Statistical signal processing (5CTA0)

Lecture 1, part A

Lecturer: Simona Turco

Electrical Engingeering, Signal Processing Systems group

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Introduction to the course

Lecture 1, Part A

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Motivation

Why statistical signal processing?

random vs **deterministic**

Why random signals?

- **Signal (time-series):**
  - observations are ordered in time and adjacent observations are dependent
  - Focus on discrete time signals (sequences)

- Successive observation are dependent → we can predict them
- If prediction are exact, the series is deterministic
- If prediction are not exact, the series is stochastic
- The degree of predictability is determined by the dependence between consecutive observations

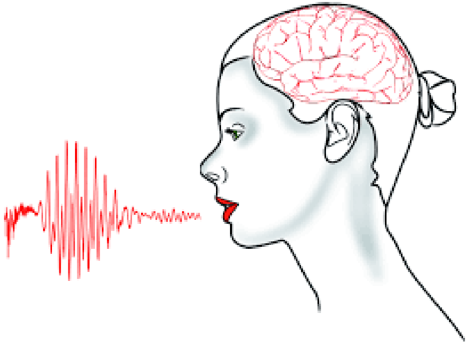
Why statistical signal processing?

“Although random signals are evolving in time in an unpredictable manner, their average statistical properties exhibit considerable regularity”

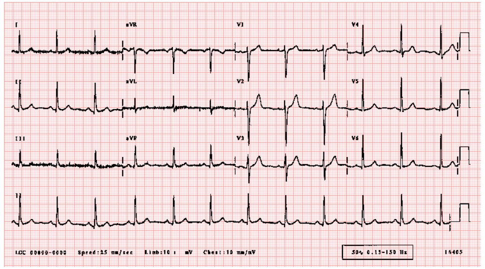
- Described mathematically by using
  - Theory of probability
  - Random variables
  - Stochastic processes

Examples of random signals

Speech

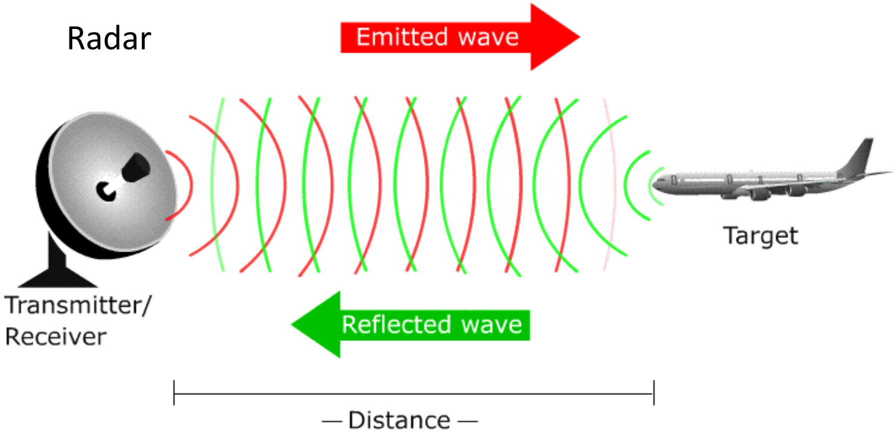


Bio-signals



Geophysical

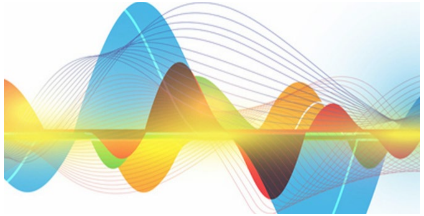
Radar



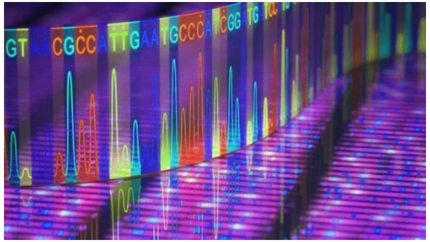
Applications

Communication, information and control theory

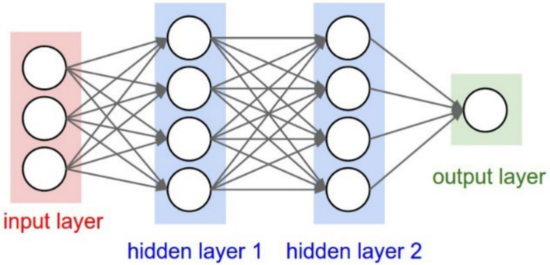
Signal processing



Biostatistics and bioinformatics



Machine learning



Economy and Finance

Content overview

Part 1
Random Variables and Random Signals

Part 2
Estimation Theory

Part 3
Spectral Estimation

Part 4
Detection Theory

Part 1: Random variables and Random Signals

Part 1

Random Variables and Random Signals

Lecture 1: Probability and Random Variables
Lecture 2: Random vectors, Random processes and random signals
Lecture 3: Rational signal models

Part 1: Random variables and Random Signals

Part 1

Random Variables and Random Signals

Lecture 1: Probability and Random Variables

Part A: Probability

Part B: Random variables

Probability

Lecture 1, Part A

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Lecture 1, part A: Probability

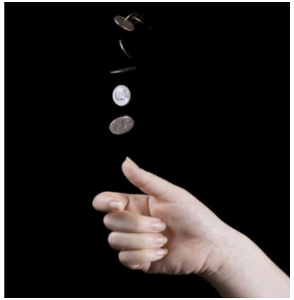
Outline

- Introduction and basic definitions
- Conditional probability
- Law of total probability
- Bayes theorem
- Independence

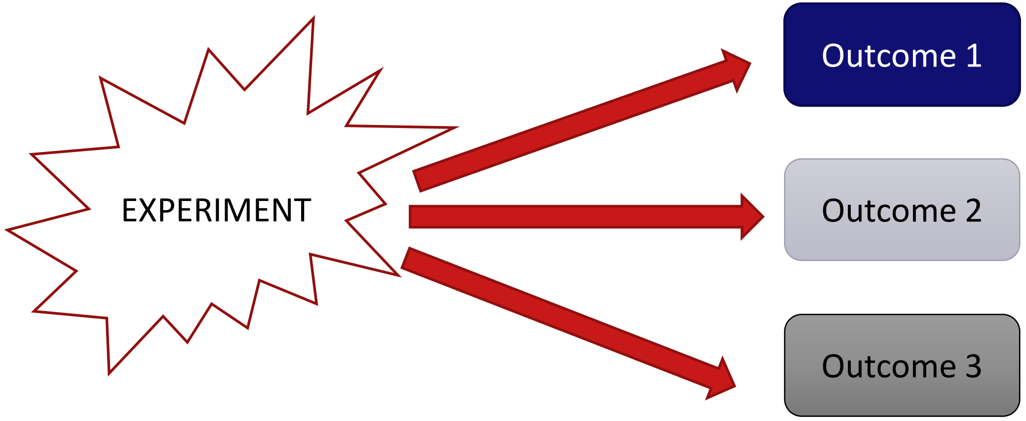
Probability: introduction

- The word “stochastic” comes from the Greek word for “random” of “chance”.

