भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Mathematical Foundations for Data Science (Probability)

Probability Measures on Discrete Sample Spaces, Conditional Probability, Bayes’ Theorem, Independence of Events, Random Variables

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

17 August 2024

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability Measures on Discrete Spaces

Discrete Sample Spaces

Let Ω be a non-empty, discrete (finite or countably infinite) sample space. Then, Ω may be represented as

$$\Omega = \{\omega_1, \omega_2, \ldots\}$$

In this case, we simply take $\mathscr{F} = 2^{\Omega}$

Probability Assignment for Discrete Sample Spaces

Given $(\Omega, \mathscr{F})$, we define $\mathbb{P} : \mathscr{F} \to [0, 1]$ as

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}), \quad A \in \mathscr{F},$$

while making sure that the assignment $\mathbb{P}$ satisfies $\sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 1$.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Examples

- $\Omega = \{H, T\}, \quad \mathscr{F} = 2^{\Omega} = \left\{\emptyset, \Omega, \{H\}, \{T\}\right\}$

Examples

- $\Omega = \{H, T\}, \quad \mathscr{F} = 2^{\Omega} = \left\{\emptyset, \Omega, \{H\}, \{T\}\right\}$

$$\mathbb{P}(\{H\}) = p = 1 - \mathbb{P}(\{T\}), \quad p \in [0, 1]$$

- $\Omega = \mathbb{N}, \quad \mathscr{F} = 2^{\Omega}$

Examples

- $\Omega = \{H, T\}, \quad \mathscr{F} = 2^\Omega = \Big\{\emptyset, \Omega, \{H\}, \{T\}\Big\}$

$$\mathbb{P}(\{H\}) = p = 1 - \mathbb{P}(\{T\}), \quad p \in [0, 1]$$

- $\Omega = \mathbb{N}, \quad \mathscr{F} = 2^\Omega$

$$\mathbb{P}(\{k\}) = \dots \qquad \text{such that} \qquad \sum_{k=1}^{\infty} \mathbb{P}(\{k\}) = 1.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Examples

- $\Omega = \{H, T\}, \quad \mathscr{F} = 2^{\Omega} = \Big\{\emptyset, \Omega, \{H\}, \{T\}\Big\}$

$$\mathbb{P}(\{H\}) = p = 1 - \mathbb{P}(\{T\}), \quad p \in [0, 1]$$

- $\Omega = \mathbb{N}, \quad \mathscr{F} = 2^{\Omega}$

$$\mathbb{P}(\{k\}) = \ldots \qquad \text{such that} \qquad \sum_{k=1}^{\infty} \mathbb{P}(\{k\}) = 1.$$

  – $\mathbb{P}(\{k\}) = p(1-p)^{k-1}, \quad k \in \Omega \quad (p \in [0, 1]$, Geometric measure)

- $\Omega = \mathbb{N} \cup \{0\}, \quad \mathscr{F} = 2^{\Omega}$

  – $\mathbb{P}(\{k\}) = e^{-\lambda} \dfrac{\lambda^k}{k!}, \quad k \in \Omega \quad (\lambda > 0$, Poisson measure)

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

$\mathbb{P}_B$ is a Valid Probability Measure

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

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

$\mathbb{P}_B$ is a Valid Probability Measure

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

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

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

$\mathbb{P}_B$ is a Valid Probability Measure

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

- $\mathbb{P}_B(\emptyset) = 0$
- $\mathbb{P}_B(\Omega) = 1$
- For any mutually disjoint collection of sets $A_1, A_2, \ldots \in \mathscr{F}$,

$$\mathbb{P}_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}_B(A_i).$$

Conditional Probability – Properties

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

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

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

Conditional Probability – Properties

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

- (Law of Total Probability)
  Let $B_1, B_2, \ldots \in \mathscr{F}$ be a *partition* of Ω, i.e., $B_i \cap B_j = \emptyset$ for all $i \neq j$, and $\bigcup_{i=1}^{\infty} B_i = \Omega$. Then, for any $A \in \mathscr{F}$,

