Mathematical Methods for Automation Engineering M
– Introduction to Probability Theory –

Andrea Mentrelli

Department of Mathematics &
Alma Mater Research Center on Applied Mathematics, AM^2
University of Bologna

andrea.mentrelli@unibo.it

A.Y. 2024/25

Deterministic Experiment vs. Random Experiment

Let us set up an *experiment* to measure the free fall time t of a heavy body from a height h

- we repeat the experiment several times, and we notice that the results are all quite similar
- the more accurate the clock, the more similar the results

There is a law allowing us to **foresee** the outcome of such an experiment:

$$t = \sqrt{\frac{2h}{g}}$$

Deterministic experiment is an experiment that we alredy expect the outcomes.

This is a **deterministic experiment**

Deterministic Experiment vs. Random Experiment

Let us set up an experiment to measure the lifetime T of the electric bulbs assembled by a production line

- repeating the experiment several times (with different bulbs), we find out that the lifetime T may vary considerably
- the uncertainty on T has nothing to do with the precision of our instruments

We do not have a reasonably easy way to foresee the outcome of this experiment → because of property variations

Nonetheless, we find out that there is a *statistical regularity* that allows us to **approximately foresee** the features of the collected data

This is a **random experiment**

Definition (Random experiment)

We call *random experiment*, an experiment that has several possible (predictable) outcomes, among which the one that will actually occur cannot be predicted

Peculiar features of random experiments
(may be present in largely varying degrees)

$t = [0, +\infty[$

- **Uncertainty of the outcome**. The possible outcomes are either countable ($\rightarrow$ roll of a die) or uncountable ($\rightarrow$ lifetime of a light bulb)
- **Repeatibility of the experiment**. Many experiments are (assumed to be) repeatable, others are not
- **Equal probabilities of the possible outcomes**. Sometimes the possible outcomes are *equally likely*, other times are not

In the real world

- many experiments are **not repeatable**
- outcomes are seldom **equally likely**

Focus of the theory of probability

- **Originally**: uncertain, repeatable, equally likely outcomes
- **Nowadays**: uncertain experiments

Introduction to Probability Theory

The first attempts to define probability were made by **Pascal** and **Fermat**, but only in 1812 **Laplace** gave the first formal definition

Definition (Classical definition of probability)

Given a well-identified random experiment, for which there are n equally likely possible outcomes, the probability of an event$^{(*)}$ E made of m outcomes is

well suited for gambling

$$\mathcal{P}(E) = \frac{m}{n}$$

outcomes

all possible outcomes

(*) An event E is a set of outcomes. For example, in the experiment "rolling a die", the event *"even"* is made up of three outcomes (2, 4 and 6)

Some **remarks on the classical definition of probability**

- well-suited for gambling
- the property $0 \leqslant \mathcal{P}(E) \leqslant 1$ holds

$P(E) = \frac{m}{n}$

if n is infinite, the formula is not possible

but...

- it does not work when the possible outcomes are **infinite**
- **equally likely outcomes** are assumed and...
- ...it is **tautological**

Introduction to Probability Theory

Richard von Mises proposed in 1928 the so-called **frequentist definition of probability**

clearly explains what our experiment is

Richard von Mises (1883–1953)

Definition (Frequentist probability)

Given a well-identified and repeatable random experiment, let E be an event and let $f_n(E)$ be the number of occurrences of E in a sequence of n repetitions of the experiment *(absolute frequency)*. The probability of E is defined as the limit of the *relative frequency* as n increases

$$\mathcal{P}(E) = \frac{f_n(E)}{n} \qquad n \to \infty$$

not an analytical calculus limit

n. of occurences

number of repetition of experiments

let's say n is a very large number

Some **remarks on the frequentist definition of probability**

- it enjoys all the virtues of the classical definition
- the drawbacks of the classical definition are avoided

but...

- repeatability of the experiment is required
- the "limit" is not an analytical one, so how do we know that it exists? and how can we be sure that, even if it exists, it will be the same for each possible sequence of repetitions of the experiment?

