Selected Topics in Information Theory

Lecture 1

Samir M. Perlaza
INRIA, Lyon, France

Current Version: September 22, 2019. Lyon, France.

Contents

Part I

Theoretical Foundations

1 Algebra of Sets

A *set* is a collection of objects referred to as *elements*. In the following, sets are denoted by caligraphic letters, e.g., $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ and the elements of a given set are listed within braces "$\{\}$". When the number of elements in a set is finite, they can be listed explicitly, e.g., $\{0, 1\}$ is the set of binary digits. Note that the use of ellipses "..." is rather common when the elements follow a particular pattern, e.g., $\{0, 1, 2, 3, \ldots, 9\}$ denotes the set of decimal digits, which contains ten elements. Some particular notations, different from calligraphic letters, are also used to denote sets. For instance,

- $\emptyset = \{\}$ is the empty set, a set without elements;
- $\mathbb{R}$ is the set of all real numbers;
- $\mathbb{N} = \{0, 1, 2, \ldots, \}$ is the set of natural numbers; and
- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots, \}$ is the set of integers.

Given a set, there exists a specific notation that allows specifying whether or not an element belongs to the set. This notation is formalized hereunder.

DEFINITION 1.1 (Membership) An element a that is in $\mathcal{A}$ is said to belong to $\mathcal{A}$ and such membership relation is denoted by $a \in \mathcal{A}$. The opposite is denoted by $a \notin \mathcal{A}$.

From Definition 1.1 and given that both binary numbers 0 and 1 are elements of the set of natural numbers, it follows that $0 \in \mathbb{N}$ and $1 \in \mathbb{N}$.

When the number of elements of a set is infinite the description can be made using the ellipses or using a description of the elements, e.g., $\{x :$ "description of x" $\}$ or $\{x \in \mathcal{O} :$ "description of x" $\}$, where $\mathcal{O}$ is a set that contains all the elements. Other particular notations different from calligraphic letters used to denote sets are:

- $\mathbb{C} = \{a + bi : a \in \mathbb{R}, b \in \mathbb{R}, \text{ and } i = \sqrt{-1}\}$ is the set of complex numbers; and
- $\mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{Z}, \text{ and } q \neq 0\right\}$ is the set of rational numbers.

DEFINITION 1.2 (Cardinality) The cardinality of a set $\mathcal{A}$ is a natural number, denoted by $|\mathcal{A}|$, representing the number of elements in $\mathcal{A}$.

The cardinality of a set can be finite or infinite. The set of natural numbers satisfies $|\mathbb{N}| = \infty$, whereas the set of binary digits $|\{0, 1\}| = 2$. The notion of cardinality implies that the elements of a set can be counted. This holds clearly when the cardinality is finite, nonetheless, even when the cardinality is infinite in some cases the elements of a set can be counted. This observation leads to distinguishing between two types of sets: *countable* and *uncountable* sets.

DEFINITION 1.3 (Countable and uncountable sets) A set $\mathcal{A}$ is said to be countable if and only if there exists an injective function $f : \mathcal{A} \to \mathbb{N}$. When such a function f exists and it is also bijective, the set $\mathcal{A}$ is said to be countably infinite. Otherwise, the set $\mathcal{A}$ is said to be uncountable.

Note that the sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ are countable, more specifically, countably infinite. Alternatively, the sets $\mathbb{R}$ and $\mathbb{C}$ are uncountable. Every finite set is countable and thus, the designation "countable" is preferred instead of the designation "countably finite". Alternatively, every uncountable set is infinite. Nonetheless, sets whose cardinality is infinite might be either countable or uncountable and thus, the designations "uncountable set" and "countably infinite set" is often needed.

The following relations allow comparing two sets.

DEFINITION 1.4 (Comparison) Given two sets $\mathcal{A}$ and $\mathcal{B}$,

- the set $\mathcal{A}$ is said to be a **subset** of $\mathcal{B}$, denoted by $\mathcal{A} \subseteq \mathcal{B}$ or $\mathcal{B} \supseteq \mathcal{A}$, if and only if for all $a \in \mathcal{A}$, it holds that $a \in \mathcal{B}$;
- the set $\mathcal{A}$ is said to be a **proper subset** of $\mathcal{B}$, denoted by $\mathcal{A} \subset \mathcal{B}$ or $\mathcal{B} \supset \mathcal{A}$, if and only if $\mathcal{A} \subseteq \mathcal{B}$ and there exists at least one element $b \in \mathcal{B}$ such that $b \notin \mathcal{A}$;
- the set $\mathcal{A}$ is said to be **identical** to $\mathcal{B}$, denoted by $\mathcal{A} = \mathcal{B}$, if and only if $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{A} \supseteq \mathcal{B}$. The opposite is denoted by $\mathcal{A} \neq \mathcal{B}$.

From Definition 1.4, the following holds:

$$\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}. \tag{1.1}$$

1.1 Basic Operations

An operation between sets generates another set. Some of these operations are described hereunder.

DEFINITION 1.5 (Unions and Intersections) Given two sets $\mathcal{A}$ and $\mathcal{B}$,

- the **union** of the sets $\mathcal{A}$ and $\mathcal{B}$, denoted by $\mathcal{A} \cup \mathcal{B}$, is a set that contains all the elements of $\mathcal{A}$ and $\mathcal{B}$;
- the **intersection** of the sets $\mathcal{A}$ and $\mathcal{B}$, denoted by $\mathcal{A} \cap \mathcal{B}$, is a set that contains the common elements between $\mathcal{A}$ and $\mathcal{B}$;

The union and the intersection of sets possess some properties that are listed hereunder.

THEOREM 1.6 (Properties) *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be some sets. Then,*

- $\mathcal{A} \cup \mathcal{B} = \mathcal{B} \cup \mathcal{A}$ *and* $\mathcal{A} \cap \mathcal{B} = \mathcal{B} \cap \mathcal{A}$ *(Commutative Property)*
- $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C} = (\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C} = \mathcal{A} \cup (\mathcal{B} \cup \mathcal{C})$ *and* $\mathcal{A} \cap \mathcal{B} \cap \mathcal{C} = (\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$ *(Associative Property)*
- $(\mathcal{A} \cup \mathcal{B}) \cap \mathcal{C} = (\mathcal{A} \cap \mathcal{C}) \cup (\mathcal{B} \cap \mathcal{C})$ *and* $(\mathcal{A} \cap \mathcal{B}) \cup \mathcal{C} = (\mathcal{A} \cup \mathcal{C}) \cap (\mathcal{B} \cup \mathcal{C})$ *(Distributive Property)*
- $\mathcal{A} \cap \mathcal{A} = \mathcal{A} \cup \mathcal{A} = \mathcal{A}$. *(Idempotent Property)*

Proof See Homework of Lecture 1. □

Often, operations among sets are performed with respect to a set that contains all the elements involved in the operation. Such a "reference" is known as the universal set and it is often denoted by $\mathcal{O}$. Taking this into account, operations such as the complement of a set can be defined.

DEFINITION 1.7 (Complements) Given two subsets $\mathcal{A}$ and $\mathcal{B}$ of the universal set $\mathcal{O}$,

- the **complement** of the set $\mathcal{A}$, denoted by $\mathcal{A}^{\mathsf{c}}$, is a set that contains all the elements in $\mathcal{O}$ except those in $\mathcal{A}$;
- the **difference** of the sets $\mathcal{A}$ and $\mathcal{B}$, denoted by $\mathcal{A} \setminus \mathcal{B}$, is a set that contains all the elements of $\mathcal{A}$ except those in $\mathcal{B}$.

The simple operations of unions, intersection and complements establish the foundations of the algebra of sets. The following results are easily obtained from Definition 1.4, Definition 1.5 and Definition 1.7.

