Introduction to Measure Theory

Jan Ertl

June 6, 2018

"Mathematics is the art of giving the same name to different things."

— Henri Poincaré

Acknowledgments

Thanks to Evan Chen, whose LaTeX-style I am using. My approach to the subject is inspired by Professors Wenjia Jing and Robert Fefferman, from whom I learned measure theory.

§1 Measure Spaces

Try to recall how you had to compute areas for the first time in elementary school - you likely had an exercise of an irregular shape that you could decompose into subsets whose area you already knew. You might have also had such computation tasks for which it was easiest to calculate the area of a shape that contained the object of interest, and subtract those smaller chunks from this larger area that did not belong to the shape of interest.

In your calculation you implicitly assumed that if you can measure a (finite) set of shapes, you can measure their disjoint union, with the assigned area being the sum of the areas of well-known shapes. Likewise, you assumed that if you can measure an encompassing area, and you can measure certain subsets of it, you can also measure their complement relative to the space. These observations about elementary school area computations shall serve as a first motivation for the following definitions:

Definition 1.1 (algebra, σ-algebra)**.** Let X be a nonempty set. A family $\mathscr{A}$ of subsets of X is said to be an **algebra** if

- $\emptyset \in \mathscr{A}$
- $\mathscr{A}$ is closed under complements and finite unions, i.e. if $A \in \mathscr{A}$, then $A^c = X - A \in \mathscr{A}$, and if $A_1, A_2, \cdots, A_N \in \mathscr{A}$, then $\bigcup_{n=1}^{N} A_n \in \mathscr{A}$

We say $\mathscr{A}$ is a **σ-algebra** if it is also closed under countable unions, i.e.

- if $A_1, A_2, \cdots \in \mathscr{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$

The additional condition on σ-algebras might be motivated by the observation that we can write many common sets, such as a circle, as the (almost) disjoint union of cubes inscribed in it.

Remark 1.2. In these notes, I will use $\mathcal{A}$ for algebras, and $\mathcal{S}$ for σ-algebras. Many other texts use $\mathscr{F}$ to denote both because algebras used to be called fields.

Exercise 1.3. Let $X = \mathbb{R}$. Define

- $\mathcal{G}_1 = \{A : A \subset \mathbb{R}, A \text{ open}\}$
- $\mathcal{G}_2 = \{A : A$ is a finite union of intervals of the form (a,b], (-∞, b], and (a,∞) $\}$
- $\mathcal{G}_3 = \{A : A$ is a countable union of (a,b], (-∞, b], and (a,∞) $\}$

Which of these are algebras? Which are even σ-algebras?

Remark 1.4. If $\mathcal{A}$ is a σ-algebra of X, then it is closed under countable intersection as well.

Proof. Suppose $A_n \in \mathcal{A}, \forall n \in \mathbb{N}$. Note that $(\bigcap_{n=1}^{\infty} A_n)^c = \bigcup_{n=1}^{\infty} A_n^c \in \mathscr{A}$, so $\bigcap_{n=1}^{\infty} A_n$ itself is in $\mathcal{A}$. The equality follows from basic set theory (see Demorgan's Laws). $\square$

Example 1.5

Suppose $X = \{$possible outcome of a coin flip $\} = \{H, T\}$. Recall that in probability theory, an event is a set of outcomes, i.e. a subset of the sample space. Then note that the event space $\{\emptyset, \{H\}, \{T\}, \{H, T\}\}$, which happens to be the power set of the sample space, is a σ-algebra.

Example 1.6

The **power set of X** is defined as $\mathcal{P}(X) = \{A : A \subset X\}$. It is easy to show that the power set is a σ-algebra. The set $\{\emptyset, X\}$ is also a σ-algebra.

Exercise 1.7. Explain why these are the 'largest' and 'smallest' possible σ-algebras, respectively, in the following sense: if $\mathcal{A}$ is any σ-algebra, then $\{\emptyset, X\} \subset \mathcal{A} \subset \mathcal{P}(X)$.

