# Measure Theory and Probability Theory

# Contents

# Chapter 1

# Measures

## 1.1 Class of sets

If $2^\Omega$ denotes the power set of the set $\Omega$ and a collection of subsets $\mathcal{C} \subset 2^\Omega$, we say $\mathcal{C}$ is defined on $\Omega$.

For convenience, let's define some operations on a collection of sets $\mathcal{C} \in 2^\Omega$.

- Given $f : \Sigma \to \Omega$, $f^{-1}(\mathcal{C}) := \{f^{-1}(A) \in \Sigma \mid A \in \mathcal{C}\}$,

- Given $B \subset \Omega$, $B \cap \mathcal{C} := \{B \cap A \mid A \in \mathcal{C}\}$.

**Definition 1.1.1 (semiring)** A nonempty family of sets $\mathcal{S}$ is called a *semiring (of sets)* if the following three statements are true for all sets $A$ and $B$,

(1) $\varnothing \in \mathcal{S}$

(2) $A, B \in \mathcal{S}$ implies $A \cap B \in \mathcal{S}$

(3) $A, B \in \mathcal{S}$ implies $A - B = \cup_{i=1}^n C_i$, where $\{C_i\}_{i=1}^n$ is a collection of pairwise disjoint sets in $\mathcal{S}$.

Since condition (3) together with $\mathcal{S} \neq \varnothing$ implies that $\varnothing \in \mathcal{S}$, conditions (1) is actually unnecessary.

**Definition 1.1.2 (ring)** A nonempty family of sets $\mathcal{R}$ is called a *ring (of sets)* if it is closed under finite union and difference. That is, the following two statements are true for all sets $A$ and $B$,

(1) $A, B \in \mathcal{R}$ implies $A \cup B \in \mathcal{R}$

(2) $A, B \in \mathcal{R}$ implies $A - B \in \mathcal{R}$.

A ring of sets $\mathcal{R}$ is a ring in the context of abstract algebra with the operations intersection(multiplication) and symmetric difference(addition), where the multiplicative identity is the set $\cup_{A \in \mathcal{R}} A$.

**Definition 1.1.3 (semialgebra[#1])** A subring $\mathcal{S}$ defined on $\Omega$ is called an *semialgebra* if it contains $\Omega$.

**Definition 1.1.4 (semialgebra[#2])** A collection of sets $\mathcal{S} \subset 2^\Omega$ is called an *semialgebra* the following two statements are true for all sets $A$ and $B$,

(1) $A, B \in \mathcal{S}$ implies $A \cap B \in \mathcal{S}$

(2) $A \in \mathcal{S}$ implies $A^c = \cup_{i=1}^n C_i$, where $\{C_i\}_{i=1}^n$ is a collection of pairwise disjoint sets in $\mathcal{S}$.

**Example 1.1.1**

$$\mathcal{C} \equiv \{(a, b] \in \mathbb{R} \mid -\infty \le a \le b < \infty\} \cup \{(a, \infty) \in \mathbb{R} \mid -\infty \le a < \infty\}$$

is a semialgebra on $\mathbb{R}$.

**Definition 1.1.5 (algebra$^{\#1}$)** A ring $\mathcal{F}$ defined on $\Omega$ is called an *algebra (of sets)* if it contains $\Omega$.

**Proposition 1.1.1** Every ring is a semiring.

*Proof.* Let $\mathcal{R}$ be a ring. First we see $\varnothing = A - A \in \mathcal{R}$. Second, it is straightforward to check

$$A \cap B = (A \cup B - (A - B)) - (B - A) \in \mathcal{R}.$$

Finally, it is a trivial fact that $A - B = C_1 \in \mathcal{R}$. Thus we show the ring $\mathcal{R}$ is also a semiring.

The following equivalent definitions of algebra are common to see.

**Definition 1.1.6 (algebra$^{\#2}$)** A collection of sets $\mathcal{F} \subset 2^\Omega$ is called an *algebra (of sets)* if

(a) $\Omega \in \mathcal{F}$,

(b) $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$,

(c) $A, B \in \mathcal{F}$ implies $A \cup B \in \mathcal{F}$.

An algebra of sets is an associative algebra over a field in the context of abstract algebra: it is a ring and also a vector space over the field $F_2$ of 2 elements, which makes it an $F_2$-algebra.

**Proposition 1.1.2** Every algebra is a ring.

*Proof.* Note it holds that

$$A - B = A \cap B^c = (A^c \cup B)^c.$$

**Example 1.1.2** Let $\Omega$ be a nonempty set, and let $\#A$ denote the number of elements of a set $A \subset \Omega$. Then

$$\mathcal{F} = \{A \subset \Omega : \text{ either } \#A \text{ is finite or } \#A^c \text{ is finite}\}$$

is an algebra defined on $\Omega$.

**Definition 1.1.7 ($\sigma$-algebra)** A class $\mathcal{F} \subset 2^\Omega$ is called a $\sigma$-*algebra* if it is an algebra and if it satisfies

(d) $A_n \in \mathcal{F}$ for $n \ge 1 \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Under the assumption that $\mathcal{F}$ is an algebra, condition (d) can be replaced by a weaker condition (d'), which leads to the following equivalent definition.

**Definition 1.1.8 ($\sigma$-algebra)** A class $\mathcal{F} \subset 2^\Omega$ is called a $\sigma$-*algebra* if it is an algebra and if it satisfies

(d') $A_n \in \mathcal{F}$, $A_n \subset A_{n+1}$ for $n \geq 1$ $\implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

**Example 1.1.3** Let $\Omega$ be a nonempty set. Then

$$\mathcal{F} = \{A \subset \Omega : \text{ either } A \text{ is a countable set or } A^c \text{ is a countable set}\}$$

is a $\sigma$-algebra defined on $\Omega$.

It is easy to check the intersection of a family of $\sigma$-algebras on $\Omega$ is also a $\sigma$-algebra. However, the union of two $\sigma$-algebras may not even be an algebra. In many instances, given an arbitrary collection of subsets of $\Omega$, one would like to extend it to a $\sigma$-algebra by adding extra subsets of $\Omega$ as few as possible. This leads to the following definition.

