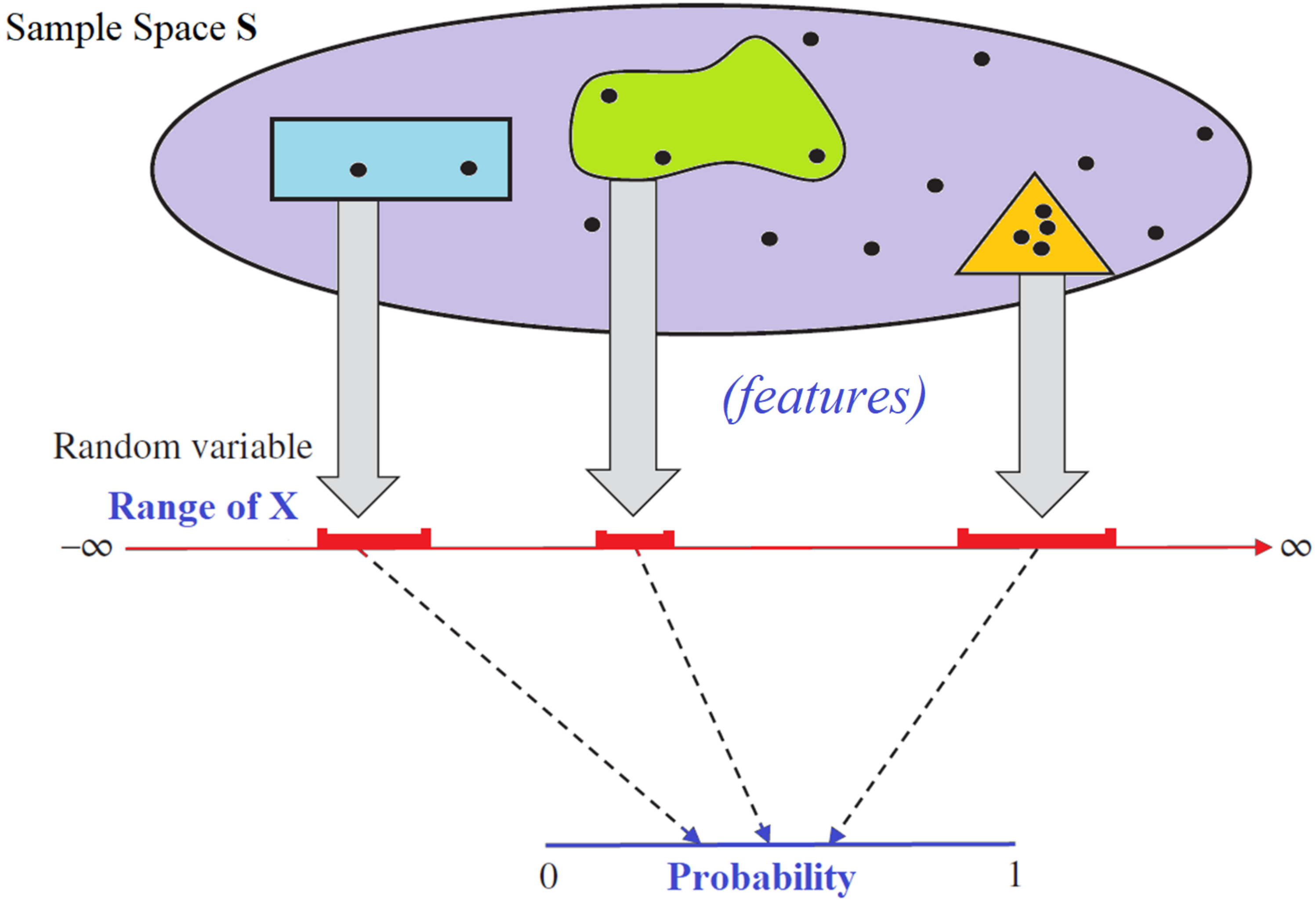
# Stochastic Process

## Basic Concepts of Probability Theory
## Module 1

Electrical Engineering Department
University of Brasília – UnB
Prof. Francisco Assis de O. Nascimento

# The universe of stochastic processes: The relationship among sample space, random variable, and probability.

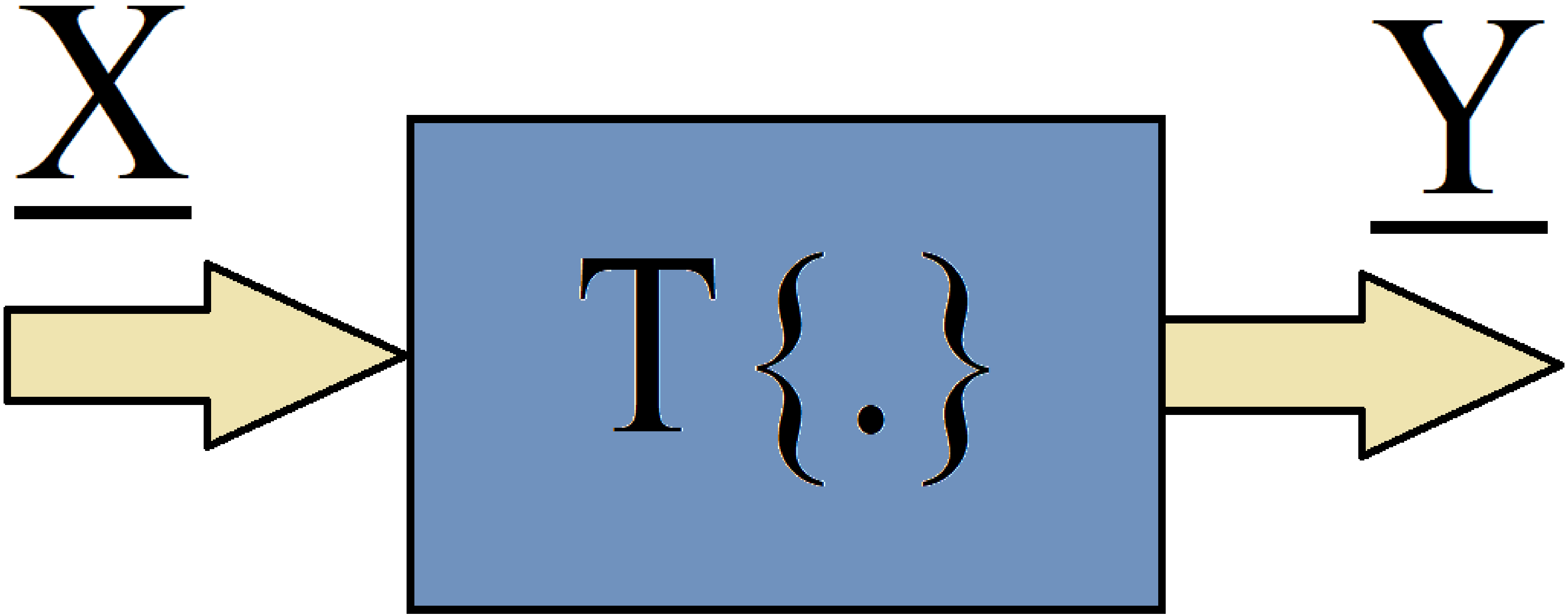


# Data and Signal

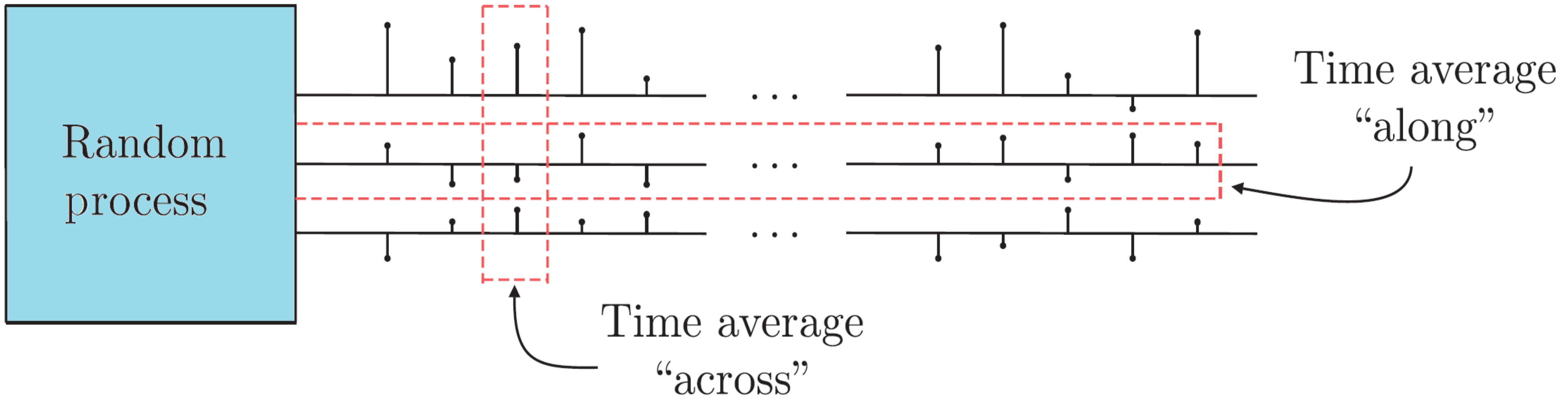
- **Data**: The data usually consists of a sequence where its elements are not related to each other by the temporal variable, a spatial variable or both.

- **Signal**: The signal is related to the variable “time”, “space”, both or other related quantity.

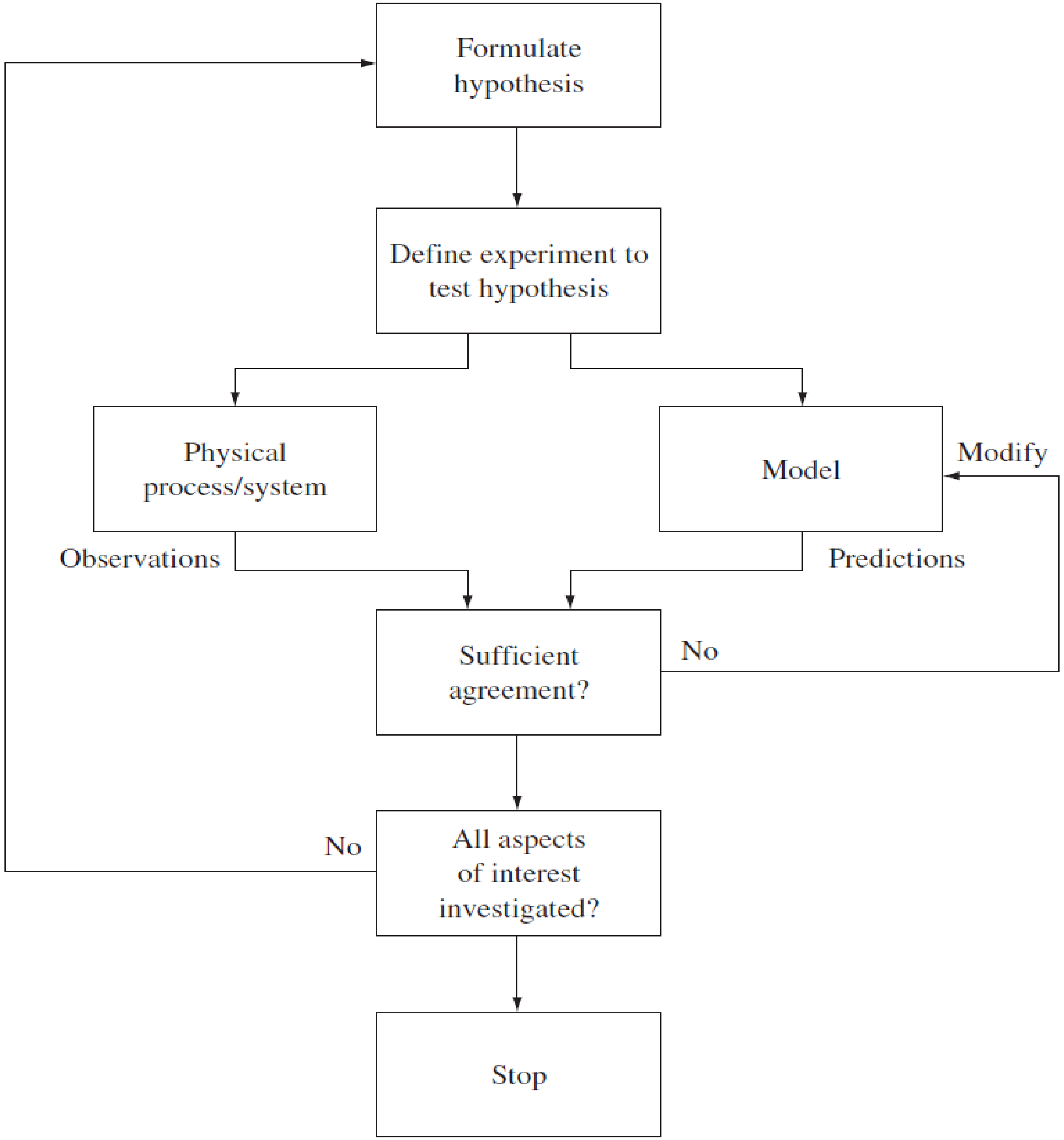
# A (Stochastic) System Overview



# How to address the problem



# Modeling Process



# Introduction: Set Theory

- Set
- Subset (A está contido em B: $A \subset B$)
- Superset (B contém A: $B \supset A$)
- Venn Diagrams
- Set Operations: union , intersection, complement, difference (subtraction) disjoint (mutually exclusive), partition, De Morgan's law, Distributive law, cartesian product (multiplication principle)
- Cardinality: Countable and Uncountable Sets
- Functions

# Random Experiments – probability concepts

- Sample space: Outcome, Event.
- Statistical regulariry behavior: number of ocorrences of any outcome in N trials, relative frequency of outcome.
- Probability: Definition of probability, Axioms of Probability, Computing Probabilities, Inclusion-Exclusion Principle.
- Discrete Probability Models.
- Continuous Probability Models.
- Conditional Probability: Chain rule for conditional probability
- Statistical Independence.

# Fiding probabilities

- Law of Total Probability.
- Bayes' Rule.
- Conditional Independence.
- Sampling: Multiplication Principle.

# A **set** is a collection of some items (countable elements).

- $A = \{♣, ◊\}$. Note that ordering does not matter, so the two sets $\{♣, ◊\}$ and $\{◊, ♣\}$ are equal. For example, we may write $♡ \notin A$. Card deck: clubs (♣), diamonds (♦), hearts (♥) and spades (♠) (in Portuguese, espadas(♠), paus(♣), copas(♥) e ouro(♦)).

- The set of natural numbers, $\mathrm{N} = \{1, 2, 3, \cdots\}$.

A **set** is a collection of some items (countable elements).

- The set of integers: $Z = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$.
- The set of rational numbers Q.

# A **set** is a collection of some items (uncountable elements).

- The set of real numbers R.

- Closed intervals on **the real line**. For example, [2, 3] is the set of all real numbers $x$ such that $2 \leq x \leq 3$.

- Open intervals on **the real line**. For example $(-1, 3)$ is the set of all real numbers $x$ such that $-1 < x < 3$.

A **set** is a collection of some items (uncountable elements).

- Similarly, [1, 2) is the set of all real numbers $x$ such that $1 \leq x < 2$.
- The set of complex numbers C is the set of numbers in the form of $(a + \mathrm{j}b)$, where $a, b \in$ R, and $\mathrm{j} = \sqrt{-1}$.

# Mathematical notation

- $A = \{x | x \text{ satisfies some property}\}$
- $A = \{x : x \text{ satisfies some property}\}$
- The symbols "|" and ":" are pronounced "such that (tal que)".

# Examples

- $C = \{x | x \in \mathrm{Z}, -2 \leq x < 10\}$, then $C = \{-2, -1, 0, \cdots, 9\}$.

- $D = \{x^2 | x \in \mathrm{N}\}$, then $D = \{1, 4, 9, 16, \cdots\}$.

- $\mathrm{Q} = \{(\mathrm{a/b}) \,|\, a, b \in \mathrm{Z}, b \neq 0\}$.

- For real numbers $a$ and $b$, where $a < b$, we can write $(a, b] = \{x \in \mathrm{R} \mid a < x \leq b\}$.

- $\mathrm{C} = \{a + \mathrm{j}b \mid a, b \in \mathrm{R}, j = \sqrt{-1}\}$.

# A **subset**

- Set *A* is a **subset** of set *B* if every element of *A* is also an element of *B*.

- We write $\boldsymbol{A \subset B}$, where “$\subset$” indicates “subset”.

- In Portuguese: “$\subset$” indicates “**está contido**”. It means: “**A está contido em B”**.