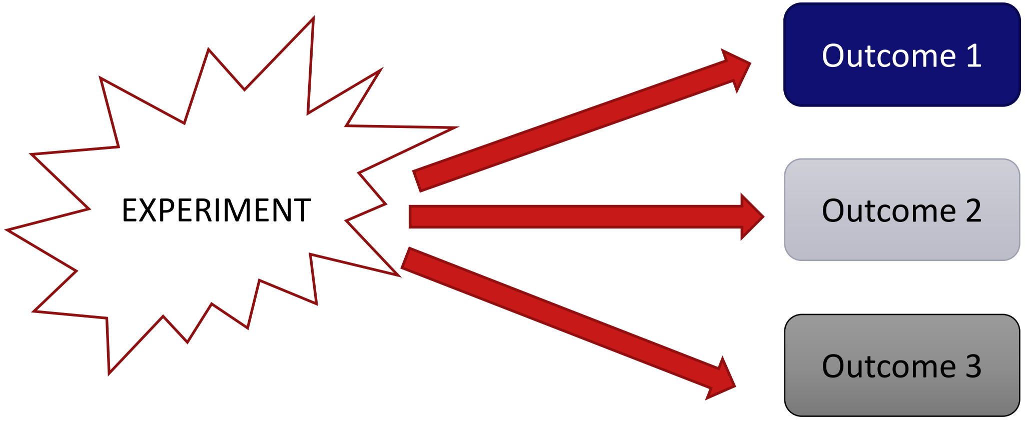
- Probability is the branch of mathematics concerned with analysis of random phenomena



Probability: definitions

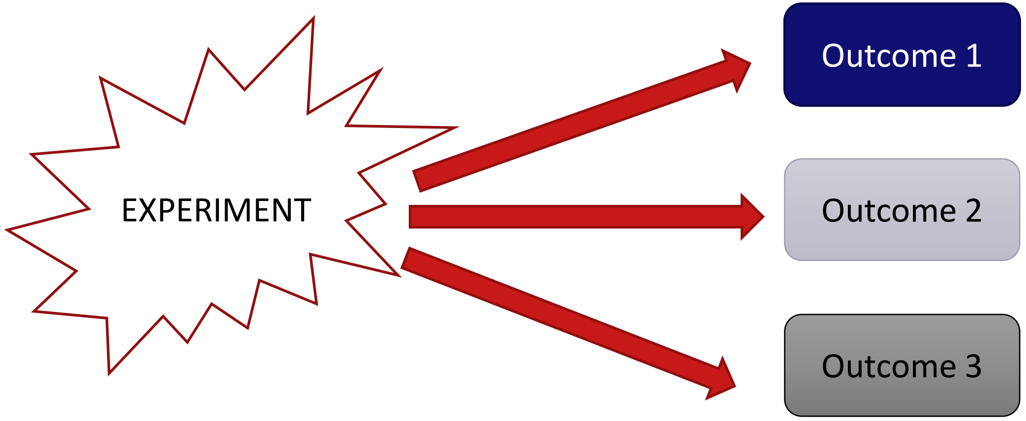


Probability: definitions



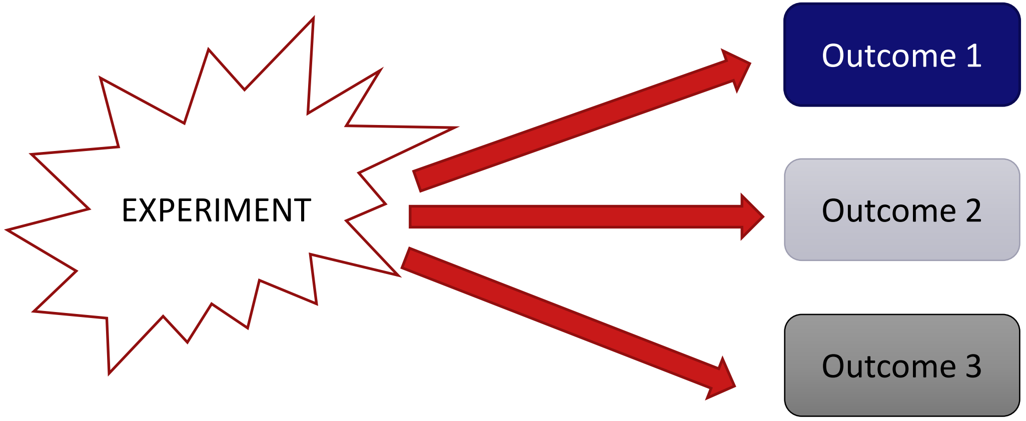
- **Experiment**: procedure that can be repeated infinitely many times

Probability: definitions



- **Experiment**: procedure that can be repeated infinitely many times
- **Observation** or **trial**: one realization of the experiment

Probability: definitions



- **Experiment**: procedure that can be repeated infinitely many times
- **Observation** or **trial**: one realization of the experiment
- **Outcome**: any possible observation of the experiment

