`MA40042` Measure Theory and Integration

Lecture 0
Introduction

- Examples and properties of measures
- Discussion of measure, probability, and integration
- Properties of measurable sets

In this lecture we will discuss some motivation about measures and their properties. Formal definitions and theorems will begin in the next lecture.

0.1 How big is a set?

Measure theory is about the following question: *How big is a set?*
For example, you'll already know some ways of measuring the size of a set:

Cardinality Let X be a nonempty set – let's say the real numbers. Then for any set $A \subset X$, the cardinality of A, which we'll write $\#(A)$, is the number of elements in A. So if $X = \mathbb{R}$, then we have $\#(\{1,2,3\}) = 3$, $\#(\{5\}) = 1$, $\#(\varnothing) = 0$, and $\#(\mathbb{Z}) = \infty$, for example.

Length Let $X = \mathbb{R}$. Then for sets $A \subset X$ we can measure A by its length $\lambda(A)$. For example, the length of an interval

$$[a,b) = \{x \in \mathbb{R} : a \leq x < b\}$$

(where $b > a$) is $\lambda\big([a,b)\big) = b - a$. We can also assign lengths to more complicated sets, such as $\lambda\big((-\infty,0)\big) = \infty$, $\lambda(\{4\}) = 0$, and $\lambda\big([0,1) \cup [3,6)\big) = 1 + 3 = 4$, for example.

Area Let $X = \mathbb{R}^2$. Then we can measure a set $A \subset \mathbb{R}$ by its area. For example, the area of a square $[a,b) \times [c,d)$ is $(b-a)(d-c)$. We can also give the area of more complicated shapes – for example, the area of a disc of radius r is πr^2.

Volume Let $X = \mathbb{R}^3$ We can measure subsets of X by their volume, in a similar way to area in $\mathbb{R}^2$.

Later in this course, we will call the cardinality the *counting measure*, and we will call length, area, and volume the *Lebesgue measure* on $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$ respectively.

0.2 What properties should a measure have?

The idea of measure theory is to generalise the ideas of cardinality, length, area, and volume to allow other 'measures' of sets.
So what properties do these examples share that we'll want any 'measure' μ to also share?

- Suppose we are working with a base set X. Then our measure μ should assign a positive real number to subsets of X, which can be 0 or ∞. **Property 1:** *A measure is a function μ from subsets of X to $[0,\infty]$.*
- The empty set $\varnothing$ has a special role to play. The empty set has cardinality 0, and its clear that the length/area/volume of the empty set should also be 0, and this should be a general rule. **Property 2:** *The measure of the empty set is $\mu(\varnothing) = 0$.*
- The above examples of measures behave well with unions. For example,

$$5 = \#\big(\{1,2,3,8,9\}\big) = \#\big(\{1,2,3\}\big) + \#\big(\{8,9\}\big) = 3 + 2,$$

  and the area of a collection non-overlapping shapes is the sum of their areas. The key here is that the sets are disjoint, in that they have no intersection. **Property 3:** *If $A_1, A_2, \ldots, A_N$ are disjoint sets, then*

$$\mu\big(A_1 \cup A_2 \cup \cdots \cup A_N\big) = \mu(A_1) + \mu(A_2) + \cdots + \mu(A_N).$$

- In fact, this is also true for countably infinite unions, like

$$\infty = \#\big(\{1,2,\ldots\}\big) = \#\big(\{1\}\big) + \#\big(\{2\}\big) + \cdots = 1 + 1 + \cdots,$$

  and

$$1 = \lambda\big([0,1)\big) = \lambda\left(\left[\frac{1}{2},1\right)\right) + \lambda\left(\left[\frac{1}{4},\frac{1}{2}\right)\right) + \lambda\left(\left[\frac{1}{8},\frac{1}{4}\right)\right) + \cdots$$
$$= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots .$$

  Property 4: *If $A_1, A_2, \ldots$ is a countably infinite sequence of disjoint sets, then*