An important σ-algebra we will return to is the Borel-σ-algebra on X, which is well-defined if X is a metric space. It is the smallest σ-algebra that contains all open sets of X. Let us make this precise.

Definition 1.8 (σ-algebra generated by a family of subsets of X)**.** Given any family $\mathcal{G}$ of subsets of X, we define

$$\sigma(\mathcal{G}) := \bigcap_{\alpha} \mathcal{S}_\alpha$$

where $\{\mathcal{S}_\alpha\}$ is the collection of σ-algebras that include $\mathcal{G}$. Note that this is well-defined because $\mathcal{P}(X) \in \{\mathcal{S}_\alpha\}$. Furthermore, $\sigma(\mathcal{G})$ is a σ-algebra. The proof is left as an easy exercise. We call $\sigma(\mathcal{G})$ the **σ-algebra generated by $\mathcal{G}$**. It is clear that it is the smallest σ-algebra containing $\mathcal{G}$.

Example 1.9 (Binary infinite trees - Sequence space)

Let $X = [0, 1]^{\mathbb{N}}$, the space of infinite binary sequences. Given any finite sequence $a = a_1 a_2 a_3 ... a_m$, define the set E_a to be the set of sequences whose first m members coincide with a. Let $\mathcal{E}$ be the collection of all these sets. We refer to the pair $(X, \sigma(\mathcal{E}))$ as **sequence space**. Note that you can also think of this as a subset of infinite coin flips.

Exercise 1.10. Prove the following Proposition:
Let $\{\mathcal{S}_\alpha\}$ be a family of σ-algebras on X. Then $\bigcap_\alpha \mathcal{S}_\alpha$ is also a σ-algebra.

We now have enough vocabulary to properly define the Borel σ-algebra:

Definition 1.11 (Borel σ-algebra)**.** Let X be a metric space, and let $\mathcal{O}$ denote the collection of open sets of X. $\sigma(\mathcal{O})$ is thus the smallest σ-algebra containing all open sets of X. We call it the **Borel σ-algebra of X** and denote it by $\mathcal{B}(X)$.

Definition 1.12 (Measurable space)**.** A pair $(X, \mathcal{S})$ with X nonempty, $\mathcal{S}$ a σ-algebra on X is called a **measurable space**.

Definition 1.13 (Measure)**.** Let $(X, \mathcal{S})$ be a measurable space. A function $\mu : \mathcal{S} \to [0, \infty]$ is called a (nonnegative) **measure** if it satisfies

(i) $\mu(\emptyset) = 0$

(ii) $\mu(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$ for any $\{A_i\}_{i=1}^\infty \subset \mathcal{S}$ s.t. $A_i \cap A_j = \emptyset \ \forall i \neq j$.

Remark 1.14. We can define signed measures $\mu : \mathcal{S} \to [-\infty, \infty]$. Our discussion here is restricted to non-negative measures, since they are natural in our applications of Lebesgue integration and probability theory.

Definition 1.15 (Measure space)**.** A triple $(X, \mathcal{S}, \mu)$ is called a **measure space** if $(X, \mathcal{S})$ is a measurable space, and μ is a measure defined on it.

Example 1.16

If $\mu(X) = 1$, then we may say μ is a probability measure, and the triple $(X, \mathcal{S}, \mu)$ is a probability space.

Exercise 1.17. Let $(X, \mathcal{S}, \mu)$ be a measure space. Prove the following:

- μ is **monotone**: if $A, B \in \mathcal{S}, A \subset B$, then $\mu(A) \leq \mu(B)$
- μ is **countably subadditive**: if $\{A_i\}_{i=1}^\infty \subset \mathcal{A}$, then $\mu(\cup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty \mu(A_i)$

The following easy exercise will come in handy when defining conditional probabilities later on:

Exercise 1.18. Let $(X, \mathcal{S}, \mu))$ be a measure space. Let $B \in \mathcal{S}$. Show that $\lambda : \mathcal{S} \to [0, \infty]$ defined by $\lambda(A) = \mu(A \cap B)$ is also a measure $(X, \mathcal{S})$.

