Mathematical foundation of information theory

Martina Tscheckl

March 1, 2016

Please send feedback to martina@tscheckl.eu.

Contents

1 Introduction: Probability Theory

01.03.2016

1.1 Definitions and Axioms

Definition (Probability space)**.** *A* probability space *is a triple* $(\Omega, \mathcal{A}, \mathbb{P})$ *where*

1. $\Omega \neq \emptyset$ *is a* set*,*
2. $\mathcal{A}$ *is a* σ-algebra *of subsets of* Ω*,*
3. $\mathbb{P}$ *is a* probability measure *on* $\mathcal{A}$*.*

Remarks. *(a) The set* Ω *is called the "sample space" and contains "all possible aspects/cases related to an experiment".*

(b) By definition, $\mathcal{A} \subset \mathcal{P}(\Omega)$ *where* $\mathcal{P}(\Omega)$ *is the collection of all subsets of* Ω*. The set* $\mathcal{P}(\Omega)$ *is called the power set of* Ω*.*

(c) The elements $A \in \mathcal{A}$ *are called "events". Events are observable (distinguishable) outcomes of an experiment.*

Examples. *(a) one coin toss:* $\Omega = \{0, 1\}$

(b) 500 *coin tosses:* $\Omega = \{0, 1\}^{500}$

(c) one coin toss + final position of coin (+ number of spins):

$$\Omega = \{0, 1\} \times \mathbb{R}^2 (\times \mathbb{N}_0)$$

Definition (Axioms)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space with* $\mathbb{P} : \mathcal{A} \to [0, 1]$*. Then we define the following axioms:*

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \implies A^C \in \mathcal{A}$
3. *If* $(A_n)_{n \in \mathbb{N}}$ *is a sequence in* $\mathcal{A}$*, then*
$$\bigcup_{i=1}^{\infty} A_n \in \mathcal{A}.$$
4. $\mathbb{P}(\Omega) = 1$
5. σ*-additivity: If* (A_n) *is a sequence in* $\mathcal{A}$ *such that* $A_k \cap A_l = \emptyset \, \forall k \neq l$ *then*
$$\mathbb{P}(\biguplus_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

Examples. *Let's consider the examples (b) and (c) from above.*

(b) 500 *coin tosses:*

$$\begin{aligned} A &= [\textit{even number of } 1s] \\ &= \{\omega = (\omega_1, \dots, \omega_{500}) \in \Omega : \omega_1 + \dots + \omega_{500} \textit{ is even}\}. \end{aligned}$$

(c) *one coin toss + final position of coin + number of spins:*

$$\begin{aligned} A &= [\textit{coin toss: } 1, \textit{ positon from origin: distance } < 2m, \textit{ spins: } 13] \\ &= \{1\} \times \mathcal{B}((0,0), 2) \times \{13\} \end{aligned}$$

where $\mathcal{B}((0,0),2) = \{(x,y) \in \mathbb{R}^2 : \sqrt{x^2+y^2} < 2\}$.

Remark.

fix this remark

1. *If* Ω *is at most countable: usually* $\mathcal{A} = \mathcal{P}(\Omega)$
2. *If* Ω *is uncountable, it is not good.*

Example. *Let* $\Omega = \mathbb{R}$*. Then* $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$*, where* $\mathcal{B}_{\mathbb{R}}$ *is the collection of all Borel sets on* $\mathbb{R}$*. So* $\mathcal{A}$ *is the smallest* σ*-algebra over* $\mathbb{R}$ *which contains all intervals.*

Proposition (Consequences of the axioms)**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space with* $\mathbb{P} : \mathcal{A} \to [0,1]$*. Then we can deduce from the axioms the following statements:*

(a) *If* $A_1, \dots, A_N \in \mathcal{A}$*, then* $\bigcup_{n=1}^{N} A_n \in \mathcal{A}$ *and since* $\mathbb{P}(\emptyset) = 0$,

$$\mathbb{P}(\biguplus_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mathbb{P}(A_n).$$

(b) *If* $(A_n)_{n\in\mathbb{N}}$ *is a finite or countable sequence in* $\mathcal{A}$*, then*

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}.$$

(c) *Let* $A, B \in \mathcal{A}$*. Then*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

(d) *Let* $A, B \in \mathcal{A}$*. If* $A \subset B$*, then*

$$\mathbb{P}(A) \leq \mathbb{P}(B).$$

(e) *"Continuity": If* $(A_n)_{n\in\mathbb{N}}$ *is a sequence in* $\mathcal{A}$ *such that* $\forall n \; A_n \subset A_{n+1}$*, then*

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Proof.

fix/complete proof

(c) Use $A \cup B = A \uplus (B \cap A^C)$.

add pic

(e)

add pic

$$
\begin{aligned}
A_N &= A_1 \cup (A_2 \setminus A_1) \cup \cdots \cup (A_N \setminus A_{N-1}) \\
&= A_1 \cup \bigcup_{n=2}^{N} (A_n \setminus A_{n-1}) \\
A &= A_1 \uplus \biguplus_{n=2}^{\infty} \underbrace{(A_n \setminus A_{n-1})}_{\in \mathcal{A}}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A_1) + \sum_{n=2}^{\infty} \mathbb{P}(An \setminus A_{n-1}) \\
&= \lim_{N\to\infty} (\mathbb{P}(A_1) + \sum_{n=2}^{N} \mathbb{P}(A_n \setminus \mathbb{P}(A_{n-1})) \\
&= \lim_{N\to\infty} \mathbb{P}(A_N)
\end{aligned}
$$

$\square$

02.03.2016