भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Conditional Probability, Bayes' Theorem, Independence of Events, Independence of σ-Algebras, Borel–Cantelli Lemma, Random Variables, CDF and its Properties

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

19/22 August 2024

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

$\mathbb{P}_B$ 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 $\bigcup_{i=1}^{\infty} B_i = \Omega$. Further, let $\mathbb{P}(B_i) > 0$ for all $i \in \mathbb{N}$. 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.

Remarks:

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

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

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

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.

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 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

Borel–Cantelli Lemma

Borel–Cantelli Lemma – Part 1

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

Lemma (Borel–Cantelli Lemma, Part 1)

Suppose $A_1, A_2, \ldots \in \mathscr{F}$ are such that $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < +\infty$. Then,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Borel–Cantelli Lemma – Part 2

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

Lemma (Borel–Cantelli Lemma, Part 2)

Suppose $A_1, A_2, \ldots \in \mathscr{F}$ are independent and satisfy $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = +\infty$. Then,

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

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

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F} \qquad \forall B \in \mathscr{B}(\mathbb{R}).$$

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

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F} \qquad \forall B \in \mathscr{B}(\mathbb{R}).$$

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 $\mathscr{F}$

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

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F} \qquad \forall B \in \mathscr{B}(\mathbb{R}).$$

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

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathscr{F} \qquad \forall B \in \mathscr{B}(\mathbb{R}).$$

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 $\mathbb{P}$

A Figure to Keep in Mind

$$
\begin{tikzcd}
(\Omega, \mathscr{F}) \arrow[rr, bend left, "X"] & & (\mathbb{R}, \mathscr{B}(\mathbb{R})) \arrow[d] \\
X^{-1}(B) \arrow[u] & & B \arrow[ll, "\text{pre-image}"']
\end{tikzcd}
$$

$$\forall B \in \mathscr{B}(\mathbb{R}), \quad X^{-1}(B) \in \mathscr{F}$$

Figure: A pictorial representation of the definition of an $\mathscr{F}$-measurable function

Simpler, Yet Equivalent, Definition of Random Variable

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

Recall from your homework that $\mathscr{B}(\mathbb{R}) = \sigma(\mathscr{D})$, where

$$\mathscr{D} = \left\{(-\infty, x] : x \in \mathbb{R}\right\}.$$

Definition (Random Variable)

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

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{F} \qquad \forall x \in \mathbb{R}.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Random Variable Simplified

$$
\begin{tikzcd}
(\Omega, \mathscr{F}) \arrow[r, "X"] & (\mathbb{R}, \mathscr{B}(\mathbb{R})) \arrow[d] \\
X^{-1}((-\infty, x]) \arrow[u] & (-\infty, x] \arrow[l, "\text{pre-image}"']
\end{tikzcd}
$$

$$\forall x \in \mathbb{R}, \quad X^{-1}((-\infty, x]) \in \mathscr{F}$$

Figure: Simplified, yet equivalent, definition of random variable

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

- $\Omega = [0,1], \qquad \mathscr{F} = \left\{\emptyset, \Omega, A, A^c\right\}$ 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 $\mathscr{F}$?

Indicator Functions

Fix a sample space Ω.
Fix a subset $A \subseteq \Omega$.

Definition (Indicator Function)

The indicator function of set A is a function $X : \Omega \to \mathbb{R}$ such that

$$X(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \in A^c. \end{cases}$$

The indicator function of set A is commonly denoted as $\mathbf{1}_A$

Exercise

Show that the function $\mathbf{1}_A$ is $\mathscr{F}$-measurable if and only if $A \in \mathscr{F}$.

Probability Law of a Random Variable

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

Definition (Probability Law)

Given a random variable $X : \Omega \to \mathbb{R}$ with respect to $\mathscr{F}$, its probability law $\mathbb{P}_X$ is a probability measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ defined as

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) \quad \forall B \in \mathscr{B}(\mathbb{R}).$$

Remarks:

- $\mathbb{P}_X$ is sometimes called the pushforward of $\mathbb{P}$ under the random variable X

Probability Law of a Random Variable

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

Definition (Probability Law)

Given a random variable $X : \Omega \to \mathbb{R}$ with respect to $\mathscr{F}$, its probability law $\mathbb{P}_X$ is a probability measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ defined as

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) \quad \forall B \in \mathscr{B}(\mathbb{R}).$$

Remarks:

