Probability

Probability is the branch of mathematics dealing with numerical description of how likely something is to happen.

In other words, probability is the measure of the likelihood that an event will occur.

Definition 1. (1) *A set E is said to be countable if there exists a bijective map from the set of natural numbers $\mathbb{N}$ to E.*
(2) *A set E is said to be uncountable if it is neither finite nor countable.*

Example 2. (1) *Define $f : \mathbb{N} \longrightarrow \mathbb{N}$ by $f(n) = n$. Clearly f is one -one and onto. Thus, $\mathbb{N}$ is countable.*
(2) *Let $\mathbb{Z}$ denotes the set of integers. Define $f : \mathbb{N} \longrightarrow \mathbb{Z}$ by*

$$f(n) = \begin{cases} \frac{n-1}{2}, & \text{if } n \text{ is odd} \\ -\frac{n}{2}, & \text{if } n \text{ is even.} \end{cases}$$

Clearly f is one-one and onto. Thus, $\mathbb{Z}$ is countable.
(3) *The set of all rational numbers $\mathbb{Q}$ is also countable. Prove!*
(4) *The set of real numbers $\mathbb{R}$ as well as intervals (excluding one point set) in $\mathbb{R}$ are uncountable. Prove!*

Definition 3 (Random experiment). *A random experiment is an experiment in which*

(1) *the set of all possible outcomes of the experiment is known in advance;*
(2) *the outcome of a particular trial of the experiment cannot be predicted in advance;*
(3) *the experiment can be repeated under identical conditions.*

Definition 4 (Sample Space). *The set of all possible outcomes of a random experiment is called the sample space. We will denote the sample space of a random experiment by $\mathcal{S}$. For example:*

(1) *For tossing a fair (unbiased) coin, the sample space $\mathcal{S}$ is $\{H, T\}$, where H means that the outcome of the toss is a head and T means that it is a tail.*
(2) *For rolling a fair die, the sample space $\mathcal{S}$ is $\{1, 2, 3, 4, 5, 6\}$.*
(3) *For simultaneously flipping a coin and rolling a die, the sample space $\mathcal{S}$ is $\{H, T\} \times \{1, 2, 3, 4, 5, 6\}$.*
(4) *For flipping two coins, the sample space $\mathcal{S}$ is $\{(H, H), (H, T), (T, H), (T, T)\}$.*
(5) *For rolling two dice, the sample space $\mathcal{S}$ is $\{(i, j) : i, j \in \{1, 2, 3, 4, 5, 6\}\}$.*

Definition 5 (σ-algebra). *A non-empty collection $\mathcal{F}$ of subsets of $\mathcal{S}$ is called a σ-algebra (or σ-field) if*

(1) $\mathcal{S} \in \mathcal{F}$;
(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;
(3) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Event and Event space: An event is a subset of the sample space $\mathcal{S}$. We say that the event E occurs when the outcome of the random experiment lies in E.

In general, any subset of $\mathcal{S}$ is not necessarily an event, rather an event is a special subset. The event space (set of all events), denoted by Σ, is a subset of the power set of $\mathcal{S}$. An event space must be a σ-algebra.

In the next remark the event space will be fixed for different sample spaces. This will be used throughout the course.

Remark 6. (1) ***If the sample space $\mathcal{S}$ is a finite or a countable set, then we will take $\Sigma = \mathcal{P}(\mathcal{S})$, where $\mathcal{P}(\mathcal{S})$ is the power set of $\mathcal{S}$.***

(2) *Let $\mathbb{B}_{\mathbb{R}}$ denote the set which contains all open intervals, closed intervals, countable unions of open intervals, countable unions of closed intervals, countable intersections of open intervals, and countable intersections of closed intervals.*

If the sample space $\mathcal{S} = \mathbb{R}$, then we will take the event space $\Sigma = \mathbb{B}_{\mathbb{R}}$

(3) *Let I be any interval. Let $\mathbb{B}_I$ denote a set which contains all open intervals contained in I, all closed intervals contained in I, all countable unions of open intervals contained in I, all countable unions of closed intervals contained in I, all countable intersections of open intervals contained in I, and all countable intersections of closed intervals contained in I.*

If the sample space $\mathcal{S} = I$, then we will take the event space $\Sigma = \mathbb{B}_I$.

For any two events E and F, the event $E \cup F$ consists of all outcomes that are either in E or in F, i.e., the event $E \cup F$ will occur if either E or F occurs. The Event $E \cap F$ consists of all outcomes which are both in E and F, i.e., the event $E \cap F$ will occur if both E and F occur.

