

# Probability theory

Survey statistics and data analytics

Ágnes Backhausz

`agnes.backhausz@ttk.elte.hu`

office hours: Monday 10-11, Thursday 10-11, D 3-415  
(except for 6 Oct)

2025/2026. fall semester

# Goals

- **foundations of mathematical statistics**: to learn the basics of the methods that are used to analyze data from random samples
- foundations of probability theory: **events, random variables, expectation, variance, correlation coefficient and related concepts**
- **behavior of the average**: laws of large numbers, central limit theorem
- development of problem-solving skills (at the problem-solving course)

# Goals

- **foundations of mathematical statistics**: to learn the basics of the methods that are used to analyze data from random samples
- foundations of probability theory: **events, random variables, expectation, variance, correlation coefficient and related concepts**
- **behavior of the average**: laws of large numbers, central limit theorem
- development of problem-solving skills (at the problem-solving course)

Exam: written exam of 100 points; limits (2-5): 40, 55, 60, 75

Material: [moodle4.elte.hu](https://moodle4.elte.hu)

# Literature

- Ross: A First Course in Probability
- David Stirzaker, Geoffrey Grimmett: One thousand exercises in probability.

# Probability and statistics

## Goals:

- analysis of data obtained from surveys, experiments
- estimation of unknown quantities based on the measurements
- checking or rejecting hypotheses
- modelling random processes
- prediction of future processes based on observations from the past

# Probability and statistics

## Goals:

- analysis of data obtained from surveys, experiments
- estimation of unknown quantities based on the measurements
- checking or rejecting hypotheses
- modelling random processes
- prediction of future processes based on observations from the past

## Applications

- statistics in social sciences: analysis of surveys
- statistics in natural sciences: analysis of measurements, experiments
- prediction: social, economical, financial processes
- actuarial science (insurance)

# Short history of probability theory

- **problem of division**, 1494: how to divide the stake if the gamblers have to stop playing (solution: Pascal, 1656)
- Cardano's book on games with dies, 1564 (published in 1663)
- calculation of life annuity, de Witt, Haley, 1671
- laws of large numbers, Jacob Bernoulli, 1713
- first half of 19th century: de Moivre, Bayes, Gauss, Poisson, Buffon
- end of 19th century: Chebyshev, Markov, Ljapunov
- **axiomatic foundation**: Kolmogorov, 1933

# Short history of probability theory

## Applications in the 20th century

- stochastic processes (Wiener, 1923)
- mathematical statistics (Fisher, 1925)
- game theory (Neumann, 1928)
- information theory (Shannon, 1948)
- time series
- financial processes (Black–Scholes, 1973)
- hierarchical learning algorithms → artificial intelligence

# Events and their probabilities

**Goal:** modelling random events

What can be the **probability** of **an event**?

# Events and their probabilities

**Goal:** modelling random events

What can be the **probability** of **an event**?

rain tomorrow in Budapest



a randomly chosen person is left-handed



a randomly chosen person has a university degree



inflation decreases in June



event A



# Events and their probabilities

**Goal:** modelling random events

What can be the **probability** of **an event**?

rain tomorrow in Budapest



a randomly chosen person is left-handed



a randomly chosen person has a university degree



inflation decreases in June



event A



**the probability of an event depends on  
the choice of the model**

# Events and their probabilities

**Goal:** modelling random events

What can be the **probability** of **an event**?

rain tomorrow in Budapest



a randomly chosen person is left-handed



a randomly chosen person has a university degree



inflation decreases in June



event A



**the probability of an event depends on  
the choice of the model**

how do we take our observations into account

# Events and their probabilities

**Goal:** modelling random events

What can be the **probability** of **an event**?

rain tomorrow in Budapest

→ 15%

a randomly chosen person is left-handed

→ 10%

a randomly chosen person has a university degree

→ 22%

inflation decreases in June

→ 44%

event A

→  $P(A)$

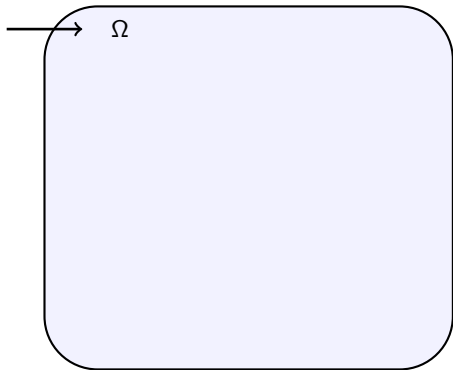
**the probability of an event depends on  
the choice of the model**

how do we take our observations into account  
if we choose a different model, the values will be different

