

Measure Theory

This appendix provides a concise overview of the key concepts and results from Measure Theory, it covers more than actually used throughout the book. Some proofs are sketched, especially long but simple points are stated with little details. In this case however precise references are given. We refer to the excellent (Bogachev 2007) for a comprehensive treatment of measure theory.

Measure Spaces

💡 Takeaway of Section

The concepts of measure, volume, perimeter, area, probability and so on, are mathematically formalized by a triple $(\Omega, \mathcal{F}, \mu)$. Ω corresponds to the space where we want to *measure*, it can be any non-empty set. \mathcal{F} is the actual **family of subsets** of Ω that we want to measure, the so called *measurable events*. \mathcal{F} should satisfy some intuitive conditions (e.g if we can measure the occurrence of A and B , we can measure the non-occurrence of A or the occurrence of $A \cup B$). μ is the actual measure (or volume, probability etc.), so it is a map that associates to each measurable set $A \in \mathcal{F}$, a number $\mu(A)$. The basic property is that $\mu(A \sqcup B) = \mu(A) + \mu(B)$.

There are two problems to keep in mind:

- We want $\mu(A \sqcup B) = \mu(A) + \mu(B)$ to hold also for infinite (countable) sums and unions. Because in Math people like to take limits.
- We want to identify measures without the need to specify their values on *each* measurable set. For instance, on the real numbers, we can decide that the length of an intervals $(a, b]$ is $b - a$, and that should be enough to characterize the way we measure lengths on \mathbb{R} .

These two problems complicate matters a bit. In this section we give some basic definitions and results to tackle those problems later.

Sigma-Algebras and Measurable Spaces

Just as topologies provides a rigorous framework for concepts like proximity and convergence, σ -algebras and filtrations formalize the notions of *available information* and *measurability*.

Definition 0.1 (Pi-System). A collection \mathcal{P} of subsets of a set Ω is a **π -system** if it satisfies

1. \mathcal{P} is non-empty.
2. For $A, B \in \mathcal{P}$, $A \cap B \in \mathcal{P}$.

A collection \mathcal{P} of subsets of a set Ω is a **quasi π -system**, or **$q\pi$ -system** for short, if it satisfies

1. \mathcal{P} is non-empty.
2. For $A, B \in \mathcal{P}$, there exists finitely many *pairwise disjoint* $C_1, \dots, C_n \in \mathcal{P}$ such that $A \cap B = \cup_i C_i$.

Definition 0.2 (Lambda-System). A collection \mathcal{L} of subsets of a set Ω is a λ -**system** if it satisfies the following properties:

1. $\emptyset \in \mathcal{L}$.
2. If $A \in \mathcal{L}$, then $A^c \in \mathcal{L}$ (closed under complementation).
3. If $(A_n)_{n=1}^{\infty}$ is a sequence of pairwise disjoint sets in \mathcal{L} , then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{L}$ (closed under countable disjoint unions).

Definition 0.3 (sigma-algebra). A σ -**algebra** on a set Ω is a collection \mathcal{F} of subsets of Ω that is both a $q\pi$ -system and λ -system. Equivalently it satisfies:

1. $\emptyset \in \mathcal{F}$.
2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (closed under complementation).
3. If $(A_n)_{n=1}^{\infty}$ is a sequence of sets in \mathcal{F} , then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ (closed under countable unions).

A **measurable space** is a pair (Ω, \mathcal{F}) , where Ω is a non-empty set and \mathcal{F} is a σ -algebra on Ω .

Remark. The intersection of any collection of σ -algebras (respectively π -systems, λ -systems) on Ω is also a σ -algebra (respectively a π -system, a λ -system). This allows us to define the σ -algebra generated by a family of sets \mathcal{C} , denoted by $\sigma(\mathcal{C})$, as the weakest (also called *coarsest* or *smallest*) σ -algebra containing \mathcal{C} (as well as the π -system and the λ -system generated by \mathcal{C}). It is the intersection of all σ -algebras containing \mathcal{C} . More generally, by the *weakest* σ -algebra such that a given property holds we mean the intersection of all such σ -algebras.

Example 0.1 (Trivial and Discrete sigma-algebras). For any set Ω , the power set (the set of all subsets of Ω) is a σ -algebra, called the **discrete** σ -algebra. This is the strongest (maximal) σ -algebra on Ω .

The collection $\{\emptyset, \Omega\}$ is also a σ -algebra, called the **trivial** σ -algebra. This is the weakest σ -algebra on Ω .

Example 0.2 (Borel sigma-algebras). If (Ω, \mathcal{T}) is a topological space, the σ -algebra $\sigma(\mathcal{T})$ generated by the open sets (equivalently the closed sets) is called the Borel σ -algebra. Its elements are called **Borel sets** or **Borelians**. When \mathcal{T} is understood, the Borel σ -algebra is often denoted $\sigma(\Omega)$. This is usually the case on Euclidean spaces, so that $\sigma(\mathbb{R})$ denotes the Borel σ -algebra generated by the standard topology on \mathbb{R} .

If (Ω, \mathcal{T}) is a *second-countable* topological space, and \mathcal{B} is a base of the topology \mathcal{T} , then the Borel σ -algebra coincides with the one generated by \mathcal{B} , $\sigma(\mathcal{T}) = \sigma(\mathcal{B})$. For instance, on \mathbb{R} the Borel σ -algebra is generated by intervals.

Example 0.3 (Classes of intervals). If $\Omega = \mathbb{R}$, intervals of the form $(a, b]$ form a π -system. The same can be said for open or closed intervals, or intervals of the form $[a, b)$. Similar statements hold for rectangles in \mathbb{R}^2 .

If $\Omega = S^1$, the class of closed arcs of Ω is a $q\pi$ -system but not a π -system. As opposed to the case $\Omega = \mathbb{R}$, the intersection of two arcs can be a union of two arcs.

Theorem 0.1 (Pi-Lambda Theorem). If \mathcal{P} is a $q\pi$ -system and \mathcal{L} is a λ -system such that $\mathcal{P} \subset \mathcal{L}$, then $\sigma(\mathcal{P}) \subset \mathcal{L}$. In other words, the weakest σ -algebra and the weakest λ -system generated by a $q\pi$ -system coincide.

Theorem 0.1. With no loss of generality, we can assume \mathcal{L} to be the weakest λ -system containing \mathcal{P} . For $A \in \mathcal{L}$ define

$$\mathcal{L}_A := \{B \subset \Omega : A \cap B \in \mathcal{L}\}$$

Notice that for $A \in \mathcal{L}$

- $\emptyset \in \mathcal{L}_A$.
- If $B \in \mathcal{L}_A$, then $A \cap B^c = (A^c \cup (A \cap B))^c \in \mathcal{L}$. Namely \mathcal{L} is closed under complementation.
- \mathcal{L}_A is closed under countable disjoint union, since $A \cap (\bigcup_i B_i) = \bigcup_i (A \cap B_i)$.

- If $B \in \mathcal{P}$, then $A \cap B \in \mathcal{L}$ as a finite disjoint union of elements of $\mathcal{P} \subset \mathcal{L}$.

Therefore \mathcal{L}_A is a λ -system containing \mathcal{P} , and by the minimality of \mathcal{L} , we conclude $\mathcal{L}_A \supset \mathcal{L}$. Which in turns means that, for $A, B \in \mathcal{L}$, it holds $A \cap B \in \mathcal{L}$. In other words \mathcal{L} is a π -system and a λ -system, and thus a σ -algebra. \square

Sometimes it is convenient to use algebras and monotone classes instead of σ -algebras, π -systems and λ -systems.

Definition 0.4 (Algebra). A collection \mathcal{A} of subsets of a set Ω is an **algebra** (or **field**) if it satisfies the following properties:

1. $\Omega \in \mathcal{A}$.
2. If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ (closed under complementation).
3. If $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$ (closed under finite unions).

