भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Lecture 10: Independence of Events, Borel–Cantelli Lemma, Conditional Probability, Law of Total Probability, Bayes' Theorem

Karthik P. N.
Assistant Professor, Department of AI
Email: pnkarthik@ai.iith.ac.in

02 September 2025

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Recap: Construction of Lebesgue Measure

Probability Assignment for Uncountable Sample Spaces

- $\Omega = (0, 1)$
- As before, suppose we start by assigning probabilities to all **singleton subsets**
- More specifically, let

$$\mathbb{P}(\emptyset) = 0, \qquad \mathbb{P}(\Omega) = 1, \qquad \mathbb{P}(\{\omega\}) = p_\omega, \quad \omega \in \Omega.$$

- **What is $\mathbb{P}((0, \frac{1}{2}))$?**
  This cannot be derived from the probabilities of singleton subsets!

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

An Important Result from Measure Theory

Theorem

Suppose Ω is an uncountable set, and $\mathscr{F} = 2^{\Omega}$.
If $\mathbb{P}$ is a valid probability measure on $\mathscr{F}$ (satisfying the three axioms of probability), then there exists a **countable subset** $S \subseteq \Omega$ such that $\mathbb{P}(S) = 1$.
Furthermore, for any $A \in \mathscr{F}$, we have

$$\mathbb{P}(A) = \sum_{\omega \in A \cap S} \mathbb{P}(\{\omega\}).$$

Takeaway

When Ω is uncountable, the only interesting probability measures on 2^{Ω} are discrete measures!

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Example 1: Lebesgue Measure on $\Omega = (0,1)$

- Let $(\Omega, \mathscr{F}) = ((0,1), \mathscr{B}(0,1))$
- Consider the collection

$$\mathscr{S} = \Big\{(a,b] : \ 0 \le a \le b \le 1\Big\}.$$

  Observe that:
  - $\emptyset \in \mathscr{S}$
  - $\mathscr{S}$ is **closed under finite intersections**
  - For any $A, B \in \mathscr{S}$, the set $A \setminus B$ may be expressed as

$$A \setminus B = \bigsqcup_{i=1}^{n} C_i,$$

    for some disjoint sets $C_1, \ldots, C_n \in \mathscr{S}$
- The collection $\mathscr{S}$ is called a **semiring**

Example 1: Lebesgue Measure on $\Omega = (0, 1)$

- Consider the collection
$$\mathscr{S} = \Big\{(a, b] : \;\; 0 \le a \le b \le 1\Big\}.$$
- Let $m : \mathscr{S} \to [0, 1]$ be an assignment satisfying the following properties:
  - $m(\emptyset) = 0$
  - $m(\Omega) = 1$
  - $m((a, b]) = b - a$
  - **Finite additivity**

Caratheodory's Extension Theorem

There exists a unique extension of m to the whole of $\mathscr{B}(0, 1)$.
The extended measure is called the **Lebesgue measure** on $\mathscr{B}(0, 1)$, denoted by λ.
In particular,
$$\lambda(A) = m(A) \qquad \forall\, A \in \mathscr{S}.$$

Example 2: Lebesgue Measure on $\Omega = \mathbb{R}$

- Consider the collection

$$\mathscr{S} = \Big\{(a,b] : \ -\infty \leq a \leq b < +\infty\Big\}.$$

- Let $m : \mathscr{S} \to [0, \ +\infty]$ be an assignment satisfying the following properties:
  - $m(\emptyset) = 0$
  - $m(\Omega) = +\infty$
  - $m((a,b]) = b - a$
  - **Finite additivity**

Caratheodory's Extension Theorem

There exists a unique extension of m to the whole of $\mathscr{B}(\mathbb{R})$.
The extended measure is called the **Lebesgue measure** on $\mathscr{B}(\mathbb{R})$, denoted by λ.
In particular,

$$\lambda(A) = m(A) \qquad \forall\, A \in \mathscr{S}.$$

Properties of Lebesgue Measure on $\mathscr{B}(\mathbb{R})$

Consider the measure space $(\Omega, \mathscr{F}, \mu) = (\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$.

