Chapter 1

Set theory

1.1 Sets

Definition 1.1. A **set** is a collection of unique objects. The objects in a set are called its **elements**.

Sets are usually denoted by uppercase Roman letters, and one way to write down a set is to list its elements inside curly braces:

$$\begin{aligned} A &= \{1,\ 2,\ 3\} \\ B &= \{a,\ b,\ c\} \\ C &= \left\{a,\ \pi,\ f(x) = x^2,\ \right\} \end{aligned}$$

Literally anything could get thrown into a set. Even another set, or set of sets, or set of set of sets, etc. The only thing we would *not* write is something like $\{1,\ 2,\ 2,\ 3\}$, because the objects in a set must be unique. So no repeats.

Notational Point 1.1. $2 \in A$ is shorthand for "2 is an element of the set A," and $4 \notin A$ is shorthand for "4 is not an element of the set A." The symbol "$\notin$" is like "$\neq$."

1.2 Famous sets

Here are some examples of special sets that we give names to:

$$\begin{aligned} \mathbb{N} &= \{0,\ 1,\ 2,\ ...\} && \text{(natural numbers)} \\ \mathbb{Z} &= \{...,\ -2,\ -1,\ 0,\ 1,\ 2,\ ...\} && \text{(integers)} \\ \mathbb{Q} &= \{\text{"ratios of integers"}\} && \text{(rational numbers)} \\ \mathbb{R} &= \{\text{"}\mathbb{Q} \text{ plus } \pi,\ e,\ \sqrt{2}, \text{ and so on"}\} && \text{(real numbers)} \\ \mathbb{C} &= \{a + bi \ :\ a,\ b \in \mathbb{R},\ i^2 = -1\} && \text{(complex numbers)} \\ \varnothing &= \{\} && \text{(empty set)} \end{aligned}$$

The starting point in mathematical probability is to write down the set S of possible outcomes of a random phenomenon:

$$S = \{H, T\} \qquad \text{(Coin flip)}$$
$$S = \{1, 2, 3, 4, 5, 6\} \qquad \text{(Die roll)}$$

This is called the **sample space**. If we wish to study phenomenon where we are uncertain what will happen, the first thing we must do is specify the realm of possibilities. What *could* happen? We write this down with a set.

1.3 Subsets

Definition 1.2. A is a **subset** of B if all of the elements in A are in B. We denote this $A \subseteq B$, read "A is a subset of B."

Consider

$$A = \{1, 2\}$$
$$B = \{1, 2, 3\}$$
$$C = \{1, 2, 3\}.$$

We see that we have $A \subseteq B$, $A \subseteq C$, $B \subseteq C$, and $C \subseteq B$. The last pair of relations suggest a definition:

Definition 1.3. $B = C$ if $B \subseteq C$ and $C \subseteq B$.

Recalling our famous sets, we have

- $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$;
- $\varnothing \subseteq A$ for any set A.

In a probability context, the sample space S is the set of all possible outcomes of a random phenomenon, and subsets of the sample space are called **events**. These are the things we will try to compute the probabilities of. Recalling the die roll example, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$, and the event that "you roll an odd number" can be represented by the subset $\{1, 3, 5\}$.

1.4 Subsets of the real line

The familiar interval notation is just shorthand for subsets of the real line:

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}$$
$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$
$$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$$
$$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}.$$

These examples provide a model for an alternative method of writing down sets, especially for infinite sets where you cannot just list out all the elements. Consider writing down the set of odd numbers:

$$\begin{aligned} O &= \{1, 3, 5, ...\} && \text{(Faking it)} \\ &= \{n \in \mathbb{N} : \exists k \in \mathbb{N} \text{ s.t. } n = 2k+1\} && \text{(Completely rigorous)} \\ &= \{n \in \mathbb{N} : n \text{ is odd}\} && \text{(Good enough; less pompous)} \end{aligned}$$

The basic template is $\{x \in U : x \text{ satisfies some condition}\}$, where U is the superset that your new subset is contained in.

