Sampling with replacement: Lecture 3

Peter has **three blue** and

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

number of **all outcomes**:

5

five options on the first day

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

number of **all outcomes**:

5 5

on the first two days $5 \cdot 5 = 25$ options
each cup goes with any other, we can take a product
the cups are different
all outcomes are equally likely

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

number of **all outcomes**:

5 5 5

on the first three days $5 \cdot 5 \cdot 5 = 125$ options
each cup goes with any other, we can take a product
the cups are different
all outcomes are equally likely

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

number of **all outcomes**:

$$5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad \longrightarrow \quad \boxed{5^7}$$

there are 5^7 equally likely options during a week
each cup goes with any other, we can take a product
the cups are different
all outcomes are equally likely

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

a good option is when he drinks from blue cups until Thursday:

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

a good option is when he drinks from blue cups until Thursday:

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

a good option is when he drinks from blue cups until Thursday:

$$\longrightarrow \boxed{3^4 \cdot 2^3}$$

$3 \cdot 3 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \cdot 2 = 3^4 \cdot 2^3$ such options
each cup goes with any other, we can multiply

the cups are different

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

another good option is when he chooses blue on days $1., 3., 4., 6.$:

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

another good option is when he chooses blue on days $1., 3., 4., 6.$:

3 2 3 3 2 3 2

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

another good option is when he chooses blue on days $1., 3., 4., 6.$:

$\longrightarrow \boxed{3^4 \cdot 2^3}$

$3 \cdot 2 \cdot 3 \cdot 3 \cdot 2 \cdot 3 \cdot 2 = 3^4 \cdot 2^3$ such options
each cup goes with any other, we can multiply

the cups are different

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

number of good outcomes:

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

number of good outcomes:

number of colorings

number of outcomes with a given coloring

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

number of good outcomes:

$$\boxed{\text{number of colorings}} \times \boxed{\text{number of outcomes with a given coloring}}$$

$$\boxed{\binom{7}{4}} \times \boxed{3^4 \cdot 2^3} = 35 \cdot 3^4 \cdot 2^3$$

the number of ways to choose four days out of seven

the same for each coloring

Sampling with replacement: Lecture 3

Peter has **three blue** and **two green** coffee cups.
Each morning, he chooses a cup randomly,
forgetting the previous choices and taking
all five with the same probability.

What is the probability that **during one week**
he drinks from a blue cup exactly for times?

5^7 equally likely outcomes

$$\boxed{\mathbb{P}(\text{blue cup exactly four times}) = \frac{\binom{7}{4}\cdot 3^4\cdot 2^3}{5^7},}$$

that is, $\frac{35\cdot 81\cdot 8}{78125} = 29,03\%$.

Sampling with replacement

- There are N balls in a box, among which M is red, the others are blue.
- **By putting the balls back after each draw**, we choose n balls, forgetting all previous events, and choosing all N balls with the same probability.
- The probability that there are k red balls among the n chosen ones:

Sampling with replacement

- There are N balls in a box, among which M is red, the others are blue.
- **By putting the balls back after each draw**, we choose n balls, forgetting all previous events, and choosing all N balls with the same probability.
- The probability that there are k red balls among the n chosen ones:

$$\boxed{\mathbb{P}(\text{exactly } k \text{ red}) = \binom{n}{k}\frac{M^k \cdot (N-M)^{n-k}}{N^n} \qquad (k = 0, 1, 2, \ldots, n).}$$

Sampling with replacement

- There are N balls in a box, among which M is red, the others are blue.
- **By putting the balls back after each draw**, we choose n balls, forgetting all previous events, and choosing all N balls with the same probability.
- The probability that there are k red balls among the n chosen ones:

$$\boxed{\mathbb{P}(\text{exactly } k \text{ red}) = \overbrace{\binom{n}{k}}^{\text{which } k \text{ choices are red}} \frac{\overbrace{M^k}^{\text{which red balls were chosen } (k \text{ draws})} \cdot \overbrace{(N-M)^{n-k}}^{\text{which blue ones were chosen } (n-k \text{ draws})}}{\underbrace{N^n}_{\text{all outcomes}}} \qquad (k = 0, 1, 2, \dots, n).}$$

Sampling with replacement

- There are N balls in a box, among which M is red, the others are blue.
- **By putting the balls back after each draw**, we choose n balls, forgetting all previous events, and choosing all N balls with the same probability.
- The probability that there are k red balls among the n chosen ones:

which red balls were chosen (k draws)

which k choices are red

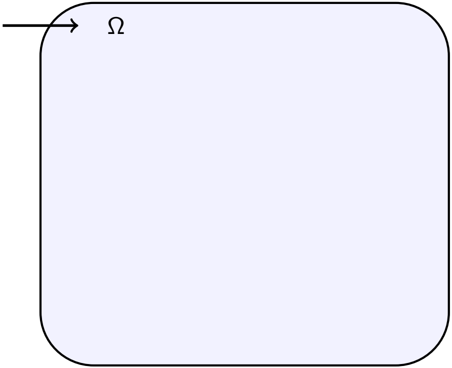
which blue ones were chosen ($n-k$ draws)

$$\boxed{\mathbb{P}(\text{exactly } k \text{ red}) = \binom{n}{k}\frac{M^k\cdot(N-M)^{n-k}}{N^n} \qquad (k=0,1,2,\dots,n).}$$

all outcomes

Remark: the number of chosen red balls has **binomial distribution**.
Since all N^n cases were uniformly likely, this is also **a classical probability space**.