THEOREM 1.8 *Given a non-empty subset $\mathcal{A}$ of a universal set $\mathcal{O}$, it holds that $a \in \mathcal{A}$ implies $a \notin \mathcal{A}^{\mathsf{c}}$.*

Proof See Homework of Lecture 1. □

THEOREM 1.9 *Given two sets $\mathcal{A}$ and $\mathcal{B}$ of a universal set $\mathcal{O}$, such that $\mathcal{A} \subseteq \mathcal{B}$, it follows that $\mathcal{A}^{\mathsf{c}} \supseteq \mathcal{B}^{\mathsf{c}}$.*

Proof See Homework of Lecture 1. □

THEOREM 1.10 *Given two subsets $\mathcal{A}$ and $\mathcal{B}$ of a universal set $\mathcal{O}$, it holds that*

$$\mathcal{A} \setminus \mathcal{B} = \mathcal{A} \cap \mathcal{B}^{\mathsf{c}}. \tag{1.2}$$

Proof See Homework of Lecture 1. □

DEFINITION 1.11 (Disjoint Sets) Given two sets $\mathcal{A}$ and $\mathcal{B}$, they are said to be disjoint if and only if

$$\mathcal{A} \cap \mathcal{B} = \emptyset. \tag{1.3}$$

THEOREM 1.12 *The union of any two sets $\mathcal{A}$ and $\mathcal{B}$ of a universal set $\mathcal{O}$ can be expressed as the union of two disjoint sets: $\mathcal{A}$ and $\mathcal{A}^{\mathsf{c}} \cap \mathcal{B}$. That is,*

$$\mathcal{A} \cup \mathcal{B} = \mathcal{A} \cup (\mathcal{A}^{\mathsf{c}} \cap \mathcal{B}). \tag{1.4}$$

Proof The proof follows from verifying that $\mathcal{A}$ and $\mathcal{A}^{\mathsf{c}} \cap \mathcal{B}$ are disjoint sets, that is,

$$\begin{aligned}\mathcal{A} \cap (\mathcal{A}^{\mathsf{c}} \cap \mathcal{B}) &= (\mathcal{A} \cap \mathcal{A}^{\mathsf{c}}) \cap \mathcal{B} \\ &= \emptyset \cap \mathcal{B} \\ &= \emptyset, \end{aligned} \tag{1.5}$$

and the fact that,

$$\begin{aligned}\mathcal{A} \cup (\mathcal{A}^{\mathsf{c}} \cap \mathcal{B}) &= (\mathcal{A} \cup \mathcal{A}^{\mathsf{c}}) \cap (\mathcal{A} \cup \mathcal{B}) \\ &= \mathcal{O} \cap (\mathcal{A} \cup \mathcal{B}) \\ &= \mathcal{A} \cup \mathcal{B}, \end{aligned} \tag{1.6}$$

where $\mathcal{O}$ is the universal set containing both $\mathcal{A}$ and $\mathcal{B}$. This completes the proof. □

Given a sequence of subsets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$ of a universal set $\mathcal{O}$, it follows from Theorem 1.12 that their union satisfies

$$\mathcal{A}_1 \cup \mathcal{A}_2 \cup \ldots = \mathcal{A}_1 \cup (\mathcal{A}_1^{\mathsf{c}} \cap \mathcal{A}_2) \cup (\mathcal{A}_1^{\mathsf{c}} \cap \mathcal{A}_2^{\mathsf{c}} \cap \mathcal{A}_3) \cup \ldots, \tag{1.7}$$

where the complement is with respect to the universal set $\mathcal{O}$ and the sets $\mathcal{A}_1$, $(\mathcal{A}_1^{\mathsf{c}} \cap \mathcal{A}_2)$, $(\mathcal{A}_1^{\mathsf{c}} \cap \mathcal{A}_2^{\mathsf{c}} \cap \mathcal{A}_3)$, $\ldots$ are disjoints sets.

1.2 Partitions and Covers

Given a non-empty set, its corresponding subsets might form two particular collections of sets: a partition or a cover.

DEFINITION 1.13 (Partition) Given a non-empty set $\mathcal{A}$, let $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$ be non-empty subsets of $\mathcal{A}$. These subsets are said to form a partition of $\mathcal{A}$ if for all pairs $(i, j) \in \{1, 2, \ldots, n\}^2$, it holds that

- $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$;
- $|\mathcal{B}_i| > 0$; and
- $\displaystyle\bigcup_{t=1}^{n} \mathcal{B}_i = \mathcal{A}$.

The empty set has exactly one partition, which corresponds to the empty set. A trivial partition of a non-empty set $\mathcal{A}$ is the set $\mathcal{A}$ itself. The smallest partition of $\mathcal{A}$, containing the proper subset $\mathcal{B}$ is simply the sets $\mathcal{B}$ and $\mathcal{B}^{\mathsf{c}}$, where the complement is with respect to $\mathcal{A}$.

DEFINITION 1.14 (Cover) Given a non-empty set $\mathcal{A}$, let $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$ be non-empty subsets of $\mathcal{A}$. These subsets are said to form a cover of $\mathcal{A}$ if for all $i \in \{1, 2, \ldots, n\}$, it holds that

- $|\mathcal{B}_i| > 0$; and

- $\bigcup_{t=1}^{n} \mathcal{B}_i = \mathcal{A}$.

Note that some authors define covers by dropping the condition that $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$ are subsets of $\mathcal{A}$ and accept that $\bigcup_{t=1}^{n} \mathcal{B}_i \supseteq \mathcal{A}$. In the case of Definition 1.13 and Definition 1.14, every partition of a set also forms a cover.

1.3 De Morgan's Laws

The following identities were introduced by Augustus de Morgan and play a key role in the algebra of sets.

THEOREM 1.15 (de Morgan Laws) *Let $\mathcal{A}$ and $\mathcal{B}$ be two sets. Then,*

$$\mathcal{A} \cup \mathcal{B} = (\mathcal{A}^{\mathsf{c}} \cap \mathcal{B}^{\mathsf{c}})^{\mathsf{c}} \ \textit{and} \tag{1.8}$$

$$\mathcal{A} \cap \mathcal{B} = (\mathcal{A}^{\mathsf{c}} \cup \mathcal{B}^{\mathsf{c}})^{\mathsf{c}}. \tag{1.9}$$

Proof See Homework of Lecture 1. □

1.4 Cartesian Products

The elements of a set are not necessarily singletons. An element of a set can be in general a tuple. Sets whose elements are tuples can obtained by an operation referred to as Cartesian product.

DEFINITION 1.16 (Cartesian Products) Given two subsets $\mathcal{A}$ and $\mathcal{B}$, their Cartesian products are denoted by $\mathcal{A} \times \mathcal{B}$ and $\mathcal{B} \times \mathcal{A}$ such that

$$\mathcal{A} \times \mathcal{B} \triangleq \{(a, b) : a \in \mathcal{A} \text{ and } b \in \mathcal{B}\} \text{ and} \tag{1.10}$$

$$\mathcal{B} \times \mathcal{A} \triangleq \{(a, b) : a \in \mathcal{B} \text{ and } b \in \mathcal{A}\}. \tag{1.11}$$

Note that the Cartesian product of $\mathcal{A}$ and $\mathcal{B}$ is a set whose elements are ordered pairs and thus, when $\mathcal{A} \neq \mathcal{B}$ it holds that $\mathcal{A} \times \mathcal{B} \neq \mathcal{B} \times \mathcal{A}$. Consider a sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$, the Cartesian product $\mathcal{A}_{i_1} \times \mathcal{A}_{i_2} \times \ldots \times \mathcal{A}_{i_n}$, with $i_t \in \{1, 2, \ldots, n\}$, is

$$\prod_{t=1}^{n} \mathcal{A}_{i_t} = \{(a_1, a_2, \ldots, a_n) : \forall t \in \{1, 2, \ldots, n\}, a_t \in \mathcal{A}_{i_t}\}. \tag{1.12}$$

When all sets are identical, i.e., $\mathcal{A}_s = \mathcal{A}$ for all $s \in \{1, 2, \ldots, n\}$, the notation can be simplified to

$$\mathcal{A}^n = \prod_{t=1}^{n} \mathcal{A} = \{(a_1, a_2, \ldots, a_n) : \forall t \in \{1, 2, \ldots, n\}, a_t \in \mathcal{A}\}. \tag{1.13}$$

The sequences of sets deserve particular attention. The following sections are devoted to some particular classes of sequences of sets and their corresponding limits.

1.5 Monotonic Sequences of Sets

Monotonic sequences of set are essentially either increasing or decreasing sequences of sets. These can be formally defined as follows.

DEFINITION 1.17 (Increasing/Decreasing Sequences) Given a set $\mathcal{A}$ and an infinite sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \ldots$, they are said to form an *increasing sequence*, whose limit is $\mathcal{A}$ if and only if

(a) $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3 \subset \ldots$ and
(b) $\bigcup_{t=1}^{\infty} \mathcal{A}_t = \mathcal{A}$.

This is denoted by $\mathcal{A}_n \uparrow \mathcal{A}$. Alternatively, they are said to form a *decreasing sequence*, whose limit is $\mathcal{A}$ if and only if

(c) $\mathcal{A}_1 \supset \mathcal{A}_2 \supset \mathcal{A}_3 \supset \ldots$ and
(d) $\bigcap_{t=1}^{\infty} \mathcal{A}_t = \mathcal{A}$.

This is denoted by $\mathcal{A}_n \downarrow \mathcal{A}$.

