Measure Theory

Bharath Ron

Table of Contents

Table of Contents 1

1 Measure Spaces 2
- 1.1 Vitali's Theorem . . . 2
- 1.2 Sigma Algebras & Measures . . . 3
  - 1.2.1 Pre-Measures & Measures . . . 5
  - 1.2.2 Completion of Measures . . . 7
- 1.3 Outer Measures . . . 8
  - 1.3.1 Caratheodory Extension Theorem . . . 8
    - 1.3.1.1 Uniqueness Theorem . . . 9
- 1.4 Borel σ-Algebras . . . 10
  - 1.4.1 Lebesgue-Stieltjes Measures . . . 10
    - 1.4.1.1 The Lebesgue Measure . . . 11

2 Lebesgue Integration 13
- 2.1 Measurable Functions . . . 13
  - 2.1.1 Approximation by Simple Functions . . . 14
- 2.2 The Lebesgue Integral . . . 15
- 2.3 Convergence of the Lebesgue Integral . . . 15
  - 2.3.1 Monotone Convergence Theorem . . . 16
    - 2.3.1.1 Linearity of the Integral . . . 16
    - 2.3.1.2 Fatou's Lemma . . . 18
  - 2.3.2 Dominated Convergence Theorem . . . 18
- 2.4 Product Measures . . . 19
  - 2.4.1 Product Sigma Algebras . . . 19
  - 2.4.2 The Fubini-Tonelli Theorem . . . 19

3 Signed Measures & Differentiation 20
- 3.1 Decomposition Theorems . . . 20
  - 3.1.1 Hahn Decomposition . . . 20
  - 3.1.2 Jordan Decomposition . . . 20
- 3.2 The Radon-Nikodym Derivative . . . 20
  - 3.2.1 Absolute Convergence . . . 20
  - 3.2.2 Radon-Nikodym Theorem . . . 20
  - 3.2.3 Lebesgue Decomposition Theorem . . . 20
- 3.3 Differentiation . . . 20
- 3.4 Radon Measures . . . 20

Bibliography 21

22

1 | Measure Spaces

A measure on a set X assigns size to subsets of X. The aim of this chapter is to identify and formalise the minimum structure needed to define such a measure. The empty set should have zero *size*, and hence we should have $\mu(\varnothing) = 0$. For disjoint sets, the total measure should be sum of measures of the individual sets.

$$\mu\Big(\coprod_{i\in F} A_i\Big) = \textstyle\sum_{i\in F} \mu(A_i),$$

where F is a finite set. If this holds when i is indexed by a countable set C, as a limit, then μ is called σ-additive. We will denote finite sets and countable sets by F and C respectively.

1.1 | Vitali's Theorem

One way to measure sets is to use geometric notions such as length, area and volume. This notion of size however cannot measure the sizes of all sets while also maintaining the geometric properties. To demostrate the contradiction consider the case $X = \mathbb{R}$ and the notion of size comes from the geometric notion of length of intervals.

Let $\mathcal{I}_{\mathbb{R}}$ denote the set of all intervals of the form $(a, b] \subset \mathbb{R}$ where $a < b$. On this set, we can define the measure of the interval to be the length of the interval $\lambda : I_{\mathbb{R}} \to [0, \infty]$ defined by,

$$\lambda\big((a, b]\big) = b - a$$

This notion of size is invariant under translations. We will now show that the expected property of translation invariance is incompatible with σ-additivity.

Firstly we need to show that this is a well-behaved and is indeed σ-additive. Let $\{(a_i, b_i]\}_{i\geq 0}$ be a collection of disjoint sets in $\mathcal{I}_{\mathbb{R}}$ with $\coprod_{i\geq 0}(a_i, b_i] = (a, b]$. Any finite subcollection can be ordered such that $a \leq a_1 < b_1 \leq \cdots \leq a_N < b_N \leq b$. This gives us, $\sum_{N\geq i\geq 0} \lambda((a_i, b_i]) \leq b-a$, since $\sum_{i\in C} \lambda((a_i, b_i]) = \lim\limits_{N\to\infty} \sum_{i=1}^{N} \lambda((a_i, b_i])$, we have,

$$\textstyle\sum_{i\in C} \lambda\big((a_i, b_i]\big) \leq b - a$$

It can be showed by induction that for any finite cover $\bigcup_{i\in C}(a_i, b_i] \supseteq (a, b]$, $b-a \leq \sum_{i\in C}(b_i - a_i)$. For the infinite cover case, one reduces it to the finite cover case, using compactness of the closed interval and then applying Heine-Borel theorem. Now since $\{(a_i, b_i]\}_{i\geq 0}$ is a cover we have,

$$b - a \leq \textstyle\sum_{i\in C} \lambda((a_i, b_i])$$

So,

$$\lambda\Big(\coprod_{i\in C}(a_i, b_i]\Big) = \textstyle\sum_{i\in C} \lambda\big((a_i, b_i]\big).$$

where C is countable. Hence λ is σ-additive.

However the problem is that the length function cannot be defined for all subsets of $\mathbb{R}$ while also respecting translation invariance, i.e., $\lambda(A+t) = \lambda(A)$ for $t \in \mathbb{R}$ and $A \subset \mathbb{R}$. This was shown by Vitali, the proof of which involves construction of the Vitali set which if λ was both translation invariant and σ-additive would have both infinite measure and be a subset of a set with finite size which is absurd.

Theorem 1.1.1. (Vitali's theorem) $\lambda : \mathcal{I}_{\mathbb{R}} \to [0, \infty]$ *cannot be extended to* $\mathcal{P}(\mathbb{R})$.

Proof

To construct the Vitali set, define an equivalence relation on $\mathbb{R}$ where $a \sim b$ iff $b - a \in \mathbb{Q}$. There are uncountably many equivalence classes as each equivalence class contains countably many elements. Choose a representative of these equivalence class that lie in $[0,1]$. This collection of representatives is called the Vitali set, denoted by V, and clearly, $V \subset [0,1]$.

Take $C = \mathbb{Q} \cap [-1,1]$, note that this is a countable set. Consider the translates of the Vitali set, $V_i = \{v + i \mid v \in V\}$ for $i \in C$. Since each element of V differ only by a irrational number by definition of V. Hence this collection of subsets is mutually disjoint. Notice that,

$$(0,1] \subseteq \coprod_{i \in C} V_i \subseteq (-1,2].$$

Expecting $\lambda(A) \leq \lambda(A) + \lambda(B \backslash A) = \lambda(B)$ for any $A \subseteq B$, and expecting the translation invariance, we should have,

$$1 = \lambda\big((0,1]\big) \leq \lambda\Big(\coprod_{i \in C} V_i\Big) \leq \lambda((-1,2]) = 3\lambda((0,1]) = 3.$$

By σ-additivity we have, $\lambda(\coprod_C V_i) = \sum_C \lambda(V_i) = \sum_C \lambda(V) = \infty \cdot \lambda(V)$. If we assume $\lambda(V)$ is finite we will have $\infty \leq 3$ and if assume $\lambda(V)$ is zero we will have $1 \leq 0$, which is absurd. $\square$

This means that the length function cannot give meaningfully measure all subsets of $\mathbb{R}$ while satisfying both translation invariance and σ-additivity. The moral is that we cannot expect all subsets to be measurable. The next best goal is then to define a notion of measure coming from the length function to a large subset $\mathcal{B}_{\mathbb{R}}$ of $\mathcal{P}(\mathbb{R})$.