Definition 0.5 (Monotone Class). A collection \mathcal{M} of subsets of a set Ω is a **monotone class** if it satisfies the following properties:

1. If $(A_n)_{n=1}^{\infty}$ is an *increasing* sequence of sets in \mathcal{M} (i.e., $A_1 \subset A_2 \subset A_3 \subset \dots$), then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$ (closed under countable increasing unions).
2. If $(B_n)_{n=1}^{\infty}$ is a *decreasing* sequence of sets in \mathcal{M} (i.e., $B_1 \supset B_2 \supset B_3 \supset \dots$), then $\bigcap_{n=1}^{\infty} B_n \in \mathcal{M}$ (closed under countable decreasing intersections).

Theorem 0.2 (Monotone Class Theorem). *Let \mathcal{A} be an algebra of subsets of a set Ω . Let \mathcal{M} be a monotone class of subsets of Ω . If $\mathcal{A} \subset \mathcal{M}$, then $\sigma(\mathcal{A}) \subset \mathcal{M}$.*

Theorem 0.2. With no loss of generality, assume that \mathcal{M} is the minimal monotone class containing \mathcal{A} . We will prove that \mathcal{M} is a σ -algebra.

For $A \in \mathcal{M}$, let $\mathcal{M}_A = \{B \subset \Omega : A \cap B \in \mathcal{M}, B^c \in \mathcal{M}\}$. It is straightforward to verify that \mathcal{M}_A is a monotone class containing \mathcal{A} . Thus $\mathcal{M}_A \supset \mathcal{M}$, which in particular implies that \mathcal{M} is a π -system. But a monotone class which is also a π -system is a σ -algebra, as can be easily checked. \square

Classes of Measurable Spaces

In this section we review a few properties that define “well-behaved” classes of measurable spaces. This can be viewed as a measurable equivalence of the various regularity properties for topological spaces.

Definition 0.6 (Atoms). In a measurable space (Ω, \mathcal{F}) , the **atoms** are the elements of \mathcal{F} obtained as equivalence classes of the equivalence relation on Ω : $x \sim y$ if and only if

$$\mathbf{1}_A(x) = \mathbf{1}_A(y) \quad \forall A \in \mathcal{F}$$

Definition 0.7 (Hausdorff and Separable spaces). The measurable space (Ω, \mathcal{F}) is called:

- **Hausdorff** if all atoms are singletons (points of Ω).
- **Separable** if \mathcal{F} is generated by a countable collection of sets of Ω .

Example 0.4. \mathbb{R} with the Borel σ -algebra is Hausdorff and separable.

As a counterexample to the Hausdorff property, consider the space $\Omega = \{a, b, c\}$ with the σ -algebra $\mathcal{F} = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$. This space is not Hausdorff because the atoms are $\{a\}$ and $\{b, c\}$.

As a counterexample to the separability, an uncountable space with the discrete σ -algebra is not separable.

Example 0.5. If (Ω, \mathcal{T}) is Hausdorff (topological sense), the Borelian measurable space $(\Omega, \sigma(\mathcal{T}))$ is Hausdorff (measurable sense).

Measurable Functions

Definition 0.8 (Measurable Function). Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces. A function $f : \Omega_1 \rightarrow \Omega_2$ is called **measurable** (or $\mathcal{F}_1/\mathcal{F}_2$ -measurable) if for every $A \in \mathcal{F}_2$, the preimage $f^{-1}(A) = \{\omega \in \Omega_1 : f(\omega) \in A\}$ is in \mathcal{F}_1 .

Proposition 0.1. Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces, and let \mathcal{C} be a collection of subsets in Ω_2 such that $\mathcal{F}_2 = \sigma(\mathcal{C})$. Then f is $\mathcal{F}_1/\mathcal{F}_2$ -measurable if and only if:

$$f^{-1}(E) \in \mathcal{F}_1, \quad \forall E \in \mathcal{C}$$

Proposition 0.1. Set operations are natural under pull-backs (that is a difficult way to say that $f^{-1}(E \cap F) = f^{-1}(E) \cap f^{-1}(F)$, and similar properties for the elementary set operations). Therefore it is easy to check that $\mathcal{F} := \{E \subset \Omega_2 : f^{-1}(E) \in \mathcal{F}_1\}$ is a σ -algebra. Since it contains \mathcal{C} , it contains \mathcal{F}_2 , namely f is $\mathcal{F}_1/\mathcal{F}_2$ -measurable. \square

If $(\Omega_2, \mathcal{F}_2)$ is a measurable spaces, a collection of functions $\mathcal{F} \subset \Omega_2^{\Omega_1}$, induces on Ω_1 a σ -algebra $\mathcal{F}_1 := \sigma(\mathcal{F}) \equiv \{f^{-1}(A) : f \in \mathcal{F}, A \in \mathcal{F}_2\}$: the weakest σ -algebra such that all functions in \mathcal{F} are measurable.

Remark. Checking functions measurability can be a daunting (and somehow boring) task. Some standard ways to easily determine that a function is measurable are:

- The composition of measurable functions is measurable (trivially).
- Continuous functions between topological spaces are measurable with respect to the respective Borel σ -algebras (from Proposition 0.1).
- Let $f_n : \Omega \rightarrow \mathbb{R} \cup -\infty \cup +\infty$ be a sequence of Borel measurable functions. The pointwise supremum and infimum of f_n is also measurable (intervals $(-\infty, c]$ generate the Borel σ -algebra and $\{\sup_n f_n \leq c\} = \bigcap_n \{f_n \leq c\}$). In particular, $\overline{\lim}_n f_n$ and $\underline{\lim}_n f_n$ are measurable.

Nontrivial examples

In general measurability issues do not receive much love in the world of Mathematics, since they are either trivial or artificial. There are of course exceptions.

A typical measurability problem is the following, which arises in optimization theory: for each “random” outcome $\omega \in \Omega$, we have several optimal strategies $A_\omega \subset E$. Can we *measurably* sample an optimal strategy? Namely, can we find a *measurable* map $f : \Omega \rightarrow E$ such that $f(\omega) \in A_\omega$ for $\omega \in \Omega$? Measurability here is crucial, since it is a minimal requirement to practically construct the function. If we have no further information, we can only assume the axiom of choice, which will give us the existence of an optimal $f(\omega) \in E$ with $f(\omega) \in A_\omega$. However, the axiom of choice is the standard way to define non-constructive objects, and typically *non-measurable*. This approach has no interest in the optimization framework. A more interesting (measurable) result is due to Kuratowski and Ryll-Nardzewski, and features many variants, see (Bogachev 2007, Theorem 6.9.3).

Theorem 0.3 (Kuratowski and Ryll-Nardzewski Selection). Let (Ω, \mathcal{F}) be a measurable space and (E, \mathcal{E}) be Polish space, where \mathcal{E} is the Borel σ -algebra. Suppose that we have a map $\Omega \ni \omega \mapsto A_\omega \in \mathcal{E}$ such that

- A_ω is a non-empty closed set of E , for all $\omega \in \Omega$.
- $\{\omega \in \Omega : A_\omega \cap O = \emptyset\} \in \mathcal{F}$ for all O open in E (weak measurability of A_ω).

Then there exists a measurable map $f: \Omega \rightarrow E$ such that $f(\omega) \in A_\omega$ for all $\omega \in \Omega$.

Another example concerns the following problem. Let (E, \mathcal{E}) be a measurable space and Ω non-empty. A function $f: \Omega \rightarrow E$ defines a σ -algebra on Ω , let's call it \mathcal{F} , defined as the weakest σ -algebra such that f is measurable. Let g be another function $g: \Omega \rightarrow E$ and assume that g is \mathcal{F} -measurable. Is it true that g is a function of f ? Namely, does a measurable $h: E \rightarrow E$ such that $g = h \circ f$ exist? One expects an affirmative answer, since for g to be \mathcal{F} -measurable, it has to be determined by f . The answer is indeed yes, but hypotheses are needed: If (E, \mathcal{E}) is a Polish space (or measurably isomorphic to a Polish space), then indeed there exists such an h .