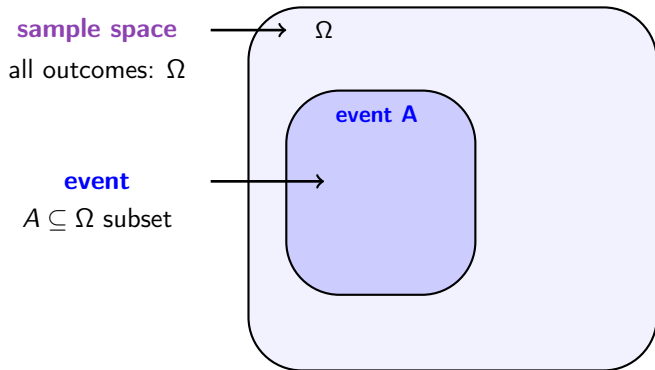
# (Kolmogorov) probability field

**sample space**

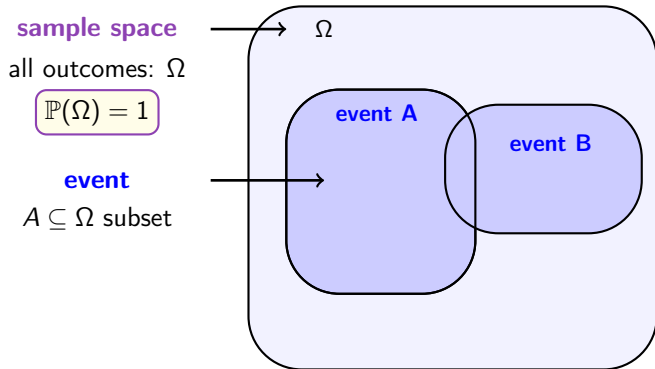
all outcomes:  $\Omega$



# (Kolmogorov) probability field



# (Kolmogorov) probability field



# (Kolmogorov) probability field

$\Omega$ : (Hungarian) **adult people**

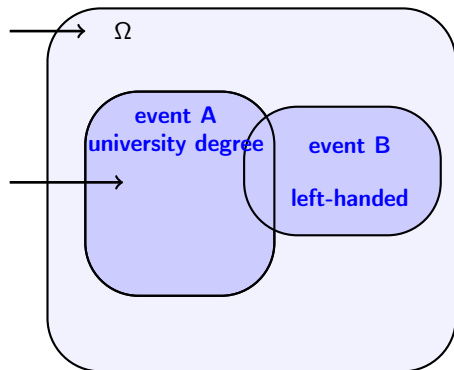
**sample space**

all outcomes:  $\Omega$

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

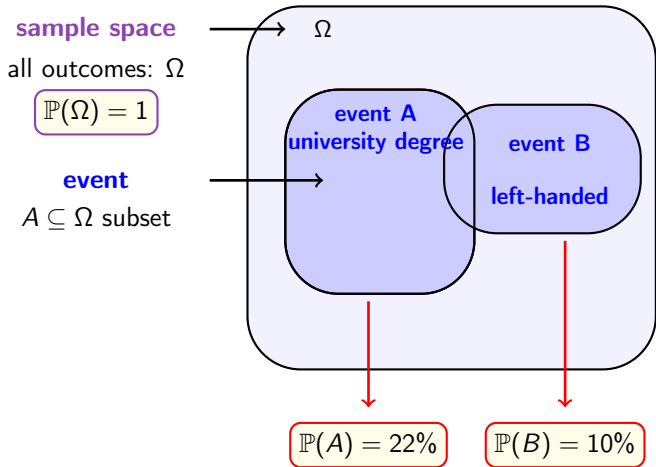
**event**

$A \subseteq \Omega$  subset



# (Kolmogorov) probability field

$\Omega$ : (Hungarian) **adult people**



# (Kolmogorov) probability field

$\Omega$ : (Hungarian) **adult people**

**sample space**

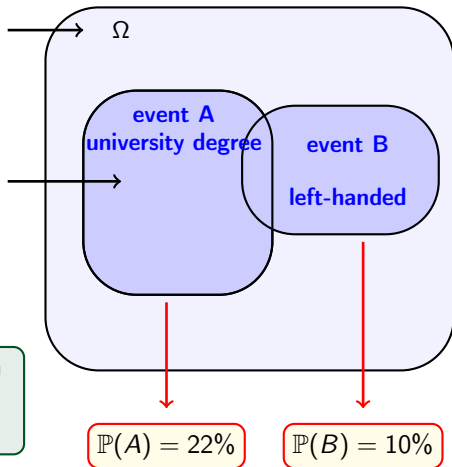
all outcomes:  $\Omega$

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

**event**

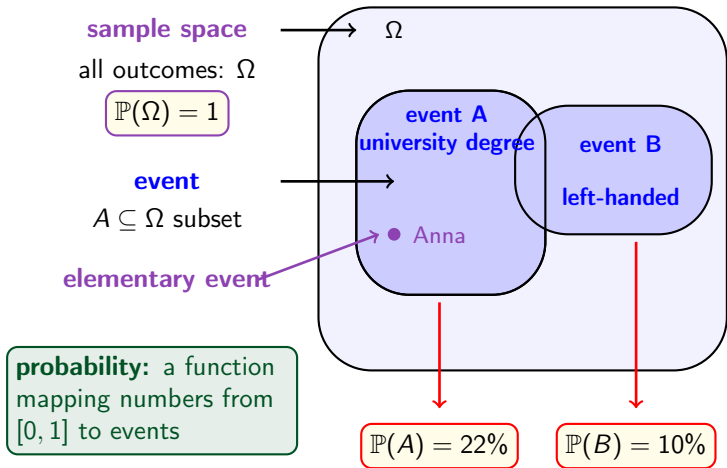
$A \subseteq \Omega$  subset

**probability:** a function mapping numbers from  $[0, 1]$  to events



# (Kolmogorov) probability field

$\Omega$ : (Hungarian) **adult people**



# Probability space/event space

- **sample space**: the set of all possible outcomes,  $\Omega$  (for example: Hungarian adults)

# Probability space/event space

- **sample space**: the set of all possible outcomes,  $\Omega$  (for example: Hungarian adults)
- **elementary event**: one single outcome, an element of  $\Omega$  (for example, Anna or Péter)

# Probability space/event space

- **sample space**: the set of all possible outcomes,  $\Omega$  (for example: Hungarian adults)
- **elementary event**: one single outcome, an element of  $\Omega$  (for example, Anna or Péter)
