mathcal{P}(\Omega)$, i.e., a function $\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$ satisfying the following properties:

$$\mathbb{P}(\varnothing) = 0, \qquad \mathbb{P}(\Omega) = 1, \qquad \mathbb{P}\left(\bigcup_{j\ge1} A_j\right) = \sum_{j\ge1} \mathbb{P}(A_j) \quad \text{for pairwise disjoint events } A_j.$$

The probability of an event $A \subseteq \Omega$ is the sum of the probabilities of the outcomes in A:

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

For a finite sample space with equally likely outcomes, $\mathbb{P}(\{\omega\}) = 1/|\Omega|$ for all $\omega \in \Omega$, so $\mathbb{P}(A) = |A|/|\Omega|$.

Basic probability rules. For any subset $A, B \subseteq \Omega$,

$$\mathbb{P}(\varnothing) = 0, \qquad \mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A),$$
$$A \subseteq B \implies \mathbb{P}(A) \le \mathbb{P}(B),$$
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B),$$
$$\mathbb{P}(A \cup B) \le \mathbb{P}(A) + \mathbb{P}(B).$$

Conditioning and independence. Let $B \subseteq \Omega$ be an event with $\mathbb{P}(B) > 0$. Conditioning on B defines a probability measure by setting

$$A \in \mathcal{P}(\Omega) \mapsto \mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \in [0, 1].$$

This probability measure $\mathbb{P}(\cdot \mid B)$ is called the *conditional probability* given B.

Events A and B are *independent* if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. When $\mathbb{P}(B) > 0$, this is equivalent to $\mathbb{P}(A \mid B) = \mathbb{P}(A)$. Disjointness and independence are different conditions.

Law of total probability Suppose $B_1, \ldots, B_m$ are mutually disjoint events whose union is Ω, and that $\mathbb{P}(B_i) > 0$ for all i. The *law of total probability* states that, for any event $A \subseteq \Omega$,

$$\mathbb{P}(A) = \sum_{i=1}^{m} \mathbb{P}(B_i)\mathbb{P}(A \mid B_i).$$

Random variables A numerical random variable is a function $X : \Omega \to \mathbb{R}$ on a probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. The distribution of X is the probability measure on $\mathcal{P}(\mathbb{R})$ given by

$$\mathbb{P}_X : A \in \mathcal{P}(\mathbb{R}) \mapsto \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}) \in [0, 1].$$

Distribution of a discrete random variable Let X be a discrete random variable, i.e., one that takes values in a countable set $\mathcal{X} = \{x_1, x_2, \ldots\}$. The distribution of X is entirely determined by the probabilities of the singletons in $\mathcal{X}$:

$$\mathbb{P}_X : A \in \mathcal{P}(\mathcal{X}) \mapsto \mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x) = \sum_{x \in A} \mathbb{P}\left(\{\omega \in \Omega : X(\omega) = x\}\right).$$

Independence of random variables Let $X_1, \ldots, X_n$ be random variables defined on the same probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Variables $X_1, \ldots, X_n$ are *mutually independent* if, for every choice of values,

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}\left(\{\omega \in \Omega : X_1(\omega) = x_1, \ldots, X_n(\omega) = x_n\}\right) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i).$$

This is the model arising from separate independent experiments; pairwise independence alone does not express mutual independence.

Expectation, variance and standard deviation. Let $X : \Omega \to X$ be a finite numerical random variable, i.e., $\mathcal{X}$ is a finite subset of $\mathbb{R}$. The expectation of X is defined as

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x\mathbb{P}(X = x).$$

Let $g : \mathcal{X} \to \mathbb{R}$ be a function. Then $g(X)$ is itself a finite numerical random variable, and its expectation is

$$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{X}} g(x)\mathbb{P}(X = x).$$

In particular, the variance of a discrete numerical variable X is defined as

$$\mathrm{Var}(X) = \sum_{x \in \mathcal{X}} (x - \mathbb{E}[X])^2 \mathbb{P}(X = x).$$

By expanding the square, one obtains the alternative formula

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

The standard deviation of X is defined as

$$\sigma_X = \sqrt{\mathrm{Var}(X)}.$$

Expectation is a weighted mean; variance measures squared dispersion, and standard deviation has the same unit as X. For real constants a, b,

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b, \qquad \mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}(X).$$

Let X, Y be finite numerical random variables defined on the same probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, and let $a, b \in \mathbb{R}$. Then