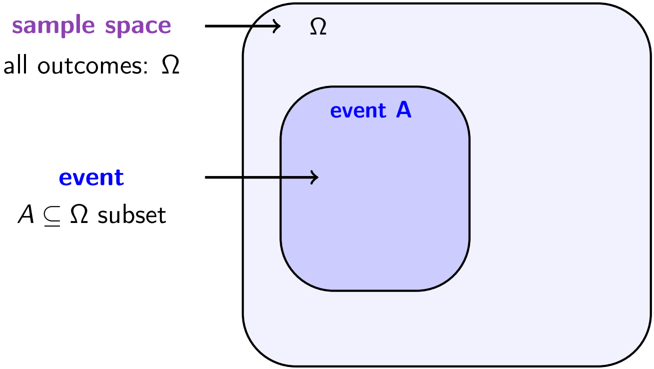
Kolmogorov probability space



sample space

all outcomes: Ω

Kolmogorov probability space



sample space

all outcomes: Ω

Ω

event A

event

$A \subseteq \Omega$ subset

Kolmogorov probability space

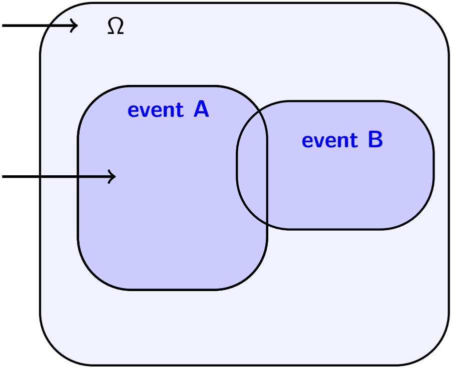
sample space

all outcomes: Ω

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

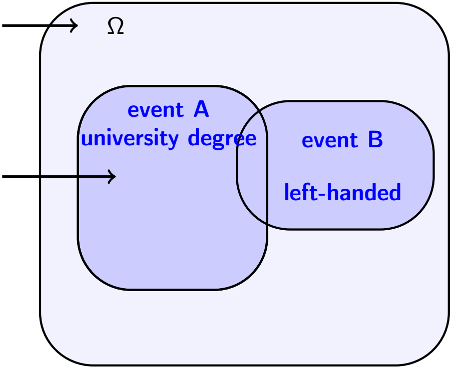
event

$A \subseteq \Omega$ subset



Kolmogorov probability space

Ω: (Hungarian) **adult people**



sample space

all outcomes: Ω

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

event

$A \subseteq \Omega$ subset

Ω

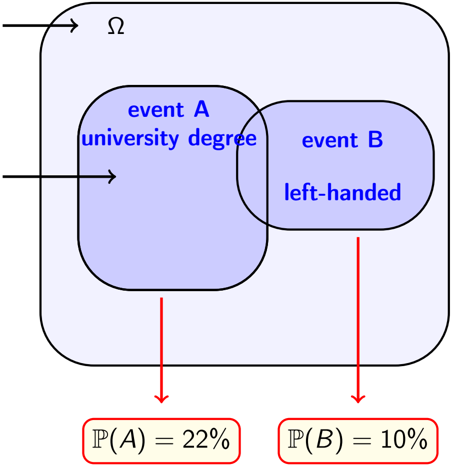
event A
university degree

event B

left-handed

Kolmogorov probability space

Ω: (Hungarian) **adult people**



sample space

all outcomes: Ω

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

event

$A \subseteq \Omega$ subset

Ω

event A
university degree

event B

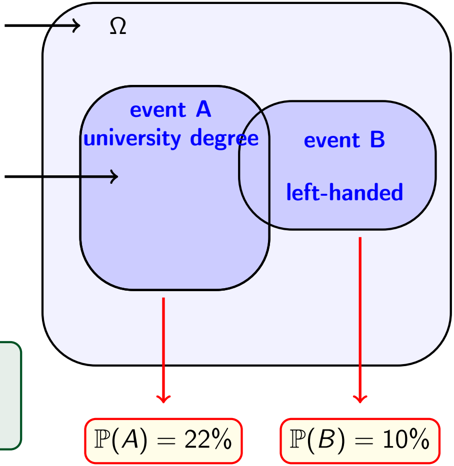
left-handed

$\boxed{\mathbb{P}(A) = 22\%}$

$\boxed{\mathbb{P}(B) = 10\%}$

Kolmogorov probability space

Ω: (Hungarian) **adult people**



sample space

all outcomes: Ω

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

event

$A \subseteq \Omega$ subset

Ω

event A
university degree

event B
left-handed

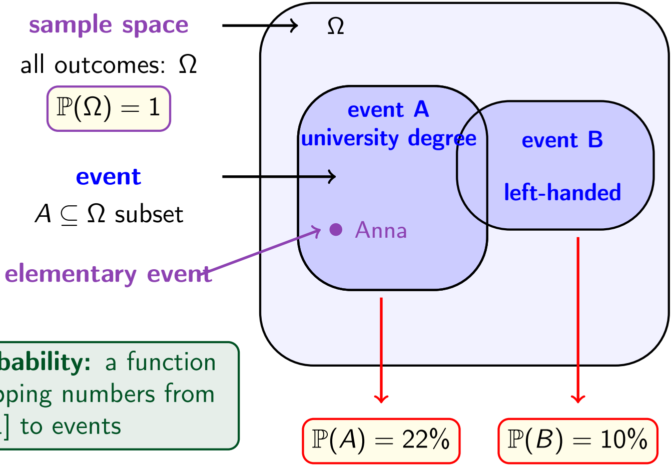
$\boxed{\mathbb{P}(A) = 22\%}$

$\boxed{\mathbb{P}(B) = 10\%}$

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

Kolmogorov probability space

Ω: (Hungarian) **adult people**



sample space

all outcomes: Ω

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

event

$A \subseteq \Omega$ subset

elementary event

$\mathbb{P}(A) = 22\%$

$\mathbb{P}(B) = 10\%$

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

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 Ω 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 Ω is 1.**

2. **Additivity:** if $A_1, A_2, \ldots$ 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 Ω is 1.**

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

   $$A_i \cap A_j = \emptyset \qquad \text{for every } 1 \leq i < j,$$

   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 Ω is 1.**

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

   $$A_i \cap A_j = \emptyset \qquad \text{for every } 1 \leq i < j,$$

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

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

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 Ω is 1.**

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

   $$A_i \cap A_j = \emptyset \qquad \text{for every } 1 \leq i < j,$$

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

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

   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 Ω 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 Ω is 1.**
- **Sample space** (all outcomes, certain event, Ω) has probability 1:

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

  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 Ω is 1.**

- **Sample space** (all outcomes, certain event, Ω) has probability 1:

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

  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**:

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

