Lecture 1

Why measure theory?

We want to answer questions like the following: what is the "total length" of an arbitrary $E \subseteq \mathbb{R}$? What about the "total area" of an arbitrary $E \subseteq \mathbb{R}^2$?
In other words, can we define a function $\mu : 2^{\mathbb{R}^d} \to [0, +\infty)$ so that $\mu(E)$ is sufficiently "nice?"
What properties would we like a function μ (called a measure) to have? Let's stick to $\mathbb{R}$ for now.

1. For $E = [a, b]$ (or (a, b)), we'd like $\mu(E) = b - a$.

2. For a sequence of disjoint intervals $I_i \subseteq \mathbb{R}$,

$$\mu\left(\bigcup_{i=1}^{n} I_i\right) = \sum_{i=1}^{n} \mu(I_i)$$

What about $\mathbb{Q} \cap [0, 1]$? What about the area under a curve?
What if E is any arbitrary set????

Pre measure theory

In the mid 1800s, Riemann first defined the Riemann integral in terms of upper and lower sums.
Fortunately, it's good enough for most "ordinary" functions a student might encounter when doing calculus for the first time.
Unfortunately, it's not good for taking limits.
For example, given $f_1, f_2, f_3, \cdots : [a, b] \to \mathbb{R}$ such that $\lim_{n\to\infty} f_n(x) =: f(x)$ exists for all x, when can we conclude that

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b \lim_{n\to\infty} f_n(x)]\,dx = \int_a^b f(x)\,dx$$

We learn in undergraduate real analysis that we may only conclude the above if the f_n converge uniformly to f.

Measure theory

Measure theory allows us to define a much more powerful theory of integration, giving us

- More integrable functions
- An integral which behaves better with limits
- An integral ideally suited for probability theory.

Our first goal will be to define a function $\mu : 2^{\mathbb{R}} \to [0, \infty)$ satisfying the following:

1. If $E_1, E_2, \ldots$ is a countable sequence of disjoint sets, then

$$\mu\left(\bigcup E_i\right) = \sum \mu(E_i)$$

If μ satisfies this, we say it is "countably additive."

2. $\mu([a, b]) = b - a$ for all such intervals.

3. μ is translation invariant, i.e. for any $t \in \mathbb{R}$,

$$\mu(E + t) = \mu(E)$$

Where "$E + t$" $:= \{x + t \mid x \in E\}$

Theorem 0.1. *(Vitali) There is no such μ.*

Proof. Suppose that such a μ exists.

Claim. *If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*

Proof. Note $B = A \coprod (B \backslash A)$, so

$$\mu(B) = \mu(A) + \mu(B \backslash A)$$

And because μ is always nonnegative, we may conclude that $\mu(B) \geq \mu(A)$. ■

Now, define an equivalence relation on $\mathbb{R}$ as follows:

$$x \sim y <=> x - y \in \mathbb{Q}$$
$$[x] := \{y \in \mathbb{R} \mid x \sim y\}$$

Claim. *Every equivalence class contains a point in $[0, 1]$.*

Proof. Homework exercise. ■

Now, for each equivalence class, choose an element in $[0, 1]$ belonging to that class. For this step, we are using choice and, it turns out, there is no way not to in this proof.
Call the resulting set A. So $A \subseteq [0, 1]$, and for any x, $[x] \cap A$ is a singleton.

Let $B = \bigcup_{q \in \mathbb{Q} \cap [-1,1]} A + q$
Note that this is a disjoint union: indeed, if $A + q$ intersects nontrivially with $A + q'$ for $q \neq q'$, then there are x, x' in A such that $x = x' + q$, and so $x \sim x'$, which by construction is impossible.

Claim. $[0,1] \subseteq B \subseteq [-1,2]$.

Proof. First, if $x \in [0,1]$, then $x = a + q$ for some $a \in A \subseteq [0,1], q = x - a \in [-1,1]$. Thus, $x \in B$.
Next, if $b \in B$, then $b = a + q$, for $q \in A = [0,1]$ and $q \in [-1,1]$, so $b \in [-1,2]$.
So we must conclude by the lemma that

$$1 = \mu([1,0]) \leq \mu(B) \leq \mu([-1,2]) = 3$$

But by the properties of μ, we also have

$$\mu(B) = \sum_{q \in \mathbb{Q} \cap [0,1]} \mu(A + q) = \sum_{q \in \mathbb{Q} \cap [0,1]} \mu(A)$$

The sum on the right hand side is either 0 or ∞. But we just showed that it is between 1 and 3, a contradiction. Therefore, $\mu(A)$ cannot be defined.

