Probability Course

Chapter 1

Basic Notions for Probability Calculation

1.1 Introduction

Probability calculation is a branch of mathematics aimed at studying random phenomena, i.e., experiments whose outcome cannot be predicted with certainty. For example, if the experiment is repeated several times, different results may be obtained. A form of indeterminacy appears in the outcome of the experiment. This form of randomness can be interpreted as our inability to conceive, explain, and use the considered physical phenomena or as a lack of information about the conditions of the experiment. Some phenomena may be inherently subject to randomness.

Examples of such experiments are numerous: one can think of the game of heads or tails, rolling a die, the sex of an unborn child, the waiting time of a customer at the post office, the lifespan of a radioactive particle.

To study these phenomena, we must first create a mathematical model. The model chosen to describe random experiments is that of a triplet commonly denoted $(\Omega, \mathcal{F}, \mathbb{P})$ and called a probability space.

1.2 Axioms of Probability Calculation

1.2.1 Notion of Random Experiment

When studying a random phenomenon, it is possible to assimilate it to a random experiment, i.e., if the same experiment is repeated several times under well-determined conditions, the result of this experiment varies and seems to obey stochastic considerations, in other words, chance. In this case, we say that we are dealing with a random experiment, usually denoted by the letter ζ.

Examples

1.a) Dice Game:

Two dice of different colors, with six faces numbered from 1 to 6, are thrown onto a flat surface. At the end of the throw, the two numbers appearing on the top faces of the dice are recorded. These two numbers represent the result of the experiment.

1.b) Heads or Tails Game:

Two coins are thrown onto a flat surface, and what is seen on the top face of each coin is recorded. If we denote F for heads and P for tails, then the result of the experiment is expressed by the two letters **F, P**.

1.c) Drawing Without Replacement:

An urn contains N balls, of which r are red and n are black. The experiment consists of drawing two balls successively from the urn without replacing the first ball drawn. The result of the experiment is the color of the ball obtained in the first draw and that obtained in the second, in that order.

1.d) Height Measurement:

The height of an individual chosen at random from a well-defined statistical population is recorded. The result of the experiment is a positive real number (after choosing a unit of measurement).

Fundamental Set Ω

A fundamental set Ω is associated with the random experiment ζ, whose elements represent all possible outcomes of the experiment ζ.

Examples

1.a) $\Omega = \{(i, j); i$ and j are integers between 1 and 6$\}$.

1.b) $\Omega = \{PP, FF, PF, FP\}$.

1.c) $\Omega = \{(a, b); a$ represents the color of the ball obtained in the first draw and b that of the second draw$\}$.

1.d) $\Omega = [0, +\infty)$. We can choose a closed interval of $\mathbb{R}$.

Remark 1.2.1. The choice of the fundamental set Ω depends on the case studied. For example, in the dice game, we can assume that each die can land on any of its edges. In this case, we will have additional possible outcomes.

Notion of Events

Consider a property Δ related to the outcome (result of the random experiment ζ). Each time ζ is performed, there are two cases: the property Δ is realized or it is not. Thus, thanks to this property, the fundamental set Ω is divided into two disjoint parts: on one side, the set E formed by the set of points $\omega, \omega \in \Omega$, representing results of ζ with the realization of the property Δ, and on the other side, the complementary set E in Ω containing the points ω of Ω that correspond to results of ζ that do not realize the property Δ. We say that the set E is the event related to the property Δ. We also say that E is the event "E is realized". We immediately see that E (the complement of E in Ω) is also an event, referred to as "E is not realized".

1.3 Operations on Events

We extend the parallelism between set notions and probabilistic notions. In addition to the complementarity operation defined earlier, there are other important operations on the fundamental set Ω related to the same experiment ζ.

a) Inclusion

Let A and B be two events of Ω. If the experiment that realizes event A necessarily realizes event B, we say that A implies B and write $A \subset B$.

b) Intersection

Let A and B be two subsets representing two events of Ω. The event "A and B" is expressed by the intersection of the subsets A and B of Ω, called the intersection event of A and B, and we write: "A and B" $= A \cap B$.

Definition 1. If $A \cap B = \emptyset$, we say that events A and B are incompatible, i.e., their simultaneous realization is impossible.

c) Union

If A and B are two events of Ω, then the event "A or B" is the event represented by the union of the two subsets A and B of Ω. We write "A or B" $= A \cup B$.

