Applied Probability for Computer Science

Isadora Antoniano-Villalobos
isadora.antoniano@unive.it

Master in Computer Science
Ca' Foscari University of Venice

Academic year 2023/2024

Overview of Elementary Probability

Sample space, outcomes and events

- Baron textbook **Chapter 1** is a suggested reading if you are interested in a panoramic view of probability and statistics as tools for Computer Sciences
- Baron textbook **Chapter 2: Probability** is part of the course prerequisites, but we will start with a quick overview to remember the main concepts

DEFINITION 2.1

A collection of all elementary results, or **outcomes** of an experiment, is called a **sample space**.

DEFINITION 2.2

Any set of outcomes is an **event**. Thus, events are subsets of the sample space.

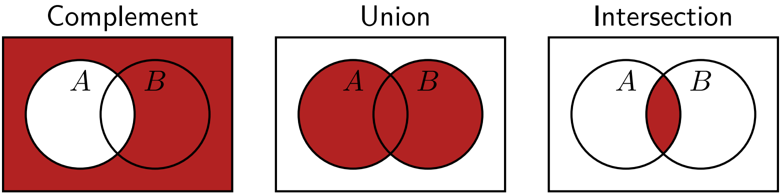
Sample space, outcomes and events

Common Notation:

Ω **or** S	for the sample space
$\emptyset$	for the empty set/event. *Note:* $\emptyset \subset \Omega$ for any Ω
A, B, E	and other capital letters for events *Note:* $E \subset \Omega$
ω **or** s	for individual outcomes *Note:* $\omega \in \Omega$ is an outcome; $\{\omega\} \subset \Omega$ is an event
$\mathbb{P}[E] = \mathbf{P}[E]$	for the probability of an event

- **Example 2.4** A tossed die can produce one of 6 possible outcomes: 1 dot, through 6 dots. Each outcome is an event. There are other events: observing an even number of dots, an odd number of dots, a number of dots less than 3, etc.

Sets and set operations



$\bar{A} = A^{\mathsf{c}}$ $\qquad$ $A \cup B$ $\qquad$ $A \cap B$

☞ $A \setminus B = A \cap \bar{B}$ defines the set difference

☞ De Morgan's laws establish the relation between union, intersection and complementation:

$$\overline{E_1 \cup \ldots \cup E_n} = \bar{E}_1 \cap \ldots \cap \bar{E}_n; \quad \overline{E_1 \cap \ldots \cap E_n} = \bar{E}_1 \cup \ldots \cup \bar{E}_n$$

Sets and set operations

DEFINITION 2.7

Events A and B are **disjoint** if their intersection is empty,

$$A \cap B = \varnothing.$$

Events $A_1, A_2, A_3, \ldots$ are **mutually exclusive** or **pairwise disjoint** if any two of these events are disjoint, i.e.,

$$A_i \cap A_j = \varnothing \quad \text{for any} \quad i \neq j\,.$$

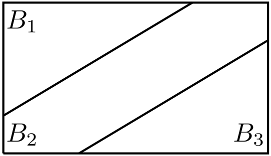
DEFINITION 2.8

Events $A, B, C, \ldots$ are **exhaustive** if their union equals the whole sample space, i.e.,

$$A \cup B \cup C \cup \ldots = \Omega.$$

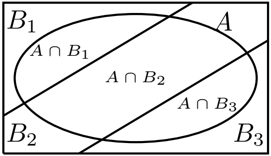
Sets and set operations

- A collection of mutually exclusive, exhaustive events is called a **partition** of the sample space



- Any event $A \subset \Omega$ can be written of the union of its intersections the elements of a partition:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3)$$



Axiomatic Probability

DEFINITION 2.9

A collection $\mathfrak{M}$ of events is a **sigma-algebra** on sample space Ω if

(a) it includes the sample space,

$$\Omega \in \mathfrak{M}$$

(b) every event in $\mathfrak{M}$ is contained along with its complement; that is,

$$E \in \mathfrak{M} \quad \Rightarrow \quad \overline{E} \in \mathfrak{M}$$

(c) every finite or countable collection of events in $\mathfrak{M}$ is contained along with its union; that is,

$$E_1, E_2, \ldots \in \mathfrak{M} \quad \Rightarrow \quad E_1 \cup E_2 \cup \ldots \in \mathfrak{M}.$$

Axiomatic Probability

- $\mathfrak{M} = \{\emptyset, \Omega\}$ is the smallest possible sigma-algebra, called **degenerate**
- $\mathfrak{M} = 2^{\Omega} = \{E : E \subset \Omega\}$ is the largest possible sigma-algebra, called the **power set**

  Note: When $\Omega \subseteq \mathbb{N}$ is **countable**, we most commonly use $\mathfrak{M} = 2^{\Omega}$, but when $\Omega \subseteq \mathbb{R}$ is **uncountable** we use the **Borel sigma-algebra** $\mathfrak{M} = \mathfrak{B}$, much smaller than the power set (see Example 2.12 of the *Baron*).

Axiomatic Probability

DEFINITION 2.10

Assume a sample space Ω and a sigma-algebra of events $\mathfrak{M}$ on it. **Probability**

$$\boldsymbol{P} \ : \ \mathfrak{M} \to [0,1]$$

is a function of events with the domain $\mathfrak{M}$ and the range $[0,1]$ that satisfies the following two conditions,

(Unit measure) The sample space has unit probability, $\boldsymbol{P}(\Omega) = 1$.