- **Probability of difference**:

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

(Kolmogorov) probability space

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

- the **sample space** Ω is a non-empty set;

(Kolmogorov) probability space

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

- the **sample space** Ω 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 $A_1, A_2, \ldots \in \mathcal{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ (a countable union of events is also an event);
  - (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** Ω 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
  1. $\Omega \in \mathcal{A}$;
  2. if $A_1, A_2, \ldots \in \mathcal{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ (a countable union of events is also an event);
  3. 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 $\mathbb{P} : \mathcal{A} \to [0, 1]$, such that
  1. $\mathbb{P}(\Omega) = 1$, that is, the sample space has probability 1;
  2. if $A_1, A_2, \ldots \in \mathcal{A}$ 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.

Finite probability space

Suppose that there are finitely many outcomes: $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, and $\mathcal{A}$ is the set of all subsets of Ω.

Notation: $p_j = \mathbb{P}(\{\omega_j\})$ is the probability of the jth outcome. By additivity, we have

$$1 = \mathbb{P}(\Omega) = p_1 + p_2 + \ldots + p_n = \sum_{j=1}^{n} p_j,$$

that is, the probability of the elementary events sums up to 1. Furthermore,

$$\mathbb{P}(A) = \mathbb{P}\Bigg(\bigcup_{j:\, \omega_j \in A} \{\omega_j\} \Bigg) = \sum_{j:\, \omega_j \in A} p_j,$$

meaning that

the probability of an event is the sum of the probabilities corresponding to the elementary events that it consists of.

Classical probability space

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

- Ω is a finite set;
- $\mathcal{A}$ is the set of all subsets of Ω;
- **all elementary events are equally likely**, that is,

$$\mathbb{P}(\omega_j) = p_j = \frac{1}{n} \qquad \text{for all } j = 1, 2, \ldots, n.$$

Then $(\Omega, \mathcal{A}, \mathbb{P})$ is a **classical probability space**. In this case, for every event $A \in \mathcal{A}$ we have

$$\mathbb{P}(A) = \frac{k}{n} = \frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n},$$

where k is the size of A, and n is the total number of elementary events (outcomes).

Classical probability space

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

- Ω is a finite set;
- $\mathcal{A}$ is the set of all subsets of Ω;
- **all elementary events are equally likely**, that is,

$$\mathbb{P}(\omega_j) = p_j = \frac{1}{n} \qquad \text{for all } j = 1, 2, \ldots, n.$$

Then $(\Omega, \mathcal{A}, \mathbb{P})$ is a **classical probability space**. In this case, for every event $A \in \mathcal{A}$ we have

$$\mathbb{P}(A) = \frac{k}{n} = \frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n},$$

where k is the size of A, and n is the total number of elementary events (outcomes).

Examples: sampling with and without replacement, in both cases, elementary events were equally likely

Example: dice rolls

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

Example: dice rolls

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

Dice, people, objects etc. are always different.