# Examples: A subset

- If $E = \{1, 4\}$ and $C = \{1, 4, 9\}$, then $E \subset C$.

- $\mathrm{N} \subset \mathrm{Z}$.

- $\mathrm{Q} \subset \mathrm{R}$.

- $\emptyset = \{\}$ is the **null set** or the **empty set**. For any set $A$, $\emptyset \subset A$.

# A superset

- If *set* $B$ is a **superset** of $A$, we can write $B \supset A$.

- In Portuguese: " $\supset$ " indicates "**contém**". It means: **B contém A**.

# Examples: A superset

- If $\boldsymbol{E} = \{1, 4\}$ and $\boldsymbol{C} = \{1, 4, 9\}$, then $\boldsymbol{C} \supset \boldsymbol{E}$.
- $Z \supset N$.
- $R \supset Q$.
- $\emptyset = \{\}$ is the **null set** or the **empty set**. For any set $A$, $A \supset \emptyset$.

# Equal sets

- Two sets are equal if they have the exact same elements.

- $A = B$ if and only if $\boldsymbol{A \subset B}$ and $\boldsymbol{B \subset A}$.

- Example:
  - $\{1, 2, 3\} = \{3, 2, 1\}$
  - $\{a, a, b\} = \{a, b\}$

# Universal set (sample set)

- Universal set is the **set** of all things that we could possibly consider in the context we are studying.

- Every **set A** is a **subset** of the **universal set**.

- In **probability theory** the **universal set** is called:

**sample set *S*** (**espaço amostral**).

# Examples of **Universal set (sample set)**

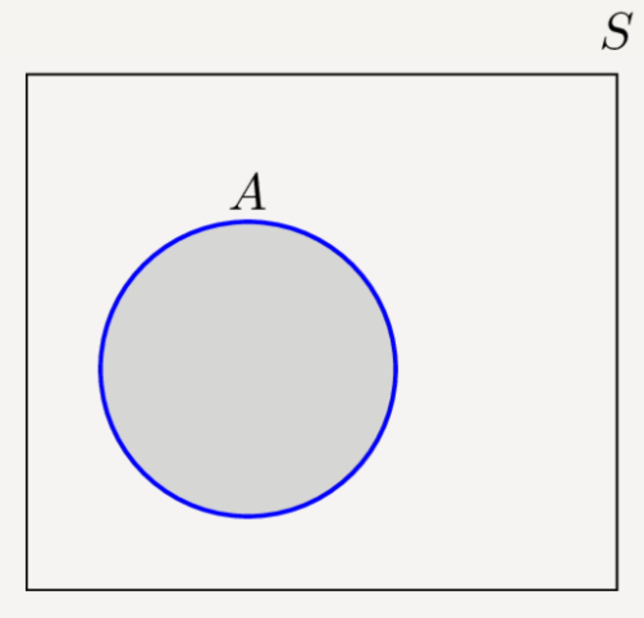
- Rolling a die - **sample set (espaço amostral de jogo dados):**

$$\boldsymbol{S} = \{1, 2, 3, 4, 5, 6\}$$

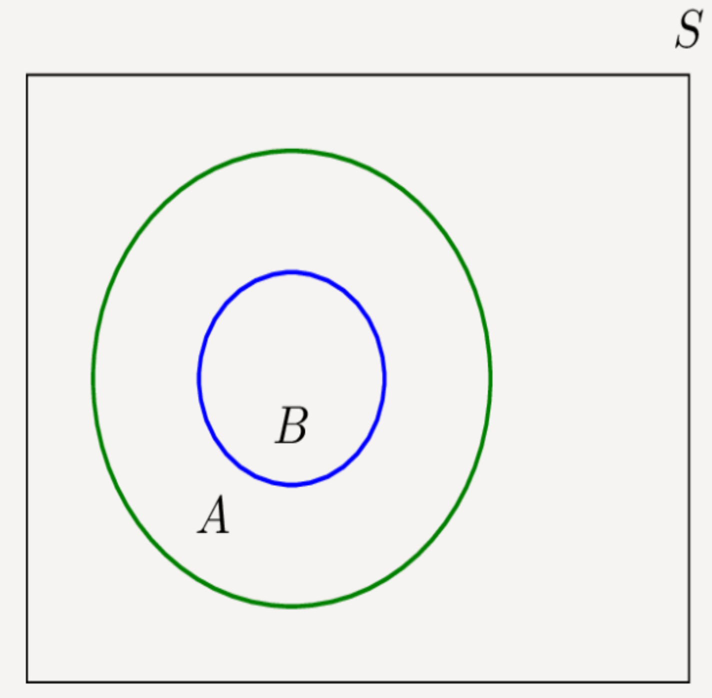
- Tossing a coin - **sample set** (espaço amostral de cara ou coroa):

$$\boldsymbol{S} = \{H, T\} \text{ (}H\text{ for heads and }T\text{ for tails)}$$

# Venn Diagrams




A – A subset
(**an event**)

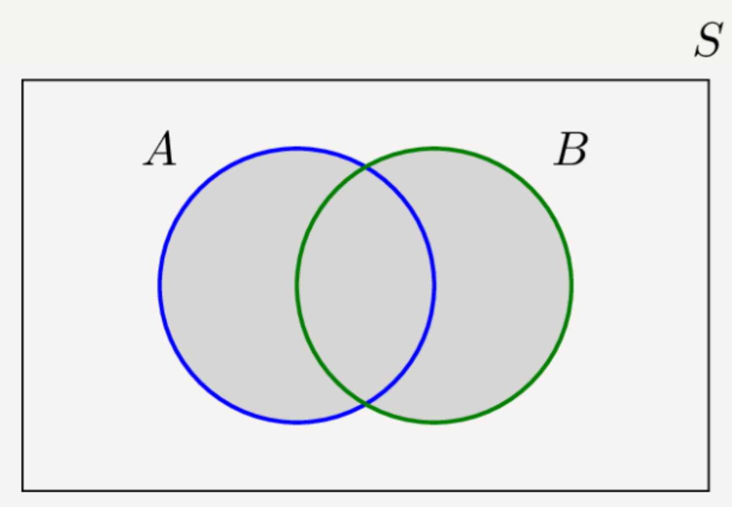



S - sample set

$B \subset A$

# Venn Diagrams: **Set** operations: **union** of two sets

It is ludic.




$B \cup A$

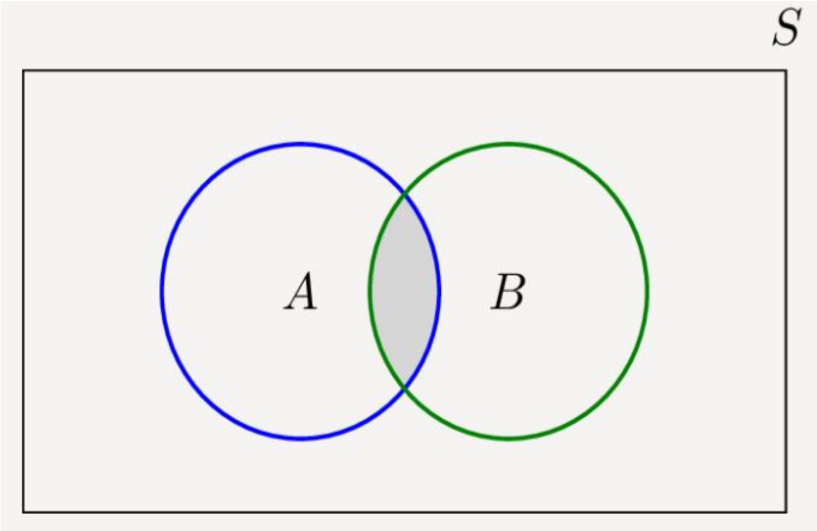
- Analogously if A={1, 2, 3} and B = {2, 3, 5, 6}

$x \in (A \cup B) \Leftrightarrow \{x | x \in A \text{ or } x \in B\}$ ("$\Leftrightarrow$" means "***if only if***")

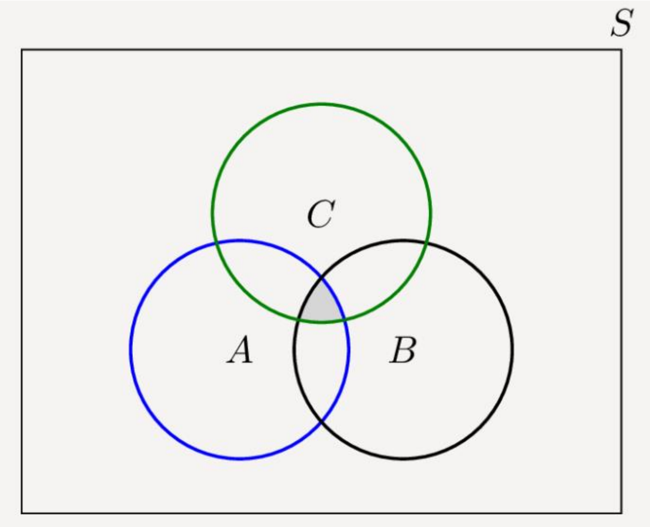
Thus, $C = (A \cup B) = \{1, 2, 3, 5, 6\}$

- Similarly, we can generalize: $A_1 \cup A_2 \cup A_3 \dots A_N = \bigcup_{k=1}^{N} A_k$

# **Set** operations: **intersection** of two sets




$B \cap A$




$A \cap B \cap C$

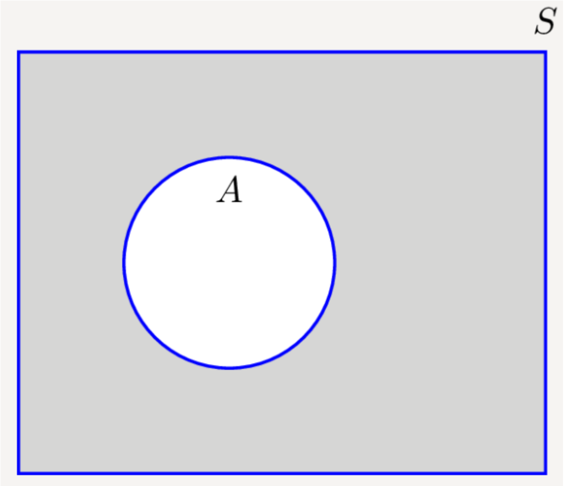
- Analogously if A={1, 2, 3} and B = {2, 3, 5, 6}

$$x \in (A \cap B) \Leftrightarrow \{x | x \in A \ \boldsymbol{and} \ x \in x \in B\}$$

Thus, $C = (A \cap B) = \{\, 2, 3\}$

- Similarly, $A_1 \cap A_2 \cap A_3 \ldots A_N = \bigcap_{k=1}^{N} A_k$

# Set operations: complement of a set




$$\bar{A} = A^c$$

- $A^c = S - A$

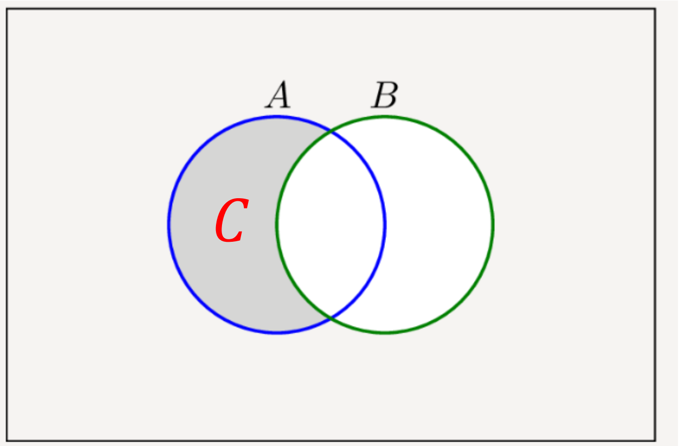
# Set operations: difference of two sets

$$C = (A - B) = \{x | x \in A \textbf{ and } x \notin B\} = A \cap B^c$$

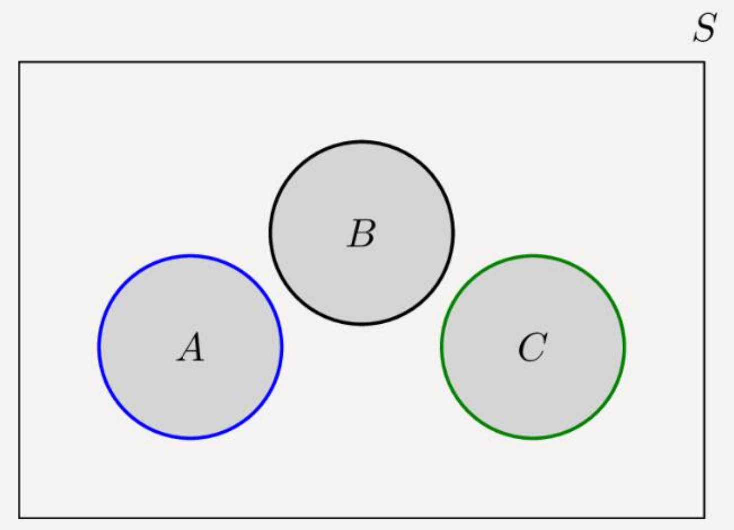
- If A={1, 2, 3} and B = {2, 3, 5, 6}

$$C = (A - B) = \{1\}$$

$$C = A - B$$



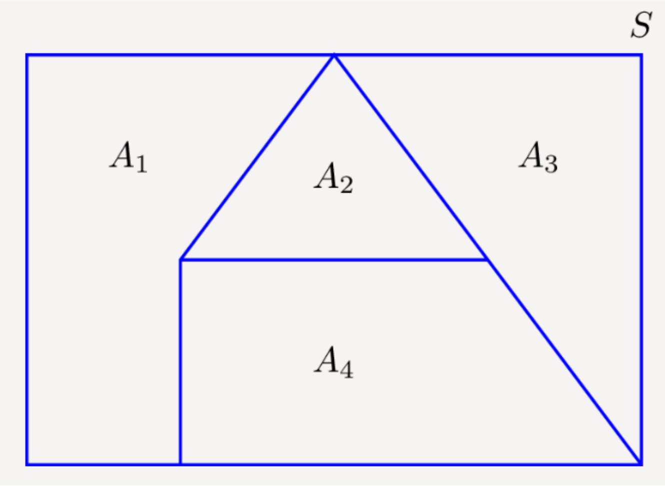
# **Mutually exclusive** or **disjoint** sets (conjuntos **mutuamente exclusivos** ou **disjuntos**)




$$A \cap B \cap C = \phi$$

## **Partition** of a set

$$\bigcup_{k=1}^{N} A_k = S \quad \bigcap_{k=1}^{N} A_k = \phi$$



# De Morgan's law

$$(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_N)^C = A_1^C \cap A_2^C \cap A_3^C \cap \cdots \cap A_N^C$$

$$\boldsymbol{\left(\bigcup_{k=1}^{N} A_k\right)^c = \bigcap_{k=1}^{N} A_k^c}$$

$$(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_N)^C = A_1^C \cup A_2^C \cup A_3^C \cup \cdots \cup A_N^C$$

$$\left(\bigcap_{k=1}^{N} A_k\right)^c = \bigcup_{k=1}^{N} A_k^c$$

- Commutative law

$$A \cup B = B \cup A \quad \text{and} \quad A \cap B = B \cap A$$

- Associative law

$$A \cup (B \cup C) = (A \cup B) \cup C \quad \text{and} \quad A \cap (B \cap C) = (A \cap B)$$

- Distributive law

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \text{and}$$

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

# Example:

- If the universal set is given by $S = \{1, 2, 3, 4, 5, 6\}$ , and $A = \{1, 2\}$, $B = \{2, 4, 5\}$,$C = \{1, 5, 6\}$ are three sets, find the following sets:
  - a. $A \cup B$
  - b. $A \cap B$
  - c. $\bar{A}$
  - d. $\bar{B}$
  - e. Check De Morgan's law by finding

    $(A \cup B)^c$ and $A^c \cap B^c$.
  - f. Check the distributive law by finding

    $A \cap (B \cup C)$ and $(A \cap B) \cup (A \cap C)$.

Solution: $S = \{1, 2, 3, 4, 5, 6\}$; $A = \{1, 2\}$, $B = \{2, 4, 5\}$, $C = \{1, 5, 6\}$

- a. $A \cup B = \{1, 2, 4, 5\}$.
- b. $A \cap B = \{2\}$.
- c. $\bar{A} = \{3, 4, 5, 6\}$ ($\bar{A}$ consists of elements that are in $\boldsymbol{S}$ but not in $\boldsymbol{A}$).
- d. $\bar{B} = \{1, 3, 6\}$.
- e. We have: $\boldsymbol{(A \cup B)^c} = \{1, 2, 4, 5\}^c = \{3, 6\}$,

  which is the same as $\boldsymbol{A^c \cap B^c} = \{3, 4, 5, 6\} \cap \{1, 3, 6\} = \{3, 6\}$.
- f. We have $\boldsymbol{A \cap (B \cup C)} = \{1, 2\} \cap \{1, 2, 4, 5, 6\} = \{1, 2\}$,

  which is the same as $\boldsymbol{(A \cap B) \cup (A \cap C)} = \{2\} \cup \{1\} = \{1, 2\}$.

