Probability theory is the mathematical study of randomness. Today we will recall some definitions and results from measure theory. Every experiment or observation results in an outcome or sample point. Let \( \Omega \) be the **sample space** consisting of all “outcomes”. We will use \( \omega \) to denote an outcome from the sample space \( \Omega \). Let \( 2^{\Omega} \) denote the power set of \( \Omega \), consisting of all subsets of \( \Omega \). First we define field and \( \sigma \)-field.

### 1.1 Fields and \( \sigma \)-fields

**Definition 1.1.** Let \( \mathcal{F} \) be a collection of subsets of \( \Omega \). \( \mathcal{F} \) is called a **field** or **algebra** if the following holds

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

(ii) \( A \in \mathcal{F} \) implies that \( A^c \in \mathcal{F} \) and

(iii) \( A, B \in \mathcal{F} \) implies that \( A \cup B \in \mathcal{F} \).

**Definition 1.2.** Let \( \mathcal{F} \) be a collection of subsets of \( \Omega \). \( \mathcal{F} \) is called a **\( \sigma \)-field** or **\( \sigma \)-algebra** if the following holds

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

(ii) \( A \in \mathcal{F} \) implies that \( A^c \in \mathcal{F} \) and

(iii) if \( A_1, A_2, \ldots \in \mathcal{F} \) is countable sequence of sets then \( \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \).

Here and in what follows, countable means finite or countably infinite. Clearly every \( \sigma \)-field is a field. Since \((\bigcap_{i=1}^{\infty} A_i)^c = \bigcup_{i=1}^{\infty} A_i^c\), every \( \sigma \)-field is also closed under countable intersection. The tuple \((\Omega, \mathcal{F})\), where \( \mathcal{F} \) is a \( \sigma \)-field is called a **measurable space**. We will use the notation \( \cup \) to **emphasize** disjoint union. Now let us give some examples of field and \( \sigma \)-field.

**Example 1.3.** \((\Omega, \mathcal{F})\) with \( \mathcal{F} = \{\emptyset, \Omega\} \). \( \mathcal{F} \) is the smallest \( \sigma \)-field on \( \Omega \), called the trivial \( \sigma \)-field.

**Example 1.4.** \((\Omega, \mathcal{F})\) with \( \mathcal{F} = 2^{\Omega} \). \( 2^{\Omega} \) is the largest \( \sigma \)-field on \( \Omega \).

**Example 1.5.** Let \( \Omega \) be any sample space. Let \( \mathcal{A} \) be the collection of all countable subsets and co-countable (complement is countable) subsets of \( \Omega \). Check that \((\Omega, \mathcal{A})\) is a \( \sigma \)-field. We will call it the **countable co-countable \( \sigma \)-field** on \( \Omega \).

**Example 1.6.** Let \( \Omega = \{1, 2, \ldots\} \). Let \( \mathcal{A} \) be the collection of all finite subsets and co-finite (complement is finite) subsets of \( \Omega \). Check that \((\Omega, \mathcal{A})\) is a field. Note that \( \{2\}, \{4\}, \{6\}, \ldots \in \mathcal{A} \) but \( \{2, 4, 6, \ldots\} = \bigcup_{i=1}^{\infty} \{2i\} \notin \mathcal{A} \). Thus \( \mathcal{A} \) is not a \( \sigma \)-field.

In general \( \sigma \)-field is a complicated object. But it is easy to construct one.
**Definition 1.7.** Given a collection \( \mathcal{C} \) of subsets of \( \Omega \) (need not be a field) we define \( \sigma(\mathcal{C}) \), the \( \sigma \)-field generated by \( \mathcal{C} \), as the smallest \( \sigma \)-field containing \( \mathcal{C} \). Similarly, we define \( A(\mathcal{C}) \), the field generated by \( \mathcal{C} \), as the smallest field containing \( \mathcal{C} \).

**Lemma 1.8.** For any \( \mathcal{C} \), \( \sigma(\mathcal{C}) \) and \( A(\mathcal{C}) \) exist.

**Proof.** We will only prove the existence of \( \sigma(\mathcal{C}) \). The other part is similar. Let \( \Lambda \) be the collection of all \( \sigma \)-fields containing \( \mathcal{C} \). Note that \( \Lambda \) is non-empty as \( 2^\Omega \), the power-set of \( \Omega \) is in \( \Lambda \). Define

\[ \mathcal{F} = \bigcap_{\mathcal{G} \in \Lambda} \mathcal{G}. \]

Clearly \( \mathcal{C} \subseteq \mathcal{F} \) and \( \mathcal{F} \) is contained in all \( \sigma \)-fields containing \( \mathcal{C} \). We claim that \( \mathcal{F} \) is a \( \sigma \)-field (Verify the axioms of \( \sigma \)-field!). This completes the proof.■

Even if \( \mathcal{C} \) is a “nice” collection of sets, \( \sigma(\mathcal{C}) \) can become quite complicated. We give some examples.