- **sample space**: possible sequences. Their number:

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

- The sequences are equally likely: each has probability 1/36.
- Number of good sequences: 6.

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

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

Counting and notation

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots 2 \cdot 1; \; \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots 2 \cdot 1}.$$

- n objects have $n!$ different orders.
- from n objects, we can choose k different $n(n-1)\ldots(n-k+1)$ many ways, if order matters: n ways for the first, $n-1$ for the second (independently of the first one), $n-2$ for the third (independently of the first two), and so on
- from n object a group of size k can be choosen $\binom{n}{k}$ many ways (order does not matter here)
- if for n experiments, we have k options for each, then the total number of outcomes (with order) is

  $$n \cdot n \cdot \ldots \cdot n = n^k.$$

  For example, two dice rolls have 36 possibilities, three rolls have $6 \cdot 6 \cdot 6$, and n have 6^n.

Properties

$$\binom{n}{k} = \binom{n}{n-k}; \quad \binom{n}{0} = 1; \binom{n}{1} = n; \quad \binom{n}{2} = \frac{n(n-1)}{2}$$

$$\binom{n}{N} = 0, \text{ if } N > n.$$

$(x+y)^2 = x^2 + 2xy + y^2$ is generalized to the **binomial theorem**:

$$(x+y)^n = x^n + nx^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{k}x^k y^{n-k} + \ldots + nxy^{n-1} + y^n.$$

Consequence for $x = y = 1$:

$$2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{k} + \ldots + \binom{n}{n-1} + \binom{n}{n}.$$

Multinomial/polynomial theorem

- $(x+y)^2 = x^2 + 2xy + y^2$
- $(x+y+z)^2 = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz$
- $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$

Multinomial/polynomial theorem

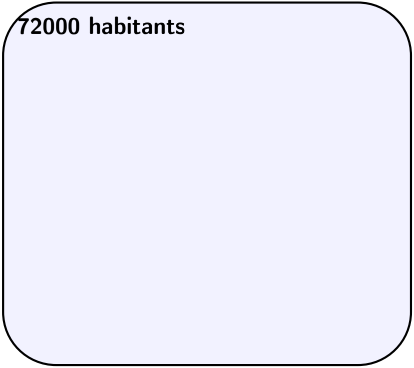
- $(x+y)^2 = x^2 + 2xy + y^2$
- $(x+y+z)^2 = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz$
- $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$
- in general:

$$(x_1 + x_2 + \ldots + x_k)^n = \sum \frac{n!}{i_1! \cdot i_2! \cdot \ldots \cdot i_k!} \cdot x_1^{i_1} \cdot x_2^{i_2} \cdot \ldots \cdot x_k^{i_k},$$

  where we sum up to all sequences of positive integers $(i_1, i_2, \ldots, i_k)$, for which $i_1 + i_2 + \ldots + i_k = n$.

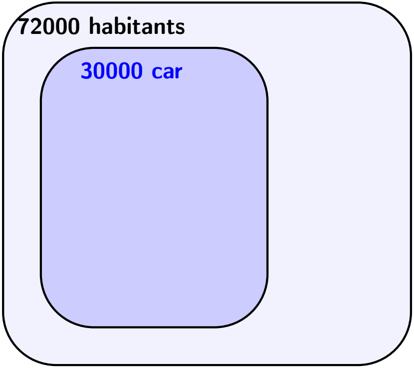
Probability of the union: example

In a city with **72000** habitants



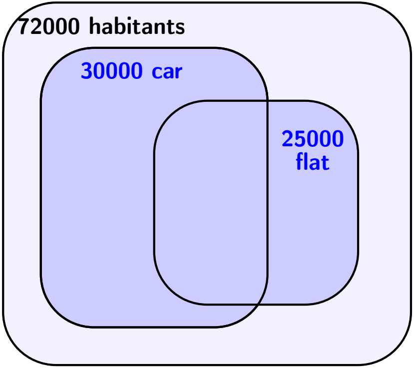
Probability of the union: example

In a city with **72000** habitants **30000** have a car,



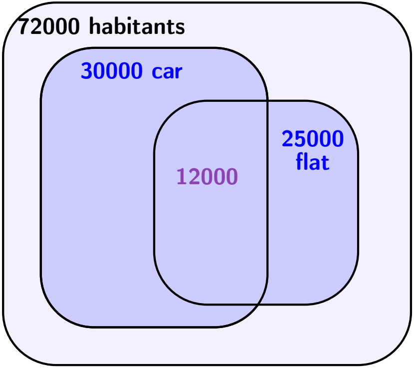
Probability of the union: example

In a city with **72000** habitants **30000** have a car, **25000** have a flat,



Probability of the union: example

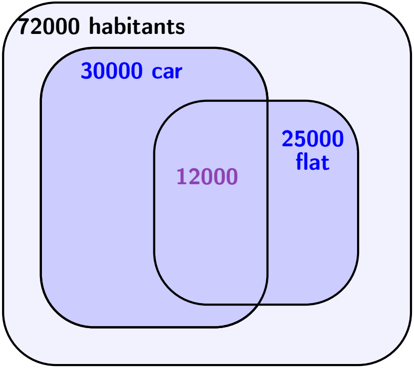
In a city with **72000** habitants **30000** have a car, **25000** have a flat, **12000** have both a car and a flat.



