Mathematics Handout

Sets

A **set** is a collection of anything that may be thought of as an individual object. A member x of a set X is called an "element" of X. We write $x \in X$. Sets are denoted by curly brackets.

Examples:

- Coin Toss: $X = \{\text{Heads}, \text{Tails}\}$.
- Natural numbers: $\mathbb{N} = \{1, 2, \ldots\}$

Notation:

- We write $x \notin A$ if x is not contained in the set A.
- For two sets A, B, we write $A \subset B$ if A is entirely contained in B. A is called a strict subset of B
- For two sets A, B, we write $A \subseteq B$ if A is contained but may also be equal to B. A is called a subset of B.

A set X that contains only elements with a certain property Q, we write

$$X = \{x | x \text{ has property } Q\}$$

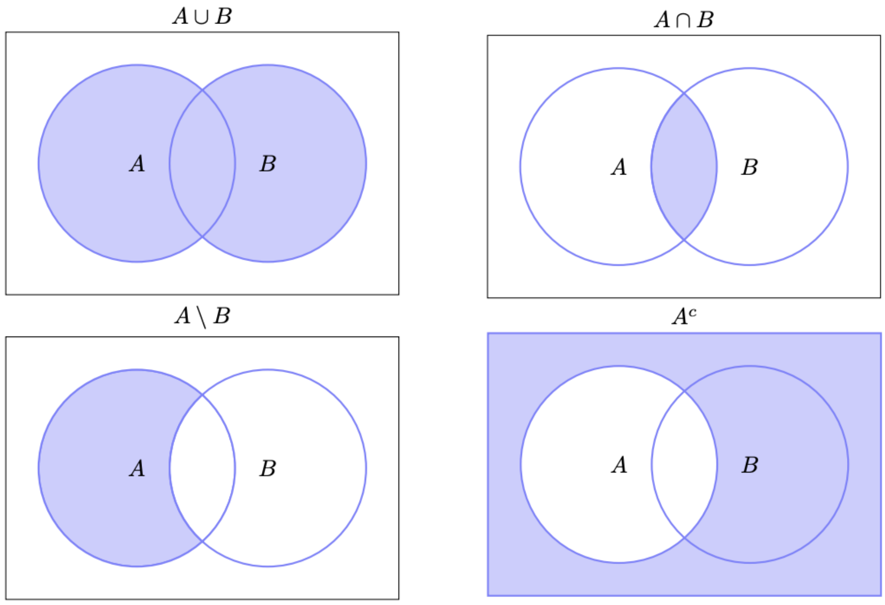
This reads: "X is the set of all objects that have property Q".

Examples:

- Rational numbers: $\mathbb{Q} = \{x | \text{ there exist } a, b \in \mathbb{N} \text{ such that } x = a/b\}$
- Even numbers $= \{x \in \mathbb{N} | \text{ is an even positive integer}\} = \{2, 4, \ldots\}$
- Interval $[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$ interval with bounds a and b

Set operations: For any two sets A, B that are subsets of a larger set U, we define

$$A \cup B := \{x \in U | x \in A \text{ or } x \in B\}$$
$$A \cap B := \{x \in U | x \in A \text{ and } x \in B\}$$
$$A \setminus B := \{x \in U | x \in A \text{ and } x \notin B\}$$
$$A^c := \{x \in U | x \notin A\}$$



Cross Products: Let $S_1, \ldots, S_n$ be arbitrary sets. A "profile" is an ordered list $(x_1, \ldots, x_n)$ where x_1 is taken from S_1, x_2 from S_2, etc. The set of all profiles is called the cross-product of S_1 to S_n, denoted by

$$S_1 \times S_2 \times \ldots \times S_n := \{(x_1, \ldots, x_n) | x_i \in S_i \text{ for all } i = 1, \ldots, n\}$$

Note that a cross-product of sets is itself a set. For a profile x in a cross product of n sets, we may write equivalently $x = (x_1, \ldots, x_n) = (x_j)_{j=1}^n$

Example:

- Set $X_1 = X_2 = [0, 1]$. Then, $X_1 \times X_2 = [0, 1]^2$ (a square).

Maxima and Minima

The maximum of a set $X \subseteq \mathbb{R}$ is the set of largest elements in X. We write

$$\max X := \{x \in X | x \geq x' \text{ for all } x' \in X\}$$

Analogously, the minimum of X is set of the smallest elements in X,

$$\min X := \{x \in X | x \leq x' \text{ for all } x' \in X\}$$

The maximum of a real valued function $u : X \to \mathbb{R}$ is the greatest value in $\{u(x) | x \in X\}$, i.e.,

$$\max_{x \in X} u(x) := \{u(x) | x \in X \text{ and } u(x) \geq u(x') \text{ for all } x' \in X\}$$

The minimum of u is the maximum of $-u$.

