

Chapter 1

Mathematical Review

Set theory is now generally accepted as the foundation of modern mathematics, and it plays an instrumental role in the treatment of probability. Unfortunately, a simple description of set theory can lead to paradoxes, while a rigorous axiomatic approach is quite tedious. In these notes, we circumvent these difficulties and assume that the meaning of a set as a collection of objects is intuitively clear. This standpoint is known as naive set theory. We proceed to define relevant notation and operations.

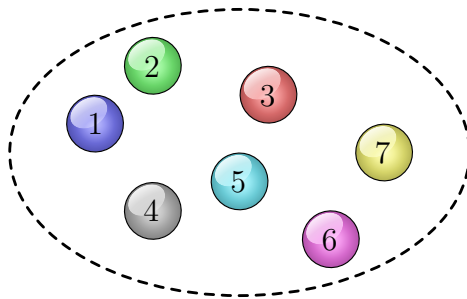


Figure 1.1: This is an illustration of a generic set and its elements.

A *set* is a collection of objects, which are called the *elements* of the set. If an element x belongs to a set S , we express this fact by writing $x \in S$. If x does not belong to S , we write $x \notin S$. We use the equality symbol to denote *logical identity*. For instance, $x = y$ means that x and y are symbols denoting the same object. Similarly, the equation $S = T$ states that S and T are two symbols for the same set. In particular, the sets S and T contain precisely the

same elements. If x and y are different objects then we write $x \neq y$. Also, we can express the fact that S and T are different sets by writing $S \neq T$.

A set S is a *subset* of T if every element of S is also contained in T . We express this relation by writing $S \subset T$. Note that this definition does not require S to be different from T . If $S \subset T$ and S is different from T , then S is a *proper subset* of T , which we indicate by $S \subsetneq T$. Moreover, $S = T$ if and only if $S \subset T$ and $T \subset S$. In fact, this latter statement outlines a methodology to show that two sets are equal.

There are many ways to specify a set. If the set contains only a few elements, one can simply list the objects in the set;

$$S = \{x_1, x_2, x_3\}.$$

The content of a set can also be enumerated whenever S has a countable number of elements,

$$S = \{x_1, x_2, \dots\}.$$

Usually, the way to specify a set is to take some collection T of objects and some property that elements of T may or may not possess, and to form the set consisting of all elements of T having that property. For example, starting with the integers \mathbb{Z} , we can form the subset S consisting of all even numbers,

$$S = \{x \in \mathbb{Z} | x \text{ is an even number}\}.$$

More generally, we denote the set of all elements that satisfy a certain property P by $S = \{x | x \text{ satisfies } P\}$. The braces are to be read as “the set of” while the symbol $|$ stands for the words “such that.”

It is convenient to introduce two special sets. The *empty set*, denoted by \emptyset , is a set that contains no elements. The *universal set* is the collection of all objects of interest in a particular context, and it is represented by Ω . Once a universal set Ω is specified, we need only consider sets that are subsets of Ω . In the context of probability, Ω is often called the *sample space*. The *complement* of a set S , relative to the universal set Ω , is the collection of all objects in Ω that do not belong to S ,

$$S^c = \{x \in \Omega | x \notin S\}.$$

For example, we have $\Omega^c = \emptyset$.

1.1 Elementary Set Operations

The field of probability makes extensive use of set theory. Below, we review elementary set operations, and we establish basic terminology and notation. Consider two sets, S and T .

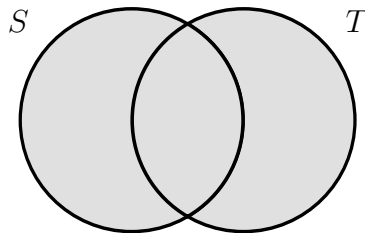


Figure 1.2: This is an abstract representation of two sets, S and T . Each contains the elements in its shaded circle and the overlap corresponds to elements that are in both sets.

The *union* of sets S and T is the collection of all elements that belong to S or T (or both), and it is denoted by $S \cup T$. Formally, we define the union of these two sets by $S \cup T = \{x | x \in S \text{ or } x \in T\}$.

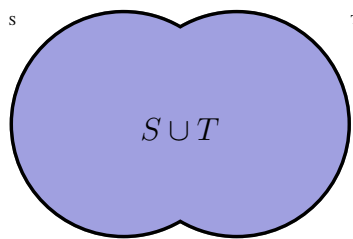


Figure 1.3: The union of sets S and T consists of all elements that are contained in S or T .

The *intersection* of sets S and T is the collection of all elements that belong to both S and T . It is denoted by $S \cap T$, and it can be expressed mathematically as $S \cap T = \{x | x \in S \text{ and } x \in T\}$.

When S and T have no elements in common, we write $S \cap T = \emptyset$. We also express this fact by saying that S and T are *disjoint*. More generally, a collection of sets is said to be disjoint if no two sets have a common element.

