



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

Conditional Probabilities

Conditional Probability Measure

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

Definition (Conditional Probability)

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

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

Then, \mathbb{P}_B is a valid probability measure on (Ω, \mathcal{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 $(\Omega, \mathcal{F}, \mathbb{P})$.

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



\mathbb{P}_B is a Valid Probability Measure

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

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

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

\mathbb{P}_B is a Valid Probability Measure

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

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

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

Conditional Probability – Properties

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

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

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

Conditional Probability – Properties

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

- (Law of Total Probability)

Let $B_1, B_2, \dots \in \mathcal{F}$ 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 $A \in \mathcal{F}$,

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

Conditional Probability – Properties

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

- (Bayes' Theorem)

Let $B_1, B_2, \dots \in \mathcal{F}$ be as before. For any $A \in \mathcal{F}$ 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 $(\Omega, \mathcal{F}, \mathbb{P})$.

- (Decomposition Rule for Conditional Probabilities)

Let $A_1, A_2, \dots \in \mathcal{F}$. Then,

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

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

Independence

Independence of Events

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

Definition (Independence of Events)

Events $A, B \in \mathcal{F}$ 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

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

Independence of Events

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

Definition (Independence of Events)

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

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

- 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 $(\Omega, \mathcal{F}, \mathbb{P})$

Definition (Independence of σ -Algebras)

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

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

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

Borel–Cantelli Lemma

Borel-Cantelli Lemma - Part 1

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

Lemma (Borel-Cantelli Lemma, Part 1)

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

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

Borel-Cantelli Lemma - Part 2

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

Lemma (Borel-Cantelli Lemma, Part 2)

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

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

Random Variables

Random Variables

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

Definition (Random Variables)

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

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

Remarks:

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

Random Variables

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

Definition (Random Variables)

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

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

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 \mathcal{F}

Random Variables

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

Definition (Random Variables)

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

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

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 \mathcal{F}
- If X is a random variable with respect to \mathcal{F} , it is called an **\mathcal{F} -measurable function**

Random Variables

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

Definition (Random Variables)

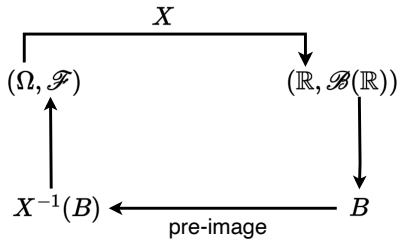
A function $X : \Omega \rightarrow \mathbb{R}$ is called a **random variable** with respect to \mathcal{F} if

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

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 \mathcal{F}
- If X is a random variable with respect to \mathcal{F} , it is called an **\mathcal{F} -measurable function**
- The definition of a random variable does not involve \mathbb{P}

A Figure to Keep in Mind



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

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

Simpler, Yet Equivalent, Definition of Random Variable

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

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

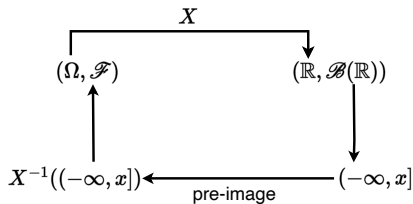
$$\mathcal{D} = \left\{ (-\infty, x] : x \in \mathbb{R} \right\}.$$

Definition (Random Variable)

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

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

Random Variable Simplified



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

Figure: Simplified, yet equivalent, definition of random variable

Examples

- $\Omega = \{1, 2, \dots, 6\}$, $\mathcal{F} = \{\emptyset, \Omega\}$, $X(\omega) = \omega$

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

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

Examples

- $\Omega = [0, 1]$, $\mathcal{F} = \left\{ \emptyset, \Omega, A, A^c \right\}$ for a fixed $A \subseteq \Omega$

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

Examples

- $\Omega = \{1, 2, 3, 4, 5\}, \quad \mathcal{F} = \sigma \left(\left\{ \{1\}, \{2, 3\} \right\} \right)$
What functions X are random variables with respect to \mathcal{F} ?

Examples

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

What functions X are random variables with respect to \mathcal{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 $X : \Omega \rightarrow \mathbb{R}$ 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 \mathcal{F} -measurable if and only if $A \in \mathcal{F}$.

Probability Law of a Random Variable

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

Definition (Probability Law)

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

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

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 $(\Omega, \mathcal{F}, \mathbb{P})$

Definition (Probability Law)

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

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

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 $(\Omega, \mathcal{F}, \mathbb{P})$

Definition (Probability Law)

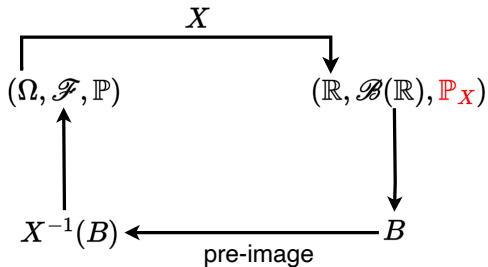
Given a random variable $X : \Omega \rightarrow \mathbb{R}$ with respect to \mathcal{F} , its **probability law** \mathbb{P}_X is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined as

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

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 $B \in \mathcal{B}(\mathbb{R})$ gives the full probabilistic description of X

Completing the Picture

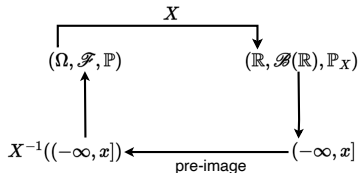


$$\mathbb{P}_X(B) = \mathbb{P} \circ X^{-1}(B) = \mathbb{P}(X^{-1}(B)) \quad \forall B \in \mathcal{B}(\mathbb{R})$$

Figure: Pictorial representation of probability law

Cumulative Distribution Function

Cumulative Distribution Function (CDF)



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

Definition (Cumulative Distribution Function)

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

Given a random variable $X : \Omega \rightarrow \mathbb{R}$ with respect to \mathcal{F} , its **cumulative distribution function (CDF)** $F_X : \mathbb{R} \rightarrow [0, 1]$ is defined as

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

CDF \longleftrightarrow Probability Law

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

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

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

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

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 $(\Omega, \mathcal{F}, \mathbb{P})$

Let $X : \Omega \rightarrow \mathbb{R}$ be a random variable with respect to \mathcal{F} with CDF F_X

- $\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1$
- (**Monotonicity**) If $x \leq y$, then $F_X(x) \leq F_X(y)$
- (**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.