Product of measurable spaces

Definition 0.9 (Product Sigma-Algebra). Let T be an arbitrary index set. For each $t \in T$, let $(\Omega_t, \mathcal{F}_t)$ be a measurable space. For $J \subset T$, define the product space

$$\Omega_J := \prod_{t \in J} \Omega_t$$

and let $\mathcal{F}_J = \prod_{t \in J} \mathcal{F}_t$ be the **product σ -algebra** on Ω_J , defined as the σ -algebra generated by the cylinder sets of the form $\prod_{t \in J} A_t$, where $A_t \in \mathcal{F}_t$ for all $t \in J$ and $A_t = \Omega_t$ for all but finitely many t .

In the previous Definition 0.9, we restricted to a subset $J \subset T$ so that we can also define a σ -algebra on Ω_T (not just on Ω_J) induced by a subset $J \subset T$. This is the weakest σ -algebra on Ω_T such that the canonical projection $\pi_J^T: \Omega_T \rightarrow \Omega_J$ is measurable. That is, we can consider on Ω_T the σ -algebra $\{(\pi_J^T)^{-1}(A), A \in \mathcal{F}_J\}$. With a somehow confusing notation, this σ -algebra is usually still denoted \mathcal{F}_J .

Measures

Definition 0.10 (Measure). Let (Ω, \mathcal{F}) be a measurable space (i.e., Ω is a set and \mathcal{F} is a σ -algebra on Ω). A **measure** μ on (Ω, \mathcal{F}) is a function $\mu: \mathcal{F} \rightarrow [0, \infty]$ such that:

1. $\mu(\emptyset) = 0$.
2. (Countable Additivity) If $(A_n)_{n=1}^\infty$ is a sequence of pairwise disjoint sets in \mathcal{F} (i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$), then

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

A triple $(\Omega, \mathcal{F}, \mu)$, where (Ω, \mathcal{F}) is a measurable space and μ a measure, is called a **measure space**.

A measure μ is a **probability measure** if $\mu(\Omega) = 1$. In such a case $(\Omega, \mathcal{F}, \mu)$ is called a **probability space**.

A measure μ is **σ -finite** if there exists a countable collection of measurable sets $\Omega_i \in \mathcal{F}$ such that $\Omega = \bigcup_i \Omega_i$ and $\mu(\Omega_i) < \infty$.

A set function with properties similar to a measure, but defined on a collection of set which is not necessarily a σ -algebra is called a **content**. In the literature, this is also called a **premeasure**, in particular when the σ -algebra is replaced by an algebra.

Definition 0.11 (Content). Let Ω be a non-empty set and \mathcal{A} a collection of subsets of Ω with $\emptyset \in \mathcal{A}$. A map $\mu: \mathcal{A} \rightarrow [0, \infty]$ is an **outer content** if

- $\mu(\emptyset) = 0$ (or equivalently $\mu(\emptyset) < \infty$).
- If $A, A_1, \dots, A_N \in \mathcal{A}$ with $A \subset \bigcup_{i=1}^N A_i$, then $\mu(A) \leq \sum_{i=1}^N \mu(A_i)$.

μ is a **content** if the last condition is replaced by the stronger

- $\mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$ for any finite collection $(A_i)_{i=1}^n$ of pairwise disjoint elements of \mathcal{A} such that $\cup_{i=1}^n A_i \in \mathcal{A}$.

μ is a σ -**additive content** if the last condition holds on countable collections:

- $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ for any countable collection $(A_i)_i$ of pairwise disjoint elements of \mathcal{A} such that $\cup_i A_i \in \mathcal{A}$.

Proposition 0.2. Let (Ω, \mathcal{F}) be a measurable space, and let μ and ν be two probability measures on \mathcal{F} . If \mathcal{P} is a π -system such that $\sigma(\mathcal{P}) = \mathcal{F}$, and $\mu(A) = \nu(A)$ for all $A \in \mathcal{P}$, then $\mu = \nu$ on \mathcal{F} .

Proposition 0.2. It is easy to check that the collection of sets on which the two probability measures coincide is a λ -system. Since it contains a π system that generates \mathcal{F} , they coincide on the entire σ -algebra by Theorem 0.1. \square

Example 0.6. Let $\Omega = \mathbb{R}$, and let $\mathcal{P} = \{(-\infty, a] : a \in \mathbb{R}\}$. \mathcal{P} is a π -system. The σ -algebra generated by \mathcal{P} , $\sigma(\mathcal{P})$, is the Borel σ -algebra on \mathbb{R} . Let \mathcal{L} be the collection of all subsets $A \subset \mathbb{R}$ such that A or A^c is countable or has a countable complement. Then \mathcal{L} is a λ -system, and $\mathcal{P} \subset \mathcal{L}$. By the Pi-Lambda Theorem, $\sigma(\mathcal{P}) \subset \mathcal{L}$.

Furthermore, by Proposition 0.2, if two probability measures μ and ν on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ agree on all intervals of the form $(-\infty, a]$, then $\mu = \nu$. The (cumulative) distribution function F_μ of a probability measure μ is defined by $F_\mu(x) = \mu((-\infty, x])$, so this shows that the distribution function uniquely determines the probability measure.

Pushforward of Measures

If (Ω, \mathcal{F}) and (E, \mathcal{E}) are measurable spaces, then a measurable $f: \Omega \rightarrow E$ can be lifted to a linear map $f_\#$ mapping measures on Ω to measures on E as

$$f_\# \mu := \mu \circ f^{-1}$$

which means, for $A \in \mathcal{E}$

$$(f_\# \mu)(A) := \mu(\{x \in \Omega : f(x) \in A\})$$

Notice that $f_\#$ maps probabilities to probabilities.

Remark. If we let $\wp(\Omega)$ be the spaces of probability measures over Ω , we can consider an injection $x \ni \Omega \mapsto \delta_x \in \wp(\Omega)$, where the **Dirac mass at x** is defined as the probability measure $\delta_x(A) = \mathbf{1}_A(x)$. The map $f_\#$ then features the following properties

- $f_\# \delta_x = \delta_{f(x)}$
- $f_\#(\alpha\mu + (1 - \alpha)\nu) = \alpha f_\# \mu + (1 - \alpha) f_\# \nu$

$$\begin{array}{ccc}
 \wp(\Omega) & \xrightarrow{\circ f^{-1}} & \wp(E) \\
 \uparrow \delta_x & & \uparrow \delta_e \\
 \Omega & \xrightarrow{f} & E
 \end{array}$$

We will see when [discussing topologies of probabilities](#) that these two properties characterize $f_\#$ among continuous maps $\wp(\Omega) \rightarrow \wp(E)$.

Continuity of Finite Measures

Given a sequence (or a net) (A_n) , we can define

$$\overline{\lim}_n A_n := \bigcap_{k \in \mathbb{N}} \bigcup_{n \geq k} A_n, \quad \underline{\lim}_n A_n := \bigcup_{k \in \mathbb{N}} \bigcap_{n \geq k} A_n$$

Whenever $\overline{\lim}_n A_n = \underline{\lim}_n A_n$ we say that A_n converges and denote $\lim_n A_n$ the limit. Notice that monotone sequences converge.

Remark. Equivalently, $\overline{\lim}_n A_n$ is the set of points $x \in \Omega$ such that $x \in A_n$ for infinitely many n . $\underline{\lim}_n A_n$ is the set of $x \in \Omega$ such that $x \notin A_n$ for finitely many n .

