

02-680 Module 2

Essentials of Mathematics and Statistics

Aidan Jan

September 2, 2025

Number Sets

Booleans

A boolean is a data type that takes two values: True and False.

- Sometimes we use 1 and 0 instead
- Can also be thought of as yes/no, on/off, etc.
- can be represented using 1 bit.

Integers

An integer is a number with **no fractional** part.

- We use the symbol \mathbb{Z} to represent the set of **all possible integers**
 - There is a lot of them (infinitely many!), so unlike booleans we can't just list all integers.
 - Can be positive or negative or zero. (As opposed to **natural numbers**, represented by \mathbb{N} , and includes only positive integers)
 - The number of bits needed to represent an integer increases with the value.

Rational Numbers

A rational number is one that can be represented by a **ratio** between two integers. That is, it can be written in form $\frac{m}{n}$, where $m, n \in \mathbb{Z}$.

- We use the symbol \mathbb{Q} to represent the set of all **rational numbers**.
- Real numbers that are not rational are called **irrational numbers**.

Real Numbers

A real number is one that can be used to **measure a continuous one-dimensional quantity** such as a length, duration, or temperature.

- Includes all integers, rational numbers, and all numbers "between" rational numbers.
- We use the symbol \mathbb{R} to represent the set of **all real numbers**.
- Typically we need more bits to store real-valued numbers
 - proportional to the value and precision
 - many representations

Complex Numbers

Complex numbers were invented in order to allow for taking the square root of negative numbers

- Each complex number has a **real** and **imaginary** part, and is written as $a + bi$, where $a, b \in \mathbb{R}$
- We denote the set of all **complex numbers** as \mathbb{C} .
- Note that \mathbb{R} is a subset of \mathbb{C} , where all values of b are 0.

Set Theory

Why does set theory matter?

Suppose we define a property like:

"Genes that are significantly upregulated under condition A "

Then we collect all the genes that match this property. This collection is a **set**.

We may also define a set by listing its members. Single-cell transcriptomics may identify:

Cell type set = {T cell, B cell, Macrophage, Fibroblast}

A **sample space** is the **collection** of all possible outcomes of an experiment.

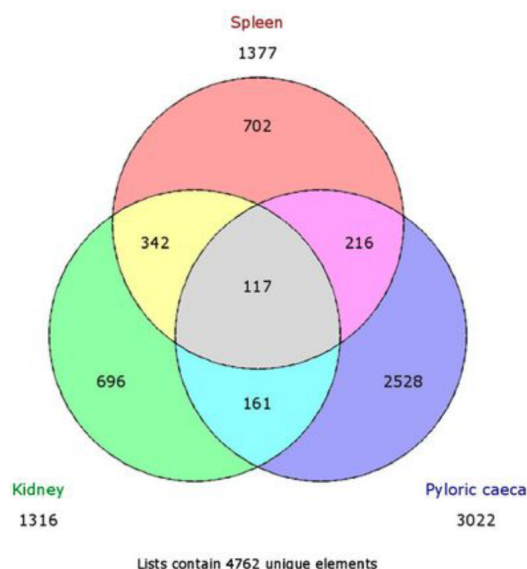
- The sample space of an experiment can be thought of as a set, or collection, of different possible outcomes; and each outcome can be thought of as a point, or an element, in the sample space.

Venn Diagrams

Venn diagrams can be used to compare sets. Each circle represents a set, and members common to multiple sets are included in the overlap of the circles.

Below is an example using differentially expressed genes (DEGs):

Common genes between the three tissues



What is a Set?

A **set** is an **unordered collection** of objects. We have already been talking about these in the abstract:

- Set of all integers \mathbb{Z}
- Set of all real numbers \mathbb{R}

We write a set using braces "{" and "}"

- use capital letters as names (when applicable)
- if we wanted to define the set of all algebraic operators we may use $O = \{+, -, \cdot, \div\}$
- or maybe the set of all prime numbers: $P = \{2, 3, 5, 7, \dots\}$

We have two standard ways of explicitly defining a set.

- First is simple (exhaustive) enumeration:

$$A = \{\text{"Welcome"}, \text{"to"}, \text{"02-680"}\}$$

In this case the set contains 3 elements, each of which is a string.

- The second way is to define a set by abstraction:

$$B = \{x^2 | x = 2 \vee x = 3\}$$

We could say that $4 \in B$, but $16 \notin B$.

- The definition after the $|$ is a statement written in **propositional logic**.

Set Membership

For a set S and an object x , the expression $x \in S$ is true when x is one of the objects in the set S .

- We read $x \in S$ as " x is an element of S ", or " x is in S ".
- There is also the notion of non-membership
 - if the set P is the set of all primes, then $4 \notin P$.
 - that is, 4 is not in P , or 4 is not prime.

Set Cardinality

For a set S , $|S|$ is the number of distinct elements in S .

Often we want to know how large a set is so we can compare sets.

