Events and Probabilities

Let $\Omega \neq \phi$ be the set of all possible outcomes of an experiment. Call Ω a sample space.

Definition Let $\Omega \neq \phi$. A collection $\mathcal{A}$ of subsets of Ω is a σ-Algebra (or σ-field) if

(1) $\phi \in \mathcal{A}$

(2) $A \in \mathcal{A} \Rightarrow A^c := \Omega \setminus A \in \mathcal{A}$

(3) $A_k \in \mathcal{A}\ (k=1,2,\cdots) \Longrightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$.

If $\mathcal{A}$ is a σ-algebra of subsets of Ω, then

- $\Omega \in \mathcal{A}$.
- $A_k \in \mathcal{A}\ (k=1,2,\cdots) \Rightarrow \bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$.
- $A_1, \cdots, A_n \in \mathcal{A} \Rightarrow \bigcup_{k=1}^{n} A_k \in \mathcal{A}$, $\bigcap_{k=1}^{n} A_k \in \mathcal{A}$.
- $A, B \in \Omega \Rightarrow A \setminus B \in \mathcal{A}$, $A \triangle B = (A \setminus B) \cup (B \setminus A) \in \mathcal{A}$.

Definition Let $\Omega \neq \phi$ and $\mathcal{A}$ a σ-algebra of subsets of Ω. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{A})$ is a function $\mathbb{P} : \mathcal{A} \to [0,1]$ satisfying

- $\mathbb{P}(\phi) = 0$ and $\mathbb{P}(\Omega) = 1$.
- If $A_k \in \mathcal{A}\ (k=1,2,\cdots)$, $A_j \cap A_k = \phi\ \forall j \neq k$,

  then $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$

Call the triple $(\Omega, \mathcal{A}, \mathbb{P})$ a probability space. If $A \in \mathcal{A}$, call $\mathbb{P}(A)$ the probability of A.

Meanings of Ω, $\mathcal{A}$, and $\mathbb{P}$

$\omega \in \Omega$: an elementary event / an outcome of an experiment

$A \in \mathcal{A}$: event that some outcome in A occurs

ϕ: empty set, an impossible event

Ω: sample space. certain event

$\mathbb{P}(A)$: the probability that the event A occurs

Proposition Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

(1) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A) \quad \forall A \in \mathcal{A}$.

(2) If $A_k \in \mathcal{A}$ $(k=1,\cdots,n)$ and $A_j \cap A_k = \phi$ $(j,k=1,2,\cdots,n, j \neq k)$ then

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

(3) If $A, B \in \mathcal{A}$ and $A \subseteq B$ then

$$\mathbb{P}(A) \leq \mathbb{P}(B).$$

(4) If $A_k \in \mathcal{A}$ $(k=1,2,\cdots)$ ~~are pairwise disjoint~~, then

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

(The right-hand side can be $+\infty$.)

(5) If $A, B \in \mathcal{A}$ then

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

(6) More generally, if $A_k \in \mathcal{A}$ $(k=1,\cdots,n)$. then

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{i} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n+1} \mathbb{P}(A_1 \cap \cdots \cap A_n)$$

(7) If $A_k \in \mathcal{A}$ $(k=1,2,\cdots)$, $A_1 \subseteq A_2 \subseteq \cdots$ Then $\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \mathbb{P}(A_k)$.
If $B_k \in \mathcal{A}$ $(k=1,2,\cdots)$, $B_1 \supseteq B_2 \supseteq \cdots$ Then $\mathbb{P}\left(\bigcap_{k=1}^{\infty} B_k\right) = \lim_{k\to\infty} \mathbb{P}(B_k)$.

Example 1. $\Omega = \{\omega_1, \cdots, \omega_n\}$ — a finite sample space. $\mathcal{A} = \mathcal{P}(\Omega) = \{$all subsets of $\Omega\}$. Let $p_i \in [0,1]$ $(i=1,\cdots,n)$ with $\sum_{i=1}^{n} p_i = 1$. Define $\mathbb{P} : \mathcal{A} \to [0,1]$ by $\mathbb{P}(\phi) = 0$ and

$$\mathbb{P}(\{\omega_{i_1}, \cdots, \omega_{i_k}\}) = \sum_{j=1}^{k} p_{i_j}$$

for any k: $1 \le k \le n$, any $i_1, \cdots, i_k \in \{1, \cdots, n\}$, distinct.