- $\mathbb{P}_X$ is sometimes called the pushforward of $\mathbb{P}$ under the random variable X
- $\mathbb{P}_X$ is sometimes denoted as $\mathbb{P} \circ X^{-1}$

Probability Law of a Random Variable

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

Definition (Probability Law)

Given a random variable $X : \Omega \to \mathbb{R}$ with respect to $\mathscr{F}$, its probability law $\mathbb{P}_X$ is a probability measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ defined as

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) \quad \forall B \in \mathscr{B}(\mathbb{R}).$$

Remarks:

- $\mathbb{P}_X$ is sometimes called the pushforward of $\mathbb{P}$ under the random variable X
- $\mathbb{P}_X$ is sometimes denoted as $\mathbb{P} \circ X^{-1}$
- $\mathbb{P}_X(B)$ for every $B \in \mathscr{B}(\mathbb{R})$ gives the full probabilistic description of X

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Completing the Picture

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
(\Omega, \mathscr{F}, \mathbb{P}) \arrow[r, "X"] & (\mathbb{R}, \mathscr{B}(\mathbb{R}), {\color{red}\mathbb{P}_X}) \arrow[d] \\
X^{-1}(B) \arrow[u] & B \arrow[l, "\text{pre-image}"']
\end{tikzcd}
$$

$$
{\color{red}\mathbb{P}_X}(B) = \mathbb{P} \circ X^{-1}(B) = \mathbb{P}(X^{-1}(B)) \quad \forall B \in \mathscr{B}(\mathbb{R})
$$

Figure: Pictorial representation of probability law

Cumulative Distribution Function

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Cumulative Distribution Function (CDF)

$$
\begin{tikzcd}
(\Omega, \mathscr{F}, \mathbb{P}) \arrow[r, "X"] & (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P}_X) \arrow[d] \\
X^{-1}((-\infty, x]) \arrow[u] & (-\infty, x] \arrow[l, "\text{pre-image}"']
\end{tikzcd}
$$

$$\color{red}{F_X}\color{black}(x) = \mathbb{P}_X((-\infty, x]) = \mathbb{P}(X^{-1}((-\infty, x])), \quad x \in \mathbb{R}$$

Definition (Cumulative Distribution Function)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Given a random variable $X : \Omega \to \mathbb{R}$ with respect to $\mathscr{F}$, its cumulative distribution function (CDF) $F_X : \mathbb{R} \to [0, 1]$ is defined as

$$F_X(x) = \mathbb{P}_X((-\infty, x]) = \mathbb{P}(X^{-1}(\infty, x]), \qquad x \in \mathbb{R}.$$

CDF $\longleftrightarrow$ Probability Law

- If we know $\mathbb{P}_X = \{\mathbb{P}_X(B) : B \in \mathscr{B}(\mathbb{R})\}$, then we can extract the CDF $F_X : \mathbb{R} \to [0, 1]$ by using the formula

$$F_X(x) = \mathbb{P}_X((-\infty, x]), \qquad x \in \mathbb{R}.$$

- Given the CDF $F_X : \mathbb{R} \to [0, 1]$, let

$$\mathbb{P}_X((-\infty, x]) = F_X(x), \qquad x \in \mathbb{R}.$$

Then, there exists a unique extension of $\mathbb{P}_X$ to all Borel subsets of $\mathbb{R}$
For a proof of this, see [Folland, 1999, Theorem 1.16]

Notation

- $\{\omega \in \Omega : X(\omega) \leq x\} = \{X \leq x\}$
- $\mathbb{P}_X((-\infty, x]) = \mathbb{P}(X^{-1}((-\infty, x])) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\}) = \mathbb{P}(X \leq x)$

Properties of CDF

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$
Let $X : \Omega \to \mathbb{R}$ be a random variable with respect to $\mathscr{F}$ with CDF F_X

- $\lim_{x\to-\infty} F_X(x) = 0, \qquad \lim_{x\to+\infty} F_X(x) = 1$
- (Monotonicity) If $x \le y$, then $F_X(x) \le F_X(y)$
- (Right-Continuity) F_X is right-continuous, i.e., for all $x \in \mathbb{R}$,

$$\lim_{\varepsilon \downarrow 0} F_X(x+\varepsilon) = F_X(x).$$

References

Folland, G. B. (1999).
Real analysis: modern techniques and their applications, volume 40.
John Wiley & Sons.