Proposition 0.3. Let $(\Omega, \mathcal{F}, \mu)$ be a finite measurable space, A_n a measurable sequence. It is not hard to check that $\underline{\lim}_n A_n$, $\overline{\lim}_n A_n$ are measurable and

$$\underline{\lim}_n \mu(A_n) \geq \mu(\underline{\lim}_n A_n), \quad \overline{\lim}_n \mu(A_n) \leq \mu(\overline{\lim}_n A_n) \quad (1)$$

In particular whenever A_n converges, $\lim_n \mu(A_n) = \mu(\lim_n A_n)$.

Denote now $A \Delta B$ the symmetric difference of A and B . Consider the *equivalence* relation on \mathcal{F} given by $A \sim B$ whenever $\mu(A \Delta B) = 0$. It is easily seen that, if $A \sim A'$ and $B \sim B'$, then $\mu(A \Delta B) = \mu(A' \Delta B')$. Thus the function

$$d_\mu(A, B) := \mu(A \Delta B)$$

passes to the quotient \mathcal{F} / \sim , and it is easily seen to be a distance, called the μ -Boolean distance. From Equation 1, we see that $(\mathcal{F} / \sim, d_\mu)$ is a **complete** metric space.

Regular Measure Spaces

Definition 0.12 (Regular Measure). Let (Ω, \mathcal{F}) be a measurable space where Ω is a topological space. A measure μ on (Ω, \mathcal{F}) is

- **Outer regular** if, for every $A \in \mathcal{F}$,

$$\mu(A) = \inf\{\mu(U) : U \supset A, U \text{ open and measurable}\}$$

- **Inner regular** if, for every $A \in \mathcal{F}$

$$\mu(A) = \sup\{\mu(C) : C \subset A, C \text{ compact and measurable}\}$$

- **regular** if it is both outer and inner regular. That is, for every $A \in \mathcal{F}$ and $\varepsilon > 0$, there exist an open set $U \in \mathcal{F}$ and a compact set $K \in \mathcal{F}$ such that $K \subset A \subset U$ and $\mu(U \setminus K) < \varepsilon$.

If Ω is a Polish space (a separable, completely metrizable topological space) equipped with its Borel σ -algebra, see Example 0.2, then every finite measure on $(\Omega, \mathcal{B}(\Omega))$ is regular. Similar statements hold on compact metric spaces, and on locally compact second-countable Hausdorff spaces. More generally, every locally finite measure on a Polish space is regular. In other words, under mild assumptions on the topology of Ω , regularity of *all* finite Borel measures holds.

Complete Measure Spaces

Definition 0.13 (Negligible set). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. A subset $A \subset \Omega$ is called **negligible** if it is contained in a set of measure 0.

Definition 0.14 (Complete Measure Space). A measure space $(\Omega, \mathcal{F}, \mu)$ is **complete** if every negligible set is measurable (and has necessarily measure 0).

Definition 0.15 (Almost Everywhere Equivalence). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. We say that a property holds **almost everywhere** (μ -a.e., or simply a.e. when the measure is clear from context) if it holds for all $\omega \in \Omega \setminus N$, where N is a μ -negligible set.

Two measurable functions $f, g : \Omega \rightarrow \mathbb{R}$ are **almost everywhere equivalent** (or μ -a.e. equivalent) if $f(\omega) = g(\omega)$ for μ -a.e. $\omega \in \Omega$; that is, if $\mu(\{\omega : f(\omega) \neq g(\omega)\}) = 0$. We often treat a.e. equivalent functions as identical, especially in the context of integration.

Proposition 0.4 (Completion of a Measure Space). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Then there exists a complete measure space $(\Omega, \overline{\mathcal{F}}, \overline{\mu})$, called the **completion** of $(\Omega, \mathcal{F}, \mu)$, such that:

1. $\mathcal{F} \subset \overline{\mathcal{F}}$.
2. $\overline{\mu}(A) = \mu(A)$ for all $A \in \mathcal{F}$.
3. For every $E \in \overline{\mathcal{F}}$, there exist sets $A, B \in \mathcal{F}$ such that $A \subset E \subset B$ and $\mu(B \setminus A) = 0$.

Moreover, the completion is unique in the following sense: If $(\Omega, \mathcal{F}', \mu')$ is another complete measure space satisfying (1), (2) and (3), then $\mathcal{F}' = \overline{\mathcal{F}}$ and $\mu' = \overline{\mu}$.

Proposition 0.4. Define

$$\overline{\mathcal{F}} = \{E \cup N : E \in \mathcal{F}, N \subset A \text{ for some } A \in \mathcal{F} \text{ with } \mu(A) = 0\}. \quad (2)$$

In other words, $\overline{\mathcal{F}}$ consists of sets that can be formed by taking a measurable set E and adding a negligible set. $\overline{\mathcal{F}}$ is a σ -algebra since:

1. $\Omega \in \mathcal{F} \subset \overline{\mathcal{F}}$.
2. Let $E \cup N \in \overline{\mathcal{F}}$, where $E \in \mathcal{F}$, $N \subset A$, and $\mu(A) = 0$. Then

$$(E \cup N)^c = E^c \cap N^c = (E^c \cap A^c) \cup (E^c \cap (A \setminus N)) \quad (3)$$

and $E^c \cap A^c \in \mathcal{F}$ while $(E^c \cap (A \setminus N)) \subset A$ is negligible.

3. Let $(E_i \cup N_i)_i$ be a sequence in $\overline{\mathcal{F}}$, where $E_i \in \mathcal{F}$, $N_i \subset A_i$, and $\mu(A_i) = 0$. Then $\cup_i (E_i \cup N_i) = (\cup_i E_i) \cup (\cup_i N_i)$ is also in $\overline{\mathcal{F}}$ since $\mu(\cup_i A_i) = 0$.

We next define $\overline{\mu}$ on $\overline{\mathcal{F}}$ by

$$\overline{\mu}(E \cup N) = \mu(E) \quad (4)$$

for $E \in \mathcal{F}$ and $N \subset A$ with $\mu(A) = 0$. It is easy to see that $\overline{\mu}$ is well-defined: suppose $E_1 \cup N_1 = E_2 \cup N_2$, where $E_1, E_2 \in \mathcal{F}$, $N_1 \subset A_1$, $N_2 \subset A_2$, and $\mu(A_1) = \mu(A_2) = 0$. Then $E_1 \subset E_2 \cup A_2$, thus $\mu(E_1 \setminus E_2) = 0$ and similarly $\mu(E_2 \setminus E_1) = 0$. Thus

$$\mu(E_1) = \mu(E_1 \cap E_2) = \mu(E_2). \quad (5)$$

This shows that $\overline{\mu}$ is well-defined. Now, we show that $\overline{\mu}$ is a measure on $\overline{\mathcal{F}}$.

1. $\overline{\mu}(\emptyset) = \mu(\emptyset) = 0$.

2. Let $(E_i \cup N_i)_i$ be a sequence of pairwise disjoint sets in $\overline{\mathcal{F}}$. Then

$$\bar{\mu}(\cup_i (E_i \cup N_i)) = \mu(\cup_i E_i) = \sum_i \mu(E_i) = \sum_i \bar{\mu}(E_i \cup N_i). \quad (6)$$

Therefore $\bar{\mu}$ is a measure. It is also clear from construction that it is complete. The points (1), (2), (3) are trivial, as well as uniqueness. □

Compactly approximated measures

We first introduce the notion of compact class.

Definition 0.16 (Compact Class). A family \mathcal{K} of subsets of a set Ω is called a **compact class** if, for any sequence (K_n) of elements of \mathcal{K} with $\cap_n K_n = \emptyset$, there exists a finite N such that $\cap_{n=1}^N K_n = \emptyset$.

Remark. A well known fact in elementary topology (the finite intersection property), states that, in a Hausdorff topological space, the collection of compact subsets forms a compact class. That is the origin of the name *compact class*.