Given a pair $(a, b) \in \mathbb{R}^2$ with $a < b - 2$, the open interval $]a, b[$ can be shown to be the limit of an increasing sequence of closed intervals. For instance, let $\mathcal{A}_n \triangleq \left[a + \frac{1}{n}, b - \frac{1}{n}\right]$ and note that $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3 \subset \ldots$ and $\bigcup_{t=1}^{\infty} \mathcal{A}_t =]a, b[$. Thus,

$$\left[a + \frac{1}{n}, b - \frac{1}{n}\right] \uparrow\ ]a, b[. \tag{1.14}$$

Similarly, given a pair $(a, b) \in \mathbb{R}^2$ with $a < b$, the closed interval $[a, b]$ can be shown to be formed by a decreasing sequence of open intervals with limit $[a, b]$. Assume for instance that $\mathcal{A}_n \triangleq \left]a - \frac{1}{n}, b + \frac{1}{n}\right[$ then, note that $\mathcal{A}_1 \supset \mathcal{A}_2 \supset \mathcal{A}_3 \supset \ldots$ and $\bigcap_{t=1}^{\infty} \mathcal{A}_t = [a, b]$. Thus,

$$\left]a - \frac{1}{n}, b + \frac{1}{n}\right[\downarrow [a, b]. \tag{1.15}$$

A half-open interval can be shown to be the limit of a decreasing sequence of open intervals. For instance, for all pairs $(a, b) \in \mathbb{R}$, with $a < b$, the sets $\mathcal{A}_n \triangleq]a - \frac{1}{n}, b[$ satisfy $\mathcal{A}_n \downarrow [a, b[$.

The De Morgan laws in Theorem 1.15 lead to the following implications.

THEOREM 1.18 *Consider an infinite sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \ldots$. Then,*

(i) If $\mathcal{A}_n \uparrow \mathcal{A}$, then $\mathcal{A}_n^{\mathsf{c}} \downarrow \mathcal{A}^{\mathsf{c}}$; and
(ii) If $\mathcal{A}_n \downarrow \mathcal{A}$, then $\mathcal{A}_n^{\mathsf{c}} \uparrow \mathcal{A}^{\mathsf{c}}$.

Proof To prove (i), note that if $\mathcal{A}_n \uparrow \mathcal{A}$, it follows that $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3 \subset \ldots$ and $\bigcup_{t=1}^{\infty} \mathcal{A}_t = \mathcal{A}$. From Theorem 1.9, the former implies that $\mathcal{A}_1^{\mathsf{c}} \supset \mathcal{A}_2^{\mathsf{c}} \supset \mathcal{A}_3^{\mathsf{c}} \supset \ldots$. Hence, from Theorem 1.15, it follows that

$$\mathcal{A}^{\mathsf{c}} = \left(\bigcup_{n=1}^{\infty} \mathcal{A}_n \right)^{\mathsf{c}} = \bigcap_{n=1}^{\infty} \mathcal{A}_n^{\mathsf{c}}. \tag{1.16}$$

This leads to the conclusion that $\mathcal{A}_n^{\mathsf{c}} \downarrow \mathcal{A}^{\mathsf{c}}$.

To prove (ii), note that if $\mathcal{A}_n \downarrow \mathcal{A}$, it follows that $\mathcal{A}_1 \supset \mathcal{A}_2 \supset \mathcal{A}_3 \supset \ldots$ and $\bigcap_{t=1}^{\infty} \mathcal{A}_t = \mathcal{A}$. From Theorem 1.9, the former implies that $\mathcal{A}_1^{\mathsf{c}} \subset \mathcal{A}_2^{\mathsf{c}} \subset \mathcal{A}_3^{\mathsf{c}} \subset \ldots$, whereas the latter, from Theorem 1.15, implies that

$$\mathcal{A}^{\mathsf{c}} = \left(\bigcap_{n=1}^{\infty} \mathcal{A}_n \right)^{\mathsf{c}} = \bigcup_{n=1}^{\infty} \mathcal{A}_n^{\mathsf{c}}. \tag{1.17}$$

This leads to the conclusion that $\mathcal{A}_n^{\mathsf{c}} \uparrow \mathcal{A}^{\mathsf{c}}$ and completes the proof. □

1.6 Limits of Sequences of Sets

DEFINITION 1.19 Consider a countable sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$. Then, the lower-limit of the sequence is

$$\liminf_n \mathcal{A}_n \triangleq \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} \mathcal{A}_k \tag{1.18}$$

and the upper-limit of the sequence is

$$\limsup_n \mathcal{A}_n \triangleq \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} \mathcal{A}_k. \tag{1.19}$$

Given a countable sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$ and a set $\mathcal{B}$ such that $\mathcal{B} \subseteq \liminf_n \mathcal{A}_n$, then there always exists a $k \in \mathbb{N}$, such that for all $n > k$ it holds that $\mathcal{B} \subset \mathcal{A}_n$. More specifically, $\mathcal{B} \subseteq \liminf_n \mathcal{A}_n$ if and only if for all $n \in \mathbb{N} \setminus \mathcal{N}$ it holds that $\mathcal{B} \subset \mathcal{A}_n$, with $\mathcal{N} \subset \mathbb{N}$, a finite subset.

Alternatively, given a countable sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$ and a set $\mathcal{B}$ such that $\mathcal{B} \subseteq \limsup_n \mathcal{A}_n$, then for all $k \in \mathbb{N}$, there always exists an integer $n > k$ such that $\mathcal{B} \subset \mathcal{A}_n$. More specifically, $\mathcal{B} \subseteq \limsup_n \mathcal{A}_n$ if and only if for all $n \in \mathcal{N}$ it holds that $\mathcal{B} \subset \mathcal{A}_n$, with $\mathcal{N} \subset \mathbb{N}$ an infinite set.

Consider a countable sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$ such that for all $n \in \mathbb{N}$,

$$\mathcal{A}_n = \begin{cases} \left]\frac{-1}{n}, 1\right] & \text{if } n \text{ is odd} \\ \left]-1, \frac{-1}{n}\right] & \text{if } n \text{ is even.} \end{cases} \tag{1.20}$$

Note that for all $m > 0$, the following holds:

$$\bigcup_{n=m}^{\infty} \mathcal{A}_n = \bigcup_{n=0}^{\infty} (\mathcal{A}_{m+2n} \cup \mathcal{A}_{m+2n+1}) \tag{1.21}$$

$$= \left(\bigcup_{n=0}^{\infty} \mathcal{A}_{m+2n} \right) \cup \left(\bigcup_{n=0}^{\infty} \mathcal{A}_{m+2n+1} \right) \text{ and} \tag{1.22}$$

$$\bigcap_{n=m}^{\infty} \mathcal{A}_n = \bigcap_{n=0}^{\infty} (\mathcal{A}_{m+2n} \cap \mathcal{A}_{m+2n+1}) \tag{1.23}$$

$$= \left(\bigcap_{n=0}^{\infty} \mathcal{A}_{m+2n} \right) \cap \left(\bigcap_{n=0}^{\infty} \mathcal{A}_{m+2n+1} \right). \tag{1.24}$$

Then, if m is even,

$$\bigcup_{n=m}^{\infty} \mathcal{A}_n = \left(\bigcup_{n=0}^{\infty} \left] -1, \frac{-1}{m+2n} \right] \right) \cup \left(\bigcup_{n=0}^{\infty} \left] \frac{-1}{m+2n+1}, 1 \right] \right) \tag{1.25}$$

$$= \left] -1, \frac{1}{m} \right] \cup \left] -\frac{1}{m}, 1 \right] \tag{1.26}$$

$$=]-1, 1] \text{ and} \tag{1.27}$$

$$\bigcap_{n=m}^{\infty} \mathcal{A}_n = \left(\bigcap_{n=0}^{\infty} \left] -1, \frac{-1}{m+2n} \right] \right) \cap \left(\bigcap_{n=0}^{\infty} \left] \frac{-1}{m+2n+1}, 1 \right] \right) \tag{1.28}$$

$$=]-1, 0] \cap [0, 1] \tag{1.29}$$

$$= \{0\}. \tag{1.30}$$

Alternatively, if m is odd,

$$\bigcup_{n=m}^{\infty} \mathcal{A}_n = \left(\bigcup_{n=0}^{\infty} \left] \frac{-1}{m+2n}, 1 \right] \right) \cup \left(\bigcup_{n=0}^{\infty} \left] -1, \frac{-1}{m+2n+1}, \right] \right) \tag{1.31}$$

$$= \left] -\frac{1}{m}, 1 \right] \cup \left] -1, -\frac{1}{m} \right] \tag{1.32}$$

$$=]-1, 1] \text{ and} \tag{1.33}$$

$$\bigcap_{n=m}^{\infty} \mathcal{A}_n = \left(\bigcap_{n=0}^{\infty} \left] \frac{-1}{m+2n}, 1 \right] \right) \cap \left(\bigcap_{n=0}^{\infty} \left] -1, \frac{-1}{m+2n+1}, \right] \right) \tag{1.34}$$

$$= [0, 1] \cap]-1, 0] \tag{1.35}$$

$$= \{0\}. \tag{1.36}$$

This implies that for all $m \in \mathbb{N}$,

$$\bigcup_{n=m}^{\infty} \mathcal{A}_n =]-1, 1], \text{ and} \tag{1.37}$$

$$\bigcap_{n=m}^{\infty} \mathcal{A}_n = \{0\}. \tag{1.38}$$

Hence, the upper-limit of the sequence is

$$\limsup_n \mathcal{A}_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \mathcal{A}_n = \bigcap_{m=1}^{\infty}]-1,1] =]-1,1], \tag{1.39}$$

and the lower limit of the sequence is

$$\liminf_n \mathcal{A}_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{A}_n = \bigcup_{m=1}^{\infty} \{0\} = \{0\}. \tag{1.40}$$

In general, the upper and lower limits satisfy the following identities.