- **event**: a subset of the sample space,  $A \subseteq \Omega$   
for example: people with university degree ( $A$ ) or left-handedness ( $B$ )

# Probability space/event space

- **sample space**: the set of all possible outcomes,  $\Omega$  (for example: Hungarian adults)
- **elementary event**: one single outcome, an element of  $\Omega$  (for example, Anna or Péter)
- **event**: a subset of the sample space,  $A \subseteq \Omega$   
for example: people with university degree ( $A$ ) or left-handedness ( $B$ )
- **probability**: a function mapping numbers from  $[0, 1]$  to the events (for example:  $\mathbb{P}(A) = 22\%$ )

# Probability space/event space

- **sample space**: the set of all possible outcomes,  $\Omega$  (for example: Hungarian adults)
- **elementary event**: one single outcome, an element of  $\Omega$  (for example, Anna or Péter)
- **event**: a subset of the sample space,  $A \subseteq \Omega$   
for example: people with university degree ( $A$ ) or left-handedness ( $B$ )
- **probability**: a function mapping numbers from  $[0, 1]$  to the events (for example:  $\mathbb{P}(A) = 22\%$ )

Further examples:

- sample space: Hungarian men; event: age at least 50
- two coin flips:  $\{HH, HT, TH, TT\}$ , event: different outcomes, that is,  $\{HT, TH\}$
- we toss a coin until we get heads:  
 $\{H, TH, TTH, TTTH, \dots, \text{all tosses are T}\}$

## Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

## Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) We invite 5 people to a meeting.

What is the probability that at least 4 people will attend?

## Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) We invite 5 people to a meeting.