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

- Note however, we don't (necessarily) want this to hold for uncountable unions, since

$$1 = \lambda\big([0,1)\big) \neq \sum_{x \in [0,1)} \lambda\big(\{x\}\big) = \sum_{x \in [0,1)} 0 = 0.$$

- We can even say something about unions when the sets are not disjoint: the sum of the measures is at least as big as the measure of the unions. For example,

$$5 = \#\big(\{1,2,3,4,5\}\big) \leq \#\big(\{1,2,3\}\big) + \#\big(\{3,4,5\}\big) = 3+3,$$

or

$$3 = \lambda\big([0,3)\big) \leq \lambda\big([0,2)\big) + \lambda\big([1,3)\big) = 2+2.$$

Property 5: *If $A_1, A_2, \dots$ is a finite or countably infinite sequence of disjoint sets, then*

$$\mu\left(\bigcup_n A_n\right) \leq \sum_n \mu(A_n).$$

We will later to define a measure on a set X to be a function μ satisfying these properties.

0.3 Probability

Probability theory is another place where we've seen 'the size of a set' before.

Here we write Ω for the 'sample space', and sets $A \subset \Omega$ are called 'events'. But the probability $\mathbb{P}$ acts a lot like a measure; $\mathbb{P}(A)$ tells us how likely the event A is, or how much 'possibility' it contains.

We know that the probability $\mathbb{P}$ obeys the properties above. For example, $\mathbb{P}(\varnothing) = 0$, and for disjoint (or 'mutually exclusive') events (A_n) the 'or' probabilities sum like the 'union bound'

$$\mathbb{P}\left(\bigcup_n A_n\right) \leq \sum_n \mathbb{P}(A_n).$$

Another property is that the probability that A doesn't occur is

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^\mathsf{c}),$$

This is in fact a version of a general property. If we rewrite it as

$$\mathbb{P}(A) + \mathbb{P}(A^\mathsf{c}) = \mathbb{P}(\Omega),$$

where $\mathbb{P}(\Omega) = 1$, we see this as being a version of the more general statement

$$\mu(A) + \mu(A^\mathsf{c}) = \mu(X).$$

(It's just that in the previous examples, $\mu(X)$ was ∞.

- **Property 6:** $\mu(A) + \mu(A^\mathsf{c}) = \mu(X)$.

A lot of the incentive to study measure theory comes from probability, and many of the applications of measure theory are also in probability.

0.4 Integration

Integration is also related to measure theory.

Once we have defined sets with measures, we can look at functions between those sets that are 'well-behaved' with respect to these measures. (Random variables are an example, in probability theory.)

We can also look at integrals of these functions. In some sense, the integral of a function over a set measures the area under a curve on that set. This is clearly related to 'how big' the function is on that set – or how big that set is when 'weighted' by the function.

Integrals are related to the expectation of a random variable in probability theory.

These issues will be dealt with further in the second half of the course.

0.5 Which sets have a measure?

Sometimes it makes sense to define the measure for any subset $A \subset X$. However, on other occasions, we want to restrict the measure to just some of the subsets of X.

For example, on uncountable sets like $\mathbb{R}$ or $\mathbb{R}^d$, the length/area/volume cannot properly be defined for every subset. (Those who know about the Banach–Tarski paradox already know this; we'll see a concrete example later in the course.)

The collection of 'measurable sets' will need some structure though – enough to allow us to make use of the properties of measures we listed above. Let's go through the properties and see what they demand:

Property 1 does not make any requirements.

Property 2 requires the empty set $\varnothing$ to be measurable.

Properties 3, 4, and 5 require finite or countably infinite unions of sets to be measurable.

Property 6 requires that X itself be measurable, and the complements of measurable sets be measurable.

Next time we will look at *σ-algebras*, which are the structures that satisfy these.

Matthew Aldridge
`m.aldridge@bath.ac.uk`