# Cartesian product

- A **Cartesian product** of two sets $\boldsymbol{A}$ and $\boldsymbol{B}$, written as $\boldsymbol{A \times B}$, is the set containing **ordered** pairs from $A$ and $B$.

- If $\boldsymbol{C = A \times B}$, then each element of $\boldsymbol{C}$ is of the form $\boldsymbol{(x, y)}$, where $\boldsymbol{x \in A}$ and $\boldsymbol{y \in B}$:

$$(A \times B) = \{(x, y) | \; x \in A \text{ and } y \in B\}$$

# Cartesian product

- Example: if $A = \{1, 2, 3\}$ and $B = \{H, T\}$, then

$$A \times B = \{(1, \mathrm{H}), (1, \mathrm{T}), (2, \mathrm{H}), (2, \mathrm{T}), (3, \mathrm{H}), (3, \mathrm{T})\}$$

- Note that here **the pairs are ordered**, so for our example, $(1, \mathrm{H}) \neq (\mathrm{H}, 1)$.

- Thus $\boldsymbol{A \times B}$ is **not** the same as $\boldsymbol{B \times A}$.

# Multiplication principle - Cardinality

- If you have two sets **A** and **B**, where a has **M** elements and B has N elements

$$|A \times B| = \text{MN}$$

- Note that the variables can be continuous, so:

$$if\ A = \{x|x \in R\}\ and\ B = \{x|x \in R\}\ \text{then}$$

$$(\text{A x B}) = \text{R}^2 = \text{R} \times \text{R} = \{(x, y)|x \in R, y \in R\}$$

# Cartesian product of $n$ sets $A_1, A_2, \cdots, A_n$

- $\boldsymbol{A_1 \times A_2 \times A_3 \times \cdots \times A_n} =$

$$\{(x_1, x_2, \cdots, x_n) \mid x_1 \in A_1 \text{ and } x_2 \in A_2 \text{ and } \cdots x_n \in A_n\}.$$

- The **multiplication principle** states that for finite sets $A_1, A_2, \cdots, A_n$, if

$$|\boldsymbol{A_1}| = \boldsymbol{M_1}, |\boldsymbol{A_2}| = \boldsymbol{M_2}, \cdots, |\boldsymbol{A_n}| = \boldsymbol{M_n},$$

then

$$|A_1 \times A_2 \times A_3 \times \cdots \times A_n| = M_1 M_2 M_3 \cdots M_n.$$

# Countable and Uncountable sets

- **Countable** sets = finite/infinite sets. (discrete variables)
- **Uncountable** sets = infinite sets (continuous variables)

- Set $A$ is called **countable** if one of the following is true:
  - if it is a finite set, $|A| < \infty$; or infinite
  - For example, it can be put in one-to-one correspondence with natural numbers N, in which case the set is said to be countably infinite.

# Countable and Uncountable sets

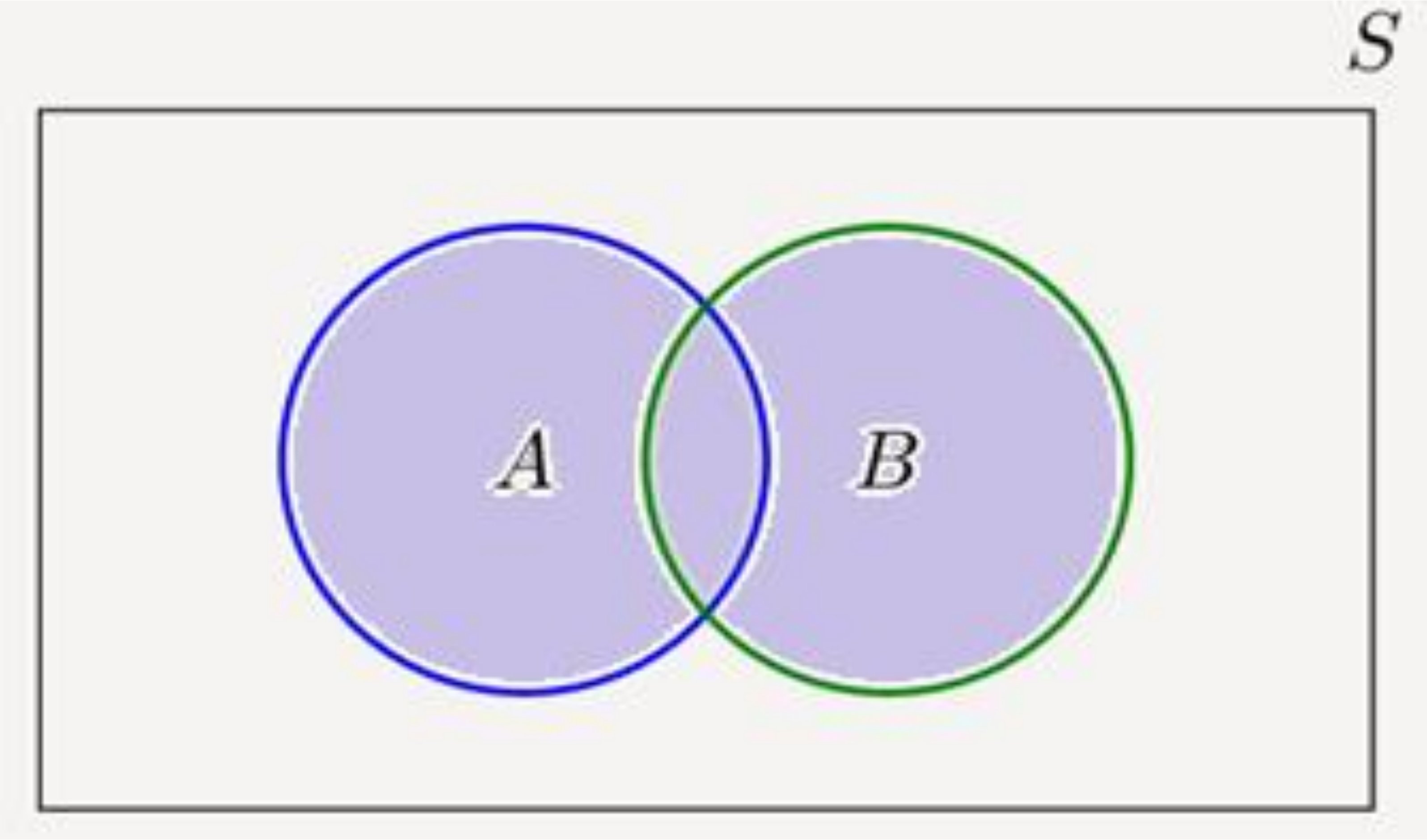
- Examples:
  - N, Z, Q, and any of their subsets are countable.
  - Any set containing an interval on the real line such as $[a, b]$, $(a, b]$, $[a, b)$, or $(a, b)$, where $a < b$ is uncountable.
- Any subset of a countable set is countable.
- Any superset of an uncountable set is uncountable.
- If $A_1, A_2, \cdots$ is a list of countable sets, then the set

$$\bigcup_{k=1}^{N} A_k = A_1 \cup A_2 \cup A_3 \ldots \cup A_N \cdots$$ is also countable.

- We call rational number every number obtained from the division (ratio) between two integers, with the divisor not zero. Every rational number can be written as a whole number, exact decimal, or repeating decimal. The way to represent a rational number is called a fraction.
- An irrational number is one that satisfies the definition, that is, a number that cannot be represented as a fraction.

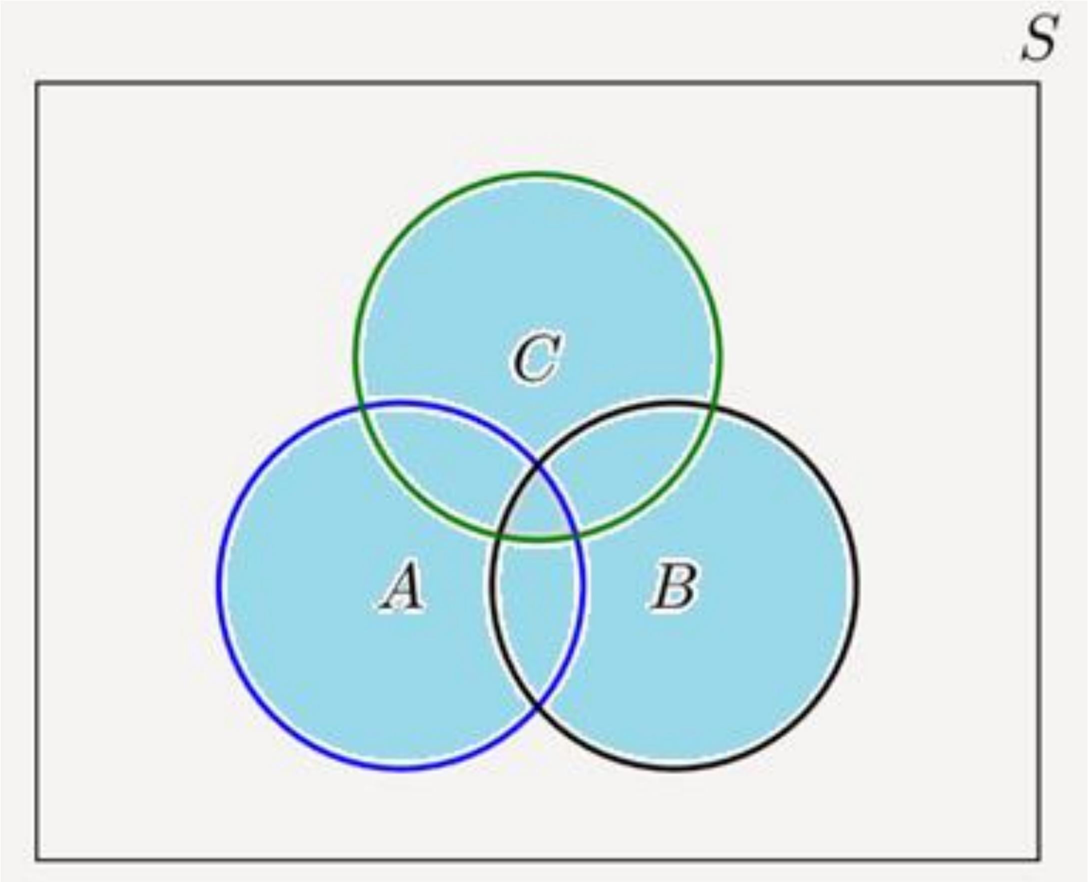
# Inclusion-exclusion principle

- $|A \cup B| = |A| + |B| - |A \cap B|$



# Inclusion-exclusion principle

- $|A \cup B \cup C| = |A|+|B|+|C|-|A \cap B|-|A \cap C|-|B \cap C|+|A \cap B \cap C|$.



# Inclusion-exclusion principle

Generally, for $n$ finite sets $A_1, A_2, A_3, \cdots, A_n$, we can write

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j|$$

$$+ \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap \cdots \cap A_n|.$$

# Exercise 1: In a class meeting

- a) there are **10 people** with **white shirts** **and** **8 people** with **red shirts**;
- b) **4 people** have **black shoes** **and** **white shirts**;
- c) 3 people have **black shoes** **and** **red shirts**;
- d) the **total number of people** with **white** **or** **red shirts** **or** **black shoes** is **21**.
- The question is : How many people have black shoes?

# Exercise 1: Solution

- Denoting:
    - White shirts – **W**
    - Red shirts – **R**
    - Black shoes – **B**
- Considering: There are no people with shirts that are both white and red (striped for example), so $|W \cap R| = 0$ and $|W \cap R \cap B| = 0$.
- What we know:
    - From a): $|W| = 10$; $|R| = 8$
    - From b): $|W \cap B| = 4$
    - From c): $|R \cap B| = 3$
    - From d): $|W \cup B \cup R| = 21$.

# Solution: Applying the inclusion-exclusion principle

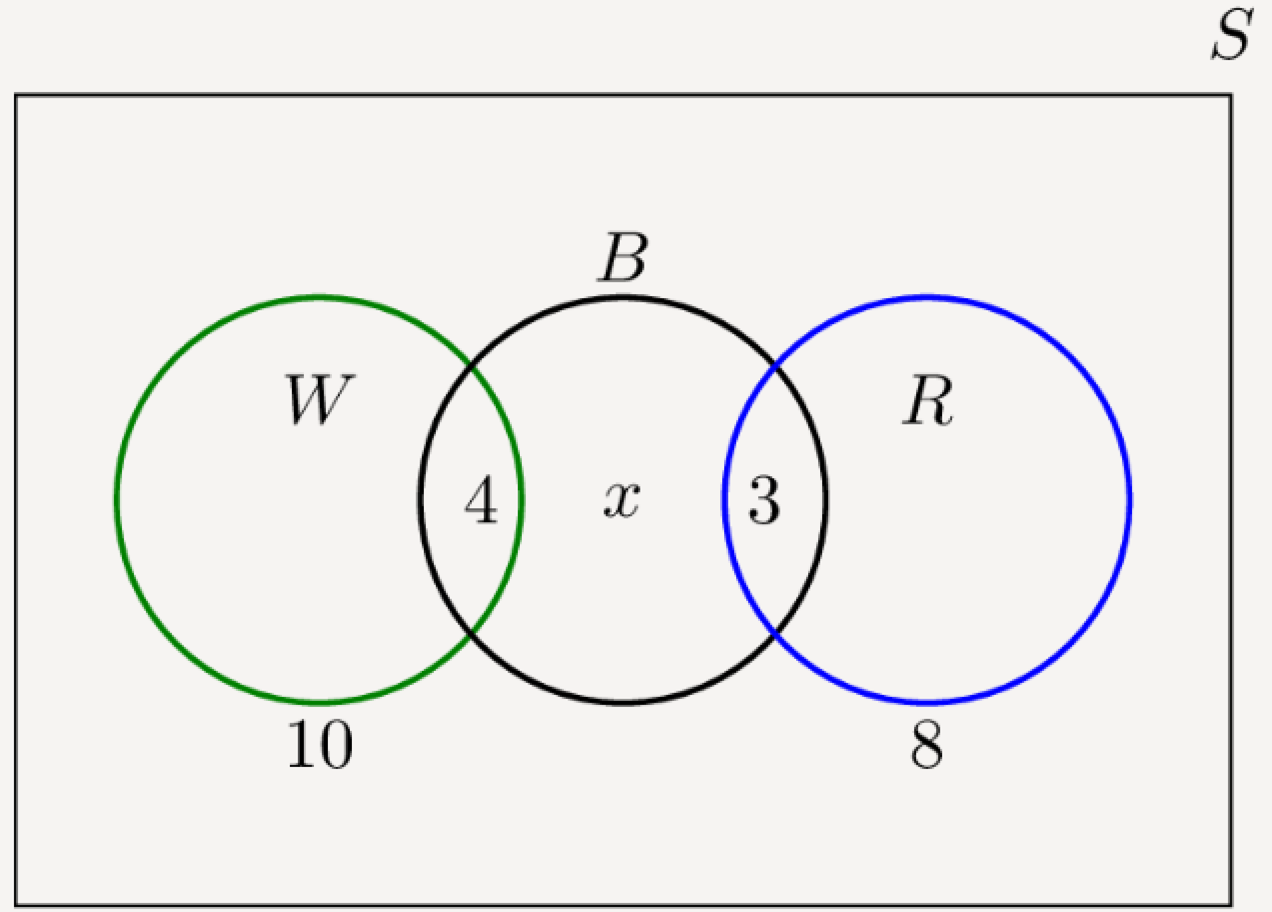
- $|\boldsymbol{W} \cup \boldsymbol{R} \cup \boldsymbol{B}| \quad = 21$

  $= |\boldsymbol{W}| + |\boldsymbol{R}| + |\boldsymbol{B}| - |\boldsymbol{W} \cap \boldsymbol{R}| - |\boldsymbol{W} \cap \boldsymbol{B}| - |\boldsymbol{R} \cap \boldsymbol{B}| + |\boldsymbol{W} \cap \boldsymbol{R} \cap \boldsymbol{B}|$

  $= 10 + 8 + |B| - 0 - 4 - 3 + 0 = 21$

- Thus $|\boldsymbol{B}| = \mathbf{10}$.