Classic, frequentist (and others...) probabilities are **interpretations that lead to the same formal framework** $\implies$ we can think of refraining from giving any interpretation and build the theory out of axioms, paying no attention to the underlying meaning

Modern **probability theory** is **axiomatic**

Andrej Nikolaevič Kolmogorov
(1903–1987)

"The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry and Algebra. This means that after we have defined the elements to be studied and their basic relations, and have stated the axioms by which these relations are to be governed, all further exposition must be based exclusively on these axioms, independent of the usual concrete meaning of these elements and their relations." (1933)

http://www.mathematik.com/Kolmogorov/

Formulation of the **axiomatic theory of probability**

- definition of the **primitives**
- formulation of the **axioms**
- proof of **theorems**, making use of the axioms

Formulation of the **axiomatic theory of probability**

- definition of the **primitives**
- formulation of the **axioms**
- proof of **theorems**, making use of the axioms

primitive

Definition (Sample space)

The set of all possible outcomes of an experiment is known as the *sample space* of the experiment and is denoted by Ω

Note that we do not know in advance the outcome of the experiment, but we *do* know the set of all the possible outcomes (i.e. the sample space Ω)

If the sample space is **discrete** (countable or not), we shall denote its elements as ω_i discrete element

sample space

$$\Omega = \{\omega_1, \omega_2, \dots\}$$

The sample space can be **continuous** as well, for instance

$$\Omega = [0, 1]$$

continuous

Example (Flipping a coin)

The sample space Ω of the experiment "flipping a coin" is discrete and countable. The outcomes are H (heads) and T (tails)

$$\omega_1 : H, \qquad \omega_2 : T; \qquad \Omega = \{\omega_1, \omega_2\}$$

sample space of tossing a coin

Example (Flipping two coins) let's say two coins are distinguishable

The sample space Ω of the experiment "flipping two coins" is discrete and countable, too

sample space of tossing 2 coins

$$\omega_1 : HH, \quad \omega_2 : TT, \quad \omega_3 : TH, \quad \omega_4 : HT; \quad \Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

Example (Tossing a die)

When tossing a die, supposing that all six sides are equally likely to appear, the sample space is (ω_k is the outcome "upturn side k")

$$\Omega = \{\omega_1, \omega_2, \dots \omega_6\}$$

Example (Flipping a coin, again and again)

Let's consider the experiment of flipping a coin until we get "heads"

$$\omega_1 : H, \quad \omega_2 : TH, \quad \omega_3 : TTH, \dots \quad \omega_n = \overbrace{TTT\dots T}^{n-1} H, \quad \dots$$

$$\Omega \equiv \{\omega_1, \omega_2, \dots\}$$

The sample set Ω is infinite and countable

Example (Random number generator) RNG

The sample space Ω of the experiment consisting in generating a random number in the interval from 0 to 1 is continuous

$$\Omega = [0, 1]$$

Example (Time to failure)

The experiment consisting in measuring the time to failure of an electronic device can be assumed to be continuous and unbounded

$$\Omega = [0, \infty[$$

Definition (Event)

We call **event** a subset of the sample space Ω. In other words, an event is a set consisting of possible outcomes of the experiment

$$E \subseteq \Omega$$

subset

Definition (Elementary event)

An **elementary event** (also called an **atomic event** or **simple event**) is an event which contains only a single outcome in the sample space ω

$$E = \{\omega\}$$

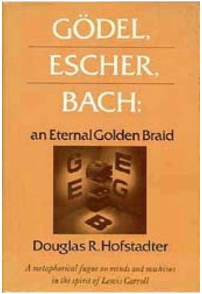
Note that "elementary event" $\neq$ "outcome"

Events and sets are isomorphic $\Rightarrow$ **algebra of events**

Isomorphism?

"The word *isomorphism* applies to two complex structures that can be mapped onto each other in a way that to each part of one structure there is a corresponding part in the other structure; the word corresponding means that the two parts play similar roles in their respective structures."

(Douglas R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, 1979)



two complex structures that map onto each other.