THEOREM 1.20 (Complements of Limits) *Consider a countable sequence of sets* $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$. *Then,*

$$\left(\limsup_n \mathcal{A}_n\right)^{\mathsf{c}} = \liminf_n \mathcal{A}_n^{\mathsf{c}}, \text{ and} \tag{1.41}$$

$$\left(\liminf_n \mathcal{A}_n\right)^{\mathsf{c}} = \limsup_n \mathcal{A}_n^{\mathsf{c}}. \tag{1.42}$$

Proof The proof is obtained using the de Morgan identities (Theorem 1.15). That is,

$$\left(\limsup_n \mathcal{A}_n\right)^{\mathsf{c}} = \left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \mathcal{A}_n\right)^{\mathsf{c}} \tag{1.43}$$

$$= \bigcup_{m=1}^{\infty} \left(\bigcup_{n=m}^{\infty} \mathcal{A}_n\right)^{\mathsf{c}} \tag{1.44}$$

$$= \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{A}_n^{\mathsf{c}} \tag{1.45}$$

$$= \liminf_n \mathcal{A}_n^{\mathsf{c}} \tag{1.46}$$

and

$$\left(\liminf_n \mathcal{A}_n\right)^{\mathsf{c}} = \left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{A}_n\right)^{\mathsf{c}} \tag{1.47}$$

$$= \bigcap_{m=1}^{\infty} \left(\bigcap_{n=m}^{\infty} \mathcal{A}_n\right)^{\mathsf{c}} \tag{1.48}$$

$$= \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \mathcal{A}_n^{\mathsf{c}} \tag{1.49}$$

$$= \limsup_n \mathcal{A}_n^{\mathsf{c}}, \tag{1.50}$$

which completes the proof. □

The following Theorem shows that the lower limit is a subset of the upper limit.

THEOREM 1.21 (Inclusions) *Consider a countable sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$. Then,*

$$\liminf_n \mathcal{A}_n \subseteq \limsup_n \mathcal{A}_n. \tag{1.51}$$

Proof Note that if $\mathcal{B} \in \liminf_n \mathcal{A}_n$, it follows that there exits an $n \in \mathbb{N}$ such that for all $k > n$, $\mathcal{B} \in \mathcal{A}_k$. This implies that for all $n \in \mathbb{N}$, there exists least one $k > n$ such that $\mathcal{B} \in \mathcal{A}_k$, which implies that $\mathcal{B} \in \limsup_n \mathcal{A}_n$. This shows that $\liminf_n \mathcal{A}_n \subseteq \limsup_n \mathcal{A}_n$. □

When the upper and lower limit are identical, it is said that a limit exits. The following Theorem introduces a couple of cases in which a limit exits.

THEOREM 1.22 *Consider a countable sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$. Then, if $\mathcal{A}_n \uparrow \mathcal{A}$ or $\mathcal{A}_n \downarrow \mathcal{A}$, it follows that*

$$\liminf_n \mathcal{A}_n = \limsup_n \mathcal{A}_n. \tag{1.52}$$

Proof Consider that $\mathcal{A}_n \uparrow \mathcal{A}$. Then, it follows that $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3 \subset \ldots$, which implies that for all $m > 0$,

(a) $\bigcup_{n=m}^{\infty} \mathcal{A}_n = \mathcal{A}$; and

(b) $\bigcap_{n=m}^{\infty} \mathcal{A}_n = \mathcal{A}_m$.

From (a), it follows that

$$\limsup_n \mathcal{A}_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \mathcal{A}_n = \bigcap_{m=1}^{\infty} \mathcal{A} = \mathcal{A}, \tag{1.53}$$

and from (b), it follows that

$$\liminf_n \mathcal{A}_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \mathcal{A}_n = \bigcup_{m=1}^{\infty} \mathcal{A}_m = \mathcal{A}. \tag{1.54}$$

The proof in the case in which $\mathcal{A}_n \downarrow \mathcal{A}$ follows similar steps, and this completes the proof. □

1.7 Algebraic Structures

1.7.1 Fields and σ-Fields

In this section, a particular class of sets is studied: sets, referred to as fields or algebras, whose elements are also sets and satisfy some particular conditions.

DEFINITION 1.23 (Field) Let $\mathscr{F}$ be a set of subsets of $\mathcal{O}$. Then, $\mathscr{F}$ is said to be a field if it is closed under complementation and finite unions, that is,

- $\mathcal{O} \in \mathscr{F}$;
- $\forall \mathcal{A} \in \mathscr{F}$, $\mathcal{A}^{\mathsf{c}} \in \mathscr{F}$; and
- for all sequences of subsets $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ in $\mathscr{F}$, $\bigcup_{t=1}^{n} \mathcal{A}_t \in \mathscr{F}$, with $n < \infty$.

DEFINITION 1.24 (σ-Field) Let $\mathscr{F}$ be a set of subsets of $\mathcal{O}$. Then, $\mathscr{F}$ is said to be a σ-field (or σ-algebra) if it is closed under complementation and countably infinite unions, that is,

- $\mathcal{O} \in \mathscr{F}$;
- $\forall \mathcal{A} \in \mathscr{F}$, $\mathcal{A}^{\mathsf{c}} \in \mathscr{F}$; and
- for all sequences of subsets $\mathcal{A}_1, \mathcal{A}_2, \ldots$ in $\mathscr{F}$, $\bigcup_{t=1}^{\infty} \mathcal{A}_t \in \mathscr{F}$.

Note that fields and σ-fields are closed under countable and infinitely countable intersections, respectively. Consider for instance the sets $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ in $\mathscr{F}$, with $n \in \mathbb{N}$, then

$$\bigcap_{i=1}^{n} \mathcal{A}_i = \left(\bigcup_{i=1}^{n} \mathcal{A}_i^{\mathsf{c}} \right)^{\mathsf{c}} \in \mathscr{F}, \tag{1.55}$$

due to closeness under complementations and (finite or countably infinite) unions.

The largest σ-field on a set $\mathcal{O}$ is the collection of all possible subsets of $\mathcal{O}$, often this collection is referred to as the *power set* of $\mathcal{O}$ and it is denoted by $2^{\mathcal{O}}$. Alternatively, the smallest σ-field on a set $\mathcal{O}$ is the collection of two sets: $\mathcal{O}$ and the empty set $\emptyset$.

Given a set $\mathcal{A} \in \mathcal{O}$, the smallest σ-field $\mathscr{F}$ on $\mathcal{O}$ containing $\mathcal{A}$ is the collection $\{\mathcal{A},\ \mathcal{A}^{\mathsf{c}},\ \mathcal{O},\ \emptyset\}$. Note that if $\mathscr{G}$ is a σ-field on $\mathcal{O}$ that contains $\mathcal{A}$, then it also contains $\mathcal{A}^{\mathsf{c}}, \mathcal{O}$ and $\emptyset$, and thus, $\mathscr{F} \subset \mathscr{G}$. Hence, the σ-field $\mathscr{F}$ is contained in any σ-field that contains $\mathcal{A}$. That is, $\mathscr{F}$ is the smallest σ-field on $\mathcal{O}$ containing $\mathcal{A}$.

Given a collection $\mathscr{S}$ of subsets of $\mathcal{O}$, the smallest σ-field containing $\mathscr{S}$ is referred to as the σ-field *induced by* $\mathscr{S}$, and it is denoted by $\sigma(\mathscr{S})$.

Given two σ-fields $\mathscr{F}$ and $\mathscr{G}$, with $\mathscr{G} \subset \mathscr{F}$, it is said that $\mathscr{G}$ is a **sub σ-field** of $\mathscr{F}$ and $\mathscr{F}$ is a **refinement** of $\mathscr{F}$.