1.2 | Sigma Algebras & Measures

σ-algebras act as the family of subsets of a set X that we can *measure*. So the definition of a σ-algebra should be such that it has the expected properties and leaves out pathologies.

Let $\Sigma(X) \subset \mathcal{P}(X)$ be a collection of subsets of X that can be *measured*. We expect the size of *nothing* to be zero, to make this precise we should first need the empty set to be measurable. Similarly, the whole set X should be measurable, could have infinite measure but it must be measurable.

$$\varnothing \in \Sigma(X), \ \ X \in \Sigma(X). \tag{empty set}$$

If A is measurable then we expect the size of A^c to be the size of X minus the size of A. So, the set A^c must be measurable.

$$A \in \Sigma(X) \Rightarrow A^c \in \Sigma(X). \tag{complements}$$

If two sets $A \in \Sigma(X)$ and $B \in \Sigma(X)$ are measurable then we expect $A \cup B$ to be measurable and if they are disjoint the total size should be sum of the individual sizes,

$$A \cup B \in \Sigma(X) \tag{finite union}$$

$\Sigma(X) \subset \mathcal{P}(X)$ is called an algebra if it satisfies conditions (empty set), (complements), and (finite union).

Since our aim is to maximize the stuff we can measure, we can extend this to countable union as well while still having nice behavior to do analysis. So, we allow countable union also to be measurable.

$$\{A_i\}_{i \in C} \subset \Sigma(X) \Rightarrow \bigcup_{i \in C} A_i \in \Sigma(X). \tag{countable union}$$

where C is countable. An algebra $\Sigma(X) \subset \mathcal{P}(X)$ that also satisfies the condition (countable union) is called a σ-algebra. The pair $(X, \Sigma(X))$ where $\Sigma(X)$ is a σ-algebra is called a measurable space. Clearly, $\mathcal{P}(X)$ and $\{\varnothing, X\}$ are σ-algebras on X and $I_{\mathbb{R}}$ is not a σ-algebra on $\mathbb{R}$. If $\Sigma(X)$ is a σ-algebra of X and $X' \subset X$ then $X' \cap \Sigma(X)$ is a σ-algebra of X'.

Directly from the definition it follows by De-Morgan's laws that standard set theoretic operations such as union, intersection, difference, symmetric difference are also in $\Sigma(X)$.

$$A, B \in \Sigma(X), \quad A \cup B,\ A \cap B,\ A \backslash B,\ A \Delta B \in \Sigma(X)$$

In order to be able to work with countable collections we need a notion of limit. Given a countable collection of sets $\{A_i\}_{i \in C} \subseteq \Sigma(X)$, $\limsup$ is the collection of elements that are in infinitely many A_is. Similarly, the $\liminf$ is the collection of elements contained in all but finitely many A_is. This implies,

$$\liminf_{i \in C} A_i \subseteq \limsup_{i \in C} A_i.$$

Since limsup and liminf can be expressed as countable unions and countable intersections[1], $\limsup$ and $\liminf$ belong to $\Sigma(X)$ if $\Sigma(X)$ is a σ-algebra. This allows us to do analysis with σ-algebras. A countable collection of subsets $\{A_i\}_{i \in C} \subset \Sigma(X)$ is said to have a limit if $\limsup_{i \in C} A_i = \liminf_{i \in C} A_i$, and it is defined to be the limit of the collection.

Suppose $\Sigma_i(X) \subset \mathcal{P}(X)$ be σ-algebras, then it can be checked that $\Sigma(X) = \cap_i \Sigma_i(X)$ is also a σ-algebra. For any collection of subsets $S \subset \mathcal{P}(X)$, the σ-algebra generated by S, denoted by $\sigma(S)$ is the intersection of all σ-algebras containing S.

$$\sigma(S) = \bigcap_{S \subset \Sigma(X)} \Sigma(X).$$

It is hence the smallest σ-algebra containing S. If $S \subset \Sigma(X)$ then $\sigma(S) \subset \Sigma(X)$. Given a collection of subsets S, we can construct $\sigma(S)$ as follows, include all the elements of S, add all complements of elements of S, add X and $\varnothing$, add all countable unions and their complements. Since any σ-algebra containing S will have these sets it's the desired intersection.

Theorem 1.2.1. *$f : Y \to X$ be a function, then,*

$$f^{-1}(\sigma(S)) = \sigma(f^{-1}(S)).$$

[1] $\limsup_{i \in C} A_i = \bigcap_{m \geq 0} \bigcup_{n \geq 0} A_n$, and $\liminf_{i \in C} A_i = \bigcup_{m \geq 0} \bigcap_{n \geq 0} A_n$.

Proof

The first step is to show that $f^{-1}(\sigma(S))$ is a σ-algebra, which is a basic set theory exercise. This σ-algebra contains $f^{-1}(S)$ and hence $\sigma(f^{-1}(S)) \subseteq f^{-1}(\sigma(S))$. To prove the other side, consider the set,

$$\Sigma = \{A \in \sigma(S) \mid f^{-1}(A) \in \sigma(f^{-1}(S))\},$$

Σ contains S, and it can be showed with some basic set theory that Σ is a σ-algebra. So $\sigma(S) \subseteq \Sigma$ and $f^{-1}(\sigma(S)) \subseteq \sigma(f^{-1}(S))$. $\square$

A useful fact to keep in mind is that the pre-image map commutes with set-operations: $f^{-1}(B^c) = (f^{-1}(B))^c$, for any $\mathcal{J}$, $f^{-1}\big(\bigcup_{i\in\mathcal{J}} B_i\big) = \bigcup_{i\in\mathcal{J}}\big(f^{-1}(B_i)\big)$,and $f^{-1}\big(\bigcap_{i\in\mathcal{J}} B_i\big) = \bigcap_{i\in\mathcal{J}}\big(f^{-1}(B_i)\big)$.

1.2.1 | Pre-Measures & Measures

Although the collection of intervals $\mathcal{I}_{\mathbb{R}}$ is not a σ-algebra, it has some nice properties. Since we intend to expend the length function on intervals to a larger class of subsets of $\mathbb{R}$ it is important to understand the properties of $\mathcal{I}_{\mathbb{R}}$. We are interested in abstracting out some properties of the length function on intervals.

The empty set can be taken to have zero length or the empty set belongs to $\mathcal{I} = \mathcal{I}_{\mathbb{R}}$.

$$\varnothing \in \mathcal{I} \tag{empty set}$$

Intersection of two intervals is again an interval, so we have,

$$A, B \in \mathcal{I} \Rightarrow A \cap B \in \mathcal{I} \tag{intersection}$$

An interval minus an interval is the disjoint union of two intervals. It is possible to do analysis with a simple generalisation, and assume disjoint union of finitely many sets. So in general we want,

$$A \backslash B = \coprod_{i\in F_B} C_i,\ \ C_i \in \mathcal{I}. \tag{disjoint union}$$

where F_B is a finite set.

A subset $\mathcal{I} \subset \mathcal{P}(X)$ that satisfies the conditions (empty set), (intersection), and (disjoint union) is called a semi-ring. Since the properties of abstract semi-rings involve boring induction based proofs it is easier and intuitive to think about these properties in terms of intervals.

If $\mathcal{I} \subseteq \mathcal{P}(X)$ is a semi-ring, and $A \in \mathcal{I}$, and $B_1, \dots B_m \in \mathcal{I}$ be a finite collection, then there exist a finite collection $C_1, \dots C_n \in \mathcal{I}$ disjoint, such that

$$A \backslash \Big(\bigcup_{j=1}^{m} B_j\Big) = \coprod_{n\geq i\geq 1} C_i.$$

It is better to visualise these properties in terms of intervals. This is saying that if we remove a finite collection of intervals out of an interval the resultant set is a disjoint union of finite collection of intervals.

