T2

Probability Theory

Salvador Ruiz Correa

August 12, 2025

PROBABILITY THEORY is the mathematical framework for quantifying uncertainty. It provides tools to model random events, reason about likelihoods, and make informed predictions based on incomplete or noisy data. In the context of artificial intelligence (AI) and machine learning (ML), probability theory is foundational—it underpins algorithms that learn from data, infer hidden patterns, and make decisions under uncertainty. From Bayesian networks and probabilistic graphical models to deep generative models and reinforcement learning, probabilistic reasoning enables systems to generalize beyond observed examples, handle ambiguity, and update beliefs as new information arrives. Without probability theory, AI would struggle to cope with the real-world variability and unpredictability that define most learning tasks.

AGENDA:

1. Probability theory and random phenomena.
2. Probability measure key features.
3. Probability theory and random phenomena.
4. Kolmogorov axioms overview.
5. Probability space: sample space, σ-algebras.

Probability Theory and Random Phenomena

PROBABILITY THEORY is a branch of mathematics that studies the behavior of *random phenomena* through a formal system of axioms. It is built upon and closely related to *measure theory*, a broader mathematical framework that generalizes the notions of length, area, and volume in $\mathbb{R}^n$ by assigning a measure to subsets of a given space.

Measure theory serves as the foundational framework for *integrating functions* over general spaces and plays a pivotal role in probability theory by enabling the rigorous definition of probability measures. Among the classical measures developed within this framework are the Jordan, Lebesgue, and Borel measures, which extend intuitive notions of length, area, and volume. More specialized constructs—such as complex measures, Haar measures (on locally compact groups), and probability measures—are designed to capture diverse properties of sets and functions across functional analysis, topology, and stochastic modeling.

Random Phenomena

Probability theory is useful for modeling *random phenomena* because it provides a mathematically rigorous way to describe uncertainty, quantify risk, and predict patterns in systems where outcomes are not deterministic.

Definition 1: Random Phenomenon

A phenomenon or procedure for generating data is *random* if
- the outcome is not predictable in advance;
- there is a predictable long-term pattern that can be described by the distribution of outcomes over very many observations.

Definition 2: Random Outcome

A *random outcome* is the result of a random phenomenon or procedure.

The outcome of an individual random experiment cannot be predicted with certainty, but the set of all *possible* outcomes is known in advance.

Box 1: Rolling a Dice

Imagine I roll a fair die privately, and I tell you if the resulting throw is odd or even.
- The possible outcomes are integers from one to six.
- The information available to you is whether the roll is odd or even.
- Probabilities are computed on the basis that each outcome is equally likely, so we have $\frac{1}{2}$ chance of obtaining odd/even.

```r
# Sample a random number in {1,2,3,4,5,6}
number = sample(1:6,1, replace=F)
if (number %% 2 > 0) {
    print("Odd number")
} else {
   print("Even number")
}
```

Figure 1: R code implementation of the random experiment described in Example 1. Outcomes have equal probability (i.e., the dice is fair).

Random Phenomena Examples

Outcomes in Health Sciences that are unpredictable can be modeled, in principle, as random phenomena. A principled modeling of these events can have a significant impact on patient care, research, and decision-making. The following are examples of random phenomena.

1. *Patient outcomes*: The response of individual patients to medical treatments or interventions can vary due to random factors. For instance, some patients may recover quickly from surgery while others with similar conditions may experience complications, and these differences can be influenced by factors that are not fully understood.
2. *Disease spread*: The spread of infectious diseases can exhibit random patterns. Factors such as the movement of infected individuals, contact patterns, and the effectiveness of preventive measures can all contribute to the randomness of disease transmission.
3. *Clinical trials*: Randomized controlled trials (RCTs) are widely used

in healthcare research to evaluate the efficacy of new treatments or interventions. The random assignment of participants to treatment and control groups helps reduce bias and accounts for random variations in outcomes.

4. *Emergency Room traffic*: The number of patients arriving at an emergency room can vary significantly from day to day or even hour to hour. Random factors like accidents, weather conditions, and disease outbreaks can influence the patient flow.
5. *Medication response*: Some patients may respond differently to the same medication due to genetic variations, environmental factors, or random fluctuations in their physiology. This can complicate medication management and dosing.
6. *Diagnostic Testing*: The results of diagnostic tests, such as blood tests or imaging scans, may show variability due to random measurement error, equipment calibration, or sample handling procedures.
7. *Disease Outbreaks*: The occurrence and severity of disease outbreaks, such as flu epidemics or COVID-19 surges, can exhibit random patterns influenced by factors like population density, vaccination rates, and individual behavior.
8. *Healthcare resource allocation*: The allocation of healthcare resources, such as hospital beds and ventilators during a public health crisis, can be influenced by random factors like the sudden surge in cases or unexpected logistical challenges.
9. *Healthcare costs*: The cost of healthcare services can vary randomly due to factors such as fluctuations in the prices of medical supplies, changes in insurance coverage, and unexpected healthcare demands.
10. *Healthcare workforce availability*: The availability of healthcare professionals, including doctors, nurses, and support staff, can be influenced by random factors like staff illnesses or scheduling conflicts.

Random Experiments

Probability theory is also useful for modeling *random experiments*.

Definition 3: Random Experiment

A *random experiment* is understood as any procedure such that when it is repeated under the same initial conditions, the result obtained is not always the same. It is characterized by three main features.

- The possible outcomes of the experiment.
- The events we can observe, i.e. the information that is revealed at the end of the experiment.
- The probabilities assigned to each event.

In practical applications, a random experiment can, in principle, be repeated numerous times under the same conditions. The outcomes of individual experiments must be independent, and must in no way be affected by any previous outcome.

Random Phenomena, Random Experiments, Measure Theory and Probability Theories

MEASURE THEORY is concerned with the problem of how to assign a size to certain sets, enabling a principle definition of a probability measure, which is the principal tool of probability theory. In daily life, assigning a size to sets is easy to do:

- count: $\{a, b, c, \ldots, x, y, x\}$ has 26 letters;
- take measurements: length (in one dimension), area (in two dimensions), volume (in three dimensions), or time;
- calculate the odds of winning the lottery.

In each case, we compare and express the result with respect to some base unit (Figure 4-1).

Figure 2: Henri Léon Lebesgue (French, 1875-1941) was a French mathematician known for his theory of integration, which was a generalization of the 17th-century concept of integration—summing the area between an axis and the curve of a function defined for that axis. His theory was published originally in his dissertation Intégrale, longueur, aire ("Integral, length, area") at the University of Nancy in 1902.

Probability Measure Key Features

Notice that triangles are more flexible than rectangles since we can represent every rectangle, and actually any odd-shaped quadrangle, as the 'sum' of two non-overlapping triangles Fig. 4-2. In doing so we have assumed a few things:

- In Fig. 4-3 we have chosen a *particular* baseline and the corresponding height arbitrarily. But the concept of *area* should not depend on such choice and the calculation this choice entails.
- The independence of the area from the way we calculate it is called well-definedness. Plainly, we have the choices shown in Fig 4-3. Notice that Fig 4-3 allows us to pick the most convenient method to work out the area.

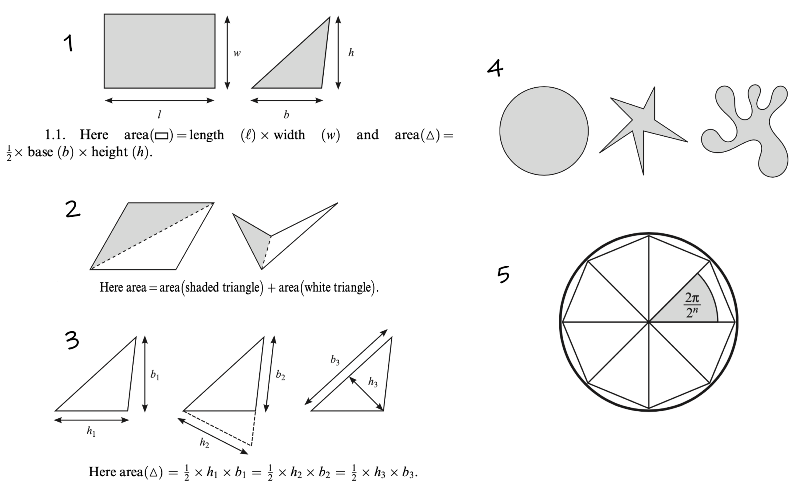

Figure 3: Measure theory is concerned with the problem of how to assign a size to certain sets. In daily life this is easy to do: we can count, take measurements, or calculate rates. In each case, we compare and express the result with respect to some base unit.