Examples

1.4.a) Let's revisit example 1.a) on the dice game and consider the following events:

- $A =$ "the two recorded numbers are odd"
- $B =$ "the two recorded numbers are even"
- $C =$ "one of the recorded numbers is even and the other is odd"

Then we have:

- $\overline{A} =$ "at least one of the two recorded numbers is even"
- $\overline{B} =$ "at least one of the two numbers is odd"
- $A \cap B = \emptyset$
- $A \cup B =$ "the two recorded numbers are either even or odd"
- $C = \overline{A \cup B} =$ "one of the two numbers is even and the other is odd"

- $\overline{A} = B \cup C$
- $\overline{B} = A \cup C$
- $B \subset \overline{A}$
- $A \subset \overline{B}$
- $\overline{A} \cap \overline{B} = C$

1.4.b) In example 1.b) (Heads or Tails game with 2 coins), the elementary event $\{PP\}$ is the intersection of event $A = \{PF, PP\}$ with event $B = \{PP, FP\}$. The event $A \cup B$ is: "heads at least once", $\overline{A \cup B}$ = "tails twice" = $\{FF\}$.

We will call certain events subsets of Ω, but not always all of them. However, if A and B are two "interesting events", then we observe that $\overline{A}$ and $A \cup B$ are also interesting. We will therefore define spaces of events that are stable under complementation or countable union.

Definition 2. Let Ω be a non-empty set. We call a σ-algebra on Ω (or σ-field) any set of subsets $\mathcal{F}$ of Ω satisfying the following three conditions:

1. $\Omega \in \mathcal{F}$
2. $\forall A, A \in \mathcal{F} \Rightarrow \overline{A} \in \mathcal{F}$ (We say that $\mathcal{F}$ is stable under complementation)
3. For any sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$ (We say that $\mathcal{F}$ is stable under countable union).

Examples

Let Ω be a non-empty set. $\mathcal{F}_1 = \{\emptyset, \Omega\}$ is a σ-algebra called the trivial σ-algebra; $\mathcal{F}_2 = \mathcal{P}(\Omega)$, the collection of all subsets of Ω, is a σ-algebra; finally, if $A \subset \Omega$, $\mathcal{F}_3 = \{A, \overline{A}, \emptyset, \Omega\}$ is a σ-algebra on Ω.

Remark 1.3.1. If $\mathcal{F}$ is a σ-algebra on Ω, then $\emptyset \in \mathcal{F}$. Moreover, $\mathcal{F}$ is stable under finite union, finite intersection, countable intersection, difference, etc. For example, let's explain why $\mathcal{F}$ is stable under countable intersection. Let $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, we have $\bigcap_{n\in\mathbb{N}} A_n = \overline{\bigcup_{n\in\mathbb{N}} \overline{A_n}}$. Since, for all n, $A_n \in \mathcal{F}$, condition (2) of the definition implies that $\overline{A_n} \in \mathcal{F}$ for all n. Condition (3) then implies that $\bigcup_{n\in\mathbb{N}} \overline{A_n} \in \mathcal{F}$; we can then apply point (2) again to obtain that $\overline{\bigcup_{n\in\mathbb{N}} \overline{A_n}} \in \mathcal{F}$.

Definition 3. Let Ω be the fundamental set associated with the random experiment ζ and $\mathcal{F}$ a σ-algebra on Ω. Then the pair $(\Omega, \mathcal{F})$ is called a **Probabilizable Space**. The elements of $\mathcal{F}$ are called events, and those of Ω are called elementary events.

Examples

In examples 1.a), 1.b), and 1.c), we can take $\mathcal{F} = \mathcal{P}(\Omega)$ and directly obtain the probabilizable space $(\Omega, \mathcal{P}(\Omega))$.

Remark 1.3.2. We would always like to work with the σ-algebra $\mathcal{F} = \mathcal{P}(\Omega)$. However, this is not possible unless Ω is a finite or countable set. If $\Omega = \mathbb{R}$, it is not possible to measure all subsets of $\mathbb{R}$ without leading to a contradiction, and we limit ourselves to a class of subsets called the Borel σ-algebra.

Definition 4. We call the Borel σ-algebra on $\mathbb{R}$, denoted $\mathbb{B}(\mathbb{R})$, the smallest σ-algebra, in the sense of inclusion, containing all intervals of $\mathbb{R}$.

