MP-208
Optimal Filtering with Aerospace Applications
Section 2.4: Probability Theory

Prof. Dr. Davi Antônio dos Santos
Instituto Tecnológico de Aeronáutica
www.professordavisantos.com

São José dos Campos - SP
2023

Contents

Probability Space…

Probability Space

Definition:

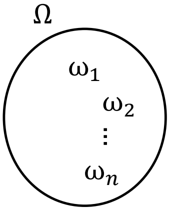
A probability space is an abstract description of a probabilistic experiment. It is usually denoted by the triplet $(\Omega, \mathcal{F}, P)$, in which:

- Ω is the sampling space of the experiment
- $\mathcal{F}$ is the σ-algebra of the experiment
- P is the probability measure

Probability Space

Sampling Space Ω:

It is the set constituted of all the possible results (called outcomes) of a certain random experiment.



Examples:

- Experiment with a coin: $\Omega = \{\text{head}, \text{tail}\}$.
- Experiment with a die: $\Omega = \{1, 2, \ldots, 6\}$.

Probability Space

σ-algebra $\mathcal{F}$:

It is a class (collection of sets) constituted of all viable subsets of Ω. These subsets are called events. $\mathcal{F}$ is such that:

- $\Omega \in \mathcal{F}$
- If $A \in \mathcal{F}$, then $\bar{A} \in \mathcal{F}$
- If $A_i \in \mathcal{F}, i = 1, 2, \ldots$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

Example:

In an experiment with a die where we want to know if an outcome is odd or even, a suitable σ-algebra is:

$$\mathcal{F} = \left\{\emptyset, \Omega, \{1, 3, 5\}, \{2, 4, 6\}\right\}$$

Probability Space

Probability Measure P:

It is a function $P : \mathcal{F} \to [0,1] \subset \mathbb{R}$ that satisfies the following three conditions:

1. $P(A) \geq 0, \ \forall A \in \mathcal{F}$
2. $P(\Omega) = 1$
3. If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

The conditions 1–3 are the so-called Axioms of Probability.

Remark:

Axiom 3 can be extented to infinitely many mutually exclusive events $A_1, A_2, A_3, \ldots$:

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

Conditional Probability. . .

Conditional Probability

Definition:

The conditional probability of an event A given another event M with non-zero probability is denoted by $P(A|M)$ and defined by:

$$P(A|M) \triangleq \frac{P(A \cap M)}{P(M)}$$

Properties:

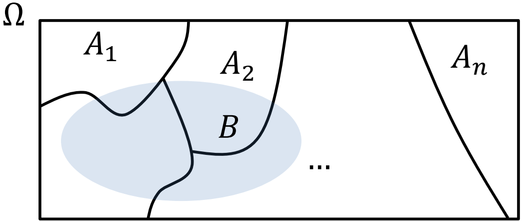
- The measure $P(\cdot|M)$ satisfies the Axioms of Probability
- If $M \subset A$, then $P(A|M) = 1$
- If $A \subset M$, then $P(A|M) \geq P(A)$

Important Identities. . .

Important Identities

Total Probability Theorem:

Consider $\mathcal{P} = [A_1, A_2, \ldots, A_n]$ as a partition of Ω and B as an arbitrary event.



The Total Probability Theorem says that:

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)$$

Important Identities

Bayes' Rule:

Consider two arbitrary events $A, B \in \mathcal{F}$. Assume that $P(A) \neq 0$ and $P(B) \neq 0$. The Bayes' rule (or theorem) says that:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Remark:

Bayes' rule is useful in statistical inference by the so-called Bayesian approach. In this application of the above result, event A is a hypothesis, event B is an evidence, $P(A|B)$ is the *a posteriori* probability (of the hypothesis, given the evidence) and $P(B|A)$ is the likelihood, which is the probability of the evidence B, assuming that the hypothesis A is true.

Important Identities

Chain Rule:

Consider the arbitrary events $A_1, ..., A_n \in \mathcal{F}$. The chain rule says that:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right) P\left(A_{n-1} | \bigcap_{i=1}^{n-2} A_i\right) \cdots P\left(A_2 | A_1\right) P\left(A_1\right)$$

Independence of Events. . .

Independence of Events

Independence of Two Events:

Two events $A_1, A_2 \in \mathcal{F}$ are said to be (statistically) independent if

$$P(A_1 \cap A_2) = P(A_1)P(A_2)$$

Independence of Three Events:

Three events $A_1, A_2, A_3 \in \mathcal{F}$ are said to be (statistically) independent if

$$P(A_i \cap A_j) = P(A_i)P(A_j), \quad i \neq j$$

and

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$

Independence of Events

Independence of n Events:

The events $A_1, A_2, ..., A_n \in \mathcal{F}$ are said to be (statistically) independent if any $k < n$ of them are independent and

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n)$$

As the independence of $k = 3$ events has been defined (in the previous slide), one can conclude that the above definition is complete.

References. . .

References

- Papoulis, A.; Pillai, S. U. **Probability, Random Variables, and Stochastic Processes**. New York: McGraw-Hill, 2002.
- Ross, S. M. **A First Course in Probability**. New York: Prentice Hall, 2002.