Mutually exclusive events: Two events E_1 and E_2 are said to be mutually exclusive if they cannot occur simultaneously, i.e., if $E_1 \cap E_2 = \emptyset$.

Similarly, we can define union and intersection of more than two events.

For any event E, the event E^c (complement of E) consists of all outcomes in the sample space S that are not in E, i.e., E^c will occur if E does not occur. For example, let $E = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$, i.e., E is the event that the sum of the dice is equal to seven, then E^c will occur if the sum of the dice is not equal to seven.

Definition 7. *Consider a random experiment with sample space $\mathcal{S}$. For each event E, we assume that a real number $P(E)$ is assigned which satisfies the following three axioms:*

(1) $0 \le P(E) \le 1$ *or simply* $P(E) \ge 0$, *for all events* E*;*
(2) $P(\mathcal{S}) = 1$*;*
(3) *If $E_1, E_2, \ldots$ is a countably infinite collection of mutually exclusive events, that is, $E_i \cap E_j = \emptyset$ for $i \neq j$, then $P(\cup_{i=1}^{\infty} E_i) = \sum\limits_{i=1}^{\infty} P(E_i)$.*

The real number $P(E)$ is known as the probability of the event E.

Remark 8. *It is clear that P is a function from the events space Σ to $[0,1]$ which satisfies the axioms* (1), (2) *and* (3). *We will call P a* ***probability function*** *and the triple $(\mathcal{S}, \Sigma, P)$ is called the probability space.*

Example 9. (1) *Let $\mathcal{S} = \{1, 2, 3, \ldots\}$. Define P on $\mathcal{P}(\mathcal{S})$ as follows:*

$$P(i) = \frac{1}{2^i}, i = 1, 2, \ldots.$$

Then P defines a probability (verify this!).

(2) *Let $\mathcal{S} = (0, \infty)$. Define P on $\mathbb{B}_{(0,\infty)}$ as follows: For each interval $I \subset \mathcal{S}$*

$$P(I) = \int_I e^{-x} dx.$$

Then P defines a probability (verify this!).

(3) *Let $\mathcal{S} = [0, 1]$. Define P on $\mathbb{B}_{[0,1]}$ as follows: For each interval $I \subset \mathcal{S}$*

$$P(I) = \text{length of } I.$$

Then P defines a probability (verify!).

Assigning Probabilities:

(1) Suppose $\mathcal{S}$ is a finite set containing n elements. Then it is sufficient to assign probability to each event containing single element. Thus for any events E, we have $P(E) = \sum_{w \in E} P(w)$. One such assignment is the equally likely assignment or the assignment of uniform probabilities. According to this assignment, $P(w) = \frac{1}{n}$, for every $w \in \mathcal{S}$ and $P(E) = \frac{\text{number of elements in } E}{n}$.

(2) If $\mathcal{S}$ is a countable set, one can not make an equally likely assignment of probabilities. It suffices to make the assignment for each event containing single element. Then for any event E, define $P(E) = \sum_{w \in E} P(w)$.

(3) If $\mathcal{S}$ is an uncountable set, then again one can not make an equally likely assignment of probabilities.

Theorem 10. *Let $(\mathcal{S}, \Sigma, P)$ be a probability space. Then*

(1) $P(\emptyset) = 0$;

(2) *For mutually exclusive events $E_1, E_2, \ldots, E_n$, we have* $P(\cup_{i=1}^n E_i) = \sum\limits_{i=1}^{n} P(E_i)$;

(3) $P(E^c) = 1 - P(E)$;

(4) *For $E_1 \subseteq E_2$, we have $P(E_1) \leq P(E_2)$ and $P(E_2 - E_1) = P(E_2) - P(E_1)$;*

(5) $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

Proof. (1) Let $E_1 = \mathcal{S}$ and $E_i = \emptyset$, $i = 2, 3, \ldots$. Then $P(E_1) = 1, E_1 = \cup_{i=1}^{\infty} E_i$ and $E_i \cap E_j = \emptyset$ for $i \neq j$. Therefore

$$\begin{aligned} 1 = P(E_1) &= P(\cup_{i=1}^{\infty} E_i) \\ &= \sum_{i=1}^{\infty} P(E_i) \\ &= 1 + \sum_{i=2}^{\infty} P(\emptyset) \end{aligned}$$

$$\Rightarrow \sum_{i=2}^{\infty} P(\emptyset) = 0$$

This shows that the constant series $\sum\limits_{i=2}^{\infty} P(\emptyset)$ converges to 0. Hence, $P(\emptyset) = 0$, otherwise the constant series $\sum\limits_{i=2}^{\infty} P(\emptyset)$ can not be convergent.