Theorem 1.19

Let μ be a measure on $(X, \mathcal{S})$. Then it is continuous from below in the sense that:

(i) $(A_1 \subset A_2 \subset A_3 \subset \cdots, A_i \in \mathcal{S}) \Rightarrow (\lim_{n\to\infty} \mu(A_n) = \mu(\cup_{i=1}^\infty A_i))$

(ii) $(A_1 \supset A_2 \supset A_3 \supset \cdots, A_i \in \mathcal{S}, \mu(A_1) < \infty) \Rightarrow (\lim_{n\to\infty} \mu(A_n) = \mu(\cap_{i=1}^\infty A_i))$

Proof. We prove (i) and leave (ii) as an exercise. Let $B_1 = A_1$, $B_i = A_i \cap A_{i-1}^c$, $i \geq 2$.

Then note $\begin{cases} A_n = A_{n-1} \cup B_n & \forall n \geq 1 \\ A_{n-1} \cap B_n = \emptyset, B_m \cap B_n = \emptyset & m \neq n \\ \cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n \end{cases}$

This means that $\cup_{n=1}^{\infty} B_n$ is a countable union of pairwise disjoint sets in $\mathcal{S}$. Hence

$$\mu(\cup_{n=1}^{\infty} A_n) = \mu(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) = \lim_{N\to\infty} \mu(A_n)$$

□

Exercise 1.20. Prove (ii).

A consequence of this continuity property leads to the Borel-Cantelli Lemma, an extremely handy tool in probability theory.

Lemma 1.21 (Borel-Cantelli Lemma)

Let μ be a measure on a σ-algebra $\mathcal{S}$. Let the sequence $\{E_i\}_{i=1}^{\infty} \subset \mathcal{S}$, and suppose $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then for $D = \{x \in X : \text{x belongs to infinitely many} E_n\}$ satisfies $\mu(D) = 0$.

Proof. They key observation here is that $D = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} E_n$. Fix $\epsilon > 0$. Since $\sum_{n=1}^{\infty} \mu(E_n) < \infty$, there is an m such that $\sum_{n=m}^{\infty} \mu(E_n) < \epsilon$, and then note

$$\mu(D) \leq \mu(\bigcup_{n\geq m} E_n) \leq \sum_{n=m}^{\infty} \mu(E_n) < \epsilon$$

□

Example 1.22 (Borel-Cantelli application)

Why is this result so useful in probability applications? Think of the E_n as events in a probability space with probability measure μ. Then $\sum_{n=1}^{\infty} \mu(E_n) < \infty$ is a necessary condition for a positive probability of an event occuring infinitely often.

Theorem 1.23 (Existence of non-measurable sets)

Assume the axiom of choice. There is no translation-invariant measure on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$!

Proof. We can return to this if people are interested and we have time at the end of class. □

Theorem 1.24 (Banach-Tarski Paradox)

Informally: One can divide the unit sphere $S \in \mathbb{R}^3$ into finitely many pieces and reassemble these using only rigid motions (translations and rotations) into two identical copies of the original sphere.

§2 Construction of Lebesgue Measure

The preceding results show us that we cannot generally find a measure on a given σ-algebra. In particular, there are sets that defy the intuitive concepts of length or volume. Instead of giving up the concept of measure, we turn back to the σ-algebras. We want to find σ-algebras for which there are well-defined measures that correspond to our intuitive notions of length, area and volume (and probability!). This means we purposefully declare some sets non-measurable, and exclude them from our measures' domain. At the same-time, we do not want to discard too many sets. The following discussion shows that given an algebra for which we have an intuitive way to assign measure, we can indeed find a σ-algebra and measure large enough to include the original algebra and the σ-algebra it generates.

We follow an abstract approach to encompass many different measure spaces. However, the application that we are focused on is the real line $\mathbb{R}$. The measure we are constructing is called the Lebesgue measure, named after the French mathematician Henri Lebesgue.