Probability of the union: example

In a city with **72000** habitants **30000** have a car,
25000 have a flat, **12000** have both a car and a flat.

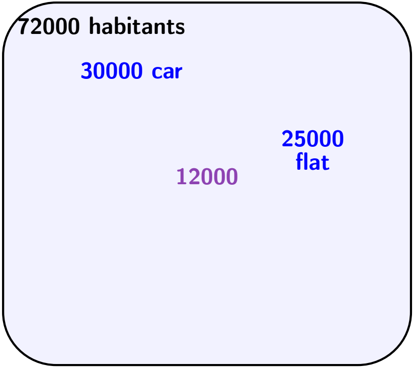
What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?



Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

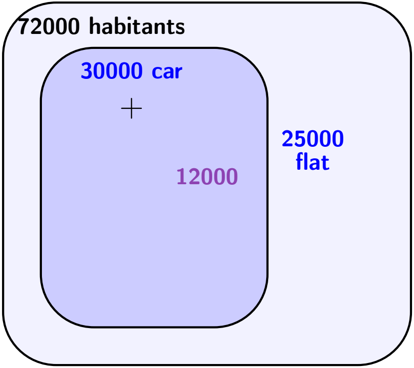
$\mathbb{P}(\textbf{at least one}) =$



Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

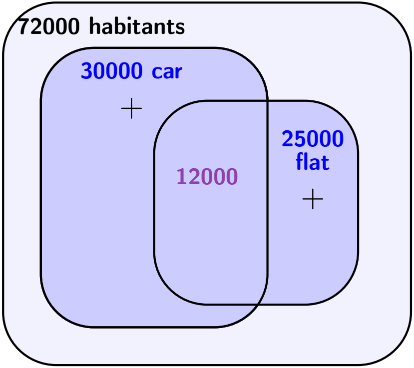
$\mathbb{P}(\textbf{at least one}) = \qquad \mathbb{P}(\textbf{car})$



Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

$$\mathbb{P}(\textbf{at least one}) = \quad \mathbb{P}(\textbf{car}) \quad + \mathbb{P}(\textbf{flat})$$

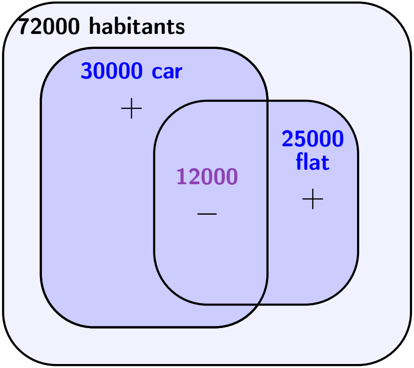


Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

$$\mathbb{P}(\textbf{at least one}) = \quad \mathbb{P}(\textbf{car}) \quad + \mathbb{P}(\textbf{flat}) \quad - \mathbb{P}(\textbf{both})$$

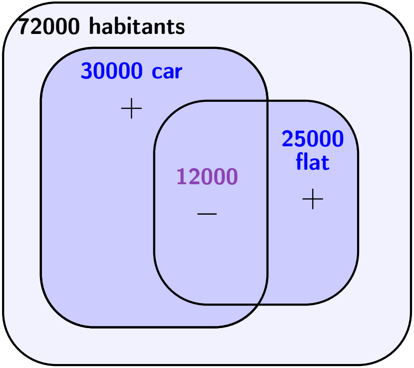
12000 habitants were counted twice, we substract this



Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

$$\mathbb{P}(\textbf{at least one}) = \mathbb{P}(\textbf{car}) + \mathbb{P}(\textbf{flat}) - \mathbb{P}(\textbf{both})$$
$$= 30000/72000 + 25000/72000 - 12000/72000$$



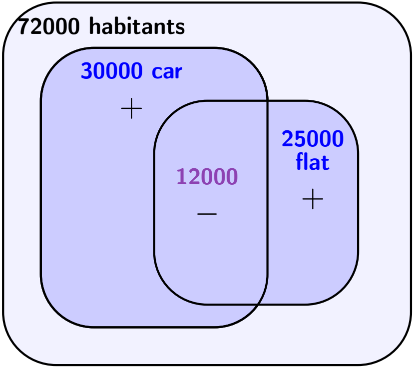
Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

$$\mathbb{P}(\textbf{at least one}) = \mathbb{P}(\textbf{car}) + \mathbb{P}(\textbf{flat}) - \mathbb{P}(\textbf{both})$$

$$= 30000/72000 + 25000/72000 - 12000/72000$$

$$= 43000/72000 = 59\%.$$



Probability of the union: example

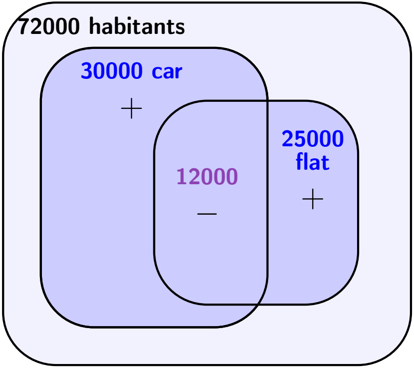
What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

