LECTURE: MEASURE THEORY

1. Introduction

The central goal of measure theory is to generalize our intuitive ideas of length, area, and volume to subsets of arbitrary sets. We will see that our new theory not only encompasses our commonsense notions of length and area but also allows us to "measure" subsets which we would not be able to do with a ruler.

Application 1: Probability The measure of a set is the probability that an event occurs. As an example, if the sample space is $C([0,1])$, we could ask ourselves: What is the measure of the subset of differentiable functions?

Application 2: Integration theory: Measure theory lets us define the Lebesgue integral, which allows us to integrate more functions, and gives us more general theorems for when we can exchange limits and integration. This resolves the problems we've seen last time.

Application 3: Nonlinear PDE: Measure theory is the main tool used to deal with nonlinear PDE, since we can't use the Fourier transform

2. Measures

Definition:

A *measure* on a set X is a function μ on subsets of X with the following properties:

(1) $\mu(E) \in [0, \infty]$

(2) $\mu(\varnothing) = 0$.

(3) Countable additivity: If the sets $\{E_k\}_{k\in\mathbb{N}}$ are disjoint subsets of X, then

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

We would also like μ to satisfy (not part of the definition)

(4) μ is invariant under symmetries such as translations, rotations, and reflections.

(5) μ agrees with our common notion of length, area, and volume. For example, we would like $\mu([0, 1]) = 1$.

Ideally, we would like μ to be defined on all subsets of X. However, it turns out to be impossible to construct such a measure that has these five properties, and there are "non-measurable sets."

Our goal is to find a collection of subsets of on which we can define a measure which is as large as possible and contains everything we care about. The collection we will need is called a σ-algebra:

Definition:

A collection $\mathcal{M}$ of subsets of X is a **σ-algebra** if

(1) $\varnothing \in \mathcal{M}$.

(2) If $E \in \mathcal{M}$ then $X \setminus E \in \mathcal{M}$.

(3) If $\{E_k\}_{k\in\mathbb{N}} \subset \mathcal{M}$, then $\bigcup_{k=1}^{\infty} E_k \in \mathcal{M}$.

An **algebra** is the same thing, except it is only closed under finite unions.

Other properties that follow include:

Fact:

(1) $X \in \mathcal{M}$

(2) $\mathcal{M}$ is closed under countable intersections

(3) $\mathcal{M}$ is closed under relative complements, i.e. if $E, F \in \mathcal{M}$, then $E \setminus F \in \mathcal{M}$.

Here are some simple examples of σ-algebras on X

Example 1:

(1) $M = \{\varnothing, X\}$ (the smallest σ-algebra)

(2) $M = \{A, A^c, \varnothing, X\}$ (the σ-algebra generated by $A \subseteq X$)

(3) $M = \{\text{all subsets of X}\}$ (the largest σ-algebra, also called the power set of X).

The first σ-algebra which is of interest is the σ-algebra generated by a collection of subsets.

Definition:

Let $\mathcal{E}$ be a collection of subsets of X. Then the **σ-algebra generated by** $\mathcal{E}$, denoted $\mathcal{M}(\mathcal{E})$, is the unique smallest σ-algebra containing $\mathcal{E}$. It is defined in one of two equivalent ways.

(1) $\mathcal{M}(\mathcal{E})$ is the intersection of all $\sigma-$algebras containing $\mathcal{E}$.

(2) If $\mathcal{N}$ is another σ-algebra containing $\mathcal{E}$, then $\mathcal{M}(\mathcal{E}) \subset \mathcal{N}$.

In other words, think of starting with $\mathcal{E}$ and take countable unions, intersections, complements etc. until you get a $\sigma-$algebra. It is worth noting that there is essentially no way of writing down an arbitrary element of $\mathcal{M}(\mathcal{E})$ in terms of the elements of $\mathcal{E}$. The most important of these σ-algebras is the Borel σ-algebra, which is generated by the open sets:

Definition:

The **Borel σ-algebra** on X, denoted $\mathcal{B}_X$, is the σ-algebra generated by the collection of open sets of X.

The Borel $\sigma-$algebra contains all open sets, all closed sets, all countable unions and countable intersections of open and closed sets (some of which may be neither open nor closed), etc. There is no nice way to write down a generic element of $\mathcal{B}_X$.

For the special case of the Borel $\sigma-$algebra on $\mathbb{R}$, we have the following nice result:

Fact:

The Borel $\sigma-$algebra on $\mathbb{R}$ is generated by any of the following:

(1) All open intervals, i.e. all intervals of the form (a, b).

(2) All intervals of one of the following forms: $[a, b], (a, b], [a, b)$.