If $\mathcal{I} \subseteq \mathcal{P}(X)$, $\mathcal{H} \subseteq \mathcal{P}(Y)$ are two semi-rings, then it can be showed that $\mathcal{I} \times \mathcal{H}$ is a semi-ring, the proof involves some basic set theory manipulations, but again it is simple to think in terms of rectangles in $\mathbb{R}^2$. We can now define pre-measure on semirings that captures the intuition we have about the length function on $\mathcal{I}_{\mathbb{R}}$.

The idea of the notion of length of intervals can now be generalised to semi-rings. Given a semi-ring $\mathcal{I} \subseteq \mathcal{P}(X)$ a pre-measure on $\mathcal{I}$ is a map, $\mu : \mathcal{I} \to [0, \infty]$ with $\mu(\varnothing) = 0$ which is σ-additive, that is

$$\mu\Big(\coprod_{i \in C} A_i\Big) = \textstyle\sum_{i \in C} \mu(A_i), \qquad (\sigma\text{-additivity})$$

where C is countable. If μ only satisfies

$$\mu\Big(\coprod_{i \in F} A_i\Big) = \textstyle\sum_{i \in F} \mu(A_i),$$

where F is finite, it is called a content on $\mathcal{I}$. We say μ is a measure if μ is σ-additive and $\mathcal{I}$ is a σ-algebra. A measure space is a triple $(X, \Sigma(X), \mu)$ where $\Sigma(X)$ is a σ-algebra of the set X and μ is a measure on $\Sigma(X)$. $(X, \Sigma(X), \mu)$ is called finite if $\mu(X) < \infty$. If $X = \bigcup_{i \in C} A_i$ where $\{A_i\}_{i \in C} \subset \Sigma(X)$ and $\mu(A_i) < \infty$ for all $i \in C$, then $(X, \Sigma(X), \mu)$ is called σ-finite.

Let $\mathcal{I}$ be a semi-ring and μ a pre-measure on $\mathcal{I}$. To study the behavior of the pre-measure μ on set theoretic operations, we reduce the sets of interests into disjoint unions of sets belonging to $\mathcal{I}$ and use the σ-additivity on the new disjoint unions.

Suppose $A, B \in \mathcal{I}$ with $B \subseteq A$, then $A = B \coprod (A \backslash B)$. Since $\mathcal{I}$ is a semi-ring we can write $A \backslash B = \coprod_{i \in F_B} C_i$ where $\{C_i\}_{F_B} \subset \mathcal{I}$ and F_B is a finite set. This is easier to visualise in terms of intervals, removal of an interval gives rise to disjoint intervals. The finite additivity gives us,

$$\mu(A) = \mu(B) + \textstyle\sum_{i \in F_B} \mu(C_i) \geq \mu(B). \qquad (\text{monotonicity})$$

Similarly, the unions of intervals can be thought of as a disjoint union of the intervals without the intersecting parts and their intersections. If $A \cup B \in \mathcal{I}$ then we have, $A = (A \cap B) \cup (A \backslash B)$. Now $A \backslash B$ is a disjoint union $\coprod_{i \in F_B} C_i$ for $C_i \in \mathcal{I}$. So, we have $A = (A \cap B) \cup \coprod_{i \in F_B} C_i$ and similarly $B = (A \cap B) \cup \coprod_{j \in F_A} D_j$ for $D_j \in \mathcal{I}$. So we have,

$$A \cup B = (A \cap B) \coprod \Big(\textstyle\coprod_{i \in F_B} C_i\Big) \coprod \Big(\textstyle\coprod_{j \in F_A} D_j\Big).$$

This gives us,

$$\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B). \qquad (\text{parallelogram})$$

Let $\{A_i\}_{i \in C} \subset \mathcal{I}$ with $\cup_{i \in C} A_i \in \mathcal{I}$ then we can write $\cup_{i \in C} A_i$ as a disjoint union by taking $A_i' = A_i \backslash (A_1 \cup \cdots \cup A_{i-1})$. Now each of these is a union of disjoint sets $A_i' = \coprod_{n \geq k} C_{k,i}$. So we have $\bigcup_{i \in C} A_i = \coprod_{k,i} C_{k,i}$ which gives,

$$\mu\Big(\bigcup_{i \in C} A_i\Big) \leq \textstyle\sum_{i \in C} \mu(A_i). \qquad (\sigma\text{-subadditivity})$$

Let $\{A_i\}_{i \in C} \subset \mathcal{I}$ and $A_i \nearrow A$ then by taking $A_i' = A_i \backslash (A_1 \cup \cdots \cup A_{i-1}) = \coprod_{m_i \geq k} C_{k,i}$. Then A can be written as a countable union,

$$A = \bigcup_{i \in C} A_i' = \bigcup_{i \in C} \coprod_{m_i \geq k} C_{k,i}$$

Let $A_N = \bigcup_{N \geq n} \bigcup_{m_i \geq k} C_{k,i}$. This is a disjoint union and hence we have,

$$\mu(A) = \textstyle\sum_{n \geq 1} \sum_{m_i \geq k} \mu(C_{k,i}) = \lim_{N \to \infty} \sum_{n=1}^{N} \sum_{k=1}^{m_i} \mu(C_{k,i}) = \lim_{N \to \infty} \mu(A_N)$$

$$\lim_{n \to \infty} \mu(A_n) = \mu(A) \qquad (\text{cont. from below})$$

Continuity from above is proved similarly. First step in most of the proofs like this involve writing the union as a disjoint union and then applying σ-additivity.

1.2.2 | Completion of Measures

A set $A \in \Sigma(X)$ is called null if $\mu(A) = 0$. A property of the space is said to hold μ-**almost everywhere** if the set where the property does not hold is a null set. The collection of all measurable μ-null sets by $\mathcal{N}_\mu$. Since $\Sigma(X)$ is a σ-algebra and assuming $\infty \cdot 0 = 0$, we assume that $\mathcal{N}_\mu$ is closed under countable unions. If $(X, \Sigma(X), \mu)$ is a measure space $\mu(A) = 0$ and $B \subset A$ with $B \in \Sigma(X)$ then the monotonicity of μ implies $\mu(B) = 0$. If all subsets of μ-null sets belong to $\Sigma(X)$ then $(X, \Sigma(X), \mu)$ is called **complete**.

It is however possible to extend the measure such that all subsets of null sets are measurable. The first step then is to construct a σ-algebra that includes all the subsets of μ-nullsets. Let

$$\overline{\Sigma}(X) = \{A \cup B \mid A \in \Sigma(X), B \subset N \in \mathcal{N}_\mu\}.$$

This includes all the subsets of μ-null sets. Since $\varnothing \in \mathcal{N}_\mu$, we also have that $\Sigma(X) \subseteq \overline{\Sigma}(X)$. We intend to find an extension of μ to $\overline{\Sigma}(X)$.