What is the probability that at least 4 people will attend?

everyone  
shows up

one person  
is missing

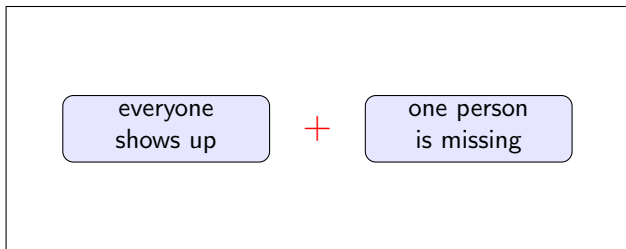
## Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) We invite 5 people to a meeting.

What is the probability that at least 4 people will attend?



$\mathbb{P}(\text{there are at least 4 people}) =$

$= \mathbb{P}(\text{everyone shows up}) + \mathbb{P}(\text{one person is missing})$

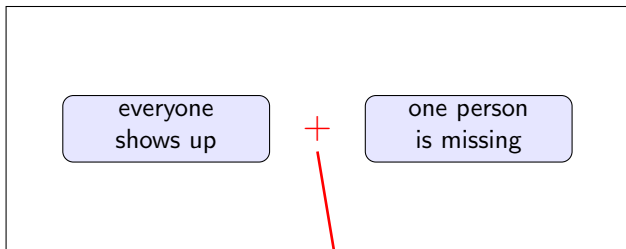
# Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) We invite 5 people to a meeting.

What is the probability that at least 4 people will attend?



$\mathbb{P}(\text{there are at least 4 people}) =$

$$= \mathbb{P}(\text{everyone shows up}) + \mathbb{P}(\text{one person is missing})$$

# Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) In general:

If events  $A$  and  $B$  are disjoint, that is, their intersection is empty, then the probability that at least one of them occurs, is the sum of their probabilities:

# Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) In general:

If events  $A$  and  $B$  are disjoint, that is, their intersection is empty, then the probability that at least one of them occurs, is the sum of their probabilities:

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

# Probability of events

1) The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

2) In general:

If events  $A$  and  $B$  are disjoint, that is, their intersection is empty, then the probability that at least one of them occurs, is the sum of their probabilities:

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

disjoint: cannot occur at the same time

union: at least one of them occurs

# Basic properties of probability

- 1 The probability of an event is always at least 0 and at most 1:

# Basic properties of probability

- 1 The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ ,

# Basic properties of probability

- ① **The probability of an event is always at least 0 and at most 1:**

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , **and the probability of the sample space  $\Omega$  is 1.**

# Basic properties of probability

- 1 The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , and the probability of the sample space  $\Omega$  is 1.

- 2 **Additivity:** if  $A_1, A_2, \dots$  are events, and no two can occur at the same time, that is,

# Basic properties of probability

- 1 The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , and the probability of the sample space  $\Omega$  is 1.

- 2 **Additivity:** if  $A_1, A_2, \dots$  are events, and no two can occur at the same time, that is,

$$A_i \cap A_j = \emptyset \quad \text{for every } i, j \geq 1,$$

then the probability of their union is the sum of their probabilities:

# Basic properties of probability

- 1 The probability of an event is always at least 0 and at most 1:

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , **and the probability of the sample space  $\Omega$  is 1.**

- 2 **Additivity:** if  $A_1, A_2, \dots$  are events, and no two can occur at the same time, that is,

$$A_i \cap A_j = \emptyset \quad \text{for every } i, j \geq 1,$$

then the probability of their union is the sum of their probabilities:

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \dots$$

# Basic properties of probability

- ① **The probability of an event is always at least 0 and at most 1:**

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , **and the probability of the sample space  $\Omega$  is 1.**

- ② **Additivity:** if  $A_1, A_2, \dots$  are events, and no two can occur at the same time, that is,

$$A_i \cap A_j = \emptyset \quad \text{for every } i, j \geq 1,$$

then the probability of their union is the sum of their probabilities:

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \dots$$

The same with another notation:

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j).$$

# Basic properties of probability

- **The probability of an event is always at least 0 and at most 1:**

# Basic properties of probability

- **The probability of an event is always at least 0 and at most 1:**

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ ,

## Basic properties of probability

- **The probability of an event is always at least 0 and at most 1:**

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , **and the probability of the sample space  $\Omega$  is 1.**

# Basic properties of probability

- **The probability of an event is always at least 0 and at most 1:**

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , **and the probability of the sample space  $\Omega$  is 1.**

- **Sample space** (all outcomes, certain event,  $\Omega$ ) has probability 1:

$$\mathbb{P}(\Omega) = 1; \quad \mathbb{P}(\emptyset) = 0,$$

while the **impossible event** (empty set) has probability 0.