**Definition 1.1.9 ($\sigma$-algebra generated by sets)** If $\mathcal{A} \subset 2^\Omega$, then the $\sigma$-*algebra generated by* $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$, is defined as

$$\sigma(\mathcal{A}) = \bigcap_{\mathcal{F} \in \mathcal{I}(\mathcal{A})} \mathcal{F}$$

where $\mathcal{I}(\mathcal{A}) = \{\mathcal{F} : \mathcal{A} \subset \mathcal{F} \text{ and } \mathcal{F} \text{ is a } \sigma\text{-algebra defined on } \Omega\}$ is the collection of all $\sigma$-algebras containing the class $\mathcal{A}$.

Note that since the power set $2^\Omega$ contains $\mathcal{A}$ and is itself a $\sigma$-algebra defined on $\Omega$, the collection $\mathcal{I}(\mathcal{A})$ is not empty and hence, the intersection in the above definition is well defined.

**Definition 1.1.10 (Borel $\sigma$-algebra)** The *Borel $\sigma$-algebra* on a topological space $\mathbb{S}$ is defined as the $\sigma$-algebra generated by the collection of open sets in $\mathbb{S}$.

**Example 1.1.4** Let $\mathcal{B}(\mathbb{R}^k)$ denote the Borel $\sigma$-algebra on the Euclidean space $\mathbb{R}^k$ with standard topology, $1 \leq k < +\infty$. Then, $\mathcal{B}(\mathbb{R}^k) = \sigma(\{A : A \text{ is an open subset of } \mathbb{R}^k\})$ is also generated by each of the following classes of sets

$$\begin{aligned}
\mathcal{O}_1 &= \{(a_1, b_1) \times \cdots \times (a_k, b_k) : -\infty \leq a_i < b_i \leq +\infty, 1 \leq i \leq k\}; \\
\mathcal{O}_2 &= \{(-\infty, x_1) \times \cdots \times (-\infty, x_k) : x_i \in \mathbb{R}, 1 \leq i \leq k\}; \\
\mathcal{O}_3 &= \{(a_1, b_1) \times \cdots \times (a_k, b_k) : a_i, b_i \in \mathbb{Q}, a_i < b_i, 1 \leq i \leq k\}; \\
\mathcal{O}_4 &= \{(-\infty, x_1) \times \cdots \times (-\infty, x_k) : x_i \in \mathbb{Q}, 1 \leq i \leq k\}.
\end{aligned}$$

**Definition 1.1.11 ($\pi$-system)** A class $\mathcal{C}$ of subsets of $\Omega$ is a $\pi$-system or a $\pi$-class if

$$A, B \in \mathcal{C} \implies A \cap B \in \mathcal{C}.$$

**Example 1.1.5** The classes $\mathcal{O}_i$ ($i = 1, 2, 3, 4$) in Example 1.1.3 are all $\pi$-systems.

**Definition 1.1.12 ($\lambda$-system)** A class $\mathcal{L}$ of subsets of $\Omega$ is a $\lambda$-system or a $\lambda$-class if

(a) $\Omega \in \mathcal{L}$,

(b) $A, B \in \mathcal{L}, A \subset B \implies B \setminus A \in \mathcal{L}$,

(c) $A_n \in \mathcal{L}, A_n \subset A_{n+1}$ for all $n \geq 1 \implies A_n \in \mathcal{L}$ for $n \geq 1$.

**Example 1.1.6** Every $\sigma$-algebra is a $\lambda$-system.

**Theorem 1.1.1** If $\mathcal{C}$ is a $\pi$-system, then $\lambda(\mathcal{C}) = \sigma(\mathcal{C})$.

## 1.2 Measurable Space

**Definition 1.2.1 (measurable space)** Given a nonempty set $\Omega$ and a $\sigma$-algebra $\mathcal{F}$ on $\Omega$, the tuple $(\Omega, \mathcal{F})$ is called a *measurable space.*

**Definition 1.2.2 (measurable map)** Given two measurable spaces $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$, the map $f : \Omega \to S$ is $\mathcal{F}/\mathcal{S}$-*measurable* if

$$f^{-1}(B) \in \mathcal{F}, \quad \forall B \in \mathcal{S}.$$

We say $f$ is measurable if the underlying $\sigma$-algebras $\mathcal{F}$ and $\mathcal{S}$ are self-evident.

Note $f^{-1}$ preserves $\sigma$-algebra, which leads to the following definition.

**Definition 1.2.3 ($\sigma$-algebra generated by measurable map)** Assume $f$ is a measurable map from $(\Omega, \mathcal{F})$ to $(S, \mathcal{S})$. The $\sigma$-algebra generated by $f$ is defined as

$$\sigma(f) := \sigma\left(f^{-1}(\mathcal{S})\right) = f^{-1}(\mathcal{S}).$$

$\sigma(f)$ is the smallest $\sigma$-algebra on $\Omega$ such that $f$ is $\mathcal{F}/\mathcal{S}$-*measurable.* That is,

$$f \text{ is } \mathcal{F}/\mathcal{S}\text{-}measurable \implies \sigma(f) \subset \mathcal{F}.$$

**Definition 1.2.4 (tensor product $\sigma$-algebra )** Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces. Denote by $\mathcal{F}_1 \otimes \mathcal{F}_2$ the $\sigma$-algebra on the Cartesian product $\Omega_1 \times \Omega_2$ generated by subsets of the form $B_1 \times B_2$, where $B_1 \in \mathcal{F}_1$ and $B_2 \in \mathcal{F}_2$, i.e.,

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma\left(\{B_1 \times B_2 : B_1 \in \mathcal{F}_1, B_2 \in \mathcal{F}_2\}\right).$$

**Definition 1.2.5 (product measurable space)** Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces. The *product measurable space* of $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ is defined as $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$.