Lemma 1.2.2. *$\overline{\Sigma}(X)$ is a σ-algebra.*

Proof

Clearly $\varnothing \in \overline{\Sigma}(X)$, because it already belonged to $\Sigma(X)$. We now have to verify that $\overline{\Sigma}(X)$ satisfies (complements) and (countable union). Both $\Sigma(X)$ and $\mathcal{N}_\mu$ are closed under countable unions. Hence so is $\overline{\Sigma}(X)$. To show that it satisfies (complements), suppose $\overline{A} \in \overline{\Sigma}$, then $\overline{A}$ is of the form $A \cup N$ for some $N \subset M$, with $\mu(M) = 0$. Assume that $A \cap M = \varnothing$ or replace A by $A \backslash N$. Then $A \cup N = (A \cup M) \cap (M^c \cup N)$. By De Morgans law, we have

$$(A \cup N)^c = (A \cup M)^c \cup (M \backslash N)$$

Since $(A \cup M)^c \in \Sigma(X)$ and $(M \backslash N) \in \mathcal{N}_\mu$, it follows that $(A \cup N)^c \in \overline{\Sigma}(X)$. □

$\overline{\Sigma}(X)$ is the smallest σ-algebra containing both $\Sigma(X)$ and $\mathcal{N}_\mu$.

Theorem 1.2.3. *There exists a unique extension $\overline{\mu}$ of μ to a complete measure on $\overline{\Sigma}(X)$.*

Proof

If $A \cup N \in \overline{\Sigma}(X)$, set $\overline{\mu}(A \cup N) = \mu(A)$. This depends on the choice of how we split $A \cup N$. So we need to show this definition is well defined. Suppose $A_1 \cup N_1 = A_2 \cup N_2$ for $A_i \in \Sigma(X)$ and $N_i \in \mathcal{N}_\mu$.

We will show the well-definedness by showing $\mu(A_1) \leq \mu(A_2)$ and $\mu(A_2) \leq \mu(A_1)$. Firstly, note that there exists some M_i with $\mu(M_i) = 0$ such that $N_i \subset M_i$. Let $M = M_1 \cup M_2 \in \mathcal{N}_\mu$. So, we have, $A_2 \subset A_1 \cup M$. This Rightarrow by subadditivity,

$$\mu(A_2) \leq \mu(A_1) + \mu(M) = \mu(A_1).$$

Similarly, we have $\mu(A_1) \leq \mu(A_2)$. So, the definition is well-defined. Since $\overline{\Sigma}(X)$ includes all the subsets of ν-null sets, and $\overline{\mu}$ is determined by μ, it is a complete measure. Uniqueness follows from positivity and subadditivity. □

We will assume that all the measures of interest to us are complete.

1.3 | Outer Measures

The length function on $\mathcal{I}_{\mathbb{R}}$ is a pre-measure (since we motivated the definition of semi-rings and pre-measures based on the properties of the length function this holds trivially). It is called the Lebesgue-Borel pre-measure and the goal is to extend it to a large enough σ-algebra. The aim for this section is to develop tools needed to extend pre-measure on semi-ring to a measure on some appropriate, large enough σ-algebra. The approach is to extend the pre-measure μ to a set function μ^* on $\mathcal{P}(X)$. From this, pick a large σ-algebra $\Sigma(\mu^*)$ on which μ^* is σ-additivity. We should also expect that $\Sigma(\mu^*)$ contains $\sigma(\mathcal{I})$.

Suppose we have a semi-ring $\mathcal{I}$ of subsets of X and we are given a pre-measure,

$$\mu : \mathcal{I} \to [0, \infty],$$

we can approximate the size of subsets of X using the subsets in $\mathcal{I}$. Let $\{A_i\}_{i\in C} \subset \mathcal{I}$ be a cover of a subset $A \subset X$ i.e., $A \subseteq \bigcup_{i\in C} A_i$. Then the size of A should be less than the sum of sizes of all subsets A_i. The quantity,

$$\mu^*(A) = \inf\left\{ \textstyle\sum_{i\in C} \mu(A_i) \;\middle|\; A \subseteq \displaystyle\bigcup_{i\in C} A_i,\;\; A_i \in \mathcal{I}\right\}$$

can be thought of as an approximation of the size of A using the subsets in $\mathcal{I}$. This is called an outer measure generated by μ. We can abstract out the nice properties of this quantity which can be later used to construct general measures.

The set function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is such that

$$\mu^*(\varnothing) = 0 \tag{empty size}$$

if $B \subseteq A$ then any cover of A will also be a cover of B, so we have,

$$\mu^*(B) \le \mu^*(A) \tag{monotonicity}$$

we expect the sum of sizes of individual subsets should be greater than the size of the union. If $\{A_i\}_{i\in C}$ is a sequence of sets in $\mathcal{P}(X)$ then, there exists a cover $\bigcup_{j\in C_i}\{O_{j,i}\}$ of A_i such that $\mu^*(A_i) \ge \sum_j \mu(O_j) - \epsilon/2^i$ for every ϵ. For example, any set from the collection of sets $\arg\inf\{\sum_j \mu(O_j)\}$ works. The union of $\{A_i\}_{i\in C}$ is contained in the union $\cup_{i,j} O_{i,j}$. Hence we have,

$$\mu^*\Big(\bigcup_{i\in C} A_i\Big) \le \textstyle\sum_{i\in C}\sum_{j\in C_i} \mu(O_{i,j}) \le \sum_{i\in C} \mu^*(A_i) + \epsilon.$$

Since the choice of ϵ was arbitrary, we have,

$$\mu^*\Big(\bigcup_{i\in C} A_i\Big) \le \textstyle\sum_{i\in C} \mu^*(A_i). \tag{σ-subadditivity}$$

where C is countable. A set map $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is called an outer measure if it satisfies the conditions (empty size), (monotonicity), and (σ-subadditivity).

1.3.1 | Caratheodory Extension Theorem

We have to now recognize the sets for which the outer measure measures the right size. Our goal now is to choose a subset of $\mathcal{P}(X)$ for which the σ-additivity holds.

What we expect from such sets is that the outer measure is a good estimate of the size. The inner measure would be the outer estimate of the complement of the set. Suppose $A \subset X$ and $Q \subset X$, we expect both estimates to be equal, i.e., $\mu^*(Q \cap A) = \mu^*(Q) - \mu^*(Q \cap A^c)$.

$$\mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \cap A^c) \ \forall Q \subseteq X. \qquad \text{(additive intersection)}$$

In such a case, A is said to additively intersect all sets with respect to μ^*. We use Q here because $\mu^*(X)$ can be infinite.
A set $A \subseteq X$ is μ^*-measurable if it satisfies (additive intersection). Let $\Sigma(\mu^*)$ denote the set of all μ^*-measurable sets. Our aim is to show that $\Sigma(\mu^*)$ is a σ-algebra and is the largest σ-algebra on which $\mu^*|_{\Sigma(\mu^*)}$ is a measure.

Theorem 1.3.1. (Carathéodory Theorem) *$\Sigma(\mu^*)$ is a σ-algebra, $\mu^*|_{\Sigma(\mu^*)}$ is a measure on $\Sigma(\mu^*)$.*

Proof

Clearly, $\varnothing \in \Sigma(\mu^*)$. For $A \in \Sigma(\mu^*)$, the complement $A^c \in \Sigma(\mu^*)$ due to the symmetry of additive intersection. We have to now show $\Sigma(\mu^*)$ also satisfies finite union and countable union. Let $A,\, B \in \Sigma(\mu^*)$. For all $Q \subseteq X$,

$$\begin{aligned}\mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \cap A^c) &= \mu^*(Q \cap A) + \mu^*(Q \cap A^c \cap B) + \mu^*(Q \cap A^c \cap B^c)\\ &\geq \mu^*((Q \cap A) \cup (Q \cap A^c \cap B)) + \mu^*(Q \cap (A \cup B)^c).\end{aligned}$$

So, $\mu^*(Q) \geq \mu^*(Q \cap (A \cup B)) + \mu^*(Q \cap (A \cup B)^c)$. Together with $\mu^*(\bigcup_{i \in C} A_i) \leq \sum_{i \in C} \mu^*(A_i)$ it's an equality and hence $A \cup B \in \Sigma(\mu^*)$. By induction it holds for all finite unions and therefore for countable union. Since μ^* is an outermeasure and is σ-additive on $\Sigma(\mu^*)$ it is a measure on $\Sigma(\mu^*)$. □

We now have to show $\Sigma(\mu^*)$ is large enough, that is, $\sigma(\mathcal{I}) \subseteq \Sigma(\mu^*)$ when μ^* is an outer measure generated by the premeasure μ on $\mathcal{I}$.