We start the construction by defining a so called pre-measure on an algebra. You can think of assigning length, area or volume to simple intervals, rectangles and

Definition 2.1 (Pre-measure)**.** $\nu : \mathcal{A} \to [0, \infty]$ is called a **pre-measure** if it satisfies:

(i) $\nu(\emptyset) = 0$

(ii) $\nu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \nu(A_n)$ for $\{A_n\} \subset \mathcal{A}$, pairwise disjoint

In one dimension, finding a

Example 2.2 (Application to Lebesgue measure on $\mathbb{R}$)

For Lebesgue measure, we will take $\mathcal{A} = \{A : A$ is a finite disjoint union of intervals of the form (a,b], (-∞, b], and (a,∞) }. It was already an exercise to show that this is an algebra. It is easy to see that ν defined by $\nu(A) = \nu(\bigcup_{i=1}^{N}(a_i, b_i]) = \sum_{i=1}^{N}(b_i - a_i)$ is a pre-measure.

Remark 2.3. We can already see that this generalizes easily to two (and thus any finite) dimensions - instead of intervals, we may use cubes and boxes. We still need to be careful how to proceed in our construction.

Definition 2.4 (Outer Measure)**.** Given X, we call $\mu^* : \mathcal{P}(X) \to [0, \infty]$ an **outer measure** if

(i) $\mu^*(\emptyset) = 0$,

(ii) it is monotone: if $A \subset B$, then $\mu^*(A) \leq \mu^*(B)$,

(iii) and it is countably subadditive: $\mu^*(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ for any $\{A_i\}_{i=1}^{\infty}$.

Based on our given pre-measure, we will now define a way to assign an outer measure. Note that by definition, the domain of outer measures is always the power set of the original space.

Theorem 2.5 (Generation of outer measure)
Let $\mathcal{A}$ be a collection of subset of X, containing $\emptyset$ and X. Let ν satisfy $\nu(\emptyset) = 0$. Then $\mu^* : \mathcal{P}(X) \to [0, \infty]$ defined as

$$\mu^*(B) := \inf\{ \sum_{n\in\mathbb{N}} \nu(A_n) : \{A_n\} \subset \mathcal{A}, B \subset \bigcup_{n\in\mathbb{N}} A_n\}$$

constitutes an outer measure on X. We call it the **outer measure generated by** ν.

Remark 2.6. Note that in the theorem, $\mathcal{A}$ need not be an algebra, and ν need not be a measure. However, if they are, they satisfy the conditions and the theorem holds, which is exactly what we want.

Example 2.7 (Application to Lebesgue measure on $\mathbb{R}$)
It is easy to show that the outer measure generated on $\mathcal{P}(\mathbb{R})$ is

$$\mu^*(B) := \inf\{ \sum_{n=1}^{\infty}(b_n - a_n) : B \subset \bigcup_{i=1}^{\infty}(a_i, b_i]\}$$

We call this the **Lebesgue outer measure**. As before, the definition generalizes into higher (finite) dimensions by taking boxes instead of intervals.

Proof. Why is μ^* well-defined? - We can always take $A_n = X$ as a cover. Let us check each condition for an outer measure:

(i) Note that the empty set is a finite cover of itself, so $\mu^*(\emptyset) = 0$.

(ii) if $A \subset B$, then note that any cover of B is also a cover of A, so taking the infimum over all possible covers yields $\mu^*(A) \leq \mu^*(B)$.

(iii) Take any $\{B_n\}_{n=1}^{\infty} \subset \mathcal{P}(X)$. Fix $\epsilon > 0$. For each B_n, there is a cover $\{A_k^n\}_{k\in\mathbb{N}} \subset \mathcal{A}$ such that
$$\sum_{k\in\mathbb{N}} \nu(A_k^n) \leq \mu^*(B_n) + \frac{\epsilon}{2^n}$$
The union of all these covers $\{A_k^n\}_{k\in\mathbb{N},n\in\mathbb{N}}$, then covers $\{B_n\}_{n=1}^{\infty}$, hence

$$\mu(\bigcup_{n\in\mathbb{N}} B_n) \leq \sum_{k\in\mathbb{N},n\in\mathbb{N}} \nu(A_k^n) \leq \sum_{n\in\mathbb{N}}(\mu^*(B_n) + \frac{\epsilon}{2^n}) \leq \sum_{n\in\mathbb{N}} \mu^*(B_n) + \epsilon$$

Since ϵ was arbitrary, this shows that μ^* is countably subadditive.