**Definition 1.2.6 (infinite product measurable space)** Let $\{(\Omega_i, \mathcal{F}_i)\}_{i \in T}$ be a collection of measurable spaces. The *infinite product measurable space* of $\{(\Omega_i, \mathcal{F}_i)\}_{i \in T}$ is defined as $\left(\prod_{i \in T} \Omega_i, \bigotimes_{i \in T} \mathcal{F}_i\right)$, where

$$\mathcal{MC}(\{\mathcal{F}_i\}_{i \in T}) := \left\{\prod_{i \in T} B_i : B_i \in \mathcal{F}_i, \ B_i = \Omega_i \text{ for all but a finite number of } i \in T\right\}$$

is called the collection of *measurable cylinders* and

$$\bigotimes_{i \in T} \mathcal{F}_i = \sigma\left(\mathcal{MC}\left(\{\mathcal{F}_i\}_{i \in T}\right)\right).$$

Product measurable space is the product in the category consisting of all measurable spaces and measurable functions and satisfies the following universal property: for every cone consisting of a measurable space $Y$ and a family of measurable functions $\{f_i : Y \to \Omega_i\}_{i \in T}$, there exists a unique measurable function

$$f : Y \longrightarrow \prod_{i \in T} \Omega_i, \qquad y \longmapsto (f_i(y))_{i \in T}$$

denoted by $\prod_{i \in T} f_i$ such that the following diagrams commute for all $i$ in $T$,

$$
\begin{array}{ccc}
Y & \xrightarrow{\ \exists! f\ } & \prod_{i \in T} \Omega_i \\
& {\scriptstyle f_i}\searrow & \big\downarrow {\scriptstyle \pi_i} \\
& & \Omega_i
\end{array}
$$

where $\pi_i : (\omega_k)_{k \in T} \mapsto \omega_i$ is the projection map. Thus we see that $\bigotimes_{i \in T} \mathcal{F}_i$ is the smallest $\sigma$-algebra on $\prod_{i \in T} \Omega_i$ such that every projection map $\pi_i$ is measurable, that is

$$\bigotimes_{i \in T} \mathcal{F}_i = \sigma\left( \bigcup_{i \in T} \sigma(\pi_i) \right).$$

Generally, we have the following proposition.

**Proposition 1.2.1**

$$\sigma\left( \prod_{i \in T} f_i \right) = \sigma\left( \bigcup_{i \in T} \sigma(f_i) \right).$$

**Proposition 1.2.2 (product and Borel $\sigma$-fields)** Let $\{S_i\}_{i=1}^{\infty}$ be a sequence of separable metric spaces. Then

$$\mathcal{B}\left( \prod_{i=1}^{\infty} S_i \right) = \bigotimes_{i=1}^{\infty} \mathcal{B}(S_i)$$

In particular, $\mathcal{B}\left(\mathbb{R}^d\right) = (\mathcal{B}(\mathbb{R}))^d$.

**Proposition 1.2.3 (convergence and limits)** Let $f_1, f_2, \ldots$ be measurable functions from a measurable space $(\Omega, \mathcal{F})$ into some metric space $(S, \rho)$. Then

 (i) $\{\omega \in \Omega : f_n(\omega) \text{ converges}\} \in \mathcal{A}$ if $S$ is complete;

 (ii) $f_n \to f$ on $\Omega$ implies that $f$ is measurable.

## 1.3 Measure

**Definition 1.3.1 (measure)** Given a measurable space $(\Omega, \mathcal{F})$, a set function $\mu : \mathcal{F} \to [0, +\infty]$ is called a *measure* if

 (a) $\mu(\varnothing) = 0$;

 (b) $\sigma$-additivity: for any countable collection $\{A_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{F}$,

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

**Definition 1.3.2 (measure space)** A *measure space* is a triple $(\Omega, \mathcal{F}, \mu)$ where $\Omega$ is a set, $\mathcal{F}$ is a $\sigma$-algebra on $\Omega$ and $\mu : \mathcal{F} \to [0, +\infty]$ is a measure.

**Proposition 1.3.1** Suppose $(\Omega, \mathcal{F}, \mu)$ is a measure space.

 (i) Monotonicity: If $A, B \in \mathcal{F}$ and $A \subset B$, then $\mu(A) \leq \mu(B)$;

 (ii) Subadditivity: For any measurable sets $\{A_i\}_{i=1}^{\infty}$ in $\mathcal{F}$,

$$\mu\left( \bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

 (iii) Continuity from below: If $\{A_i\}_{i=1}^{\infty}$ is a collection of sets in $\mathcal{F}$ that are increasing (meaning that $A_1 \subset A_2 \subset A_3 \subset \cdots$ ) then the union of the sets $A_n$ is measurable and

$$\mu\left( \bigcup_{i=1}^{\infty} A_i \right) = \lim_{i \to \infty} \mu(A_i) = \sup_{i \geq 1} \mu(A_i);$$

(iv) Continuity from above: If $\{A_i\}_{i=1}^{\infty}$ is a collection of sets in $\mathcal{F}$ that are decreasing (meaning that $A_1 \supset A_2 \supset A_3 \supset \cdots$) then the intersection of the sets $A_n$ is measurable. Furthermore, if at least one of the $A_n$ has finite measure then

$$\mu \left( \bigcap_{i=1}^{\infty} A_i \right) = \lim_{i \to \infty} \mu\left(A_i\right) = \inf_{i \geq 1} \mu\left(A_i\right).$$

**Definition 1.3.3 (pre-measure)** Let $\mathcal{S}$ be a semiring on $\Omega$. A set function $\mu_0 : \mathcal{S} \to [0, +\infty]$ is called a *pre-measure* if

(a) $\mu_0(\varnothing) = 0$;

(b) $\sigma$-additivity: for any countable collection $\{A_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{S}$,

$$\mu_0 \left( \bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

Clearly we see every measure is a pre-measure, since every $\sigma$-algebra is a semiring.

**Definition 1.3.4 ($\sigma$-finite measure)** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu$ a measure on it. The measure $\mu$ is called a *$\sigma$-finite measure* if the set $\Omega$ can be covered with at most countably many measurable sets with finite measure, which means that there are sets $A_1, A_2, \ldots \in \mathcal{F}$ with $\mu\left(A_n\right) < \infty$ for all $n \in \mathbb{N}_+$ that satisfy

$$\bigcup_{n=1}^{\infty} A_n = \Omega.$$

*$\sigma$-finite pre-measure* can be defined in a similar way.

**Definition 1.3.5 (finite measure)** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu$ a measure on it. The measure $\mu$ is called a *finite measure* if $\mu(\Omega) < \infty$.

**Definition 1.3.6 (probability measure)** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu$ a measure on it. The measure $\mu$ is called a *probability measure* if $\mu(\Omega) = 1$. In this case, $(\Omega, \mathcal{F}, \mu)$ is called a *probability measure space* or *probability space* in short.

**Definition 1.3.7 (complete measure)** A measure $\mu$ defined on the measurable space $(\Omega, \mathcal{F})$ is called *complete* if for any $A \in \mathcal{F}$ with $\mu(A) = 0$, $2^A \subset \mathcal{F}$. In this case, the measure space $(\Omega, \mathcal{F}, \mu)$ is called a complete measure space.