Probability: definitions

Experiment

1 observation/trial = tossing the coin twice

Probability: definitions

Experiment

Outcome HH

Outcome TT

Outcome TH

Outcome HT

Probability: definitions

Experiment

Outcome HH

Outcome TT

Outcome TH

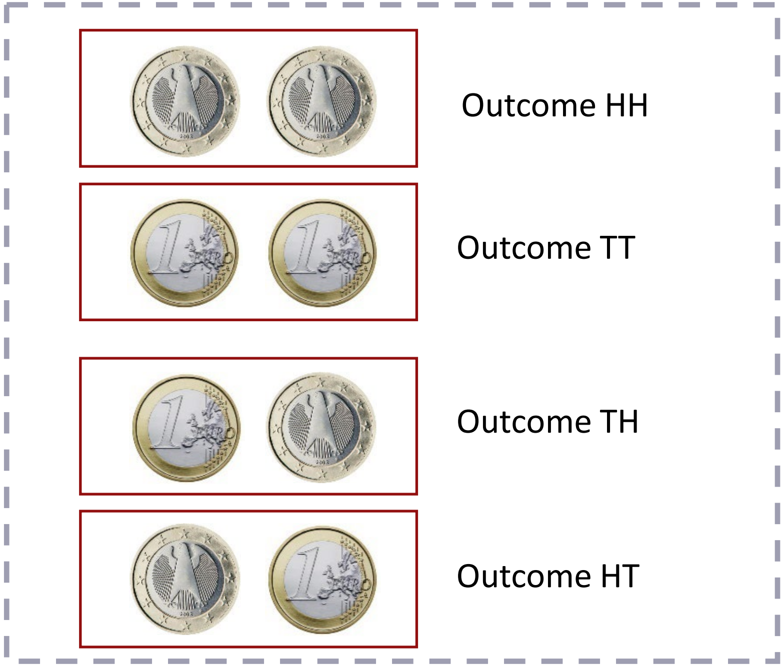
Outcome HT

SAMPLE SPACE

The **sample space**, denoted by S, is the set of all possible outcomes

Probability: definitions

Experiment



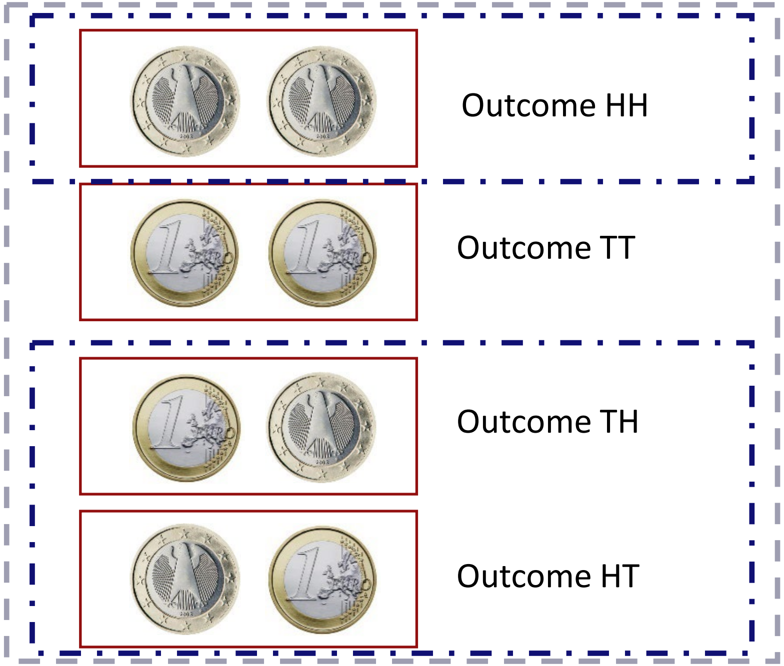
SAMPLE SPACE

The **sample space**, denoted by S, is the set of all possible outcomes

An **event** is a set of outcomes of an experiment, which can be the sample space or a subset of the sample space.

Probability: definitions

Experiment



SAMPLE SPACE

Event A:
Tossing at least one head

Probability: definitions

Experiment

Outcome HH

Outcome TT

Outcome TH

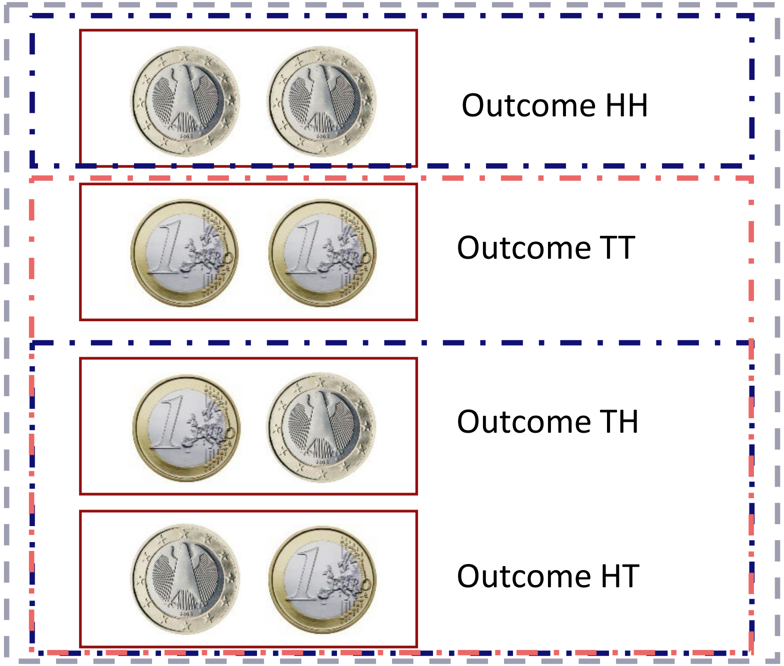
Outcome HT

SAMPLE SPACE

Event B:
tossing at least one tail

Probability: definitions

Experiment



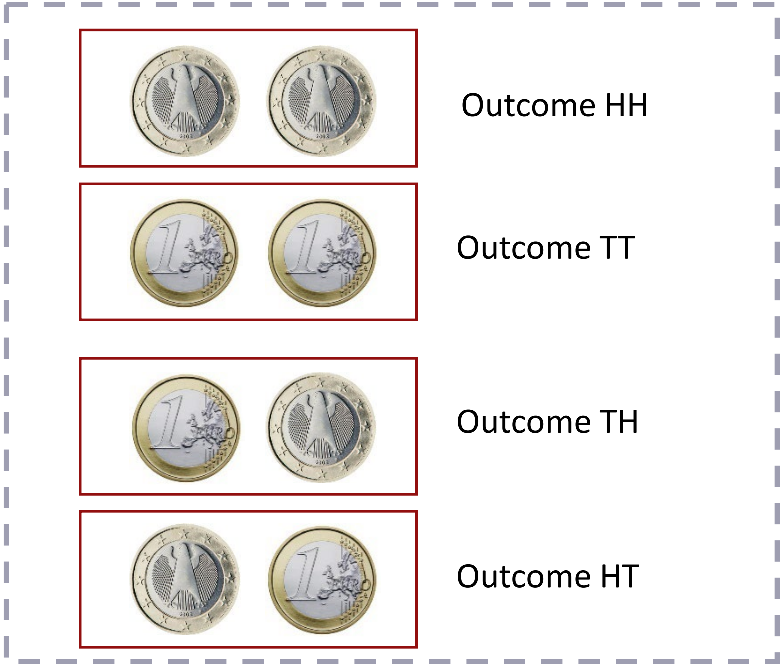
SAMPLE SPACE

Event A:
tossing at least one head

Event B:
tossing at least one tail

Probability: definitions

Experiment



SAMPLE SPACE

If an event is an empty set of outcomes, it is a **null event**, denoted by Ø

Event C:
Tossing three heads

Probability: classic definition

"The probability of an event is ***the ratio of the number of cases favorable to it to the number of all cases possible****" when nothing leads us to expect that any one of these cases should occur more than any other, which renders them, for us, equally possible*