Theorem 1.3.2. (Carathéodory's Extension Theorem) *If $\mu : \mathcal{I} \to [0, \infty]$ be a pre-measure, then, $\sigma(\mathcal{I}) \subseteq \Sigma(\mu^*)$ and $\mu^*|_{\Sigma(\mu^*)}$ is a measure.*

Proof

Let $Q \subseteq X$ with $\mu^*(Q) < \infty$ and let $\{A_i\}_{i \in C} \subset \mathcal{I}$ be a countable cover of Q. $A_i = (A_i \cap A) \cup (A_i \cap A^c)$. $A_i \cap A = B_i \in \mathcal{I}$ by intersection and $A_i \cap A^c = A_i \backslash A = \coprod_{n_i \geq j} C_{i,j}$ by disjoint union for $B_i, C_{i,j} \in \mathcal{I}$.

$$\mu^*(A_i) = \mu^*(B_i) + \textstyle\sum_{n_i \geq j} \mu^*(C_{i,j}).$$

So, the sum is then given by, $\sum_{i \in C} \mu^*(A_i) = \sum_{i \in C} \mu^*(B_i) + \sum_{i \in C} \sum_{n_i \geq j} \mu^*(C_{i,j}) \geq \mu^*(Q \cap A) + \mu^*(Q \cap A^c)$. Since the choice of the cover $\{A_i\}_{i \in C}$ was arbitrary we have,

$$\mu^*(Q) = \inf\{\textstyle\sum_{i \in C} \mu(A_i) \mid A \subseteq \cup_{i \in C} A_i\} \geq \mu^*(Q \cap A) + \mu^*(Q \cap A^c).$$

So, $A \in \Sigma(\mu^*)$. □

Note that the Caratheodory extension is a complete extension. Heuristically, the reason is because we have tried to build a σ-algebra by estimating sizes by the sizes of sets that contain sets. So, any subset of a measure zero set, will be measurable, and will also be measure zero.

1.3.1.1 | Uniqueness Theorem

1.4 | Borel σ-Algebras

If $(X, \mathcal{T})$ is a topological space where $\mathcal{T}$ is the collection of all open sets, the σ-algebra generated by the collection $\mathcal{T}$ is called the Borel σ-algebra, denoted by $\mathcal{B}_X$. Elements of $\mathcal{B}_X$ are called Borel sets. Since σ-algebras include countable unions, complements, and empty sets we have the following important property of Borel σ-algebras.

Lemma 1.4.1. *$\mathcal{B}_X$ includes countable unions and intersections of open and closed sets.*

An immediate corollary is that if $f : Y \to X$ is a continuous map, then, $f^{-1}(\mathcal{B}(\Gamma)) = \mathcal{B}(f^{-1}(\Gamma))$, because for continuous functions, preimage of open sets are open by definition. The statement follows for the generated σ-algebras.

A regular Borel measure is a measure in which every measurable set can be approximated from above by open measurable sets, and from below by compact measurable sets.

1.4.1 | Lebesgue-Stieltjes Measures

If $\mathcal{B}_{\mathbb{R}}$ is the Borel σ-algebra for $\mathbb{R}$, the topology on $\mathbb{R}$ is generated by all open intervals. Since $\mathcal{B}_{\mathbb{R}}$ must include all closed intervals as well. It follows that it can be generated by open intervals, closed intervals, half closed intervals, and open and closed rays. Using the fact that it can also generated be generated by half open intervals, we have,

$$\sigma(\mathcal{I}_{\mathbb{R}}) = \mathcal{B}_{\mathbb{R}}.$$

Let $\alpha : \mathbb{R} \to \mathbb{R}$ be any increasing right continuous function, that is, $\alpha(a) \leq \alpha(b)$ whenever $a \leq b$ and $\lim_{x \to a^+} \alpha(x) = \alpha(a)$ for all a. Define

$$\lambda_\alpha((a, b]) = \alpha(b) - \alpha(a).$$

Such an α is called a distribution function.

Lemma 1.4.2. *λ_α is a pre-measure on $\mathcal{I}_{\mathbb{R}}$.*

Proof

We have to verify that λ_α is (σ-additivity). Let $\{(a_i, b_i]\}_{i \geq 0}$ be a collection of disjoint sets in $\mathcal{I}_{\mathbb{R}}$ with $\coprod_{i \geq 0}(a_i, b_i] = (a, b]$. Any finite subcollection can be ordered such that $a \leq a_1 < b_1 \leq \cdots \leq a_N < b_N \leq b$. This gives us, $\sum_{N \geq i \geq 0} \lambda_\alpha((a_i, b_i]) \leq \alpha(b) - \alpha(a)$, since $\sum_{i \in C} \lambda_\alpha((a_i, b_i]) = \lim\limits_{N \to \infty} \sum_{i=1}^{N} \lambda_\alpha((a_i, b_i])$, we have,

$$\textstyle\sum_{i \in C} \lambda_\alpha((a_i, b_i]) \leq \alpha(b) - \alpha(a)$$

It can be showed by induction that for any finite cover $\bigcup_{i \in C}(a_i, b_i] \supseteq (a, b]$, $\alpha(b) - \alpha(a) \leq \sum_{i \in C} \lambda_\alpha(b_i - a_i)$. For the infinite cover case, one reduces it to the finite cover case, using compactness of the closed interval and then applying Heine-Borel theorem. Now since $\{(a_i, b_i]\}_{i \in C}$ is a cover we have,

$$\alpha(b) - \alpha(a) \leq \textstyle\sum_{i \in C} \lambda_\alpha((a_i, b_i]).$$

By right continuity of α we have,

$$\lambda_\alpha\Big(\coprod_{i\in C}(a_i,b_i]\Big) = \textstyle\sum_{i\in C}\lambda_\alpha((a_i,b_i]).$$

where C is countable. Hence λ_α is σ-additive. □

It can be verified that λ_α defines a pre-measure on $\mathcal{I}_\mathbb{R}$. α does not have to be strictly increasing and this allows us to give zero measure to certain intervals. Using this pre-measure we define an outer measure,

$$\lambda_\alpha^*(A) = \inf\Big\{\textstyle\sum_{i\in C}\lambda_\alpha(A_i) \mid A\subseteq \bigcup\limits_{i\in C} A_i,\ \ A_i\in\mathcal{I}\Big\}$$

Since $\mathcal{I}_\mathbb{R}$ generates the σ-algebra $\mathcal{B}_\mathbb{R}$, by Caratheodory extension theorem, we have $\mathcal{B}_\mathbb{R} = \sigma(\mathcal{I}_R R)\subseteq\Sigma(\lambda_\alpha^*)$ and $\lambda_\alpha^*|_{\Sigma(\lambda_\alpha^*)}$ is a measure. So we have proved the following,

Theorem 1.4.3. *Every set in $\mathcal{B}_\mathbb{R}$ is λ_α^*-measurable.*

Given a regular Borel measure μ on $\mathbb{R}$ with $\mu(K)<\infty$ for every compact set $K\subset\mathbb{R}$, define $\alpha(a)=\mu((0,x])$ for $a\geq 0$ and $\alpha(a) = -\mu((a,0])$ for $a<0$. Then α is an increasing right continuous functions. So, every positive measure on Borel σ-algebra of $\mathbb{R}$ can be thought of as a Lebesgue-Stieltjes measure.