**Definition 1.3.8 (Borel measure)** Suppose $\mathcal{B}(\mathbb{S})$ is the Borel $\sigma$-algebra on a topological space $\mathbb{S}$. A *Borel measure* is measure defined on the measurable space $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$.

**Definition 1.3.9 ($\mu$-null set)** Given a measure space $(\Omega, \mathcal{F}, \mu)$, a set $A \in \mathcal{F}$ is called a $\mu$-null set is $\mu(A) = 0$.

**Definition 1.3.10 (extension of a measure space)** The measure space $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mu})$ is called the extension of the measure space $(\Omega, \mathcal{F}, \mu)$ if $\mathcal{F} \subset \widetilde{\mathcal{F}}$ and $\widetilde{\mu}|_{\mathcal{F}} = \mu$.

**Proposition 1.3.2** Given a measure space $(\Omega, \mathcal{F}, \mu)$, define

$$\widetilde{\mathcal{F}} = \{A \cup E \in 2^{\Omega} \mid A \in \mathcal{F}, E \text{ is a subset of a } \mu\text{-null set}\}$$

and

$$\widetilde{\mu} : \widetilde{\mathcal{F}} \longrightarrow [0, +\infty],$$
$$A \cup E \longmapsto \mu(A).$$

Then we have

(1) $\widetilde{\mu}$ is well-defined,

(2) $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mu})$ is a complete measure space,

(3) $\widetilde{\mu}$ is the unique measure on $(\Omega, \widetilde{\mathcal{F}})$ that coincides with $\mu$ on $\mathcal{F}$,

(4) $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mu})$ is the smallest complete extension of $(\Omega, \mathcal{F}, \mu)$.

**Definition 1.3.11 (completion of a measure space)** Given a measure space $(\Omega, \mathcal{F}, \mu)$, define

$$\widetilde{\mathcal{F}} = \{A \cup E \in 2^\Omega \mid A \in \mathcal{F}, E \text{ is a subset of a } \mu\text{-null set}\}$$

and

$$\widetilde{\mu} : \widetilde{\mathcal{F}} \longrightarrow [0, +\infty],$$
$$A \cup E \longmapsto \mu(A).$$

The measure space $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mu})$ is called the *completion of* $(\Omega, \mathcal{F}, \mu)$.

**Definition 1.3.12 (outer measure)** A set function $\mu^* : 2^\Omega \to [0, +\infty]$ is called a *outer measure* on $\Omega$ if

(1) $\mu^*(\varnothing) = 0$;

(2) $A \subset B \subset \Omega \implies \mu^*(A) \leq \mu^*(B)$;

(3) countable subadditivity: for any countable collection $\{A_i\}_{i=1}^\infty$ of sets in $2^\Omega$,

$$\mu^* \left( \bigcup_{n=1}^\infty A_n \right) \leq \sum_{n=1}^\infty \mu^*(A_n).$$

**Definition 1.3.13 ($\mu^*$-measurable)** Suppose that $\mu^*$ is an outer measure on $\Omega$. A set $A$ is said to be $\mu^*$-*measurable* if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^* (E \cap A^c) \text{ for all } E \subset \Omega$$

**Proposition 1.3.3** Let $\Omega$ be a set, let $\mu^*$ be an outer measure on $\Omega$, and let $\mathcal{M}_{\mu^*}$ be the collection of all $\mu^*$-measurable subsets of $\Omega$. Then

(1) $\mathcal{M}_{\mu^*}$ is a $\sigma$-algebra,

(2) the restriction of $\mu^*$ to $\mathcal{M}_{\mu^*}$ is a measure on $\mathcal{M}_{\mu^*}$,

(3) $(\Omega, \mathcal{M}_{\mu^*}, \mu^*|_{\mathcal{M}_{\mu^*}})$ is a complete measure space.

**Definition 1.3.14 (induced outer measure)** Given a pre-measure $\mu_0$ on a semialgebra, $\mathcal{S} \subset 2^\Omega$, the *outer measure induced by* $\mu_0$ is the set function $\mu_0^* : 2^\Omega \to [0, +\infty]$, as

$$\mu_0^*(A) = \inf \left\{ \sum_{n=1}^\infty \mu_0 (A_n) : \{A_n\}_{n \geq 1} \subset \mathcal{S}, A \subset \bigcup_{n \geq 1} A_n \right\}.$$

Especially, given a measure $\mu$ on a measurable space $(\Omega, \mathcal{F})$, the *outer measure induced by* $\mu$ is the set function $\mu^* : 2^\Omega \to [0, +\infty]$, as

$$\mu^*(A) = \inf \{\mu (C) : C \in \mathcal{F}, A \subset C\}.$$

**Theorem 1.3.1 (Carathéodory's extension theorem)** Let $\mathcal{C}$ be a semialgebra on $\Omega$, let $\mu_0 : \mathcal{C} \to [0, +\infty]$ be a pre-measure and let $\mu_0^*$ be the outer measure induced by $\mu_0$. Then

(1) $\mathcal{C} \subset \sigma(\mathcal{C}) \subset \mathcal{M}_{\mu_0^*}$, where $\mathcal{M}_{\mu_0^*}$ denotes the collection of all $\mu_0^*$-measurable subsets of $\Omega$,

(2) $\mu_0^*|_{\mathcal{C}} = \mu_0$,

(3) $(\Omega, \sigma(\mathcal{C}), \mu_0^*|_{\sigma(\mathcal{C})})$ is a measure space. $(\Omega, \mathcal{M}_{\mu_0^*}, \mu_0^*|_{\mathcal{M}_{\mu_0^*}})$ is a complete measure space,

(4) If $\mu_0$ is $\sigma$-finite, then

    (i) $\mu_0^*|_{\sigma(\mathcal{C})}$ is the unique measure on $(\Omega, \sigma(\mathcal{C}))$ that coincides with $\mu_0$ on $\mathcal{C}$. In this case, $\mu_0^*|_{\sigma(\mathcal{C})}$ is also $\sigma$-finite,

    (ii) $(\Omega, \mathcal{M}_{\mu_0^*}, \mu_0^*|_{\mathcal{M}_{\mu_0^*}})$ is the completion of $(\Omega, \sigma(\mathcal{C}), \mu_0^*|_{\sigma(\mathcal{C})})$,

    (iii) $\mu_0^*|_{\mathcal{M}_{\mu_0^*}}$ is the unique measure on $(\Omega, \mathcal{M}_{\mu_0^*})$ that coincides with $\mu_0$ on $\mathcal{C}$.