$\square$

Theorem 2.8 (Carathéodory Construction)
Let μ^* be an outer measure on X, and consider the collection $\mathcal{M}$ of subsets $E \subset X$ such that for every $B \subset X$,

(*) $\mu^*(B) \geq \mu^*(B \cap E) + \mu^*(B \cap E^c)$

Then $\mathcal{M}$ is a σ-algebra, and $\bar{\mu} : \mathcal{M} \to [0, \infty]$ defined by $\bar{\mu} = \mu^*|_{\mathcal{M}}$ ($\bar{\mu}(E) = \mu^*(E)$ for $E \in \mathcal{M}$) is a measure on $\mathcal{M}$.

Example 2.9 (Application to Lebesgue measure on $\mathbb{R}$)
This Theorem yields that the triple $(\mathbb{R}, \mathcal{M}, \bar{\mu} := \mu^*|_{\mathcal{M}})$ is a measure space. We call $\bar{\mu}$ the **Lebesgue measure**. We call $\mathcal{M}$ the collection of **Lebesgue measurable** sets.

Exercise 2.10. Explain why (*) in the preceding theorem could be replaced by

$$\mu^*(B) = \mu^*(B \cap E) + \mu^*(B \cap E^c)$$

Proof. We need to check four properties.

1. $\emptyset \in \mathcal{M}$: (*) holds for $E = \emptyset$

2. $\mathcal{M}$ is closed under complements: This is obvious since (*) is symmetric.

3. $\mathcal{M}$ is closed under countable union:
   Let us show first that $\mathcal{M}$ is closed under finite union: let $E, F \in \mathcal{M}$. We want to show that $E \cup F \in \mathcal{M}$. Observe that

   $$B \cap (E \cup F) = (B \cap E \cap F^c) \cup (B \cap E^c \cap F) \cup (B \cap E \cap F)$$

   and because $F \in \mathcal{M}$, we have both

   $$\mu^*(B \cap E \cap F^c) + \mu^*(B \cap E \cap F) \leq \mu^*(B \cap E) \quad \text{and}$$
   $$\mu^*(B \cap E^c \cap F^c) + \mu^*(B \cap E^c \cap F) \leq \mu^*(B \cap E^c)$$

   Now using countable subadditivity, these inequalities, and finally the fact that $E \in \mathcal{M}$, we get

   $$\begin{aligned}\mu^*(B \cap (E \cup F)) + \mu^*(B \cap (E \cup F)^c) \leq &\mu^*(B \cap E \cap F^c) + \mu^*(B \cap E \cap F)\\ &+ \mu^*(B \cap E^c \cap F^c) + \mu^*(B \cap E^c \cap F)\\ \leq &\mu^*(B \cap E) + \mu^*(B \cap E^c) \leq \mu^*(B)\end{aligned}$$

   This shows closedness under finite union. Now take $\{E_n\}_{n \in \mathbb{N}} \subset \mathcal{M}$. Without loss of generality, we take the E_n to be pairwise disjoint. Define $E = \bigcup_{n \in \mathbb{N}} E_n$. By closedness under finite union, we have

$$\begin{aligned}\mu^*(B\cap(\bigcup_{n=1}^{N}E_n)) =&\mu^*(B\cap(\bigcup_{n=1}^{N}E_n)\cap E_N)+\mu^*(B\cap(\bigcup_{n=1}^{N}E_n)\cap E_N^c)\\ =&\mu^*(B\cap E_N)+\mu^*(B\cap(\bigcup_{n=1}^{N-1}E_n))\text{ iterate...}\\ =&\sum_{n=1}^{N}\mu^*(B\cap E_n)\end{aligned}$$