(3) All rays of one of the following forms $(a, \infty), [a, \infty), (-\infty, b), (-\infty, b]$.

Proof: Let $\mathcal{E}$ be one of these collections. Since $\mathcal{E} \in \mathcal{B}_{\mathbb{R}}$, $\mathcal{M}(\mathcal{E}) \in \mathcal{B}_{\mathbb{R}}$. All we have do now is show that $\mathcal{M}(\mathcal{E})$ contains all open sets.

For (1), we showed earlier that every open set in $\mathbb{R}$ is the countable union of disjoint open intervals.

For (2) you use

$$[a, b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n} \right) \text{ and } (a, b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n} \right]$$

And similarly for (3)

Definition:

A **measure space** $(X, \mu, \mathcal{M})$ is a set X, along with a $\sigma-$algebra of measurable sets on X and a measure $\mu : \mathcal{M} \to [0, \infty]$ on $\mathcal{M}$

A measure μ has many nice properties, such as a following:

Fact:

Let X be a set, and μ a measure defined on a σ-algebra $\mathcal{M}$. Then μ has the following properties.

(1) (Monotonicity) If $E \subset F$, $\mu(E) \leq \mu(F)$. In addition, if $\mu(E) < \infty$, then $\mu(F \setminus E) = \mu(F) - \mu(E)$.

(2) (Countable Subadditivity) For any sequence of sets $\{E_n\}$

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

(3) (Continuity from Below) If $\{E_n\}$ is an increasing sequence of nested sets, i.e. $E_n \subset E_{n+1}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

(4) (Continuity from Above) If $\{E_n\}$ is an decreasing sequence of nested sets, i.e. $E_n \supset E_{n+1}$, and $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

As the example $E_n = (n, \infty)$ shows, we really need $\mu(E_1) < \infty$ in (4)

3. The Lebesgue Measure

Here we'll construct the Lebesgue measure, often denoted m, which measures length of subsets of $\mathbb{R}$.

First, we define the **Lebesgue outer measure**

Motivation: Consider any subset E of $\mathbb{R}$. We can find always find a countable collection of open intervals $\{(a_n, b_n)\}$ such that

$$E \subset \bigcup_{n=1}^{\infty} (a_n, b_n).$$

In addition, it makes intuitive sense that we should have

$$\text{length of } E \leq \sum_{n=1}^{\infty} (b_n - a_n),$$

since E is contained in the union of these intervals.

Definition:

For any subset E of $\mathbb{R}$, the **Lebesgue outer measure** is the function m^*, taking values in $[0, \infty]$ (∞ is allowed!) defined by

$$m^*(E) = \inf \left\{ \sum_{n=1}^{\infty} (b_n - a_n) : E \subset \bigcup_{n=1}^{\infty} (a_n, b_n) \right\}.$$

Essentially, we are "approximating E from the outside" by open intervals, which is why we call this outer measure. The idea is that we cover E with a countable collection of open intervals and add up their lengths; the outer measure is the infimum over all such covers (which may still be infinite!)

Formally, an outer measure is defined as follows.

Definition:

An **outer measure** on a set X is a function μ^* defined on *all* subsets of X with the following properties:

(1) $\mu^*(E) \in [0, \infty]$

(2) $\mu(\varnothing) = 0$.

(3) (Monotonicity) If $E \subset F$ then $\mu^*(E) \leq \mu^*(F)$

(4) (Countable subadditivity) $\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} \mu(E_k)$.

It is important to note that the outer measure μ^* is defined on all subsets of X. Going back to the Lebesgue outer measure, we can show the following:

Fact:

(1) m^* is an outer measure

(2) If I is an interval then $m^\star(I) =$ Length of I

Next, we show that the Lebesgue outer measure of any countable set is 0.

Fact:

Let $E \subset \mathbb{R}$ be a countable set. Then $m^*(E) = 0$.

Proof: This prove uses the famous $\frac{\epsilon}{2^n}$ trick

Let $x_1, x_2, \cdots$ an enumeration of the elements in E

Let $\epsilon > 0$ be given

Cover each x_n with the open interval $I_n = (x_n - r_n, x_n + r_n)$, where

$$r_n = \left(\frac{1}{2}\right)\left(\frac{\epsilon}{2^n}\right)$$

By construction, each interval I_n has length $2r_n = \epsilon/2^n$

Since the outer measure is the inf of the total length of all such covers,

$$m^*(E) \leq \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon,$$

where we used the infinite sum of the geometric series. Since ϵ is arbitrary, $m^*(E) = 0$ □

Corollary:

$$m^\star(\mathbb{Q}) = 0$$

Next, for any outer measure μ^*, we define the concept of μ^*-measurability.