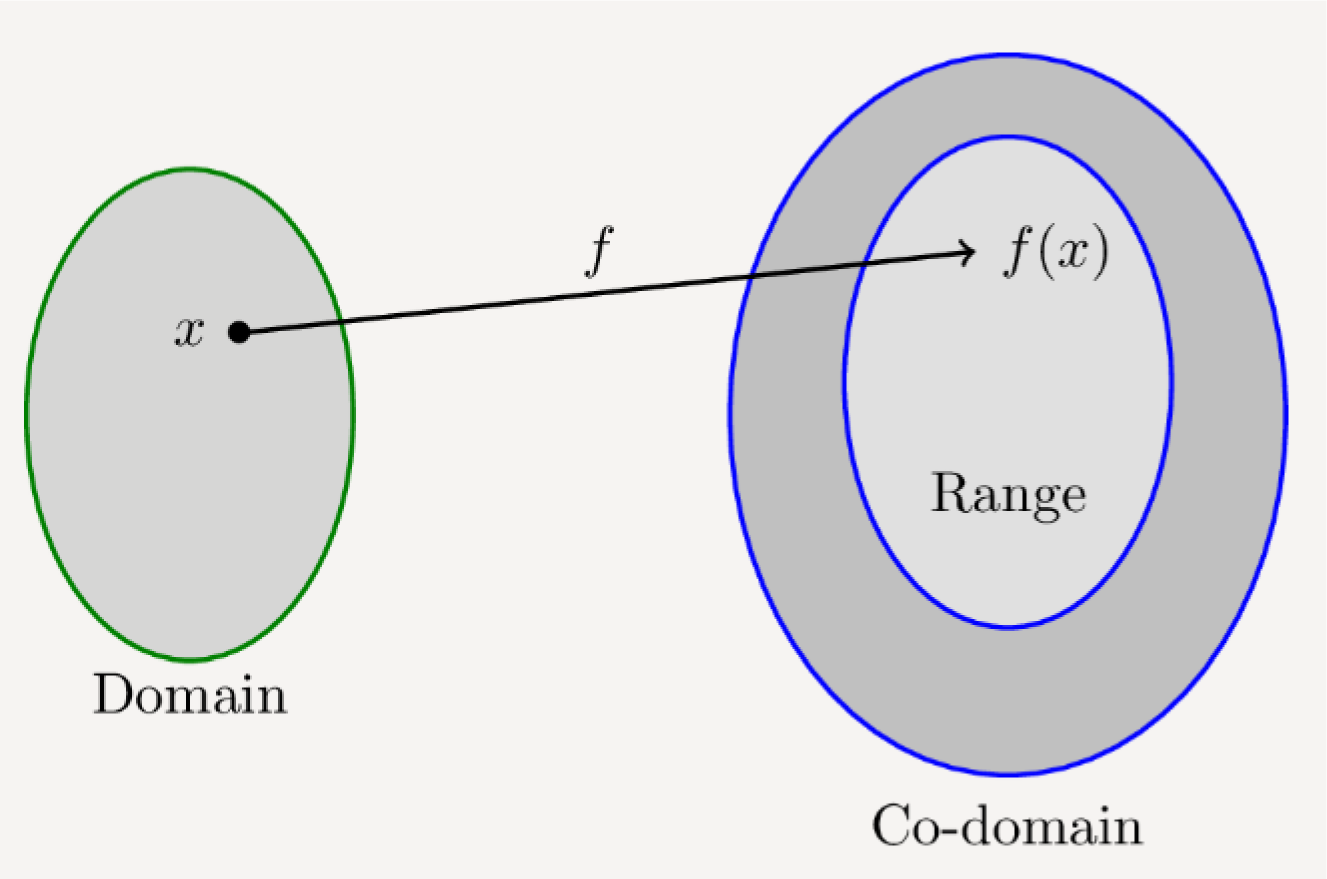
# **Solution: Inclusion-exclusion** Venn diagram.




$$21 = 10 + 8 + x \Rightarrow x = 3$$

$$\Rightarrow |B| = 4 + x + 3 = 10$$

# Functions: domain x co-domain.




$$f : A \to B.$$

- The output of a function $\boldsymbol{f : A \to B}$ always belongs to the co-domain $B$.

- Not all values in the co-domain are always covered by the function.

# Example

- Consider the function $\boldsymbol{f} : \mathbf{R} \rightarrow \mathbf{R}$, defined as $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}^2$. This function takes any real number $x$ and outputs $x^2$. For example, $\boldsymbol{f}(\mathbf{2}) = \mathbf{4}$.

- Consider the function $g : \{H, T\} \rightarrow \{0, 1\}$, defined as $g(H) = 0$ and $g(T) = 1$. This function can only take two possible inputs ***H*** or ***T***, where ***H*** is mapped to **0** and ***T*** is mapped to **1**.

# Random Experiments

- A random experiment is a process by which we observe something uncertain.

- An **outcome** is a result of a random experiment.

- An **event** be could an **outcome** or it could be a **conjunct of outcomes**.

- The set of all possible outcomes is called the **sample space** (the sample space is our *universal set*).

# Random Experiments

- When we repeat a random experiment several times, we call each one of them a **trial** (Substantivo: o processo; o ensaio; a experiência; a verificação; o julgamento – veredito).

- Em português costumamos denominar de experimento randômico, **realização** do sistema *radômico*).

# Randon experimente examples

- Random experiment: toss a coin; sample space:

  $\boldsymbol{S}$ = {***heads, tails***} or as we usually write it, {***H, T***}. Not all values in the co-domain are always covered by the function.

- Random experiment: roll a die; sample space:

$$\boldsymbol{S} = \{1, 2, 3, 4, 5, 6\}.$$

- Random experiment: observe the number of iPhones sold by an Apple store in Boston in 2023; sample space:

$$\boldsymbol{S} = \{0, 1, 2, 3, \cdots\}.$$

# Randon experimente examples

- Random experiment: observe the number of goals in a soccer match; sample space:

$$\boldsymbol{S} = \{0, 1, 2, 3, \cdots\}.$$

- Random experiment: toss a coin three times and observe the sequence of heads/tails. The sample space here may be defined as

$$\boldsymbol{S} = \{(H,H,H), (H,H,T), (H,T,H), (T,H,H), (H,T,T), (T,H,T), (T,T,H), (T,T,T)\}.$$

# Random Experiments (summary)

- **Outcome**: A result of a random experiment.
- **Sample Space**: The set of all possible outcome s.
- **Event**: A subset of the sample space.

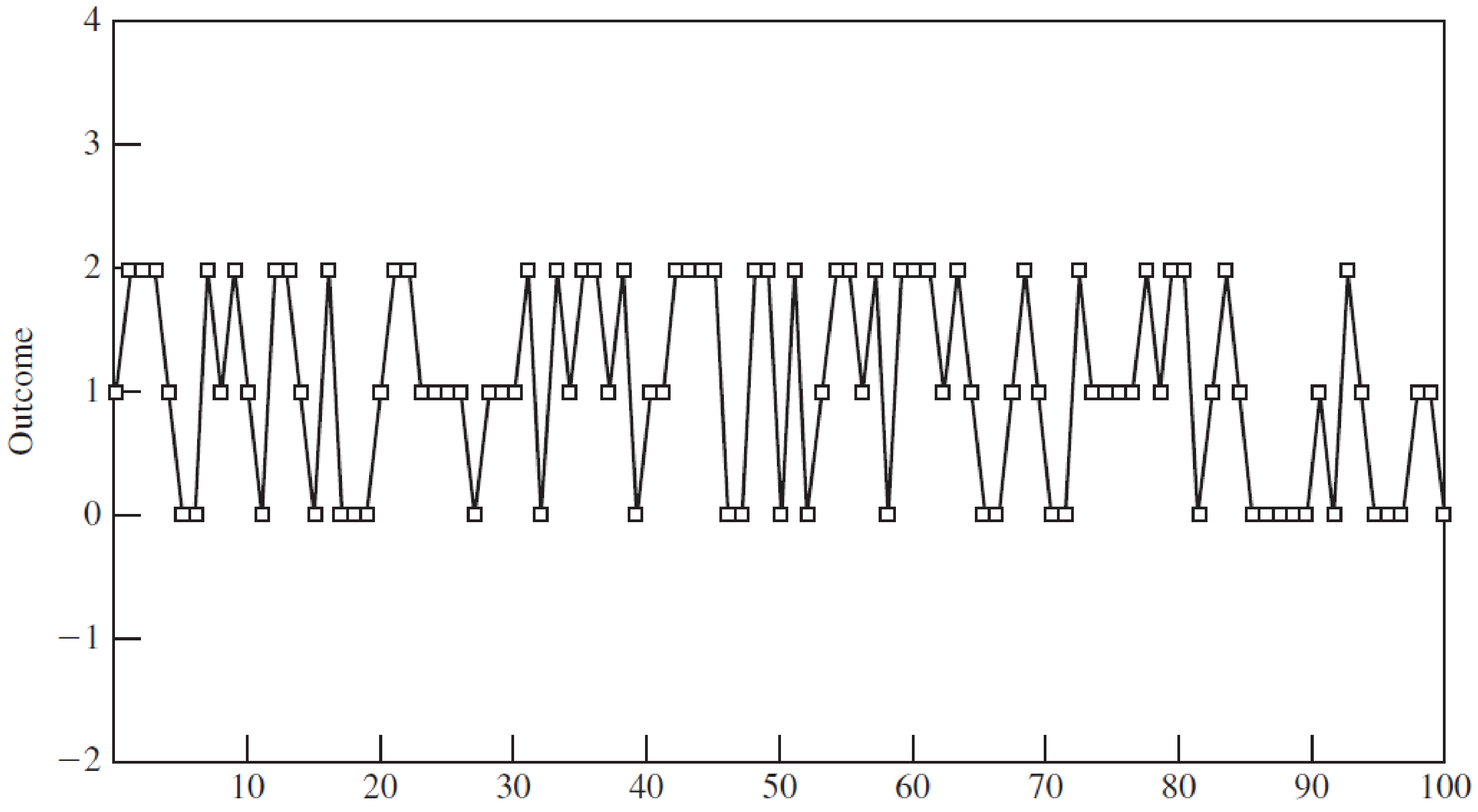
- An event is a collection of possible outcomes. In other words, an event is a subset of the sample space to which we assign a **probability**.

# Random processes in engineering

- Random processes in engineering usually involves an experiment that has a large number of outcomes - long sequences repetitions (trials).
- It is importante do define some quantities of interest of our study:
  - **Relative frequency** of outcome
  - **Statistical regularity**
  - **Probability** of any **Event (it maybe a single outcome)**

# Another example

- Supose a ball is selected from a urn. The identical balls are labelled: **0, 1, 2**. ("**Unordered Sampling with Replacement**")
- The **sample space** is S = {0, 1, 2}. Also supose a **trial (experimento)** with 100 repetitions (**outcomes**).



# **Relative frequency** of outcome

- Supose the randon experiment is repeated *n* times under identical conditions: long sequences repetitions (trials).

- **This is not always true**. In many circumstances, there is a little information about the "problem" to be modeled.

# **Relative frequency** of outcome

- Let $N_0$**(n)**, $N_1$**(n)** and $N_2$**(n)** be the number of times in which the **outcomes** are performed.

- The **relative frequency** of outcome $\boldsymbol{k}$ be defined by: $f_k(n) = \frac{N_{k(n)}}{n}$

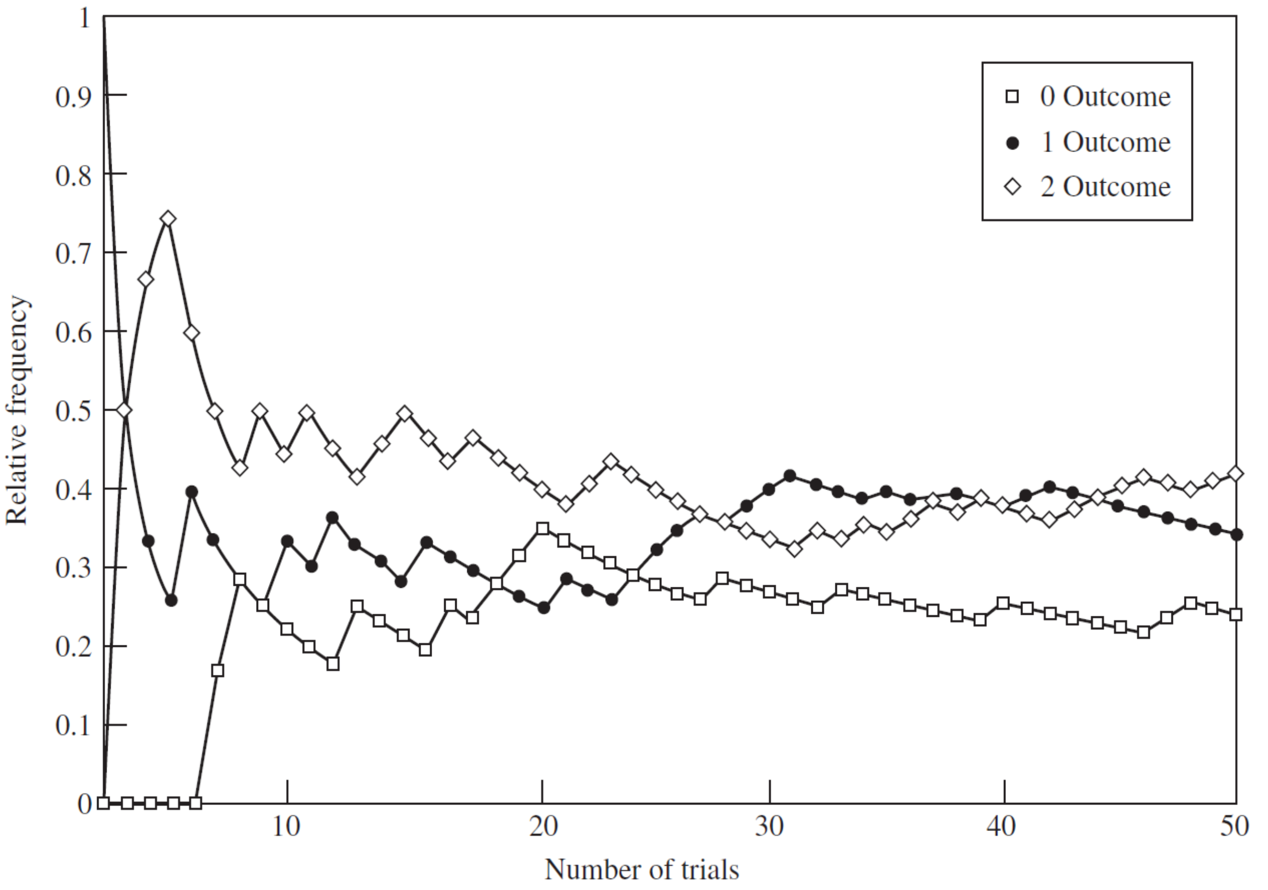
# Statistical regularity - probability

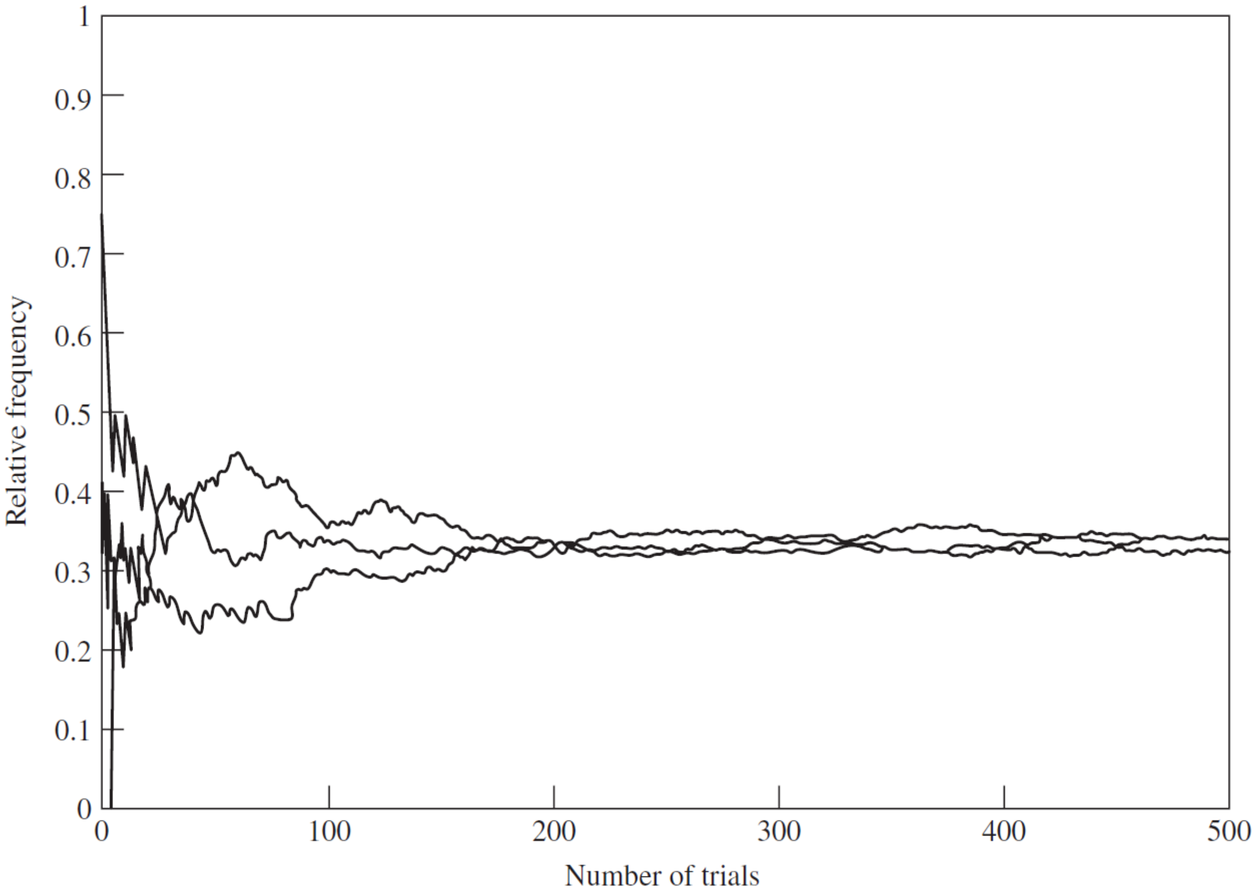
- By **statistical regularity** we mean that varies less and less about a constant value as *n* is made large,

$$\lim_{n\to\infty} f_k(n) = p_k$$

- The constant is called the **probability** of the **outcome** *k*.

