Supplementary: Why σ-algebra and measure?

Let $\Omega = [0, 1]$, the probability of the event $(a, b]$, where $0 \leq a \leq b < 1$ can be defined by

$$\mathbb{P}((a, b]) = F(b) - F(a),$$

where F is a non-decreasing and right-continuous function with

$$\lim_{x \to 0} F(x) = 0,$$
$$\lim_{x \to 1} F(x) = 1.$$

However, there are many other events like $\bigcup_{i=1}^{\infty} (a_i, b_i]$ whose probabilities we are interested in.

In particular, $\mathbb{P}$ should have the following properties:

(i) $\mathbb{P}(\Omega) = 1$.

(ii) If $A_1, A_2, \ldots$ are disjoint sets, then

$$\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} \mathbb{P}(A_i).$$

(iii) If A is congruent to B (i.e., A is B transformed by translation, rotation or reflection), then $\mathbb{P}(A) = \mathbb{P}(B)$.

Unfortunately, for these conditions to hold for *all* events would lead to inconsistency.

- To see why, define an equivalence $x \sim y$ iff $x - y$ is rational.
- Then Ω can be partitioned into equivalence classes.
- Let $N \subset \Omega$ be a subset that contains exactly one member of each equivalence class (we need the axiom of choice here).
- For each rational number $r \in \mathbb{Q} \cap [0, 1)$, let

  $$N_r = \{x + r \,:\, x \in N \cap [0, 1 - r)\} \cup \{x + r - 1 \,:\, x \in N \cap [1 - r, 1]\},$$

  i.e., N_r is N translated to the right by r with the part after $[0, 1)$ shifted to the front (wrapped around) so that $N_r \subset \Omega = [0, 1]$.

From properties (ii) and (iii), we have for any rational $r \in \mathbb{Q} \cap [0,1)$,

$$\mathbb{P}(N) = \mathbb{P}(N \cap [0, 1-r)) + \mathbb{P}(N \cap [1-r, 1)) = \mathbb{P}(N_r). \qquad (1)$$

We also have the following:

1. Every $x \in \Omega$ belongs to a N_r because if $y \in N$ is an element of the equivalence class of x, then $x \in N_r$ where $r = x - y$ if $x \geq y$ or $r = x - y + 1$ if $x < y$.
2. Every $x \in \Omega$ belongs to exactly one N_r because if $x \in N_r \cap N_s$ for $r \neq s$, then $x - r$ or $x - r + 1$ and $x - s$ or $x - s + 1$ would be distinct elements of N belonging to the same equivalence class, contradicting how we chose N.

- Therefore, Ω is the disjoint union of N_r over all rational $r \in \mathbb{Q} \cap [0,1)$.
- From properties (i) and (ii), we also have $1 = \mathbb{P}(\Omega) = \sum_r \mathbb{P}(N_r)$. But $\mathbb{P}(N_r) = \mathbb{P}(N)$ from equation (1), so the sum is either 0 if $\mathbb{P}(N) = 0$ or ∞ if $\mathbb{P}(N) > 0$, a contradiction.

- This example shows that it is impossible to define a suitable $\mathbb{P}$ for all possible events, some of which are very weird objects.
- E.g., Banach and Tarski (1924) showed that in $\mathbb{R}^n$ where $n \geq 3$, even stranger subsets can be constructed! One can cut up a tennis ball into a finite number of pieces and rearrange them to form a ball the size of Earth.
- The solution that mathematicians have come up with is to restrict to a collection of subsets *and* a $\mathbb{P}$ with "nice" properties, i.e., a σ-algebra and measure, respectively.