The area of the circle in 5 can be computed as follows.

$$\text{area}(\bigcirc) = lim_{n \longrightarrow \infty} 2^n \times \text{area}(\triangle \text{ at step } n).$$

Source: R. L. SchillingMeasures, Integrals, and Martingales, 2nd edition, Cambridge University Press, 2017.

- In Fig. 4-2 we actually use two facts:
  - the area of non-overlapping (disjoint) sets can be added, i.e.

$$\text{area}(A) = \alpha,\ \text{area}(B) = \beta\,, A \cap B = \emptyset \rightarrow, \text{area}(A \cup B) = \alpha + \beta.$$

This shows that the least we should expect from a reasonable measure μ is that it is

- well-defined, takes values in $[0, \infty]$, and $\mu(\emptyset) = 0$;
- additive, i.e., $\mu(A \cap B) = \mu(A) + \mu(B)$ whenever $A \cap B = \emptyset$;
- is invariant under congruences and translations, which is a characteristic property of length, area, and volume (the Lebesgue measure in $\mathbb{R}^n$).

Measures defined on the set of outcomes/events of a random phenomena/experiments enable a principled definition definition of probability.

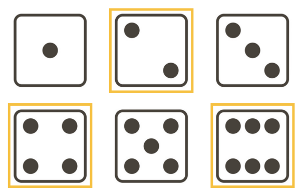

Figure 4: Measures over finite and countable sets have the properties described above. For instance, given a fair dice,

$$\mu(\{\text{⚁}, \text{⚃}, \text{⚅}\}) = \mu(\{\text{⚁}\}) + \mu(\{\text{⚃}\}) + \mu(\{\text{⚅}\}) = \frac{1}{2}.$$

Kolmogorov's Axioms

Kolmogorov's axioms are the foundations of probability theory. These axioms were introduced by Andrey Kolmogorov in 1933. These axioms remain central and have direct contributions to real-world probability cases.

These axioms are expressed using the fundamental concept of *probability space* $(\Omega, \mathscr{F}, P)$, where Ω is the sample space, $\mathscr{F}$ is a σ-algebra, and P is a probability measure. The elements of Ω are often called outcomes or *elementary events*, and the elements of $\mathscr{F}$, which are subsets of Ω, are called *events*.

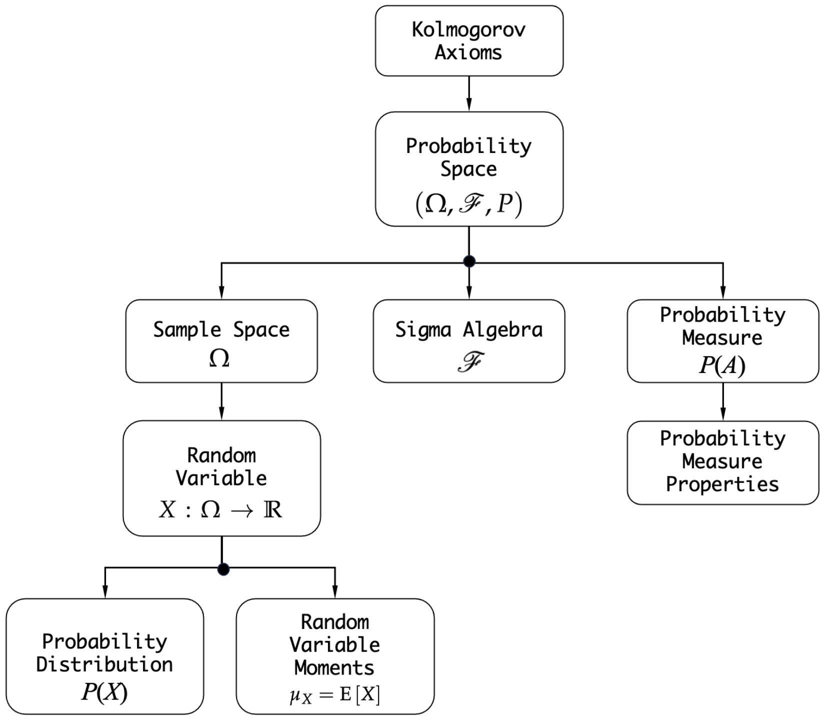

Figure 5: Kolmogorov axions principal concepts.

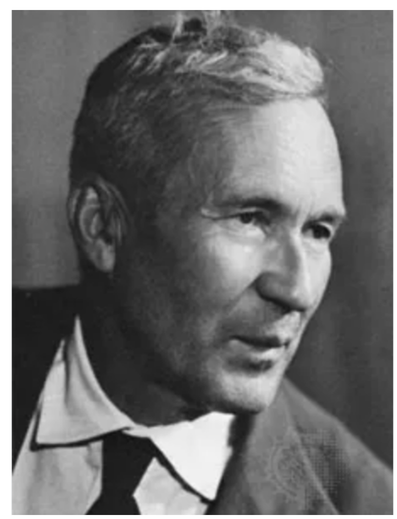

Figure 6: Andrey Nikolaevich Kolmogorov (Russian, 1903-1987) was a Soviet mathematician who contributed to the mathematics of probability theory, topology, intuitionistic logic, turbulence, classical mechanics, algorithmic information theory, and computational complexity. His contributions to probability theory provided a principled approach to the study of random phenomena.

Informally speaking a probability space is like a special playground where we play with chance and randomness. It's made up of three important parts:

- *Sample space:* This is like a list of all the possible things that can happen when we're dealing with something random. For example, if we're rolling a die, the sample space includes all the possible outcomes: 1, 2, 3, 4, 5, and 6.
- *Events*: These are like the games we play in our probability playground. Events are just groups of outcomes from the sample space. For example, the event of "getting an even number" when rolling a die includes outcomes 2, 4, and 6 (see Box 1).
- *Probability measure*: This is like a rule that tells us how likely each event is. It's like saying, "In our playground, the chance of this game happening is this much." For example, the probability of getting an even number when rolling a fair six-sided die is 1/2 because there are three even outcomes out of six possible outcomes.

Probability Space

THE PROBABILITY SPACE CONCEPT sets a principled frame to model random phenomena and random experiments. Here we provide an interpretation of each of the elements of a probability space: sample space, σ-algebra, and probability measure.

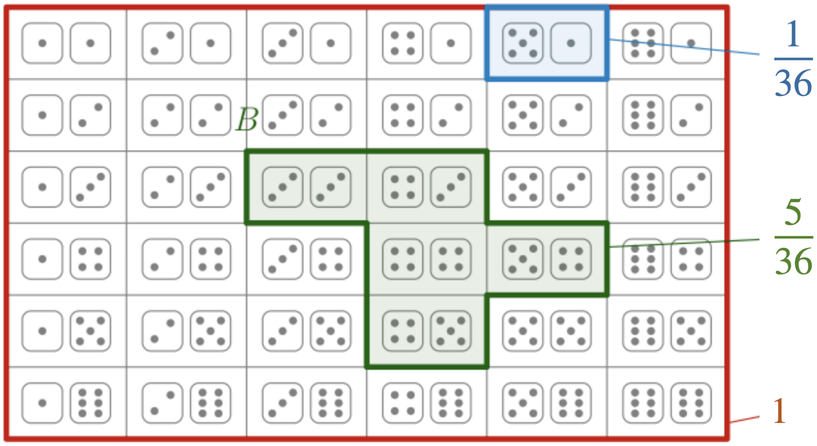

Figure 7: Probability space for throwing a die twice in succession: The sample space Ω consists of all 36 possible elementary outcomes; three different events (colored polygons) are shown with their respective probabilities (assuming that the dice are fair). Source: https://en.wikipedia.org

Definition 4: Probability Space

A probability space consists of an ordered triplet $(\Omega, \mathscr{F}, P)$, where Ω is an arbitrary *set* called *sample space*, $\mathscr{F}$ is a σ-algebra of subsets of Ω, and P is a probability measure defined on $\mathscr{F}$.