Definition 0.17. Let Ω be a non-empty set, \mathcal{A} a collection of subsets of Ω , μ an outer content on (Ω, \mathcal{A}) and \mathcal{K} a compact class of subsets of Ω . We say that μ is **compactly approximated by \mathcal{K}** if for every $A \in \mathcal{A}$ and $\varepsilon > 0$, there exist $K_\varepsilon \in \mathcal{K}$, $A_\varepsilon \in \mathcal{A}$ such that $A \setminus K_\varepsilon \subset A_\varepsilon$ and $\mu(A_\varepsilon) \leq \varepsilon$.

Example 0.7. An inner regular measure on a topological space is compactly approximated by the collection of compact sets, see Definition 0.12.

Lemma 0.1. A compactly approximated (by some compact class \mathcal{K}) outer content is continuous at 0. Namely, if $A_n \in \mathcal{A}$ is a decreasing sequence with $\cap_n A_n = \emptyset$, then $\lim_n \mu(A_n) = 0$.

Lemma 0.1. Fix $\varepsilon > 0$, and let $K_n \in \mathcal{K}$ and $B_n \in \mathcal{A}$ be such that $K_n \subset A_n$, $A_n \setminus B_n = \emptyset$ and $\mu(B_n) \leq \varepsilon 2^{-n}$. Then $\cap_n K_n \subset \cap_n A_n = \emptyset$. Thus for some finite N , $\cap_{n=1}^N K_n = \emptyset$. Therefore

$$A_N \cap (\cup_{n=1}^N B_n)^c = \cap_{n=1}^N (A_n \cap B_n^c) \subset \cap_{n=1}^N (A_n \cap (A_n^c \cup K_n)) \subset \cap_{n=1}^N K_n = \emptyset$$

Namely $A_N \subset \cup_{n=1}^N B_n$ and since μ is an outer content

$$\mu(A_N) \leq \sum_{n=1}^N \mu(B_n) \leq \varepsilon$$

□

Extension Theorems

In this section we briefly described some results that easily allow us to define measures on a space, by only defining their values on a specific class of sets. Let us consider the following problems.

- We want to define a measure on \mathbb{R} . The most intuitive approach would consist in defining its values on intervals such as $(a, b]$ (for σ -finite measures) or $(-\infty, a]$ (for finite measures, since the measure of this interval could be identically ∞ for non-finite measures). Would these values ‘be enough’ to characterize the measure on the whole σ -algebra they generate? Caratheodory’s Extension Theorem Theorem 0.4 answers this question (positively).
- We want to define a measure on functions, say on the space $E^{\mathbb{N}} = \{X : \mathbb{N} \rightarrow E\}$. Certainly we need to know the measure of sets of the form $\{X \in E^{\mathbb{N}} : (X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n)\}$. Since these sets are characterized by finitely many conditions $X_i \in A_i$, in many cases their measure can be described explicitly. If we know the measure on sets of this form, can we extend it to a measure on the whole product σ -algebra of $E^{\mathbb{N}}$? Kolmogorov’s Extension Theorem Theorem 0.5 answers this question (positively under mild conditions).

Caratheodory’s Extension Theorem

💡 Takeaway of Section

In order to define a measure μ on a measurable space (Ω, \mathcal{F}) it is enough to define:

- A **quasi semiring** \mathcal{S} , that generates the σ -algebra \mathcal{F} .
- A **content** μ_0 on \mathcal{S} , that identifies μ on the quasi semiring.
- A **compact class** \mathcal{K} on Ω , that compactly approximates μ_0 on \mathcal{S} .

Moreover if μ_0 is σ -finite on \mathcal{S} , the measure μ on (Ω, \mathcal{F}) that coincides with μ_0 on \mathcal{S} is unique.

Examples: $\Omega = \mathbb{R}$, \mathcal{F} is the Borel σ -algebra, \mathcal{S} is the collection of intervals $(a, b]$, \mathcal{K} is the class of compact sets (or simply compact intervals).

Definition 0.18 (Quasi Semiring). A **quasi semiring** of sets \mathcal{S} on a set Ω is a collection of subsets of Ω that satisfies:

1. $\emptyset \in \mathcal{S}$.
2. If $A, B \in \mathcal{S}$, then both $A \cap B$ and $A \setminus B$ are finite disjoint unions of elements of \mathcal{S} . Namely there exist *pairwise disjoint* sets B_1, \dots, B_n and *pairwise disjoint* $C_1, \dots, C_n \in \mathcal{S}$ such that $A \cap B = \cup_i B_i$, $A \setminus B = \cup_i C_i$.

Theorem 0.4 (Caratheodory’s Extension Theorem). *Let \mathcal{S} be a quasi semiring of subsets of Ω , and let $\mu_0 : \mathcal{S} \rightarrow [0, \infty]$ be a σ -additive content, see Definition 0.11. Then there exists a measure μ (called the extension of μ_0) on $\sigma(\mathcal{S})$ such that $\mu(A) = \mu_0(A)$ for all $A \in \mathcal{S}$. If μ_0 is σ -finite, then the extension is unique.*

Theorem 0.4. We give a sketch of the proof, which can be found in (Patriota 2011).

1. **Outer content.** For any $E \subset \Omega$, define the *outer content*

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu_0(A_i) \mid A_i \in \mathcal{S}, E \subset \cup_{i=1}^{\infty} A_i \right\}. \quad (7)$$

where the infimum of the empty set is understood to be $+\infty$. It is easy to see that $\mu^*(\emptyset) = 0$, and that μ^* is monotone, namely $\mu^*(A) \leq \mu^*(B)$ if $A \subset B$. We next show that $\mu^*(E) \leq \sum_{n=1}^{\infty} \mu^*(E_n)$ (subadditivity). Let $E = \bigcup_{n=1}^{\infty} E_n$. Without loss of generality, assume that $\mu^*(E_n) < \infty$ for all n . Fix $\varepsilon > 0$. For each n , there exists a sequence $(A_{n,i})_{i \in \mathbb{N}}$ in \mathcal{S} such that $E_n \subset \cup_i A_{n,i}$ and $\sum_i \mu_0(A_{n,i}) < \mu^*(E_n) + 2^{-n-1}\varepsilon$. Then $E \subset \cup_n \cup_i A_{n,i}$, so

$$\mu^*(E) \leq \sum_n \sum_i \mu_0(A_{n,i}) < \sum_n (\mu^*(E_n) + 2^{-n-1}\varepsilon) = \sum_n \mu^*(E_n) + \varepsilon \quad (8)$$

2. **The σ -algebra \mathcal{F}^* .** Define the collection of sets

$$\mathcal{F}^* := \{E \subset \Omega : \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c), \forall A \subset \Omega\} \quad (9)$$

Since μ^* is subadditive, in the last formula the equality can be equivalently replaced with $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$. It is then an elementary task to check that \mathcal{F}^* is a σ -algebra.

3. $\mathcal{F}^* \supset \mathcal{S}$. First notice that, in the Equation 9, the covering A_i can be taken to be disjoint. With this in mind, one can prove that for $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$ for all $A \in \Omega$ and $E \in \mathcal{S}$.
4. μ^* **restricted to \mathcal{F}^* is a measure.** We skip the proof of this lengthy but elementary check.
5. **Extension existence.** Since μ^* is a measure on the σ -algebra \mathcal{F}^* containing \mathcal{S} , it induces a measure on $\sigma(\mathcal{S})$. We need to show $\mu^*(S) = \mu_0(S)$ for all $S \in \mathcal{S}$. Clearly, $\mu^*(S) \leq \mu_0(S)$ by definition, while we skip the elementary but lengthy prove of the opposite inequality.
6. **Extension uniqueness.** Assume μ_0 is σ -finite, with $\Omega = \bigcup_{n=1}^{\infty} \Omega_n$, $\Omega_n \in \mathcal{S}$ increasing, and $0 < \mu_0(\Omega_n) < \infty$. Let μ and ν be two extensions of μ_0 to $\sigma(\mathcal{S})$. Since \mathcal{S} is a $q\pi$ -system, μ and ν coincide on Ω_n , namely $\mu(A \cap \Omega_n) = \nu(A \cap \Omega_n)$ for $A \in \sigma(\mathcal{S})$, in view of Theorem 0.1. Since we assumed the Ω_n monotone, one then deduces the equality taking the limit in n .