This definition requires a comment: it is not obvious a priori that we can speak of "the smallest σ-algebra" containing the intervals. However, the definition makes sense because the intersection of any family of σ-algebras on Ω is still a σ-algebra on Ω. We can therefore consider the intersection of all σ-algebras on $\mathbb{R}$ containing the intervals, which by construction becomes "the smallest σ-algebra".

Definition 5. We call a complete system of events any sequence $E_1, E_2, ..., E_n$ of events of $\mathcal{F}$ that are pairwise incompatible and such that $\bigcup_{i=1}^{n} E_i = \Omega$. The sets $(E_i)_{1\leq i\leq n}$ form a partition of Ω.

1.4 Notion of Probability

Definition 6. Let $(\Omega, \mathcal{F})$ be a probabilizable space and $\mathbb{P}$ an application from $\mathcal{F}$ to $[0, 1]$. $\mathbb{P}$ is a probability on $(\Omega, \mathcal{F})$ if it satisfies the following conditions:

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

2. For any sequence $(A_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{F}$ (events) that are pairwise incompatible, i.e., $A_n \cap A_m = \emptyset$ for $n \neq m$, we have:
$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$$
This property is called σ-additivity.

Definition 7. Let $(\Omega, \mathcal{F})$ be a probabilizable space and $\mathbb{P}$ a probability on this space. The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space or space of probabilities.

Remark 1.4.1. For any random experiment described by a probabilizable space $(\Omega, \mathcal{F})$, there are many possible probabilities $\mathbb{P}$ that can be defined on this space. However, in practice, the choice of the probability $\mathbb{P}$ is determined either by natural conditions or by experimental considerations.

Example 1.4.1. A die is rolled on a flat surface, and the result of interest is the number indicated by the die. The fundamental set will typically be $\Omega = \{1, 2, 3, 4, 5, 6\}$. We use as the σ-algebra the set $\mathcal{P}(\Omega)$, and in the case of a fair die, we define the probability law $\mathbb{P}$ such that each singleton has the same probability, and it is easy to see that $\mathbb{P}(\{i\}) = \frac{1}{6}$ for all $i \in \Omega$. It is straightforward to verify that this completely determines $\mathbb{P}$, and we then have, for any event $A \in \mathcal{P}(\Omega)$, $\mathbb{P}(A) = \frac{\mathrm{Card}(A)}{6}$.

The definition of a probability leads to the following properties:

Proposition 1.4.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If A, B are two events of $\mathcal{F}$, then:

1. $\mathbb{P}(\emptyset) = 0$
2. Additivity property: If $A \cap B = \emptyset$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
3. $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$
4. Monotonicity: If $A \subset B$, $\mathbb{P}(A) \leq \mathbb{P}(B)$
5. If $A \subset B$, $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$; in particular: $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
6. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Proof.

1. To obtain $\mathbb{P}(\emptyset) = 0$, it suffices to take $A_0 = \Omega$ and $A_n = \emptyset$ for $n \geq 1$ in the definition and apply σ-additivity.
2. Here, it also suffices to take $A_0 = A$, $A_1 = B$, and $A_n = \emptyset$ for $n \geq 2$ in the definition and apply σ-additivity.
3. We write $B = (B \setminus A) \cup (A \cap B)$ and use additivity: $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$.
4. and 5. For these two points, we apply point 3 for $A \subset B$.
5. We can write $A \cup B = A \cup B \setminus A$ and use additivity and point 3.

1.5 Some Important Probability Spaces

1.5.1 The Space Ω is Finite or Countable

As we have seen earlier, in this case, we usually assume that the σ-algebra of events $\mathcal{F}$ is $\mathcal{P}(\Omega)$, the set of all subsets of Ω, and thus the probabilizable space will be $(\Omega, \mathcal{P}(\Omega))$. The probabilities that can be defined on this space are then described by the following result:

Proposition 1.5.1. Let Ω be a finite or countable set. Let $\omega \mapsto p_\omega$ be an application from Ω to $\mathbb{R}_+$ such that:
$$\sum_{\omega\in\Omega} p_\omega = 1$$
For any $A \subset \Omega$, we then define:
$$\mathbb{P}(A) = \sum_{\omega\in A} p_\omega$$
Then $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ is a probability space (i.e., the application $\mathbb{P}$ defined by the last formula is a probability on $(\Omega, \mathcal{P}(\Omega))$). Conversely, any probability $\mathbb{P}$ on $(\Omega, \mathcal{P}(\Omega))$ is of the previous type, with $p_\omega = \mathbb{P}(\{\omega\})$.