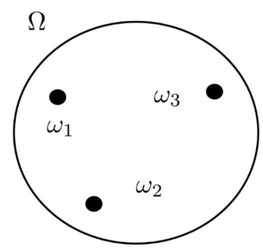

Figure 8: Sample space Ω containing three possible outcomes ω_1, ω_2, and ω_3.

Sample Space Ω

Loosely speaking, sample space in probability is like a list of all the things that could possibly happen when you're dealing with something random, like flipping a coin or rolling a die. Imagine you're flipping a coin. The sample space for this is just a list of the two things that can happen: "heads" and "tails." This is the sample space associated with flipping a coin. So, a sample space is like a simple list of all the different outcomes you might get when you do something random. It helps you see what could happen and helps you figure out the chances or probabilities of each outcome.

Definition 5: Sample Space

A *sample space* is a set Ω containing all possible outcomes of a random phenomenon or procedure. An *outcome* ω is an element in Ω, i.e., $\omega \in \Omega$ (which we may or may not observed).

- Rolling a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Flipping two coins: $\Omega = \{\mathsf{HH}, \mathsf{HT}, \mathsf{TH}, \mathsf{TT}\}$.

- Type in "random.rand()" on a Python shell: $\Omega = [0, 1]$.
- The temperature in my room from today to time T: $\Omega = \{f(t) : f(t) \geq -273.1^o\mathrm{C} \in [0, \mathrm{T}]\}$, where $f(t)$ is a continuous function.

Definition 6: Event

An *event* A is represented by a subset of Ω. After the realization of a random experiment we say that "A happens" if $\omega \in A$.

- Getting a die even roll: $A = \{2, 4, 6\}$.
- Getting the same outcome in two coin flips: $A = \{\mathrm{HH}, \mathrm{TT}\}$.
- "random.rand()" gives a number larger than 0.5: $A = (0.5, 1]$.
- The temperature in my room of at time T is above 20^oC : $A = \{T > 20^oC\}$.

Sample Space Types

Sample spaces can be classified according to their cardinality, i.e., the number of elements, #Ω, forming the set.

Definition 7: Types of Sample Spaces

- A sample space consisting of a finite or a *countably infinite* number of elements is called a *discrete sample space*.
- When the sample space includes all the numbers in an interval of the real line, it is called a *continuous sample space*.

The following are examples of sample spaces.

1. *Coin toss*: When you flip a fair coin, the sample space consists of two possible outcomes: heads (H) or tails (T). $\Omega = \{$ H, T$\}$.
2. *Rolling a six-sided die*: When you roll a standard six-sided die, the sample space includes the numbers 1 through 6. $\Omega = \{$ 1, 2, 3, 4, 5, 6$\}$.
3. *Flipping two coins*: If you flip two coins simultaneously, the sample space includes all possible combinations of outcomes for each coin. $\Omega = \{$ {H,H}, {H,T}, {T,H}, {T,T}$\}$.
4. *Rolling two dice*: Rolling two six-sided dice results in a sample space that includes all possible combinations of the dice values. $\Omega = \{$ {1,1}, {1,2}, ...{6,6}$\}$. (Figure 4).
5. *Rolling a six-sided die and flipping a coin*: If you roll a die and flip a coin, the combined sample space includes all possible pairs of outcomes.

$$\Omega = \{ \text{\{1, H\}, \{1, T\}, \{2, H\}, \{2, T\}, ..., \{6, H\}, \{6, T\} .} \}$$

6. *Diagnostic test*: In a diagnostic test for a disease, the sample space could consist of four outcomes $\omega_1 = \mathrm{TP}$, $\omega_2 = \mathrm{FP}$, $\omega_3 = \mathrm{TN}$ and

$\omega_4 = \text{FN}$, so $\Omega = \{\text{TP, FP, TN, FN}\}$. Key: True Positive, TP; False Positive, FP; True Negative, TN; and False Negative FN.

7. *Medication dosage*: Imagine a situation where a doctor is deciding on the dosage of a particular medication for a patient. The sample space would consist of all possible dosages that the doctor can prescribe, ranging from the lowest possible dose to the highest. $\Omega = \{\text{Low, Medium, High}\}$.
8. *Hospital stay length*: When a patient is admitted to a hospital, the length of their stay can vary. The sample space for the length of stay could include various possibilities, such as short stays, average stays, and long stays. $\Omega = \{\text{Short, Average, Long}\}$.
9. *Patient discharge status*: After receiving treatment, a patient may be discharged under different conditions. The sample space for discharge status might include being discharged in good health, with ongoing treatment needs, or with a referral to a specialist. $\Omega = \{\text{Discharged, Ongoing treatment, Referred}\}$.
10. *Surgical outcomes*: In the case of a surgical procedure, the sample space could include different possible outcomes, such as successful surgery, or unsuccessful surgery. $\Omega = \{\text{Successful, Unsuccessful}\}$.
11. *Disease progression*: For patients with chronic diseases, the progression of the disease can vary. The sample space for disease progression could include different stages of the disease, such as mild, moderate, or severe. $\Omega = \{\text{Mild, Moderate, Severe}\}$.
12. *Emergency room triage*: In an emergency room, patients are triaged based on the severity of their condition. The sample space for triage levels might include various levels of urgency, such as critical, urgent, or non-urgent. $\Omega = \{\text{Critical, Urgent, Non-urgent}\}$.
13. *Appointment scheduling*: When scheduling patient appointments, there can be various time slots available. The sample space for appointment scheduling would consist of all the possible time slots.

$$\Omega = \{\text{8:00 AM, 10:00 AM, 12:00 PM, 2:00 PM, 4:00 PM}\}.$$

14. *Psychological testing:* In psychological assessments, the sample space may consist of all possible test scores or responses. For example, when conducting an IQ test, the sample space includes all potential scores that individuals might achieve.
15. *Survey responses:* When conducting surveys or questionnaires in psychology, the sample space represents all possible responses to each question. For instance, in a survey about people's feelings, the sample space could include options like "happy," "neutral," "sad," and so on.
16. *Behavioral observations:* When observing and recording behaviors in psychological studies, the sample space could encompass various

behavioral categories. For example, in a study on child behavior, the sample space might include categories like "playing," "crying," "listening," and "talking."

17. *Emotional responses:* In studies of emotional responses, the sample space can describe the full range of possible emotions that individuals might experience, such as "joy," "anger," "fear," "disgust," "surprise," and "sadness."
18. *Blood pressure measurement*: When measuring blood pressure, both systolic and diastolic pressures are continuous variables, and their sample spaces consist of real numbers within specified ranges. For example, systolic pressure might fall within the range of 90 mmHg to 180 mmHg, while diastolic pressure might range from 60 mmHg to 120 mmHg. Therefore

$$\Omega = [90\,\mathsf{mmHg}, 180\,\mathsf{mmHg}],$$

and

$$\Omega = [60\,\mathsf{mmHg}, 120\,\mathsf{mmHg}],$$

for systolic and diastolic pressures, respectively.
19. *Weather forecast*: A weather forecast might have different categories like sunny, cloudy, rainy, or snowy. $\Omega = \{ \text{Sunny, Cloudy, Rainy, Snowy}\}$.
20. *Temperature measurement*: When measuring temperature using a thermometer, the sample space includes all possible real numbers within a specified temperature range. For instance, the temperature could be any real number within a range of -100^oC to 100^oC. $\Omega = [-100^oC, 100^oC]$.
21. *Inventory management:* The sample space could represent the different levels of inventory for a product. For example,

$$\Omega = \{ \text{High Inventory, Moderate Inventory, Low Inventory}\}.$$

22. *Market Research:* When conducting market research, the sample space may represent customer preferences. For instance,

$$\Omega = \{ \text{Product A Preferred, Product B Preferred, No Preference}\}.$$

