Chapter 1

Independence

We enter the realm of probability theory at this point, where we define independence of events and random variables. In the following, $(\Omega, \mathcal{A}, \mathbf{P})$ is a probability space and the sets $A \in \mathcal{A}$ are the events. As soon as constructing probability spaces has become routine, the concrete probability space will lose its importance and it will be only the random variables that will interest us.

1.1 Independence of Events

If there is partial information on the outcome of a random experiment, the probabilities for the possible events may change.

Example 1.1 (Rolling a die once)**.** We throw a die and consider the events

$$\begin{aligned} A &:= \{\text{the face shows an odd number}\}. \\ B &:= \{\text{the face shows three or smaller}\}. \end{aligned}$$

We model the experiment with $\Omega = \{1, \ldots, 6\}$, $\mathcal{A} = 2^\Omega$ and $\mathbf{P} = \mathcal{U}_\Omega$. Clearly, $\mathbf{P}[A] = \frac{1}{2}$ and $\mathbf{P}[B] = \frac{1}{2}$. However, what is the probability that A occurs if we already know that B occurs? Then

$$A = \{1, 3, 5\} \quad \text{and} \quad B = \{1, 2, 3\},$$

and intuitively

$$\mathbf{P}[A \,|\, B] := \frac{\#(A \cap B)}{\#(B)} = \frac{2}{3}.$$

$\square$

Motivated by this example, we make the following definition.

Definition 1.2 (Conditional probability)**.** Let B be an event. We define the **conditional probability given** B for any $A \in \mathcal{A}$ by

$$\mathbf{P}[A \,|\, B] = \begin{cases} \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} & \text{if } \mathbf{P}[B] > 0, \\ 0 & \text{otherwise.} \end{cases}$$

□

Property 1.3. Let $B \in \mathcal{A}$ with $\mathbf{P}[B] > 0$. Then $\mathbf{P}[\,\cdot \,|\, B]$ is a probability measure on $(\Omega, \mathcal{A})$.

Proof. Exercise. □

We consider two events A and B as (stochastically) independent if the occurrence of B does not change the probability that A occurs, i.e.,

$$\mathbf{P}[B] > 0 \quad \Rightarrow \quad \mathbf{P}[A \,|\, B] = \mathbf{P}[A].$$

Thanks to the following property we can define (stochastic) independence of two events in different ways.

Property 1.4. Let A and B be two events. Then the following propositions are equivalent:

(i) $\mathbf{P}[A \cap B] = \mathbf{P}[A] \cdot \mathbf{P}[B]$;

(ii) $\mathbf{P}[B] > 0 \quad \Rightarrow \quad \mathbf{P}[A \,|\, B] = \mathbf{P}[A]$;

(iii) $\mathbf{P}[A] > 0 \quad \Rightarrow \quad \mathbf{P}[B \,|\, A] = \mathbf{P}[B]$.

Proof. Exercise. □

Example 1.5 (Rolling a die twice)**.** Consider the random experiment of rolling a die twice. We model the experiment with $\Omega = \{1, \ldots, 6\}^2$, $\mathcal{A} = 2^\Omega$ and $\mathbf{P} = \mathcal{U}_\Omega$.

(i) If an event A_1 depends only on the outcome of the first roll and an event A_2 depends only on the outcome of the second roll, then A_1 and A_2 should be independent. Formally, let $\tilde{A}_1, \tilde{A}_2 \subset \{1, \ldots, 6\}$ and let

$$A_1 = \tilde{A}_1 \times \{1, \ldots, 6\} \quad \text{and} \quad A_2 = \{1, \ldots, 6\} \times \tilde{A}_2.$$

Finally,

$$\mathbf{P}[A_1 \cap A_2] = \frac{\#(\tilde{A}_1 \times \tilde{A}_2)}{36} = \frac{\#\tilde{A}_1}{6} \cdot \frac{\#\tilde{A}_2}{6} = \mathbf{P}[A_1] \cdot \mathbf{P}[A_2].$$

(ii) Stochastic independence can occur also in less obvious situations. For instance, let A be the event where the sum of the two rolls is odd,

$$A = \{\, (\omega_1, \omega_2) \in \Omega \;;\; \omega_1 + \omega_2 \in \{3, 5, 7, 9, 11\} \,\},$$

and let B be the event where the first roll gives at most three,

$$B = \{\, (\omega_1, \omega_2) \in \Omega \;;\; \omega_1 \in \{1, 2, 3\} \,\}.$$