A σ-field that plays a key role in the following sections is the Borel σ-field.

DEFINITION 1.25 (Borel σ-Field) The smallest σ-field on $\mathbb{R}$ containing all open intervals (a, b) for all pairs $(a, b) \in \mathbb{R}^2$, with $a \leqslant b$, is called the Borel σ-field.

The Borel σ-field in the reals is denoted by $\mathscr{B}(\mathbb{R})$. Alternatively, the Borel σ-field in a specific interval $\mathcal{A} \in \mathscr{B}(\mathbb{R})$ is denoted by

$$\mathscr{B}(\mathcal{A}) \triangleq \{\mathcal{A} \cap \mathcal{B} : \mathcal{B} \in \mathscr{B}(\mathbb{R})\}. \tag{1.56}$$

Note that in Section 1.5, it has been shown that closed intervals or semi-closed intervals can be obtained as the limit of decreasing sequences of open sets. Similarly, by the closeness under complementations, it could be verified that $\mathscr{B}(\mathbb{R})$

also contains the sets $]-\infty, a[$, $]-\infty, a]$, $]b, \infty[$ and $[b, \infty[$. From this perspective, the Borel σ-field also contains all possible closed and semi-closed intervals in $\mathbb{R}$.

THEOREM 1.26 *Let $\mathscr{F}$ and $\mathscr{G}$ be two σ-fields of $\mathcal{O}$. Then, $\mathscr{F} \cap \mathscr{G}$ is also a σ-field of $\mathcal{O}$.*

Proof See Homework of Lecture 1. □

THEOREM 1.27 *Let $\mathscr{F}$ and $\mathscr{G}$ be two σ-fields of $\mathcal{O}$. Then, $\mathscr{F} \cup \mathscr{G}$ is not necessarily a σ-field of $\mathcal{O}$.*

Proof See Homework of Lecture 1. □

DEFINITION 1.28 (Filtration) An infinite sequence of σ-fields $\mathscr{F}_1, \mathscr{F}_2, \ldots$ defined on a set $\mathcal{O}$, such that $\mathscr{F}_i \subset \mathscr{F}_{i+1}$ for all $i \in \mathbb{N} \setminus \{0\}$, is a filtration.

DEFINITION 1.29 (Limit σ-Field) Given a filtration $\mathscr{F}_1, \mathscr{F}_2, \ldots$ defined on a set $\mathcal{O}$, the limit σ-field is $\sigma\left(\{\mathscr{F}_i : i \in \mathbb{N} \setminus \{0\}\right)$.

2 Measure Theory

2.1 Measurable Spaces and Measurable Functions

DEFINITION 2.1 (Measurable Space) Given a set $\mathcal{O}$ and a σ-field $\mathscr{F}$ on $\mathcal{O}$, the pair $(\mathcal{O}, \mathscr{F})$ is said to be a measurable space.

DEFINITION 2.2 (Product of Measurable Spaces) Let $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ be two measurable spaces. The product of these measurable spaces is a measurable space denoted by $(\mathcal{A}, \mathscr{F}) \times (\mathcal{B}, \mathscr{G})$ such that

$$(\mathcal{A}, \mathscr{F}) \times (\mathcal{B}, \mathscr{G}) \triangleq (\mathcal{A} \times \mathcal{B}, \sigma(\mathscr{F} \times \mathscr{G})), \tag{2.1}$$

where $\sigma(\mathscr{F} \times \mathscr{G})$ is the smallest σ-field on $\mathcal{A} \times \mathcal{B}$ containing $\mathscr{F} \times \mathscr{G}$.

DEFINITION 2.3 (Measurable Function) Let $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ be two measurable spaces. The function $f : \mathcal{A} \to \mathcal{B}$ is said to be measurable relative to $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ if for all $\mathcal{G} \in \mathscr{G}$,

$$f^{-1}(\mathcal{G}) \in \mathscr{F}. \tag{2.2}$$

The verification of whether or not a function is measurable might be tedious and thus, the following theorem easies this task in the case in which the target σ-field is induced by a particular collection of sets.

THEOREM 2.4 *Let $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ be two measurable spaces, such that $\mathscr{G} = \sigma(\mathcal{C})$, for some set of subsets $\mathcal{C}$. Then, a function $f : \mathcal{A} \to \mathcal{B}$ is said to be measurable relative to $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ if for all $\mathcal{G} \in \mathcal{C}$,*

$$f^{-1}(\mathcal{G}) \in \mathscr{F}. \tag{2.3}$$

Proof See Homework of Lecture 1. □

THEOREM 2.5 *Consider a measurable function f with respect to $(\mathcal{A}, \mathscr{E})$ and $(\mathcal{B}, \mathscr{F})$. Consider also a measurable function g with respect to $(\mathcal{B}, \mathscr{F})$ and $(\mathcal{C}, \mathscr{G})$. Then, the composition $g \circ f$ is measurable with respect to $(\mathcal{A}, \mathscr{E})$ and $(\mathcal{C}, \mathscr{G})$.*

Proof See Homework of Lecture 1. □

A function f that is measurable with respect to $(\mathcal{A}, \mathscr{F})$ and $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is often called Borel measurable on $(\mathcal{A}, \mathscr{F})$. Nonetheless, when $\mathcal{A} = \mathbb{R}^k$, for some $k > 0$, it simply said that f is Borel measurable and it is assumed that $\mathscr{F} = \mathscr{B}\left(\mathbb{R}^k\right)$.

DEFINITION 2.6 (Isomorphic Measurable Spaces) Two measurable spaces $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ are said to be isomorphic if there exists a bijective function $f : \mathcal{A} \rightarrow \mathcal{B}$ that is measurable with respect to $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$ and its functional inverse f^{-1} is measurable with respect to $(\mathcal{B}, \mathscr{G})$ and $(\mathcal{A}, \mathscr{F})$. If it exists, f is referred to as an isomorphism of $(\mathcal{A}, \mathscr{F})$ and $(\mathcal{B}, \mathscr{G})$.

DEFINITION 2.7 (Standard Measurable Spaces) A measurable space is said to be standard if it is isomorphic to a measurable space $(\mathcal{A}, \mathscr{B}(\mathcal{A}))$, with $\mathcal{A} \in \mathscr{B}(\mathbb{R})$.

Given a measure space $(\mathcal{O}, \mathscr{F})$, the elements of $\mathcal{O}$ are referred to as *outcomes*, whereas those in $\mathscr{F}$ are referred to as *events*. These denominations are often related to the fact that measurable spaces are the building blocks of *probability theory*. From this perspective, given an experiment, the set $\mathcal{O}$ contains all the "outcomes" that might be observed after the experiment. A particular "event" is a subset of "outcomes". More specifically, it is a set in $\mathscr{F}$. In order to determine whether or not an "event" $\mathcal{A} \in \mathscr{F}$ has taken place, all the corresponding outcomes must be verified. That is, all the outcomes of the experiment must be elements of the event of $\mathcal{A}$. The intuition for the requirement of closeness under complementations follows from the fact that if a given event is verifiable so is the same event not taking place. The intuition for closeness under unions stems from the fact that events can be jointly verified.

A refinement of these intuitions leads to the notion of measure, which is reminiscent to the notion of a distance in a metric space, for instance.

2.2 Measures

DEFINITION 2.8 (Measure) A measure on a σ-field $\mathscr{F}$ is a non-negative real-valued function $\mu : \mathscr{F} \rightarrow \mathbb{R}_+$ such that for all finite or countably infinite sequences of disjoint sets $\mathcal{A}_1, \mathcal{A}_2, \ldots$ in $\mathscr{F}$,

$$\mu(\mathcal{A}_1 \cup \mathcal{A}_2 \cup \ldots) = \mu(\mathcal{A}_1) + \mu(\mathcal{A}_2) + \ldots. \tag{2.4}$$

Note that a measure is always positive but it is not necessarily finite. This observation is formalized by the following definitions.

DEFINITION 2.9 (Finite Measure) Given a measure μ on the measurable space $(\mathcal{O}, \mathscr{F})$, it is said to be a finite measure if $\mu(\mathcal{O}) < \infty$.

DEFINITION 2.10 (σ-Finite Measures) Given a measure μ on the measurable space $(\mathcal{O}, \mathscr{F})$, it is said to be a σ-finite if there exists an infinite sequence $\mathcal{A}_1, \mathcal{A}_2, \ldots,$ of sets in $\mathscr{F}$ such that $\bigcup_{t=1}^{\infty} \mathcal{A}_t = \mathcal{O}$ and for all $n \in \mathbb{N} \setminus \{0\}$, $\mu(\mathcal{A}_n) < \infty$.