□

Remark. It is worth noting that for the uniqueness of the extension, it is necessary that μ_0 is σ -finite. It is not enough (for uniqueness) that an extension of μ_0 is σ -finite.

Proposition 0.5. *If a content μ on quasi semiring \mathcal{S} is compactly approximated, then it is σ -additive.*

Proposition 0.5. Suppose that $A_1, \dots, A_m \in \mathcal{S}$, $B_1, \dots, B_n \in \mathcal{S}$, are two pairwise disjoint collections of elements of \mathcal{S} such that $A := \bigcup_i A_i = \bigcup_j B_j$. Then $C_{i,j} := A_i \cap B_j$ can be written as a disjoint union of $K_{i,j}$ elements $C_{i,j,k}$ of \mathcal{S} . Therefore, since μ is additive

$$\sum_i \mu(A_i) = \sum_{i,j} \sum_k \mu(C_{i,j,k}) = \sum_j \mu(B_j) := \mu(A)$$

This allows us to consistently extend $\mu(\cdot)$ to a content on the collection \mathcal{S}^u of disjoint unions of elements of \mathcal{S} . We can define an outer content μ^* on the collection \mathcal{S}^* of finite (possibly non disjoint) unions of elements of \mathcal{S} , as in Equation 7. We have already remarked in the proof of Theorem 0.4, that in Equation 7 we can restrict to pairwise disjoint A_i , see (Patriota 2011, Proposition 1.1), therefore $\mu^*(A) = \mu(A)$ for $A \in \mathcal{S}^u$.

It is easy to check that μ^* on \mathcal{S}^* can be compactly approximated (see (Bogachev 2007 Proposition 1.12.4)). From Lemma 0.1, μ^* is continuous at 0 in \mathcal{S}^u .

Let now $A_i \in \mathcal{S}$ be a pairwise disjoint sequence such that $\bigcup_i A_i =: A \in \mathcal{S}$. Then $\bigcup_{i>n} A_i = \bigcup_{i \leq n} (A \cap A_i^c)$ and each $A \cap A_i^c$ is in \mathcal{S}^u since \mathcal{S} is a quasi semi-ring. Therefore $\bigcup_{i>n} A_i \in \mathcal{S}^*$ and

$$\mu(\bigcup_{i \geq 1} A_i) = \mu^*(\mu(\bigcup_{i \geq 1} A_i)) \leq \mu^*(\bigcup_{i>n} A_i) + \mu^*(\bigcup_{i \leq n} A_i) = \mu^*(\bigcup_{i>n} A_i) + \sum_i \mu(\bigcup_{i \leq n} A_i)$$

and $\mu^*(\bigcup_{i>n} A_i) \rightarrow 0$ as $n \rightarrow \infty$.

□

Example 0.8 (Lebesgue measure). On $\Omega = \mathbb{R}$, consider the semiring \mathcal{A} of intervals $[a, b)$, and the content $\hat{\lambda}([a, b)) = b - a$. Then $\hat{\lambda}$ is compactly approximated by the compact class of finite closed intervals, thus it is σ -additive on \mathcal{A} by Proposition 0.5. By Theorem 0.4, $\hat{\lambda}$ extends uniquely to a σ -additive measure λ_{Borel} on the Borel σ -algebra of \mathbb{R} . Recalling Proposition 0.4, the completion of the Borel σ -algebra w.r.t. λ_{Borel} is called the Lebesgue σ -algebra; the completion λ of λ_{Borel} is called the Lebesgue measure.

Kolmogorov's Extension Theorem

💡 Takeaway of Section

If we want to define a probability measure on a space of functions or sections, it is enough to define its finite-dimensional projections, provided these projections are regular enough (compactly approximated). This is a minimal requirement which always holds on reasonable topological spaces, like Polish spaces.

Kolmogorov's Extension Theorem provides conditions for the existence of a probability measure on an infinite product space, given a consistent family of measures on finite-dimensional projections.

Theorem 0.5 (Kolmogorov's Extension Theorem). *Let T be an arbitrary index set. For each $t \in T$, let $(\Omega_t, \mathcal{F}_t)$ be a measurable space. For $J \subset T$, define the product space*

$$\Omega_J := \prod_{t \in J} \Omega_t \quad (10)$$

and let \mathcal{F}_J be the product σ -algebra on Ω_J , see Definition 0.9. For subsets $I \subset J \subset T$, let $\pi_I^J : \Omega_J \rightarrow \Omega_I$ denote the canonical projection map.

Suppose that for each finite subset $J \Subset T$, we are given a compactly approximated (see Definition 0.17) probability measure μ_J on $(\Omega_J, \mathcal{F}_J)$. Then the following are equivalent

- The **consistency condition** holds, namely for any finite subsets $I \subset J \Subset T$

$$\mu_I = (\pi_I^J)_* \mu_J, \quad (11)$$

where $(\pi_F^G)_* \mu_G$ is the pushforward measure defined by $((\pi_F^G)_* \mu_G)(A) = \mu_G((\pi_F^G)^{-1}(A))$ for $A \in \mathcal{F}_F$.

- There exists a (necessarily unique) probability measure μ on the product space $\Omega_T := \prod_{t \in T} \Omega_t$ such that

$$\mu_J = (\pi_J^T)_* \mu \quad (12)$$

for all finite subsets $J \Subset T$.

Remark 0.1. The consistency condition is clearly necessary for the existence of a measure μ on Ω_T satisfying Equation 12. Indeed its pushforwards $(\mu_F)_{F \Subset T}$ defined by the canonical projections trivially satisfy Equation 11, since $\pi_G^F \circ \pi_F^T = \pi_G^T$ and therefore

$$\mu_F := \mu \circ (\pi_G^T)^{-1} = \mu \circ (\pi_F^T)^{-1} \circ (\pi_G^F)^{-1} = \mu_G \circ (\pi_F^G)^{-1} \quad (13)$$

The theorem thus states that the compactly approximated condition implies the opposite implication.

Theorem 0.5. Let \mathcal{R} be the semiring of subsets of Ω_T of the form

$$\prod_{t \in J} A_t \times \prod_{t \in T \setminus J} \Omega_t$$

for any finite $J \Subset T$ and $(A_t \in \mathcal{F}_t)_{t \in J}$. We define a set function μ on \mathcal{R} by $\mu(C_B^J) = \mu_J(B)$. The consistency condition ensures that μ is well-defined and that μ is a content, see Definition 0.11 on \mathcal{R} .

Let \mathcal{K} be the collection of subsets of Ω_T of the form

$$\prod_{t \in J} K_t \times \prod_{t \in T \setminus J} \Omega_t$$

for any finite $J \in T$ and $(K_t \in \mathcal{K}_t)_{t \in J}$. \mathcal{K} is easily seen to be a compact class, see Definition 0.16. Moreover μ is compactly approximated by \mathcal{K} on \mathcal{R} . By Lemma 0.1, μ is continuous at 0 on \mathcal{R} . This in turn yields that μ is σ -additive in view of Proposition 0.5. Since \mathcal{R} is a semiring, and thus a quasi semiring, we conclude by Theorem 0.4. \square

Integration and Lebesgue Spaces

💡 Takeaway of Section

Measure-theoretic integrals provide a general definition that includes sums, series, classical Riemann integrals, expected values of random variables, and more. It works as one would expect, that is like elementary sums (e.g. it is linear, monotone). The precise mathematical characterization however, allows us to be very precise with convergence, see Theorem 0.12.