The argmax of u over X is the element of X that attains the maximum of u. We write

$$\arg\max_{x \in X} u(x) := \max\{x \in X | u(x) \geq u(x') \text{for all } x' \in X\}$$

The argmin of u is the element of X that attains the maximum of $-u$

"not i notation": Let $s_{-i} := (s_1, ..., s_{i-1}, s_{i+1}, ..., s_n)$ be the profile of all elements from s_1 to s_n except i (short: x not i). Then for all $i \in N$ we write $x = (s_i, s_{-i})$.

Operators

Summation

Let $x_1, x_2, \ldots, x_n$ be any sequence of real numbers. For the sum over these numbers, we write

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$$

Equivalently, if $X = \{x_1, x_2, \ldots, x_n\}$, we may write $\sum_{x \in X} x$.

Product

Let $x_1, x_2, \ldots, x_n$ be any sequence of real numbers. For the sum over these numbers, we write

$$\prod_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n$$

Equivalently, if $X = \{x_1, x_2, \ldots, x_n\}$, we may write $\sum_{x \in X} x$.

Basic Calculus

In mathematics, a limit is the value that a function (or sequence) "approaches" as the input (or index) "approaches" some value. Limits are essential to calculus (and mathematical analysis in general) and are used to define continuity, derivatives, and integrals.

In formulas, a limit L that is attained when the argument x of a real valued function f approaches point a is written as

$$\lim_{x \to a} f(a) = L.$$

Definition (Continuity) A real-valed function f that maps from a set $U \subset \mathbb{R}$ to $\mathbb{R}$ is said to be *continuous* at $a \in U$ if $\lim_{x \to a} f(x) = f(a)$. The function f is called continuous, if it is continuous at any point on U.

Definition (Differentiability) A real-valed function $f : U \subset \mathbb{R} \to \mathbb{R}$ is said to be differentiable at $a \in U$ if its derivative, defined by

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exist.

Occasionally, one writes $df(x)/dx$ for the derivative instead of $f'(x)$ (this is useful when the function under consideration does not have it's own name, like the sum $f(x) + g(x)$.)

Chain rule: Suppose $g : \mathbb{R} \to \mathbb{R}$ is differentiable at x and $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a = g(x)$. Then the derivative of $f(g(x))$ exist and is given by

$$\frac{d}{dx}\big(f(g(x))\big) = f'(g(x))g'(x).$$

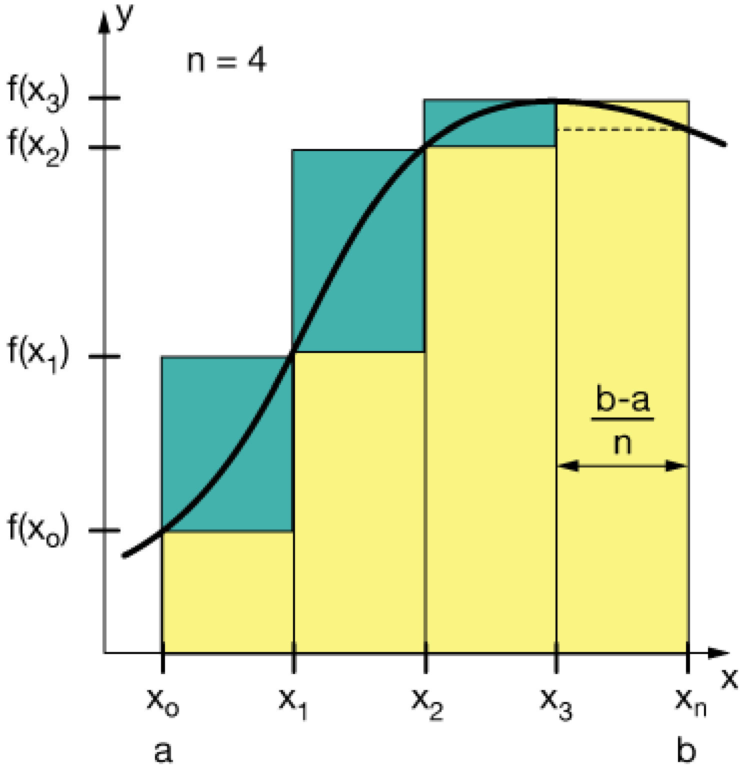
The opposite of taking the derivate of a function, which represents the slope of the graph of that function at any point, we have the integral, which represents the area underneath the graph. The integral is the reverse operation of the derivative in the sense that the integral of the derivative of a function is the function itself.

The following is a basic version of the Riemann integral.

Definition (Definite Integral) Given a function $f(x)$ that is continuous on the interval $[a, b]$, divide the interval into n subintervals of equal width, $\Delta x_n = (b-a)/n$, and from the i-th interval choose a point, $x_{i,n}$ (where $i = 1, \dots, n$). Then the definite integral of $f(x)$ from a to b is defined as

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i,n})\Delta x_n$$

The following diagram illustrates this



To practice solving integrals, here is a good internet source here.