∎

So, if this is impossible, which criterion should we weaken to make it possible?
If we weaken the first to get finite additivity, we run into problems for $d \geq 3$, for example the Banach-Tarski paradox.
If we weaken the other two, then μ is no longer compactible with the usual notion of "length."

Two good choices

- Given a measure on a family of sets, it extends to an outer measure on all sets.
- Similarly, given an outer measure, you can single out "nice sets" on which it is a measure.

What kind of family of subsets should we restrict to?
Let X be a set.

Definition 0.1. $\mathcal{A}$ is an <u>algebra of subsets of X</u> if $\mathcal{A} \neq \varnothing$, and

- $$E_1, \ldots, E_n \in \mathcal{U} \implies \bigcup_{i=1}^{n} E_i \in \mathcal{U}$$

  In other words, it is "closed under finite unions."

- $$E \in \mathcal{U} \implies X \backslash E \in \mathcal{U}$$

  In other words it is "closed under compliments."

- $$\varnothing, X \in \mathcal{A}$$

Lemma 1. *If $\mathcal{A}$ is an algebra of subsets, then $\mathcal{A}$ is closed under finite intersections.*

Proof. Homework 2

Example 0.1. (i) $\mathcal{A} = 2^X$

(ii) $\mathcal{A} = \{\varnothing, X\}$

(iii) $\mathcal{A}$ = all finite OR cofinite subsets of X (cofinite means the complement is finite).

Definition 0.2. A σ-algebra $\mathcal{A}$ is an algebra that is closed under countable unions.

Remark. σ-algebras are closed under countable intersections.

Example 0.2. Above, (i) and (ii) are σ-algebras, but (iii) is not.

Proposition 1. *Given any family $\mathcal{E}$ of subsets of X, there is a smallest σ-algebra $\mu(\mathcal{E})$ containing $\mathcal{E}$, meaning that if $\mathcal{F}$ is a σ-algebra containing $\mathcal{E}$, then $\mu(\mathcal{E}) \subseteq \mathcal{F}$.*

Lecture 2

Definition 0.3. Given a nonempty set X and $\mathcal{M}$ a σ-algebra of subsets of X, we call $(X, \mathcal{M})$ a measurable space.

Recall:

Proposition 2. *Given any family $\mathcal{E}$ of subsets of X, there is a smallest σ-algebra $\mathcal{M}(\mathcal{E})$ containing $\mathcal{E}$, meaning that if $\mathcal{F}$ is a σ-algebra containing $\mathcal{E}$, then $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{F}$.*

Proof. We begin with a claim.

Claim. Given any nonempty collection $\mathcal{C}$ of σ-algebras on X, then

$$\cap\mathcal{C} := \{E \subseteq X \mid E \in \mathcal{A} \forall \mathcal{A} \in \mathcal{C}\}$$

is a σ-algebra.

Proof. Homework 2

■

Let $\mathcal{C} = \{\mathcal{A} \mid \mathcal{A}$ is a σ-algebra on X and $\mathcal{E} \subseteq \mathcal{A}\}$. $\mathcal{C}$ is nonempty, because $2^X \in \mathcal{C}$. By the claim, $\cap\mathcal{C}$ is a σ-algebra. By the definition of $\mathcal{C}$, $\mathcal{E} \subseteq \mathcal{C}$ and for any σ-algebra $\mathcal{A}$ such that $\mathcal{E} \subseteq \mathcal{A}$, $\cap\mathcal{C} \subseteq \mathcal{A}$.
Thus $\mathcal{M}(\mathcal{E}) = \cap\mathcal{C}$ is the smallest σ-algebra containing $\mathcal{E}$.

■

Remark. Intuitively, $\mathcal{M}(\mathcal{E})$ is a σ-algebra containing the sets in $\mathcal{E}$ by "going from the outside in," starting with σ-algebras that are "too big" and taking intersections.
Recall: a topology τ is a collection of subsets of a set X (called open sets), which is closed under arbitrary unions and finite intersections, and $X, \varnothing \in \tau$
Let (X, τ) be a topoological space.