23. *Financial Investments:* In the context of investment decisions, the sample space might represent investment outcomes like

$$\Omega = \{ \text{Profit, Break-even, Loss}\}.$$

24. *Project Management:* For project management, the sample space could depict project outcomes such as

$$\Omega = \{\text{On Time and On Budget, Delayed but On Budget, Delayed and Over Budget}\}.$$

25. *Product Launch Success:* When launching a new product, the sample space may represent the possible outcomes for success, like

$$\Omega = \{ \text{High Sales, Moderate Sales, Low Sales}\}.$$

26. *Employee Performance Evaluation:* In performance evaluations, the sample space might represent performance levels, such as

$$\Omega = \{\text{ High Inventory, Moderate Inventory, Low Inventory}\}.$$

27. *Customer Satisfaction:* Customer satisfaction surveys often use sample spaces like

$$\Omega = \{\text{Very Satisfied, Satisfied, Neutral, Dissatisfied, Very Dissatisfied}\}.$$

28. *Risk Assessment:* In risk assessment, the sample space may include potential risks or events, like

$$\Omega = \{\text{Market Downturn, Regulatory Changes, Supplier Issues}\}.$$

29. *Marketing Campaign Effectiveness:* For assessing the effectiveness of a marketing campaign, the sample space could represent the different customer responses, such as

$$\Omega = \{\text{Conversion, Click-Through, No Response}\}.$$

30. *Production Quality Control:* In quality control, the sample space might represent the quality of products, like

$$\Omega = \{\text{Defective, Passes Quality Control, Exceptional Quality}\}.$$

31. *Sample Space of Continuous Functions $C([a,b]$:* The sample space of continuous functions often denoted as $C([a,b])$, represents the set of all functions f(x) defined on the closed interval [a, b] that are continuous over that interval. In mathematical notation:

$$\Omega = \{f(x)|f(x)\text{is continuous on } [a,b]\}.$$

32. *The sample space of texts:* In the context of natural language and textual data, represents the set of all possible text strings that can be generated within a specific language or character encoding. The sample space of texts is essentially the universe of all potential textual documents, messages, or sequences of characters. Here are some key points to consider regarding the sample space of texts:
    - *Character set*: The sample space of texts depends on the character set used. For example, in English, it would be the set of all combinations of English letters (both uppercase and lowercase), digits, punctuation marks, and special characters.
    - *Infinite nature*: The sample space of texts is typically considered infinite, as there is no upper limit on the length of text strings that can be generated. Texts can be of varying lengths, from a single character to entire books or larger documents.

- *Variability*: The sample space encompasses a wide variety of texts, ranging from common words and phrases to unique combinations and rare or gibberish sequences of characters. It includes valid and meaningful texts as well as invalid, meaningless, or nonsensical ones.
- *Natural language*: The sample space is most relevant in the context of natural language, where it represents the potential for human communication in written form. It includes text in languages other than English, each with its own set of characters and rules.
- *Encoding and formatting*: The sample space is influenced by text encoding standards and formatting rules. For instance, plain text documents, HTML, JSON, XML, or any other text-based data format contribute to the diversity within the sample space.
- *Applications*: The sample space of texts is fundamental in natural language processing (NLP), text analysis, information retrieval, and various data science applications that involve textual data.
- *Size and complexity*: Due to the infinite nature of the sample space, it is practically impossible to exhaustively list or describe all possible text strings. The complexity increases with the size of the character set and the length of the text.
- *Machine earning*: In machine learning and NLP tasks, understanding the sample space of texts is important for tasks like text generation, text classification, and text mining. Models and algorithms are trained on data from this sample space to perform specific tasks.

The nature of the sample space of natural language can be appreciated by understanding how deep learning architectures for Large Language Models (LLMs), conduct natural language processing tasks. Models such as transformers, often use *embeddings* and *positional encoding* to convert text sequences into number sequences.

- *Word embeddings*: Transformers use word embeddings to convert input tokens (e.g., words) into continuous vector representations. These embeddings capture semantic information, allowing the model to understand the meaning of words. Word embeddings help transformers handle a large vocabulary efficiently and generalize better because they learn to represent words with similar meanings in similar vector spaces.
- *Positional encoding*: Transformers don't have a built-in notion of word order or position, which is essential for understanding sequences (e.g., sentences or documents). Positional encoding provides this information. Positional encoding is added to the word embeddings to convey the position of words in a sequence. It allows the model to distinguish between the same

word in different positions and learn the sequential relationships between words. In summary, embeddings and positional encoding in transformers play a critical role in enabling these models to handle sequences of data effectively, whether it's natural language text or other sequential data. Embeddings capture the meaning of words or tokens, while positional encoding imparts information about the order of tokens in the sequence. This combination enables transformers to excel in tasks like machine translation, text generation, and more.

Box 2: Application Example (Injection System Faliure?)

A spacecraft's motor receives fuel from an injection system, denoted by F. If this injection system fails, a master alarm (M) is triggered on the commander's main dashboard. One possible cause of failure in the injection system is the occurrence of extreme vibrations, represented by V, which the spacecraft may experience during operation. However, it is important to note that extreme vibrations can erroneously trigger the master alarm even if the injection system itself is functioning correctly.

- Write down the sample space for this example.
- Write down the outcomes related to the events:
- A = "The injection system failed."
- B = "The injection system failed and the master alarm is activated."
- C = "The injection system failed, the master alarm is activated, and the spacecraft experiences extreme vibrations."
- Solution.

1. Sample space and outcome probabilities:

Ω's oucomes	$P(\omega)$
$\omega_1 = \mathsf{V^cFA}$	α_1
$\omega_2 = \mathsf{V^cFA^c}$	α_2
$\omega_3 = \mathsf{V^cF^cA}$	α_3
$\omega_4 = \mathsf{V^cF^cA^c}$	α_4
$\omega_5 = \mathsf{VFA}$	α_5
$\omega_6 = \mathsf{VFA^c}$	α_6
$\omega_7 = \mathsf{VF^cA}$	α_7
$\omega_8 = \mathsf{VF^cA^c}$	α_8

2. Events:

A's outcomes	$P(\omega)$
$\omega_5 = \mathsf{VFA}$	α_5
$\omega_6 = \mathsf{VFA^c}$	α_6
$\omega_7 = \mathsf{VF^cA}$	α_7
$\omega_8 = \mathsf{VF^cA^c}$	α_8

B 's outcomes	$P(\omega)$
$\omega_5 = \mathsf{VFA}$	α_5
$\omega_6 = \mathsf{VFA^c}$	α_6

C's outcomes	$P(\omega)$
$\omega_5 = \mathsf{VFA}$	α_5

Box 3: Sample Space Exercise

- Mr. Holmes now lives in Los Angeles. One morning when Homes leaves his house, he realizes that his grass is wet (H). Is it due to rain (R), or has he forgotten to turn off the sprinkler (S)? Next, he notices that the grass of his neighbor, Dr. Watson, is also wet (W).
  1. Write down the sample space for this example.
  2. Write down the outcomes related to the events:
     - A = "Holmes grass is wet."
     - B = "Holmes forgot to turn the sprinkler off."
     - C = "Holmes forgot to turn the sprinkler off and Watson's grass is wet."
- Solution.
  1. Sample space and outcome probabilities:

Ω's oucomes	$P(\omega)$
$\omega_1 = \mathrm{H^cW^cS^cR^c}$	α_1
$\omega_2 = \mathrm{H^cW^cS^cR}$	α_2
$\omega_3 = \mathrm{H^cW^cSR^c}$	α_3
$\omega_4 = \mathrm{H^cW^cSR}$	α_4
$\omega_5 = \mathrm{H^cWS^cR^c}$	α_5
$\omega_6 = \mathrm{H^cWS^cR}$	α_6
$\omega_7 = \mathrm{H^cWSR^c}$	α_7
$\omega_8 = \mathrm{H^cWSR}$	α_8
$\omega_9 = \mathrm{HW^cS^cR^c}$	α_9
$\omega_{10} = \mathrm{HW^cS^cR}$	α_{10}
$\omega_{11} = \mathrm{HW^cSR^c}$	α_{11}
$\omega_{12} = \mathrm{HW^cSR}$	α_{12}
$\omega_{13} = \mathrm{HWS^cR^c}$	α_{13}
$\omega_{14} = \mathrm{HWS^cR}$	α_{14}
$\omega_{15} = \mathrm{HWSR^c}$	α_{15}
$\omega_{16} = \mathrm{HWSR}$	α_{16}

  2. Events:

A's outcomes	$P(\omega)$
$\omega_9 = \mathrm{HW^cS^cR^c}$	α_9
$\omega_{10} = \mathrm{HW^cS^cR}$	α_{10}
$\omega_{11} = \mathrm{HW^cSR^c}$	α_{11}
$\omega_{12} = \mathrm{HW^cSR}$	α_{12}
$\omega_{13} = \mathrm{HWS^cR^c}$	α_{13}
$\omega_{14} = \mathrm{HWS^cR}$	α_{14}
$\omega_{15} = \mathrm{HWSR^c}$	α_{15}
$\omega_{16} = \mathrm{HWSR}$	α_{16}

B 's outcomes	$P(\omega)$
$\omega_3 = \mathrm{H^cW^cSR^c}$	α_3
$\omega_4 = \mathrm{H^cW^cSR}$	α_4
$\omega_7 = \mathrm{H^cWSR^c}$	α_7
$\omega_8 = \mathrm{H^cWSR}$	α_8
$\omega_{11} = \mathrm{HW^cSR^c}$	α_{11}
$\omega_{12} = \mathrm{HW^cSR}$	α_{12}
$\omega_{15} = \mathrm{HWSR^c}$	α_{15}
$\omega_{16} = \mathrm{HWSR}$	α_{16}

C's outcomes	$P(\omega)$
$\omega_7 = \mathrm{H^cWSR^c}$	α_7
$\omega_8 = \mathrm{H^cWSR}$	α_8
$\omega_{15} = \mathrm{HWSR^c}$	α_{15}
$\omega_{16} = \mathrm{HWSR}$	α_{16}

Box 4: Application Example (Was it the burglar?)

Mary lives in San Francisco City. One afternoon, she is driving back home and receives a phone call from her neighbor Jane. She told her that her house alarm was set off (A). While driving, she also heard on the radio (R) that a small earthquake (E) hit the city. Small earthquakes sometimes activate the alarm, and perhaps this is the reason why the alarm was sounding.

- Write down the sample space for this example.
- Write down the outcomes related to the events:
- A = "A burglar (B) broke into the house."
- B = "A burglar broke into the house and the alarm was set off."
- C = "A burglar broke into the house and the alarm was set off and a small earthquake hit the city."

σ-algebras

To gain some intuition of what a σ-algebra is, imagine you have a collection of things, like a number of different events or sets of outcomes in probability. A sigma algebra is like a special container or group that holds these things together in an organized way (Figure 19).

Definition 8: σ-algebra

A σ-algebra $\mathscr{F}$ on a set Ω is a family of subsets of Ω such that:

- $\Omega \in \mathscr{F}$
- $A \in \mathscr{F} \implies A^c \in \mathscr{F}$
- $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F} \implies \bigcup A_n \in \mathscr{F}$

The set $A \in \mathscr{F}$ is said to be *measurable* or *$\mathscr{F}$-measurable*.

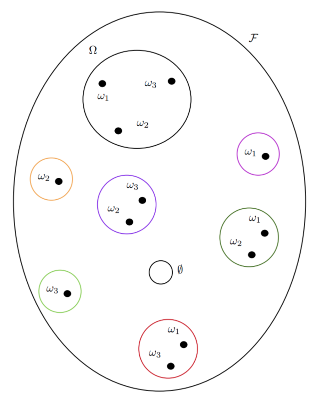

Figure 9: The event space is a σ-algebra on the sample space Ω. It is a subset of the power set, $\mathcal{P}(\Omega)$, whose elements, called events, satisfy some regularity conditions. When is finite and discrete, the power set is a valid σ-algebra that can be used as the event space. The σ-algebra satisfies $\Omega \in \mathcal{F}$; the sample space is an event called the *sure* event. $A \in F \implies A^c \in \mathcal{F}$, i.e., $\mathcal{F}$ is *closed* under complementation. $A_1, A_2, \ldots, \bigcup A_n \in \mathscr{F}$, so F is closed under *countable unions*. Source: http://bit.ly/3s11bq5.

To be a σ-algebra, this container has to follow a few rules:

1. It must always contain the "whole thing." In other words, it includes everything you're interested in. For example, if you're thinking about all possible outcomes when rolling a die, the σ-algebra would definitely include "rolling a 1," "rolling a 2," and so on, all the way up to "rolling a 6."
2. It should also include the "opposites" of things. Let's say you toss a coin and have "getting a head" as one of your events. The sigma σ-algebra should also have "not getting a head" (which means getting a tail) inside it.
3. When you look inside the σ-algebra, you should find all the combinations and possibilities of the things you're interested in. So, if you have "rolling an even number" and "rolling a prime number,"

you want the sigma-algebra to include things like "rolling an even number AND rolling a prime number."

In summary, a sigma-algebra is like a special container that holds all the possible outcomes and their opposites in an organized way, helping you study and understand different events and probabilities in a systematic manner. The following are examples of σ-algebras.

1. Let X be a set.
2. $\{\emptyset, X\}$ is a σ-algebra (the minimal σ-algebra in X).
3. $\mathscr{P}(X)$, the power set of X (the maximal σ-algebra in X).
4. $\{\emptyset, X, B, B^c\}$, $B \subset X$ is a σ-algebra.
5. $X = \{\mathsf{HH}, \mathsf{HT}, \mathsf{TH}, \mathsf{TT}\} = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.

$$\begin{aligned}
\mathscr{A}_{\text{min}} &= \{\emptyset, \{\omega_1, \omega_2, \omega_3, \omega_4\}\}. \\
\mathscr{A}_{\text{max}} &= \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}.\{\omega_1, \omega_4\}, \\
&\quad \{\omega_2, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}, \\
&\quad \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_3, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}, X\}. \\
\mathscr{A}_1 &= \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\}^c, X\}, \{\omega_1, \omega_2\}^c = \{\omega_3, \omega_4\}.
\end{aligned}$$

6. Is $\omega_1 \neq \{\omega_1\}$?.

Exercises

1. Construct a σ-algebra for the set $\Omega = \{⚀, ⚁\}$.

$$\mathscr{F} = \{\Omega, \emptyset, \{⚀\}, \{⚁\}\}$$

2. Consider the set $\Omega = \{⚀, ⚁, ⚂\}$.
   - Construct the *minimal* σ-algebra for the set Ω.

   $$\mathscr{F} = \{\Omega, \emptyset\}.$$

   - Construct the *maximal* σ-algebra for the set Ω (i.e., Ω's power set, $\mathscr{P}(\Omega)$).

   $$\mathscr{F} = \{\emptyset, \{⚀\}, \{⚁\}, \{⚂\}, \{⚀, ⚁\}, \{⚀, ⚂\}, \{⚁, ⚂\}, \Omega\}.$$

   - Construct a σ-algebra for the set Ω that is neither maximal or minimal.

   $$\mathscr{F} = \{\Omega, \emptyset, B, B^c\}, B \subset \Omega$$

   - *Remark:* no set is an element of itself. For example,

   $$⚀ \neq \{⚀\}.$$

3. *Emotional responses:* In studies of emotional responses, the sample space can describe the full range of possible emotions that individuals might experience, such as "joy," "anger," "fear," "disgust,"

"surprise," and "sadness." Write down the sample space Ω associated with the given outcomes. Let

$$\mathscr{F} = \{\Omega, \emptyset, B, B^c\},$$

where $B = \{\text{joy}\}$. Is $\mathscr{F}$ is a σ-algebra?