Daniel Bernoulli
(1700 – 1782)


Pierre-Simone Laplace
(1749 – 1827)

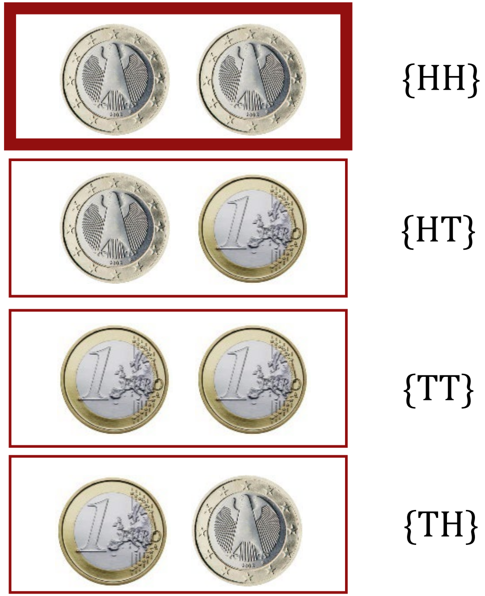
If a random experiment can result in N **mutually exclusive** and **equally likely** outcomes, and if event A results from the occurrence of N_A of these outcomes, then the probability of A is defined as

$$\Pr[A] = \frac{N_A}{N}$$

Probability: classic definition

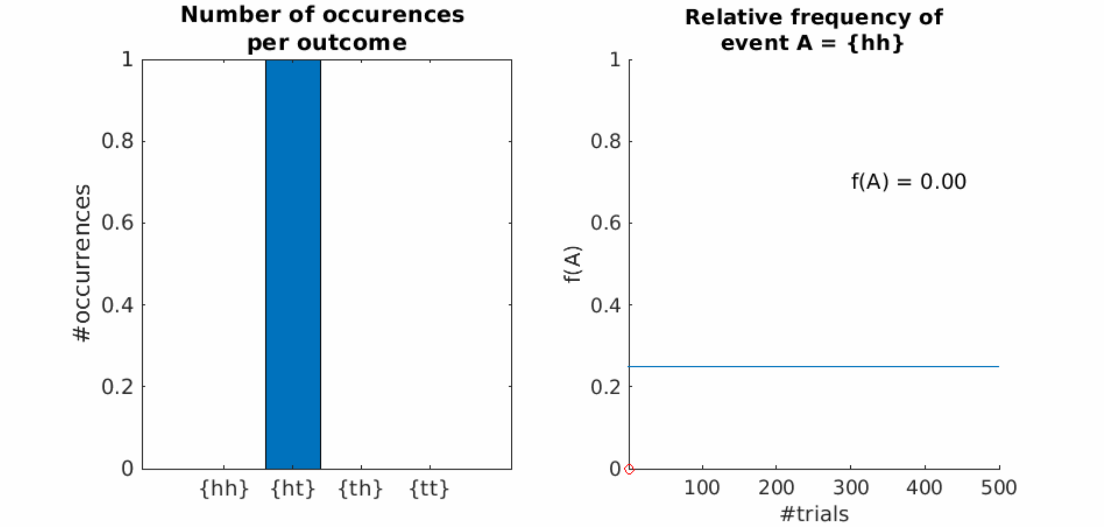
What is the probability of obtaining two heads?

$$\Pr[\{HH\}] = 1/4$$



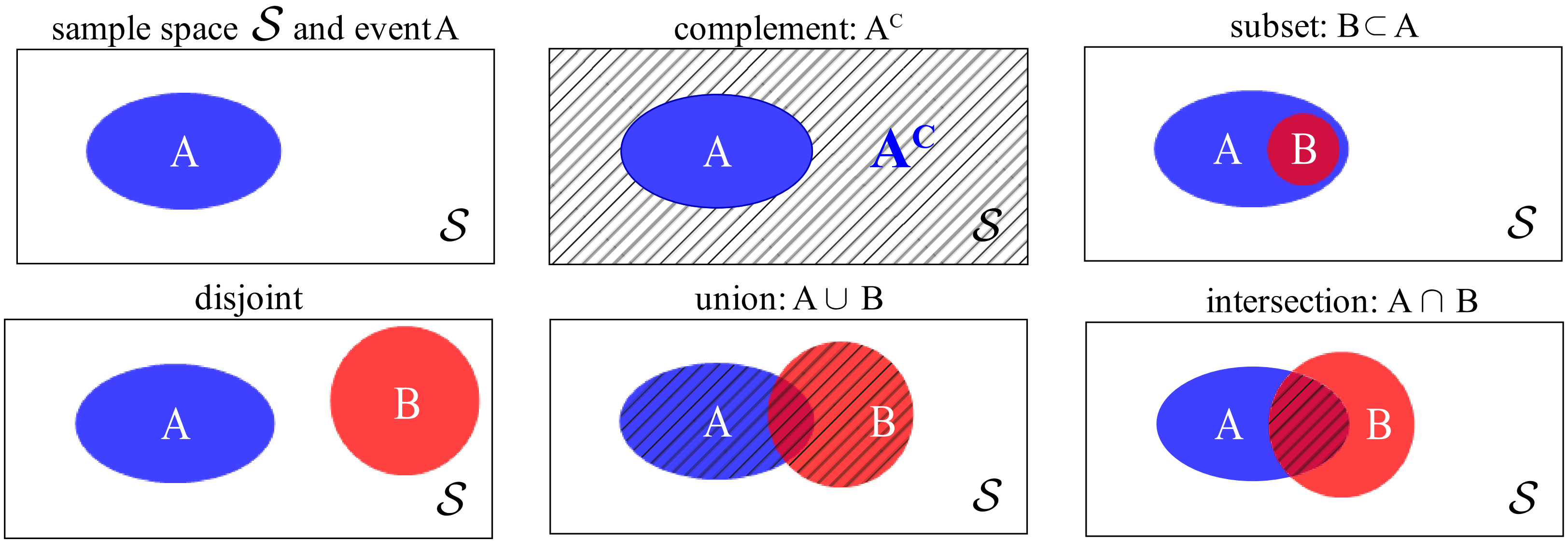
Probability: frequentist definition

$$f_A = \frac{\text{number of occurrences of event } A}{\text{total number of observations}} = \frac{N(A)}{N}$$