Definition 0.4. The Borel σ-algebra of X, denoted $\mathcal{B}X$, is the σ-algebra generated by the open subsets of X. Its members are known as Borel sets.
What do the Borel sets look like? Let's go from the "inside out."
Let $\mathcal{F}$ =open sets in X, $\mathcal{F}^\sigma$ all countable unions of sets in $\mathcal{F}$, $\mathcal{F}^\delta$ all countable intersections, and $\overline{\mathcal{F}}$ complements of sets in $\mathcal{F}$.
To build Borel sets:

$$\mathcal{F} \to \mathcal{F}^\delta \cup \overline{\mathcal{F}^\delta} \to \cdots \to \mathcal{B}X$$

To learn more, look up the "Borel hierarchy."

Proposition 3. *The Borel σ-algebra on $\mathbb{R}$, which we denote $\mathcal{B}_\mathbb{R}$, is generated by each of the following.*

(i) Open intervals $\mathcal{E}_1 = \{(a, b) \mid a < b, a, b \in \mathbb{R}\}$

(ii) Closed intervals $\mathcal{E}_2 = \{[a, b] \mid a \leq b, a, b, \in \mathbb{R}\}$

(iii) Half-open intervals $\mathcal{E}_3 = \{[a, b) \mid a < b, a, b, \in \mathbb{R}\}$

(iv) Open rays $\mathcal{E}_4 = \{(a, \infty) \mid a \in \mathbb{R}\}$

(v) Closed rays $\mathcal{E}_5 = \{[a, \infty) \mid a \in \mathbb{R}\}$

That is, $\mathcal{M}(\mathcal{E}_i) = \mathcal{B}_\mathbb{R}$ *for any* $i \in \{1, \ldots, 5\}$.

Proof. Homework 2

■

Let

- $\{(X_1, \mathcal{M}_i)\}_{i=1}^\infty$ be a collection of measurable spaces.

$$X := \prod_{i=1}^\infty X_i$$

- π_i be the projection $X \to X_i$

Example 0.3. If $(X_i, \mathcal{M}_i) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, for $i \in \{1, \ldots, n\}$. Then $X = \mathbb{R}^n$.

Definition 0.5. The product σ-algebra

$$\bigotimes_{i \in \mathbb{N}} \mathcal{M}_i := \mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{M}_i\}\right)$$

Our goal is to show that $\mathcal{B}_{\mathbb{R}^n} = \otimes_{i=1}^n \mathcal{B}_{\mathbb{R}}$.

Proposition 4. *Given $\mathcal{E}_i \subseteq 2^{X_i}$ such that $X_i \in \mathcal{E}_i$, let $\mathcal{M}_i = \mathcal{M}(\mathcal{E}_i)$. Then*

$$\bigotimes_{i \in \mathbb{N}} \mathcal{M}_i = \mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{E}_i\}\right)$$

Note: If $\mathcal{E} \subseteq \mathcal{F}$, then $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}(\mathcal{F})$. If $\mathcal{E} \subseteq \mathcal{M}(\mathcal{F})$, then $\mathcal{M}(\mathcal{E}) \subseteq \mathcal{M}\mathcal{F}$.
Recall: Given a function $f : X \to Y$ between arbitrary nonempty sets, then

(i) $f^{-1}(\cup_{i \in \mathbb{N}} E_i) = \cup_{i \in \mathbb{N}} f^{-1}(E_i)$ for all $E_i \subseteq X$.

(ii) $f^{-1}(E^c) = (f^{-1}(C))^c$, for all $E \subseteq X$.

Proof. By the first statement of the preceding note,

$$\mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{E}_i\}\right) \subseteq \mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{M}_i\}\right) = \bigotimes_{i \in \mathbb{N}} \mathcal{M}_i$$

For equality, it suffices to show

$$\mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{M}_i\}\right) \subseteq \mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{E}_i\}\right)$$

Let $\mathcal{M}\left(\{\prod_{i \in \mathbb{N}} E_i \mid E_i \in \mathcal{E}_i\}\right) = \mathcal{A}$.
Note that

$$\begin{aligned}\prod_{i \in \mathbb{N}} E_i &= \{x \in X \mid \pi_i(x) \in E_i \forall i\} \\ &= \bigcap_{i \in \mathbb{N}} \{x \in X : \pi_i(x) \in E_i\} \\ &= \bigcap_{i \in \mathbb{N}} \pi_i^{-1}(E_i)\end{aligned}$$

Because $\mathcal{A}$ is a σ-algebra, it suffices to show that $\pi_i^{-1}(E_i) \in \mathcal{A}$ for all $E_i \in \mathcal{M}_i$.