A particular case of bounded measures is that of probability measures. A measure μ on a σ-field $\mathscr{F}$ of the set $\mathcal{O}$ is said to be a *probability measure* if it satisfies $\mu(\mathcal{O}) = 1$. It is also said to be *concentrated on* $\mathcal{A}$, if $\mu(\mathcal{A}^{\mathsf{c}}) = 0$.

Consider a measurable space $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Then, the **Lebesgue measure** $\mu : \mathscr{B}(\mathbb{R}) \to \mathbb{R}$ is for all $(a, b) \in \mathbb{R}^2$, with $a \leqslant b$,

$$\mu\left(]a, b[\right) = \mu\left(]a, b]\right) = \mu\left([a, b[\right) = \mu\left([a, b]\right) = a - b. \tag{2.5}$$

Often, the Lebesgue measure is associated with the length of an interval. Nonetheless, a Lebesgue measure can also be defined in measurable spaces of the form $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$, with $n > 0$. When, $n = 2$ the Lebesgue measure is associated to the area, when $n = 3$ it is associated to the volume, *et cetera desunt*.

Consider a measurable space $(\mathcal{O}, \mathscr{F})$ and consider the following measures:

- Let a be an element of $\mathcal{O}$, i.e., $a \in \mathcal{O}$. Then, the **Dirac measure** with respect to a, denoted by $\delta_a : \mathscr{F} \to \{0, 1\}$ is

$$\delta_a(\mathcal{A}) = \begin{cases} 1 & \text{if} \quad a \in \mathcal{A} \\ 0 & \text{if} \quad a \notin \mathcal{A}. \end{cases} \tag{2.6}$$

- The function $\mu : \mathscr{F} \to \mathbb{N}$ is a **counting measure** if for all $\mathcal{A} \in \mathscr{F}$

$$\mu(\mathcal{A}) = |\mathcal{A}|\,. \tag{2.7}$$

The following theorem introduces some properties of measures.

THEOREM 2.11 *Let μ be a measure on the σ-field $\mathscr{F}$. Then,*

(a)$\mu(\emptyset) = 0$*;*
(b)$\forall\,(\mathcal{A}, \mathcal{B}) \in \mathscr{F}^2$, $\mu(\mathcal{A} \cup \mathcal{B}) + \mu(\mathcal{A} \cap \mathcal{B}) = \mu(\mathcal{A}) + \mu(\mathcal{B})$*;*
(c)$\forall\,(\mathcal{A}, \mathcal{B}) \in \mathscr{F}^2$, *with* $\mathcal{A} \subset \mathcal{B}$, $\mu(\mathcal{B}) = \mu(\mathcal{A}) + \mu(\mathcal{B} \setminus \mathcal{A})$.

THEOREM 2.12 *Consider a σ-field $\mathscr{F}$ on a set $\mathcal{O}$ and let μ be a measure on $\mathscr{F}$. Consider also an infinite sequence of subsets $\mathcal{A}_1, \mathcal{A}_2, \ldots,$ in $\mathscr{F}$. Then,*

(a)if $\mathcal{A}_n \uparrow \mathcal{A}$, $\lim_{n\to\infty} \mu(\mathcal{A}_n) = \mu(\mathcal{A})$*;*
(b)if $\mathcal{A}_n \downarrow \mathcal{A}$ *and* $\mu(\mathcal{O}) < \infty$, $\lim_{n\to\infty} \mu(\mathcal{A}_n) = \mu(\mathcal{A})$.

DEFINITION 2.13 (Measure Space) Given a measurable space $(\mathcal{O}, \mathscr{F})$ and a measure μ on $\mathscr{F}$, the triplet $(\mathcal{O}, \mathscr{F}, \mu)$ is said to be a measure space.

A measure space $(\mathcal{O}, \mathscr{F}, \mu)$ whose measure μ is a probability measure is called a *probability space*.

2.3 Integration

Given a measure space $(\mathcal{O}, \mathscr{F}, \mu)$ and a Borel measurable function f on $(\mathcal{O}, \mathscr{F})$, the integral of the function f with respect to μ, often referred to as **Lebesgue Integral** is denoted by

$$\int_{\mathcal{O}} f \mathrm{d}\mu, \text{ or } \int_{\mathcal{O}} f(x)\mu(\mathrm{d}x), \text{ or } \int_{\mathcal{O}} f(x)\mathrm{d}\mu(x), \text{ or } \mu f, \text{ or } \mu(f). \tag{2.8}$$

Nonetheless, the notation used in the following would be $\int_{\mathcal{O}} f \mathrm{d}\mu$.

Given an arbitrary function $f : \mathcal{O} \to \mathbb{R}$, Borel measurable with respect to $(\mathcal{O}, \mathscr{F})$, its positive part and **negative part** are **non-negative part** denoted by $f^+ : \mathcal{O} \to \mathbb{R}_+$ and $f^- : \mathcal{O} \to \mathbb{R}_+$, respectively. That is, for all $x \in \mathcal{O}$,

$$f^+(x) = \max\{f(x), 0\} \text{ and} \tag{2.9}$$

$$f^-(x) = -\min\{f(x), 0\}. \tag{2.10}$$

THEOREM 2.14 *Let f be an arbitrary Borel measurable function on $(\mathcal{O}, \mathscr{F})$. Then, the functions f^+ and f^- are both Borel measurable functions on $(\mathcal{O}, \mathscr{F})$.*

Proof See Homework of Lecture 1. □

DEFINITION 2.15 (Simple Functions) Consider a mesurable space $(\mathcal{O}, \mathscr{F})$. Then, a function $f : \mathcal{O} \to \mathbb{R}$ is said to be simple if and only if it is Borel measurable with respect to $(\mathcal{O}, \mathscr{F})$ and it takes finitely many different values.

Note that every simple function $f : \mathcal{O} \to \mathbb{R}$ defined on a measurable space $(\mathcal{O}, \mathscr{F})$ can be written as follows:

$$f(x) = \sum_{t=1}^{m} a_t \mathbb{1}_{\{x \in \mathcal{A}_t\}}, \tag{2.11}$$

where $m \in \mathbb{N}$ is finite, $(a_1, a_2, \ldots, a_m) \in \mathbb{R}$ and $\mathcal{A}_1, \mathcal{A}_2, \ldots \mathcal{A}_m$ are disjoint sets in $\mathscr{F}$.

DEFINITION 2.16 (Increasing and Decreasing Sequences of Functions) Consider a sequence of Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$, denoted by f_1, f_2, f_3, ... The sequence is said to be increasing if for all (m, n) with $m < n$, it holds that for all $x \in \mathcal{O}$,

$$f_m(x) < f_n(x). \tag{2.12}$$

Alternatively, the sequence is said to be decreasing if for all (m, n) with $m < n$, it holds that for all $x \in \mathcal{O}$,

$$f_m(x) > f_n(x). \tag{2.13}$$

The following is a fundamental property of Borel measurable functions in terms of increasing sequences of simple functions.

THEOREM 2.17 *Given a mesurable space $(\mathcal{O}, \mathscr{F})$, any non-negative Borel measurable function f on $(\mathcal{O}, \mathscr{F})$ is the limit of an increasing sequence of non-negative, finite simple functions.*

Proof The proof is by construction. Consider the step functions $f_n : \mathcal{O} \to \mathbb{R}$ defined as follows:

$$f_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if} \quad \frac{k-1}{2^n} \leqslant f(x) < \frac{k}{2^n} \text{ for some } k \in \{1, 2, \ldots, n2^n\} \\ n & \text{if} \quad f(x) \geqslant n \\ 0 & \text{if} \quad f(x) = 0. \end{cases} \tag{2.14}$$

Note that, for all $n \in \mathbb{N}$, f_n is a non-negative quantizer of the function f with resolution $\frac{1}{2^n}$ and span n. Thus, it is simple. Moreover, for all $x \in \mathcal{O}$, $\lim_{n\to\infty} h_n(x) = h(x)$. □

Theorem 2.17 leads to the following more general result.

THEOREM 2.18 *Given a mesurable space $(\mathcal{O}, \mathscr{F})$, any arbitrary Borel measurable function f on $(\mathcal{O}, \mathscr{F})$ is the limit of a sequence of finite simple functions f_1, f_2, ..., such that for all $n \in \mathbb{N} \setminus \{0\}$ and for all $x \in \mathcal{O}$, $|f_n(x)| < |f(x)|$.*

Using these notations, the integral of the function f with respect to μ is defined hereunder.