The Equiprobable Case

Consider the particular case of Proposition 1.5.1 where Ω is finite with $\text{Card}(\Omega) = n$ and the elementary events are equiprobable (i.e., they all have the same probability). In this case, it is easy to determine the probability $\mathbb{P}$. Indeed, since all $p_\omega = \mathbb{P}(\{\omega\})$ are equal, it is easy to see that:

$$\forall \omega \in \Omega,\ p_\omega = \frac{1}{\text{Card}(\Omega)} = \frac{1}{n}$$

Then, in this case, if $A \subset \Omega$, we have:

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = \frac{\text{Number of favorable cases}}{\text{Total number of cases}}$$

Thus, in this case, probability calculations reduce to counting problems. This probability $\mathbb{P}$ is called the uniform probability on Ω.

Example 1.5.1. Consider the random experiment $\zeta =$ "a die is rolled twice". Then $\Omega = \{(i,j); 1 \leq i \leq 6 \text{ and } 1 \leq j \leq 6\}$. If the die is fair, then it is natural to say that the elementary events are equiprobable, i.e., $P(\{(i,j)\}) = \frac{1}{36}$ for all $(i,j) \in \Omega$. If A is the event "the sum of the recorded numbers is equal to 6", then $P(A) = \sum_{(i,j)\in A} P(\{(i,j)\}) = \frac{5}{36}$, since here $A = \{(1,5),(5,1),(2,4),(4,2),(3,3)\}$.

Remark 1.5.1. Attention: an impossible event has probability 0, a certain event has probability 1. The converse is not true.

To show this more clearly, consider the experiment $\zeta =$ "we search for a hidden coin in the hands of several people and stop when it is found". Then Ω consists of all sequences of the form $(EEE...ER)$ where E denotes a failure in the attempt to find the coin and R denotes success. In the probabilizable space $(\Omega, \mathcal{P}(\Omega))$, we define: For any $\omega \in \Omega$ such that $\omega = \underbrace{EE...ER}_{n+1}$, $\mathbb{P}(\{\omega\}) = q^n p$, $n \in \mathbb{N}$ with $q = \mathbb{P}(E)$ and $p = \mathbb{P}(R)$; $q = 1 - p$ ($0 < p < 1$ is the probability of finding the coin in each attempt). It is easy to see that: $\sum_{n=0}^{+\infty} q^n p = p\frac{1}{1-q} = 1$. If A represents the event "the coin is found at the third attempt at least", then $\mathbb{P}(A) = \sum_{n=2}^{+\infty} q^n p = q^2$. We note that the event $(EE...E...E......)$ has a probability of 0 with respect to the defined probability, yet the possibility that this event occurs exists since it is part of the set $\mathcal{P}(\Omega)$ (but this event is experimentally unrealizable).

1.6 Conditional Probabilities and Independence

1.6.1 Conditional Probabilities

The notion of conditional probability allows us to take into account the information we have to update the probability we assign to an event. In other words, we are in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let A and B be two events, each with a non-zero probability of occurring. We are interested in the new probability of one of them if we are assured that the other has occurred.

Example 1.6.1. Consider the random experiment $\zeta =$ "a fair die is rolled once". Then $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, and $\mathbb{P}$ is the uniform probability. Let A: "the face is even" $= \{2,4,6\}$ and B: "the face is less than or equal to 3" $= \{1,2,3\}$. We know that $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$. But if we are assured of the occurrence of A (i.e., the face is even), the probability of B seems to drop to $\frac{1}{3}$.

Now, we formally define this concept to make it easily usable in calculations in a safe and efficient manner.