# Statistical regulariry behavior



# Properties of relative frequency

- A random experimente with a sample space:

$$S = \{1, 2, \ldots, K\}.$$

- The **number of ocorrences** of any **outcome** in **n trials** could be computed by:

$$0 \leq N_k(n) \leq n \qquad \text{for } k = 1, 2, \ldots, K$$

- The **relative frequencies** are a number between zero and one:

$$0 \leq f_k(n) \leq 1 \qquad \text{for } k = 1, 2, \ldots, K$$

# Properties of relative frequency

- The sum of the **number of occurrences** of all possible **outcomes** must be ***n***:

$$\sum_{k=1}^{K} N_k(n) = n$$

- The sum of all the **relative frequencies** are equal to:

$$\sum_{k=1}^{K} f_k(n) = 1.$$

# Event

- The relative frequency of an event is the sum of the relative frequencies of the associated outcomes
- Let ***C*** be the event ***A* or *B*** occurs where *A* and *B* are two (outcomes) events that cannot occur simultaneously; The number of times when ***C*** occurs is

$$N_C(n) = N_A(n) + N_B(n) \quad f_C(n) = f_A(n) + f_B(n).$$

# Event

- Example: The number of experiments in which the outcome is an even-numbered ball. The relative frequency of the event is thus

$$f_E(n) = \frac{N_E(n)}{n} = \frac{N_0(n) + N_2(n)}{n} = f_0(n) + f_2(n)$$

# Probability

- **Axioms of Probability** (an **axiom** or postulate is a statement that is taken to be true, to serve as a premise or starting point for further reasoning and arguments):

  **Axiom 1**: For any event $A$, $\boldsymbol{P(A) \geq 0}$.

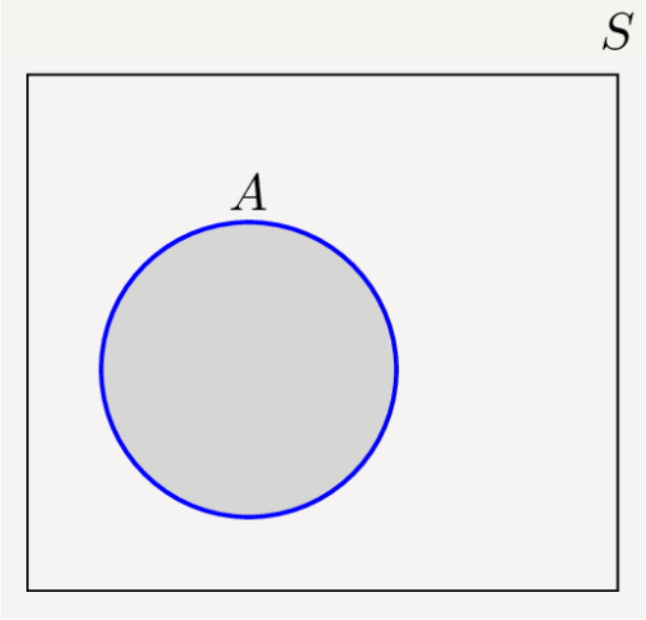
  **Axiom 2**: Probability of the sample space $S$ is $\boldsymbol{P(S) = 1}$.

  **Axiom 3**: If $A_1, A_2, A_3, \cdots$ are **disjoint events**, then

$$P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

**Axiom 1**: For any event $A$, $\boldsymbol{P(A) \geq 0}$.

**Axiom 2**: Probability of the sample space $S$ is $\boldsymbol{P(S) = 1}$.




A – A subset (an event)

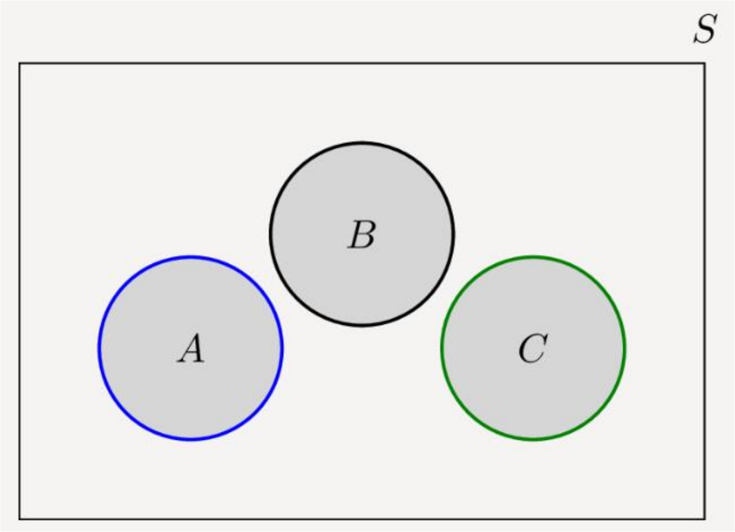
$$P(A) = \frac{|A|}{|S|} \geq 0$$

$$P(S) = \frac{|S|}{|S|} = 1$$

S - sample set

**Axiom 3**: If $A_1, A_2, A_3, \cdots$ are **disjoint events**, then

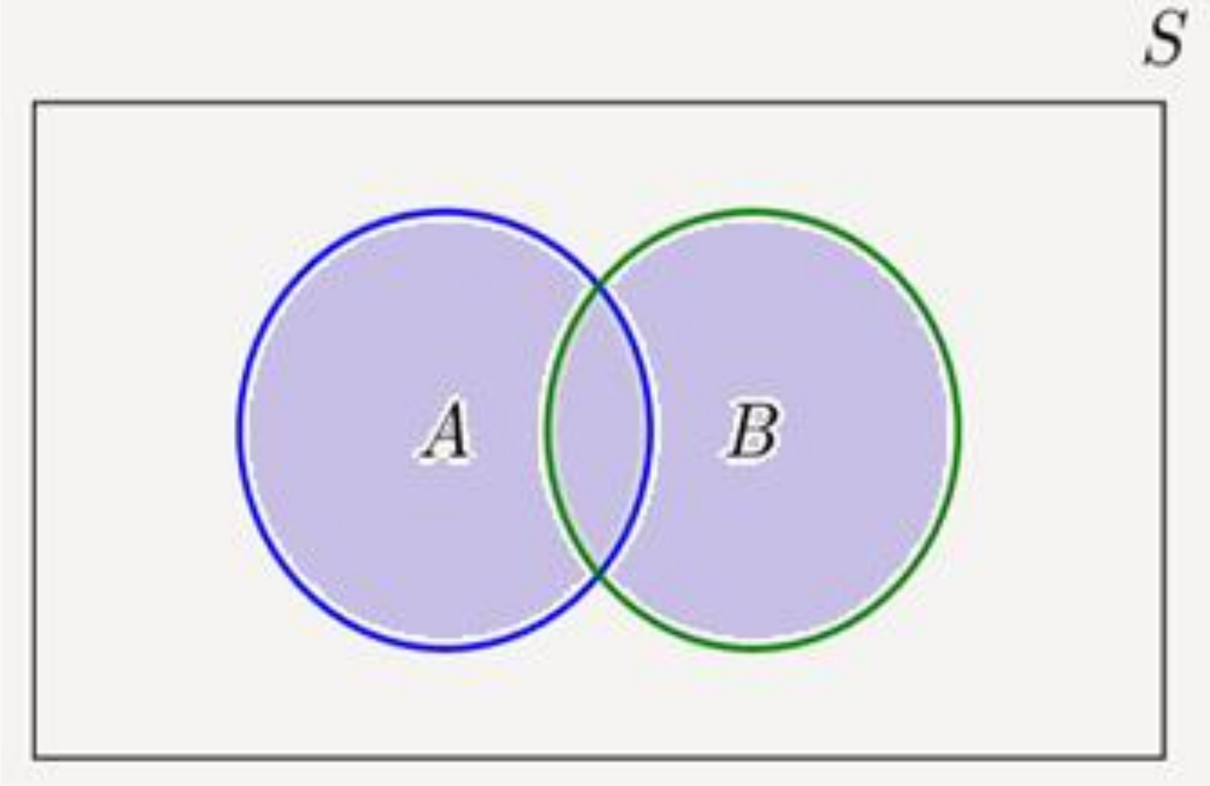
$$P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$




$$A \cap B \cap C = \phi$$

$$P(A \cup B \cup C) = \frac{|A|}{|S|} + \frac{|B|}{|S|} + \frac{|C|}{|S|} = \frac{|A|+|B|+|C|}{|S|} = P(A) + P(B) + P(C)$$

# Observe




- $|A \cup B| = |A| + |B| - |A \cap B|$

$$P(A \cup B) = \frac{|A| + |B| - |A \cap B|}{|S|} =$$

$$P(A) + P(B) - \mathrm{P}(A \cap B)$$

$$P(A \cup B) \neq \frac{|A|}{|S|} + \frac{|B|}{|S|} = P(A) + P(B)$$

# Probability

- Notation:

$$P(A \cap B) = P(A \text{ and } B) = P(A, B)$$

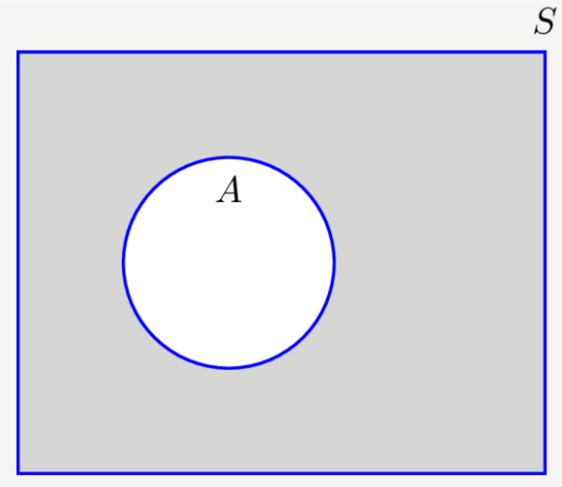
$$P(A, B) = P(A \cap B) = \frac{|A \cap B|}{|S|}$$

$$P(A \cup B) = P(A \text{ or } B)$$

$$P(A \cup B) = P(A \text{ or } B) = \frac{|A \cup B|}{|S|}$$

# From the axioms of probability, it is true

- For any event $A$, $P(A^c) = 1 - P(A)$.




$$\bar{A} = A^c$$

- $A^c = S - A$

$$P(A^c) = \frac{|S|}{|S|} - \frac{|A|}{|S|} = 1 - \mathrm{P(A)}$$

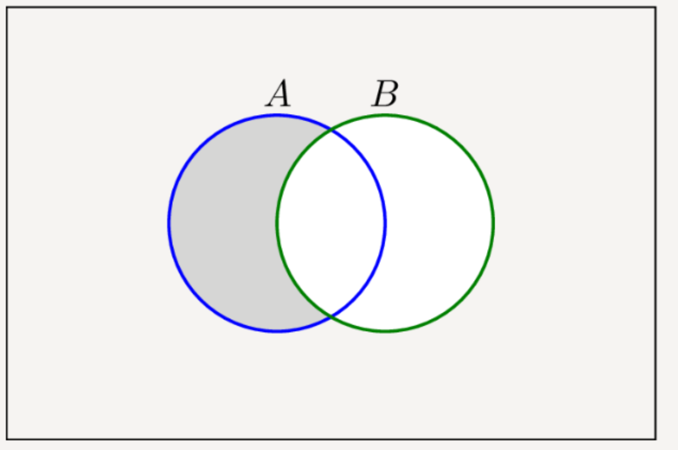
- The probability of the empty set is zero, i.e., $P(\emptyset) = 0$.

# From the axioms of probability, it is true

- For any event $A$, $P(A) \leq 1$.

- $P(A - B) = P(A) - P(A \cap B)$.

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




$C = A - B = A - A \cap B$

- If $A \subset B$ then $P(A) \leq P(B)$.

# Inclusion-exclusion principle

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
  (inclusion-exclusion principle for $n = 2$).

- $P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A \cap B \cap C)$,
  (inclusion-exclusion principle for $n = 3$).

# Inclusion-exclusion principle

- Generally, for $n$ events $A_1, A_2, \cdots, A_N$, we have

$$P\left(\bigcup_{k=1}^{N} A_k\right) = \sum_{i=1}^{N} P(A_k) - \sum_{i<j} P\left(A_i \cap A_j\right) + \sum_{i<j<k} P\left(A_i \cap A_j \cap A_k\right) - \ldots + (-1)^{N-1} P\left(\bigcap_{k=1}^{N} A_k\right)$$

# Event Classes - summary

- A **probability model is specified** by identifying the sample space ***S***, the event class of interest, and an initial probability assignment, a "**probability law**", from which the probability of all events can be computed.

# Event Classes - summary

- The sample space ***S*** specifies the set of all possible outcomes.

- If the sample space **S** has a countable (maybe it is infinite) number of elements, *S* is **discrete**; if it is uncountable *S* is **continuous** otherwise.

# Event Classes - summary

- Events are subsets of ***S*** that result from specifying conditions that are of interest in the particular experiment.

- When ***S*** is **discrete**, events consist of the union of elementary events.

# Event Classes - summary

- When $S$ is **continuous**, events consist of the **union or intersection of intervals in the continuous approach** (a real line, or an area and so on).

# Event Classes - summary

- The **axioms of probability** specify a set of properties that must be satisfied by the **probabilities of events.**

(An **axiom** or **postulate** is a statement that is taken to be true, to serve as a premise or starting point for further reasoning and arguments).

# Event Classes - summary

- The corollaries that follow from the axioms provide rules for computing the probabilities of events in terms of the probabilities of other related events.

# Ramdom Process

- Discrete Probability Models

- Continuous Probability Models

# Examples of discrete and continuous Random experiments

- ***Experiment 1***: Select a ball from an urn containing balls numbered 1 to 50. Note the number of the ball. $S_1 = \{1, 2, \ldots, 50\}$

- ***Experiment 2***: Select a ball from an urn containing balls numbered 1 to 4. Suppose that balls 1 and 2 are black and that balls 3 and 4 are white. Note the number and color of the ball you select. $S_2 = \{(1, b), (2, b), (3, w), (4, w)\}$

- ***Experiment 3***: Toss a coin three times and note the sequence of heads and tails.

$$S_3 = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}, \text{TTH}, \text{THT}, \text{HTT}, \text{TTT}\}$$

- ***Experiment 4***: Toss a coin three times and note the number of heads.

$$S_4 = \{0, 1, 2, 3\}$$

# Examples of discrete and continuous Random experiments

- ***Experiment 5***: Count the number of voice packets containing only silence produced from a group of $N$ speakers in a 10-ms period.