Introduction to Probability Theory

Algebra of events: given the isomorphism between sets and events, we borrow the following concepts/terminology used in set theory

- **union** (or **sum**) C of the events A and B

a new event named C

$$C = A \cup B \qquad (C = A + B)$$

union

events are interchangable with sets

(events ≈ sets)

- **intersection** (or **product**) C of the events A and B

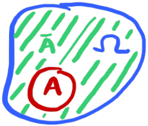
$$C = A \cap B \qquad (C = AB)$$

intersection

- **complement** (or **negation**) of the event A

all other outcomes that do not include in an event

$$\overline{A}$$



- **difference** of the events A and B

$$C = A - B = A - A \cap B = A \cap \overline{B}$$

Note: $\overline{\Omega} = \emptyset$, $\overline{\emptyset} = \Omega$ ($\emptyset$ is called *null event*)

Algebra of events: given the isomorphism between sets and events, we borrow the following concepts/terminology used in set theory

1. **Idempotence**

$$A \cup A = A \qquad A \cap A = A$$

2. **Associative laws**

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

3. **Commutative laws**

$$A \cup B = B \cup A \qquad A \cap B = B \cap A$$

4. **Distributive laws**

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

Algebra of events: given the isomorphism between sets and events, we borrow the following concepts/terminology used in set theory

5. **Identity laws**

$$A \cup \emptyset = A \qquad A \cap \Omega = A$$
$$A \cup \Omega = \Omega \qquad A \cap \emptyset = \emptyset$$

6. **Complement laws**

$$A \cup \overline{A} = \Omega \qquad A \cap \overline{A} = \emptyset$$
$$\overline{\overline{A}} = A \qquad \overline{\Omega} = \emptyset \qquad \overline{\emptyset} = \Omega$$

7. **DeMorgan's laws**

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

Definition (Mutually exclusive, or incompatible, events)

The events $A, B \subseteq \Omega$ are *mutually exclusive* (or *incompatible*) if they cannot occur simultaneously

no elements in common

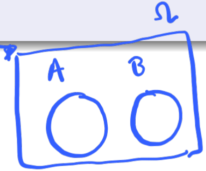
$$\boxed{A \cap B = \emptyset}$$

The events $E_i \subseteq \Omega$ $(i = 1, 2, \ldots)$ are *mutually exclusive* if

mutually exclusive, no intersection

$$E_i \cap E_j = \emptyset, \qquad \forall i \neq j$$

for any



Definition (Exhaustive events)

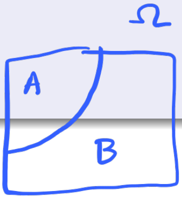
The events $A, B \subseteq \Omega$ are *exhaustive* if their union event is the sample space Ω

$$\boxed{A \cup B = \Omega}$$

The events $E_i \subseteq \Omega$ $(i = 1, 2, \dots)$ are *exhaustive* if

$$\bigcup_{i=1}^{\infty} E_i = \Omega$$

exhaustive



Definition (Partition)

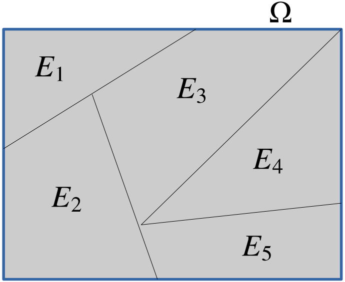
A set of exhaustive and mutually exclusive events is called a *partition* of Ω

The events $E_i \subseteq \Omega$ $(i = 1, 2, \ldots n)$ are a partition of Ω if and only if

$E_i \cap E_j = \emptyset$ → mutually exclusive

$\bigcup_{i=1}^{n} E_i = \Omega$ → exhaustive events

- $\bigcup_{i=1}^{n} E_i = \Omega$ (exhaustive events)
- $E_i \cap E_j = \emptyset \quad \forall i \neq j$ (mutually exclusive events)



if we collect all of the events, we get the sample space

→ no overlapping for each event

collection / set of events

Definition (Field)

only n is a finite number

A set of events $\mathcal{E}$ is a *field* if it is closed with respect to negation and union

field

$\bigcup_{k=1}^{n} A_i \in \mathcal{E}$

- $A \in \mathcal{E} \quad \Rightarrow \quad \overline{A} \in \mathcal{E}$ — the negation is a part of $\mathcal{E}$ too
- $A, B \in \mathcal{E} \quad \Rightarrow \quad \boxed{A \cup B \in \mathcal{E}}$ $\Rightarrow A \cap B \in \mathcal{E}$, $A - B \in \mathcal{E}$

belongs to

If $\mathcal{E}$ is a field, and $A, B \in \mathcal{E} \implies A \cap B,\ A - B \in \mathcal{E}$

