Chapter 0

Review of set theory

Set theory plays a central role in the theory of probability. Thus, we will open this course with a quick review of those notions of set theory which will be used repeatedly.

0.1 Sets

Throughout this chapter, let $X$ be a set, which we will call the whole space. The elements of $X$ will often be called points. If $A$ is a subset of $X$ and $x$ is a point in $X$, then the notation

$$x \in A$$

means that $x$ belongs to $A$, i.e., it is one of the points in $A$. If $x$ does not belong to $A$, we write

$$x \notin A.$$  

Whenever the identity of the whole space $X$ is clear, we will refer to subsets of $X$ simples as sets.

If $A$ and $B$ are sets, the notations

$$A \subset B \quad \text{or} \quad B \supset A$$

mean that $A$ is a subset of $B$, i.e., every point in $A$ is also a point in $B$. Two sets are called equal if they contain exactly the same set of points, i.e.,

$$A \subset B \quad \text{and} \quad B \subset A.$$
When considering the collection of all subsets of \( X \), we always include the empty set, which is denoted by \( \emptyset \). By definition,

\[
\forall x \in X, \quad x \notin \emptyset.
\]

For every set \( A \), it always holds that

\[
\emptyset \subseteq A \subseteq X.
\]

We will often consider sets of sets. To avoid confusion, we will use the term a \textit{collection} (コレクション) of sets, rather than a set of sets; however remember, a collection is nothing but a set. Thus, if \( C \) is a collection of sets and \( A \) is a set, then \( A \in C \) means that the set \( A \) is one of the elements of the collection of sets \( C \). The notation \( A \subseteq C \), on the other hand, is meaningless.

\textit{Example}: Let \( X \) be a whole space. The collection of all its subsets is denoted by \( 2^X \); it is called the \textit{power set} of \( X \) ( мощность множества). This notation is suggestive, as if \( X \) has finite cardinality \( |X| \), then

\[
|2^X| = 2^{|X|}.
\]

\textbf{Definition 0.1} A sequence \((A_n)\) of sets is called increasing (возрастает) if

\[
A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots.
\]

It is called decreasing (убывает) if

\[
A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots.
\]

In either case it is called monotone.

To denote sets, we will use curly brackets. For example, if \( x, y \in X \), then

\[
\{x, y\}
\]

denotes the set whose only elements are \( x \) and \( y \). If \((x_k)_{k=1}^n\) is a finite sequence of distinct points in \( X \), we will denote the set whose only elements are the elements of that sequence by

\[
\{x_k : k = 1, \ldots, n\}.
\]
If \((x_k)\) is an infinite sequence of distinct points in \(X\), we will denote the set whose only elements are the elements of that sequence by
\[
\{x_k : k \in \mathbb{N}\}.
\]

Suppose that for every \(x \in X\), \(\pi(x)\) is a proposition assuming values either True or False. Then,
\[
\{x \in X : \pi(x)\}
\]
denotes the set of all points in \(X\) for which the proposition \(\pi(x)\) is True. More generally, if \(\pi_1(x), \pi_2(x), \ldots\) is a sequence of propositions regarding the point \(x\), then
\[
\{x \in X : \pi_1(x), \pi_2(x), \ldots\}
\]
denotes the set of all points in \(X\) for which all the propositions \(\pi_k(x)\) are True.

**Example:** Let \(X = \mathbb{N}\) and for \(n \in \mathbb{N}\) let
\[
\pi(n) = \begin{cases} 
  \text{True} & \text{n is even} \\
  \text{False} & \text{n is odd}.
\end{cases}
\]

Then,
\[
\{n \in \mathbb{N} : \pi(n)\}
\]
denotes the set of even integers. In many instances, we will use the more intuitive notation
\[
\{n \in \mathbb{N} : \text{n is even}\}.
\]

As a tautological, yet useful remark: every set \(A\) can be expressed using the curly bracket notation as
\[
A = \{x \in X : x \in A\}.
\]