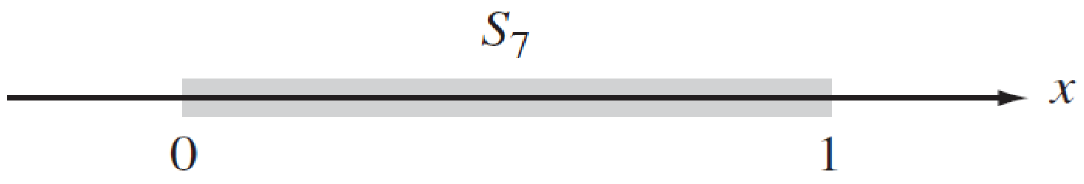
$$S_5 = \{0, 1, 2, \ldots, N\}$$

- ***Experiment 6***: A block of information is transmitted repeatedly over a noisy channel until an error-free block arrives at the receiver. Count the number of transmissions required.

$$S_6 = \{1, 2, 3, \ldots\}$$

- ***Experiment 7***: Pick a number at random between zero and one.

$$S_7 = \{x : 0 \leq x \leq 1\} = [0, 1]$$



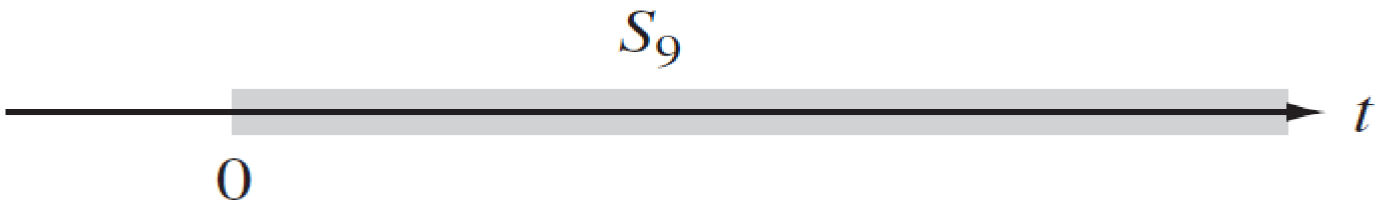

- ***Experiment 8***: Measure the time between page requests in a Web server.

$$S_8 = \{t : t \geq 0\} = [0, \infty)$$

# Examples of discrete and continuous Random experiments

- ***Experiment 9***: Measure the lifetime of a given computer memory chip in a specified environment.

$$S_9 = \{t : t \geq 0\} = [0, \infty)$$




- ***Experiment 10***: Determine the value of an audio signal at time $t_1$.

$$S_{10} = \{v : -\infty < v < \infty\} = (-\infty, \infty)$$

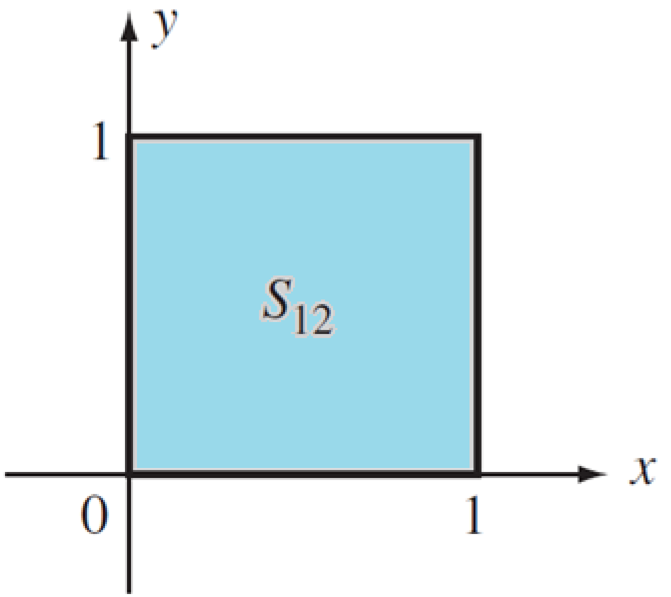
- ***Experiment 11***: Determine the values of an audio signal at times $t_1$ and $t_2$.

$$S_{11} = \{(v_1, v_2) : -\infty < v_1 < \infty \text{ and } -\infty < v_2 < \infty\}$$

# Examples of discrete and continuous Random experiments

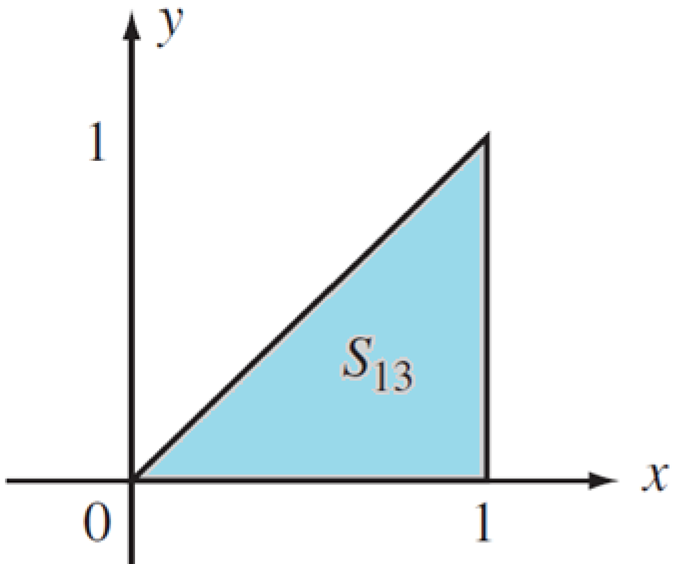
- ***Experiment 12***: Pick two numbers at random between zero and one.

$$S_{12} = \{(x, y): 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1\}$$




- ***Experiment 13***: Pick a number $X$ at random between zero and one, then pick a number $Y$ at random between zero and $X$.

$$S_{13} = \{(x, y): 0 \leq y \leq x \leq 1\}$$



# Examples of discrete and continuous Random experiments

- ***Experiment 14***: A system component is installed at time. For let as long as the component is functioning and let after the component fails.

$S_{14}$ = set of functions $X(t)$ for which $X(t) = 1$ for $0 \leq t < t_0$ and $X(t)$

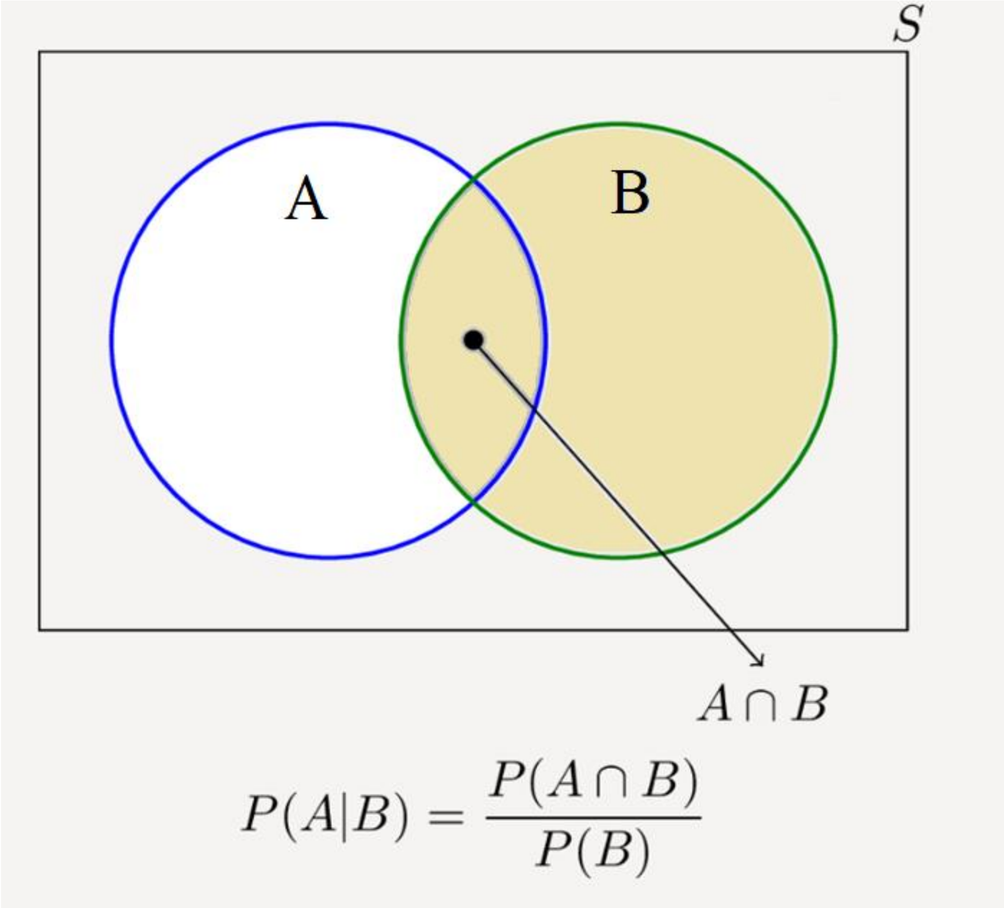
where $t_0 > 0$ is the time when the component fails.

# Conditional probability

- If $A$ and $B$ are two events in a sample space $S$, then the **conditional probability of** $A$ **given** $B$ *can be computed by*

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

$$=\frac{P(A\cap B)}{P(B)}=\frac{\boldsymbol{P(A,B)}}{\boldsymbol{P(B)}};\ \boldsymbol{P(B)>0}.$$



# Statistical Independence

- Two events *A* and *B of the sample space S* are independent if

$$P(A \cap B) = \mathrm{P(A, B)} = P(A)P(B).$$

- Note That:

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = \mathrm{P(A)}$$

# Exercício 2

- Consider a random experiment: Outcome: (observing the day's weather) + (rolling a die). The day's weather has three possibilities: sunny, cloudy and rainy. The die has 6 possibilities. Consider all possible equiprobable outcomes.

- a) What is the **size** of the sample space?
- b) What is the probability on a rainy day to get a 6 by rolling a die?
- c) Are the weather of the day and the rolling a die result independent? Can you use statistical Independence expression?
- d) Consider an event where we have a **sunny day** and when the die is rolled the result is **less than 3**. What is the probability of this event happening?

# Statistical Independence

- Three events $A$, $B$, and $C$ *of the sample space* $S$ are independent, so

$$P(A \cap B) = P(A)P(B),$$

$$P(A \cap C) = P(A)P(C),$$

$$P(B \cap C) = P(B)P(C),$$

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

# Statistical Independence

- *If in the* ***sample space S*** *there are* ***N events***

$$A_1, A_2, \cdots, A_N.$$

- *We say that they are* **independent** if we must have

$$P(A_i \cap A_j) = P(A_i)P(A_j),$$

$$\text{for all distinct } i, j \in \{1, 2, \cdots, N\}$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k),$$

$$\text{for all distinct } i, j, k \in \{1, 2, \cdots, N\}$$

# Statistical Independence - **Lemma 1**

- If $A$ and $B$ are independent events, then

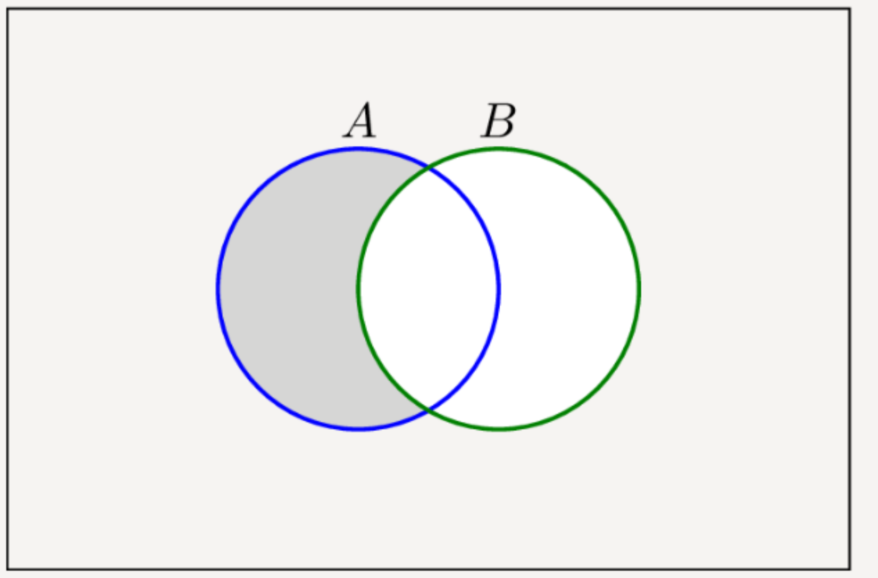
  $A$ and $B^c$ are independent,

  $A^c$ and $B$ are independent,

  $A^c$ and $B^c$ are independent

Example: Proof that events $A$ and $B^c$ are independent knowing that A and B are independent

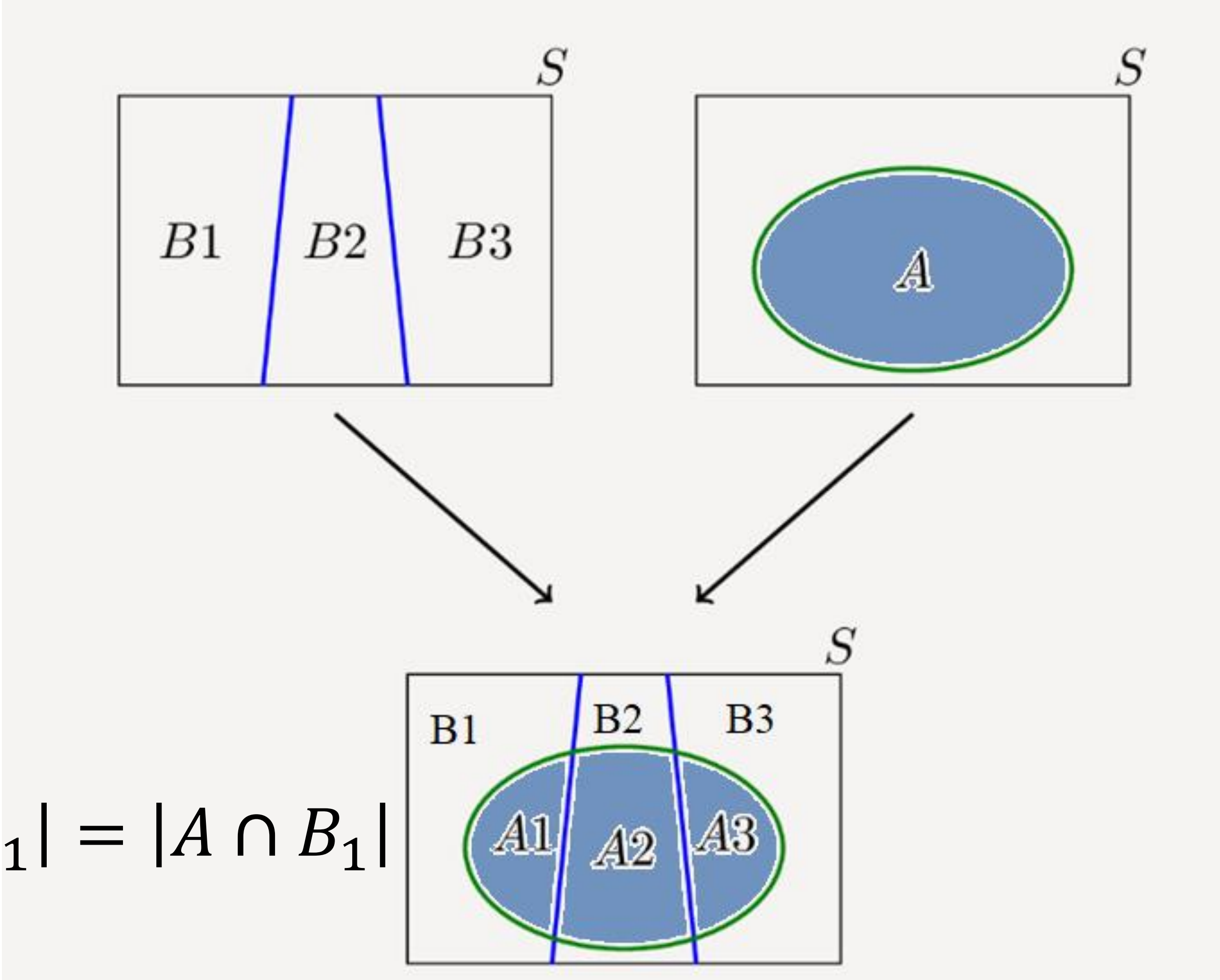
$$
\begin{aligned}
P(A,B^c) &= P(A - B) \text{ (from Venn Diagram)} \\
&= P(A) - P(A,B) = P(A) - P(A)P(B) \\
&= P(A)(1 - P(B)) \\
&= P(A)P(B^c).
\end{aligned}
$$