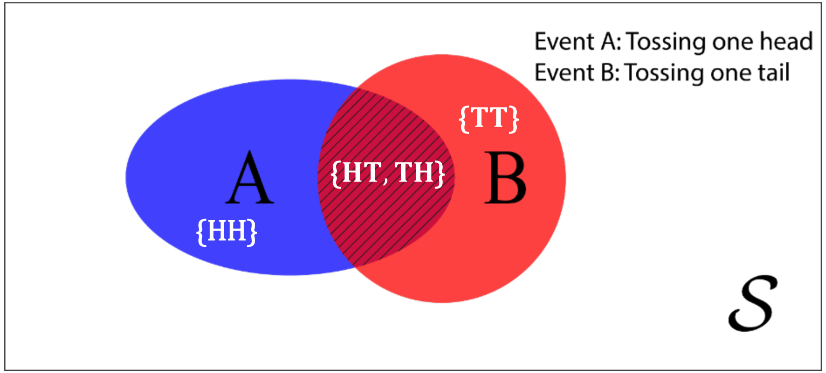


$$\Pr[A] = \lim_{N\to\infty} f_A = \lim_{N\to\infty} \frac{N(A)}{N}$$

Introduction: Venn diagrams



Probability Axioms



$$\begin{aligned} \Pr[A] &= \Pr[\{HH, HT, TH\}] \\ \Pr[B] &= \Pr[\{TT, HT, TH\}] \\ \Pr[\mathcal{S}] &= \Pr[\{HH, TT, HT, TH\}] = 1 \end{aligned}$$

$$\Pr[A] = \Pr[\{HH, HT, TH\}] =$$
$$= \Pr[\{HH\}] + \Pr[\{HT\}] + \Pr[\{TH\}] =$$
$$= 1/4 + 1/4 + 1/4 = 3/4$$

Probability Axioms:

1. $0 \leq \Pr[A] \leq 1$
2. $\Pr[\mathcal{S}] = 1$
3. $\Pr[A_1 \cup A_2 \cup \ldots \cup A_M] = \Pr[A_1] + \Pr[A_2] + \ldots + \Pr[A_M]$ $\qquad A_1, \ldots, A_M$, set of disjoint events

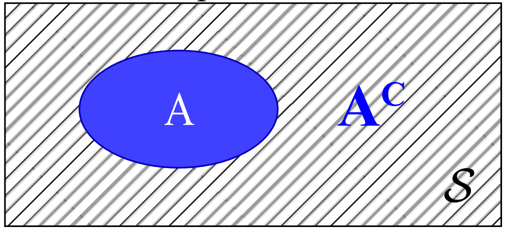
Consequences of probability axioms

- $\Pr[\emptyset] = 0$

Consequences of probability axioms

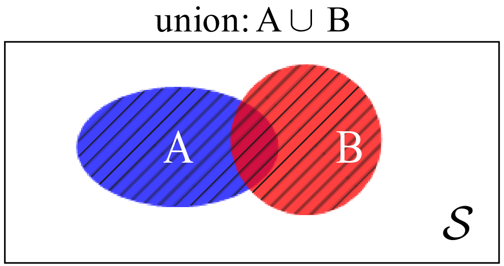
- $\Pr[\emptyset] = 0$
- $\Pr[A^C] = 1 - \Pr[A]$

complement: A^C



Consequences of probability axioms

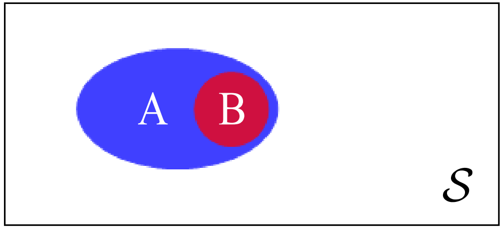
- $\Pr[\emptyset] = 0$
- $\Pr[A^C] = 1 - \Pr[A]$
- For any events A and B, $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$



Consequences of probability axioms

- $\Pr[\emptyset] = 0$
- $\Pr[A^C] = 1 - \Pr[A]$
- For any events A and B, $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$
- If $B \subseteq A$ it holds that $\Pr[B] \leq \Pr[A]$

subset: B⊂A

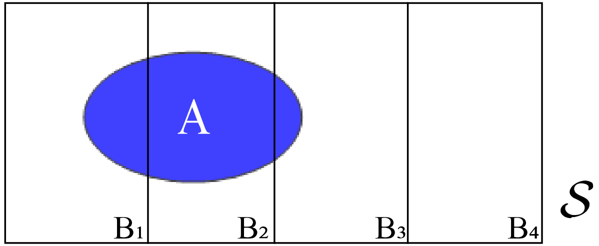


Consequences of probability axioms

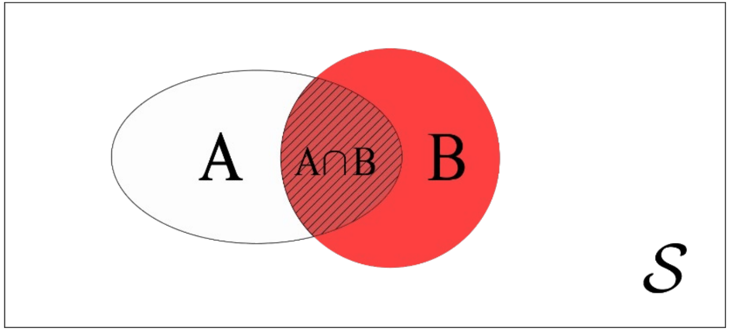
- $\Pr[\emptyset] = 0$
- $\Pr[A^C] = 1 - \Pr[A]$
- For any events A and B, $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$
- If $B \subseteq A$ it holds that $\Pr[B] \leq \Pr[A]$
- For any event A and event space $\{B_1, B_2, \ldots, B_m\}$, it holds that

$$\Pr[A] = \sum_{i=1}^{m} \Pr[A \cap B_i]$$

event space {B₁, B₂, B₃, B₄} and event A



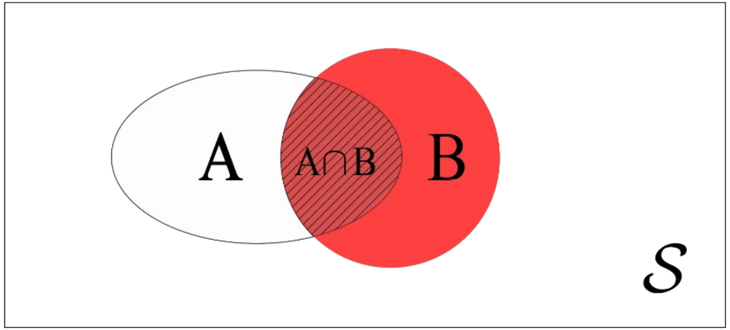
Conditional probability



$$\Pr[A|B] = \frac{\Pr[AB]}{\Pr[B]} = \frac{\Pr[A \cap B]}{\Pr[B]}$$

- $\Pr[A]$ is the *a priori* knowledge about the occurrence of the event A, before an experiment takes place.
- The conditional probability $\Pr[A|B]$ is the knowledge about the occurrence of A when we know that B has occurred (*a posteriori*)
- The event *B* becomes the sample space