Finally, for every \(x \in X\), \(\{x\}\) is a subset of \(X\), i.e., an element of \(2^X\); such a subset is called a *singleton* (תֵּכֶּנֶּט). Be careful to distinguish \(x\) from \(\{x\}\).
0.2 Unions and intersections

Let $C$ be a collection of subsets of $X$. The subset of $X$ including all those points that belong to \textit{at least one} set of the collection $C$ is called the \textit{union} ($\cup$) of the collection, and it is denoted by

$$
\bigcup C \quad \text{or} \quad \bigcup \{A : A \in C\}.
$$

In the case where the collection $C$ includes only two sets, e.g., $C = \{A, B\}$, we simply write

$$
\bigcup C = A \cup B.
$$

If

$$
C = \{A_i : i = 1, \ldots, n\}
$$

is a finite collection of sets, we write

$$
\bigcup C = \bigcup_{i=1}^{n} A_i.
$$

If

$$
C = \{A_i : i \in \mathbb{N}\}
$$

is a countable collection of sets, we write

$$
\bigcup C = \bigcup_{i=1}^{\infty} A_i.
$$

Finally, if

$$
C = \{A_\gamma : \gamma \in \Gamma\}
$$

is a non-countable collection of sets, where $\Gamma$ is an index set, we write

$$
\bigcup C = \bigcup_{\gamma \in \Gamma} A_\gamma.
$$

By convention if $C$ is an empty collection of sets, then

$$
\bigcup C = \emptyset.
$$

It is easy to see that whenever $A \subset B$, then

$$
A \cup B = B.
$$
In particular, it always holds that
\[ \emptyset \cup A = A \quad \text{and} \quad A \cup X = X. \]

Moreover, if \( (A_n) \) is an increasing sequence of sets, then
\[ \bigcup_{k=1}^{n} A_k = A_n. \]

Let \( C \) be a collection of subsets of \( X \). The subset of \( X \) including all those points that belong to all sets of the collection \( C \) is called the intersection (union) of the collection, and it is denoted by
\[ \bigcap C \quad \text{or} \quad \bigcap \{A : A \in C\}. \]

In the case where the collection \( C \) includes only two sets, e.g., \( C = \{A, B\} \), we simply write
\[ \bigcap C = A \cap B. \]

If
\[ C = \{A_i : i = 1, \ldots, n\} \]

is a finite collection of sets, we write
\[ \bigcap C = \bigcap_{i=1}^{n} A_i. \]

If
\[ C = \{A_i : i \in \mathbb{N}\} \]

is a countable collection of sets, we write
\[ \bigcap C = \bigcap_{i=1}^{\infty} A_i. \]

Finally, if
\[ C = \{A_\gamma : \gamma \in \Gamma\} \]

is a non-countable collection of sets, where \( \Gamma \) is an index set, we write
\[ \bigcap C = \bigcap_{\gamma \in \Gamma} A_\gamma. \]
By convention if $\mathcal{C}$ is an empty collection of sets then

$$\bigcap \mathcal{C} = X.$$

It is easy to see that whenever $A \subset B$, then

$$A \cap B = A.$$

In particular, it always holds that

$$\emptyset \cap A = \emptyset \quad \text{and} \quad A \cap X = A.$$

Moreover, if $(A_n)$ is a decreasing sequence of sets, then

$$\bigcap_{k=1}^{n} A_k = A_n.$$

**Definition 0.2** Two sets $A$ and $B$ are called disjoint (🌀) if they have no points in common, i.e., $A \cap B = \emptyset$. A disjoint collection of sets is a collection of sets such that every two distinct sets in this collection are disjoint.

Since unions of disjoint collections of sets play an important role in the theory of probability, we will denote the union of a disjoint collection by $\cup$. For example,

$$A \cup B$$

denotes the union of the disjoint sets $A$ and $B$.