# Law of Total Probability

- If $B_1, B_2, B_3, \cdots$ are a **partition** of the sample space $S$, then for any event $A$ we have

$$P(A) = \sum_k P(A, B_k) = \sum_k P(A|B_k)P(B_k)$$



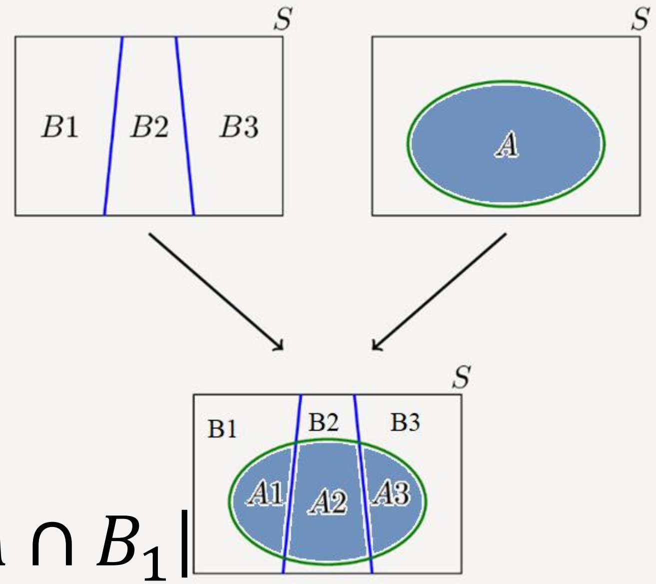

$$|A_1| = |A \cap B_1|$$

# Demonstration - Law of Total Probability

- It is known that: $|A| = |A_1| + |A_2| + |A_3|$
- Dividing on both sides by |S| results in $\frac{|A|}{|S|} = \frac{|A_1|}{|S|} + \frac{|A_2|}{|S|} + \frac{|A_3|}{|S|}$
- The same as: $P(A) = P(A_1) + P(A_2) + P(A_3)$ or

$$P(A) = \frac{|A \cap B_1|}{|S|} + \frac{|A \cap B_2|}{|S|} + \frac{|A \cap B_3|}{|S|}$$

$$= P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3)$$




- From the conditional probability

$$\mathrm{P}(A|B) = \frac{P(A \cap B)}{P(B)} \text{ then } P(A \cap B) = P(A|B)P(B)$$

- Finally, we can write

$$|A_1| = |A \cap B_1|$$

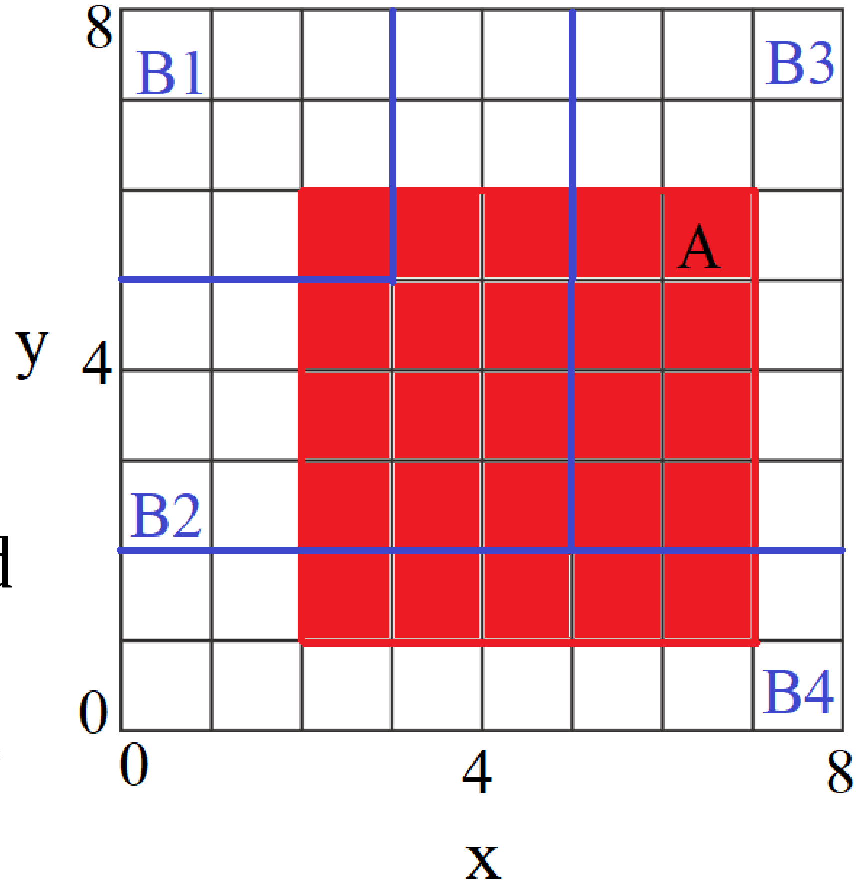
$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)$$

# Law of Total Probability

- The **Law of Total Probability** for k **partitions** of the sample space $S$, and a designed event $A$, finaly, it can be written as

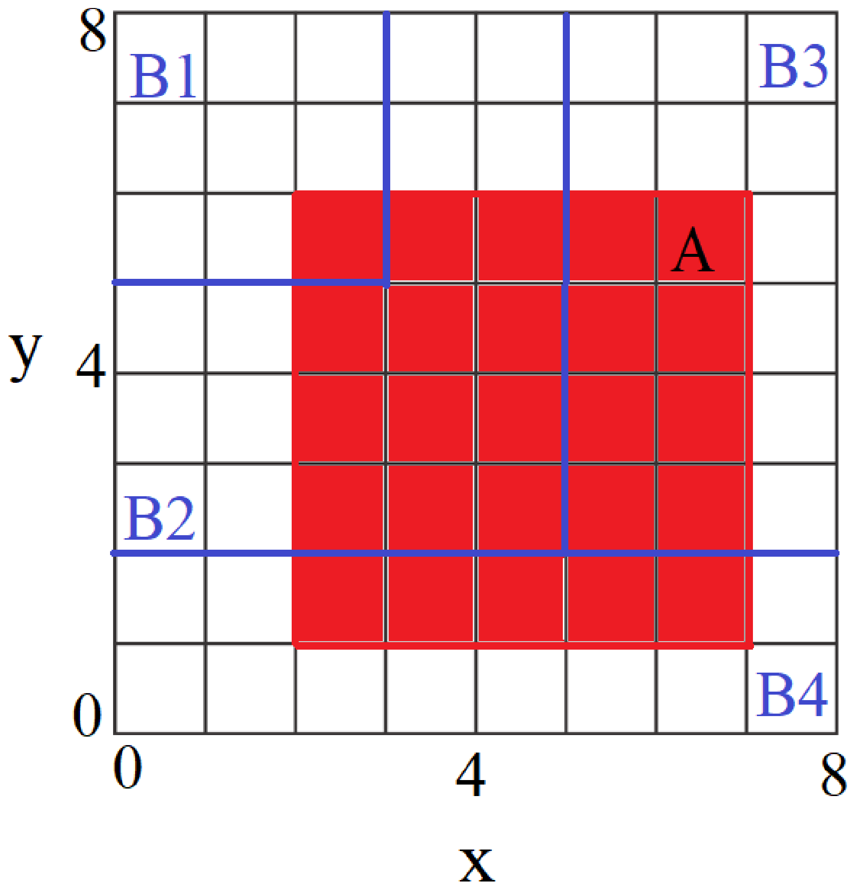
$$P(A) = \sum_{k} P(A \cap B_k) = \sum_{k} P(A|B_k)P(B_k)$$

# Exercise 3




- A cannon shoots iron filings (*limalha de ferro*) onto an 8 x 8 cm$^2$ bulkhead (*anteparo*). This target is mounted perpendicular to the trajectory of the iron filings. Assume that the probability of hitting any point on the bulkhead is equal (don't ask me how).
- 1- Consider that the red rectangle drawn on the screen corresponds to my event of interest.
- 2 - Assume that the rectangular target screen has been divided into **partitions** which are drawn in blue. As shown in the figure.
- Determine the probability of the iron filings hitting the red rectangle using the law of total probability.

# Exercise 3: Solution – Given the sample space S and the partition B1, B2, B3 and B4, calculate P(A) using the Law of Total Probability



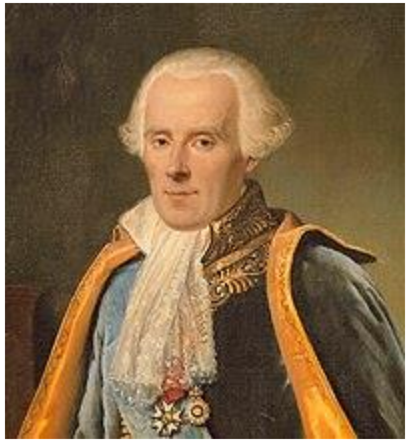

- $P(A) = |A|/|S| = 25/64 = 0.3906$ (by the definition)
- $P(B_1) = 9/64 = 0.1406$; $P(A|B_1) = 1/9 = 0.1111$
- $P(B_2) = 21/64 = 0.3281$; $P(A|B_2) = 11/21 = 0.5238$
- $P(B_3) = 18/64 = 0.2814$; $P(A|B_3) = 8/18 = 0.4444$
- $P(B_4) = 16/64 = 0.2500$; $P(A|B_4) = 5/16 = 0.3125$

$$P(A) = P(A|B_1)\,P(B_1) + P(A|B_2)\,P(B_2) + P(A|B_3)\,P(B_3) + P(A|B_4)\,P(B_4)$$

$$= 0.016 + 0.1719 + 0.1251 + 0.078 = 0.3906$$

# Bayes' theorem

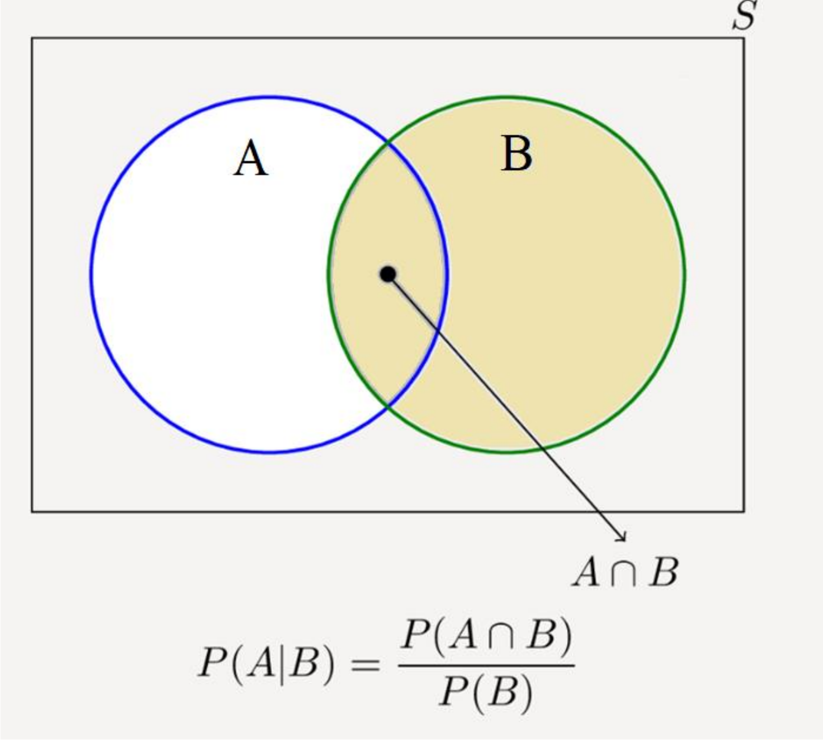
- Bayes' theorem is named after the English pastor and mathematician **Thomas Bayes** (1701 – 1761).
- Thomas Bayes was the first to provide an equation that would allow new evidence to update the probability of an event from a priori knowledge (or the initial belief in the occurrence of an event).
- Bayes' unpublished manuscript was significantly edited by Richard Price before being read posthumously at the Royal Society.
- Bayes' theorem was later developed by Pierre-Simon Laplace, who was the first to publish a modern formulation in 1812 in his book *Analytical Theory of Probability*.



# Bayes' law

- We found that **conditional probability of *A* given *B*** (P(A|B)) *as*:

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




- Similarly, we can derive P(B|A)

$$P(B|A) = \frac{|\mathrm{B} \cap A|}{|A|} = \frac{\frac{|\mathrm{B} \cap A|}{|S|}}{\frac{|A|}{|S|}} = \frac{P(\mathrm{B} \cap A)}{P(A)} = \frac{P(B,A)}{P(A)}$$

- However,

$$|\mathrm{B} \cap A| = |\mathrm{A} \cap B|$$

$$P(\mathrm{B},A) = P(A,B)$$

# Bayes' law - continuation

- We can write

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

- The **Bayes' Rule to compute** P(B|A) can be written as

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

# Bayes' law - continuation

- If $C_1, C_2, C_3, \cdots$ form a **partition** of the **sample space $S$** (*classes*), and $x$ is any event with $P(x) \neq 0$, we have

$$P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)}$$

- Using the Law of Total Probability, the Bayes' Rule can, finaly, be written as

$$P(C_k|x) = \frac{P(x|C_k)P(C_k)}{\sum_k P(x|C_k)P(C_k)}$$

# Conditional Independence

- From the definition of conditional probability,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A,B)}{P(B)}, P(B) > 0$$

- Two events *A* and *B* are independent if $P(A \cap B) = P(A)P(B)$, or equivalently, $P(A|B) = P(A)$.
- **Definition**: If events $A$ and $B$ are **conditionally independent** given an event $C$ with P(C) > 0:

$$P(A,B|C) = P(A|C)P(B|C)$$

# Conditional Independence

- If P(B|C), P(C) ≠ 0. If A and B are conditionally independent given C, we obtain

$$P(A|B,C) = \frac{P(A,B,C)}{P(B,C)} = \frac{P(A,B,C)/P(C)}{P(B,C)/P(C)}$$

$$= \frac{P(A,B|C)}{P(B|C)}$$

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

# Exercise 4

- Consider a random experiment: Outcome: (observing the day's weather) + (rolling a die). The day's weather has three possibilities: sunny, cloudy and rainy. The die has 6 possibilities. Consider all possible equiprobable outcomes. Suppose an event C is when the die is rolled it gives a value greater than 1 and the day weather can be anything. ($|C| = 5\text{x}3 = \mathbf{15}$).
- Consider the Event A where we have a sunny day, and the outcome of the dice game can be anything. Event B when the die is rolled, and the result is less than 3, and the day weather can be anything.
- a)What is the probability of this event happening A, since the event C is happened?
- b)What is the probability of the events A and B happening, since the event C is happened?