Next, note that we once again have by (*) that for any N,

$$\begin{aligned}\mu^*(B) =&\mu^*(B\cap(\bigcup_{n=1}^{N}E_n))+\mu^*(B\cap(\bigcup_{n=1}^{N}E_n)^c)\\ \geq&\sum_{n=1}^{N}\mu^*(B\cap E_n)+\mu^*(B\cap(\bigcup_{n=1}^{\infty}E_n)^c)\end{aligned}$$

where the inequality derives from the observation that $(\bigcup_{n=1}^{N}E_n)^c)$ is a decreasing chain. Note that the left-hand side of the equation is independent of N, so it also holds for the limit. Thus

$$\begin{aligned}\mu^*(B) \geq&\sum_{n=1}^{\infty}\mu^*(B\cap E_n)+\mu^*(B\cap(\bigcup_{n=1}^{\infty}E_n)^c)\\ \geq&\mu^*(B\cap(\bigcup_{n=1}^{\infty}E_n))+\mu^*(B\cap(\bigcup_{n=1}^{\infty}E_n)^c)\end{aligned}$$

This completes the proof. $\mathcal{M}$ is a closed under countable union, and is therefore a σ-algebra.

4. Finally, we need to show that $\bar{\mu}$ is a measure: Take the first of the two lines above, and replace B by $\bigcup_{n=1}^{\infty}E_n$. This yields $\mu^*(\bigcup_{n=1}^{\infty}E_n)\geq\sum_{n=1}^{\infty}\mu^*(E_n)$. Note that countable subadditivity yields the reverse inequality. Hence equality holds, which means that $\bar{\mu}$ is a measure.

□

Remark 2.11. We have not seen whether the algebra and pre-measure we started with were preserved, in the sense that $\mathcal{A}\subset\mathcal{M}$ and $\bar{\mu}|_{\mathcal{A}}=\nu$. Thankfully, this is the case.

Theorem 2.12 (Carathéodory Extension Theorem - Existence)
Let $(\mathcal{A},\nu)$ be an algebra - premeasure pair on X. Let μ^* denote the outer measure generated by ν, and $\mathcal{M}$ denote the σ-algebra from the Carathéodory construction. Let $\sigma(\mathcal{A})$ be the σ-algebra generated by $\mathcal{A}$ and $\mu:=\mu^*|_{\sigma(\mathcal{A})}$. Then $\sigma(\mathcal{A})\subset\mathcal{M}$, and $\mu|_{\mathcal{A}}=\mu^*|_{\mathcal{A}}=\nu$.

Example 2.13 (Application to Lebesgue measure on $\mathbb{R}$)

Carathéodory proves that not only is $(\mathbb{R}, \mathcal{M}, \bar{\mu} := \mu^*|_{\mathcal{M}})$ a measure space, but that the Borel-algebra $\mathcal{B}(\mathbb{R})$ is a subset of $\mathcal{M}$, and the Lebesgue measure corresponds to the intuition we have on intervals and their countably infinite combinations. We call $\mu := \mu^*|_{\sigma(\mathcal{A})}$ the **Borel measure**.

Exercise 2.14. Why is it true that the Borel-algebra $\mathcal{B}(\mathbb{R})$ is a subset of $\mathcal{M}$? Hint: Carathéodory does most of the work - you only need to show that $\sigma(\mathcal{A}) = \sigma(\mathcal{O})$.

Theorem 2.15 (Carathéodory Extension Theorem - Uniqueness)

In the notation of the previous theorem, if X can be written as the union of countably many sets of finite pre-measure in the original algebra, i.e. $X = \bigcup_{n\in\mathbb{N}} A_n, A_n \in \mathcal{A}, \nu(A_n) < \infty$, then the extension of the pre-measure is unique.

Remark 2.16. The condition on X above is called **σ-finiteness**.

Example 2.17 (Application to Lebesgue measure on $\mathbb{R}$)

The Lebesgue measure is a natural extension, and is unique. This is all we have ever wanted in life.

§3 Lebesgue Integration

Proof. □

Definition 3.1.

Definition 3.2.

Exercise 3.3.

Theorem 3.4

Lemma 3.5

Proof. □

Definition 3.6.

Remark 3.7.