- For the set of all algebraic operators described before $|O| = 4$
- For the set of all possible bit values B , $|B| = 2$.

but sometimes, we can't actually count how many there are. For example,

- $|\mathbb{Z}| = \infty$
- but, we still know that $|\mathbb{Z}| < |\mathbb{R}|$, since we know that \mathbb{Z} is a subset of \mathbb{R} .

In the examples mentioned earlier:

$$A = \{\text{"Welcome"}, \text{"to"}, \text{"02-680"}\}$$

$$B = \{x^2 | x = 2 \vee x = 3\}$$

$$|A| = 3 \text{ and } |B| = 2.$$

Note that if we define another set:

$$C = \{(2 + 2), (2 - 2), (2 \cdot 2), (2 \div 2), (2^2)\}$$

the cardinality $|C| = 3$. It is not 5 since cardinality only includes unique elements.

Set Construction

As described earlier, there are two ways to construct a set. Exhaustive enumeration and set abstraction.

- Note that not all sets need to be assigned a name.
- Remember that sets are **unordered**, so the following sets are equivalent.
 - $\{2 + 2, 2 \cdot 2, 2 \div 2, 2 - 2\}$
 - $\{0, 1, 4\}$
 - $\{4, 0, 1\}$

Set Abstraction

Let U be a set of possible elements called the **universe**. Let $P(x)$ be a condition, also called a **predicate**, that

- for every element $x \in U$
- $P(x)$ is **true** or **false**.

then we can write $\{x \in U : P(x)\}$

- which is all of the objects from $x \in U$
- for which $P(x)$ is **true**.

For example:

- set of even prime numbers: $\{x \in \mathbb{Z} \geq 1 : x \text{ is prime and } x \text{ is even}\}$
- set of primes between 10 and 20: $\{y \in \mathbb{Z} : y \text{ is prime and } 10 \leq y \leq 20\}$
- set of bits: $\{b \in \mathbb{Z} : b^2 = b\}$

For convenience, sometimes we don't use the universe if it's obvious

- this is particularly helpful when the predicate determines the universe
- In that case we could write something like $\{x : P(x)\}$

There is, of course, multiple ways to write these sets as well.

- Two digit perfect squares: $\{n \in \mathbb{Z} : \sqrt{n} \in \mathbb{Z} \text{ and } 10 \leq n \leq 99\}$
- $\{n^2 : n \in \mathbb{Z} \text{ and } 10 \leq n^2 \leq 99\}$
- $\{n^2 : n \in \{4, 5, 6, 7, 8, 9\}\}$ are the same set

Special Sets

There are some sets that we use that are so important we define them on their own, they also usually have **special symbols**.

The **Empty Set**, also called the **null set**, is a set that contains no elements.

$$\emptyset = \{\}$$

Since there are no elements in \emptyset , which we can write as $\forall x : x \notin \emptyset$, or $\nexists x : x \in \emptyset$, $|\emptyset| = 0$.

- Any set with **no elements** can be reduced to \emptyset
- Can also be defined as $\{x : \text{False}\}$ (similar to an if statement with "false" in the condition)

Additionally, we have common sets for groups of numbers.

- \mathbb{N} : natural numbers (positive whole numbers)
- \mathbb{Z} : integers (all whole numbers)
- \mathbb{Q} : rational numbers (numbers representable as fractions)
- \mathbb{R} : real numbers (decimal numbers)
- $\{0, 1\}$: booleans (no special symbol, represents two states)
- Σ : characters (sometimes, "alphabet")

Sometimes, we want to slightly restrict these sets, so sometimes you may see something like \mathbb{Z}^+ , which is equivalent to $x \in \mathbb{Z} | x > 0$, or $\mathbb{Z}^{\geq 0}$, which is equivalent to $x \in \mathbb{Z} | x \geq 0$.

Operators

We have four standard operators used to compare sets, referred to as **set operators**.

- **Union** (\cup): The union of two sets S and T , denoted $S \cup T$, is the set of all elements in either S or T (or both).

$$S \cup T = \{x : x \in S \text{ or } x \in T\}$$

- **Intersection** (\cap): The intersection of two sets S and T , denoted $S \cap T$, is the set of all elements in both S and T .

$$S \cap T = \{x : x \in S \text{ and } x \in T\}$$

- **Set Difference** (\setminus): Similar to numerical subtraction, given two sets S and T , $S \setminus T$ is the set of elements in S but not T .

$$S \setminus T = \{x | x \in S \wedge x \notin T\}$$

- For set difference, order matters. $S \setminus T$ and $T \setminus S$ are different sets.
- Sometimes, set difference is represented using the minus ($-$) symbol instead of backslash (\setminus).