Measure-theoretic integral

Let $(\Omega, \mathcal{F}), \mu$ be a measure space and let $f: \Omega \rightarrow [0, \infty)$. If $f = \sum_j a_j 1_{A_j}$ takes finitely-many values a_1, \dots, a_n , respectively on the measurable sets A_1, \dots, A_n , then f is called **simple** or **discrete**. For such functions we set

$$\mu(f) \equiv \int f d\mu := \sum_j a_j \mu(A_j)$$

$\mu(f)$ is easily seen to depend on f and not on the values (a_j) and measurable (A_j) used to represent it as a sum.

For any (not necessarily simple) function $f: \Omega \rightarrow [0, \infty]$, we have two reasonable definitions for the **integral of f with respect to μ** :

$$\mu(f) \equiv \int_{\Omega} f(x) d\mu(x) := \sup \{ \mu(g) : g \leq f, g \text{ simple} \} \quad (14)$$

$$\mu(f) \equiv \int_{\Omega} f(x) d\mu(x) := \inf \{ \mu(g) : g \geq f, g \text{ simple} \} \quad (15)$$

Definition 0.19 (Integral). If $f: \Omega \rightarrow [0, \infty]$ is measurable the r.h.s. of Equation 14 and Equation 15 coincide (as can be easily checked), and the quantity $\mu(f)$ is called the **integral of f with respect to μ** .

If $f = f^+ - f^-$ is not necessarily positive, one then sets $\mu(f) = \mu(f^+) - \mu(f^-)$, provided the two integrals are not both $+\infty$.

If $\mu(|f|) < \infty$, f is called **μ -integrable**, or **integrable** if the measure μ is understood from the context.

It is immediate to verify that the integral is

- **linear:** $\mu(\alpha f + g) = \alpha \mu(f) + \mu(g)$.
- **monotone:** if $f \geq g$ then $\mu(f) \geq \mu(g)$.
- **compatible with the μ -a.e. equivalence relation:** if $f = g$ μ -a.e. then $\mu(f) = \mu(g)$.

The notation $\mu(f)$ is particularly useful to stress the linearity of the operation, as in some context measures and functions are in duality. Notice that the integral may be $+\infty$ even if f is finite.

Example 0.9. If Ω is a countable space endowed with the discrete σ -algebra (or more in general if the measurable space (Ω, \mathcal{F}) has at most countably many atoms), then the integral writes as a (possibly finite, at most countable)

$$\mu(f) := \sum_{x \in \Omega} f(x) \mu(\{x\})$$

Example 0.10. If $\Omega = \mathbb{R}$, μ is the Lebesgue measure and f is Riemann integrable, then $\mu(f) = \int f(x) dx$ coincides with the classical Riemann integral.

Theorem 0.6 (Fubini's Theorem). *Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two σ -finite measure spaces, and let $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$ be the product measure space. Let $f: \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ be a measurable function.*

If f is integrable $\int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \otimes \mu_2) < \infty$ then the iterated integrals are equal to the integral over the product space:

$$\int_{\Omega_1} \left(\int_{\Omega_2} f(x, y) d\mu_2(y) \right) d\mu_1(x) = \int_{\Omega_2} \left(\int_{\Omega_1} f(x, y) d\mu_1(x) \right) d\mu_2(y) = \int_{\Omega_1 \times \Omega_2} f(x, y) d(\mu_1 \otimes \mu_2)(x, y)$$

Lebesgue spaces

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. For $p \in (0, \infty]$, the quotient space (here \sim denotes the a.e. equivalence Definition 0.15)

$$L^p(\Omega, \mathcal{F}, \mu) := \left\{ f: \Omega \rightarrow \mathbb{R} \text{ measurable} : \int |f|^p d\mu < \infty \right\} / \sim, \quad p < \infty$$

$$L^\infty(\Omega, \mathcal{F}, \mu) := \{ f: \Omega \rightarrow \mathbb{R} \text{ measurable} : \exists C > 0 : |f| \leq C \mu\text{-a.e.} \} / \sim, \quad p = \infty$$

is called the Lebesgue space with exponent p .

Proposition 0.6. *For $p \in [1, \infty]$, $L^p(\Omega, \mathcal{F}, \mu)$ is a Banach space when equipped with the norm*

$$\|f\|_{L^p} := \left(\int |f|^p d\mu \right)^{1/p}, \quad p < \infty$$

$$\|f\|_{L^\infty} := \inf \{ C \geq 0 : |f| \leq C \mu\text{-a.e.} \}, \quad p = \infty$$

Moreover if $\mu(\Omega) < \infty$, $\|f\|_{L^\infty} = \lim_{p \rightarrow \infty} \|f\|_{L^p}$.

Decomposition of Measures

💡 Takeaway of Section

If μ is a reference measure on a measurable space, we can decompose any other measure as follows.

$$\nu = \varrho \mu + \nu^s$$

where ν^s is singular with respect to μ , in other words ν only weights sets of μ -measure 0. While $\varrho \mu$ should be interpreted as in Equation 16.

Definition 0.20 (Absolutely Continuous and Singular Measures). Let (Ω, \mathcal{F}) be a measurable space. A measure ν is said to be **absolutely continuous** with respect to a measure μ (denoted $\nu \ll \mu$) if for every $A \in \mathcal{F}$, $\mu(A) = 0$ implies $\nu(A) = 0$. If $\nu \ll \mu$ and $\mu \ll \nu$, namely if they have the same sets of 0 measure, they are called **equivalent**.

On the opposite side, μ and ν are **mutually singular** (denoted $\mu \perp \nu$) if there exists a set $A \in \mathcal{F}$ such that $\mu(A) = 0$ and $\nu(A^c) = 0$.

Remark. On \mathbb{R} with the Borel σ -algebra, a measure ν is absolutely continuous with respect to Lebesgue measure μ if and only if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any finite collection of disjoint intervals (a_i, b_i) with $\sum_i (b_i - a_i) < \delta$, we have $\sum_i \nu((a_i, b_i)) < \varepsilon$. This corresponds to the elementary notion of absolute continuity for functions.

Let $(\Omega, \mathcal{F}, \mu)$ be a σ -finite measure space, and $\varrho: \Omega \rightarrow [0, \infty]$ is measurable, then the function

$$\nu(A) := \int_A \varrho d\mu, \quad A \in \mathcal{F} \quad (16)$$

is a measure (Ω, \mathcal{F}) , as can be easily checked using the monotone convergence Theorem 0.10. It satisfies

$$\int f d\nu = \int f \varrho d\mu \quad f \geq 0, \text{ measurable} \quad (17)$$

This measure is denoted by $\nu = \varrho\mu$ or, in view of Equation 17, the notation $d\nu = \varrho d\mu$ is also used.

Theorem 0.7 (Radon-Nikodym derivative). Let $(\Omega, \mathcal{F}, \mu)$ be a σ -finite measure space, and let ν be another measure on (Ω, \mathcal{F}) . The following are equivalent

- ν is absolutely continuous with respect to μ , $\nu \ll \mu$.
- There exists a measurable function $\varrho: \Omega \rightarrow [0, \infty]$ such that $\nu = \varrho\mu$, namely Equation 16 holds.

The function ϱ is called the **Radon-Nikodym derivative** of ν with respect to μ , and is denoted by $\frac{d\nu}{d\mu}$, and it is unique up to μ -a.e. identification.

Moreover, if any of the two equivalent conditions above holds, the following are also equivalent

- ϱ is finite, namely takes values in $[0, \infty)$ (up to its irrelevant definition on a μ -negligible set).
- ν is σ -finite.