4. *Surgical outcomes*: In the case of a surgical procedure, the sample space could include different possible outcomes, such as successful surgery, complications, or the need for additional surgeries.

$$\Omega = \{\text{Successful surgery, Complications, Additional surgery}\}$$

Let $\mathscr{F} = \{\Omega, \emptyset, A, A^c\}$ where $A = \{\text{Successful surgery}\}$. Show that $\mathscr{F}$ is a σ-algebra?

5. *Product sales:* When launching a new product or analyzing sales data, the sample space represents all possible sales outcomes, including the range of sales volumes and revenue generated.

$$\Omega = \{\text{Low sales, Moderate sales, High sales}\}$$

Construct a σ-algebra for Ω.

6. *Emergency room triage*: In an emergency room, patients are triaged based on the severity of their condition. The sample space for triage levels might include various levels of urgency, such as critical, urgent, or non-urgent.

$$\Omega = \{\text{Critical, Urgent, Non-urgent}\}.$$

What is the maximal σ-algebra for Ω?

Box 5: Intuitions behind a σ-algebra

- Represents the information that will be revealed to us after the realization of a random outcome.
- Contains all events that can be verified if they happened or not after ω has been realized.

- Roll a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$\mathscr{F}_1 = \{\emptyset, \Omega, \{1, 2, 3\}, \{4, 5, 6\}\}.$$
$$\mathscr{F}_{\max} = \mathscr{P}(\Omega) = \text{the set of all subsets of } \Omega.$$

are both σ-algebras. $\mathscr{F}_1$ contains information on whether the roll is strictly less than four or not, and $\mathscr{F}_2$ contains information on the exact outcome. (Why?)

- $\Omega = \{\mathsf{HH}, \mathsf{HT}, \mathsf{TH}, \mathsf{TT}\}$.

$$\mathscr{F}_1 = \{\emptyset, \Omega, \{\mathsf{HH}, \mathsf{HT}\}, \{\mathsf{TH}, \mathsf{TT}\}\}.$$

- $\mathscr{F}_1$ is a σ-algebra containing information on the outcome of the first flip. (Why?)

*σ-algebra Generators**

Given a set of events or outcomes, a σ-algebra generator refers to the smallest σ-algebra that contains these events. In other words, it's the collection of all possible events that can be formed by combining the original set of events through set operations like unions, intersections, and complements.

Definition 9: σ-algebra Generators*

σ-algebra generators are defined as follows:

- For every system of sets $\mathscr{G} \in \mathscr{P}(\Omega)$ there exist a smallest σ-algebra containing $\mathscr{G}$. This is the σ-algebra generated by $\mathscr{G}$, denoted by $\sigma(\mathscr{G})$, and $\mathscr{G}$ is called the generator.
- Notice that the intersection $\bigcap_{i \in I} \mathscr{A}_i$ of arbitrarily many σ-algebras in Ω is again a σ-algebra in Ω.
- Except in a few simple examples, it is hard to write down explicitly a generated σ-algebra.

Box 3 shows an example of a σ-algebra generator and the corresponding σ-algebra.

Box 6: σ-algebra Generators Example*

- Let $A, B \in \Omega$, $A \bigcap B = \emptyset$. Define $\mathscr{G} = \{A, B\}$.

$$\sigma(\mathscr{G}) = \{\emptyset, A, B, A \cup B, A^c, B^c, (A \cup B)^c, \Omega\}.$$

Prove this statement.

In an experiment of rolling a die, suppose we are interested in knowing whether the outcome belongs to a low-range $(1, 2)$, mid-range $(3, 4)$, or high-range $(5, 6)$. What is the minimal information required? In this example, the required minimal information is "encoded" by the σ-algebra generated by $\mathscr{G} = \{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$:

$$\sigma(\mathscr{G}) = \{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \{3, 4, 5, 6\}, \{1, 2, 5, 6\}, \{1, 2, 3, 4\}, \Omega\}.$$

If we are interested in the exact outcome of the die, then take $\mathscr{G} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$ and $\sigma(\mathscr{G}) = \mathscr{P}(\Omega)$.

Box 7: Application Example (Injection System Faliure?)

A spacecraft's motor receives fuel from an injection system, denoted by F. If this injection system fails, a master alarm (M) is triggered on the commander's main dashboard. One possible cause of failure in the injection system is the occurrence of extreme vibrations, represented by V, which the spacecraft may experience during operation. However, it is important to note that extreme vibrations can erroneously trigger the master alarm even if the injection system itself is functioning correctly.

- Write down the sample space for this example.
- Write down the sigma algebra generated by event A and provide an intuitive interpretation.
- A = "The injection system failed."
- Solution.

1. Sample space and outcome probabilities:

Ω's oucomes	$P(\omega)$
$\omega_1 = \mathrm{V^cFA}$	α_1
$\omega_2 = \mathrm{V^cFA^c}$	α_2
$\omega_3 = \mathrm{V^cF^cA}$	α_3
$\omega_4 = \mathrm{V^cF^cA^c}$	α_4
$\omega_5 = \mathrm{VFA}$	α_5
$\omega_6 = \mathrm{VFA^c}$	α_6
$\omega_7 = \mathrm{VF^cA}$	α_7
$\omega_8 = \mathrm{VF^cA^c}$	α_8

2. Events:

A's outcomes	$P(\omega)$
$\omega_5 = \mathrm{VFA}$	α_5
$\omega_6 = \mathrm{VFA^c}$	α_6
$\omega_7 = \mathrm{VF^cA}$	α_7
$\omega_8 = \mathrm{VF^cA^c}$	α_8

Box 8: Sample Space Exercise

- Mr. Holmes now lives in Los Angeles. One morning when Homes leaves his house, he realizes that his grass is wet (H). Is it due to rain (R), or has he forgotten to turn off the sprinkler (S)? Next, he notices that the grass of his neighbor, Dr. Watson, is also wet (W).
  1. Write down the sample space for this example.
  2. Write down the sigma algebra generated by event C and provide an intuitive interpretation.
     - C = "Holmes forgot to turn the sprinkler off and Watson's grass is wet."
- Solution.
  1. Sample space and outcome probabilities:

Ω's oucomes	$P(\omega)$
$\omega_1 = \mathsf{H^cW^cS^cR^c}$	α_1
$\omega_2 = \mathsf{H^cW^cS^cR}$	α_2
$\omega_3 = \mathsf{H^cW^cSR^c}$	α_3
$\omega_4 = \mathsf{H^cW^cSR}$	α_4
$\omega_5 = \mathsf{H^cWS^cR^c}$	α_5
$\omega_6 = \mathsf{H^cWS^cR}$	α_6
$\omega_7 = \mathsf{H^cWSR^c}$	α_7
$\omega_8 = \mathsf{H^cWSR}$	α_8
$\omega_9 = \mathsf{HW^cS^cR^c}$	α_9
$\omega_{10} = \mathsf{HW^cS^cR}$	α_{10}
$\omega_{11} = \mathsf{HW^cSR^c}$	α_{11}
$\omega_{12} = \mathsf{HW^cSR}$	α_{12}
$\omega_{13} = \mathsf{HWS^cR^c}$	α_{13}
$\omega_{14} = \mathsf{HWS^cR}$	α_{14}
$\omega_{15} = \mathsf{HWSR^c}$	α_{15}
$\omega_{16} = \mathsf{HWSR}$	α_{16}

  2. Event:

C's outcomes	$P(\omega)$
$\omega_7 = \mathsf{H^cWSR^c}$	α_7
$\omega_8 = \mathsf{H^cWSR}$	α_8
$\omega_{15} = \mathsf{HWSR^c}$	α_{15}
$\omega_{16} = \mathsf{HWSR}$	α_{16}

