# 02-680 Module 2

Essentials of Mathematics and Statistics

Aidan Jan

August 28, 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:

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

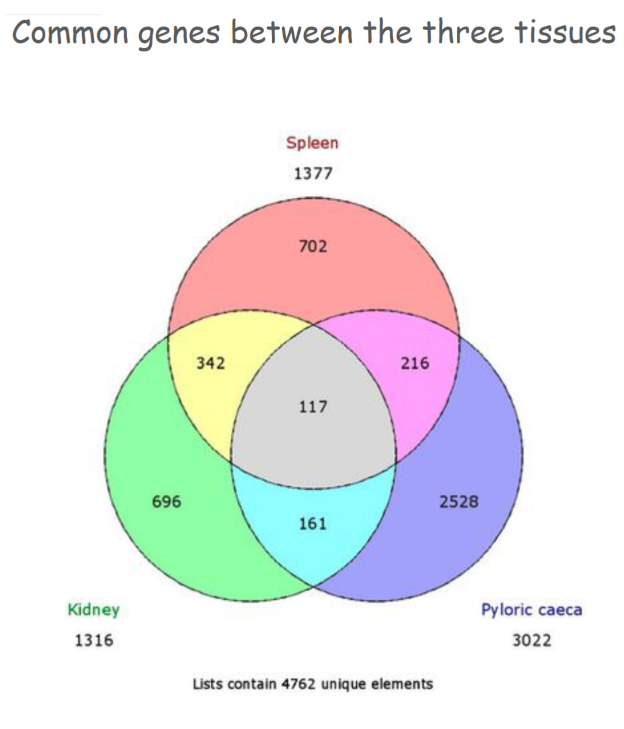
A **sample space** is the **collection** of all possible outcomes of an experiment.

- 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):



## 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:

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

  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:

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

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

$|A| = 3$ 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