# Basic properties of probability

- **The probability of an event is always at least 0 and at most 1:**

$$0 \leq \mathbb{P}(A) \leq 1$$

holds for every event  $A$ , **and the probability of the sample space  $\Omega$  is 1.**

- **Sample space** (all outcomes, certain event,  $\Omega$ ) has probability 1:

$$\mathbb{P}(\Omega) = 1; \quad \mathbb{P}(\emptyset) = 0,$$

while the **impossible event** (empty set) has probability 0.

- **Probability of complement:** if  $A \subseteq \Omega$  is an event, then the probability that  **$A$  does not occur:**

$$\mathbb{P}(\bar{A}) = \mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A).$$

- **Probability of difference:**

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

# Axiomatic foundation of probability

$\Omega$ : event space (set of all possible outcomes)

$\mathcal{A}$ : set of events

$\mathbb{P}$ : **probability**, with the following properties:

- **mapping number to events**, more precisely, for  $A \in \mathcal{A}$  we have  $\mathbb{P}(A) \geq 0$
- **additive**: if  $A_1, A_2, \dots \in \mathcal{A}$  are events, and they are pairwise disjoint, then

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \dots$$

- $\mathbb{P}(\Omega) = 1$ , that is, the sample space has probability 1

The triad  $(\Omega, \mathcal{A}, \mathbb{P})$  is called **a probability space**.

## (Kolmogorov) probability space

The triad  $(\Omega, \mathcal{A}, \mathbb{P})$  is a (Kolmogorov) **probability space**, if

- the **sample space**  $\Omega$  is a non-empty set;

## (Kolmogorov) probability space

The triad  $(\Omega, \mathcal{A}, \mathbb{P})$  is a (Kolmogorov) **probability space**, if

- the **sample space**  $\Omega$  is a non-empty set;
- $\mathcal{A} \subseteq \mathcal{P}(\Omega)$  is the **set of events**: for every  $A \in \mathcal{A}$  we have  $A \subseteq \Omega$  such that
  - i)  $\Omega \in \mathcal{A}$ ;
  - ii) if  $A_1, A_2, \dots \in \mathcal{A}$ , then  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$  (a countable union of elements in  $\mathcal{A}$  is also in  $\mathcal{A}$ );
  - iii) if  $A \in \mathcal{A}$ , then  $\Omega \setminus A \in \mathcal{A}$  (that is, the complement of a set in  $\mathcal{A}$  is also in  $\mathcal{A}$ ).

## (Kolmogorov) probability space

The triad  $(\Omega, \mathcal{A}, \mathbb{P})$  is a (Kolmogorov) **probability space**, if

- the **sample space**  $\Omega$  is a non-empty set;
- $\mathcal{A} \subseteq \mathcal{P}(\Omega)$  is the **set of events**: for every  $A \in \mathcal{A}$  we have  $A \subseteq \Omega$  such that
  - ❶  $\Omega \in \mathcal{A}$ ;
  - ❷ if  $A_1, A_2, \dots \in \mathcal{A}$ , then  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$  (a countable union of elements in  $\mathcal{A}$  is also in  $\mathcal{A}$ );
  - ❸ if  $A \in \mathcal{A}$ , then  $\Omega \setminus A \in \mathcal{A}$  (that is, the complement of a set in  $\mathcal{A}$  is also in  $\mathcal{A}$ ).
- **probability** is a function  $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ , such that
  - ❶  $\mathbb{P}(\Omega) = 1$ , that is, the sample space has probability 1;
  - ❷ if  $A_1, A_2, \dots \in \mathcal{A}$  and for every  $1 \leq i < j$  we have  $A_i \cap A_j = \emptyset$ , then

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n),$$

that is, the probability of the union of countably many, pairwise disjoint events is the sum of the probabilities.

## Example: sum of two rolls

We roll two fair dice. What is the probability that the sum is 7?

## Example: sum of two rolls

We roll two fair dice. What is the probability that the sum is 7?

**Dice, coins, people etc. are always different and considered different.**

- **sample space:** possible series of rolls. Their number:

## Example: sum of two rolls

We roll two fair dice. What is the probability that the sum is 7?