Although it might seem that these two events are entangled in some way, they are stochastically independent. Indeed, it is not difficult to check that $\mathbf{P}[A] = \mathbf{P}[B] = \frac{1}{2}$ and $\mathbf{P}[A \cap B] = \frac{1}{4}$. □

What is the condition for three events A, B, C to be independent? Of course, any of the pairs (A, B), (B, C) and (A, C) has to be independent. However, we have to make sure also that the simultaneous occurrence of B and C does not change the probability that A occurs, i.e.,

$$\mathbf{P}[B \cap C] > 0 \quad \Rightarrow \quad \mathbf{P}[A \,|\, B \cap C] = \mathbf{P}[A].$$

Thanks to the following property we can define (stochastic) independence of three events in different ways.

Property 1.6. Let A, B and C be three events. Then the following propositions are equivalent:

(i) (a1) $\mathbf{P}[A \cap B] = \mathbf{P}[A] \cdot \mathbf{P}[B]$,
(b1) $\mathbf{P}[B \cap C] = \mathbf{P}[B] \cdot \mathbf{P}[C]$,
(c1) $\mathbf{P}[A \cap C] = \mathbf{P}[A] \cdot \mathbf{P}[C]$,
(d1) $\mathbf{P}[A \cap B \cap C] = \mathbf{P}[A] \cdot \mathbf{P}[B] \cdot \mathbf{P}[C]$;

(ii) (a2) $\mathbf{P}[B] > 0 \quad \Rightarrow \quad \mathbf{P}[A \,|\, B] = \mathbf{P}[A]$,
(b2) $\mathbf{P}[C] > 0 \quad \Rightarrow \quad \mathbf{P}[B \,|\, C] = \mathbf{P}[B]$,
(c2) $\mathbf{P}[C] > 0 \quad \Rightarrow \quad \mathbf{P}[A \,|\, C] = \mathbf{P}[A]$,
(d2) $\mathbf{P}[B \cap C] > 0 \quad \Rightarrow \quad \mathbf{P}[A \,|\, B \cap C] = \mathbf{P}[A]$;

(iii) (a3) Se satisface (a2), (b2), (c2),
(b3) $\mathbf{P}[A \cap C] > 0 \quad \Rightarrow \quad \mathbf{P}[B \,|\, A \cap C] = \mathbf{P}[B]$.

(iv) (a4) Se satisface (a2), (b2), (c2),
(b4) $\mathbf{P}[A \cap B] > 0 \quad \Rightarrow \quad \mathbf{P}[C \,|\, A \cap B] = \mathbf{P}[C]$.

Proof. Exercise. □

Example 1.7 (Rolling a die three times)**.** We roll a die three times. We model the experiment with $\Omega = \{1, \ldots, 6\}^3$, $\mathcal{A} = 2^\Omega$ and $\mathbf{P} = \mathcal{U}_\Omega$.

(i) If for any $i = 1, 2, 3$ the event A_i depends only on the outcome of the ith roll, then the events A_1, A_2 and A_3 should be independent. Indeed, let $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3 \subset \{1, \ldots, 6\}$ and let

$$A_1 = \tilde{A}_1 \times \{1, \ldots, 6\} \times \{1, \ldots, 6\},$$

$$A_2 = \{1, \ldots, 6\} \times \tilde{A}_2 \times \{1, \ldots, 6\},$$
$$A_3 = \{1, \ldots, 6\} \times \{1, \ldots, 6\} \times \tilde{A}_3.$$

Finally, if $i \neq j$, $\mathbf{P}[A_i \cap A_j] = \frac{\#\tilde{A}_i \cdot \#\tilde{A}_j \cdot \#\{1,\ldots,6\}}{216} = \frac{\#\tilde{A}_i}{6} \cdot \frac{\#\tilde{A}_j}{6} \cdot \frac{\#\{1,\ldots,6\}}{6} = \mathbf{P}[A_i] \cdot \mathbf{P}[A_j]$ and $\mathbf{P}[A_1 \cap A_2 \cap A_3] = \frac{\#(\tilde{A}_1 \times \tilde{A}_2 \times \tilde{A}_3)}{216} = \frac{\#\tilde{A}_1}{6} \cdot \frac{\#\tilde{A}_2}{6} \cdot \frac{\#\tilde{A}_3}{6} = \mathbf{P}[A_1] \cdot \mathbf{P}[A_2] \cdot \mathbf{P}[A_3]$.

