I. Background on Probability

A prob. space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where

① Ω is a set called the sample space

② $\mathcal{F}$ is a collection of subsets of Ω satisfying

"Questions we want to ask in our experiment"

(i) $\Omega \in \mathcal{F}$

(ii) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$

(iii) If $A_i \in F$ then $\bigcup_i A_i \in \mathcal{F}$

③ $\mathbb{P} : \mathcal{F} \to [0,1]$

(i) $0 \leq \mathbb{P}(A) \leq 1 \quad \forall \ A \in \mathcal{F}$

(ii) $\mathbb{P}(\Omega) = 1$

(iii) If $A_1, A_2, \ldots$ is a sequence of mutually exclusive events then

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

Mutually Exclusive

$A_i \cap A_j = \emptyset, \ i \neq j$

Specific Cases

① Discrete probability spaces

- When Ω is countable (in particular, finite)

what is the σ-algebra?

$\mathcal{F} = 2^{\Omega}$ is the power set of Ω, i.e. the set of all subsets of Ω

How to build $\mathbb{P}$?

We are going to use a probability mass function (PMF)

$$p : \Omega \to [0,1]$$

s.t.

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

Then the probability measure induced by the pmf p is:

$$\mathbb{P}_p(A) := \sum_{\omega \in A} p(\omega) \qquad A \in \mathcal{F}$$

- For discrete Ω, and given a PMF p, we will always think of the probability space $(\Omega, 2^{\Omega}, \mathbb{P}_p)$

② Continuous Probability Space

When $\Omega = \mathbb{R}^d$, how to build $\mathcal{F}$?

- Optimally, we would like to take $\mathcal{F} = 2^{\mathbb{R}^d}$

! However, it is impossible to define any "natural" probability measure with $\mathcal{F} = 2^{\mathbb{R}^d}$

$$\mu(A) \neq \mu(x + A) \qquad A \subseteq \mathbb{R}, \quad x \in \mathbb{R}$$

$\forall$ sets

violates translation invariant measure

Thus, we have to pick a smaller $\mathcal{F}$, that is "cleverly constructed" to include all interesting subsets of $\mathbb{R}^d$, but exclude pathological ones.

- Borel σ-Algebra
  - When $d=1$, $\Omega = \mathbb{R}$.
  - Given a subset $A \subseteq \Omega$, the σ-Algebra generated by A, denoted by $\sigma(A)$ is

$$\sigma(A) = \{\emptyset, \Omega, A, A^c\}$$

smallest σ-Algebra generated by A

- For $A, B \subseteq \Omega$,

$$\sigma(A,B) = \{\phi, \Omega, A, B, A^c, B^c, A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c, A \cap B \cap A^c, A \cap B \cap B^c, \dots\}$$

- Generally, given a collection $\{A_i\}_{i \in I}$ of subsets $A_i \subseteq \Omega$,

$\sigma(\{A_i\}_{i \in I})$ is the smallest σ-algebra that contains $\{A_i\}_{i \in I}$.

To generate $\sigma(\{A_i\}_{i \in I})$ we collect all the (countable) unions / intersections / complement of the generating set $\{A_i\}_{i \in I}$

For the Borel σ-algebra on $\mathbb{R}$, denoted by $\mathcal{B}(\mathbb{R})$ this is generated by all half-infinite intervals of $\mathbb{R}$.

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(-\infty, a] \mid a \in \mathbb{R}\})$$

In general, for $d \geq 1$

$$\mathcal{B}(\mathbb{R}^d) = \{(-\infty, a_1] \times (-\infty, a_2) \times \dots \times (-\infty, a_d) \mid a_1, \dots, a_d \in \mathbb{R}\}$$

In this class
Whenever $\Omega = \mathbb{R}^d$, we take $\mathcal{F} = \mathcal{B}(\mathbb{R}^d)$

How to build $\mathbb{P}$?

Provided an integrable function $f: \mathbb{R}^d \to \mathbb{R}_{\geq 0}$

s.t.

$$\int_{\mathbb{R}^d} f(x)\,dx = 1 \qquad \left(\begin{array}{l}\text{this function is called a} \\ \text{probability density function}\end{array}\right)$$

We can build a probability measure $\mathbb{P}_f$:

$$\mathbb{P}_f(B) = \int_B f(x)\,dx \qquad B \in \mathcal{B}(\mathbb{R}^d)$$

$\Rightarrow$ The probability space is

$$(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_f)$$

Notation: We will use $\mathcal{P}(\Omega)$ to denote the set of all prob. measures over Ω

- If Ω is countable then $\mathcal{F} = 2^{\Omega}$
- If $\Omega = \mathbb{R}^d$, then $\mathcal{F} = \mathcal{B}(\mathbb{R}^d)$

Properties of probability space

(i) $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) = 0$

(ii) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ Law of Complement Probability

(iii) $\mathbb{P}(A) \leq \mathbb{P}(B)$ for $A \subseteq B$ Monotonicity

(iv)

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_n\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$ Union Bound

(v) (i) increasing Continuity of Probability

Let $A_1 \subseteq A_2 \subseteq \dots$ be increasing events to an event $A \in \mathcal{F}$ $\left(\bigcup_{i=1}^{\infty} A_i = A\right)$

then

$$\lim_{n \to \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$$

(ii) decreasing

let $A_1 \supseteq A_2 \supseteq \dots$ be decreasing events $A \in \mathcal{F}$

$\left(\bigcap_{n=1}^{\infty} A_n = A\right)$

then $$\lim_{n \to \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$$

Conditional Probability

Given $(\Omega, \mathcal{F}, \mathbb{P})$ and $A \in \mathcal{F}$ w/ $\mathbb{P}(A) > 0$, can define

$$\mathbb{P}(\cdot \mid A) : \mathcal{F} \to [0,1] \text{ by}$$

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \quad \forall\, B \in \mathcal{F}$$

One can show that $(\Omega, \mathcal{F}, \mathbb{P}(\cdot \mid A)\}$ is a prob. space

Bayes Theorem

Let $A, B \in \mathcal{F}$, $\mathbb{P}(A), \mathbb{P}(B) > 0$, then

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\, \mathbb{P}(B)}{\mathbb{P}(A)}$$