Any operation on $A, B \in \mathcal{E}$ gives a result that belongs to $\mathcal{E}$ as well; however, this property does not necessarily hold for union/intersection of infinite events

any operation of elements inside $\mathcal{E}$ doesn't produce events outside $\mathcal{E}$

Definition (Borel field)

Given a field $\mathcal{E}$, if the union/intersection of an infinite number of events of $\mathcal{E}$ belong to $\mathcal{E}$, then $\mathcal{E}$ is a *Borel field*

Example

In the experiment "tossing a die", the sample space is $\Omega = \{\omega_1, \ldots \omega_6\}$. Is the following class a field?

null event

$$\Big\{\emptyset, \quad \Omega, \quad \overbrace{\{\omega_1, \omega_3, \omega_5\}}^{\text{odd number events}}, \quad \overbrace{\{\omega_2, \omega_4, \omega_6\}}^{\text{even number events}}, \quad \overbrace{\{\omega_1\}}^{\text{atomic event}}\Big\} = \mathcal{E}\,?$$

No, it is not. Even thought $\{\omega_1\}$ and $\{\omega_2, \omega_4, \omega_6\}$ belong to the class, their union $\{\omega_1\} \cup \{\omega_2, \omega_4, \omega_6\} = \{\omega_1, \omega_2, \omega_4, \omega_6\}$ does not belong to the class.

1. Is $A \in \mathcal{E} \Rightarrow \bar{A} \in \mathcal{E}$?
2. Is $A, B \in \Rightarrow A \cup B \in \mathcal{E}$?

Example

If $\{\omega_1\}$ is removed from the previous class, we are left with a field

for every A, B must belongs to $\mathcal{E}$ ($\forall A, B \in \mathcal{E}$)

Example

must be clarified first

Given the sample space $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, a possible fields over it is

$$\mathcal{E} = \{\, \emptyset,\ \{\omega_1, \omega_2\},\ \{\omega_3, \omega_4\},\ \{\omega_1, \omega_2, \omega_3, \omega_4\}\,\}$$

but what is the most general field?

The most general field is the one containing all the possible subsets of Ω (including $\emptyset$ and Ω itself)

empty — one element — two element

$$\mathcal{E} = \{\, \emptyset,\ \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3\},$$
$$\{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_3, \omega_4\},$$
$$\{\omega_2, \omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3, \omega_4\}\,\}$$

three element

four element, also the Ω

$\boxed{\text{class} = \text{set}}$

Definition (Probability)

The probability is a number associated to an event

Remarks

- In the axiomatic framework, the probability is a primitive concept (therefore, it is vague)
- We accept that probability is a number because we intuitively assume that probability implies an order (we place an event in between two extreme events, namely the *null* event and the *certain* event)

Formulation of the **axiomatic theory of probability**

- definition of the **primitives**
- formulation of the **axioms**
- proof of **theorems**, making use of the axioms

Axiomatic Theory of Probability

we're going to assign the probability to the event NOT the outcome

outcome is NOT a mathematical object

- A sample space Ω is given
- A field $\mathcal{E}$ over Ω is given

field definition

1. $A \in \mathcal{E} \quad \Rightarrow \quad \overline{A} \in \mathcal{E}$
2. $A, B \in \mathcal{E} \quad \Rightarrow \quad A \cup B,\ A \cap B,\ A - B \in \mathcal{E}$

$\mathcal{E}$ includes all the possible events (≡ all possible subsets of Ω)

- There exists a function $\mathcal{P}$ associating to each event of $\mathcal{E}$ a number, for which the following properties hold (**Kolmogorov's axioms**)

probability properties

1. $\mathcal{P}(A) \geqslant 0 \qquad \forall A \in \mathcal{E}$
2. $\mathcal{P}(\Omega) = 1$
3. $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) \qquad \forall A, B \in \mathcal{E} \quad \text{with} \quad A \cap B = \emptyset$