(ii) Consider now the events

$$A_1 := \{\, (\omega_1, \omega_2, \omega_3) \in \Omega \,;\; \omega_1 = \omega_2 \,\},$$
$$A_2 := \{\, (\omega_1, \omega_2, \omega_3) \in \Omega \,;\; \omega_2 = \omega_3 \,\},$$
$$A_3 := \{\, (\omega_1, \omega_2, \omega_3) \in \Omega \,;\; \omega_1 = \omega_3 \,\}.$$

Then $\#A_1 = \#A_2 = \#A_3 = 36$; hence $\mathbf{P}[A_1] = \mathbf{P}[A_2] = \mathbf{P}[A_3] = \frac{1}{6}$. Furthermore, $\#(A_i \cap A_j) = 6$ if $i \neq j$; hence $\mathbf{P}[A_i \cap A_j] = \frac{1}{36}$ if $i \neq j$. On the other hand, we have $\#(A_1 \cap A_2 \cap A_3) = 6$, thus $\mathbf{P}[A_1 \cap A_2 \cap A_3] = \frac{1}{36} \neq \mathbf{P}[A_1] \cdot \mathbf{P}[A_2] \cdot \mathbf{P}[A_3]$. Therefore, the events A_1, A_2 and A_3 are not independent. □

In order to define independence of larger families of events, we have to request the validity of product formulas, like in propositions (i) of properties 1.4 and 1.6, not only for pairs and triplets but for all finite subfamilies of events. We thus make the following definition.

Definition 1.8 (Independence of events)**.** Let I be an arbitrary index set and let $(A_i)_{i \in I}$ be an arbitrary family of events. The family $(A_i)_{i \in I}$ is called **independent** if for any finite subset $J \subset I$ the product formula holds:

$$\mathbf{P}\left[\bigcap_{j \in J} A_j\right] = \prod_{j \in J} \mathbf{P}[A_j].$$

□

The most prominent example of an independent family of infinitely many events is given by the perpetuated independent repetition of a random experiment:

Example 1.9. Let E be a nonempty finite set (the set of possible outcomes of the individual experiment) and let $(p_e)_{e\in E}$ be a probability vector. We model the experiment with $\Omega = E^{\mathbb{N}}$, $\mathcal{A}$ be the σ-algebra generated by the "cylinder sets" and $\mathbf{P} = \sum_{e\in E} p_e\delta_e{}^{\otimes\mathbb{N}}$ be the product measure. Let $\tilde{A}_i \subset E$ for any $i \in \mathbb{N}$, and let A_i be the event where $\tilde{A}_i$ occurs in the ith experiment; that is,

$$A_i = \{\, (e_1, e_2, \dots) \in \Omega\ ;\ e_i \in \tilde{A}_i \,\} = \bigcup_{(e_1,\dots,e_i)\in E^{i-1}\times\tilde{A}_i} [e_1, \dots, e_n].$$

Intuitively, the family $(A_i)_{i\in\mathbb{N}}$ should be independent if the definition of independence makes any sense at all. We check that this is indeed the case. *Exercise.* □

If A and B are independent, then A^c and B also are independent since $\mathbf{P}[A^c\cap B] = \mathbf{P}[B]-\mathbf{P}[A\cap B] = \mathbf{P}[B]-\mathbf{P}[A]\mathbf{P}[B] = (1-\mathbf{P}[A])\mathbf{P}[B] = \mathbf{P}[A^c]\mathbf{P}[B]$. We generalize this observation in the following property.

Property 1.10. Let I be an arbitrary index set and let $(A_i)_{i\in I}$ be an arbitrary family of events. Define $B_i^0 = A_i$ and $B_i^1 = {A_i}^c$ for $i \in I$. Then the following three statements are equivalent.

(i) The family $(A_i)_{i\in I}$ is independent.

(ii) There is an $\alpha \in \{0,1\}^I$ such that the family $(B_i^{\alpha_i})_{i\in I}$ is independent.

(iii) For any $\alpha \in \{0,1\}^I$, the family $(B_i^{\alpha_i})_{i\in I}$ is independent.

Proof. Exercise. □

Example 1.11 (Euler product formula)**.** The Riemann zeta function, $\zeta(s)$, is a function of a complex variable s that analytically continues the sum of the infinite series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

which converges when the real part of s is greater than 1. The Euler product formula is a representation of the Riemann zeta function as an infinite product:

$$\zeta(s) = \prod_{p\in\mathcal{P}} \frac{1}{1-p^{-s}},$$

where $\mathcal{P} := \{\, p \in \mathbb{N}\ ;\ p \text{ is prime} \,\}$. We check that this is indeed the case for $s \in]1, +\infty[$. *Exercise.* □

If we roll a die infinitely often, what is the chance that the face shows a six infinitely often? This probability should equal one; otherwise, there would be a last point in time when we see a six and after which the face only shows a number one to five. However, this is not very plausible. The following theorem confirms the conjecture mentioned above and also gives conditions under which we cannot expect that infinitely many of the events occur.