- $\lambda(\{x\}) = 0$ for all $x \in \mathbb{R}$
- $\lambda(a, b) = \lambda((a, b]) = \lambda([a, b)) = \lambda([a, b]) = b - a$
- $\lambda(\mathbb{Q}) = 0$
- **Exercise:** $\lambda(K) = 0$, where K denotes the Cantor set

Independence of Events

Independence of Events

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

Definition (Independence of Events)

Events $A, B \in \mathscr{F}$ are said to be independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

We write $A \perp\!\!\!\perp B$ as a shorthand notation to denote that A and B are independent.

Some Tidbits on Independence

- Suppose $A \in \mathscr{F}$ is such that $\mathbb{P}(A) = 0$
  - $A \perp\!\!\!\perp A$
  - $A \perp\!\!\!\perp B$ for all $B \in \mathscr{F}$
- Suppose $A \in \mathscr{F}$ is such that $\mathbb{P}(A) = 1$
  - $A \perp\!\!\!\perp A$
  - $A \perp\!\!\!\perp B$ for all $B \in \mathscr{F}$
- If $A \perp\!\!\!\perp B$, then:
  - $A^{\complement} \perp\!\!\!\perp B$
  - $A \perp\!\!\!\perp B^{\complement}$
  - $A^{\complement} \perp\!\!\!\perp B^{\complement}$
- **Can an event be independent of itself? Yes!**

Independence of Multiple Events

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

Definition (Independence of Events)

- Events $A_1, A_2, \ldots, A_n \in \mathscr{F}$ are said to be independent if for all $\mathcal{I}_0 \subseteq \{1, 2, \ldots, n\}$,

$$\mathbb{P}\left(\bigcap_{i \in \mathcal{I}_0} A_i\right) = \prod_{i \in \mathcal{I}_0} \mathbb{P}(A_i).$$

- Let $\mathcal{I}$ be an arbitrary index set. A collection of events $\{A_i : i \in \mathcal{I}\}$ is independent if for every finite subset $\mathcal{I}_0 \subseteq \mathcal{I}$, the collection of events $\{A_i : i \in \mathcal{I}_0\}$ is independent.

Borel–Cantelli Lemma

Borel–Cantelli Lemma

Lemma (Borel–Cantelli Lemma)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

1. Suppose $A_1, A_2, \ldots \in \mathscr{F}$ are such that $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < +\infty$. Then,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0. \qquad \left(\limsup_{n\to\infty} A_n = A_{\text{limsup}} = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k\right)$$

2. Suppose $A_1, A_2, \ldots \in \mathscr{F}$ are **independent** and satisfy $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$. Then,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1.$$

Proof of Borel–Cantelli Lemma, Part 1

- Suppose $A_1, A_2, \ldots \in \mathscr{F}$ are such that $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < +\infty$
- We then have

$$
\begin{aligned}
\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k\right) &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k\right) && \text{(continuity of probability)} \\
&\leq \lim_{n\to\infty} \sum_{k\geq n} \mathbb{P}(A_k) && \text{(union bound)} \\
&= 0
\end{aligned}
$$

- We thus proved that $\mathbb{P}(A_{\mathrm{limsup}}) = 0$

Indian Institute of Technology Hyderabad

Proof of Borel–Cantelli Lemma, Part 2

- Suppose $A_1, A_2, \ldots \in \mathscr{F}$ are **independent** and satisfy $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$
- For each $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{k\geq n} A_k^{\complement}\right) &= \prod_{k\geq n} \mathbb{P}(A_k^{\complement}) \qquad \text{(independence of } A_n^{\complement}, A_{n+1}^{\complement}, \ldots) \\
&= \prod_{k\geq n} (1 - \mathbb{P}(A_k)) \leq \prod_{k\geq n} \exp\left(-\mathbb{P}(A_k)\right) \qquad (1 - x \leq \exp(-x)\ \forall\, x \geq 0) \\
&= \exp\left(-\sum_{k\geq n} \mathbb{P}(A_k)\right) = 0.
\end{aligned}$$