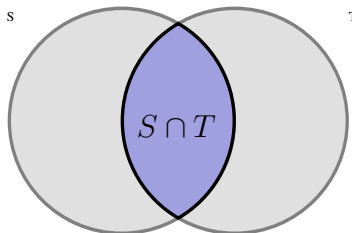


Figure 1.4: The intersection of sets S and T only contains elements that are both in S and T .

A collection of sets is said to form a *partition* of S if the sets in the collection are disjoint and their union is S .

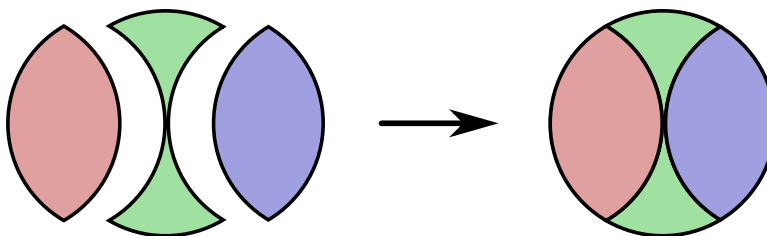


Figure 1.5: A partition of S is a collection of sets that are disjoint and whose union is S .

The *difference* of two sets, denoted by $S - T$, is defined as the set consisting of those elements of S that are not in T , $S - T = \{x | x \in S \text{ and } x \notin T\}$. This set is sometimes called the complement of T relative to S , or the complement of T in S .

So far, we have looked at the definition of the union and the intersection of two sets. We can also form the union or the intersection of arbitrarily many sets. This is defined in a straightforward way,

$$\bigcup_{i \in \mathbb{I}} S_i = \{x | x \in S_i \text{ for some } i \in \mathbb{I}\}$$

$$\bigcap_{i \in \mathbb{I}} S_i = \{x | x \in S_i \text{ for all } i \in \mathbb{I}\}.$$

The index set \mathbb{I} can be finite or infinite.

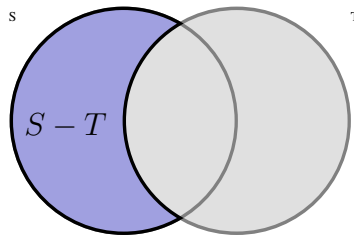


Figure 1.6: The complement of T relative to S contains all the elements of S that are not in T .

1.2 Additional Rules and Properties

Given a collection of sets, it is possible to form new ones by applying elementary set operations to them. As in algebra, one uses parentheses to indicate precedence. For instance, $R \cup (S \cap T)$ denotes the union of two sets R and $S \cap T$, whereas $(R \cup S) \cap T$ represents the intersection of two sets $R \cup S$ and T . The sets thus formed are quite different.

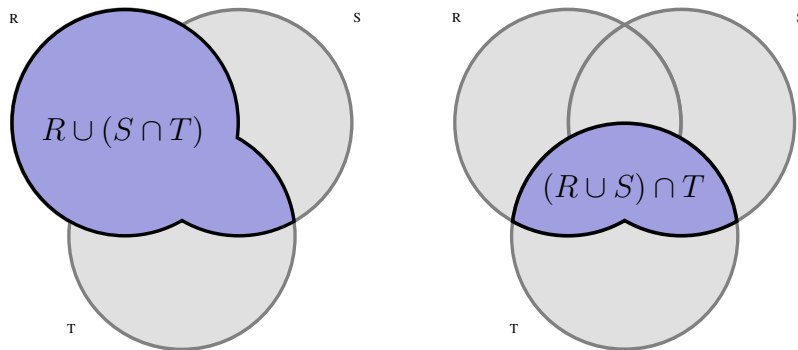


Figure 1.7: The order of set operations is important; parentheses should be employed to specify precedence.

Sometimes, different combinations of operations lead to a same set. For instance, we have the following distributive laws

$$R \cap (S \cup T) = (R \cap S) \cup (R \cap T)$$

$$R \cup (S \cap T) = (R \cup S) \cap (R \cup T).$$

Two particularly useful equivalent combinations of operations are given by

De Morgan's laws, which state that

$$\begin{aligned} R - (S \cup T) &= (R - S) \cap (R - T) \\ R - (S \cap T) &= (R - S) \cup (R - T). \end{aligned}$$

These two laws can also appear in different forms,

$$\begin{aligned} \left(\bigcup_{i \in \mathbb{I}} S_i \right)^c &= \bigcap_{i \in \mathbb{I}} S_i^c \\ \left(\bigcap_{i \in \mathbb{I}} S_i \right)^c &= \bigcup_{i \in \mathbb{I}} S_i^c \end{aligned}$$

when multiple sets are involved. To establish the first equality, suppose that x belongs to $(\bigcup_{i \in \mathbb{I}} S_i)^c$. Then x is not contained in $\bigcup_{i \in \mathbb{I}} S_i$. That is, x is not an element of S_i for any $i \in \mathbb{I}$. This implies that x belongs to S_i^c for all $i \in \mathbb{I}$, and therefore $x \in \bigcap_{i \in \mathbb{I}} S_i^c$. We have shown that $(\bigcup_{i \in \mathbb{I}} S_i)^c \subset \bigcap_{i \in \mathbb{I}} S_i^c$. The converse inclusion is obtained by reversing the above argument. The second law can be obtained in a similar fashion.