Theorem 1.12 (Borel–Cantelli lemma)**.** Let $A_1, A_2, \ldots$ be events. It holds:

(i) $\sum_{n=1}^{+\infty} \mathbf{P}[A_n] < +\infty \quad \Rightarrow \quad \mathbf{P}[A^*] = 0.$

(ii) $(A_n)_{n\in\mathbb{N}}$ is independent, $\sum_{n=1}^{+\infty} \mathbf{P}[A_n] = +\infty \quad \Rightarrow \quad \mathbf{P}[A^*] = 1.$

Remark 1.13. Let $A_1, A_2, \ldots$ be independent events, then only the probabilities $\mathbf{P}[A^*] = 0$ and $\mathbf{P}[A^*] = 1$ could show up. Thus the Borel–Cantelli lemma belongs to the class of so-called 0–1 laws. □

Example 1.14. We throw a die again and again, then the probability of seeing a six infinitely often is one. *Exercise.* □

Example 1.15 (The necessity of the independent hypothesis in the Borel–Cantelli lemma)**.** *Exercise.* □

Example 1.16. Let $\Lambda \in]0, +\infty[$ and $0 \leq \lambda_n \leq \Lambda$ for every $n \in \mathbb{N}$. Let X_n, $n \in \mathbb{N}$, be Poisson random variables with parameters λ_n. Then

$$\mathbf{P}[X_n \geq n \quad \text{i.o.}] = 0.$$

Exercise. □

(*Proof of Borel–Cantelli lemma.*) *Exercise.* □

Now we extend the notion of independence from families of events to families of classes of events.

Definition 1.17 (Independence of classes of events)**.** Let I be an arbitrary index set and let $\mathcal{E}_i \subset \mathcal{A}$ for all $i \in I$. The family $(\mathcal{E}_i)_{i\in I}$ is called independent if, for any finite subset $J \subset I$ and any choice of $E_j \in \mathcal{E}_j$, $j \in J$, we have

$$\mathbf{P}\left[\bigcap_{j\in J} E_j\right] = \prod_{j\in J} \mathbf{P}[E_j]. \tag{1.1}$$

□

Example 1.18. Let E be a nonempty finite set (the set of possible outcomes of the individual experiment) and let $(p_e)_{e\in E}$ be a probability vector. Let $(\Omega, \mathcal{A}, \mathbf{P})$ be the probability space of example 1.9. For $i \in \mathbb{N}$ define

$$\mathcal{E}_i := \{\{\omega \in \Omega \,;\, \omega_i \in A\} \,;\, A \subset E\}\,.$$

We claim that the family $(\mathcal{E}_i)_{i\in\mathbb{N}}$ is independent. *Exercise.* □

Theorem 1.19. Let I be an arbitrary index set and let $\mathcal{E}_i \subset \mathcal{A}$ for all $i \in I$.

(i) If I is finite, and for any $i \in I$ let $\mathcal{E}_i \subset \mathcal{A}$ with $\Omega \in \mathcal{E}_i$. Then

$$(\mathcal{E}_i)_{i\in I} \text{ is independent} \quad \Leftrightarrow \quad (1.1) \text{ holds for } J = I.$$

(ii) $(\mathcal{E}_i)_{i\in I}$ is independent $\quad \Leftrightarrow \quad$ $(\mathcal{E}_j)_{j\in J}$ is independent for all finite $J \subset I$.

(iii) If $\mathcal{E}_i$ is a π-system for every $i \in I$, then

$$(\mathcal{E}_i)_{i\in I} \text{ is independent} \quad \Leftrightarrow \quad (\sigma(\mathcal{E}_i))_{i\in I} \text{ is independent.}$$

(iv) If K be an arbitrary index set and $(I_k)_{k\in K}$ are mutually disjoint subsets of I, then

$$(\mathcal{E}_i)_{i\in I} \text{ is independent} \quad \Rightarrow \quad (\cup_{i\in I_k}\mathcal{E}_i)_{k\in K} \text{ is independent.}$$

Proof. *Exercise.* □