Conditional probability

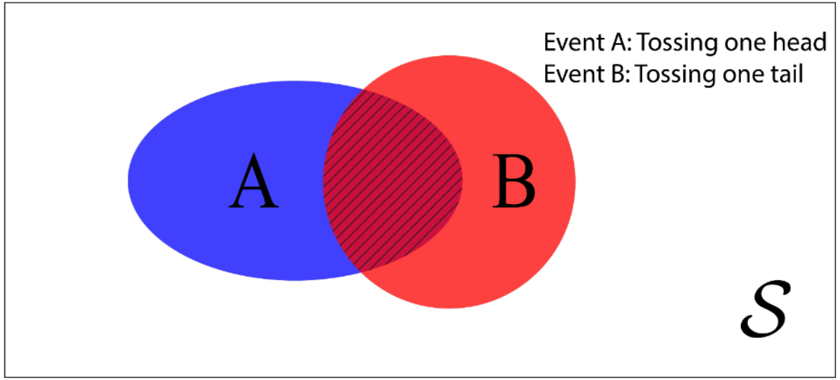


$$\Pr[A|B] = \frac{\Pr[AB]}{\Pr[B]} = \frac{\Pr[A \cap B]}{\Pr[B]}$$

Properties of conditional probability:

1. $\Pr[A|B] \geq 0$
2. $\Pr[B|B] = 1$
3. $\Pr[A|B] = \Pr[A_1|B] + \Pr[A_2|B] + \ldots + \Pr[A_M|B]$ $\qquad$ $A_1, \ldots, A_M$, set of disjoint events

Example Conditional Probability



Double head

{HH}

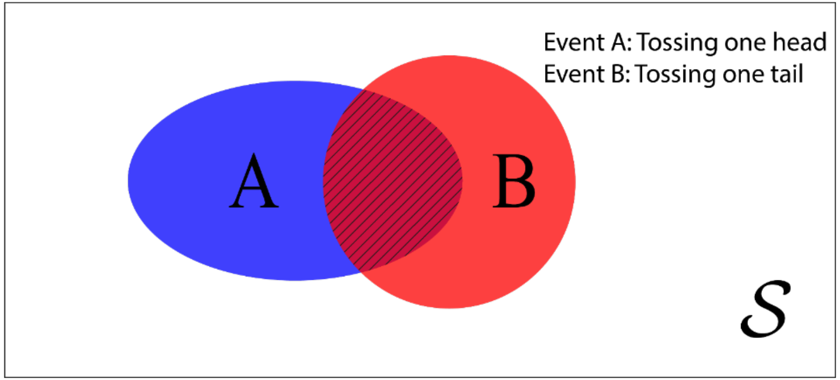
First toss head

{HH, HT}

First toss tail

{TT, TH}

Example Conditional Probability



Double head

$\Pr[\{HH\}] = 1/4$

First toss head

$\Pr[\{HH, HT\}] = 1/2$

First toss tail

$\Pr[\{TT, TH\}] = 1/2$

Example Conditional Probability

- Probability of having two heads in two-coin tosses, given the first toss is a head:

$$\Pr[\{\mathrm{HH}\}|\{\mathrm{HH},\mathrm{HT}\}] = \frac{\Pr[\{\mathrm{HH}\} \cap \{\mathrm{HH},\mathrm{HT}\}]}{\Pr[\{\mathrm{HH},\mathrm{HT}\}]} = \frac{1/4}{1/2} = 1/2$$

- Probability of having two heads in two-coin tosses, given the first toss is a tail:

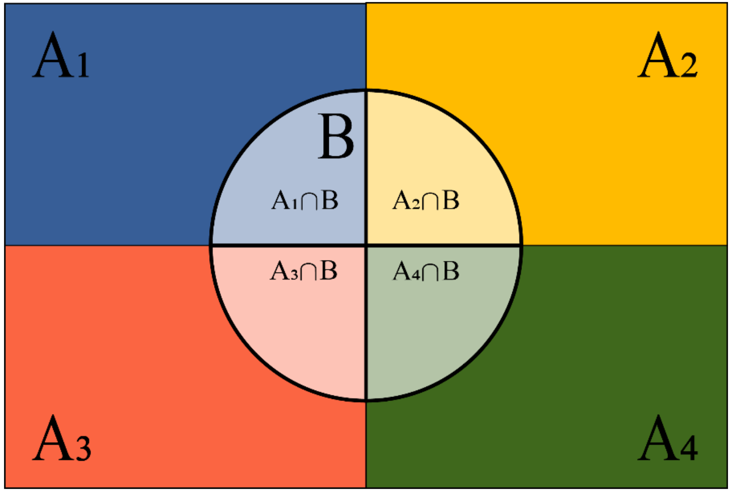
$$\Pr[\{\mathrm{HH}\}|\{\mathrm{TH},\mathrm{TT}\}] = \frac{\Pr[\{\mathrm{HH}\} \cap \{\mathrm{TT},\mathrm{TH}\}]}{\Pr[\{\mathrm{TT},\mathrm{TH}\}]} = \frac{0}{1/2} = 0$$

$$\Pr[A \mid B] = \frac{\Pr[AB]}{\Pr[B]} = \frac{\Pr[A \cap B]}{\Pr[B]}$$

Law of total probability

Sample space partitioned into pairwise disjoint events A_i.

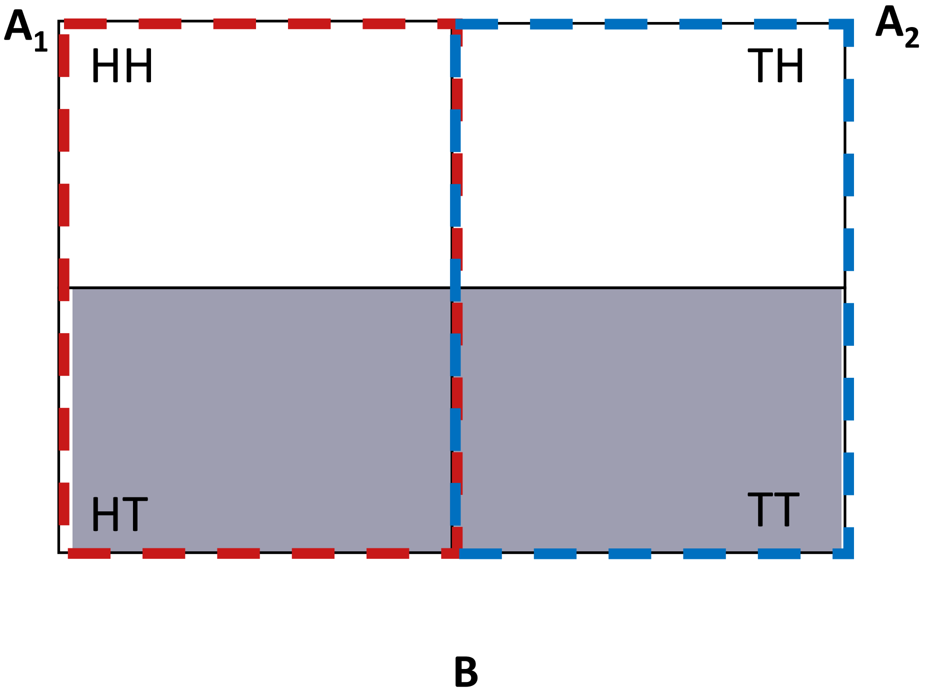
$$\Pr[B] = \sum_i \Pr[A_i \cap B] = \sum_i \Pr[B \mid A_i]\Pr[A_i]$$