1.5 Set operations

The three main operations on sets are:

- (**Union**) $A \cup B$ is the set of elements in either A or B (or both!):

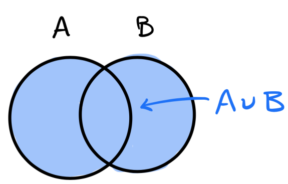

Figure 1.1

- (**Intersection**) $A \cap B$ is the set of elements in both A and B:

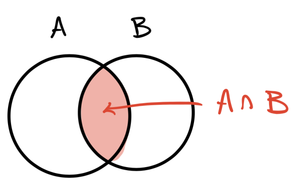

Figure 1.2

  It could be the case that sets A and B do not have any elements in common. Then $A \cap B = \emptyset$, and we call A and B **disjoint**.

- (**Complement**) If U denotes some overarching "reference set," and $A \subseteq U$, then $A^C = \{x \in U : x \notin A\}$. So, the stuff *not in* A:

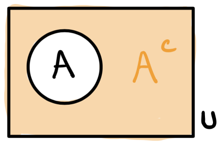

Figure 1.3

1.6 Algebraic properties

The basic set operations obey the following algebraic properties:

$$
\begin{array}{ll}
\text{Commutative} & A \cup B = B \cup A \\
& A \cap B = B \cap A \\
\\
\text{Associative} & (A \cup B) \cup C = A \cup (B \cup C) \\
& (A \cap B) \cap C = A \cap (B \cap C) \\
\\
\text{Distributive} & (A \cup B) \cap C = (A \cap C) \cup (B \cap C) \\
& (A \cap B) \cup C = (A \cup C) \cap (B \cup C) \\
\\
\text{De Morgan's Laws} & (A \cup B)^C = A^C \cap B^C \\
& (A \cap B)^C = A^C \cup B^C.
\end{array}
$$

1.7 Proof techniques

The set identities above are actually theorems that we can prove. Here are some techniques:

1.7.1 Non-rigorous proof: Venn diagrams

Theorem 1.1. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

Demonstration. See Figure 1.4. We draw a picture representing the left-hand side of the equality, we draw a picture representing the right-hand side, and we recognize that in either case we drew the same thing. □

1.7.2 Rigorous proof: element chasing

Theorem 1.2. $(A \cup B)^C = A^C \cap B^C$.

Proof. We prove subsethood in both directions:

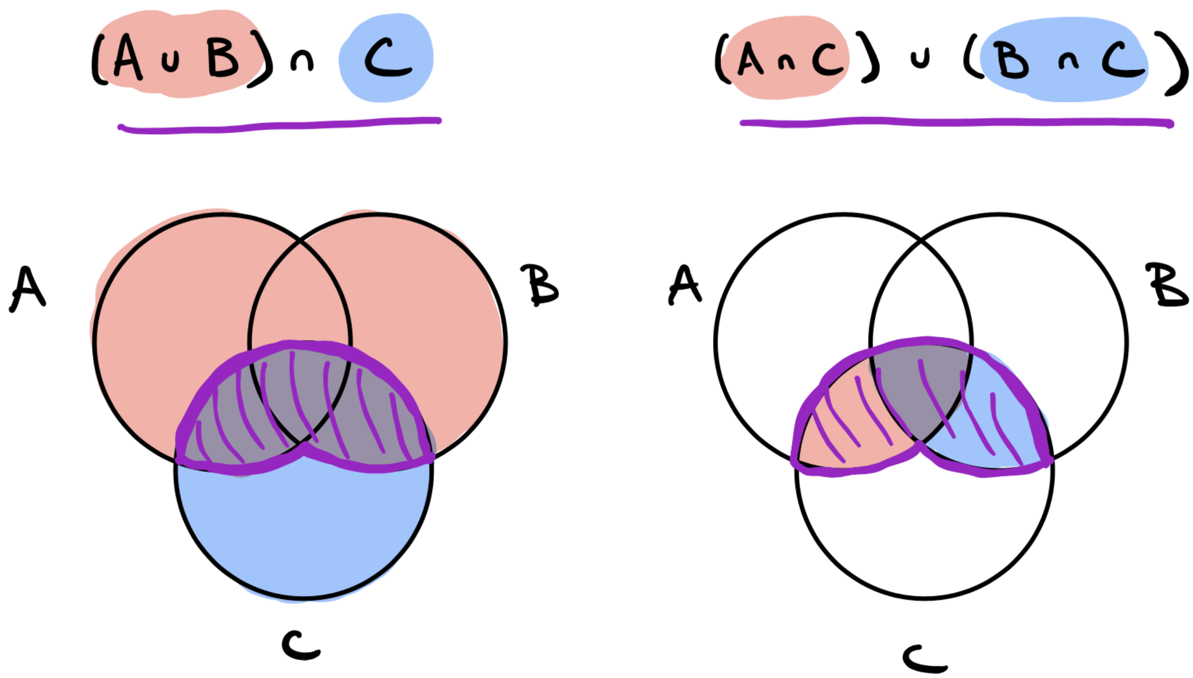

Figure 1.4

- Show $(A \cup B)^C \subseteq A^C \cap B^C$. Let $x \in (A \cup B)^C$ be an arbitrary element. Then

$$\begin{aligned} x \in (A \cup B)^C &\implies x \notin A \cup B \\ &\implies x \notin A \text{ and } x \notin B \\ &\implies x \in A^C \text{ and } x \in B^C \\ &\implies x \in A^C \cap B^C. \end{aligned}$$

  Since x is an arbitrary element, we have the result for all elements in $(A \cup B)^C$, which necessarily implies subsethood.

- Show $A^C \cap B^C \subseteq (A \cup B)^C$. Let $x \in A^C \cap B^C$ be an arbitrary element. Then

$$\begin{aligned} x \in A^C \cap B^C &\implies x \in A^C \text{ and } x \in B^C \\ &\implies x \notin A \text{ and } x \notin B \\ &\implies x \notin A \cup B \\ &\implies x \in (A \cup B)^C. \end{aligned}$$

  Since x is an arbitrary element, we have the result.

$\square$

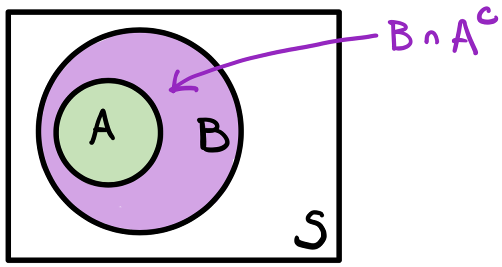

Figure 1.5

1.7.3 Rigorous proof: plain ol' algebra

Going forward, we are free to take for granted identities like the ones in Section 1.6. Armed with them, it is possible to prove subsequent set identities by simply chaining together algebraic manipulations. Consider the situation displayed in Figure 1.5 where $A \subseteq B \subseteq S$. The picture motivates the conjecture that $B = A \cup (B \cap A^C)$, which we can prove algebraically:

Theorem 1.3. If $A \subseteq B$, then $B = A \cup (B \cap A^C)$.

Proof 1. We start from B and arrive at $A \cup (B \cap A^C)$:

$$\begin{aligned} B &= B \cap S \\ &= B \cap (A \cup A^C) && (A \cup A^C = S) \\ &= (B \cap A) \cup (B \cap A^C) && \text{(distributive property)} \\ &= A \cup (B \cap A^C). && (B \cap A = A) \end{aligned}$$

□

Along the way, we used a distributive property, and these two facts:

- $A \cup A^C = S$;
- $C = C \cap D$ whenever $C \subseteq D$.

In fact, we used the second one twice. Strictly speaking, a proper element-chasing proof is required to establish both facts, but it's quick and easy. Once that's done, you can take these things for granted when composing algebraic proofs like the one above (just make sure you indicate what identities you're invoking along the way).

Proof 2. We start with $A \cup (B \cap A^C)$ and arrive at B:

$$\begin{aligned} A \cup (B \cap A^C) &= (A \cup B) \cap (A \cup A^C) && \text{(distributive property)} \\ &= B \cap (A \cup A^C) && (A \cup B = B) \\ &= B \cap S && (A \cup A^C = S) \\ &= B. \end{aligned}$$

□

Here we used the fact that $C \cup D = D$ whenever $C \subseteq D$. Simply for illustrative purposes, I presented two proofs of the same thing, starting from either side of the alleged equality. You of course don't have to do that. One proof is plenty.