1.4.1.1 | The Lebesgue Measure

Now that we have developed some of the basic abstract nonsense to measure sets, we apply this to the interval length measure, the Lebesgue measure. The Lebesgue measure on $\mathcal{B}_\mathbb{R}$ is the measure when α is the identity function,

$$\lambda : (a,b]\mapsto b-a.$$

For the case of $\mathbb{R}^n$, $\mathcal{B}_{\mathbb{R}^d}$ is generated by cubes of the form $\prod_{j=1}^d(a_j,b_j]$. The collection of all cubes forms a semi-ring, denoted by $\mathcal{I}_\mathbb{R}^d$ and the pre-measure is the volume function, $\lambda^d:\mathcal{I}_\mathbb{R}^d\to[0,\infty]$ given by,

$$\lambda^d\Big(\textstyle\prod_{j=1}^d(a_j,b_j]\Big) := \textstyle\prod_{j=1}^d(b_j-a_j) \tag{L}$$

We have two σ-algebras of interest to us, $\sigma(\mathcal{I}_\mathbb{R}) = \mathcal{B}_{\mathbb{R}^d}$ and $\Sigma(\lambda^{d*}) = \mathcal{L}_{\mathbb{R}^d}$. $\Sigma(\lambda^{d*})$ is called the Lebesgue σ-algebra on $\mathbb{R}^d$. By Carathéodory extension theorem we have $\mathcal{B}_{\mathbb{R}^d}\subseteq\mathcal{L}_{\mathbb{R}^d}$. $(\mathbb{R}^d,\mathcal{L}_{\mathbb{R}^d},\lambda^d)$ is the completion of $(\mathbb{R}^d,\mathcal{B}_{\mathbb{R}^d},\lambda^d)$.

$\mathbb{R}^d$ comes equipped with a self-action, given by translation. The translation keeps the Lebesgue pre-measure invariant, and hence Lebesgue measure is invariant under translations. The converse is also true.

Theorem 1.4.4. (Translation Invariance) *λ^d is the unique translation invariant measure on $\mathcal{L}_{\mathbb{R}^d}$, upto scaling.*

Proof

We will prove that if μ is a translation invariant measure with $\mu((0,1]^d) = 1$ then $\mu = \lambda^d$. Note that it is sufficient to show that they coincide for $\mathcal{I}_{\mathbb{R}^d}$, the uniqueness of extension guarantees that the two are the same on $\mathcal{L}_{\mathbb{R}^d}$.

It's enough to prove for cubes with rational edge lengths. Note that any cube can be obtained by tranlating cubes of the form $\prod_{j=1}^d (0, m_j/n_j]$ for $m_j, n_j \in \mathbb{N}$. By translation invariance the volume of this cube is,

$$\mu(\prod_{j=1}^{d}\Big(0, m_j/n_j]\Big) = \big(\prod_{j=1}^d m_j\big) \cdot \mu\Bigg(\prod_{j=1}^{d}(0, \tfrac{1}{n_j}]\Bigg).$$

Hence it suffices to show that

$$\mu\Big(\prod_{j=1}^{d}(0, \tfrac{1}{n_j}]\Big) = \lambda^d\Big(\prod_{j=1}^{d}(0, \tfrac{1}{n_j}]\Big) = \prod_{j=1}^d \tfrac{1}{n_j}$$

This is the case because due to translation invariance, and $\mu((0,1]^d) = 1$ and

$$1 = \mu\big((0,1]^d\big) = \big(\prod_{j=1}^d n_j\big) \cdot \mu\Big(\prod_{j=1}^{d}(0, \tfrac{1}{n_j}]\Big).$$

□

2 | Lebesgue Integration

From here on, λ will denote the Lebesgue measure on $\mathbb{R}$.

2.1 | Measurable Functions

Suppose we have a map $f : X \to Y$, we would like to understand when we can use measure on X to measure subsets of Y. We can assign a set $A \subset Y$ the measure of the set $f^{-1}(A)$ if $f^{-1}(A)$ to be measurable. If $(X, \Sigma(X))$ and $(Y, \Sigma(Y))$ are two measurable spaces, we say a map $f : X \to Y$ is measurable if

$$f^{-1}(\Sigma(Y)) \subseteq \Sigma(X) \tag{measurable}$$

If $(X, \Sigma(X), \mu)$ be a measure space we can push forward the measure, and in such a case,

$$f_*\mu = \mu \circ f^{-1} \tag{pushforward}$$

is a measure on $\Sigma(Y)$, called pushforward measure of μ under f.

Because,

$$f^{-1}(\sigma(\mathcal{I})) = \sigma(f^{-1}(\mathcal{I}))$$

it suffices to show that for a generator $\mathcal{I}$ of $\Sigma(Y)$, $f^{-1}(E) \in \Sigma(X)$ for all $E \in \mathcal{I}$. If the σ-algebras are Borel σ-algebras then any continuous function $f : X \to Y$ is measurable because $f^{-1}(U)$ is open in X by definition of continuity. Composition $f \circ g$ of Borel measurable functions are Borel measurable. When it comes to integration, we are usually interested in functions with values in real or complex numbers. The σ-algebra we are interested in is $\mathcal{L}_{\mathbb{R}^d}$. In this case however Borel measurability does not imply Lebesgue measurabililty. Because $f^{-1}(U)$ can be a set in $\mathcal{L}_{\mathbb{R}^d}$, and there is no guarantee that $g^{-1}(f^{-1}(U))$ lies in $\mathcal{L}_{\mathbb{R}^d}$.

Lemma 2.1.1. *Let $(X, \Sigma(X))$ be a measurable space, $f : X \to \mathbb{R}$ is Borel measurable if*

$$f^{-1}((a, \infty)), f^{-1}([a, \infty)), f^{-1}((-\infty, a)), f^{-1}((-\infty, a]) \in \Sigma(X).$$

Proof

Since the σ-algebra $\mathcal{B}_{\mathbb{R}}$ can be generated by half-open and half-closed rays, it follows that, if f is measurable then, the preimage of these are in $\Sigma(X)$. $\square$

We will use this as the criterion for measurability of real valued functions. We will denote the real measurable functions on X by $M^1(X)$. We will denote non-negative measurable functions by $M^1_+(X)$. A complex valued function is measurable if both of its real and imaginary parts are measurable.

Using the measurability of compositions of Borel measurable functions, we can consider for any two measurable functions $f, g : X \to \mathbb{R}$, we can consider the composite functions, $F(x) = (f(x), g(x))$, $\phi(z, w) = z + w$, and $\psi(z, w) = z \cdot w$. Since $\mathcal{B}_{\mathbb{R}^2} = \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$, it follows that the compositions of these functions are also measurable. The compositions are the maps $f + g$ and $f \cdot g$ respectively. Hence we have proved the following,

Theorem 2.1.2.

$$f, g \in M^1(X) \Rightarrow f + g, f \cdot g \in M^1(X).$$

Let $\{f_i\}_{i \in C}$ be a sequence of real valued measurable functions. For every $f_i^{-1}((a, \infty])$ is measurable for each f_i. Similarly for $f_i^{-1}([-\infty, a))$. It follows that countable unions and intersections of these measurable sets are also measurable.