Definition:

Let μ^* be an outer measure on X. Then a subset $A \subset X$ is μ^**-measurable* if for every test set $E \subset X$,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A).$$

Intuitively a set A is $\mu^\star-$measurable if you can split any arbitrary "test set" E into two "nice pieces" using A.

To complete the construction of the Lebesgue measure, we will use the Carathéadory Extension Theorem:

Carathéadory Extension Theorem:

Let μ^* be an outer measure on X, and let $\mathcal{M}$ be the collection of μ^*-measurable sets. Then

(1) $\mathcal{M}$ is a σ-algebra.

(2) μ^* ,when restricted to $\mathcal{M}$, is a measure, which we designate μ.

(3) $\mathcal{M}$ contains all μ^*-null sets, i.e. all sets N with $\mu^*(N) = 0$.

Proof-Sketch: (will be skipped)

STEP 1: Show $\mathcal{M}$ is closed under complements and finite unions, thus is an algebra.

STEP 2: Show μ^* is finitely additive on $\mathcal{M}$.

STEP 3: Show $\mathcal{M}$ is closed under countable disjoint unions. This implies it is closed under countable unions, and thus is a σ-algebra. To see why this works, take any sequence of sets $\{E_n\}$. Construct F_n from E_n by deleting anything which overlaps with $E_1, \ldots, E_{n-1}$. We can get from E_n to F_n, or vice versa, by using a finite number of set operations. In addition, by construction, $\cup E_n = \cup F_n$

STEP 4: Show μ^* is countably additive on $\mathcal{M}$, hence a measure on $\mathcal{M}$.

STEP 5: Show if $\mu^*(N) = 0$ then $N \in \mathcal{M}$ □

The last property is particularly important. If a set has measure 0, we can, in general, ignore it. This property guarantees that if we construct a measure this way, all μ^*-null sets are actually in the σ-algebra.

Lebesgue measure on $\mathbb{R}$

We can apply Carathéadory Extension Theorem on the Lebesgue outer measure m^* to get the Lebesgue measure on $\mathbb{R}$. Let $\mathcal{L}$ be the resulting $\sigma-$algebra of m^*-measurable sets, and let m be the measure we obtain by restricting m^* to $\mathcal{L}$.

All that is left is to show that $\mathcal{L}$ contains the entire Borel $\sigma-$algebra. The easiest way to do this is to show that all open intervals are μ^*-measurable. Since $\mathcal{L}$ is a $\sigma-$algebra, it must then contain the entire Borel $\sigma-$algebra. But beware that $\mathcal{L}$ is larger than the Borel $\sigma-$algebra.

We can then show that the Lebesgue measure m is translation-invariant.

4. Measurable Functions

We can now define measurable functions, which are the analog of random variables in probability.

Definition:

Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be sets together with $\sigma-$algebras. Then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$**-measurable** (or just **measurable**) if, for all $F \in \mathcal{N}$, $f^{-1}(F) \in \mathcal{M}$.

This is analogous to our definition of continuous functions, except we use measurable sets instead of open sets.

In particular, compositions of measurable functions are measurable.

Although the definition of measurability can be applied to functions between any two arbitrary measure spaces, we will only consider real-valued functions from now on.

Definition:

(1) $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable if f is $(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{R}})$-measurable.

(2) $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable if f is $(\mathcal{L}, \mathcal{B}_{\mathbb{R}})$-measurable, where $\mathcal{L}$ is the Lebesgue σ-algebra on $\mathbb{R}$.

The main issue with the definition of measurability is that it is hard to verify, since we don't have a good way to characterize what sets belong to a given σ-algebra. Luckily, it suffices to check the measurability criterion on a set which generates a σ-algebra

Fact:

Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measure spaces, and suppose $\mathcal{N}$ is the σ-algebra generated by the collection of sets $\mathcal{E}$. Then $f : X \to Y$ is measurable if and only if $f^{-1}(E) \subset \mathcal{M}$ for all $E \in \mathcal{E}$.

Proof-Sketch:

($\Longrightarrow$) This follows from the definition of measurability, since $\mathcal{E} \subset \mathcal{N}$.

($\Longleftarrow$) Define the set

$$\mathcal{H} = \{A \in \mathcal{N} : f^{-1}(A) \in \mathcal{M}\}.$$

We want to show that $\mathcal{H}$ contains $\mathcal{N}$. From our initial assumption, $\mathcal{H}$ contains $\mathcal{E}$. Next, we show that $\mathcal{H}$ is a σ-algebra, by using the definition of a σ-algebra and the fact the the inverse image operator f^{-1} commutes with set operations. Since $\mathcal{H}$ is a σ-algebra containing $\mathcal{E}$, it must also contain $\mathcal{N}$, since $\mathcal{N}$ is the σ-algebra generated by $\mathcal{E}$ □

We have the following important corollary.