(Sigma-additivity) For any finite or countable collection of *mutually exclusive* events $E_1, E_2, \ldots \in \mathfrak{M}$,

$$\boldsymbol{P}\left\{E_1 \cup E_2 \cup \ldots\right\} = \boldsymbol{P}(E_1) + \boldsymbol{P}(E_2) + \ldots.$$

☞ All the rules of probability are consequences from this definition, which allows the calculation of probabilities for all events of interest.

Axiomatic Probability

- $\mathbb{P}[\emptyset] = 0$
- $\mathbb{P}\left[\bar{A}\right] = 1 - \mathbb{P}[A]$
- $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$ for any events A and B
  Note: The extension of this result to the union of n events gives rise to the **Inclusion-exclusion formula**
- $\mathbb{P}[A \cup B] \leq \mathbb{P}[A] + \mathbb{P}[B]$ for any events A and B
  Note: The extension of this result to the union of n events gives rise to the **Bonferroni inequalities**

Conditional Probability and Independence

DEFINITION 2.11

Events $E_1, \ldots, E_n$ are **independent** if they occur independently of each other, i.e., occurrence of one event does not affect the probabilities of others.

- In other words, events $\{E_i\}_{i=1}^n$ are independent if and only if

$$\mathbb{P}\left[\bigcap_{i=1}^n E_i\right] = \prod_{i=1}^n \mathbb{P}\left[E_i\right]$$

- Formally,

$$\mathbb{P}\left[A|B\right] = \frac{\mathbb{P}\left[A \cap B\right]}{\mathbb{P}\left[B\right]}$$

It follows that, in general

$$\mathbb{P}\left[A \cap B\right] = \mathbb{P}\left[B\right]\mathbb{P}\left[A|B\right]$$

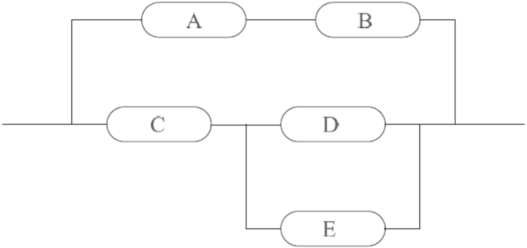
DEFINITION 2.16

Events A and B are **independent** if occurrence of B does not affect the probability of A, i.e.,

$$\boldsymbol{P}\{A \mid B\} = \boldsymbol{P}\{A\}.$$

- Formally, A and B are independent if and only if $\mathbb{P}[A|B] = \mathbb{P}[A]$. Which, in turn, happens if and only if $\mathbb{P}[A \cap B] = \mathbb{P}[B]\,\mathbb{P}[A]$
- **Example 2.20:** Calculate the reliability of this system if each component is operable with probability 0.92 independently of the other components

Conditional Probability and Independence



DEFINITION 2.15

Conditional probability of event A given event B is the probability that A occurs when B is *known to occur*.

- **Example 2.35 (Diagnostics of computer codes).** A new computer program consists of two modules. The first module contains an error with probability 0.2. The second module is more complex; it has a probability of 0.4 to contain an error, independently of the first module. An error in the first module alone causes the program to crash with probability 0.5. For the second module, this probability is 0.8. If there are errors in both modules, the program crashes with probability 0.9. Suppose the program crashed.What is the probability of errors in both modules?

- **Bayes Rule** is a direct consequence of the definition of conditional probability:

$$\mathbb{P}\left[B|A\right] = \frac{\mathbb{P}\left[AB\right]}{\mathbb{P}\left[A\right]}$$

The numerator can be substitued by:

$$\mathbb{P}\left[AB\right] = \mathbb{P}\left[A|B\right]\mathbb{P}\left[B\right]$$

Conditional Probability and Bayes Theorem

- And the denominator by the expression obtained from the **Law of Total Probability**:
  If $B_1, \ldots, B_k$ is a partition of Ω, then:

Law of Total Probability	$\boldsymbol{P}\{A\} = \sum_{j=1}^{k} \boldsymbol{P}\{A \mid B_j\} \boldsymbol{P}\{B_j\}$ In case of two events ($k = 2$), $\boldsymbol{P}\{A\} = \boldsymbol{P}\{A \mid B\} \boldsymbol{P}\{B\} + \boldsymbol{P}\{A \mid \overline{B}\} \boldsymbol{P}\{\overline{B}\}$

- In general, we obtain **Bayes Rule**: for every $i = 1, \ldots, k$,

$$\mathbb{P}[B_i|A] = \frac{\mathbb{P}[A|B_i]\,\mathbb{P}[B_i]}{\sum_{j=1}^{k} \mathbb{P}[A|B_j]\,\mathbb{P}[B_j]}$$

- In particular, since B and $\bar{B}$ always define a partition of Ω, it follows that:

Bayes Rule for two events

$$\boxed{\boldsymbol{P}\{B \mid A\} = \frac{\boldsymbol{P}\{A \mid B\}\boldsymbol{P}\{B\}}{\boldsymbol{P}\{A \mid B\}\boldsymbol{P}\{B\} + \boldsymbol{P}\{A \mid \overline{B}\}\boldsymbol{P}\{\overline{B}\}}}$$

- Have a look at **Example 2.34**!