Definition 11. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and B an event with a non-zero probability, and let A be an event. We call the conditional probability of A given B, denoted $\mathbb{P}(A|B)$ or $\mathbb{P}_B(A)$, the number:

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

Example 1.6.2. Mr. and Mrs. H have two children, one of whom is a girl. What is the probability that the other is a boy? Same question if we assume that the girl is the eldest? To solve this problem rigorously, we need to construct an appropriate probability space. We assume that a newborn is a boy or a girl with equal probability and independently of the sex of their siblings. Thus, a family of two children is in one of the four configurations of $\Omega = \{FG, GF, FF, GG\}$, each with a probability of 1/4. In the first case, we assume that the event $A_1 = \{FG, GF, FF\}$ is realized, and we seek the probability that there is a boy (and a girl) in the family, i.e., the conditional probability $\mathbb{P}(B_1|A_1)$ with $B_1 = \{FG, GF\}$. We obtain:

$$\mathbb{P}(B_1|A_1) = \frac{\mathbb{P}(B_1 \cap A_1)}{\mathbb{P}(A_1)} = \frac{\mathbb{P}(\{FG, GF\})}{\mathbb{P}\{FG, GF, FF\}} = \frac{2/4}{3/4} = \frac{2}{3}$$

In the second case, we seek $\mathbb{P}(B_2|A_2)$ where $A_2 = \{FG, FF\}$ and $B_2 = \{FG\}$. We obtain $\mathbb{P}(B_2|A_2) = \frac{1}{2}$.

Proposition 1.6.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and B an event with a non-zero probability. Then the application:

$$\mathbb{P}(\cdot|B) : \mathcal{F} \to \mathbb{R}_+$$
$$A \mapsto \mathbb{P}(A|B)$$

is a probability on $\mathcal{F}$.

Proof.

1. We have $\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$.

2. Let $(A_n)_{n \in \mathbb{N}}$ be a sequence of events of $\mathcal{F}$ that are pairwise incompatible, i.e., $A_n \cap A_m = \emptyset$ for $n \neq m$. We have:

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n | B\right) = \frac{\mathbb{P}\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} (A_n \cap B)\right)}{\mathbb{P}(B)}$$

The events $(A_n \cap B)_{n \in \mathbb{N}}$ are themselves pairwise incompatible, and thus:

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n | B\right) = \frac{\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} (A_n \cap B)\right)}{\mathbb{P}(B)} = \sum_{n \in \mathbb{N}} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n|B)$$

Therefore, $\mathbb{P}(\cdot|B)$ is a probability on $\mathcal{F}$.

Proposition 1.6.2. (Formula of Compound Probabilities)
If A and B are two events with non-zero probabilities, then:

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

Generalization: Let $A_1, A_2, ..., A_n$ be n events with strictly positive probabilities. Then:

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

Proof. Easy.

Proposition 1.6.3. (Formula of Total Probabilities)
Let B be an event such that $\mathbb{P}(B) > 0$ and $\mathbb{P}(\overline{B}) > 0$. Then:

$$\forall A \in \mathcal{F} : \mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\overline{B})\mathbb{P}(\overline{B})$$

Generalization: Let $B_1, B_2, ..., B_n$ be a complete system of events for Ω such that $\mathbb{P}(B_j) > 0$ for $j = 1, 2, ..., n$. Then:

$$\forall A \in \mathcal{F} : \mathbb{P}(A) = \sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)$$

Proof. We have $\forall A \in \mathcal{F} : A = A \cap \Omega$ and on the other hand $\Omega = \bigcup_{j=1}^{n} B_j$, so:

$$\mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \mathbb{P}\left(A \cap \left(\bigcup_{j=1}^{n} B_j\right)\right) = \mathbb{P}\left(\bigcup_{j=1}^{n} (A \cap B_j)\right)$$

The events $(A \cap B_j)_{j=1,n}$ are pairwise incompatible, so:

$$\mathbb{P}(A) = \sum_{j=1}^{n} \mathbb{P}(A \cap B_j) = \sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)$$

Stochastic Independence of Events

Independence of Two Events:

Idea: If A and B are two events, we say they are independent if the knowledge of the occurrence of one does not change the probability of the occurrence of the other, i.e., $\mathbb{P}(A|B) = \mathbb{P}(A)$ or $\mathbb{P}(B|A) = \mathbb{P}(B)$. To define this notion even for events with zero probability, we have the following definition:

Definition 12. Two events A and B are said to be independent if and only if:

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

We often obtain independent events when we repeat an experiment without the first experiment interfering with the second. For example, this is the case when we play heads or tails twice. Here is another example.