- Taking limits as $n \to \infty$ on either sides, we get

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n^{\complement}\right) = 0 \iff \mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Independence of σ-Algebras

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

Definition (Independence of σ-Algebras)

Let $\mathscr{F}_1, \mathscr{F}_2 \subseteq \mathscr{F}$ be sub-σ-algebras of $\mathscr{F}$. Then, $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \qquad \forall A \in \mathscr{F}_1,\ B \in \mathscr{F}_2.$$

More generally, for an arbitrary index set $\mathcal{I}$, the collection sub-σ-algebras $\{\mathscr{F}_i : i \in \mathcal{I}\}$ are said to be independent if for all choices of $A_i \in \mathscr{F}_i, i \in \mathcal{I}$, the events $\{A_i : i \in \mathcal{I}\}$ are independent.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Conditional Probabilities

Conditional Probability Measure

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Definition (Conditional Probability)

Given $B \in \mathscr{F}$ such that $\mathbb{P}(B) > 0$, define

$$\mathbb{P}_B : \mathscr{F} \to [0, 1] \qquad \text{via} \qquad \mathbb{P}_B(A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad A \in \mathscr{F}.$$

Then, $\mathbb{P}_B$ is a valid probability measure on $(\Omega, \mathscr{F})$, and is called the **conditional probability measure** conditioned on the event B.

Notation: $\mathbb{P}_B(A)$ is denoted more commonly as $\mathbb{P}(A|B)$.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

$\mathbb{P}_B$ is a Valid Probability Measure on $(\Omega, \mathscr{F})$

- $\mathbb{P}_B(\emptyset) = 0$

- $\mathbb{P}_B(\Omega) = 1$

- For any mutually disjoint collection of sets $A_1, A_2, \ldots \in \mathscr{F}$,

$$\mathbb{P}_B\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}_B(A_n).$$

Conditional Probability – Properties

- Fix $B \in \mathscr{F}$ such that $0 < \mathbb{P}(B) < 1$. Then, for any $A \in \mathscr{F}$,

$$\mathbb{P}(A) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^{\mathsf{c}}) \cdot \mathbb{P}(B^{\mathsf{c}}).$$

Conditional Probability – Properties

- **(Law of Total Probability)**

  Suppose $B_1, B_2, \ldots \in \mathscr{F}$ form a **partition** of Ω, i.e.,

$$B_i \cap B_j = \emptyset \quad \forall i \neq j, \qquad \bigsqcup_{n \in \mathbb{N}} B_n = \Omega.$$

  Then, for any $A \in \mathscr{F}$,

$$\mathbb{P}(A) = \sum_{n:\, \mathbb{P}(B_n) > 0} \mathbb{P}(A | B_n) \cdot \mathbb{P}(B_n).$$

Conditional Probability – Properties

- **(Bayes' Theorem)**

  Suppose $B_1, B_2, \ldots \in \mathscr{F}$ form a **partition** of Ω, i.e.,

$$B_i \cap B_j = \emptyset \quad \forall i \neq j, \qquad \bigsqcup_{n \in \mathbb{N}} B_n = \Omega.$$

  For any $A \in \mathscr{F}$ such that $\mathbb{P}(A) > 0$,

$$\mathbb{P}(B_n|A) = \begin{cases} \dfrac{\mathbb{P}(A|B_n) \cdot \mathbb{P}(B_n)}{\sum\limits_{j:\, \mathbb{P}(B_j)>0} \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}, & \mathbb{P}(B_n) > 0, \\ 0, & \mathbb{P}(B_n) = 0. \end{cases}$$

Conditional Probability – Properties

- **(Chain Rule)**

  Let $A_1, A_2, \ldots \in \mathscr{F}$. Then,

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \mathbb{P}(A_1)\cdot\mathbb{P}(A_2|A_1)\cdot\mathbb{P}(A_3|A_1\cap A_2)\cdots$$

$$= \mathbb{P}(A_1)\cdot\prod_{n\geq 2}\mathbb{P}\left(A_m \,\middle|\, \bigcap_{j=1}^{n-1} A_j\right),$$

provided each of the conditional probabilities on the right-hand side is defined.