$$\boxed{\mathbb{P}(A \cup B) = \quad \mathbb{P}(\textbf{car}) \quad + \mathbb{P}(\textbf{flat}) \quad - \mathbb{P}(\textbf{both})}$$

$$= 30000/72000 + 25000/72000 - 12000/72000$$

$$= 43000/72000 = 59\%.$$

↑ **union**



Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

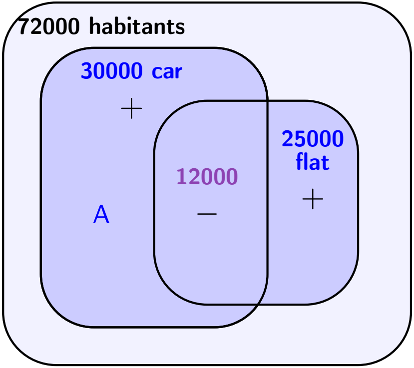
$$\boxed{\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(\textbf{flat}) - \mathbb{P}(\textbf{both})}$$

$$= 30000/72000 + 25000/72000 - 12000/72000$$

$$= 43000/72000 = 59\%.$$

↑

union



Probability of the union: example

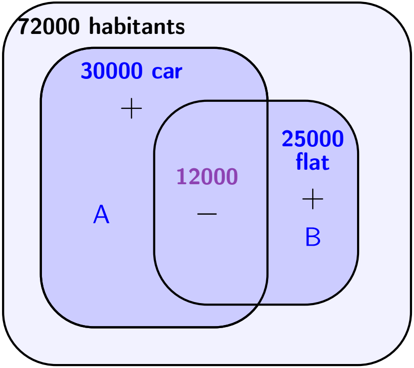
What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

$$\boxed{\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(\textbf{both})}$$

$$= 30000/72000 + 25000/72000 - 12000/72000$$

$$= 43000/72000 = 59\%.$$

↑ **union**



Probability of the union: example

What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

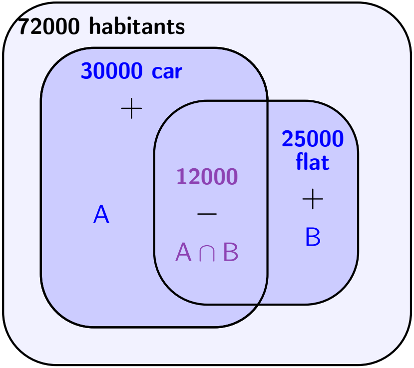
$$\boxed{\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)}$$

$$= 30000/72000 + 25000/72000 - 12000/72000$$

$$= 43000/72000 = 59\%.$$

↑ **union**

↑ **intersection**



Probability of the union: example

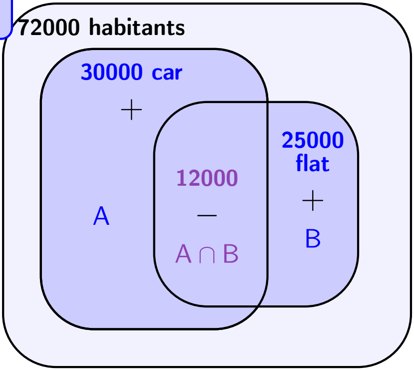
What is the probability that a uniformly randomly chosen habitant has **at least one of** a car and a flat?

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

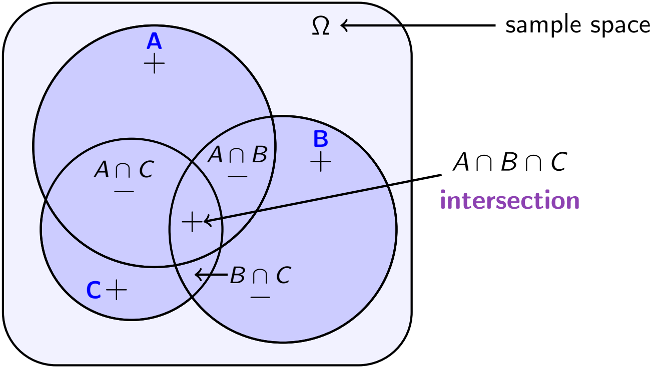
$$= 30000/72000 + 25000/72000 - 12000/72000$$

$$= 43000/72000 = 59\%.$$

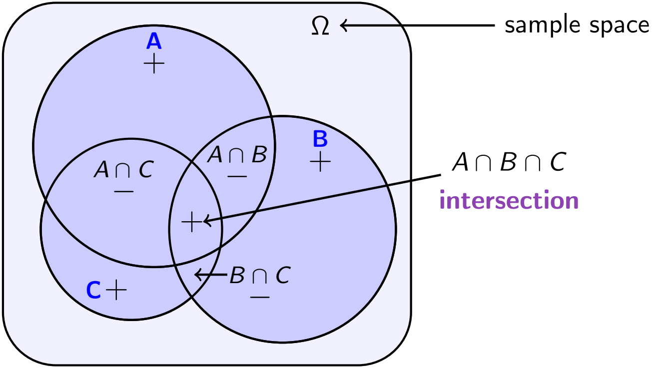
inclusion-exclusion formula



Inclusion-exclusion formula for three events



Inclusion-exclusion formula for three events



The probability of the **union** (at least one occurs):