Constructing a σ-Algebra from Overlapping Generator Sets

Let the sample space be all 3-bit binary strings:

$$\Omega = \{000, 001, 010, 011, 100, 101, 110, 111\}$$

Define the generator $\mathcal{G} = \{G_1, G_2, G_3\}$ as follows:

- $G_1 = \{x \in \Omega : \text{first bit is } 1\}$
- $G_2 = \{x \in \Omega : \text{number of 1s is even}\}$
- $G_3 = \{x \in \Omega : \text{last bit is } 1\}$

The membership table is:

x	$x \in G_1$	$x \in G_2$	$x \in G_3$	Membership Vector
000	0	1	0	$(0,1,0)$
001	0	0	1	$(0,0,1)$
010	0	0	0	$(0,0,0)$
011	0	1	1	$(0,1,1)$
100	1	0	0	$(1,0,0)$
101	1	1	1	$(1,1,1)$
110	1	1	0	$(1,1,0)$
111	1	0	1	$(1,0,1)$

Each unique membership vector defines an atom:

$$
\begin{aligned}
A_{(0,1,0)} &= \{000\} \\
A_{(0,0,1)} &= \{001\} \\
A_{(0,0,0)} &= \{010\} \\
A_{(0,1,1)} &= \{011\} \\
A_{(1,0,0)} &= \{100\} \\
A_{(1,1,1)} &= \{101\} \\
A_{(1,1,0)} &= \{110\} \\
A_{(1,0,1)} &= \{111\}
\end{aligned}
$$

These atoms form a partition of Ω, and the σ-algebra generated by $\mathcal{G}$ is:

$$\sigma(\mathcal{G}) = \left\{ \bigcup_{i \in I} A_i : I \subseteq \{1, \ldots, 8\} \right\}$$

Excercise

Let the sample space be the set of all 4-bit binary strings: $\Omega = \{0000, 0001, 0010, \ldots, 1111\}$. That is,

$$\Omega = \{x \in \{0,1\}^4\},$$

with $|\Omega| = 16$. Define the generator:

$$\mathcal{G} = \{A_1, A_2\}, \quad \text{where}$$

- $A_1 = \{x \in \Omega : x \text{ has even parity}\}$
- $A_2 = \{x \in \Omega : x \text{ starts with } 1\}$
- Construct the the smallest σ-algebra on Ω that contains both A_1 and A_2.
- Explicitly list the elements of A_1 and A_2.
- Determine all atoms (minimal non-empty elements) of $\mathcal{F}$.
- List all subsets in $\mathcal{F}$.
- Compare $\mathcal{F}$ with the power set $\mathcal{P}(\Omega)$.
- Discuss how this σ-algebra could be used in a probabilistic model.

Solution:

Let the sample space be:

$$\Omega = \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111\}$$

Define the generator sets:

- $A_1 = \{x \in \Omega : x \text{ has even parity}\}$ (i.e., an even number of 1s)
- $A_2 = \{x \in \Omega : x \text{ starts with } 1\}$
- Membership table:

x	$x \in A_1$	$x \in A_2$	Membership Vector
0000	1	0	$(1, 0)$
0001	0	0	$(0, 0)$
0010	0	0	$(0, 0)$
0011	1	0	$(1, 0)$
0100	0	0	$(0, 0)$
0101	1	0	$(1, 0)$
0110	1	0	$(1, 0)$
0111	0	0	$(0, 0)$
1000	0	1	$(0, 1)$
1001	1	1	$(1, 1)$
1010	1	1	$(1, 1)$
1011	0	1	$(0, 1)$
1100	1	1	$(1, 1)$
1101	0	1	$(0, 1)$
1110	0	1	$(0, 1)$
1111	1	1	$(1, 1)$

We group elements by membership vector:

$$\begin{aligned}
A_{(1,0)} &= \{0000, 0011, 0101, 0110\} \\
A_{(0,0)} &= \{0001, 0010, 0100, 0111\} \\
A_{(0,1)} &= \{1000, 1011, 1101, 1110\} \\
A_{(1,1)} &= \{1001, 1010, 1100, 1111\}
\end{aligned}$$

These atoms form a partition of Ω, and the σ-algebra generated by $\mathcal{G} = \{A_1, A_2\}$ is:

$$\sigma(\mathcal{G}) = \left\{ \bigcup_{i \in I} A_i : I \subseteq \{1, 2, 3, 4\} \right\}$$

The power set $\mathcal{P}(\Omega)$ contains $2^{16} = 65{,}536$ subsets. Our σ-algebra contains only 16 subsets.

Probabilistic interpretation This σ-algebra partitions Ω based on two binary features:

- Parity (even vs. odd)
- First bit (0 vs. 1)

This structure is useful in models where outcomes are grouped by these features, such as:

- Binary classification tasks

- Feature-based probability assignments
- Simplified Markov models or decision trees
  Now, define the generator $\mathcal{G} = \{G_1, G_2\}$:
- $G_1 = \{x \in \Omega : \text{first bit is } 1\}$
- $G_2 = \{x \in \Omega : \text{number of 1s is even}\}$
  The membership table is:

x	$x \in G_1$	$x \in G_2$	Membership Vector
000	0	1	$(0, 1)$
001	0	0	$(0, 0)$
010	0	0	$(0, 0)$
011	0	1	$(0, 1)$
100	1	0	$(1, 0)$
101	1	1	$(1, 1)$
110	1	1	$(1, 1)$
111	1	0	$(1, 0)$

Group elements by membership vector to get the atoms $\sigma(\mathcal{G})$:

$$\begin{aligned} A_{(0,0)} &= \{001, 010\} \\ A_{(0,1)} &= \{000, 011\} \\ A_{(1,0)} &= \{100, 111\} \\ A_{(1,1)} &= \{101, 110\} \end{aligned}$$

These atoms form a partition of Ω, and the σ-algebra generated by $\mathcal{G}$ is:

$$\sigma(\mathcal{G}) = \left\{ \bigcup_{i \in I} A_i : I \subseteq \{1, 2, 3, 4\} \right\}$$

Definition 10: *Filtration*

Let $\mathscr{F}$ be a σ-algebra. An increasing sequence

$$\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}_n \subset \cdots \subset \mathscr{F}$$

is a filtration of sub-σ-algebras of $\mathscr{F}$. The sequence is indexed by non-negative natural numbers. Other indexing sets can be used for defining a filtration (v.gr. $\mathbb{R}$).

A filtration essentially is a mathematical model that represents partial knowledge about an outcome. The filtrations tell us whether an event happened or not. One may envision the 'filtration process' as a sequence of filters, each filter providing us a more detailed view of the events in Ω.

Example 1 of a Probability Filtration

As a second example, we introduce a problem concerning the price of given financial stock S_n at time n. The initial stock price is defined by S_0 and we have a time horizon composed of three discrete time steps ($n = 1, 2, 3$).

- Possible outcomes at time $n = 0$

$$X = \{uuu, uud, udu, udd, duu, dud, ddu, ddd\},$$

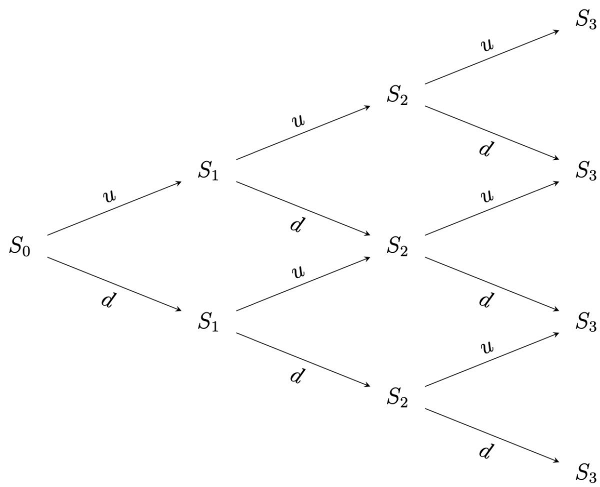

Figure 10: Stock prices.