The fact that integration and differentiation are inverse operation is an important result, commonly known as the "Fundamental Law of Calculus" and its formal statement is as follows.

Theorem: Let f be a real-valued function on a closed interval $[a, b]$ and suppose F is a function that has f as it derivative in $[a, b]$:

$$F'(x) = f(x).$$

If f is Riemann integrable on $[a, b]$ then

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

Optimization

We call x^* a (global) maximum of a real-valued function $f : \mathbb{R} \to \mathbb{R}$ if

$$x^* = \max\{x | f(x) \geq f(x') \text{ for all } x' \in \mathbb{R}\}.$$

Theorem Suppose $f : \mathbb{R} \to \mathbb{R}$ is a differentiable real-valued function. If x^* is a global maximum of f, then

$$f'(x^*) = 0.$$

Intuition: at every (interior!!) maximum of a differentiable function, the derivative is zero. It may be the case that the derivative is zero, but at that point f is not a maximum!

Theorem Suppose $f : \mathbb{R} \to \mathbb{R}$ is a twice differentiable and concave real-valued function. If $f'(x^*) = 0$, then x^* is a global maximum of f.

Invertibility

Definition (Invertible Function): A function f that maps from a set $U \in \mathbb{R}$ to another set $Y \in \mathbb{R}$ is called invertible if for any point $y \in Y$, there exists exactly one point $x \in U$

such that $y = f(x)$. If f is invertible, then there exists a function, called the "inverse" of f such that $g(f(x)) = x$ for all $x \in U$.

Intuition: the inverse is simply the reverse operation of f. To find the inverse, you set $y := f(x)$ and then solve the resulting equation for x. A function may fail to be invertible for two reasons:

- some points in $y \in Y$ are not reached by f
- there are several points in U that map to the same point in Y.

Example: Any strictly increasing or strictly decreasing function on $\mathbb{R}$ is invertible.

Inverse Function Theorem: Suppose $f : U \to Y$ is invertible at x and let $g : Y \to U$ be its inverse. If f is differentiable on U, then g is differentiable on Y, and for any $x \in U$, the following identity holds:

$$f'(x) = \frac{1}{g'(f(x))}.$$

Example: Let $f(x) = x^2$ on $U = (0, \infty)$. The derivative is $f'(x) = 2x$. The inverse is $g(y) = \sqrt{y}$ which is defined on $Y = (0, \infty)$. We have $g'(y) = \frac{1}{2\sqrt{y}}$. Then

$$\frac{1}{g'(f(x))} = 2\sqrt{f(x)} = 2x = f'(x)$$

for all $x \in U$.

Probability distributions

A *probability distribution* specifies probabilities of occurrence of different possible outcomes in an experiment.

Definition: A **probability distribution** of a random value x which takes values in a set Ω is a function Pr that assigns a probability $\mathrm{Pr}(E) \in [0, 1]$ to each subset $E \subseteq \Omega$, such that

- $\mathrm{Pr}(\Omega) = 1.$

- $\Pr(\emptyset) = 0$
- $\Pr(E_1 \cup E_2 \cup \ldots) = \Pr(E_1) + \Pr(E_2) + \ldots$

Example (Coin Toss): Suppose a random variable denotes the outcome of a fair coin toss ("the experiment"). Possible realisations: "heads" and "tails". Probability distribution (fair coin): "heads" and tails, each with probability 0.5

Probability distributions are generally divided into two classes.

- A **discrete** probability distribution is for random variables with finite or countable realisations and it is represented by a list of the probabilities, one for each outcome.
- A **continuous** probability distribution is for random variables with realisations on a continuum, and it is described by probability density functions.

A **finite probability distribution** over X is a vector $p = (p_1, p_2, \ldots, p_n)$ of non-negative numbers whose sum is equal to 1.

A **continuous probability distribution** over an interval $[a, b] \subset \mathbb{R}$ is described by a continuous cumulative distribution function (c.d.f.) F, which has the properties:

1. $F(a) = 0$
2. $F(b) = 1$
3. F is (weakly) increasing on $[a, b]$

Here $F(x)$ represents the probability that the random variable is equal to or lower than x. If F is differentiable, its derivate $f = F'$ is called the (probability) density function of the distribution.

Example: The so called "uniform distribution" is a distribution over the unit interval $[0, 1]$ which has the c.d.f. $F(x) = x$ and density $f(x) = 1$.

Suppose the random variable X has the set of possible realisations $(x_1, \ldots, x_n)$ and the probability distribution $p = (p_1, \ldots, p_n)$.

The **expectation** of a finite random variable X on a finite set $\{x_1, \ldots, x_n\}$ wth distribution $(p_1, \ldots, p_n)$ is defined as

$$E[X] = \sum_{i=1}^{n} p_i x_i.$$

The **expectation** of a continuous random variable X over an interval $[a, b]$ with density f is defined as

$$E[X] = \int_a^b f(x)x', dx.$$