Slides

Probability Theory

Chapter 1

Where do we use Probability Theory for?

Ch1: Introduction to the probability concept

Organization of this part of the module

- **Combined tutorials and lectures ("colstructions"):** First questions about the home made exercises will be discussed, where after there is time to solve exercises Second hour: new theory illustrated with examples. At home: make a summary and solve exercises.

- **4 homework assignments: each HWA will count 5% in the end result,** if beneficial for the student**) .** HWA's are **not mandatory**: submit your own solutions (no plain copying!) on one A4 and do not send e-mails Aim: pass the course by regularly working on it.

- **Blackboard-info:** slides, solutions, assignments, sample tests with solutions, study aids, etc.

Goals of this course

- **Giving proper models of real life situations where "chance" plays a role: the** model enables as to compute desired probabilities and expected values.
- **Probability Theory:** if the probability model is fully specified we can, for example, compute the probability of certain outcomes of experiments.
  Application e.g. in performance analysis of computer systems.
- Probability Theory is also a preparation for **Statistics:**
  if the model is not fully known, we will try to make statements about the population on the basis of a (small) random sample, drawn from the (large) population.
  Concepts as Level of confidence or significance are based on probability statements.
  Applications: users tests and simulation of performance

Model: sample space S and probability P

- **<u>Example 1</u>: one roll with a dice:**
  Outcomes are 1, 2, 3, 4 , 5 or 6 → $S = \{1, 2, \ldots, 6\}$
  Event A = "even number" = {2, 4, 6}
  A occurs in 3 out of 6 outcomes: $P(A) = \frac{3}{6}$

  Condition for the correctness of this calculation:
  The dice is "fair": every outcome is **equally likely**.

- **<u>Example 2</u>:** We observe the times (in *min*)between the consecutively arriving customers: 24 interarrival times.

$$S = \{(x_1, \ldots, x_{24}) \mid x_i \geq 0, i = 1, 2, \ldots, 24\}$$

A = "the totale interval of time is greater than 50 *min*."

$$S = \{(x_1, \ldots, x_{24}) \; in \; S \;\mid\; \sum_{i=1}^{24} x_i \geq 50\}$$

Concepts (1)

- **Stochastic experiment:** outcome depends on chance
- **Sample space *S*** = {all outcomes}: an outcome ***s*** ∈ ***S***
- **Event *A* ⊂ *S*:** *A* is a subset of *S*

  Special events:

  The largest is ***S*** itself: the **certain** event and

  the smallest is ø, the **impossible** event

  *Note that in this course **A** ⊂ **S** means the same as A ⊆ **S***

  (math A)

- Properties: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$A \cup B = A \cup \left(\overline{A} \cap B\right)$ and $A = (A \cap B) \cup \left(A \cap \overline{B}\right)$

De Morgan`s rules:

$\overline{A \cup B} = \overline{A} \cap \overline{B}$ and $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Concepts (2)

- The ***P***robability of an event A: $\boldsymbol{P(A)}$
- Evident properties that should be true:

$$\mathbf{0 \leq P(A) \leq 1}, \qquad \boldsymbol{P(ø) = 0} \text{ and } \boldsymbol{P(S) = 1}$$

- **the frequency-interpretation** of e.g. $P(A) = 0.80$:

"If the stochastic experiment is repeated many times, *A* will occur in about in 80% of the repetions".

- This notion reflects an "experimentally determined probability": the **relative frequency of an event *A*** is the proportion of occurrences of A in n repetitions of the experiment: $\boldsymbol{f_n(A) \approx \frac{n(A)}{n}}$

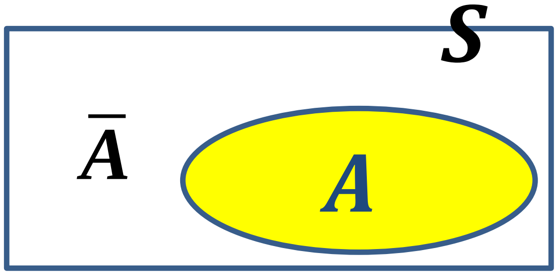
- **Experimental law of large numbers**

for large n we have $\boldsymbol{f_n(A) \approx P(A)}$

Basic rules of probability (1)

- The **complement** of *A*: $\overline{A} = \{s \in S |\ s \notin A\}$

  say "not-A"



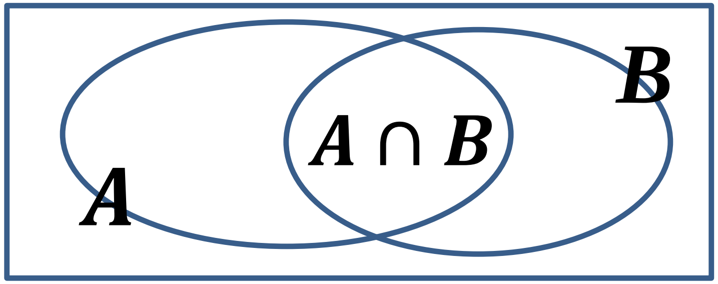
- The **complement rule:**

  $$P(\overline{A}) = 1 - P(A)$$

- The **intersection** $A \cap B = AB$:

  set of outcomes in *A* and *B*.