↳ these make sense only we only work on field ($\mathcal{E}$)

Remark. The third axiom is readily extended to the case of n incompatible events

e.g. $P(E_1 \cup E_2) = P(E_1 \cup E_2) = P(E_1) + P(E_2)$

$$\mathcal{P}\left(\cup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mathcal{P}\left(E_i\right) \qquad E_i \cap E_j = \emptyset \quad \forall i \neq j$$

The extension to the case of an infinity of incompatible events

$$\mathcal{P}\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathcal{P}\left(E_i\right) \qquad E_i \cap E_j = \emptyset \quad \forall i \neq j$$

is not straightforward. This has to be seen as an additional axiom which holds when $\mathcal{E}$ is a Borel field

Definition (Probability space)

The tuple $(\Omega, \mathcal{E}, \mathcal{P})$ is called *probability space*

a collection of

Formulation of the **axiomatic theory of probability**

- definition of the **primitives**
- formulation of the **axioms**
- proof of **theorems**, making use of the axioms

Introduction to Probability Theory

Theorem (Probability of negation)

The probability of non-occurrence of the event E, is the 1-complement of the probability of the event E

$$\mathcal{P}\left(\overline{E}\right) = 1 - \mathcal{P}\left(E\right)$$

Proof. Since E and $\overline{E}$ are mutually exclusive ($E \cap \overline{E} = \emptyset$) and exhaustive ($E \cup \overline{E} = \Omega$), we have

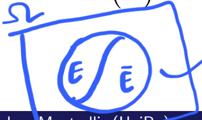
$$P(\Omega) = P(E \cup \overline{E}) = \underbrace{P(E) + P(\overline{E})}$$

$$1 = \mathcal{P}\left(\Omega\right) = \mathcal{P}\left(E \cup \overline{E}\right) = \mathcal{P}\left(E\right) + \mathcal{P}\left(\overline{E}\right)$$

$$1 = P(E) + P(\overline{E})$$

Therefore: $\mathcal{P}\left(\overline{E}\right) = 1 - \mathcal{P}\left(E\right)$ □

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



$E \cap \overline{E} = \emptyset$

$E \cup \overline{E} = \Omega$

Theorem (Probability of the null event)

The probability of the null event is zero

$$\mathcal{P}(\emptyset) = 0$$

Proof. Making use of the previous theorem and of the second axiom, we have

$$\mathcal{P}(\emptyset) = \mathcal{P}\left(\overline{\Omega}\right) = 1 - \mathcal{P}(\Omega) = 0$$

$\square$

Theorem (Probability is not greater than 1)

The probability of any event E is less than or equal to 1

$$\mathcal{P}\left(E\right) \leqslant 1$$

Proof. We have already proven that $\mathcal{P}\left(E\right) = 1 - \mathcal{P}\left(\overline{E}\right)$. Making use of the first axiom, i.e. $\mathcal{P}\left(\overline{E}\right) \geqslant 0$, we have

$$\mathcal{P}\left(E\right) = 1 - \mathcal{P}\left(\overline{E}\right) \implies \mathcal{P}\left(E\right) \leqslant 1$$

$\square$

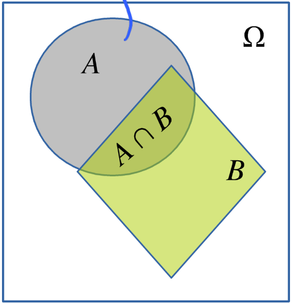
From the first axiom and this theorem, it follows: $0 \leqslant \mathcal{P}\left(E\right) \leqslant 1 \ \forall E$

Theorem (Probability of the union of events)

The probability of the union of any two events A and B is

$$\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B) \quad \forall A, B$$

not mutually exclusive



Proof. $A \cup B$ and B can be written in terms of union of mutually exclusive events

$$\underbrace{A \cup B = A \cup (\overline{A} \cap B) \qquad B = (A \cap B) \cup (\overline{A} \cap B)}_{\text{to show these are mutually exclusive}}$$

Therefore

$$\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(\overline{A} \cap B)$$
$$\mathcal{P}(B) = \mathcal{P}(A \cap B) + \mathcal{P}(\overline{A} \cap B)$$

and so: $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$ □

Remark. If A and B are mutually exclusive ($A \cap B = \emptyset$), from the previous theorem the third axiom is recovered: $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B)$

Note that this theorem, *obviously*, is not a proof of the axiom (why?)