For simplicity, $(\Omega, \sigma(\mathcal{C}), \mu_0^*|_{\sigma(\mathcal{C})})$ and $(\Omega, \mathcal{M}_{\mu_0^*}, \mu_0^*|_{\mathcal{M}_{\mu_0^*}})$ can be denoted as $(\Omega, \sigma(\mathcal{C}), \mu_0^*)$ and $(\Omega, \mathcal{M}_{\mu_0^*}, \mu_0^*)$.

**Definition 1.3.15 (product measure)** Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two measure spaces. A measure $\mu$ on the measurable space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ is said to be a *product measure* of $\mu_1$ and $\mu_2$ if it satisfies the property

$$(\mu)(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2)$$

for all $B_1 \in \Omega_1$, $B_2 \in \Omega_2$. In this case, the measure space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu)$ is called the product measure space of $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$.

**Proposition 1.3.4** Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two measure spaces. If the measure $\mu_1$ and $\mu_2$ are $\sigma$-finite, the product measure of $\mu_1$ and $\mu_2$ uniquely exists and denoted by $\mu_1 \times \mu_2$.

**Definition 1.3.16 (product probability measure)** Let $\{(\Omega_i, \mathcal{F}_i, \mathrm{P}_i)\}_{i \in T}$ be a collection of probability measure spaces. The map

$$\mathrm{P}_0 : \mathcal{MC}(\{\mathcal{F}_i\}_{i \in T}) \longrightarrow [0, 1]$$
$$\prod_{i \in T} B_i \longmapsto \prod_{i \in T : \mathrm{P}_i(B_i) < 1} \mathrm{P}_i(B_i)$$

can be uniquely extended to a probability measure $\mathrm{P}$ on the product measurable space $\left( \prod_{i \in T} \Omega_i, \bigotimes_{i \in T} \mathcal{F}_i \right)$. The *product probability space* of $\{(\Omega_i, \mathcal{F}_i, \mathrm{P}_i)\}_{i \in T}$ is defined as $\left( \prod_{i \in T} \Omega_i, \bigotimes_{i \in T} \mathcal{F}_i, \mathrm{P} \right)$.

## 1.3.1 Lebesgue-Stieltjes measure

**Definition 1.3.17 (Lebesgue-Stieltjes measure on $\mathbb{R}$)** Given nondecreasing function $F : \mathbb{R} \to \mathbb{R}$ and semialgebra

$$\mathcal{C} \equiv \{(a, b] \in \mathbb{R} \mid -\infty \le a \le b < \infty\} \cup \{(a, \infty) \in \mathbb{R} \mid -\infty \le a < \infty\}$$

we can define a pre-measure $\mu_F$ on $\mathcal{C}$ as follows

$$\mu_F((a, b]) = F(b+) - F(a+),$$
$$\mu_F((a, \infty)) = F(\infty) - F(a+).$$

The measure space $\left( \mathbb{R}, \mathcal{M}_{\mu_F^*}, \mu_F^* \right)$ is called a *Lebesgue-Stieltjes measure space* and $\mu_F^*$ is the *Lebesgue-Stieltjes measure* generated by $F$.

# Chapter 2

# Integration

Assume that all the functions in this chapter are from $\Omega$ to $\mathbb{R}$.

## 2.1 The definition of integration

### 2.1.1 The integration of nonnegative simple functions

**Definition 2.1.1 (simple function)** Given a measurable space $(\Omega, \mathcal{F})$, a simple function is a function of the form

$$f(\omega) = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i}(\omega), \quad a_i \in \mathbb{R},$$

where $A_i$ are disjoint subsets of $\Omega$ that belong to $\mathcal{F}$.

**Definition 2.1.2 (integration of nonnegative simple function)** Given a measure space $(\Omega, \mathcal{F}, \mu)$, the integral of the nonnegative simple function $f(x) = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i}(\omega)$ with respect to $\mu$ is defined as

$$\int_{\Omega} f d\mu := \sum_{i=1}^{m} a_i \mu_{A_i}.$$

Note that if $f(\omega) = \sum_{j=1}^{m} b_j \mathbf{1}_{B_j}(\omega)$ where $B_i$ are disjoint subsets of $\Omega$ that belong to $\mathcal{F}$, there must be

$$A_i \cap B_j \neq 0 \implies a_i = b_j.$$

By $\sigma$-additivity of the measure $\mu$, we have

$$\sum_{i=1}^{m} a_i \mu(A_i) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i \mu(A_i \cap B_j)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} b_j \mu(A_i \cap B_j) = \sum_{j=1}^{n} b_j \mu(B_j).$$

Hence $\int_{\Omega} f d\mu$ does not depend on the representation of $f$ used in its definition.

### 2.1.2 The integration of nonnegative measurable functions

**Proposition 2.1.1** Let $\{f_n\}_{n \geq 1}$ and $\{g_n\}_{n \geq 1}$ be two sequences of nonnegative simple functions on $(\Omega, \mathcal{F}, \mu)$ to such that if as $n \to \infty$, $f_n(\omega) \uparrow f(\omega)$ and $g_n(\omega) \uparrow f(\omega)$ for all

$\omega \in \Omega$, then

$$\lim_{n \to \infty} \int_{\Omega} f_n d\mu = \lim_{n \to \infty} \int_{\Omega} g_n d\mu.$$

**Definition 2.1.3 (integration of nonnegative measurable function)** Given a measure space $(\Omega, \mathcal{F}, \mu)$, the integral of the nonnegative measurable function $f$ with respect to $\mu$ is defined as

$$\int_{\Omega} f d\mu := \lim_{n \to \infty} \int_{\Omega} f_n d\mu$$

where $\{f_n\}_{n \geq 1}$ is any sequence of nonnegative simple functions such that $f_n(\omega) \uparrow f(\omega)$ for all $\omega \in \Omega$.

### 2.1.3 The integration of measurable functions

**Definition 2.1.4 (integration of measurable function)** Given a measure space $(\Omega, \mathcal{F}, \mu)$. Let $f^+ = f\mathbf{1}_{\{f \geq 0\}}$ and $f^- = -f\mathbf{1}_{\{f < 0\}}$. The integral of $f$ with respect to $\mu$ is defined as

$$\int_{\Omega} f d\mu = \int_{\Omega} f^+ d\mu - \int_{\Omega} f^- d\mu$$

provided that at least one of the integrals on the right side is finite.