1.3 Cartesian Products

There is yet another way to create new sets from existing ones. It involves the notion of an *ordered pair* of objects. Given sets S and T , the *cartesian product* $S \times T$ is the set of all ordered pairs (x, y) for which x is an element of S and y is an element of T , $S \times T = \{(x, y) | x \in S \text{ and } y \in T\}$.

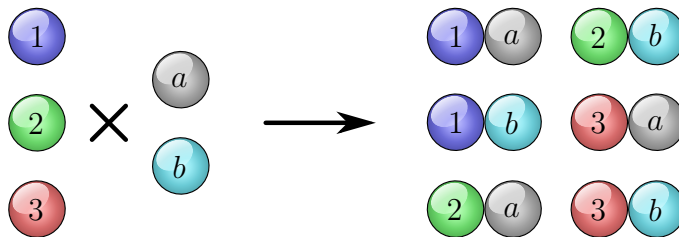


Figure 1.8: Cartesian products can be used to create new sets. In this example, the sets $\{1, 2, 3\}$ and $\{a, b\}$ are employed to create a cartesian product with six elements.

1.4 Functions

A *function* is a special type of relation that assigns exactly one value to each input. A common representation for a function is $f(x) = y$, where f denotes the rule of correspondence. The *domain* of a function is the set over which this function is defined; that is, the collection of arguments that can be used as input. The *codomain* is the set into which all the outputs of the function are constrained to lie. To specify these sets explicitly, the notation

$$f : X \rightarrow Y$$

is frequently used, where X indicates the domain and Y is the codomain. In these notes, we adopt an intuitive point of view, with the understanding that a function maps every argument to a unique value in the codomain. However, a function can be defined formally as a triple (X, Y, F) where F is a structured subset of the Cartesian product $X \times Y$.

Example 1. Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ where $f(x) = x^2$. In this case, the domain \mathbb{R} and the codomain \mathbb{R} are identical. The rule of correspondence for this function is $x \mapsto x^2$, which should be read “ x maps to x^2 .”

Example 2. An interesting function that plays an important role in probability is the indicator function. Suppose S is a subset of the real numbers. We define the indicator function of set S , denoted $\mathbf{1}_S : \mathbb{R} \rightarrow \{0, 1\}$, by

$$\mathbf{1}_S(x) = \begin{cases} 1, & x \in S \\ 0, & x \notin S. \end{cases}$$

In words, the value of the function $\mathbf{1}_S(\cdot)$ indicates whether its argument belongs to S or not. A value of one represents inclusion of the argument in S , whereas a zero signifies exclusion.

The *image* of function is the set of all objects of the form $f(x)$, where x ranges over the elements of X ,

$$\{f(x) \in Y \mid x \in X\}.$$

The image of $f : X \rightarrow Y$ is sometimes denoted by $f(X)$ and it is, in general, a subset of the codomain. Similarly, the *preimage* of a set $T \subset Y$ under $f : X \rightarrow Y$ is the subset of X defined by

$$f^{-1}(T) = \{x \in X | f(x) \in T\}.$$

It may be instructive to point out that the preimage of a singleton set can contain any number of elements,

$$f^{-1}(\{y\}) = \{x \in X | f(x) = y\}.$$

This set, $f^{-1}(\{y\})$, is sometimes called the *level set* of y .

A function is *injective* or *one-to-one* if it preserves distinctness; that is, different elements from the domain never map to a same element in the codomain. Mathematically, the function $f : X \rightarrow Y$ is injective if $f(x_1) = f(x_2)$ implies $x_1 = x_2$ for all $x_1, x_2 \in X$. The function f is *surjective* or *onto* if its image is equal to its codomain. More specifically, a function $f : X \rightarrow Y$ is surjective if and only if, for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$. Finally, a function that is both one-to-one and onto is called a *bijection*. A function f is bijective if and only if its inverse relation f^{-1} is itself a function. In this case, the preimage of a singleton set is necessarily a singleton set. The inverse of f is then represented unambiguously by $f^{-1} : Y \rightarrow X$, with $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$.

1.5 Set Theory and Probability

Set theory provides a rigorous foundation for modern probability and its axiomatic basis. It is employed to describe the laws of probability, give meaning to their implications and answer practical questions. Becoming familiar with basic definitions and set operations is key in understanding the subtleties of probability; it will help overcome its many challenges. A working knowledge of set theory is especially critical when modeling measured quantities and evolving processes that appear random, an invaluable skill for engineers.

Further Reading

1. Bertsekas, D. P., and Tsitsiklis, J. N., *Introduction to Probability*, Athena Scientific, 2002: Section 1.1.
2. Gubner, J. A., *Probability and Random Processes for Electrical and Computer Engineers*, Cambridge, 2006: Section 1.2.