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ &-\mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \\ &+\mathbb{P}(A \cap B \cap C).\end{aligned}$$

Inclusion-exclusion formula

For two events. The probability that at least one of A and B occurs:

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

For three events. The probability that at least one of A, B and C occurs:

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \\ &\quad - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)\end{aligned}$$

Inclusion-exclusion formula

For two events. The probability that at least one of A and B occurs:

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

For three events. The probability that at least one of A, B and C occurs:

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \\ - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)\end{aligned}$$

In general: the union of $A_1, \ldots, A_n$ (at least one event occurs) has probability

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_n) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \\ - \ldots - \mathbb{P}(A_{n-1} \cap A_n) + \mathbb{P}(A_1 \cap A_2 \cap A_3) + \ldots\end{aligned}$$

Inclusion-exclusion formula

The probability that at least one of the events $A_1, A_2, \dots, A_n$ occurs:

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_n) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \\ - \ldots - \mathbb{P}(A_{n-1} \cap A_n) + \mathbb{P}(A_1 \cap A_2 \cap A_3) + \ldots$$

That is:

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) &= \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{1 \le i_1 < i_2 \le n} \mathbb{P}(A_{i_1} \cap A_{i_2}) + \\ &+ \sum_{1 \le i_1 < i_2 < i_3 \le n} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \\ &- \ldots + (-1)^{n+1} \mathbb{P}(A_1 \cap \ldots \cap A_n) = \\ &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}).\end{aligned}$$

Inclusion-exclusion formula: example

Peter collects toys that are sold in a chocolate egg. There are ten types of toys, each egg contains one uniformly at random. What is the probability that his collection is complete (contains all ten types) after buying 20 chocolate eggs?

Inclusion-exclusion formula: example

Peter collects toys that are sold in a chocolate egg. There are ten types of toys, each egg contains one uniformly at random. What is the probability that his collection is complete (contains all ten types) after buying 20 chocolate eggs?

Let A_i be the event that type i is missing from the collection ($i = 1, \ldots, 10$). The probability that all types are present:

Inclusion-exclusion formula: example

Peter collects toys that are sold in a chocolate egg. There are ten types of toys, each egg contains one uniformly at random. What is the probability that his collection is complete (contains all ten types) after buying 20 chocolate eggs?

Let A_i be the event that type i is missing from the collection ($i = 1, \ldots, 10$). The probability that all types are present: $1 - \mathbb{P}(A_1 \cup \ldots \cup A_{10})$.

Inclusion-exclusion formula: example

Peter collects toys that are sold in a chocolate egg. There are ten types of toys, each egg contains one uniformly at random. What is the probability that his collection is complete (contains all ten types) after buying 20 chocolate eggs?

Let A_i be the event that type i is missing from the collection ($i = 1, \ldots, 10$). The probability that all types are present: $1 - \mathbb{P}(A_1 \cup \ldots \cup A_{10})$.

First term of the inclusion-exclusion formula:

$$\mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_{10}) = 10 \cdot \frac{9^{20}}{10^{20}},$$

because A_j means that j cannot occur, but all the other 9 can be chosen.

Inclusion-exclusion formula: example

Similarly, the next sum, with a minus sign:

$$\mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_3) + \ldots + \mathbb{P}(A_9 \cap A_{10}) = \binom{10}{2} \cdot \frac{8^{20}}{10^{20}},$$

because the number of terms is the same as the pairs from $1, 2, \ldots, 10$, and $A_i \cap A_j$ means that i and j cannot occur, so we can choose from 8 types for every chocolate egg.

Inclusion-exclusion formula: example

Similarly, the next sum, with a minus sign:

$$\mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_3) + \ldots + \mathbb{P}(A_9 \cap A_{10}) = \binom{10}{2} \cdot \frac{8^{20}}{10^{20}},$$

because the number of terms is the same as the pairs from $1, 2, \ldots, 10$, and $A_i \cap A_j$ means that i and j cannot occur, so we can choose from 8 types for every chocolate egg.

Altogether:

$$\mathbb{P}(A_1 \cup \ldots \cup A_{10}) = 10 \cdot \frac{9^{20}}{10^{20}} - \binom{10}{2} \cdot \frac{8^{20}}{10^{20}} + \binom{10}{3} \cdot \frac{7^{20}}{10^{20}} - \ldots =$$

$$= \sum_{k=1}^{10} (-1)^{k+1} \binom{10}{k} \frac{(10-k)^{20}}{10^{20}} = 78,53\%.$$

Inclusion-exclusion formula: example

Similarly, the next sum, with a minus sign:

$$\mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_3) + \ldots + \mathbb{P}(A_9 \cap A_{10}) = \binom{10}{2} \cdot \frac{8^{20}}{10^{20}},$$

because the number of terms is the same as the pairs from $1, 2, \ldots, 10$, and $A_i \cap A_j$ means that i and j cannot occur, so we can choose from 8 types for every chocolate egg.