**Definition 2.1.5 (integrable function)** Given a measure space $(\Omega, \mathcal{F}, \mu)$, a measurable function $f$ is integrable with respect to $\mu$ if

$$\int_{\Omega} |f| d\mu$$

exists, or equivalently both $\int_{\Omega} f^+ d\mu$ and $\int_{\Omega} f^- d\mu$ are finite.

## 2.2 $L^p$-space

**Definition 2.2.1 (linear space $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$)** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $0 < p \leq \infty$. Let

$$\|f\|_p := \begin{cases} \left( \int_{\Omega} |f|^p d\mu \right)^{\min\left\{\frac{1}{p}, 1\right\}}, & 0 < p < \infty, \\ \inf \{M : |f| < M \text{ a.e.}\}, & p = \infty. \end{cases}$$

Then

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mu) := \{f : \|f\|_p < \infty\}$$

is a linear space.

**Definition 2.2.2 (Banach space $L^p(\Omega, \mathcal{F}, \mu)$)** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $1 \leq p \leq \infty$. Then

$$\mathcal{N} := \{f : f = 0 \text{ a.e.}\} = \ker\left( \|\cdot\|_p \right).$$

is a subspace of $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ and

$$L^p(\Omega, \mathcal{F}, \mu) := \mathcal{L}^p(\Omega, \mathcal{F}, \mu)/\mathcal{N}$$

is a Banach space with the norm

$$\|\cdot\|_p : L^p(\Omega, \mathcal{F}, \mu) \longrightarrow [0, \infty),$$
$$f + \mathcal{N} \longmapsto \|f\|_p.$$

For simplicity, we can identify $f$ with its equivalence class $f + \mathcal{N}$.

**Theorem 2.2.1 (Hölder's inequality)** Let $(\Omega, \Sigma, \mu)$ be a measure space and let $p, q \in [1, \infty]$ with $1/p + 1/q = 1$. Then for all measurable real- or complex-valued functions $f$ and $g$ on $\Omega$,

$$\|fg\|_1 \le \|f\|_p \|g\|_q$$

If, in addition, $p, q \in (1, \infty)$ and $f \in L^p(\Omega, \Sigma, \mu)$ and $g \in L^q(\Omega, \Sigma, \mu)$, then Hölder's inequality becomes an equality if and only if $|f|^p$ and $|g|^q$ are linearly dependent in $L^1(\Omega, \Sigma, \mu)$, meaning that there exist real numbers $\alpha, \beta \ge 0$, not both of them zero, such that $\alpha |f|^p = \beta |g|^q$ $\mu$-almost everywhere.

The numbers $p$ and $q$ above are said to be Hölder conjugates of each other.

## 2.3 Convergence of measurable function

**Definition 2.3.1 (measurable function)** Let $(\Omega_1, \Sigma_1)$ and $(\Omega_2, \Sigma_2)$ be measurable spaces, meaning that $\Omega_1$ and $\Omega_2$ are sets equipped with respective $\sigma$-algebras $\Sigma_1$ and $\Sigma_2$. A function $f : \Omega_1 \to \Omega_2$ is said to be a *measurable function*, if for every $E \in \Sigma_2$, the preimage of $E$ under $f$ is measurable in $\Sigma_1$, i.e.

$$f^{-1}(E) := \{\omega \in \Omega_1 | \ f(\omega) \in E\} \in \Sigma_1, \ \ \forall E \in \Sigma_2.$$

If $f : \Omega_1 \to \Omega_2$ is measurable, we sometimes write

$$f \colon (\Omega_1, \Sigma_1) \to (\Omega_2, \Sigma_2)$$

to emphasize the dependency on the $\sigma$-algebras $\Sigma_1$ and $\Sigma_2$. In this section, we focus on a sort of special measurable functions:

$$f \colon (\Omega, \Sigma) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

**Definition 2.3.2 (convergence in measure)** Let $f, f_n$ $(n \in \mathbb{N}) : \Omega \to \mathbb{R}$ be measurable functions on a measure space $(\Omega, \Sigma, \mu)$. The sequence $f_n$ is said to converge in measure to $f$ if for every $\varepsilon > 0$,

$$\lim_{n \to \infty} \mu \left( \{\omega \in \Omega : |f(\omega) - f_n(\omega)| \ge \varepsilon\} \right) = 0,$$

# Chapter 3

# Differentiation

## 3.1 Signed measures

**Definition 3.1.1 (signed measure)** Given a measurable space $(\Omega, \mathcal{F})$, a set function $\mu : \mathcal{F} \to (-\infty, +\infty]$ or $\mu : \mathcal{F} \to [-\infty, +\infty)$ is called a *signed measure* if

(a) $\mu(\varnothing) = 0$;

(b) $\sigma$-additivity: for any countable collection $\{A_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{F}$,

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

**Definition 3.1.2 (finite signed measure)** A signed measure $\mu$ on measurable space $(\Omega, \mathcal{F})$ is *finite* if $\mu(\mathcal{F}) \in (-\infty, +\infty)$.

**Example 3.1.1** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, let $f$ belong to $L^1(\Omega, \mathcal{F}, \mu)$, and define a function $\nu$ on $\mathcal{F}$ by $\nu(A) = \int_A f \, d\mu$. Then the linearity of the integral and the dominated convergence theorem imply that $\nu$ is a signed measure on $(\Omega, \mathcal{F})$. Note that such a signed measure is the difference of the positive measures $\nu_1$ and $\nu_2$ defined by $\nu_1(A) = \int_A f^+ d\mu$ and $\nu_2(A) = \int_A f^- d\mu$.

**Definition 3.1.3 (positive set)** Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu$ be a signed measure on it.

1. $P$ is a *positive set* for $\mu$, or equivalently $P$ is a $\mu$-*positive set*, if for every $E \in \mathcal{F}$ such that $E \subseteq P$, one has $\mu(E) \geq 0$.

2. $N$ is a *negative set* for $\mu$, or equivalently $N$ is a $\mu$-*negative set*, if for every $E \in \mathcal{F}$ such that $E \subseteq N$, one has $\mu(E) \leq 0$.