Unions and intersections satisfy the following important properties:

1. Both are commutative and associative.
2. Intersections are distributive over unions,

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

3. Unions are distributive over intersections,

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$
0.3 Complements

Let $A$ be a set. It complement ($מֶשֶלֶל$) is the set of points which do not belong to $A$,

$$A^c = \{x : x \notin A\}.$$  

Complementation satisfies the following properties:

1. For all sets $A$,
   $$A \cup A^c = X \quad \text{and} \quad A \cap A^c = \emptyset.$$  
2. For all sets $A$, $(A^c)^c = A$.  
3. $\emptyset^c = X$ and $X^c = \emptyset$.  
4. If $A \subset B$ then $A^c \supset B^c$.  
5. For every (possibly non-countable) collection $C$ of sets,
   $$\left(\bigcup\{A : A \in C\}\right)^c = \bigcap\{A^c : A \in C\},$$
   and
   $$\left(\bigcap\{A : A \in C\}\right)^c = \bigcup\{A^c : A \in C\}.$$  

Let $A$ and $B$ be sets. The difference between $A$ and $B$ is the set of all those points that are in $A$ but not in $B$, namely,

$$A \setminus B = \{x : x \in A \quad \text{and} \quad x \notin B\} = A \cap B^c.$$  

0.4 Limits

Let $(A_n)$ be a sequence of sets. The superior limit ($בָּלָא לַתֶּלֶמֶל$) of this sequence is defined as the set of points that belong to infinitely many of those sets:

$$\limsup_n A_n = \{x : \text{for all } k \text{ there exists an } n \geq k \text{ such that } x \in A_n\}.$$  

If $x \in \limsup_n A_n$ we say that $x \in A_n$ infinitely often ($כְָאָסָף נְשָׁה$).  
The inferior limit ($בָּלָא לְתַתֶּלֶמֶל$) of this sequence is defined as the set of points that belong to all but finitely many of those sets:

$$\liminf_n A_n = \{x : \text{there exists } k \text{ such that } x \in A_n \text{ for all } n \geq k\}.$$
If \( x \in \limsup_n A_n \) we say that \( x \in A_n \) eventually. Note that both superior and inferior limits always exist. Also, it always holds that
\[
\liminf_n A_n \subset \limsup_n A_n.
\]
In the event that the superior limit and the inferior limit of a sequence of sets coincide, we call this set the limit of the sequence, namely,
\[
\lim A_n = \limsup_n A_n = \liminf_n A_n.
\]

By definition, for \( k \in \mathbb{N} \),
\[
\bigcup_{n=k}^\infty A_n = \{ x : \text{there exists an } n \geq k \text{ such that } x \in A_n \},
\]
hence,
\[
\bigcap_{k=1}^\infty \bigcup_{n=k}^\infty A_n = \limsup_n A_n.
\]
Likewise, for \( k \in \mathbb{N} \),
\[
\bigcap_{n=k}^\infty A_n = \{ x : x \in A_n \text{ for all } n \geq k \},
\]
hence,
\[
\bigcup_{k=1}^\infty \bigcap_{n=k}^\infty A_n = \liminf_n A_n.
\]

**Proposition 0.1** If \( (A_n) \) is a monotone sequence of sets, then it has a limit. Specifically, if \( (A_n) \) is increasing then
\[
\lim A_n = \bigcup_{n=1}^\infty A_n.
\]
and if \( (A_n) \) is decreasing then
\[
\lim A_n = \bigcap_{n=1}^\infty A_n.
\]
Proof: Let \((A_n)\) be increasing. Then for every \(k \in \mathbb{N}\),
\[
\bigcup_{n=k}^{\infty} A_n = \bigcup_{n=1}^{\infty} A_n \quad \text{and} \quad \bigcap_{n=k}^{\infty} A_n.
\]
Thus,
\[
\limsup_{n \to \infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} A_n,
\]
and
\[
\liminf_{n \to \infty} A_n = \bigcup_{k=1}^{\infty} A_k.
\]
Let \((A_n)\) be decreasing. Then for every \(k \in \mathbb{N}\),
\[
\bigcap_{n=k}^{\infty} A_n = \bigcap_{n=1}^{\infty} A_n \quad \text{and} \quad \bigcup_{n=k}^{\infty} A_n.
\]
Thus,
\[
\limsup_{n \to \infty} A_n = \bigcap_{k=1}^{\infty} A_k,
\]
and
\[
\liminf_{n \to \infty} A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} A_n.
\]