Corollary:

Every continuous function $f : (X, \mathcal{B}_X) \to (Y, \mathcal{B}_Y)$ is measurable.

We can also define measurability in terms of the generators of the Borel σ-algebra on $\mathbb{R}$. The infinite rays are more useful than the intervals in this case.

Corollary:

A function $(X, \mathcal{M}) \to \mathbb{R}$ is measurable if and only if one of the following is true.

(1) $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

(2) $f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

(3) $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

(4) $f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

The fourth one is most useful in practice. If we call the set $f^{-1}(a)$ the level set of a, then the set

$$f^{-1}((-\infty, a]) = \{x : f(x) \leq a\}$$

can be called the sublevel set of a. Thus, for measurability of real-valued functions, we only have to check the sublevel sets.

Note: This is why often in probability you deal with sets of the form $\{X \leq c\}$ where X is a random variable

Fact:

Let f, g and $\{f_n\}_{n\in\mathbb{N}}$ be measurable functions from $(X, \mathcal{M})$ to $\mathbb{R}$. Then the following are measurable.

(1) $f + g$ and fg

(2) $\max(f, g)$ and $\min(f, g)$

(3) $\sup_n f_n(x)$ and $\inf_n f_n(x)$

(4) $\limsup_{n\to\infty} f_n(x)$ and $\liminf_{n\to\infty} f_n(x)$.

(5) $\lim_{n\to\infty} f_n(x)$, provided the limit exists.

5. LIMINF AND LIMSUP

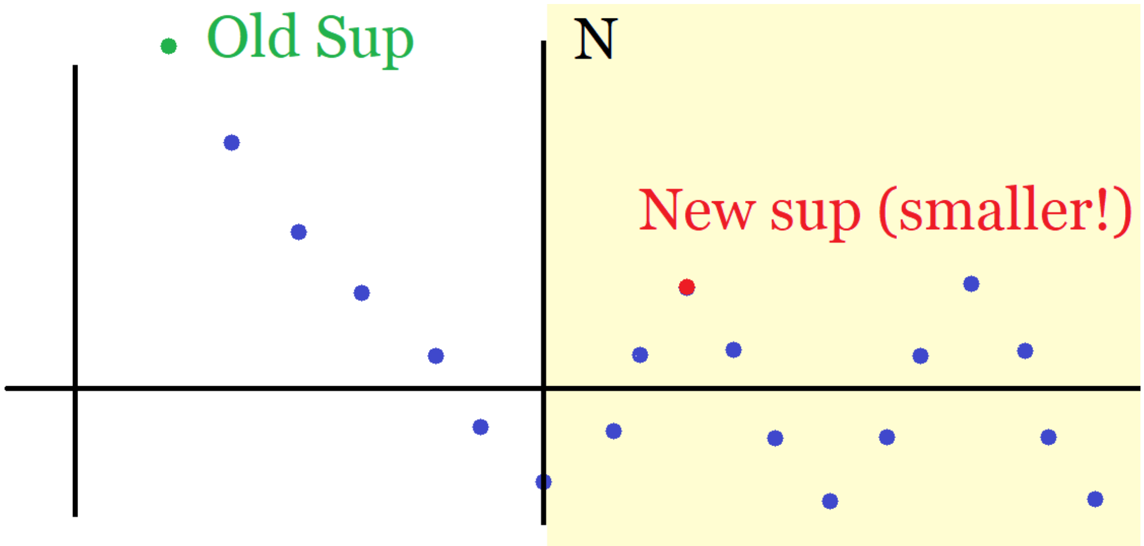
Definition:

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\sup_{k\geq n} x_k\right) \qquad \limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\inf_{k\geq n} x_k\right)$$

We can think of the limsup as the limit of the sequence $\{y_n\}$, given by

$$y_n = \sup_{k\geq n} x_k$$

This is a decreasing sequence, since we are taking the supremum over fewer and fewer terms.



The limit superior of $\{x_n\}$ is the limit of this sequence, which is the infimum since the sequence is decreasing.

Similarly, the limit inferior is the limit of an increasing sequence.

As long as we allow the values $\pm\infty$, $\limsup_{n\to\infty} x_n$ and $\liminf_{n\to\infty} x_n$ exist for all real-valued sequences, and

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$$

Fact:

$\lim_{n\to\infty} x_n$ exists if and only if $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$