Theorem (Inclusion/exclusion identity)

The previous theorem is readily generalized to the case of n events $A_1, A_2, \ldots A_n$*:*

$$\mathcal{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathcal{P}(A_i) - \sum_{i<j} \mathcal{P}(A_i \cap A_j) + \sum_{i<j<k} \mathcal{P}(A_i \cap A_j \cap A_k) + \ldots + (-1)^{n+1} \mathcal{P}(A_1 \cap A_2 \cap \ldots \cap A_n)$$

For three events:

$$\boxed{\begin{gathered}\mathcal{P}(A \cup B \cup C) = \mathcal{P}(A) + \mathcal{P}(B) + \mathcal{P}(C) \\ -\mathcal{P}(A \cap B) - \mathcal{P}(A \cap C) - \mathcal{P}(B \cap C) + \mathcal{P}(A \cap B \cap C)\end{gathered}}$$

Example

We pick two items from two production lines, L_1 and L_2, of the same facility. With probability 0.05, the component from line L_1 will be faulty; with probability 0.04, the component from line L_2 will be faulty; and with probability 0.03, they will be both faulty. What is the probability that none of the items is faulty?

$P(C_1) = 0.05$ $P(C_2) = 0.04$

$P(C_1 \cap C_2) = 0.03$ Asked: $P(\bar{C}_1 \cap \bar{C}_2)$?

Let C_k denote the event that the component from L_k is faulty ($k = 1, 2$). The probability that at least one the items is faulty is

$$\mathcal{P}(C_1 \cup C_2) = \underset{0.05}{\mathcal{P}(C_1)} + \underset{0.04}{\mathcal{P}(C_2)} - \underset{0.03}{\mathcal{P}(C_1 \cap C_2)} = 0.05 + 0.04 - 0.03 = 0.06$$

Because the event that none of the items is faulty is the complement of the event that at least one of them is faulty, we obtain the result

$$\mathcal{P}\left(\overline{C_1} \cap \overline{C_2}\right) = \mathcal{P}\left(\overline{C_1 \cup C_2}\right) = 1 - \mathcal{P}(C_1 \cup C_2) = 0.94$$

Example

If we were to ask someone what they thought the chances were of

- rain today → T as the event : P(T) = 0,3
- rain tomorrow → W as the event : P(W) = 0,4
- rain both today and tomorrow $P(T \cap W) = 0{,}2$
- rain either today or tomorrow $P(T \cup W) = 0{,}6$

it is quite possible that, after some deliberation, they might give 30%, 40%, 20%, and 60% as answers. Should we accept these answers?

because not all numbers are independent

$$\mathcal{P}(T) = 0.3 \quad \mathcal{P}(W) = 0.4 \quad \mathcal{P}(T \cap W) = 0.2 \quad \mathcal{P}(T \cup W) = 0.6$$

but

0,3 + 0,4 − 0,2 = 0,5 → not consistent w/ the axioms

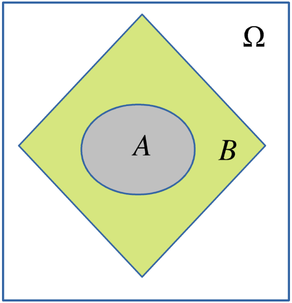
$$\mathcal{P}(T \cup W) = \mathcal{P}(T) + \mathcal{P}(W) - \mathcal{P}(T \cap W) = 0.3 + 0.4 - 0.2 \neq 0.6$$

Such answers are not consistent with the axioms of probability (One possibility we could accept is 30%, 40%, 10%, and 60%)

Theorem (Probability of a subset)

Let A and B be two events. If A is a subset of B, then the probability of A is less than or equal to the probability of B

$$A \subset B \qquad \Longrightarrow \qquad \mathcal{P}(A) \leqslant \mathcal{P}(B)$$



Proof. The event B can be written as the union of two mutually exclusive events

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

thus, from axioms I and III

$$\mathcal{P}(B) = \mathcal{P}(A) + \underbrace{\mathcal{P}\left(\overline{A} \cap B\right)}_{\geq 0} \geqslant \mathcal{P}(A)$$

only possible 0 or positive

□

In the axiomatic theory of probability, the problem of the **measure of probability** is not a critical one

The only requirement is **compatibility with axioms**

Ok, but... how to measure the probability from a practical viewpoint? How probabilities are evaluated?