Example 2 Let $\Omega = \mathbb{R}^n$ and $\mathcal{A} = \mathcal{B}(\mathbb{R}^n)$: the Borel σ-algebra of $\mathbb{R}^n$ (i.e., the smallest σ-algebra containing all the open subsets of $\mathbb{R}^n$). Let $x_0 \in \mathbb{R}^n$. Define

$$\mathbb{P}(A) = \begin{cases} 1 & \text{if } x_0 \in A, \\ 0 & \text{if } x_0 \notin A. \end{cases}$$

Then $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P})$ is a probability space. $\mathbb{P}$ is called the Dirac mass at x_0, denoted δ_{x_0}.

Example 3 Let $f \in L^1(\mathbb{R}^n)$ with $f \ge 0$ on $\mathbb{R}^n$ and

$$\int_{\mathbb{R}^n} f(x)\,dx = 1.$$

Define $\mathbb{P}_f : \mathcal{B}(\mathbb{R}^n) \to [0,1]$ by

$$\mathbb{P}(A) = \int_A f(x)\,dx \quad \forall A \in \mathcal{B}(\mathbb{R}^n).$$

Then $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P})$ is a probability space.

We now fix a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition Let $A, B \in \mathcal{A}$. Suppose $\mathbb{P}(B) > 0$. The **conditional probability** that A occurs given that B occurs is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Proposition

(1) If $A, B \in \mathcal{A}$ with $0 < \mathbb{P}(B) < 1$ then

$$\mathbb{P}(A) = \mathbb{P}(A|B)\,\mathbb{P}(B) + \mathbb{P}(A|B^c)\,\mathbb{P}(B^c)$$

(2) Let $B_1, \cdots, B_n \in \mathcal{A}$ be a partition of Ω, i.e., $B_i \cap B_j = \phi \; (\forall i \neq j)$ and $\bigcup_{k=1}^{n} B_k = \Omega$. Suppose all $\mathbb{P}(B_i) > 0$. $(i = 1, \cdots, n)$. Then

$$\mathbb{P}(A) = \sum_{k=1}^{n} \mathbb{P}(A|B_k)\,\mathbb{P}(B_k).$$

Proposition Let $A_1, \cdots, A_n \in \mathcal{A}$ be such that $\mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) > 0$. Then

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2|A_1)\,\mathbb{P}(A_3|A_1 \cap A_2) \cdots \cdot \mathbb{P}(A_n | A_1 \cap \cdots \cap A_{n-1}).$$

A very useful identity:

$$\mathbb{P}(A \cap B | C) = \mathbb{P}(A | B \cap C)\,\mathbb{P}(B|C)$$

where $A, B, C \in \mathcal{A}$ and $\mathbb{P}(B \cap C) > 0$.

<u>Definition</u> Two events $A, B \in \mathcal{A}$ are <u>independent</u>, if
$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B).$$
More generally, a family of events $\{A_i : i \in I\}$ (each $A_i \in \mathcal{A}$, $i \in I$) is <u>independent</u>, if
$$\mathbb{P}\Big(\bigcap_{i \in J} A_i\Big) = \prod_{i \in J} \mathbb{P}(A_i)$$
for any finite subset J of I.

<u>The product probability space</u>

Let $(\Omega_k, \mathcal{A}_k, \mathbb{P}_k)$ $(k = 1, \cdots, n)$ be probability spaces. Let $\Omega = \Omega_1 \times \cdots \times \Omega_n$. Let $\mathcal{A}$ be the smallest σ-algebra of subsets of Ω that contains all the subsets of Ω ~~of~~ in the form
$$A_1 \times A_2 \times \cdots \times A_n,$$
where $A_k \in \mathcal{A}_k$ $(k = 1, \cdots, n)$. Then, there exists a unique probability measure $\mathbb{P} : \Omega \to [0,1]$ such that
$$\mathbb{P}(A_1 \times \cdots \times A_n) = \mathbb{P}_1(A_1) \cdots \mathbb{P}(A_n)$$
for $A_k \in \mathcal{A}_k$ $(k = 1, \cdots, n)$. The product probability space is then defined to be $(\Omega, \mathcal{A}, \mathbb{P})$.