We can now consider the supremum of the collection, $\sup_{i \in C} f_i(x)$,

$$\big(\sup_{i \in C} f_i\big)^{-1}((a, -\infty]) = \bigcup_{i \in C} f_i^{-1}((a, \infty])$$

Hence it is a measurable function. Similarly,

$$\big(\inf_{i \in C} f_i\big)^{-1}([-\infty, a)) = \bigcap_{i \in C} f_i^{-1}([-\infty, a)).$$

Similarly, define $h_k(x) = \sup_{i > k} f_i(x)$, h_k is measurable for each k. $\limsup_{i \in C} f_i(x) = \big(\inf_k h_k\big)(x)$. Hence we have proved the following,

Theorem 2.1.3.

$$\{f_i\}_{i \in C} \subset M^1(X) \Rightarrow \sup_i f_i, \inf_i f_i, \limsup_i f_i, \liminf_i f_i \in M^1(X).$$

Since the inverse image under $\lim_i f_i$ of generators can be written as limsups or liminfs, if $f(x) = \lim_i f_i(x)$ exists for all x, f is measurable. We say two measurable functions f and g are measurable if the set $\{x \mid f(x) \neq g(x)\}$ is measure zero. Similarly $f_i \to f$ almost everywhere if $\{x \mid \lim_i f_i(x) \neq f(x)\}$ is measure zero.

2.1.1 | Approximation by Simple Functions

Let $(X, \Sigma(X))$ be a measurable space. If $E \in \Sigma(X)$, the characteristic function of E is defined by,

$$\chi_E(x) = \begin{cases} 1, & x \in E \\ 0, & x \notin E \end{cases}.$$

A simple function f is a function of the form,

$$f(x) = \sum_{i \in F} c_i \chi_{E_i}(x)$$

where $c_i \in \mathbb{R}$ and $E_i \in \Sigma(X)$ where F is a finite set.

By dividing up the domain of a measurable function into a finite union, we can approximate non-negative measurable functions by non-negative simple functions. Let $f : X \to \mathbb{R}$

be a non-negative measurable function. Divide up $\mathbb{R}$ into a finite union of disjoint subsets, consisting of intervals

$$\epsilon_{i,n} = \left[\tfrac{i-1}{2^n}, \tfrac{i}{2^n}\right)$$

To keep this collection finite, let $f_n = [n, \infty)$. $\{\epsilon_{i,n}\} \cup f_n$ is a finite collection, such that $\coprod_i \epsilon_{i,n} \coprod f_n = \mathbb{R}^+$. The non-negative measurable function can be approximated by simple functions,

$$s_n = \sum_{i=1}^{n2^n} \tfrac{i-1}{2^n} \chi_{f^{-1}(\epsilon_{i,n})} + n\chi_{f^{-1}(f_n)}.$$

For all x such that $f(x) > n$, $s_n(x) = n$, for all x with $(i-1)/2^n \geq f(x) \geq i/2^n$, $s_n(x)$ takes the value $(i-1)/2^n$.

$$s_n(x) \to f(x)$$

for all x as $n \to \infty$.

2.2 | The Lebesgue Integral

The Lebesgue integral is defined for simple functions, and extended to general measurable function by taking a limit. Let $(X, \Sigma(X), \mu)$ be a measure space. Let $s = \sum_{n \geq i \geq 0} c_i \chi_{E_i}$ be a simple function. Then the integral of s is defined to be,

$$\int s \equiv \int s \, d\mu := \sum_{n \geq i \geq 0} c_i \mu(E_i).$$

whenever $c_i = 0$ and $\mu(E_i) = \infty$, we use the convention $0 \cdot \infty = 0$. Note that this is welldefined. Because if $s = \sum_{n \geq i \geq 0} c_i \chi_{E_i} = \sum_{m \geq i \geq 0} b_j \chi_{E_j}$

If f is a non-negative measurable function, then the integral of f is defined to be,

$$\int f d\mu = \sup \left\{ \int s \mid 0 \leq s \leq f, \ s \, \text{simple} \right\}.$$

If $f : X \to \mathbb{R}$ is any measurable function, let $f^+ = \sup(f, 0)$ and $f^- = \sup(-f, 0)$. Provided $\int f^+ d\mu$ and $\int f^+ d\mu$ are not both infinite, the integral of f is defined by,

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

We say a measurable function f is integrable if

$$\int |f| d\lambda < \infty. \tag{integrable}$$

The collection of all integrable functions on X is denoted by $L^1(X)$, the collection of non-negative integrable functions is denoted by $L^1_+(X)$.

2.3 | Convergence of the Lebesgue Integral

We have defined the integral as a limit, of the integral for simple functions. To study the properties of the integral, it is important to be able to understand how they behave under taking limits, so we can extend properties of integral from the properties of the integral for simple functions.

We say a sequence of functions $\{f_i\}_{i\in C} : X \to \mathbb{R}$ converges to $f : X \to \mathbb{R}$ if,

$$\left|f_i(x) - f(x)\right| \to 0$$

as $i \to \infty$ for all $x \in X$. We are interested in the behavior of the integral for under taking limits. Note that many of the properties of this notion of convergence of integral requires the functions to be integrable.

2.3.1 | Monotone Convergence Theorem

The monotone convergence theorem relates limits of integral of non-negative increasing sequences of functions to the integral of the limit of the functions. This allows us to extend properties of simple functions to non-negative functions, which can then be extended to real valued functions.

Theorem 2.3.1. (Monotone Convergence Theorem) *If $\{f_i\}$ is a sequence of non-negative measurable functions such that $f_i(x) \le f_{i+1}(x)$ for all i and x.*

$$\lim_{i\to\infty} f_i(x) = f(x) \forall x \Rightarrow \lim_{i\to\infty} \int f_i d\lambda = \int f d\lambda.$$

Proof

Since $\{f_i\}$ are non-negative functions, $\int f_i$ is an increasing sequence of numbers. Let the limit be L. Since limits of measurable functions is also measurable, $\lim_{i\to\infty} f_i$ is a measurable function. We have to show $L = \int f$. We will show by sandwiching $L \le \int f \le L$.

Let $s = \sum_{m\ge j} s_j \chi_{S_j}$ be a non-negative simple function. Let A_i be the set of all x such that $0 \le cs(x) \le f_i(x) \le f(x)$ for each $c \in (0, 1)$. Then $\{A_i\}_{i\in I}$ is an increasing collection and $A_i \to X$.

$$\int f_i d\lambda \ge \int_{A_i} f_i d\lambda \ge c \int s d\lambda = c \int_{A_i} \textstyle\sum_{m\ge j} s_j \chi_{S_j} = c \sum_{m\ge j} s_i \lambda(S_j \cap A_i).$$

As $i \to \infty$ the right hand side converges to,

$$c \textstyle\sum_{m\ge j} s_i \lambda(S_j) = c \displaystyle\int s d\lambda.$$

So, $L \ge c \int s$ for all $c \in (0, 1)$. Since c was arbitrarily chosen from $(0, 1)$, it also holds for 1, and hence we have, $L \ge \int s$. Since the integral $\int f$ is defined to be the supremum of integral for all simple functions s with $s \le f$, it follows that, $L \ge \int f$. So, we have, $L \ge \int f$. $\square$

2.3.1.1 | Linearity of the Integral

Linearity of integral follows because we defined integral as a limit of integral of simple functions. Let f and g be two simple functions, given by, $f = \sum_{i\in I} f_i \chi_{F_i}$, and $g = \sum_{j\in J} g_j \chi_{G_j}$. Without loss of generality, assume that $\{F_i\}_{i\in I}$ and $\{G_j\}_{j\in J}$. Their sum is then given by,

$$f + g = \textstyle\sum_{i\in I} \sum_{j\in J} (f_i + g_j) \chi_{F_i \cap G_j},$$

and by definition of the Lebesgue integral,

$$\begin{aligned}\int (f+g)d\lambda &= \sum_{i\in I}\sum_{j\in J}(f_i+g_j)\lambda F_i\cap G_j \\ &= \sum_{i\in I}\sum_{j\in J} f_i\lambda(F_i\cap G_j) + \sum_{i\in J}\sum_{j\in J} g_j\lambda(F_i\cap G_j) \\ &= \sum_{i\in I} f_i\lambda F_i + \sum_{j\in J} g_j\lambda G_j = \int f\, d\lambda + \int g\, d\lambda.\end{aligned}$$

Using monotone convergence theorem this extends to measurable functions.