**Theorem 3.1.1 (Hahn Decomposition Theorem)** Let $(\Omega, \mathcal{F})$ be a measurable space, and let $\mu$ be a signed measure on $(\Omega, \mathcal{F})$. Then there are disjoint subsets $P$ and $N$ of $\Omega$ such that $P$ is a positive set for $\mu$, $N$ is a negative set for $\mu$, and $\Omega = P \cup N$. Let

$$\mu^+ : \mathcal{F} \longrightarrow [0, \infty],$$
$$A \longmapsto \mu(A \cap P)$$

$$\mu^- : \mathcal{F} \longrightarrow [0, \infty],$$
$$A \longmapsto -\mu(A \cap N)$$

$\mu$ is the difference of two measures, that is

$$\mu = \mu^+ - \mu^-,$$

at least one of which is finite.

The *variation* of the signed measure $\mu$ is the measure $|\mu|$ defined by $|\mu| = \mu^+ + \mu^-$. It is easy to check that

$$|\mu(A)| \le |\mu|(A), \quad \forall A \in \mathcal{F}.$$

**Definition 3.1.4 (absolutely continuous)** Let $(\Omega, \mathcal{F})$ be a measurable space, and let $\mu$ and $\nu$ be signed measures on $(\Omega, \mathcal{F})$. Then $\nu$ is *absolutely continuous* with respect to $\mu$ if for all $A \in \mathcal{F}$,

$$\mu(A) = 0 \implies \nu(A) = 0.$$

One usually writes $\nu \ll \mu$ to indicate that $\nu$ is absolutely continuous with respect to $\mu$.

**Definition 3.1.5 (concentrated on a measurable set)** Let $(\Omega, \mathcal{F})$ be a measurable space and $E \in \mathcal{F}$.

1. A measure $\mu$ on $(\Omega, \mathcal{F})$ is *concentrated on $E$* if $\mu(E^c) = 0$.

2. A signed or complex measure $\mu$ on $(\Omega, \mathcal{F})$ is *concentrated on $E$* if the variation $|\mu|$ of $\mu$ is concentrated on $E$, or equivalently, if each $\mathcal{F}$-measurable subset $A$ of $E^c$ satisfies $\mu(A) = 0$.

**Definition 3.1.6 (mutually singular)** Suppose that $\mu$ and $\nu$ are measures (or signed measures) on $(\Omega, \mathcal{F})$. Then $\mu$ and $v$ are *mutually singular* if there is an $\mathcal{F}$-measurable set $E$ such that $\mu$ is concentrated on $E$ and $\nu$ is concentrated on $E^c$.

One usually writes $\mu \perp \nu$ to indicate that $\mu$ and $\nu$ are mutually singular.

**Theorem 3.1.2 (Lebesgue Decomposition Theorem)** Let $(\Omega, \mathcal{F})$ be a measurable space.

1. Let $\mu$ be a measure on $(\Omega, \mathcal{F})$, and let $\nu$ be a $\sigma$-finite measure on $(\Omega, \mathcal{F})$. Then there are unique measures $\nu_a$ and $\nu_s$ on $(\Omega, \mathcal{F})$ such that

   (a) $\nu_a \ll \mu$,
   (b) $\nu_s \perp \mu$,
   (c) $\nu = \nu_a + \nu_s$.

   In this case, $\nu_a$ and $\nu_s$ are $\sigma$-finite measures and called the absolutely continuous and singular parts of $\nu$.

2. Let $\mu$ be a measure on $(\Omega, \mathcal{F})$, and let $\nu$ be a finite signed measure on $(\Omega, \mathcal{F})$. Then there are unique finite signed measures $\nu_a$ and $\nu_s$ on $(\Omega, \mathcal{F})$ such that

   (a) $\nu_a \ll \mu$,
   (b) $\nu_s \perp \mu$,
   (c) $\nu = \nu_a + \nu_s$.

**Theorem 3.1.3 (Radon-Nikodym Theorem)** Let $(\Omega, \mathcal{F})$ be a measurable space.

1. Let $\mu$ and $\nu$ be $\sigma$-finite measures on $(\Omega, \mathcal{F})$. If $v \ll \mu$, then there is an $\mathcal{F}$-measurable function $g : \Omega \to [0, +\infty)$ such that $\nu(A) = \int_A g d\mu$ holds for each $A \in \mathcal{F}$. The function $g$ is unique up to $\mu$-almost everywhere equality.

2. Let $\mu$ be a $\sigma$-finite measure on $(\Omega, \mathcal{F})$, and let $\nu$ be a finite signed measure on $(\Omega, \mathcal{F})$. If $\nu \ll \mu$, then there exists unique $g \in L^1(\Omega, \mathcal{F}, \mu)$ and satisfies $\nu(A) = \int_A g d\mu$ for each $A \in \mathcal{F}$.

# Chapter 4

# Probability Space

## 4.1  Basic Properties

**Proposition 4.1.1** Let $A_1, A_2, \ldots, A_n$ be events in a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Then

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}(-1)^{k-1} \sum_{1 \leq j_1 < j_2 < \cdots j_k \leq n} \mathrm{P}\left(A_{j_1} A_{j_2} \cdots A_{j_k}\right).$$

In particular, we have

$$\mathrm{P}\left(A_1 \cup A_2\right) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 A_2).$$

## 4.2  Independence

**Definition 4.2.1 (independence of events)** Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. A finite number of events $A_1, \cdots, A_n \in \mathcal{F}$ are *independent* if

$$\mathrm{P}\left(A_1 \cdots A_n\right) = \mathrm{P}\left(A_1\right) \cdots \mathrm{P}\left(A_n\right).$$

An infinite collection of events $\{A_i\}_{i \in I}$ is independent if for every finite subset $J$ of $I$, we have

$$\mathrm{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathrm{P}\left(A_i\right).$$

**Definition 4.2.2 (independence of sets of events)** Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. A finite number of sets of events $\mathcal{A}_1, \cdots, \mathcal{A}_n \subset \mathcal{F}$ are independent if for any $A_1 \in \mathcal{A}_1, \cdots, A_n \in \mathcal{A}_n$, $A_1, \cdots, A_n$ are independent.

An infinite collection of sets of events $\{\mathcal{A}_i\}_{i \in I}$ is independent if for every finite subset $J = \{i_1, \cdots, i_j\}$ of $I$, $A_{i_1}, \cdots, A_{i_j}$ are independent.

**Definition 4.2.3 (independence of random variables)** A collection of random variables $\{X_i\}_{i \in I}$ is independent if $\{\sigma\left(X_i\right)\}_{i \in I}$ is independent.