(2) Let $E_i = \emptyset$, $i = n+1, n+2, \ldots$. Then $E_i \cap E_j = \emptyset$ for $i \neq j$ and $P(E_i) = 0$, $i = n+1, n+2, \ldots$. Therefore, $P(\cup_{i=1}^n E_i) = P(\cup_{i=1}^{\infty} E_i) = \sum\limits_{i=1}^{\infty} P(E_i) = \sum\limits_{i=1}^{n} P(E_i)$ (since $P(E_i) = 0$, $i = n+1, n+2, \ldots$).

(3) Since $E \cup E^c = \mathcal{S}$ and $E \cap E^c = \emptyset$, $1 = P(E \cup E^c)$. Thus $1 = P(E) + P(E^c)$ (by using (2)). Hence $P(E^c) = 1 - P(E)$.

(4) Since $E_2 = E_1 \cup (E_2 - E_1)$ and $E_1 \cap (E_2 - E_1) = \emptyset$, $P(E_2) = P(E_1 \cup (E_2 - E_1)) = P(E_1) + P(E_2 - E_1)$. This implies that $P(E_2 - E_1) = P(E_2) - P(E_1)$.

(5) Since $E_1 \cup E_2 = E_1 \cup (E_2 - E_1)$ and $E_1 \cap (E_2 - E_1) = \emptyset$, $P(E_1 \cup E_2) = P(E_1 \cup (E_2 - E_1)) = P(E_1) + P(E_2 - E_1)$ $\ldots(i)$
Also since $E_2 = (E_1 \cap E_2) \cup (E_2 - E_1)$ and $(E_1 \cap E_2) \cap (E_2 - E_1) = \emptyset$, $P(E_2) = P(E_1 \cap E_2) + P(E_2 - E_1) \Rightarrow P(E_2 - E_1) = P(E_2) - P(E_1 \cap E_2)$ $\ldots(ii)$
Thus, by equation (i) and (ii), we have

$$\boxed{P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)}$$

□

Inclusion-exclusion identity: For events E_1, E_2 and E_3 we have

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3) &= P((E_1 \cup E_2) \cup E_3) = P(E_1 \cup E_2) + P(E_3) - P((E_1 \cup E_2) \cap E_3) \\ &= P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P((E_1 \cap E_3) \cup (E_2 \cap E_3)) \\ &= P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_1 \cap E_3) - P(E_2 \cap E_3) \\ &+ P(E_1 \cap E_2 \cap E_3) \end{aligned}$$

Inductively, for any n events $E_1, E_2, \ldots, E_n$, we have
$P(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum\limits_{i=1}^{n} P(E_i) - \sum\limits_{i<j} P(E_i \cap E_j) + \sum\limits_{i<j<k}^{n} P(E_i \cap E_j \cap E_k) - \cdots + (-1)^{n+1} P(E_1 \cap E_2 \cap \cdots \cap E_n)$.

This identity is known as the inclusion-exclusion identity.

Exhaustive events: The countable collection $\{E_i \mid i \in \wedge\}$ of events is said to be exhaustive if $P(\cup_{i \in \wedge} E_i) = 1$, where $\wedge$ is an index set.

Definition 11. *Let $(\mathcal{S}, \Sigma, P)$ be a probability space and (E_n) be a sequence of events in Σ.*

(1) *We say that sequence (E_n) is increasing (written as $E_n \uparrow$) if $E_n \subseteq E_{n+1}$, $n = 1, 2, \ldots$;*
(2) *We say that sequence (E_n) is decreasing (written as $E_n \downarrow$) if $E_{n+1} \subseteq E_n$, $n = 1, 2, \ldots$;*
(3) *We say that the sequence (E_n) is monotone if either $E_n \uparrow$ or $E_n \downarrow$.*
(4) *If $E_n \uparrow$, we define $Lim_{n\to\infty} E_n = \cup_{n=1}^{\infty} E_n$*
(5) *If $E_n \downarrow$, we define $Lim_{n\to\infty} E_n = \cap_{n=1}^{\infty} E_n$*

Theorem 12. *(**Continuity of Probability**) Let (A_n) be a sequence of monotone events. Then*

$$P(Lim_{n\to\infty} E_n) = \lim_{n\to\infty} P(E_n),$$

where $\lim\limits_{n\to\infty} P(E_n)$ denotes the limit of the real sequence $(P(E_n))$.

Proof. Exercise □