Claim. *Let* $\mathcal{F}_i := \{E_i \subseteq X_i \mid \pi_i^{-1}(E_i) \in \mathcal{A}\}$.
This is a σ*-algebra.*

Proof.

$$\bigcup_{i=1}^{\infty} \pi_i^{-1}(E_i) = \pi_i^{-1}(\bigcup_{i=1}^{\infty} E_i)$$

So $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}_i$.
Similarly,

$$\pi_i^{-1}(E^c) = (\pi_i^{-1}(E))^c$$

so $E \in \mathcal{F}_i$ implies $E^c \in \mathcal{F}_i$.

Lecture 3

Because $X_i \in \mathcal{E}_i$and $\pi_i^{-1}(E_i) = X_1 \times X_2 \times \cdots \times E_i \times \cdots$, we know $\pi_i^{-1}(E_i) \in \mathcal{A}$ for all $E_i \in \mathcal{E}_i$.
In other words, $\mathcal{E}_i \subseteq \mathcal{F}_i$. Since $\mathcal{F}_i$ is a σ-algebra, $\mathcal{M}(\mathcal{E}_i) = \mathcal{M}_i \subseteq \mathcal{F}_i$.
Thus, $\pi_i^{-1}(\mathcal{E}_i) \in \mathcal{A}$ for all $E_i \in \mathcal{M}_i$ and all i.

■

In order to characterize the Borel product σ-algebra, it will be convenient to assume our undderlying spaces have a metric that induces the topology.
Let $(X_i, d_i), i = 1, \ldots, n$ be metric spaces. Let

$$X = \prod_{i=1}^{n} X_i$$

Endow the product space with the metric

$$d_{\max}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) = \max_{i=1,\ldots,n} (d_1(x_1, y_1), d_2(x_2, y_2), \ldots, d_n(x_n, y_n))$$

Theorem 0.2. *Given metric spaces* $X_1, X_2, \ldots, X_d$ *and their product*

$$X = \prod_{i=1}^{d} X_i$$

endowed with the metric $d_{\max}$*, then* $\otimes_{i=1}^{n} \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$
If the X_i *are all seperable, then* $\otimes_{i=1}^{n} \mathcal{B}_{X_i} = \mathcal{B}_X$
Remark. Since the definition of $\mathcal{B}_X$ only depends on the topology of X, then this statement holds even if $d_{\max}$ is replaced by an equivalent metric, where "equivalent" means "generates the same topology."

Remark. $d_{\max}$ is convenient because:

$$\begin{aligned} B_r(x_1, \ldots, x_m) &= \{(y_1, \ldots, y_n) | d_{\max}(\vec{x}, \vec{y}) < r\} \\ &= \{(y_2, \ldots, y_n) \mid d_i(x_i, y_i) < r \forall i\} \\ &= \prod_{i=1}^{n} B_r(X_i) \end{aligned}$$

Recall:
Fact 1: If $X_1, \ldots, X_m$ are seperable, so is $\prod_{i=1}^{m} X_i$.
Fact 2: In a seperable metric space, every open set can be written as a countable union of balls, $\mathcal{U} = \cup_{i=1}^{\infty} B_i$
Fact 3: $\{\prod_{i=1}^{n} E_i \mid E_i \subseteq X_i, \text{open}\} \subseteq \{\text{open subsets of } X\}$

Proof. : By the previous proposition, $\otimes_{i=1}^{n} \mathcal{B}_{X_i}$ is generated by

$$\{\prod_{i=1}^{n} E_i \mid E_i \subseteq X_{\text{open}}\} \subseteq \{\text{open subsets of } X\}$$