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

For **independent** variables X and Y, the variance of a linear combination is

$$\mathrm{Var}(aX + bY) = a^2 \,\mathrm{Var}(X) + b^2 \,\mathrm{Var}(Y).$$

Bernoulli trials and the binomial distribution. Let $X : \Omega \to \{0, 1\}$ be a random variable defined on a probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Then X is called a Bernoulli random variable, and the probability $p := \mathbb{P}(X = 1)$ is called its success probability. The Bernoulli distribution of X is

$$\mathbb{P}_X : A \in \mathcal{P}(\{0, 1\}) \mapsto \sum_{x \in A} \mathbb{P}(X = x) = \sum_{x \in A} p^x (1 - p)^{1-x}.$$

The distribution is fully determined by the success probability p. When X is a Bernoulli random variable with success probability p, we write $X \sim \mathrm{Bern}(p)$.

Characteristics of a Bernoulli random variable. Let Y be a Bernoulli random variable with success probability p. Then

$$\mathbb{E}[Y] = p, \quad \mathrm{Var}(Y) = p(1 - p).$$

Binomial random variable and binomial distribution. Let $n \in \mathbb{N}^*$ and $p \in [0, 1]$. Let $Y_1, \ldots, Y_n$ be independent Bernoulli random variables with success probability p, defined on a probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Define $X = Y_1 + \cdots + Y_n$. Then X is called a binomial random variable with parameters n and p. The distribution of the random variable $X : \Omega \to [\![0, n]\!]$ is called the binomial distribution with parameters n and p, denoted by $\mathrm{Bin}(n, p)$. $\mathrm{Bin}(n, p)$ is defined by

$$\mathrm{Bin}(n, p) = \mathbb{P}_X : A \in \mathcal{P}([\![0, n]\!]) \mapsto \sum_{x \in A} \binom{n}{x} p^x (1 - p)^{n-x},$$

where, for all $x \in [\![0, n]\!]$, $\binom{n}{x} = \dfrac{n!}{x!(n - x)!} = \dfrac{n(n - 1) \cdots (n - x + 1)}{x!}$ is the binomial coefficient, i.e., the number of subsets of size x in a set of size n.

Characteristics of a binomial random variable. Let $X \sim \mathrm{Bin}(n, p)$ for some $n \in \mathbb{N}^*$ and $p \in [0, 1]$. Then

$$\mathbb{E}[X] = np, \quad \mathrm{Var}(X) = np(1-p), \quad \sigma_X = \sqrt{np(1-p)}.$$

Samples, sums and means. A sample of size $n \geq 1$ is a list $X_1, \ldots, X_n$ of independent, identically distributed random variables (i.i.d.), each with mean μ and variance σ^2. For $S_n = \sum_{i=1}^n X_i$ and $M_n = S_n/n$,

	Expectation	Variance	Standard deviation
S_n	$n\mu$	$n\sigma^2$	$\sigma\sqrt{n}$
M_n	μ	σ^2/n	$\sigma/\sqrt{n}$

For Bernoulli samples, M_n is the observed success frequency.

Chebyshev's inequality and concentration. For every $\delta > 0$ and every finite-valued random variable X with mean μ and variance σ^2,

$$\mathbb{P}(|X - \mu| \geq \delta) \leq \frac{\sigma^2}{\delta^2}, \qquad \mathbb{P}(|M_n - \mu| \geq \delta) \leq \frac{\sigma^2}{n\delta^2}.$$

The second inequality follows from the first and $\mathrm{Var}(M_n) = \sigma^2/n$. For $\sigma > 0$ and $k > 0$, deviations of at least $k\sigma$ have probability at most $1/k^2$. These universal upper bounds may be far above the exact probabilities; replace an upper bound above 1 by 1. For a precision $\delta > 0$ and risk $\alpha \in]0, 1[$, a sufficient sample size is

$$n \geq \left\lceil \frac{\sigma^2}{\alpha\delta^2} \right\rceil, \qquad n \geq 1,$$

ensures $\mathbb{P}(|M_n - \mu| < \delta) \geq 1 - \alpha$. For Bernoulli variables, $\sigma^2 = p(1-p) \leq 1/4$, so $n \geq \lceil 1/(4\alpha\delta^2) \rceil$ suffices even if p is unknown. To require a strict risk bound, take $n > \sigma^2/(\alpha\delta^2)$.

Law of large numbers. For every fixed $\delta > 0$,

$$\lim_{n \to \infty} \mathbb{P}(|M_n - \mu| \geq \delta) = 0.$$