Example: Let \(X\) be the set of integers, \(\mathbb{N}\), and let
\[
A_n = \begin{cases} 
\{2, 4, 6, \ldots\} & \text{if } n \text{ is even} \\
\{1, 3, 5, \ldots\} & \text{if } n \text{ is odd}.
\end{cases}
\]
Then,
\[
\limsup_{n \to \infty} A_n = X \quad \text{and} \quad \liminf_{n \to \infty} A_n = \emptyset,
\]
i.e., the sequence \((A_n)\) does not have a limit.

Example: Let again \(X = \mathbb{N}\) and let
\[
A_k = \{k^j : j = 0, 1, 2, \ldots\},
\]
namely,
\[
A_1 = \{1\} \quad A_2 = \{1, 2, 4, \ldots\} \quad A_3 = \{1, 3, 9, \ldots\},
\]
etc. Then,
\[
\lim_{k \to \infty} A_k = \{1\}.
\]
0.5 Algebras and \( \sigma \)-algebras of sets

**Definition 0.3** A collection \( \mathcal{C} \) of sets is called an algebra of sets (אלגברה של \( \mathcal{C} \) קבוצות) if it satisfies the following properties:

1. If \( A \in \mathcal{C} \) then \( A^c \in \mathcal{C} \).
2. If \( A, B \in \mathcal{C} \) then \( A \cup B \in \mathcal{C} \).
3. \( X \in \mathcal{C} \).

**Proposition 0.2** Let \( \mathcal{C} \) be an algebra of sets. Then,

1. \( \emptyset \in \mathcal{C} \).
2. If \( A_1, \ldots, A_n \in \mathcal{C} \) then \( \bigcup_{i=1}^{n} A_i \in \mathcal{C} \).
3. If \( A_1, \ldots, A_n \in \mathcal{C} \) then \( \bigcap_{i=1}^{n} A_i \in \mathcal{C} \).
4. Let \( A, B \in \mathcal{C} \). Then their difference \( A \setminus B \) is in \( \mathcal{C} \).

**Proof**: Since \( \mathcal{C} \) is closed under complementation and \( X \in \mathcal{C} \),

\[ \emptyset = X^c \in \mathcal{C}. \]

The second assertion follows by induction. The third assertion follows from the second assertion and de Morgan’s Law,

\[ \bigcap_{i=1}^{n} A_i = \left( \bigcup_{i=1}^{n} A_i^c \right)^c \in \mathcal{C}. \]

The fourth assertion holds because

\[ A \setminus B = A \cap B^c. \]

Thus, an algebra of sets is a collection of sets closed under finitely many set-theoretic operations of union, intersection and complementation.
Example: Let $A$ be a set. Then, the collection of sets

$$\mathcal{C} = \{\emptyset, A, A^c, X\}$$

is an algebra of sets.

Example: The collection of all subsets $2^X$ is an algebra of sets.

Example: Suppose that $X$ is an infinite set. The collection of all finite subsets of $X$ is not an algebra of sets. Neither is the collection of all infinite subsets of $X$.

Definition 0.4 A collection $\mathcal{C}$ of sets is called a $\sigma$-algebra of sets if it is an algebra of sets, and in addition, if $(A_n)$ is a sequence of sets in $\mathcal{C}$, then

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{C}.$$ 

Proposition 0.3 Let $\mathcal{C}$ be a $\sigma$ algebra of sets. If $(A_n)$ is a sequence of sets in $\mathcal{C}$, then

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{C}.$$ 

Proof: This is an immediate consequence of the identity

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c.$$ 

That is, a $\sigma$-algebra of sets is closed with respect to countably many set-theoretic operations.