Thus $\otimes_{i=1}^{n} \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$.
Now, suppose $X_1, \ldots, X_n$ are seperable. By facts 1 and 2, every open subset of X can be written as a countable union of balls.
To prove $\otimes_{i=1}^{n} \mathcal{B}_{X_i} = \mathcal{B}_X$, it suffices to show that

$$\{\text{open subsets of } X\} \subseteq \otimes_{i=1}^{n} \mathcal{B}_{X_i}$$

The left hand side is equal to $\{\cup_{j=1}^{\infty} B_j \mid B_j \subseteq X_{\text{open}} \text{ball}\}$, and the right hand side is equal to $\mathcal{M}\left(\{\prod_{i=1}^{n} E_i \mid E_i \text{ open}\}\right)$
This will hold, as long as we can show $B_j \in \mathcal{M}\left(\{\prod_{i=1}^{n} E_i \mid E_i \text{ open}\}\right)$
Since X is endowed with $d_{\max}$, we know that any open ball in X can be expressed as $B = \prod_{i=1}^{n} B_i$, where $B_i \subseteq X_i$ is a ball. This gives the result. ■

Now, it is finally time to talk about measures.

Measures

Call $(X, \mathcal{M})$ a measurable space when X is a set and $\mathcal{M}$ is a σ-algebra on X.

Definition 0.6. A <u>measure</u> on a measurable space $(X, \mathcal{M})$ is a function $\mu : \mathcal{M} \to [0, +\infty]$ such that

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

(ii) If $\{E_i\}$ is a countable disjoint collection of sets, then

$$\mu(\bigcup E_i) = \sum \mu(E_i)$$

This is called "countable (disjoint) additivity"

Example 0.4. (Dirac mass/Dirac measure)
Let $(X, \mathcal{M}) = (X, 2^X)$.
Fix $x_0 \in X$ and define

$$\mu(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases}$$

Example 0.5. (Counting measure)
Let $(X, \mathcal{M}) = (X, 2^X)$. Define

$$\mu(A) = |A| = \text{ the number of elements in } A$$

Given a measurable space $(X, \mathcal{M})$ and a measure μ, we call $(X, \mathcal{M}, \mu)$ a <u>measure space</u> and $E \in \mathcal{M}$ a <u>measurable set</u>

Theorem 0.3. *For any measure space $(X, \mathcal{M}, \mu)$ and measurable sets $A, B, A_1, A_2, \cdots \in \mathcal{M}$,*

(i) $A \subseteq B \implies \mu(A) \leq \mu(B)$. This is called "monotonicity"

(ii) $\mu(\bigcup_i A_i) \leq \sum_i \mu(A_i)$. This is called "(countable) sub additivity)."

(iii) If $A_i \subseteq A_{i+1}$, then $\mu(\bigcup_i A_i) = \lim_{n\to\infty} \mu(A_i)$. This is called "continuity from below."

(iv) If $A_{i+1} \subseteq A_i$ for all i, and $\mu(A_i) < \infty$, then $\mu(\bigcap A_i) = \lim_{i\to\infty} \mu(A_i)$. This is called "continuity from above."

Remark. For (iv), why do we need the additional hypothesis $\mu(A_1) < \infty$?.
Consider the counting measure on $(\mathbb{N}, 2^{\mathbb{N}})$, and $A_i = \{n \in \mathbb{N} \mid n \geq i\}$, which satisfies $A_{i+1} \subseteq A_i$, but it fails $\mu(A_1) < \infty$:

$$0 = \mu(\varnothing) = \mu(\cap_{i=1}^{\infty} A_i) \neq \lim_{i\to\infty} \mu(A_i) = +\infty$$

Proof. (i) Since $A \subseteq B$, $B = A \cup (B\backslash A)$, so $\mu(B) = \mu(A) + \mu(B\backslash A)$ by countable additivity. $\mu(B\backslash A) \geq 0$, so (i) follows.

(ii) Define $B_1 = A_1$, $B_2 = A_2 \backslash A_1$, $B_3 = A_3 \backslash (A_1 \cup A_2), \dots, B_n = A_n \backslash (\cup_{i=1}^{n-1} A_i)$. Then $\cup_i A_i = \cup_i B_i$, so by countable disjoint additivity,

$$\mu(\cup_i A_i) = \mu(\cup_i B_i) = \sum_i \mu(B_i) \leq \sum_i \mu(A_i)$$

(iii) Define $B_1 = A_1$, and $B_i = A_i \backslash A_{i-1}$. Then $A_n = \cup_{i=1}^n B_i$, so $\cup_{i=1}^\infty A_i = \cup_{i=1}^\infty B_i$. Thus $\mu(A_n) = \mu(\cup_{i=1}^n B_i) = \sum_{i=1}^n \mu(B_i)$. Consequently,

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

(iv) Next time!