Theorem 2.3.2.

$$f,g\in L^1(X)\Rightarrow \int_A (f+g)d\lambda = \int_A f\, d\lambda + \int_A g\, d\lambda.$$

Proof

For measurable non-negative functions, f and g, take s_n and t_n to be non-negative simple functions increasing to f and g respectively. Then by the above it follows that

$$\begin{aligned}\int (f+g)d\lambda &= \lim_{n\to\infty}\int (s_n+t_n)d\lambda \\ &= \lim_{n\to\infty}\int s_n d\lambda + \lim_{n\to\infty}\int t_n d\lambda && \text{(by MCT)} \\ &= \int \lim_{n\to\infty} s_n d\lambda + \int \lim_{n\to\infty} t_n d\lambda = \int f d\lambda + \int g d\lambda.\end{aligned}$$

Similarly, for real valued measurable functions f and g, we have,

$$\int |f+g|d\lambda \le \int \big(|f|+|g|\big)d\lambda = \int |f|d\lambda + \int |f|d\lambda < \infty$$

so, $f+g$ is integrable, and we can write $f+g=(f+g)^+-(f+g)^- = f^+-f^-+g^+-g^-$. Then applying the linearity for non-negative functions we get linearity for real measurable functions. If f and g are complex measurable, then we can apply the linearity to the real and complex parts to prove linearity. So, we have proved the following theorem. We get the $\int_A$ by multiplying the functions f and g by χ_A. □

Theorem 2.3.3. $\{f_i\}_{i\in I}$ *be non-negative measurable functions,*

$$\int \big(\textstyle\sum_{\infty\ge i\ge 0} f_i\big)d\lambda = \displaystyle\sum_{\infty\ge i\ge 0}\int f_i d\lambda.$$

Proof

Let $F_N = \sum_{N\ge i} f_i$. Since each f_i is non-negative, this is a sequence of non-negative measurable functions such that $F_j(x)\le F_{j+1}(x)$ for all x and $F_N\to\sum_{\infty\ge i} f_i$. Hence (2.3.1) is applicable.

$$\begin{aligned}\int\sum_{\infty\ge i} f_i &= \int \lim_{N\to\infty}\sum_{N\ge i} f_i = \int \lim_{N\to\infty} F_N \\ &= \lim_{N\to\infty}\int F_N = \lim_{N\to\infty}\sum_{\infty\ge i}\int f_i = \sum_{\infty\ge i}\int f_i.\end{aligned}$$

We used the linearity of the Lebesgue integral in the third step. □

$\int$ can be thought of as a linear functional on the vector space of measurable functions.

2.3.1.2 | Fatou's Lemma

The monotone convergence theorem requires the sequence to be increasing. The Fatou lemma allows us to talk about limits for the more general case.

Theorem 2.3.4. (Fatou's Lemma) *Let $\{f_i\}_{i\in C}$ be non-negative measurable functions. Then*

$$\int \liminf_{n\to\infty} f_i \leq \liminf_{n\to\infty} \int f_i.$$

Proof

The idea is to reduce the problem so monotone convergence theorem is applicable. Let $g_j = \inf_{i\geq j} f_i$. Since $g_j(x) \leq g_{j+1}(x)$ for all x, and $g_j \to \liminf_i f_i$.

We can now apply monotone convergence theorem to $\{g_j\}$.

$$\int \liminf_{i\to\infty} f_i = \int \lim_{j\to\infty} g_j = \lim_{j\to\infty} \int g_j$$

Now,

$$\int g_j \leq \inf_{i\geq j} \int f_i.$$

So, by taking the limit $j \to \infty$ we have $\lim_j \int g_j \leq \liminf_i \int f_i$,

$$\int \liminf_{i\to\infty} f_i = \lim_{j\to\infty} \int g_j \leq \liminf_{i\to\infty} \int f_i.$$

□

Suppse $\{f_i\}$ is a sequence of real measureable functions with $f_i \to f$, and $\sup_i \int |f_i| \leq K < \infty$. Then, $\{|f_i|\}$ is a sequence of non-negative measurable functions and $|f_i| \to |f|$.

$$\int \liminf_{n\to\infty} |f_i| \leq \liminf_{n\to\infty} \int |f_i|.$$

2.3.2 | Dominated Convergence Theorem

The other prominent convergence theorem of the Lebesgue integration theory is the dominated convergence theorem.

Theorem 2.3.5. (Dominated Convergence Theorem) *Let $\{f_i\}_{i\in C} \subset L^1$ with $f_i(x) \to f(x)$ for all x.*

$$\exists g \in L^1, |f_i(x)| \leq |g(x)| \Rightarrow \lim_{i\to\infty} \int f_i \to \int f.$$

Proof

We have to use g so we can apply monotone convergence theorem or Fatou's lemma. Consider $f_i + g \geq 0$. Then $\{f_i + g\}_{i \in C}$ is a sequence of non-negative measurable functions. So, Fatou's lemma is applicable. By linearity of the integral we have,

$$\int f + \int g = \int (f+g) \leq \liminf_{i \to \infty} \int (f_i + g) = \liminf_{i \to \infty} \int f_i + \int g.$$

$$\Rightarrow \int f \leq \liminf_{i \to \infty} \int f_i.$$

Similarly we can apply it to $\{f_i - g\}_{i \in C}$ and we have,

$$\int g - \int f = \int (g - f) \leq \liminf_{i \to \infty} \int (g - f_i) = \int g - \liminf_{i \to \infty} \int f_i.$$

$$\Rightarrow \int f \geq \limsup_{i \to \infty} \int f_i.$$

Combining these with the fact that $\liminf \leq \limsup$ the theorem is proved. $\square$

2.4 | Product Measures

2.4.1 | Product Sigma Algebras

We can construct measure on the product spaces using measures on the individual spaces. If $\{(X_i, \Sigma(X_i))\}_{i \in \mathcal{J}}$ is a collection of measurable spaces, then

2.4.2 | The Fubini-Tonelli Theorem

3 | Signed Measures & Differentiation

3.1 | Decomposition Theorems

3.1.1 | Hahn Decomposition

3.1.2 | Jordan Decomposition

3.2 | The Radon-Nikodym Derivative

3.2.1 | Absolute Convergence

3.2.2 | Radon-Nikodym Theorem

3.2.3 | Lebesgue Decomposition Theorem

3.3 | Differentiation

3.4 | Radon Measures

Summary

[1] G B Folland, Real Analysis; Modern Techniques and Their Applications John Wiley & Sons, Inc.