- **Complement** ($\bar{}$): While we can write something like \bar{S} , it is normally written as \bar{S} . This represents the set of elements that are not in S .

$$\bar{S} = \sim S := \{x \in U | x \notin S\}$$

Properties of Union and Intersection

Instead of writing long unions and intersections over multiple sets, we can use indexed notation, just like we do with summation and products.

$$\bigcup_{i=1}^n S_i = S_1 \cup S_2 \cup \cdots \cup S_n$$
$$\bigcap_{i=1}^n S_i = S_1 \cap S_2 \cap \cdots \cap S_n$$

Since a union combines two sets, the following is true:

$$\max\{|S|, |T|\} \leq |S \cup T| \leq |S| + |T|$$

Since an intersection only contains elements present in both sets, the following is true:

$$0 \leq |S \cap T| \leq \min\{|S|, |T|\}$$

Arithmetics and Sets

We previously saw the use of sums and products like

$$\sum_{i=3}^5 2^i = 8 + 16 + 32 = 56$$

but we can do the same using a set. Define $S = \{3, 4, 5\}$, then we can say:

$$\sum_{x \in S} 2^x = 8 + 16 + 36 = 56$$

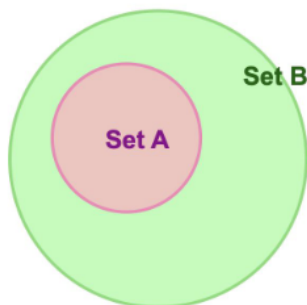
We can do the same with product, max, and min:

$$\prod_{x \in S} x \quad \max_{x \in S} x \quad \min_{x \in S} x$$

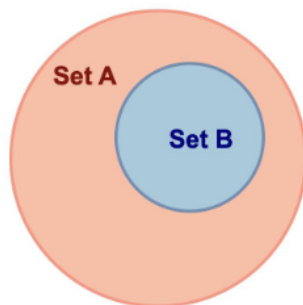
Comparing Sets

We often need to know if one set is contained within another. (notice while similar this is different from saying one is larger than another.)

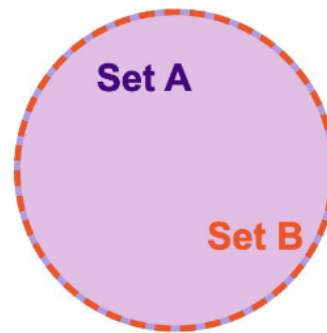
- Subset $A \subseteq B$: A is fully contained within B .
- Superset $A \supseteq B$: B is fully contained within A .
- Equality $A = B$: A contains all elements of, but no more than, B .



$$A \subseteq B$$



$$A \supseteq B$$



$$A = B$$

Subset (\subseteq)

We say one set is a **subset** of another if all of its elements are in both sets.

$$S \subseteq T \Leftrightarrow \forall x \in S : x \in T$$

A set S is a subset of a set T , written $S \subseteq T$, if every $x \in S$ is also an element of T .

- In other words, $S \subseteq T$ is equivalent to $S - T = \{\}$.

Superset (\supseteq)

If a set contains another set, we can say the original is a **superset** of the other.

$$S \supseteq T \Leftrightarrow \forall x \in T : x \in S$$

Equality ($=$)

If two sets contain the same elements, they are equal.

$$S = T \Leftrightarrow (\forall x \in S : x \in T) \wedge (\forall y \in T : y \in S)$$

S is equal to T if and only if every element of S is also in T , and every element of T is also in S .

$$S = T \Leftrightarrow (S \subseteq T) \wedge (S \supseteq T)$$

Proper Subset and Proper Supersets

We also have the proper subset (\subset) and proper superset (\supset) operators, which represent subset and superset (but not equality).

- We define a proper subset as

$$S \subset T \Leftrightarrow (\forall x \in S : x \in T) \wedge (\exists y \in T : y \notin S)$$

- A set S is a proper subset of a set T , written $S \subset T$, if $S \subseteq T$ and $S \neq T$. In other words, $S \subset T$ whenever $S \subseteq T$, but $T \not\subseteq S$.

- We define a proper superset as

$$S \supset T \Leftrightarrow (\forall x \in T : x \in S) \wedge (\exists y \in S : y \notin T)$$

- S is a proper superset of T if and only if every element of T is in S , and there exists at least one element in S that is not in T .

Power Sets

The **power set** represents the collection of all subsets of a given set. It is a set of sets.

$$P(S) := \{T | T \subseteq S\}$$

The **power set** of S is denoted by (S) .

For example, the power set of $0, 1, 2$:

$$\{0, 1, 2\} = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, \{0, 1, 2\}\}$$

Also, a power set of a power set is a set of set of sets...

Cardinality of a Power Set

An interesting fact is that the size of the power set of a set is

$$|P(S)| = 2^{|S|}$$

Think of each element of $P(S)$ as a binary number of length $|S|$. If the i -th bit (or position) is 1, then the i -th element from S is in that subset. Then, each binary number represents a unique subset of S .

Power Set of the Empty Set

The cardinality of the power set of the empty set is 1. ($2^0 = 1$). This is because even though the empty set is empty, it still has one subset, which is itself.

$$P(\emptyset) = \{\emptyset\}$$

Note that $\{\emptyset\} \neq \emptyset$. This is a set containing the empty set, not the empty set itself.