Lecture 1: Measurable space, measure and probability

Random experiment: uncertainty in outcomes

Ω: sample space or outcome space; a set containing all possible outcomes

Definition 1.1. Let $\mathcal{F}$ be a collection of subsets of a sample space Ω. $\mathcal{F}$ is called a σ-field (or σ-algebra) if and only if it has the following properties.
(i) The empty set $\emptyset \in \mathcal{F}$.
(ii) If $A \in \mathcal{F}$, then the complement $A^c \in \mathcal{F}$.
(iii) If $A_i \in \mathcal{F}$, $i = 1, 2, ...$, then their union $\cup A_i \in \mathcal{F}$.

$\mathcal{F}$ is a set of sets
Two trivial examples: $\mathcal{F}$ contains $\emptyset$ and Ω only and $\mathcal{F}$ contains all subsets of Ω
Why do we need to consider other σ-field?
$\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$, where $A \subset \Omega$

$\mathcal{C}$ = a collection (set) of subsets of Ω
$\sigma(\mathcal{C})$: the smallest σ-field containing $\mathcal{C}$ (called the σ-field generated by $\mathcal{C}$)
$\sigma(\mathcal{C}) = \mathcal{C}$ if $\mathcal{C}$ itself is a σ-field
$\Gamma = \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-field on } \Omega \text{ and } \mathcal{C} \subset \mathcal{F}\}$
$\sigma(\mathcal{C}) = \cap_{\mathcal{F}\in\Gamma}\mathcal{F}$

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

$\mathcal{R}^k$: the k-dimensional Euclidean space ($\mathcal{R}^1 = \mathcal{R}$ is the real line)
$\mathcal{B}^k$: the Borel σ-field on $\mathcal{R}^k$; $\mathcal{B}^k = \sigma(\mathcal{O})$, $\mathcal{O}$ is the collection of all open sets
$C \in \mathcal{B}^k$, $\mathcal{B}_C = \{C \cap B : B \in \mathcal{B}^k\}$ is the Borel σ-field on C

Measure: length, area, volume...

Definition 1.2. Let $(\Omega, \mathcal{F})$ be a measurable space. A set function ν defined on $\mathcal{F}$ is called a *measure* if and only if it has the following properties.
(i) $0 \le \nu(A) \le \infty$ for any $A \in \mathcal{F}$.
(ii) $\nu(\emptyset) = 0$.
(iii) If $A_i \in \mathcal{F}$, $i = 1, 2, ...$, and A_i's are disjoint, i.e., $A_i \cap A_j = \emptyset$ for any $i \neq j$, then

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

$(\Omega, \mathcal{F})$ a measurable space; $(\Omega, \mathcal{F}, \nu)$ a measure space

If $\nu(\Omega) = 1$, then ν is a probability measure (we usually use notation P instead of ν)

A measure ν may take ∞ as its value
(1) For any $x \in \mathcal{R}$, $\infty + x = \infty$, $x\,\infty = \infty$ if $x > 0$, $x\,\infty = -\infty$ if $x < 0$, and $0\,\infty = 0$;
(2) $\infty + \infty = \infty$;
(3) $\infty^a = \infty$ for any $a > 0$;
(4) $\infty - \infty$ or ∞/∞ is not defined

Examples:

$$\nu(A) = \begin{cases} \infty & A \in \mathcal{F}, A \neq \emptyset \\ 0 & A = \emptyset. \end{cases}$$

Counting measure. Let Ω be a sample space, $\mathcal{F}$ the collection of all subsets, and $\nu(A)$ the number of elements in $A \in \mathcal{F}$ ($\nu(A) = \infty$ if A contains infinitely many elements). Then ν is a measure on $\mathcal{F}$ and is called the *counting measure*.
Lebesgue measure. There is a unique measure m on $(\mathcal{R}, \mathcal{B})$ that satisfies $m([a, b]) = b - a$ for every finite interval $[a, b]$, $-\infty < a \leq b < \infty$. This is called the *Lebesgue measure*. If we restrict m to the measurable space $([0, 1], \mathcal{B}_{[0,1]})$, then m is a probability measure.

Proposition 1.1. Let $(\Omega, \mathcal{F}, \nu)$ be a measure space.
(i) (Monotonicity). If $A \subset B$, then $\nu(A) \leq \nu(B)$.
(ii) (Subadditivity). For any sequence $A_1, A_2, \ldots$,

$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \nu(A_i).$$

(iii) (Continuity). If $A_1 \subset A_2 \subset A_3 \subset \cdots$ (or $A_1 \supset A_2 \supset A_3 \supset \cdots$ and $\nu(A_1) < \infty$), then

$$\nu\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} \nu\left(A_n\right),$$

where

$$\lim_{n\to\infty} A_n = \bigcup_{i=1}^{\infty} A_i \quad \left(\text{or } = \bigcap_{i=1}^{\infty} A_i\right).$$

Let P be a probability measure. The *cumulative distribution function* (c.d.f.) of P is defined to be

$$F(x) = P\left((-\infty, x]\right), \quad x \in \mathcal{R}$$

Proposition 1.2. (i) Let F be a c.d.f. on $\mathcal{R}$. Then
(a) $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$;
(b) $F(\infty) = \lim_{x\to\infty} F(x) = 1$;
(c) F is nondecreasing, i.e., $F(x) \leq F(y)$ if $x \leq y$;
(d) F is right continuous, i.e., $\lim_{y\to x, y>x} F(y) = F(x)$.
(ii) Suppose that a real-valued function F on $\mathcal{R}$ satisfies (a)-(d) in part (i). Then F is the c.d.f. of a unique probability measure on $(\mathcal{R}, \mathcal{B})$.