Altogether:

$$\mathbb{P}(A_1 \cup \ldots \cup A_{10}) = 10 \cdot \frac{9^{20}}{10^{20}} - \binom{10}{2} \cdot \frac{8^{20}}{10^{20}} + \binom{10}{3} \cdot \frac{7^{20}}{10^{20}} - \ldots =$$

$$= \sum_{k=1}^{10} (-1)^{k+1} \binom{10}{k} \frac{(10-k)^{20}}{10^{20}} = 78,53\%.$$

This was the probability of the event that at least one type is missing, and hence the probability that the collection is complete is $100 - 78,53 = 21,47\%$.

Independence

Which events can be considered independent? Anne is a randomly chosen participant for a survey.

Anne has a car

rain in Buda tomorrow

Anne's income is more than the average

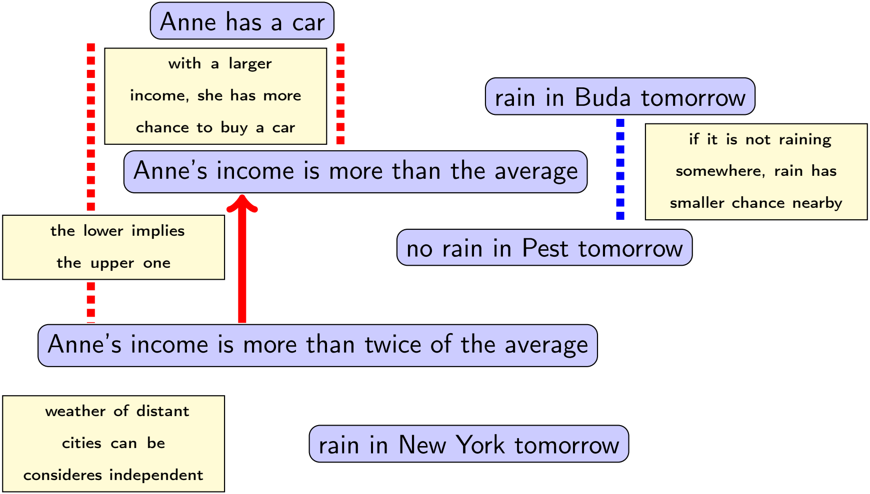
no rain in Pest tomorrow

Anne's income is more than twice of the average

rain in New York tomorrow

Independence

Which events can be considered independent? Anne is a randomly chosen participant for a survey.



Anne has a car

with a larger income, she has more chance to buy a car

rain in Buda tomorrow

if it is not raining somewhere, rain has smaller chance nearby

Anne's income is more than the average

the lower implies the upper one

no rain in Pest tomorrow

Anne's income is more than twice of the average

weather of distant cities can be consideres independent

rain in New York tomorrow

Independent events

Events $A, B \in \mathcal{A}$ are **independent**, if

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

that is, **the probability of the intersection is the product of the probabilities**.

Independent events

Events $A, B \in \mathcal{A}$ are **independent**, if

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

that is, **the probability of the intersection is the product of the probabilities**.

For several events, all subsets have to satisfy this property.

Events $A_1, A_2, \ldots \in \mathcal{A}$ are **independent**, if for every $k \geq 1$ and $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ we have

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \ldots \mathbb{P}(A_{i_k}).$$

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **second roll is 6**?

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **second roll is 6**?

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

A: first roll is 6; **B**: second roll is 6; $A \cap B$: both rolls are 6

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **second roll is 6**?

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

A: first roll is 6; **B**: second roll is 6; A ∩ B: both rolls are 6

$$\frac{1}{36} = \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) = \frac{1}{6} \cdot \frac{1}{6}.$$

Events A and B are **independent**.

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **the sum is** 10?

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **the sum is** 10?

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

A: first roll is 6; **C**: the sum is 10; $A \cap C$: the first roll is 6, the second one is 4

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **the sum is** 10?

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

A: first roll is 6; **C**: the sum is 10; A ∩ C: the first roll is 6, the second one is 4

$$\frac{1}{36} = \mathbb{P}(A \cap C) \neq \mathbb{P}(A) \cdot \mathbb{P}(C) = \frac{1}{6} \cdot \frac{1}{12}.$$

Events A and C are **not independent**.

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **the sum is** 7?

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **the sum is** 7?

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

A: first roll is 6; **D**: second roll is 7; $A \cap D$: first roll is 6, second is 1

Independence: example

We take two fair dice rolls. Are the following events independent: **first roll is 6**; **the sum is** 7?

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

A: first roll is 6; **D**: second roll is 7; A ∩ D: first roll is 6, second is 1

$$\frac{1}{36} = \mathbb{P}(A \cap D) = \mathbb{P}(A) \cdot \mathbb{P}(D) = \frac{1}{6} \cdot \frac{1}{6}.$$

Events *A* and *D* are **independent**.