**Dice, coins, people etc. are always different and considered different.**

- **sample space:** possible series of rolls. Their number:

$$6 \cdot 6 = 36; \text{ both values can chosen } 6 \text{ ways.}$$

## Example: sum of two rolls

We roll two fair dice. What is the probability that the sum is 7?

**Dice, coins, people etc. are always different and considered different.**

- **sample space:** possible series of rolls. Their number:

$$6 \cdot 6 = 36; \text{ both values can chosen } 6 \text{ ways.}$$

- All outcomes have the same probability: each has probability  $1/36$ .
- Number of outcomes where the sum is 7:

## Example: sum of two rolls

We roll two fair dice. What is the probability that the sum is 7?

**Dice, coins, people etc. are always different and considered different.**

- **sample space:** possible series of rolls. Their number:

$$6 \cdot 6 = 36; \text{ both values can chosen 6 ways.}$$

- All outcomes have the same probability: each has probability  $1/36$ .
- Number of outcomes where the sum is 7: 6.

11	12	13	14	15	<b>16</b>
21	22	23	24	<b>25</b>	26
31	32	33	<b>34</b>	35	36
41	42	<b>43</b>	44	45	46
51	<b>52</b>	53	54	55	56
<b>61</b>	62	63	64	65	66

Hence  $\mathbb{P}(\text{the sum is 7}) = 6/36 = 1/6$ .

## Example: two dice rolls

- **elementary events:** possible outcomes of the experiment, a series of rolls, for example: **15** or **22**
- **sample space:** the set of elementary events; the set  $\Omega$  has 36 elements, as follows:  $\Omega = \{11, 12, 13, \dots, 16, 21, \dots, 26, \dots, 61, \dots, 66\}$ .

## Example: two dice rolls

- **elementary events:** possible outcomes of the experiment, a series of rolls, for example: **15** or **22**
- **sample space:** the set of elementary events; the set  $\Omega$  has 36 elements, as follows:  $\Omega = \{11, 12, 13, \dots, 16, 21, \dots, 26, \dots, 61, \dots, 66\}$ .
- **event:** subsets of the sample space  $\Omega$   
for example: the sum is 7, that is,  $A = \{16, 25, 34, 43, 52, 61\}$   
or: the sum is at most 3, that is,  $B = \{11, 12, 21\}$   
 $\mathcal{A}$ : set of events, in this case all subsets of  $\Omega$   
total number of events:  $|\mathcal{A}| = 2^{36}$

## Example: two dice rolls

- **elementary events:** possible outcomes of the experiment, a series of rolls, for example: **15** or **22**
- **sample space:** the set of elementary events; the set  $\Omega$  has 36 elements, as follows:  $\Omega = \{11, 12, 13, \dots, 16, 21, \dots, 26, \dots, 61, \dots, 66\}$ .
- **event:** subsets of the sample space  $\Omega$   
for example: the sum is 7, that is,  $A = \{16, 25, 34, 43, 52, 61\}$   
or: the sum is at most 3, that is,  $B = \{11, 12, 21\}$   
 $\mathcal{A}$ : set of events, in this case all subsets of  $\Omega$   
total number of events:  $|\mathcal{A}| = 2^{36}$
- **probability:** a function  $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$  mapping numbers from  $[0, 1]$  to the events. For example:

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}; \quad \mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{3}{36} = \frac{1}{12}.$$

## Connection to relative frequency

Let  $\Omega$  be the sample space, the set of all possible outcomes, and  $A \subseteq \Omega$  an event, a subset of  $\Omega$ .

The **relative frequency** of event  $A$  out of  $n$  experiments:

$$r(A) = \frac{\text{number of occurrences of } A}{\text{number of experiments}} = \frac{\text{number of occurrences of } A}{n}.$$

Relative frequency satisfies the following properties:

- **it maps numbers from  $[0, 1]$  to events**, that is, for  $A \in \mathcal{A}$  we have  $0 \leq r(A) \leq 1$
- **additive**: if  $A_1, A_2, \dots \in \mathcal{A}$  are events, and they are pairwise disjoint, then

$$r(A_1 \cup A_2 \cup A_3 \cup \dots) = r(A_1) + r(A_2) + r(A_3) + \dots$$

- $r(\Omega) = 1$ , that is, the relative frequency of  $\Omega$  is 1

Hence relative frequency satisfies the properties of probability mentioned above.