- The corresponding sigma-algebra generated by these outcomes are:

$$\mathscr{F}_0 = \mathscr{G}(X) = \{\emptyset, X\},$$

  At $n = 0$, all paths are possible. Thus, the event set $A = \{uuu, uud, udu, udd, ddd, ddu, dud, duu\}$ – with the sequences describing the movement per time step – contains all possible paths; that is $X = A$.
- Possible outcomes at time $n = 1$

$$\{A_{u\cdot\cdot}, A_{d\cdot\cdot}\} = \{\{uuu, uud, udu, udd\}, \{duu, dud, ddu, ddd\}\},$$

  The corresponding sigma-algebra generated by these outcomes are:

$$\begin{aligned}\mathscr{F}_1 &= \mathscr{G}(\{A_{u\cdot\cdot}, A_{d\cdot\cdot}\}) = \{\emptyset, A_{u\cdot\cdot}, A_{d\cdot\cdot}, A_{u\cdot\cdot} \cup A_{d\cdot\cdot}\} = \{\emptyset, A_{u\cdot\cdot}, A_{d\cdot\cdot}, \Omega\} \\ &= \{\emptyset, A_{u\cdot\cdot}, A^c_{u\cdot\cdot}, X\}\end{aligned}$$

- At $n = 1$, we know that the stock price went either up or down. The corresponding events can be defined by $A_{u\cdot\cdot} = \{uuu, uud, udu, udd\}$ and $A_{d\cdot\cdot} = \{ddd, ddu, dud, duu\}$.
- Notice: if the price went up, we know our sample-path will be in $A_{u\cdot\cdot}$ and not in $A_{d\cdot\cdot}$.
- At $n = 2$, we have four event sets: $A_{uu\cdot} = \{uuu, uud\}$, $A_{ud\cdot} = \{udu, udd\}$, $A_{du\cdot} = \{duu, dud\}$, $A_{dd\cdot} = \{ddu, ddd\}$. Observe that the information is getting increasingly fine-grained.
- At $n = 3$, we obviously know the exact price path that has been followed.

Box 9: *FiltrationExample*

- The filtration is as follows: $\mathscr{F}_0 = \mathscr{G}(X) = \{\emptyset, X\}$,
- Possible outcomes at time $n = 1$

$$\{A_{u\cdot\cdot}, A_{d\cdot\cdot}\} = \{\{uuu, uud, udu, udd\}, \{duu, dud, ddu, ddd\}\},$$

- The corresponding sigma-algebra generated by these outcomes are:

$$\begin{aligned}\mathscr{F}_1 &= \mathscr{G}(\{A_{u\cdot\cdot}, A_{d\cdot\cdot}\}) \\ &= \{\emptyset, A_{u\cdot\cdot}, A_{d\cdot\cdot}, A_{u\cdot\cdot} \cup A_{d\cdot\cdot}\} \\ &= \{\emptyset, A_{u\cdot\cdot}, A_{d\cdot\cdot}, \Omega\} \\ &= \{\emptyset, A_{u\cdot\cdot}, A^c_{u\cdot\cdot}, X\}\end{aligned}$$

- Possible outcomes at time $n = 2$

$$\begin{aligned}\{A_{uu\cdot}, A_{ud\cdot}, A_{du\cdot}, A_{dd\cdot}\} = \{&\{uuu, uud\}, \{udu, udd\}, \\ &\{duu, dud\}, \{ddu, ddd\}\}.\end{aligned}$$

- The corresponding sigma-algebra generated by these outcomes are:

$$\begin{aligned}\mathscr{F}_2 &= \mathscr{G}(\{A_{uu\cdot}, A_{ud\cdot}, A_{du\cdot}, A_{dd\cdot}\}) \\ &= \{\emptyset, A_{u\cdot\cdot}, A_{d\cdot\cdot}, A_{uu\cdot}, A_{ud\cdot}, A_{du\cdot}, A_{dd\cdot}, A^c_{uu\cdot}, A^c_{ud\cdot}, A^c_{du\cdot}, A^c_{dd}, \\ &A_{uu\cdot} \cup A_{du\cdot}, A_{uu\cdot} \cup A_{dd\cdot}, A_{ud\cdot} \cup A_{du\cdot}, A_{ud\cdot} \cup A_{dd\cdot}, X\}.\end{aligned}$$

- Let's compute $\mathscr{G}(\{A_{uu\cdot}, A_{ud\cdot}, A_{du\cdot}, A_{dd\cdot}\})$.

$$\begin{aligned}A_{uu\cdot} &= \{uuu, uud\}, \\ A_{ud\cdot} &= \{udu, udd\}, \\ A_{du\cdot} &= \{duu, dud\}, \\ A_{dd\cdot} &= \{ddu, ddd\}, \\ A^c_{uu\cdot} &= A_{ud\cdot} \cup A_{du\cdot} \cup A_{dd\cdot}, \\ A^c_{ud\cdot} &= A_{uu\cdot} \cup A_{du\cdot} \cup A_{dd\cdot}, \\ A^c_{du\cdot} &= A_{uu\cdot} \cup A_{ud\cdot} \cup A_{dd\cdot}, \\ A^v_{dd\cdot} &= A_{uu\cdot} \cup A_{ud\cdot} \cup A_{du\cdot},\end{aligned}$$

$$\begin{aligned}A_{u\cdot\cdot} &= A_{uu\cdot} \cup A_{ud\cdot}, \\ A_{\cdot u\cdot} &= A_{uu\cdot} \cup A_{du\cdot}, \\ &A_{uu\cdot} \cup A_{dd\cdot}, \\ &A_{ud\cdot} \cup A_{du\cdot}, \\ A_{\cdot d\cdot} &= A_{ud\cdot} \cup A_{dd\cdot}, \\ A_{d\cdot\cdot} &= A_{du\cdot} \cup A_{dd\cdot}, \\ \Omega &= A_{u\cdot\cdot} \cup A_{d\cdot\cdot}.\end{aligned}$$

- In summary:

$$\begin{aligned}\mathscr{F}_0 &= \{\emptyset, X\}, \\ \mathscr{F}_1 &= \{\emptyset, A_{u\cdot\cdot}, A^c_{u\cdot\cdot}, X\}, \\ \mathscr{F}_2 &= \mathscr{G}(\{A_{uu\cdot}, A_{ud\cdot}, A_{du\cdot}, A_{dd\cdot}\}) \\ &= \{\emptyset, A_{u\cdot\cdot}, A_{d\cdot\cdot}, A_{uu\cdot}, A_{ud\cdot}, A_{du\cdot}, A_{dd\cdot}, A^c_{uu\cdot}, A^c_{ud\cdot}, A^c_{du\cdot}, A^c_{dd}, \\ &A_{uu\cdot} \cup A_{du\cdot}, A_{uu\cdot} \cup A_{dd\cdot}, A_{ud\cdot} \cup A_{du\cdot}, A_{ud\cdot} \cup A_{dd\cdot}, \Omega\}. \\ \mathscr{F}_3 &= \mathscr{G}(X) = \mathscr{P}(X)\end{aligned}$$

$$\boxed{\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \mathscr{F}_3}$$

Example 1 of a Probability Filtration

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a discrete-time stochastic process $(X_t)_{t=0}^3$ representing three coin tosses, where:

$$X_t = \begin{cases} 1 & \text{if the } t\text{-th toss is heads} \\ 0 & \text{if the } t\text{-th toss is tails} \end{cases}$$

Define a filtration $(\mathcal{F}_t)_{t=0}^3$ such that $\mathcal{F}_t$ represents the information available up to time t:

- $\mathcal{F}_0 = \{\emptyset, \Omega\}$ (no information yet)
- $\mathcal{F}_1 = \sigma(X_1)$ (information from the first toss)
- $\mathcal{F}_2 = \sigma(X_1, X_2)$ (information from the first two tosses)
- $\mathcal{F}_3 = \sigma(X_1, X_2, X_3)$ (full information from all tosses)

This filtration satisfies:

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \mathcal{F}$$

Each $\mathcal{F}_t$ is a σ-algebra that grows with time, modeling the accumulation of information. Provide the details of this example.