Example: The collection of all subsets $2^X$ is a $\sigma$-algebra of sets.
Proposition 0.4 Let \( C, \gamma \in \Gamma \) be a (possible non-countable) family of \( \sigma \)-algebras of sets. Then, their intersection,
\[
\mathcal{F} = \bigcap \{ C_\gamma : \gamma \in \Gamma \}
\]
is also a \( \sigma \)-algebra of sets.

Proof: Let \((A_n)\) be a sequence of sets in \( \mathcal{F} \). By definition, \( A_n \in C_\gamma \) for every \( \gamma \in \Gamma \). Since \( C_\gamma \) is a \( \sigma \)-algebra, then
\[
X \in C_\gamma, \quad A_n^c \in C_\gamma \quad \text{and} \quad \bigcup_{n=1}^{\infty} A_n \in C_\gamma.
\]
This holds for every \( \gamma \in \Gamma \), then
\[
X \in \mathcal{F}, \quad A_n^c \in \mathcal{F} \quad \text{and} \quad \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.
\]

Definition 0.5 Let \( C \) be a non-empty collection of sets. The \( \sigma \)-algebra generated by \( C \) is the intersection of all those \( \sigma \)-algebras containing \( C \).
\[
\sigma(C) = \bigcup \{ A : A \text{ is a } \sigma\text{-algebra and } C \subset A \}.
\]
Since \( 2^X \) is a \( \sigma \)-algebra containing \( C \), this intersection is not empty. By Proposition 0.4, \( \sigma(C) \) is a \( \sigma \)-algebra.

Example: Let \( A \) be a set. Then,
\[
\sigma(\{ A \}) = \{ \emptyset, A, A^c, X \}.
\]

Exercise 0.1 Let \( A, B \) be sets. Find \( \sigma(\{A, B\}) \).
0.6 Inverse functions

In calculus, you learned that a function \( f : X \to Y \) is invertible only if it is both one-to-one and onto; in this case, we can define

\[
f^{-1} : Y \to X.
\]

An inverse function, however, can always be defined as a mapping between sets. For every function \( f : X \to Y \) we can define

\[
f^{-1} : 2^Y \to 2^X
\]

by

\[
f^{-1}(A) = \{ x \in X : f(x) \in A \}, \quad A \subset Y.
\]

An important property of the inverse function \( f^{-1} \) is that it preserves (commutes with) set-theoretic operations:

**Proposition 0.5** Let \( f : X \to Y \). Then,

1. For every \( A \subset Y \)
   \[
   (f^{-1}(A))^c = f^{-1}(A^c).
   \]
2. If \( A, B \subset Y \) are disjoint so are \( f^{-1}(A), f^{-1}(B) \subset X \).
3. \( f^{-1}(Y) = X \).
4. If \( A_n \subset Y \) is a sequence of subsets, then
   \[
   f^{-1}\left(\bigcap_{n=1}^{\infty} A_n\right) = \bigcap_{n=1}^{\infty} f^{-1}(A_n).
   \]

**Proof**: Just follow the definitions. For example,

\[
x \in f^{-1}(A) \quad \text{iff} \quad f(x) \in A,
\]

hence

\[
x \in (f^{-1}(A))^c \quad \text{iff} \quad f(x) \notin A \quad \text{iff} \quad f(x) \in A^c \quad \text{iff} \quad x \in f^{-1}(A^c).
\]
The fact that inverse functions commute with set-theoretic operations has the following implication. Let $X$ and $Y$ be sets and let $f : X \to Y$. Let $\mathcal{F} \subset 2^Y$ be a $\sigma$-algebra of subsets of $Y$, then

$$\{ f^{-1}(A) : A \in \mathcal{F} \}$$

is a $\sigma$-algebra of subsets of $X$. 