Example: Total Probability

- Two disjoint events:
  - **A_1** = First toss a head
  - **A_2** = First toss a tail
- Consider the event:
  - **B** = Second toss a tail

$$\Pr[B] = \Pr[A_1 \cap B] + \Pr[A_2 \cap B]$$



Bayes theorem

- From the definition of conditional probability

$$\Pr[A \mid B]\Pr[B] = \Pr[A \cap B] = \Pr[B \cap A] = \Pr[B \mid A]\Pr[A].$$

$$\underbrace{\Pr[B \mid A]}_{\text{posterior}} = \frac{\overbrace{\Pr[A \mid B]}^{\text{likelihood}}\overbrace{\Pr[B]}^{\text{prior}}}{\underbrace{\Pr[A]}_{\text{evidence}}}.$$

"how to update or revise the strengths of evidence-based beliefs in light of new evidence a posteriori"

Where is J from?

European person by pseudonym of J

Description:
J is blond and tall. He/she preferably moves around by bike. His/her lunch is typically simple and convenient: she/he eats bread and cheese and a glass of milk. Sometimes, when it's cold, he/she also likes to have a cup of soup. His/Her favorite is made of peas.

Is J Dutch?

Where is J from?

European person by pseudonym of J

Description:
J is blond and tall. He/she preferably moves around by bike. His/her lunch is typically simple and convenient: she/he eats bread and cheese and a glass of milk. Sometimes, when it's cold, he/she also likes to have a cup of soup. His/Her favorite is made of peas.

Where is J from?