  say "A ànd B"



- The **union** of *A* and *B*: $A \cup B$

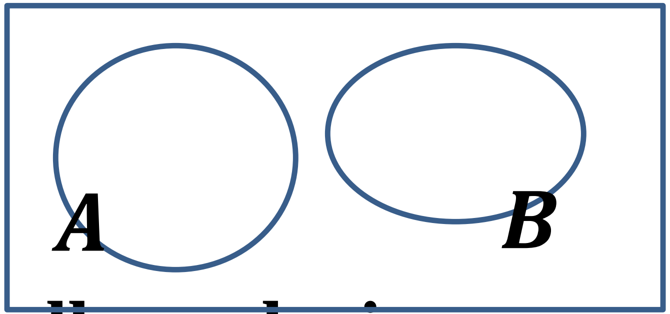
  say "A or B (or both)"

- **General addition rule:**

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

Basic rules of probability (2)

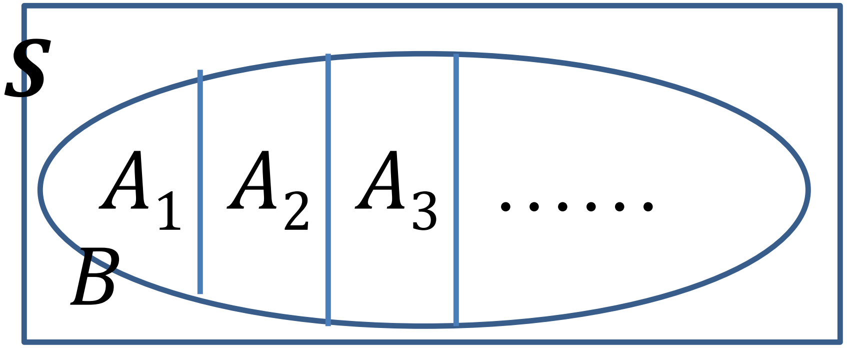
- **Mutually exclusive** or **disjoint** events.
  $\boldsymbol{A \cap B = \emptyset}$, so $\boldsymbol{P(A \cap B) = 0}$



- (Special) **addition rule for mutually exclusive events:**
- A series of events $\boldsymbol{A_1, A_2}$, … **are mutually exclusive** if $\boldsymbol{A_i \cap Aj = \emptyset}$**,** for each pair (i, j)
- A **partition** of B:

$$B = \bigcup_i A_i$$

such that A_1, A_2, …
are mutually exclusive.



Computing probabilities if all **outcomes are equally likely**: **count** the favourable and the total number.

- This principle can be applied often when random samples are conducted: an arbitrary choice from a finite population: every outcome of the population is equally likely.
- **Probability definition of Laplace**: compute the probability of A by counting the proportion of favourable outcomes.

$$\boldsymbol{P}(\boldsymbol{A}) = \frac{\boldsymbol{N}(\boldsymbol{A})}{\boldsymbol{N}(\boldsymbol{S})} = \frac{\textbf{Number of favourable outcomes of A}}{\textbf{Total number of outcomses in S}}$$

- A pair **(S, P)** of a finite sample space S and its Laplace probability P is called a **symmetric probability space.**

What conditions should we impose on P for any (arbitrary) probability space (S, P) , e.g. if $N(S)$ is finite?

Foundation Probability Theory: **Kolmogorov`s axioms**

A function ***P***, that assigns a real number ***P*(*A*)** to each event ***A*** **Probability (measure)** if:

1. $\boldsymbol{P(A) \geq 0}$
2. $\boldsymbol{P(S) = 1}$
3. If $A_1, \dots, A_n$ or $A_1, A_2, \dots$ are mutually exclusive, then

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

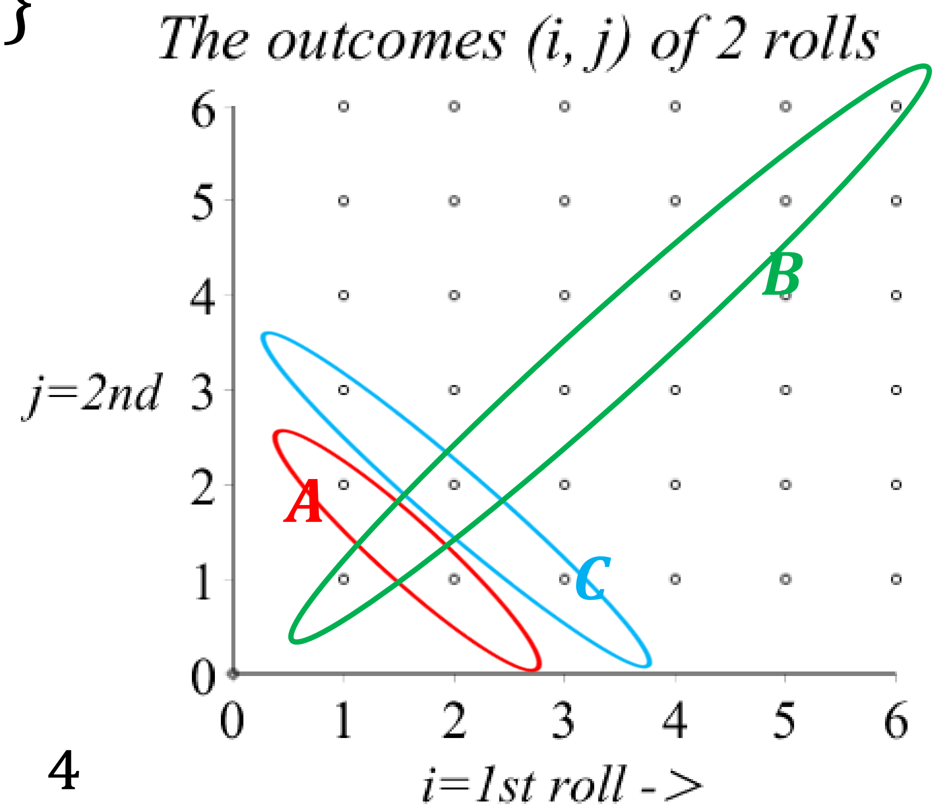
All other desirable properties can be derived from the axioms:

- $\boldsymbol{P(\emptyset) = 0}$ and $\boldsymbol{0 \leq P(A) \leq 1}$
- $\boldsymbol{P(\overline{A}) = 1 - P(A)}$
- $\boldsymbol{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$
- **If $\boldsymbol{A \subset B}$, then $\boldsymbol{P(A) \leq P(B)}$**

Examples (1)

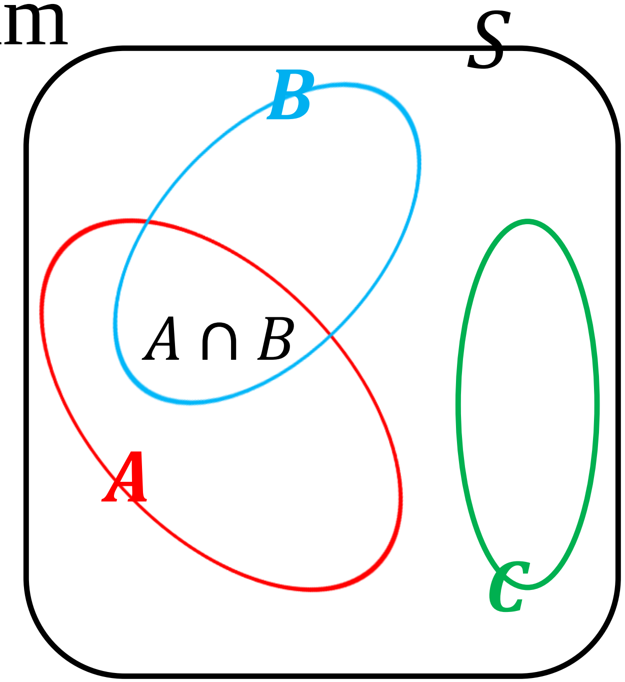
Example: roll a dice twice:

- $S = \{(i, j) | i, j = 1,2, \dots, 6\}$
- $A =$ "total $= 4$"
  $= \{(1,3), 2,2), (3,1)\}$
  $P(A) = \frac{N(A)}{N(S)} = \frac{3}{36}$
- $B =$ "$1^{st} = 2^{nd}$ *roll*"
  $P(B) = \frac{6}{36}$
- $C =$ "total $= 5$": $P(C) = \frac{4}{36}$
- $P(\overline{A}) = \frac{33}{36} = 1 - P(A)$



Examples (2)

Example (continues): Venn-diagram



- $P(A \cap B) = P\big((2,2)\big) = \frac{1}{36}$
- $P(A \cup B) = P(A) + P(B) - P(AB)$
  $\frac{8}{36} = \frac{3}{36} + \frac{6}{36} - \frac{1}{36}$
- A and C are mutually exclusive,
  B and C as well,
  but A and B are not: $P(AB) \neq 0$
- Since $A = AB \cup A\overline{B}$, we have: $P(A) = P(AB) + P(A\overline{B})$,
  so $P\big(A\overline{B}\big) = P(A) - P(AB) = \frac{3}{36} - \frac{1}{36} = \frac{2}{36}$
  (correct, since $A\overline{B} = \{(1,3), (3,1)\}$)

Examples (3)

Example: The probability of a prize in a lottery is (about) 10% at each draw

a. What is the probability of the first win at the tenth draw?

b. What is the probability that the first win after the 20th draw?

Solutions: each outcome is the number of draws, needed for the first win: $S = \{1, 2, 3, \ldots\ldots\}$

a. $P(10) = 0.9 \cdot \cdots \cdot 0.9 \cdot 0.1 = 0.9^9 \cdot 0.1 \approx 3.9\%$

Reasoning: at each draw you will have a success probability of 0.1 and a 0.9 probability of no prize. "10" will occur if the in the first 9 draws no prize is won and at the tenth draw a prize is won.

b. $A = \{21, 22, \ldots\}$, so $P(A) = 0.9^{20} \approx 12.2\%$

A occurs if in all first 20 draws you win no prize.