## 4.3  Conditional Expectation

**Definition 4.3.1 (conditional expectation)** Let $X$ be a $\mathcal{F}$-measurable random variable on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ such that $\mathrm{E}[|X|] < \infty$. Given a $\sigma$-algebra $\mathcal{G} \subset \mathcal{F}$, a random variable $Z$ that is $\mathcal{G}$-measurable and satisfies

$$\mathrm{E}[X \mathbf{1}_A] = \mathrm{E}[Z \mathbf{1}_A] \quad \text{for all } A \in \mathcal{G}$$

is called the *conditional expectation* of $Y$ given $\mathcal{G}$ and is written as $\mathrm{E}[X \,|\, \mathcal{G}]$.

In probability theory, we show that conditional expectation exists and is unique up to absolutely surely equality. If not pointed out explicitly, all equalities and inequalities involving conditional expectation are considered to hold absolutely surely.

**Proposition 4.3.1 (projection in $L^2(\Omega, \mathcal{F}, \mathrm{P})$)** Assume $X \in L^2(\Omega, \mathcal{F}, \mathrm{P})$ and $\mathcal{G}$ is a sub-$\sigma$-algebra of $\mathcal{F}$. The set

$$S_{\mathcal{G}} = \{Y \in L^2(\Omega, \mathcal{F}, \mathrm{P}) : \sigma(Y) \in \mathcal{G}\}$$

is a closed subspace of $L^2(\Omega, \mathcal{F}, \mathrm{P})$. The conditional expectation mapping

$$\mathrm{E}[\cdot \,|\, \mathcal{G}] : L^2(\Omega, \mathcal{F}, \mathrm{P}) \longrightarrow S_{\mathcal{G}}$$
$$X \longmapsto \mathrm{E}[X \,|\, \mathcal{G}]$$

is a projection to the subspace $S_{\mathcal{G}}$ in $L^2(\Omega, \mathcal{F}, \mathrm{P})$.

**Proposition 4.3.2** Let $X$, $Y$, $X_n$ be integrable $\mathcal{F}$-measurable random variables on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Assume $\mathcal{G}, \mathcal{H}$ are $\sigma$-algebras such that $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$.

1. If $a$, $b$ are constants, then $\mathrm{E}[aX + bY \,|\, \mathcal{G}] = a\,\mathrm{E}[X \,|\, \mathcal{G}] + b\,\mathrm{E}[Y \,|\, \mathcal{G}]$.

2. If $X$ equals a constant $a$, then $\mathrm{E}[X \,|\, \mathcal{G}] = a$.

3. If $X \geq Y$, then $\mathrm{E}[X \,|\, \mathcal{G}] \geq \mathrm{E}[Y \,|\, \mathcal{G}]$.

4. $|\mathrm{E}[X \,|\, \mathcal{G}]| \leq \mathrm{E}[|X| \,|\, \mathcal{G}]$.

5. If $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function and $\phi(X)$ is integrable, then $\phi\left(\mathrm{E}[X \,|\, \mathcal{G}]\right) \leq \mathrm{E}[\phi(X) \,|\, \mathcal{G}]$.

6. If $\lim\limits_{n \to \infty} X_n = X$ and $|X_n| \leq X$, then $\lim\limits_{n \to \infty} \mathrm{E}[X_n \,|\, \mathcal{G}] = \mathrm{E}[X \,|\, \mathcal{G}]$.

7. $\mathrm{E}[\mathrm{E}[X \,|\, \mathcal{G}]] = \mathrm{E}[X]$.

8. $\mathrm{E}[\mathrm{E}[X \,|\, \mathcal{H}] \,|\, \mathcal{G}] = \mathrm{E}[\mathrm{E}[X \,|\, \mathcal{G}] \,|\, \mathcal{H}] = \mathrm{E}[X \,|\, \mathcal{H}]$

9. If $X$ and $\mathcal{G}$ are independent, that is, whenever $A \in \sigma(X)$ and $B \in \mathcal{G}$, $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$, then $\mathrm{E}[X \,|\, \mathcal{G}] = X$.

10. If $Z$ is $\mathcal{G}$-measurable and $ZX$ is integrable, then $\mathrm{E}[ZX \,|\, \mathcal{G}] = Z\,\mathrm{E}[X \,|\, \mathcal{G}]$

*Proof.* 1. If $a$, $b$ are constants, then $\mathrm{E}[aX + bY \,|\, \mathcal{G}] = a\,\mathrm{E}[X \,|\, \mathcal{G}] + b\,\mathrm{E}[Y \,|\, \mathcal{G}]$.

## 4.4 Conditional Probability

**Definition 4.4.1 (conditional probability)** Let $A, B$ be two events and $\mathrm{P}(B) > 0$. Then

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(AB)}{\mathrm{P}(B)}.$$

Alternatively, we can define the conditional probability as

$$\mathrm{P}(A \mid B) = \mathrm{E}[\mathbf{1}_A \,|\, \sigma(\mathbf{1}_B)] = \mathrm{E}[\mathbf{1}_A \,|\, \sigma(\{B\})].$$

# Appendix

**Definition 4.4.2 (associative algebra)** Let $R$ be a fixed commutative ring. An associative $R$-algebra (or more simply, an $R$-algebra) is an additive abelian group $A$ which has the structure of both a ring and an $R$-module in such a way that the ring multiplication $* : A \times A \to A$ and the scalar multiplication $\cdot : R \times A \to A$ is compatible in the way that

$$r \cdot (x * y) = (r \cdot x) * y = x * (r \cdot y)$$

for all $r \in R$ and $x, y \in A$.

**Definition 4.4.3 (Polish space)** a Polish space is a separable completely metrizable topological space; that is, a space homeomorphic to a complete metric space that has a countable dense subset.

**Proposition 4.4.1 (convergence of random variable sequences)**

$$1 < p < q \begin{cases} X_n \xrightarrow{L^q} X & X_n \xrightarrow{a.s.} X \\ & \qquad\qquad\qquad\qquad X_n \xrightarrow{p} X \dashrightarrow X_n \xrightarrow{d} X \\ X_n \xrightarrow{L^p} X \dashrightarrow X_n \xrightarrow{L^1} X \quad \text{UI} \end{cases}$$

18

# Bibliography

[1] Athreya, Krishna B., and Soumendra N. Lahiri. *Measure theory and probability theory*. Vol. 19. New York: Springer, 2006.