**Example 1.9.** Let \( \Omega \) be a sample space and \( \mathcal{C} \) be the collection of all singleton subsets of \( \Omega \). We claim that \( \sigma(\mathcal{C}) \) is the countable co-countable \( \sigma \)-field on \( \Omega \).

**Definition 1.10.** Let \( \Omega \) be a topological space with \( \mathcal{C} \) being the collection of all open subsets of \( \Omega \). The \( \sigma \)-field \( \mathcal{B} = \sigma(\mathcal{C}) \) is called the **Borel** \( \sigma \)-field on \( \Omega \).

We will denote the Borel \( \sigma \)-field of \( \mathbb{R}^n \) by \( \mathcal{B}^n \). When \( n = 1 \) we will simply use \( \mathcal{B} \).

### 1.2 Measure

**Definition 1.11.** A **measure** is a nonnegative countable additive set function on a measurable space \((\Omega, \mathcal{F})\), i.e., a function \( \mu : \mathcal{F} \to \mathbb{R} \) such that

(i) \( \mu(A) \geq 0 \) for all \( A \in \mathcal{F} \), and

(ii) if \( A_i \in \mathcal{F} \) is a countable sequence of disjoint sets (i.e., \( A_i \cap A_j = \emptyset \) for all \( i \neq j \)), then

\[ \mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i). \]

A **Probability measure** \( P \) is a measure with \( P(\Omega) = 1 \). A **Probability space** is a triple \((\Omega, \mathcal{F}, P)\) where \( \mathcal{F} \) is a \( \sigma \)-field and \( P \) is a probability measure on \((\Omega, \mathcal{F})\).

We define a pre-measure \( \nu \) as a nonnegative countable additive set function on \((\Omega, A)\), where \( A \) is only a field, i.e., a function \( \mu : A \to \mathbb{R} \) such that

(i) \( \nu(A) \geq 0 \) for all \( A \in A \), and

(ii) if \( A_i \in A \) is a countable sequence of disjoint sets (i.e., \( A_i \cap A_j = \emptyset \) for all \( i \neq j \)) such that \( \bigcup_{i=1}^{\infty} A_i \in A \), then

\[ \nu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \nu(A_i). \]
Abusing notation, we will also call a pre-measure as a measure (which will be clear from the context).

It is easy to check that for a measure (or pre-measure) $\mu$ we have $\mu(\emptyset) = 0$ (by condition (ii), taking $A_1 = A_2 = \cdots = \emptyset$). First we mention some properties of a measure (also true for pre-measure).

**Theorem 1.12.** Let $\mu$ be a measure on the measurable space $(\Omega, \mathcal{F})$. Then the following hold.

(a) (Monotonicity) If $A \subseteq B$ then $\mu(A) \leq \mu(B)$.

(b) (Sub-additivity) If $A \in \bigcup_{i=1}^{\infty} A_i$ then $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$.

(c) (Continuity from below) If $A_i \uparrow A$ (i.e., $A_1 \subseteq A_2 \subseteq \cdots$ and $A = \bigcup_{i=1}^{\infty} A_i$) then $\mu(A_i) \uparrow \mu(A)$.

(d) (Continuity from above) If $A_i \downarrow A$ (i.e., $A_1 \supseteq A_2 \supseteq \cdots$ and $A = \bigcap_{i=1}^{\infty} A_i$) then $\mu(A_i) \downarrow \mu(A)$.

Before going to the proof, we explain an important “disjointification” operation for a sequence of sets that will be used repeatedly. Given a sequence of sets $A_1, A_2, A_3, \ldots$ (which can be arbitrary) if we define

$$B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus (A_1 \cup A_2), B_4 = A_4 \setminus (A_1 \cup A_2 \cup A_3), \ldots$$

then $B_i$’s are disjoint and $\bigcup_{i=1}^{m} B_i = \bigcup_{i=1}^{m} A_i$ for all $m \geq 1$. Now we proceed to the proof.