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i).$$

Conditional Probability – Properties

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

- (Bayes' Theorem)
  Let $B_1, B_2, \ldots \in \mathscr{F}$ be as before. For any $A \in \mathscr{F}$ such that $\mathbb{P}(A) > 0$,

$$\mathbb{P}(B_i | A) = \frac{\mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}{\sum_{j=1}^{\infty} \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}.$$

Conditional Probability – Properties

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

- (Decomposition Rule for Conditional Probabilities)
  Let $A_1, A_2, \ldots \in \mathscr{F}$. Then,

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdots$$
$$= \mathbb{P}(A_1) \cdot \prod_{i=2}^{\infty} \mathbb{P}\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right),$$

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

Independence

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.

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.

Remarks:

- The definition of independence does not involve conditional probabilities
- If $\mathbb{P}(B) > 0$, then

$$A \perp\!\!\!\perp B \quad \Longleftrightarrow \quad \mathbb{P}(A|B) = \mathbb{P}(A).$$

Independence of 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.

Random Variables

Random Variables

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

Definition (Random Variables)

A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F} \qquad \forall x \in \mathbb{R}.$$

Random Variables

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

Definition (Random Variables)

A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F} \qquad \forall x \in \mathbb{R}.$$

Remarks:

- A random variable is neither random nor a variable; it is a deterministic function

Random Variables

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

Definition (Random Variables)

A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F} \qquad \forall x \in \mathbb{R}.$$

Remarks:

- A random variable is neither random nor a variable; it is a deterministic function
- The definition of a random variable is closely tied to the underlying σ-algebra $\mathscr{F}$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Random Variables

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

Definition (Random Variables)

A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F} \qquad \forall x \in \mathbb{R}.$$

Remarks:

- A random variable is neither random nor a variable; it is a deterministic function
- The definition of a random variable is closely tied to the underlying σ-algebra $\mathscr{F}$
- If X is a random variable with respect to $\mathscr{F}$, it is called an $\mathscr{F}$-measurable function

Random Variables

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

Definition (Random Variables)

A function $X : \Omega \to \mathbb{R}$ is called a random variable with respect to $\mathscr{F}$ if

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F} \qquad \forall x \in \mathbb{R}.$$

Remarks:

- A random variable is neither random nor a variable; it is a deterministic function
- The definition of a random variable is closely tied to the underlying σ-algebra $\mathscr{F}$
- If X is a random variable with respect to $\mathscr{F}$, it is called an $\mathscr{F}$-measurable function
- The definition of a random variable does not involve $\mathbb{P}$

Examples

- $\Omega = \{1, 2, \ldots, 6\}, \qquad \mathscr{F} = \Big\{\emptyset, \Omega\Big\}, \qquad X(\omega) = \omega$

  Is X a random variable with respect to $\mathscr{F}$?

Examples

- $\Omega = \{1, 2, \ldots, 6\}, \qquad \mathscr{F} = \Big\{\emptyset, \Omega\Big\}, \qquad X(\omega) = \omega$

  Is X a random variable with respect to $\mathscr{F}$?

- What functions X are random variables with respect to $\mathscr{F}$?

Examples

- $\Omega = \{1, 2, 3, 4, 5\}, \qquad \mathscr{F} = \left\{\emptyset, \Omega, A, A^c\right\}$ for a fixed $A \subseteq \Omega$

  What functions X are random variables with respect to $\mathscr{F}$?

Examples

- $\Omega = \{1, 2, 3, 4, 5\}, \qquad \mathscr{F} = \sigma\left(\left\{\{1\}, \{2, 3\}\right\}\right)$

  What functions X are random variables with respect to $\mathscr{F}$?

Examples

- $\Omega = \mathbb{N}, \qquad \mathscr{F} = 2^{\Omega}$
  What functions X are random variables with respect to $\mathscr{F}$?