DEFINITION 2.19 (Lebesgue Integral) Given a measurable space $(\mathcal{O}, \mathscr{F})$ and a Borel measurable function $f : \mathcal{O} \to \mathbb{R}$ with respect to $(\mathcal{O}, \mathscr{F})$, the integral of the function f with respect to μ is defined as follows:

- when f is simple, that is, for all $x \in \mathcal{O}$, $f(x) = \sum_{t=1}^{m} a_t \mathbb{1}_{\{x \in \mathcal{A}_t\}}$, for some finite $m \in \mathbb{N}$, $(a_1, a_2, \ldots, a_n) \in \mathbb{R}$ and $\mathcal{A}_1, \mathcal{A}_2, \ldots \mathcal{A}_m$ disjoint sets in $\mathscr{F}$, then

$$\int_{\mathcal{O}} f \mathrm{d}\mu \triangleq \sum_{t=1}^{m} a_t \mu(\mathcal{A}_t); \tag{2.15}$$

- when f is non-negative Borel measurable on $(\mathcal{O}, \mathscr{F})$, then,

$$\int_{\mathcal{O}} f \mathrm{d}\mu \triangleq \sup \left\{ \int_{\mathcal{O}} g \mathrm{d}\mu : g \text{ is simple and } \forall x \in \mathcal{O}, 0 \leqslant g(x) \leqslant f(x) \right\}; \text{and} \tag{2.16}$$

- when f is an arbitrary Borel measurable function on $(\mathcal{O}, \mathscr{F})$, then,

$$\int_{\mathcal{O}} f \mathrm{d}\mu \triangleq \int_{\mathcal{O}} f^{+} \mathrm{d}\mu - \int_{\mathcal{O}} f^{-} \mathrm{d}\mu. \tag{2.17}$$

The Lebesgue integral $\int_{\mathcal{O}} f \mathrm{d}\mu$ of an arbitrary function f, Borel measurable with respect to $(\mathcal{O}, \mathscr{F})$ is said to exist if both $\int_{\mathcal{O}} f^{+} \mathrm{d}\mu < \infty$ and $\int_{\mathcal{O}} f^{-} \mathrm{d}\mu < \infty$. This condition is required because otherwise the operation in (2.17) would be $\infty - \infty$, which is undetermined. This immediately implies that if the function f is either non-negative or non-positive, then the Lebesgue integral always exists due to the fact that one of its positive or negative parts is always zero. In this particular case, the integral always exists but it can be either ∞ or $-\infty$.

A condition that ensures the existence of the Lebesgue integral $\int_{\mathcal{O}} f \mathrm{d}\mu$ is $\int_{\mathcal{O}} |f| \mathrm{d}\mu < \infty$, which is known as the **absolute integrability** condition on f with respect to μ.

When the integral of f with respect to the measure μ is over a particular set $\mathcal{A} \in \mathscr{F}$ other than $\mathcal{O}$, i.e., $\displaystyle\int_{\mathcal{A}} f \mathrm{d}\mu$, it follows that:

$$\int_{\mathcal{A}} f \mathrm{d}\mu = \int_{\mathcal{O}} g \mathrm{d}\mu, \tag{2.18}$$

where $g : \mathcal{O} \to \mathbb{R}$ is for all $x \in \mathcal{O}$, $g(x) = f(x)\mathbb{1}_{\{x \in \mathcal{A}\}}$.

The following theorem compares the integrals $\displaystyle\int_{\mathcal{A}} f \mathrm{d}\mu$ and $\displaystyle\int_{\mathcal{O}} f \mathrm{d}\mu$.

THEOREM 2.20 *Let f be a non-negative Borel measurable function with respect to $(\mathcal{O}, \mathscr{F})$ and μ a measure on $(\mathcal{O}, \mathscr{F})$. Then, the integral $\int_{\mathcal{A}} f \mathrm{d}\mu$ exists for all $\mathcal{A} \in \mathcal{O}$ and*

$$\int_{\mathcal{A}} f \mathrm{d}\mu \leqslant \int_{\mathcal{O}} f \mathrm{d}\mu. \tag{2.19}$$

Proof Assume that f is a non-negative simple function with the form in (2.15). Then, let $g : \mathcal{O} \to \mathbb{R}$ be for all $x \in \mathcal{O}$,

$$g(x) = f(x)\mathbb{1}_{\{x \in \mathcal{A}\}} = \sum_{t=1}^{m} a_t \mathbb{1}_{\{x \in \mathcal{A}_t\}} \mathbb{1}_{\{x \in \mathcal{A}\}} = \sum_{t=1}^{m} a_t \mathbb{1}_{\{x \in \mathcal{A}_t \cap \mathcal{A}\}}, \tag{2.20}$$

which is also a simple function. Hence,

$$\int_{\mathcal{A}} f \mathrm{d}\mu = \int_{\mathcal{A}} g \mathrm{d}\mu = \sum_{t=1}^{m} a_t \mu(\mathcal{A}_t \cap \mathcal{A}) \leqslant \sum_{t=1}^{m} a_t \mathbb{1}_{\{x \in \mathcal{A}_t\}} = \int_{\mathcal{O}} f \mathrm{d}\mu. \tag{2.21}$$

The proof continues with the analysis of non-negative functions (other than simple functions) using the same argument. □

This integral, which holds for arbitrary Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$, is referred to as an indefinite integral. This is formalized hereunder.

DEFINITION 2.21 Given measurable space $(\mathcal{O}, \mathscr{F})$ a set $\mathcal{A} \in \mathscr{F}$ and an arbitrary Borel measurable function $f : \mathcal{O} \to \mathbb{R}$ with respect to $(\mathcal{O}, \mathscr{F})$, the integral

$$\int_{\mathcal{A}} f \mathrm{d}\mu \tag{2.22}$$

is referred to as the indefinite integral of f with respect to μ on $\mathcal{A}$.

The denomination of indefinite integral stems from the fact that if $\mathcal{O} = \mathbb{R}$, $\mathscr{F} = \mathscr{B}(\mathbb{R})$ and μ is the Lebesgue measure in (2.22), then given an interval $\mathcal{A} = [a, x]$, it follows that

$$\int_{\mathcal{A}} f \mathrm{d}\mu = \int_{a}^{x} f(t)\mathrm{d}t, \tag{2.23}$$

where the integral on the right hand side of (2.23) is the Riemman indefinite integral.

Lebesgue integrals exhibit several properties that are reminiscent to those of Riemman integrals. The following theorem highlights one of those properties.

THEOREM 2.22 (Integration of a weighted function) *Let f be a Borel measurable function with respect to $(\mathcal{O}, \mathscr{F})$ and μ a measure on $(\mathcal{O}, \mathscr{F})$. Then, if $\int_{\mathcal{O}} f \mathrm{d}\mu$ exists, it holds that for all $c \in \mathbb{R}$, $\int_{\mathcal{O}} c f \mathrm{d}\mu$ exists and*

$$\int_{\mathcal{O}} c f \mathrm{d}\mu = c \int_{\mathcal{O}} f \mathrm{d}\mu. \tag{2.24}$$

THEOREM 2.23 *Let f and g be two Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$ and μ a measure on $(\mathcal{O}, \mathscr{F})$. Then, if for all $x \in \mathcal{O}$, $f(x) \geqslant g(x)$, it follows that $\int_{\mathcal{O}} f \mathrm{d}\mu \geqslant \int_{\mathcal{O}} g \mathrm{d}\mu$, given that both integrals exist.*

Proof See Homework of Lecture 2. □

THEOREM 2.24 *Let f be a Borel measurable function with respect to $(\mathcal{O}, \mathscr{F})$ and μ a measure on $(\mathcal{O}, \mathscr{F})$. Then, if $\int_{\mathcal{O}} f \mathrm{d}\mu$ exists, it holds that*

$$\left| \int_{\mathcal{O}} f \mathrm{d}\mu \right| \leqslant \int_{\mathcal{O}} |f| \, \mathrm{d}\mu. \tag{2.25}$$

Proof See Homework of Lecture 2. □

THEOREM 2.25 (Additivity of Integrals) *Let f and g be two Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$ and μ a measure on $(\mathcal{O}, \mathscr{F})$. Then, when the integrals $\int_{\mathcal{O}} f \mathrm{d}\mu$, $\int_{\mathcal{O}} g \mathrm{d}\mu$ and $\int_{\mathcal{O}} f + g \mathrm{d}\mu$ exist, it holds that*

$$\int_{\mathcal{O}} f + g \, \mathrm{d}\mu = \int_{\mathcal{O}} f \mathrm{d}\mu + \int_{\mathcal{O}} g \mathrm{d}\mu. \tag{2.26}$$

Proof See Homework of Lecture 2. □

Using the notion of increasing functions (Definition 2.16), the monotone convergence theorem can be stated as follows.