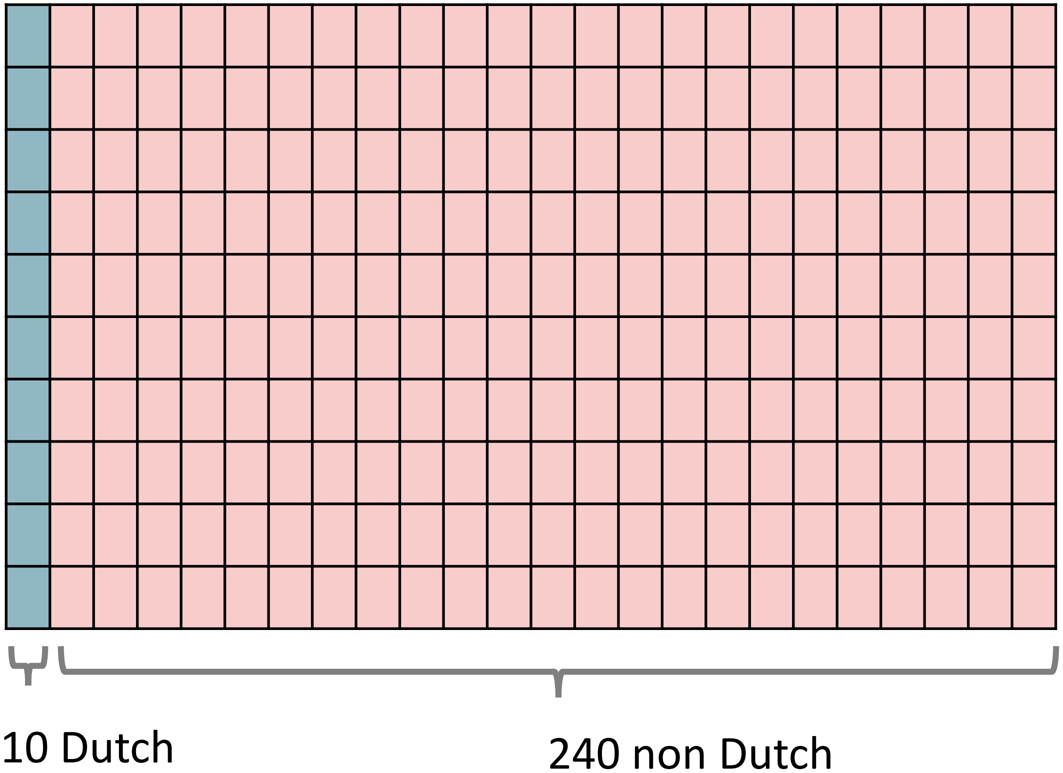
European person by pseudonym of J

European union population: 445 millions
Dutch population: 17 millions

1 in 25

Geometrical interpretation Bayes theorem

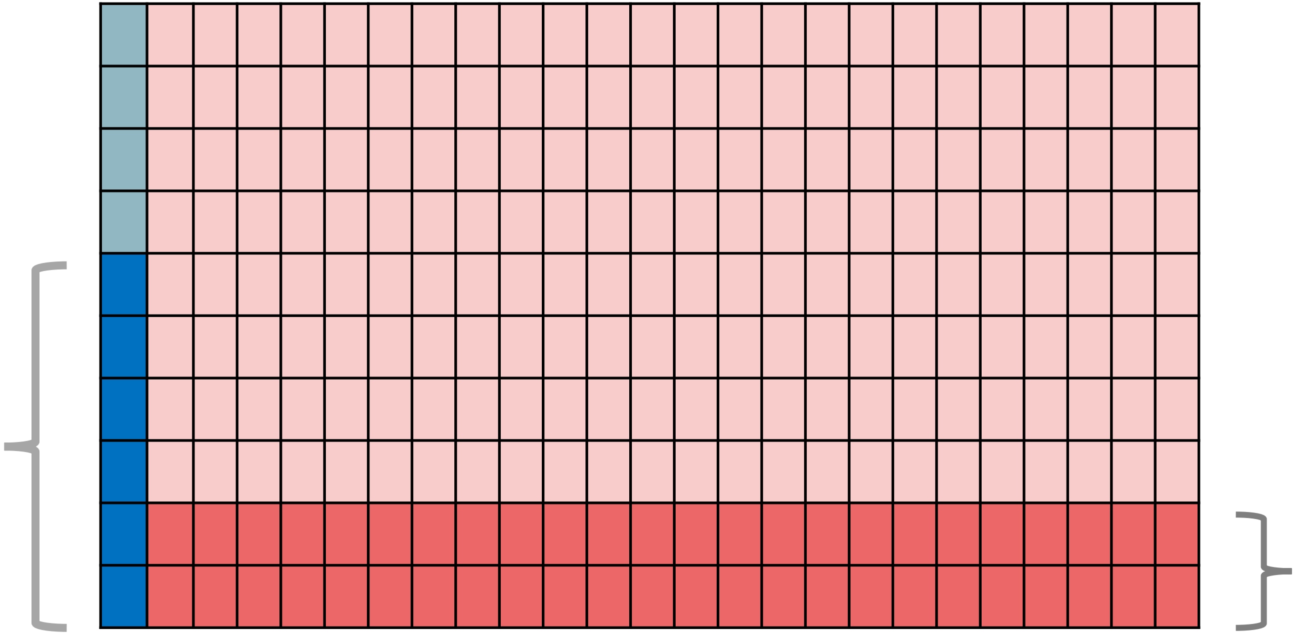
Sample of 250 people from EU



Geometrical interpretation Bayes theorem

Sample of 250 people from EU

60% of Dutch
fit description



20% of non-Dutch
fit description

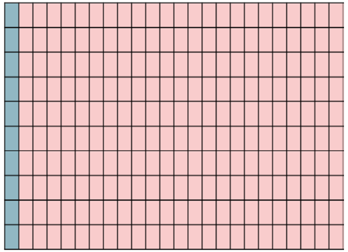
$$\Pr[\mathit{Dutch} \mid \mathit{description}] = \frac{\Pr[\mathit{description}|\mathit{Dutch}]\Pr[\mathit{Dutch}] =}{\Pr[\mathit{description}]} = \frac{6}{6+48} \approx 11\%!$$

Interpretation Bayes theorem

- **Prior**: 4% of EU population is Dutch
- **Evidence**: description, 60% of Dutch people fit description
- **Posterior**: update prior, there is a 11% chance that the person described is Dutch

Prior

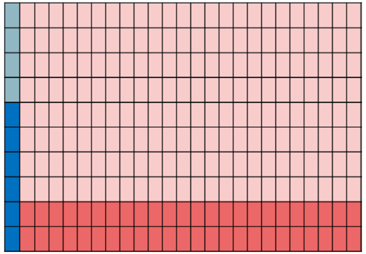
$\Pr[Dutch]$



(all possibilities)

Evidence

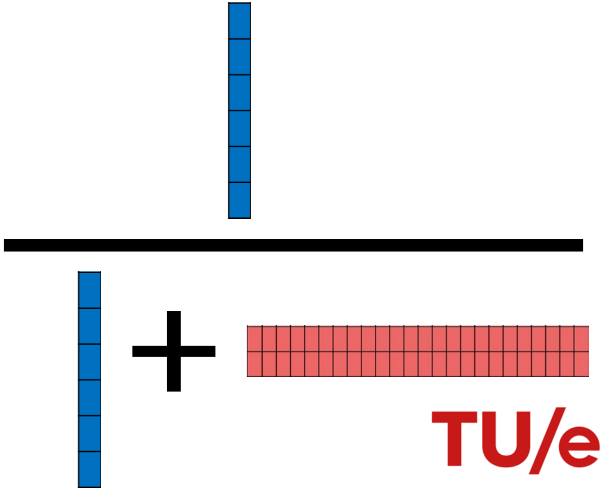
$\Pr[description]$



(possibilities fitting evidence)

Posterior

$(\Pr[Dutch \mid description])$



Independence

- Events A and B are independent if and only if:

$$\Pr[A \cap B] = \Pr[A]\Pr[B]$$

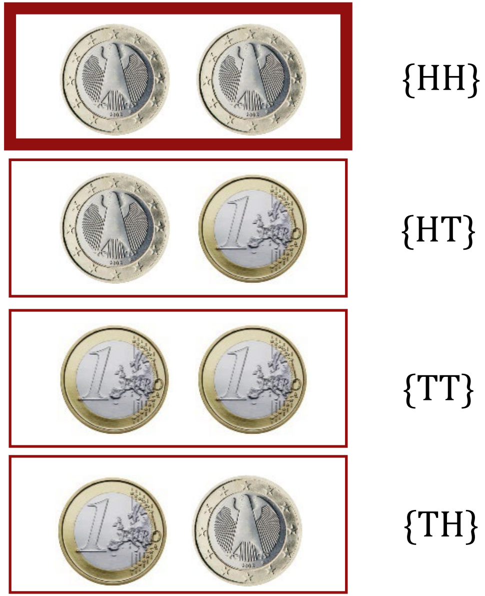
- Independent and disjoints are **NOT** synonyms

- Extension to multiple sets: Multiple sets $\{A_1, A_2, \ldots, AM\}$ are independent if and only if the following two constraints hold
  1. Every possible combination of two sets is independent
  2. It holds that $\Pr[A_1 A_2 \ldots A_M] = \Pr[A_1]\Pr[A_2]\ldots\Pr[A_M]$

Probability: classic definition

What is the probability of obtaining two heads?

$$\Pr[\{HH\}] = 1/4$$



Example independent events

Each coin toss in independent!

Double head

First toss head

Second toss head

$\Pr[\{HH, HT\}] = 1/2$

$\Pr[\{TT, TH\}] = 1/2$

$$\Pr[\{HH\}] = 1/2 \cdot 1/2 = 1/4$$

Wrap-up (I)

- **Probability** is the branch of mathematics concerned with analysis of random phenomena
- Random experiments are described in terms of ***trials (or observations)***, ***outcomes***, ***events***, and ***sample space***
- By the **classic definition**, the probability is calculated as the number of favorable outcomes over the total number of outcomes
- By the **frequentist definition**, the probability is defined as the limit for an infinitely large number of trials of the relative frequency
- Probability is governed by three **probability axioms**

Wrap-up (II)

- The **conditional probability** is the a posteriori probability of an event, given the knowledge that another event has occurred
- The **law of total probability** allows relating the *a priori* probability of an event to a sum of conditional probabilities over several distinct events, which partition the sample space
- **Bayes theorem** permits updating probabilities in view on new evidence
- Two events are **independent** when the outcome of one event does not influence the outcome of the other
- Two events are **disjoint** when they cannot occur at the same time

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Statistical signal processing (5CTA0)

Lecture 1, part A

Lecturer: Simona Turco

Electrical Engingeering, Signal Processing Systems group