Lecture 2
The Borel σ-algebra

- Review of open, closed, and compact sets in $\mathbb{R}^d$
- The Borel σ-algebra on $\mathbb{R}^d$
- Intervals and boxes in $\mathbb{R}^d$

For countable sets X, the powerset $\mathcal{P}(X)$ is a useful σ-algebra. But in larger sets, such as $\mathbb{R}$ or $\mathbb{R}^d$, the full powerset is often too big for interesting measures to exist, such as the Lebesgue length/area/volume measure. (We'll see a example of this later in the course.)

On $\mathbb{R}$ or $\mathbb{R}^d$ a useful σ-algebra is the *Borel σ-algebra*. The benefits of the Borel σ-algebra are:

- it's large enough to contain any set you're likely to come across (although we will later construct a specific set that isn't in it);
- it's not so large that useful measures (like the Lebesgue measure) fail to exist.

First, we'll need to recall some facts about open and closed sets to define this.

2.1 Open and closed sets

Recall the following definitions about open and closed sets in $\mathbb{R}^d$.

Definition 2.1. Write

$$B_d(\mathbf{x}, r) := \{\mathbf{y} \in \mathbb{R}^d : |\mathbf{y} - \mathbf{x}| < r\}$$

for the open ball of radius r about $\mathbf{x} \in \mathbb{R}^d$. (We'll suppress the subscript d when the dimension is clear from context.)

- A set $G \subset \mathbb{R}^d$ is *open* if for all $\mathbf{x} \in G$ there exists an $r > 0$ such that $B(\mathbf{x}, r) \subset G$.
- A set $F \subset \mathbb{R}^d$ is *closed* if its complement F^{c} is open.
- A set $K \subset \mathbb{R}^d$ is *compact* if every open cover of K has a finite subcover. That is, if K can be 'covered' by a (not necessarily countable) union of open sets

$$K \subset \bigcup_{\alpha \in I} G_\alpha, \qquad G_\alpha \subset \mathbb{R}^d \text{ open for all } \alpha \in I,$$

then there exists a finite subset $\{\alpha_1, \dots, \alpha_N\} \subset I$ that also covers K,

$$K \subset \bigcup_{n=1}^{N} G_{\alpha_n}.$$

Checking whether a set is compact under this definition can be difficult, but is made much easier by the following theorem (which we won't prove here).

Theorem 2.2 (Heine–Borel theorem)**.** *A set $K \subset \mathbb{R}^d$ is compact if and only if it is closed and bounded.*

Bounded here means that $K \subset B(\mathbf{x}, r)$ for some $\mathbf{x} \in \mathbb{R}^d$ and some $r \in (0, \infty)$.

2.2 $\mathcal{B}(\mathbb{R}^d)$

Fix d, and write $\mathcal{G}$, $\mathcal{F}$, and $\mathcal{K}$ for, respectively, the collection of open, closed, and compact sets in $\mathbb{R}^d$.

Definition 2.3. The *Borel σ-algebra* on $\mathbb{R}^d$ is the σ-algebra generated by the open sets, $\mathcal{B}(\mathbb{R}^d) := \sigma(\mathcal{G})$.

A set $A \in \mathcal{B}(\mathbb{R}^d)$ is called a *Borel set*.

We sometimes write $\mathcal{B}$ for $\mathcal{B}(\mathbb{R})$.

We could instead generate $\mathcal{B}(\mathbb{R}^d)$ with the closed sets:

Theorem 2.4. $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{F})$.

Proof. By the definition of $\mathcal{B}(\mathbb{R}^d)$, we need to show $\sigma(\mathcal{G}) = \sigma(\mathcal{F})$. We will do this in two parts: showing that $\sigma(\mathcal{G}) \subset \sigma(\mathcal{F})$, and that $\sigma(\mathcal{F}) \subset \sigma(\mathcal{G})$.

First, $\sigma(\mathcal{G}) \subset \sigma(\mathcal{F})$. Since $\sigma(\mathcal{G})$ is the smallest σ-algebra containing $\mathcal{G}$, and $\sigma(\mathcal{F})$ is a σ-algebra, it suffices to show that $\mathcal{G} \subset \sigma(\mathcal{F})$. Let G be an open set, and let $D = G^{\mathsf{c}}$. Since $D^{\mathsf{c}} = G$ is open, D is closed, so $D \in \mathcal{F}$. Hence $G = D^{\mathsf{c}}$ is the complement of a set in $\mathcal{F}$, so is in $\sigma(\mathcal{F})$.

Second, $\sigma(\mathcal{F}) \subset \sigma(\mathcal{G})$, for which again $\mathcal{F} \subset \sigma(\mathcal{G})$ suffices. For F closed, write $E = F^{\mathsf{c}}$. Then E is open, and $E^{\mathsf{c}} = F$ is in $\sigma(\mathcal{G})$, as desired. □

We can also show that $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{K})$; this is a homework problem.