Example 1.6.4. If we roll a fair die twice, then $\Omega = E \times E$ with $E = \{1, 2, 3, 4, 5, 6\}$. Since the die is fair, all elementary events have the same probability. Let's define the following events:

- $A = \{(i, j), i \text{ is even}\}$
- $B = \{(i, j), j \text{ is even}\}$
- $C = \{(i, j), i + j \text{ is even}\}$

It is clear that $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$, and $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A) = \mathbb{P}(A \cap C) = \mathbb{P}(C \cap A) = \mathbb{P}(B \cap C) = \mathbb{P}(C \cap B) = \frac{1}{4}$. Therefore, the events A, B, and C are pairwise independent.

Remark 1.6.1. (Attention): Do not confuse incompatibility with independence of two events. If A and B are two incompatible events, then they are independent only if $\mathbb{P}(A) = 0$ or $\mathbb{P}(B) = 0$. When $\mathbb{P}(A) \neq 0$ and $\mathbb{P}(B) \neq 0$, they are always dependent.

Proposition 1.6.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and A and B two events of $\mathcal{F}$. Then we have:

$$A \text{ and } B \text{ independent} \Leftrightarrow A \text{ and } \overline{B} \text{ independent} \Leftrightarrow \overline{A} \text{ and } B \text{ independent} \Leftrightarrow \overline{A} \text{ and } \overline{B} \text{ independent}$$

Proof. We only prove: A and B independent $\Leftrightarrow$ A and $\overline{B}$ independent (the others are similar). We have $\overline{B} = (A \cap \overline{B}) \cup (\overline{A} \cap \overline{B})$, so:

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

(since $(A \cap \overline{B})$ and $(\overline{A} \cap \overline{B})$ are incompatible). Then:

$$A \text{ and } B \text{ independent} \Leftrightarrow \mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B) \Leftrightarrow \mathbb{P}(A\cap\overline{B}) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(B)(1 - \mathbb{P}(A)) = \mathbb{P}(B)\mathbb{P}(\overline{A}) \Leftrightarrow A \text{ and } \overline{B}$$

Independence of a Family of Events

We are in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $(A_i)_{i\in I}$ be a finite or infinite family of events of $\mathcal{F}$. Saying that these events are pairwise independent is equivalent to writing: $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ for all $i, j \in I$. However, there is also the possibility that $(A_i)_{i\in I}$ are independent (we also say mutually independent), which is given by the following definition:

Definition 13. We say that the events $(A_i)_{i\in I}$ are mutually independent if for any subset $J \subseteq I$, we have:

$$\mathbb{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} \mathbb{P}(A_i)$$

This condition is stricter than pairwise independence of events. Pairwise independence corresponds to the restriction of this condition to subsets J of cardinality 2. The trap is the following: if A, B, C are three independent events, then (A, B), (A, C), and (B, C) are pairs of independent events, but the converse is false. That is, mutual independence of several events implies their pairwise independence, but the converse is false, as shown by the following example:

Example 1.6.5. Consider the random experiment ζ = "a die is rolled twice" and the following events:

- A_1 = "the first roll is 1"

- $A_2 =$ "the second roll is 1"
- $A_3 =$ "the same number appears in both rolls"

It is clear that the events A_1, A_2, and A_3 are pairwise independent. Moreover, $P(A_1) = P(A_2) = P(A_3) = \frac{1}{6}$, and $A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3$ corresponds to "obtaining the number 1 in both rolls." Thus,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2) = \frac{1}{36}$$

which is different from

$$P(A_1)P(A_2)P(A_3) = \frac{1}{216}$$

Therefore, the events A_1, A_2, and A_3 are pairwise independent but not independent as a whole.

Proposition

If $(A_j)_{j\in I}$ is a family of independent events and J is a finite subset of I, then for any $i \notin J$, we have:

$$P(A_i \mid \bigcap_{j\in J} A_j) = P(A_i)$$

In other words, knowledge of the occurrence of the events $(A_j)_{j\in J}$ has no influence on the probability of occurrence of the event A_i for $i \notin I$. This equality constitutes a second criterion of independence and gives its meaning.

Remark

In an independent family of events $(A_j)_{j\in J}$, where $J \subset I$ and J is finite, we can replace some events A_j with their complements $\bar{A}_j$ without losing the property of independence.

Proof

The proof follows directly from the definition of independence of events.