- making use of experience
- starting from the results of experiments
- they can be guessed
- they can even be sensed!

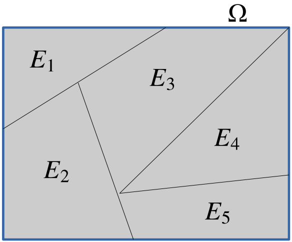
Eventually, the results must be compared to real world data and observations (the classic, frequentist and subjective approaches are still useful!) $\Rightarrow$ This is the formulation of a **model**

Definition (Equally likely partition)

A partition $\{E_1, E_2, \ldots E_n\}$ of Ω is an *equally likely partition* if all the events have the same probability

Given p the probability of each event, we have

$$
\begin{aligned}
1 = \mathcal{P}(\Omega) &= \mathcal{P}\big(\overbrace{E_1 \cup E_2 \cup \ldots \cup E_n}^{\Omega}\big) \\
&= \underbrace{\mathcal{P}(E_1) + \mathcal{P}(E_2) + \ldots + \mathcal{P}(E_n)}_{n} \\
&= n\,p
\end{aligned}
$$



therefore

$$\mathcal{P}(E_i) = p = \frac{1}{n} \qquad \forall i = 1, 2, \ldots n$$

is an element of   only one outcome

An event $A \subseteq \Omega$ is made up by m elementary equally likely events E_i

$$\mathcal{P}(A) = \mathcal{P}\left(\bigcup_{i=1}^{m} E_i\right) = \sum_{i=1}^{m} \mathcal{P}(E_i) = \frac{m}{n}$$

Thus, in a **equally likely discrete partition** of Ω, the probability of $A \subseteq \Omega$ is the ratio between the cardinality m of A (**favorable cases**) and the cardinality of Ω (**possible cases**)

$$\mathcal{P}(A) = \frac{m}{n}$$

cardinality = number of sets

The **classic probability** is formally recovered from the axioms as a particular case ($\Rightarrow$ the classic probability is compatible with the axioms)

Note that the assumption of equally likely events is arbitrary

Is the **frequentist approach** also compatible with the axioms?

$$\mathcal{P}(A) \simeq \frac{n_A}{n} \qquad (n \gg)$$

n_A: number of occurrences of A
n: total number of trials
Clearly

1st axiom

2nd axiom

$$\mathcal{P}(A) \geqslant 0 \quad \forall A \qquad \mathcal{P}(\emptyset) \simeq \frac{n_\emptyset}{n} = 0 \qquad \mathcal{P}(\Omega) \simeq \frac{n_\Omega}{n} = 1$$

Moreover, if A and B are mutually exclusive, $n_{A \cup B} = n_A + n_B$, and so

$$\mathcal{P}(A \cup B) \simeq \frac{n_{A \cup B}}{n} = \frac{n_A}{n} + \frac{n_B}{n} = \mathcal{P}(A) + \mathcal{P}(B)$$

3rd axiom

This means that we can also use the frequentistic approach to assign probabilities! (the "Law of large numbers" [see later] guarantees that the limit exists)

What if the sample space is **uncountable**? (for instance, $\Omega \equiv \mathbb{R}$)

Problem: It is not possible to assign a probability to *all* the subsets of $\mathbb{R}$ without breaking compatibility with Kolmogorov's axioms

Solution: We assign a probability only to the *measurable* events $\mathcal{E}$ (open intervals $]x_1, x_2[$, closed intervals $[x_1, x_2]$, single points $x = x_1$ and any union/intersection of these). Such a set $\mathcal{E}$ is a Borel field

Remark: This way of proceeding is perfectly legitimate and, as a matter of fact, this restriction need not concern us, as all events of any practical interest are measurable