**Proof.** (a) Let $C = B \setminus A$. Clearly, $C \in \mathcal{F}$ and $C \cup A = B$. Thus

$$\mu(B) = \mu(C \cup A) = \mu(C) + \mu(A) \geq \mu(A).$$

(b) We consider the sequence of sets $B_i = A \cap A_i$ and disjointify it to get $C_i$ such that $A = \bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$. Clearly $C_i \subseteq B_i \subseteq A_i$. Thus, using (ii) of the definition of measure and part (a) of this theorem, we get

$$\mu(A) = \mu(\bigcup_{i=1}^{\infty} C_i) = \sum_{i=1}^{\infty} \mu(C_i) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

(c) Using the disjointification operation on the sequence of sets $A_1, A_2, A_3, \ldots$, we get the disjoint sequence $B_1 = A_1, B_i = A_i \setminus A_{i-1}$. We have $A = \bigcup_{i=1}^{\infty} B_i$ and so

$$\mu(A) = \mu(\bigcup_{i=1}^{\infty} B_i) = \lim_{m \to \infty} \sum_{i=1}^{m} \mu(B_i) = \lim_{m \to \infty} \mu(\bigcup_{i=1}^{m} B_i) = \lim_{m \to \infty} \mu(A_m).$$

(d) Note that $A_1 \setminus A_n \uparrow A_1 \setminus A$. Thus part (c) of this theorem implies that $\mu(A_1 \setminus A_n) \uparrow \mu(A_1 \setminus A)$. Now part (a) shows that for $A \subseteq B$, $\mu(B \setminus A) = \mu(B) - \mu(A)$. It follows that $\mu(A_n) \downarrow \mu(A)$. ■

**Definition 1.13.** A measure $\mu$ is **finite** if $\mu(\Omega) < \infty$.

**Definition 1.14.** A measure $\mu$ is **σ-finite** if $\Omega = \bigcup_{i=1}^{\infty} A_i$ with $\mu(A_i) < \infty, \forall i$.

**Example 1.15.** Let $\Omega = \text{set of outcomes if we toss 2 dices} = \{1, 2, \cdots , 6\} \times \{1, 2, \cdots , 6\}$, and $\mathcal{F} = 2^\Omega$, then we can define measure $\mu(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{36}$.

**Example 1.16.** Let $\Omega$ be a countable set, $\mathcal{F} = 2^\Omega$. Take a sequence $P_w \geq 0, w \in \Omega$ s.t. $\sum_{w \in \Omega} P_w = 1$, then $\mu(A) = \sum_{w \in A} P_w$ is a probability measure.
Example 1.17. Let $\Omega = \mathbb{R}$ and $\mathcal{F}$ be a countable co-countable $\sigma$-field. Define

$$
\mu(A) = \begin{cases} 
0, & \text{if } A \text{ is countable} \\
1, & \text{if } A \text{ is co-countable.}
\end{cases}
$$

Proof: Clearly, $\mu(A) \geq 0$ and $\mu(\Omega) = 1$. Let $A = \bigcup_{i=1}^{\infty} A_i$. We consider two separate cases:

**Case 1:** $A$ is countable. Then each $A_i$ is countable and $\mu(A) = 0 = \sum_{i=1}^{\infty} \mu(A_i)$.

**Case 2:** $A$ is co-countable. If sets $X, Y$ are both co-countable, then $X \cap Y \neq \emptyset$. Otherwise, $X^c \cup Y^c = \Omega$ and the left side is countable while the right side is uncountable. Therefore, when $A$ is co-countable, exactly one of $A_i$ is co-countable, then $\mu(A) = 1 = 1 + 0 + \cdots = \sum_{i=1}^{\infty} \mu(A_i)$.

Example 1.18. Let $\Omega = (0,1)$ and $\mathcal{F} = \mathcal{B}(\Omega) = \sigma(\text{open subsets of } \Omega)$. Define $\mu((a,b]) = b-a$, $0 \leq a < b < 1$. How to define $\mu(A)$ in general?

Given a collection of subsets $\mathcal{C}$ one can always construct a field or $\sigma$-field containing $\mathcal{C}$. However, given a nonnegative countable additive set function on $\mathcal{C}$ (think about open sets), it might not be possible to extend it as a measure to $\sigma(\mathcal{C})$. However, this is possible under certain condition. We will mention one such condition next.

Definition 1.19. We define a **semialgebra** $\mathcal{S}$ as a collection of subsets such that

(i) $A, B \in \mathcal{S}$ implies that $A \cap B \in \mathcal{S}$ and

(ii) $A \in \mathcal{S}$ implies that $A^c$ is a finite disjoint union of sets in $\mathcal{S}$.

Note that the collection of sets

$$
\mathcal{S} = \{(a,b] \mid -\infty \leq a \leq b \leq \infty\}
$$

is a semialgebra.