Finally, for σ -finite measures $\nu \ll \mu \ll \lambda$, the **chain rule** holds

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \frac{d\mu}{d\lambda} \quad (18)$$

Theorem 0.7. We give a short sketch of the proof, we refer to (Bogachev 2007, 3.2; Conway 2012, 113) for the interested reader. If $\nu = \varrho\mu$, then certainly $\nu(A) = 0$ whenever $\mu(A) = 0$. So we only need to prove that, if $\nu \ll \mu$, there exists ϱ such that $\nu = \varrho\mu$. We can also restrict to the case where μ is finite, since one can prove the theorem for the restriction of μ to sets of finite measure, and get the σ -finite case as an immediate consequence. We also restrict to ν being finite for the sake of simplicity.

Consider the set $\mathcal{E} := \{f \in L^0(\Omega, \mu) : f\mu \leq \nu\}$ and define

$$\varrho(x) = \sup_{f \in \mathcal{E}} f(x)$$

\mathcal{E} is non-empty since the zero function is in \mathcal{E} . Moreover if $f_1, f_2 \in \mathcal{E}$, then $\max(f_1, f_2) \in \mathcal{E}$. From this, one can deduce that $\varrho \in \mathcal{E}$. Therefore one only needs to check that $\varrho\mu(A) - \nu(A) \geq 0$ for all $A \in \mathcal{F}$, to conclude. \square

Theorem 0.8 (Lebesgue Decomposition Theorem). Let $(\Omega, \mathcal{F}, \mu)$ be a σ -finite measure space, and let ν be another σ -finite measure on (Ω, \mathcal{F}) . Then there exists a unique decomposition $\nu = \nu_{ac} + \nu_s$, where ν_{ac} and ν_s are σ -finite measures on (Ω, \mathcal{F}) such that $\nu_{ac} \ll \mu$ (absolutely continuous) and $\nu_s \perp \mu$ (singular).

Theorem 0.8. Since ν and μ are σ -finite, so is $\nu + \mu$. Moreover, $\mu \ll \nu + \mu$. By the Radon-Nikodym theorem, there is a measurable function ϱ such that

$$\mu(A) = \int_A \varrho d(\nu + \mu), \quad \forall A \in \mathcal{F} \quad (19)$$

Let $B = \{\omega : \varrho(\omega) = 0\}$ and define ν_{ac} and ν_s by

$$\nu_{ac}(A) = \nu(A \cap B^c), \quad \nu_s(A) = \nu(A \cap B), \quad \forall A \in \mathcal{F} \quad (20)$$

Clearly $\nu = \nu_{ac} + \nu_s$, and $\nu_s(B^c) = 0$ and $\mu(B) = 0$. That is ν_s and μ are singular.

Let now A be such that $\mu(A) = 0$. Then

$$0 = \mu(A \cap B^c) = \int_{A \cap B^c} h d(\nu + \mu) \geq \int_{A \cap B^c} h d\nu$$

Since $h > 0$ on B^c , necessarily $\nu(A \cap B^c) = 0$. Namely $\nu_{ac} \ll \mu$.

To show uniqueness, suppose $\nu = \nu'_{ac} + \nu'_s$ is another such decomposition. Then $\nu_{ac} - \nu'_{ac} = \nu'_s - \nu_s$. The left-hand side is absolutely continuous with respect to μ , while the right-hand side is singular with respect to μ . The only measure that is both absolutely continuous and singular with respect to μ is the zero measure. \square

Convergence results

Considering σ -additivity measures is a strong restriction when compared to just additive measures. The reason for developing such a theory lies in the ensuing powerful results about convergence.

Definition 0.21 (Convergence of measurable functions). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let (E, d) be a metric space (equipped with its Borel σ -algebra). A sequence (or net) (f_n) of measurable functions $f_n : \Omega \rightarrow E$ converges to f

- **μ -a.e.** if $f_n(x) \rightarrow f(x)$ outside a set of measure 0. Namely

$$\mu \left(\{x \in \Omega : \overline{\lim}_n d(f_n(x), f(x)) \neq 0\} \right) = 0$$

- **in μ -measure** if for each $\varepsilon > 0$

$$\overline{\lim}_n \mu \left(\{x \in \Omega : \overline{\lim}_n d(f_n(x), f(x)) > \varepsilon\} \right) = 0$$

- **in $L^p(\Omega, \mathcal{F}, \mu)$** if

$$\lim_n \int d(f_n(x), f(x))^p d\mu(x) = 0$$

Theorem 0.9. Let $(\Omega, \mathcal{F}, \mu)$, (E, d) f_n, f as in Definition 0.21. The following holds

- If $f_n \rightarrow f$ μ -a.e., then $f_n \rightarrow f$ in measure.
- If $f_n \rightarrow f$ in L^p , then $f_n \rightarrow f$ in measure.

- If $f_n \rightarrow f$ in measure, then there exists a subsequence f_{n_k} that converges to f μ -a.e..
- If $f_n \rightarrow f$ in L^p , then there exists a subsequence f_{n_k} that converges to f μ -a.e..

Theorem 0.10 (Monotone Convergence Theorem). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. If $(f_n)_{n=1}^\infty$ is a sequence of non-negative measurable functions such that $f_n(\omega) \uparrow f(\omega)$ for μ -a.e. $\omega \in \Omega$, then $\int f_n d\mu \uparrow \int f d\mu$.

Theorem 0.11 (Dominated Convergence Theorem). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. If $(f_n)_{n=1}^\infty$ is a sequence of measurable functions such that $f_n(\omega) \rightarrow f(\omega)$ for μ -a.e. $\omega \in \Omega$, and there exists an integrable function g (i.e., $\int |g| d\mu < \infty$) such that $|f_n(\omega)| \leq g(\omega)$ for all n and μ -a.e. ω , then $\int f_n d\mu \rightarrow \int f d\mu$.

Theorem 0.12 (Vitali Convergence Theorem (finite measure)). Let $(\Omega, \mathcal{F}, \mu)$ be a finite measure space, and let $(f_n)_{n=1}^\infty$ be a sequence of functions in $L^p(\Omega, \mathcal{F}, \mu)$ for $1 \leq p < \infty$. Then, f_n converges to f in $L^p(\Omega, \mathcal{F}, \mu)$ iff

- The sequence (f_n) converges in measure to f .
- The sequence $(|f_n|^p)$ is uniformly integrable:

$$\lim_{M \rightarrow \infty} \sup_n \int_{|f_n| \geq M} |f_n|^p d\mu = 0$$

Theorem 0.13 (Vitali Convergence Theorem (general measure)). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and let $(f_n)_{n=1}^\infty$ be a sequence of functions in $L^p(\Omega, \mathcal{F}, \mu)$ for $1 \leq p < \infty$. Then f_n converges to f in $L^p(\Omega, \mathcal{F}, \mu)$ iff:

- The sequence (f_n) converges in measure to f .
- The functions $(|f_n|^p)$ are uniformly integrable.
- For every $\varepsilon > 0$, there exists a set E of finite measure, such that $\int_{E^c} |f_n|^p < \varepsilon$ for all n .

Theorem 0.14 (Borel-Cantelli Lemma). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and let (A_n) be a sequence of measurable sets. If $\sum_{n=1}^\infty \mu(A_n) < \infty$, then the measure of the limsup of the sets is zero:

$$\mu \left(\limsup_{n \rightarrow \infty} A_n \right) = \mu \left(\bigcap_{m=1}^\infty \bigcup_{n=m}^\infty A_n \right) = 0.$$

This means that the set of points that belong to infinitely many A_n has measure zero.

Bogachev, Vladimir I. 2007. *Measure Theory*. Springer.

Conway, J. B. 2012. *A Course in Abstract Analysis*. Graduate Studies in Mathematics. American Mathematical Society.

Patriota, Alexandre Galvao. 2011. "An Extended Version of the Caratheodory Extension Theorem." <https://api.semanticscholar.org/CorpusID:211837752>.