THEOREM 2.26 (Monotone Convergence) *Let $(\mathcal{O}, \mathscr{F}, \mu)$ be a measure space and f be a Borel measurable function with respect to $(\mathcal{O}, \mathscr{F})$. Let also f_1, f_2, f_3, $\ldots$ be an increasing sequence of Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$. Assume that for all $x \in \mathcal{O}$,*

$$\lim_{t \to \infty} f_t(x) = f(x). \tag{2.27}$$

Then, if follows that

$$\lim_{t \to \infty} \int_{\mathcal{O}} f_t(x) \mathrm{d}\mu = \int_{\mathcal{O}} f(x) \mathrm{d}\mu. \tag{2.28}$$

Proof See Homework of Lecture 2. □

THEOREM 2.27 (Fatou's Lemma) *Let $(\mathcal{O}, \mathscr{F}, \mu)$ be a measure space and f and f_1, f_2, f_3, $\ldots$ be Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$. Then,*

- *when for all $n \in \mathbb{N} \setminus \{0\}$ and for all $x \in \mathcal{O}$, $f_n(x) \geqslant f(x)$ and $\int_{\mathcal{O}} f \mathrm{d}\mu > -\infty$, it holds that*

$$\liminf_{n\to\infty} \int_{\mathcal{O}} f_n \mathrm{d}\mu \geqslant \int_{\mathcal{O}} \left(\liminf_{n\to\infty} f_n\right) \mathrm{d}\mu \tag{2.29}$$

- *when for all $n \in \mathbb{N} \setminus \{0\}$ and for all $x \in \mathcal{O}$, $f_n(x) \leqslant f(x)$ and $\int_{\mathcal{O}} f \mathrm{d}\mu < \infty$, it holds that*

$$\limsup_{n\to\infty} \int_{\mathcal{O}} f_n \mathrm{d}\mu \leqslant \int_{\mathcal{O}} \left(\liminf_{n\to\infty} f_n\right) \mathrm{d}\mu \tag{2.30}$$

Proof See Homework of Lecture 2. □

THEOREM 2.28 (Dominated Convergence) *Let $(\mathcal{O}, \mathscr{F}, \mu)$ be a measure space and f, g and f_1, f_2, f_3, ... be Borel measurable functions with respect to $(\mathcal{O}, \mathscr{F})$. Assume that for all $n \in \mathbb{N} \setminus \{0\}$ and for all $x \in \mathcal{O}$, $|f_n(x)| \leqslant g(x)$ and $\int_{\mathcal{O}} |g| \mathrm{d}\mu < \infty$ and $\lim_{n\to\infty} f_n(x) = f(x)$ almost everywhere with respect to μ. Then, $\int_{\mathcal{O}} |f| \mathrm{d}\mu < \infty$ and*

$$\lim_{n\to\infty} \int_{\mathcal{O}} f_n \mathrm{d}\mu = \int_{\mathcal{O}} f \mathrm{d}\mu. \tag{2.31}$$

Proof See Homework of Lecture 2. □

2.4 Radom-Nikodym Derivative

In order to introduce the notion of Radon-Nikodym derivative, the notion of absolute continuity of a measure with respect to another is needed.

DEFINITION 2.29 (Absolute Continuity) Given two measures μ and ν on a measurable space $(\mathcal{O}, \mathscr{F})$, ν is said to be absolutely continuous with respect to μ if for all $\mathcal{A} \in \mathscr{F}$ for which $\mu(\mathcal{A}) = 0$, it holds that $\nu(\mathcal{A}) = 0$.

THEOREM 2.30 *Given a mesurable space $(\mathcal{O}, \mathscr{F})$ and a non-negative Borel measurable function $f : \mathcal{O} \to \mathbb{R}$ with respect to $(\mathcal{O}, \mathscr{F})$, let $\nu : \mathscr{F} \to \mathbb{R}_+$ be*

$$\nu(\mathcal{A}) = \int_{\mathcal{A}} f \mathrm{d}\mu. \tag{2.32}$$

Then, ν is a measure on $(\mathcal{O}, \mathscr{F})$.

Consider two measures μ and ν on a measurable space $(\mathcal{O}, \mathscr{F})$ and a non-negative Borel measurable function f such that for all $\mathcal{A} \in \mathscr{F}$, $\nu(A)$ is the indefinite integral of f with respect to a measure μ, i.e.,

$$\nu(\mathcal{A}) = \int_{\mathcal{A}} f \mathrm{d}\mu. \tag{2.33}$$

Note that for all $\mathcal{A}$ for which $\mu(\mathcal{A}) = 0$, it follows that $\nu(\mathcal{A}) = 0$. That is, ν is absolutely continuous with respect to μ. The following theorem states the converse: if ν is absolutely continuous with respect to μ, then ν is obtained as the indefinite integral of f with respect to a measure μ, with f being a unique non-negative Boreal measurable function with respect to $(\mathcal{O}, \mathscr{F})$.

THEOREM 2.31 (Radon-Nikodym Derivative) *Consider a measurable space $(\mathcal{O}, \mathscr{F})$ and let μ and ν be measures on $(\mathcal{O}, \mathscr{F})$ such that μ is σ-finite and ν is absolutely continuous with respect to μ. Then, there exists a unique Borel measurable function $g : \mathcal{O} \to \mathbb{R}$, up to negligible sets with respect to μ, such that for all $\mathcal{A} \in \mathscr{F}$,*

$$\nu(\mathcal{A}) = \int_{\mathcal{A}} g \mathrm{d}\mu. \tag{2.34}$$

Proof See Homework of Lecture 2 □

The function g in (2.34) is often referred to as the **density** of ν with respect to μ, the **likelyhood ratio** of ν with respect to μ, or the **Radon-Nikodym derivative** of ν with respect to μ. To emphasize this denomination, it is often denoted by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

The following results are immediate extensions of Theorem 2.31.

COROLLARY *Let μ and ν be two measures on a given measurable space $(\mathcal{O}, \mathscr{F})$. Then, ν is absolutely continuous with respect to μ if and only if there exists a Borel measurable function $g : \mathcal{A} \to \mathbb{R}$ such that for all $\mathcal{A} \in \mathscr{F}$, the equality in* (2.34) *holds.*

The following theorem describe some of the properties of the Randon-Nikodym derivative.

THEOREM 2.32 *Let μ and ν be two measures on a given measurable space $(\mathcal{O}, \mathscr{F})$ with ν being absolutely continuous with respect to μ and μ being σ-finite. Then,*

- *for all $x \in \mathcal{O}$, $\frac{\mathrm{d}\mu}{\mathrm{d}\mu}(x) = 1$.*
- *if $f : \mathcal{O} \to \mathbb{R}_+$ is a non-negative Borel measurable function with respect to $(\mathcal{O}, \mathscr{F})$, it holds that for all $\mathcal{A} \in \mathscr{F}$,*

$$\int_{\mathcal{A}} f \mathrm{d}\nu = \int_{\mathcal{A}} f \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \mathrm{d}\mu; \tag{2.35}$$

- *if λ is a σ-finite measure on $(\mathcal{O}, \mathscr{F})$, μ is absolutely continuous with respect to λ, it holds that for all $x \in \mathcal{O}$*

$$\frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x) = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}(x) \frac{\mathrm{d}\mu}{\mathrm{d}\lambda}(x); \text{ and} \tag{2.36}$$

- *if μ is absolutely continuous with respect to ν, and ν is σ-finite, it holds that for all $x \in \mathcal{O}$*

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(x) \frac{\mathrm{d}\mu}{\mathrm{d}\nu}(x) = 1. \tag{2.37}$$

Proof See Homework of Lecture 1. □

THEOREM 2.33 *Let μ be a σ-finite measure on $(\mathcal{O}, \mathscr{F})$ and $\nu_1, \nu_2, \ldots, \nu_n$ be finite measures on $(\mathcal{O}, \mathscr{F})$ such that for all $k \in \{1, 2, \ldots, n\}$, ν_k is absolutely continuous with μ. Then, that for all $x \in \mathcal{O}$,*

$$\frac{\mathrm{d}\sum_{t=1}^{n} \nu_t}{\mathrm{d}\mu}(x) = \sum_{t=1}^{m} \frac{\mathrm{d}\nu_t}{\mathrm{d}\mu}(x). \tag{2.38}$$

Moreover, if ν is a measure on $(\mathcal{O}, \mathscr{F})$ such that for all $\mathcal{A} \in \mathscr{F}$, $\nu(\mathcal{A}) = \lim_{n\to\infty} \sum_{t=1}^{n} \nu_t(\mathcal{A})$, then ν is absolutely continuous with μ and

$$\lim_{n\to\infty} \frac{\mathrm{d}\sum_{t=1}^{n} \nu_t}{\mathrm{d}\mu}(x) = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}(x). \tag{2.39}$$

Proof See Homework of Lecture 1. □