# Exercise 4: Solution

- $|S| = 18.$
- Event A: It is a sunny day, and the outcome of the dice game can be anything $|A| = 6.$
- Event B: The die is rolled, and the result is less than 3, and the day weather can be anything $|B| = \mathbf{6}.$
- Event C: The die is rolled it gives a value greater than 1 and the day weather can be anything $|C| = 5\text{x}3 = 15.$
- a) The probability of a **sunny day given the event C**

$$P(A|C) = P(A) = \frac{6}{18} = \frac{1}{3}$$

# Exercise 4: Solution

- b) By the conditional Independence we have to calculate

$$P(A, B|C) = P(A|C)P(B|C)$$

- The probability of the event B given the event C

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

$|B \cap C| = 3$, then

$$P(B|C) = \frac{3}{15} = \frac{P(B,C)}{P(C)} = \frac{\frac{3}{18}}{\frac{15}{18}} = \frac{1}{5}$$

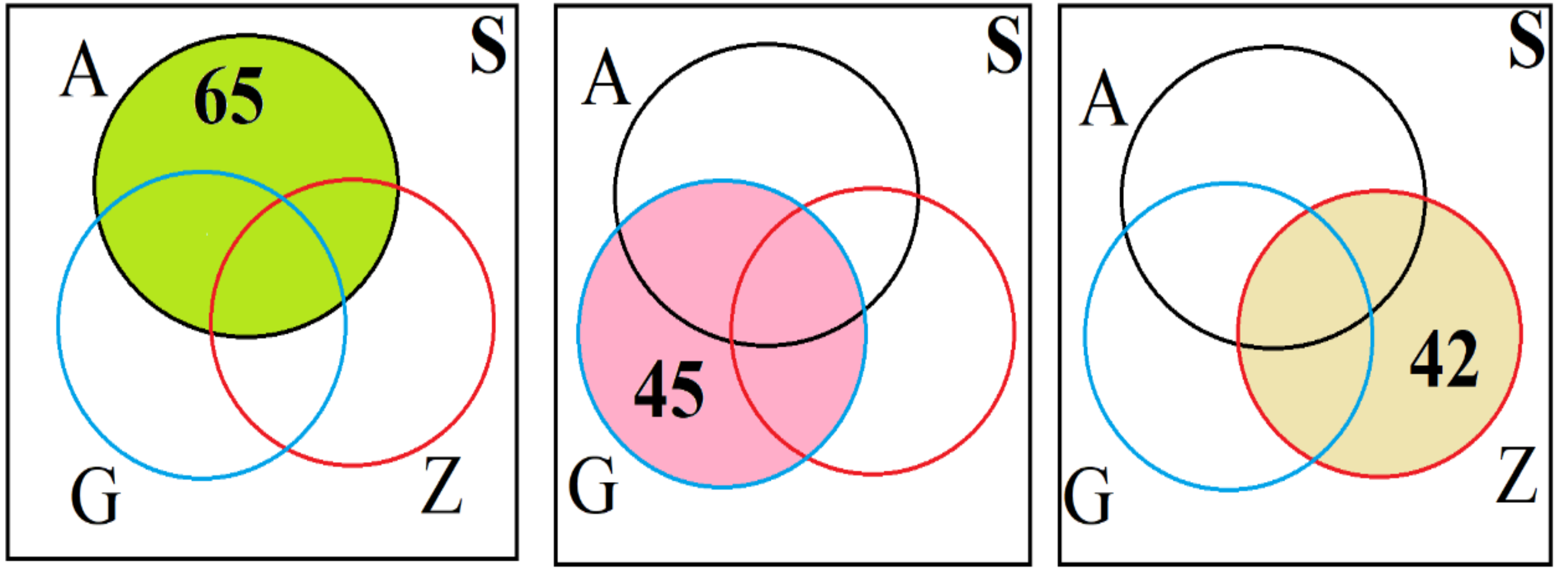
- Then $P(A,B|C) = P(A|C)P(B|C) = \frac{1}{3}\frac{1}{5} = \frac{1}{15}$

# Exercise 5

- In a research on the **stock market** **(mercado de ações)**, involving publicly traded companies, a sample was made where 120 people were interviewed, it was found that in the **stock market** **65 people owns shares of company A**, **45 people owns shares of company G** and **42 owns shares of company Z**.
- It was also verified that **20 people owns shares of companies A and G**, **25 owns shares of companies A and Z**, **15 owns shares of companies G and Z**, and **8 people hold shares of the three companies**.
  a) Find the number of people who own shares, at least, of one of the three companies.
  b) Find the number of people who own shares of a single company.
  c) Find the number of people who do not own shares of any company.

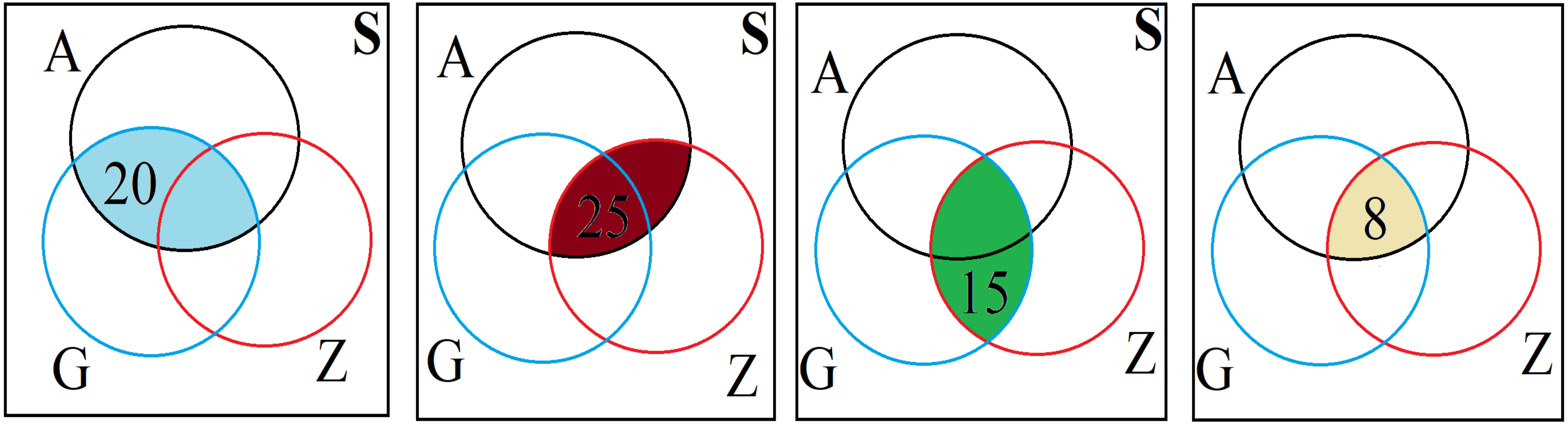
# First step, what information is given?

- 1 – A research with 120 people (sample space size **S**)
- 2 - 65 people have shares in company A, **|A| = 65**
- 3 - 45 people have shares of stock G, **|G| = 45**
- 4 - 42 people have Z shares, **|Z| = 42**



# Other information that was given

- 5 - 20 people have shares of companies A and G, $|A \cap G| = 20$
- 6 - 25 people have shares of companies A and Z, $|A \cap Z| = 25$
- 7 - 15 people have shares of companies G and Z, $|G \cap Z| = 15$
- 8 - 8 people have shares of the three companies. $|A \cap G \cap Z| = 8$

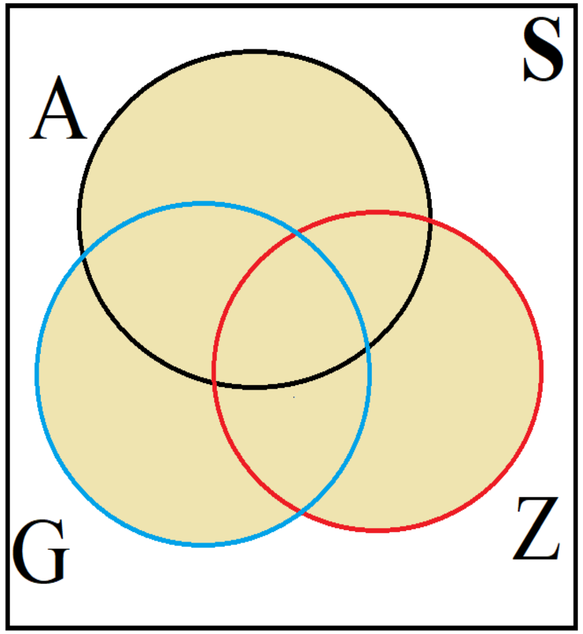


# Solution:

a) Find the number of people who own shares, at least, of one of the three companies.

- It means: We have to calculate how many people have shares of company A or (at least) company G or (at least) company Z. This means: **|A ∪ G ∪ Z|**. We can resolve the issue using the **principle of inclusion - exclusion**:

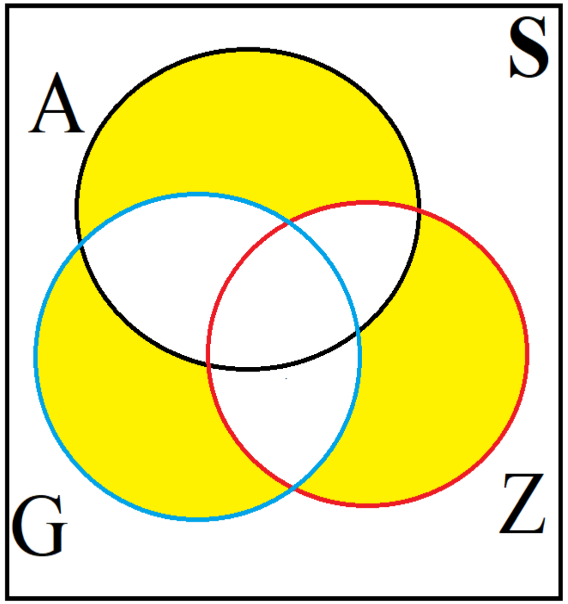
- $|A \cup G \cup Z| = |A|+|G|+|Z|-|A \cap G|-|A \cap Z|-|G \cap Z|+|A \cap G \cap Z| = 65 + 45 + 42 - 20 - 25 - 15 + 8 = 100$



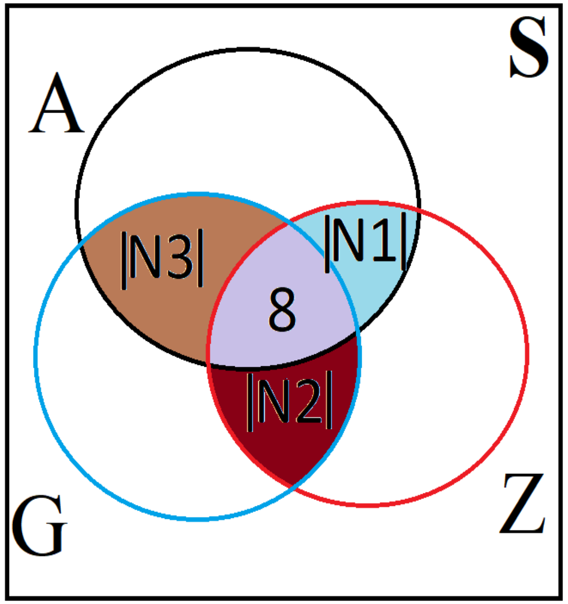

$|A \cup G \cup Z| = 100$

# Solution of item b)

- b) Find the number of people who own shares of a single company.
- The scenario is shown in the following figure: in Fig. (a) shows what we want, and in Fig. (b) shows what we have to subtract from $|A \cup G \cup Z|$ to obtain Fig. (a), that is, the number of people who own shares of a single company ($|M|$)
- $$|M| = |A \cup G \cup Z| - |N1| - |N2| - |N3| - |A \cap G \cap Z|$$



(a)



(b)

# Solution: Continuation of item b)

- Since the number of people with shares of the three companies **A**, **G** and **Z** was given, and it is equal $|A \cap G \cap Z| = 8$. Furthermore, we can compute $|N1|$, $|N2|$ and $|N3|$. $|N1|$ can be calculated by:

$$|N1| = |A \cap Z| - |A \cap G \cap Z| = 25 - 8 = 17$$

- Similarly, $|N2|$ the number of people with shares only in companies **G** and **Z** can be calculated by

$$|N2| = |G \cap Z| - |A \cap G \cap Z| = 15 - 8 = 7$$

- Following the same path, $|N3|$ the number of people with shares only in companies **A** and **Z** can be calculated by

$$|N3| = |A \cap G| - |A \cap G \cap Z| = 20 - 8 = 12$$

- Finally, we can calculate the number of people who only own shares in a single company:

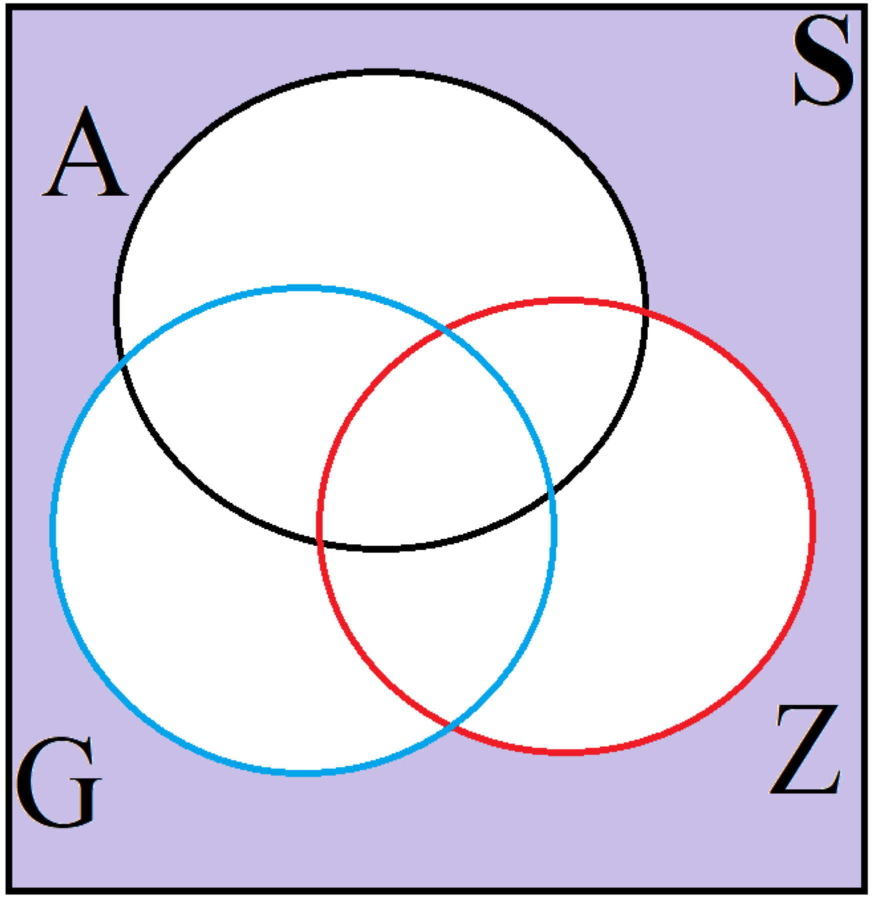
$$|M| = |A \cup G \cup Z| - |N1| - |N2| - |N3| - |A \cap G \cap Z|$$

$$= 100 - 17 - 7 - 12 - 8 = 56$$

# Solution of item c)

c) Find the number of people who do not own shares of any company.

**$|P| = |S| - |A \cup G \cup Z| = 120 – 100 = 20$**; $|S| = 120$ corresponde ao espaço amostral.



# Exercise 6

- Considering that the sample presented in exercise 5 is a random experiment that represents the behavior of an investor community. Using the results of exercise 5 as a data source:

- a) Find the probability that a person does not own any of the three shares which is part of the random experiment.
- b) Find the probability that an investor owns, simultaneously, shares of the three companies.

# Exercise 6 - continuation

- c) Knowing that an investor owns shares in company Z, what is the probability that he will also own shares in company G: P(G|Z)?
- d) Knowing that an investor owns shares of company G, what is the probability that he also owns shares of company A (you have to solve this item)?
- e) Using Bayes' law and the result of item c) of this exercise, find the probability of an investor owning shares in company Z, knowing that he owns shares in company G. That means P(Z|G).

# Exercise 6: Solution

- a) Find the probability that a person does not own any of the three shares that is part of the random experiment.

$$P(x_1) = \frac{|S| - |A \cup G \cup Z|}{|S|} = \frac{20}{120} = \frac{1}{6}$$

- b) Find the probability that an investor owns, simultaneously, shares of the three companies.

$$P(x_2) = \frac{|A \cap G \cap Z|}{|S|} = \frac{8}{120} = \frac{1}{15}$$

# Exercise 6: Solution

- c) Knowing that an investor owns shares in company Z, what is the probability that he will also own shares in company G?

$$\boldsymbol{P(x_3) = P(\mathrm{G}|\mathrm{Z}) = \frac{|G \cap Z|}{|Z\,|} = \frac{15}{42} = \frac{5}{14} = 0.357}$$

- Another solution, using the conditional probability equation:

$$P(\mathrm{G}|\mathrm{Z}) = \frac{P(G,Z)}{P(Z)}; P(G,Z) = P(G \cap Z) = \frac{|G \cap Z|}{|S|} = \frac{15}{120} = 0.125;$$

$$P(Z) = \frac{|Z|}{|S|} = \frac{42}{120} = 0.350; \text{then } \boldsymbol{P(G|Z) = \frac{0.125}{0.350} = 0.357}$$

# Exercise 6: Solution

- e) Using Bayes' law and the result of item c) of this exercise, find the probability of an investor owning shares in company Z, knowing that he owns shares in company G.

By Bayes' law $P(B|A) = \frac{P(A|B)P(B)}{P(A)} \rightarrow P(Z|G) = \frac{P(G|Z)P(Z)}{P(G)}$

From c) $P(G|Z) = 0.357$; $P(Z) = \frac{|Z|}{|S|} = \frac{42}{120} = 0.350$; $P(G) = \frac{|\boldsymbol{G}|}{|\boldsymbol{S}|} = \frac{\mathbf{45}}{\mathbf{120}} = 0.375$;

then $\boldsymbol{P(Z|G)} = \frac{\mathbf{(0.357)(0.350)}}{\mathbf{0.375}} = \mathbf{0.333}$

Another solution: $\boldsymbol{P(Z|G)} = \frac{|\mathbf{Z \cap G}|}{|\boldsymbol{G}|} = \frac{\mathbf{15}}{\mathbf{45}} = \mathbf{0.333} = \frac{\boldsymbol{P(Z,G)}}{\boldsymbol{P(G)}}$

# This is the end of module 1!