2.3 Intervals and boxes

Intervals and boxes are important sets, as we 'know' what their lengths (or areas, or volumes) should be.

Recall the definition of the semi-open interval

$$[a,b) = \{x \in \mathbb{R} : a \le x < b\}, \qquad \text{for } a < b.$$

These are convenient because we have for $a < b < c$ that the union $[a,b) \cup [b,c) = [a,c)$ is a disjoint one. We could use instead the open intervals (a,b) or closed intervals $[a,b]$, but it would be a bit fiddly later on in the course. Note though the open intervals are indeed open, and the closed intervals closed.

Definition 2.5. A set $I \subset \mathbb{R}$ is an *interval* is it is of one of the following forms:

- the empty set, $I = \varnothing$;
- a semi-open interval, $I = [a,b)$ for $a < b$;
- an infinite or semi-inifite interval, $I = (-\infty, b)$, $I = [a, \infty)$, or $I = \mathbb{R}$.

We write $\mathcal{I}$ for the set of such intervals.

Again, the empty, semi-infinite, and infinite intervals are included in the definition for later convenience.

It will be useful to assume the convention $[a,b) = \varnothing$ for $b \le a$.

Definition 2.6. A set $I \subset \mathbb{R}^d$ is an *interval box* is it is of the form

$$I = I_1 \times I_2 \times \cdots \times I_d = \prod_{i=1}^{d} I_i, \qquad \text{for } I_1, I_2, \dots, I_d \in \mathcal{I}.$$

We write $\mathcal{I}_d$ for the set of interval boxes, or just $\mathcal{I}$ when the dimension is obvious by context.

Theorem 2.7. $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{I}_d)$.

So if we want the intervals and boxes to have length/area/volume (and we do!), we will have to work in the Borel σ-algebra.

We'll do the proof long-windedly for $d = 1$, then more quickly for the general case.

Proof for $d = 1$. As before, we need to show $\mathcal{I} \subset \mathcal{B} = \sigma(\mathcal{G})$ and $\mathcal{G} \subset \sigma(\mathcal{I})$.

First $\mathcal{I} \subset \mathcal{B}$. The empty set is open. For the semi-open case, we have

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

which writes $[a,b)$ as a countable intersection of open sets. For the infinite and semi-infinite cases, we have

$$(-\infty, b) = \bigcup_{n=1}^{\infty} [-n, b) \qquad [a, \infty) = \bigcup_{n=1}^{\infty} [a, n), \qquad \mathbb{R} = \bigcup_{n=1}^{\infty} [-n, n),$$

where we've already shown $[a,b)$ is a Borel set.

Second $\mathcal{G} \subset \sigma(\mathcal{I})$. Given an open set G, we obviously have

$$G = \bigcup_{x \in G} \{x\}.$$

Further, since G is open, for each $x \in G$ there exists a $r_x > 0$ such that $B(x, r_x) \subset G$, and so

$$G = \bigcup_{x \in G} B(x, r_x) = \bigcup_{x \in G} (x - r_x, x + r_x).$$

Setting $a_x = x - r_x/2$ and $b_x = x + r_x/2$, we then have

$$G = \bigcup_{x \in G} [a_x, b_x).$$

The good news is that we've written G as a union of intervals, but the bad news is that this union might be an uncountable one. But we can pull a clever trick here. Let c_x be a rational number in (a_x, x) and d_x a rational number in (x, b_x). Then

$$G = \bigcup_{x \in G} [c_x, d_x).$$

But there are only countable many intervals with rational endpoints, so by removing repeats we can write this union as a countable one. Hence we are done. □

General case. First, we have $\mathcal{I} \subset \sigma(\mathcal{G})$, since

$$\prod_{i=1}^{d} [a_i, b_i) = \bigcap_{n=1}^{\infty} \prod_{i=1}^{d} \left(a_i - \frac{1}{n}, b_i \right),$$

and the infinite boxes are countable unions of these.

Now we show $\mathcal{G} \subset \sigma(\mathcal{I})$. For each $\mathbf{x}$ in an open set G, let $r_{\mathbf{x}}$ be such that $B(\mathbf{x}, r_{\mathbf{x}}) \subset G$. Pick $c_{\mathbf{x},i}$ and $d_{\mathbf{x},i}$ to be rational numbers in $(x_i - r_{\mathbf{x}}/2\sqrt{d}, x_i)$ and $(x_i, x_i + r_{\mathbf{x}}/2\sqrt{d})$ respectively. (The $\sqrt{d}$ ensures the box fits inside the ball.) Then

$$G = \bigcup_{x \in G} \prod_{i=1}^{d} [c_{\mathbf{x},i}, d_{\mathbf{x},i}),$$

and because there are only countably many boxes with rational coordinates, the union consists of only countably many distinct sets. □