

EECS 126 Lecture Notes

Dun-Ming Huang

Contents

1	Introduction, Probability Spaces, Conditional Probability Theorem	3
1.1	Set Operation	3
1.2	Probabilistic Model	5

Chapter 1

Introduction, Probability Spaces, Conditional Probability Theorem

Course Policies

As always, first lecture is mostly about course policies.

- Look over syllabus and course website here
- Check Ed and Gradescope
- Industrial revolution of generative AI :0

Textbook doesn't cover all lecture contents.

Concept of Probability

The concept of probability theory began with the concept of frequency, where, a specific outcome occurs at a certain frequency that denotes the probability of that outcome in some sequence of experiments. For many applications, this stays true. However, in other applications where the notion of frequency does not apply (eg., if a presidential candidate wins at a probability of 50%, does that candidate win on the frequency of 0.5 as well?). Consequently, we find a different mathematical formalism to illustrate the notion of probability. And, with the diversity of interpretations, many frameworks exist for the study of probability. In this course, via exploring numerous examples at which probability theories are applied and investigated, we develop our intuition toward the content of probability theory.

1.1 Set Operation

1.1.1 Definition of Sets

Let us begin with the definition of sets.

Definition 1.1.1. Set

A **set** is a collection of objects. The objects contained in a set is called the **element** of that set.

Sets are usually denoted using capital letters; for example, A .

1.1.2 Element-wise Operations

Suppose that an element x belongs to a set \mathcal{X} , we denote it as follows:

Symbol 1.1.1. Inclusion

The expression:

$$x \in \mathcal{X}$$

denotes that the element x belongs to a set \mathcal{X} .

Meanwhile, the expression:

$$x \notin \mathcal{X}$$

denotes that the element x does not belong to \mathcal{X} .

Usually, sets are written in the following format:

$$\mathcal{X} = \{x_1, x_2, \dots, x_n\}$$

In the case where the set contains infinite elements, we may denote as follows:

$$\mathcal{X} = \{x_1, x_2, \dots\}$$

To build a set conditionally, we may use a set-builder notation:

Symbol 1.1.2. Set-Builder Notation

$$\{x | x \text{ satisfies predicate } p\}$$

For example, the set of all integers that are divisible is denoted as:

$$\{2k | k \in \mathbb{Z}\}$$

1.1.3 Set-Set Relations

Sets can be affiliated with each other:

Definition 1.1.2. Subsets

If every element of some set \mathcal{X} is also an element of set \mathcal{Y} , we state that \mathcal{X} is a subset of \mathcal{Y} :

$$\mathcal{X} \subseteq \mathcal{Y}$$

Definition 1.1.3. Equality of Sets

Sets \mathcal{X} and \mathcal{Y} are equal if $\mathcal{X} \subseteq \mathcal{Y} \wedge \mathcal{Y} \subseteq \mathcal{X}$.

Last but not least, let us discuss the notion of universality in set theory:

Definition 1.1.4. Universal Set

A universal set Ω is the set of all elements that will be related to the application of discussion in mathematics.

On the other hand, empty sets denotes an empty set with no elements. By definition, the empty set is a subset of every set.

Sets can also be related to each other based on their elemental differences.

Definition 1.1.5. Complements

The complement of a set \mathcal{X} is defined as follows:

$$\mathcal{X}^C = \{x | x \in \omega \wedge x \notin \mathcal{X}\}$$

On the other hand, the notations “union” and “intersection” are defined as taught before.

$$\begin{aligned} \mathcal{X} \cup \mathcal{Y} &= \{x | x \in \mathcal{X} \vee x \in \mathcal{Y}\} && \text{(Union)} \\ \mathcal{X} \cap \mathcal{Y} &= \{x | x \in \mathcal{X} \wedge x \in \mathcal{Y}\} && \text{(Intersection)} \end{aligned}$$

By definition, then,

$$\begin{aligned} (S^C)^C &= S \\ S \cup \Omega &= \Omega \end{aligned}$$

1.1.4 Algebraic Rules of Set Arithmetics

There exists commutativity in the algebra of sets:

$$\forall S, T : S \cup T = T \cup S$$

There also exists associativity in the algebra of sets:

$$\forall S, T, M : S \cap (T \cup M) = (S \cap T) \cup (S \cap M)$$

And, there’s the famous DeMorgan’s Law:

$$\begin{aligned} \left(\bigcup_n S_n \right)^C &= \bigcap_n S_n^C \\ \left(\bigcap_n S_n \right)^C &= \bigcup_n S_n^C \end{aligned}$$

1.2 Probabilistic Model

There are two components to a probabilistic model:

1. A sample space Ω that refers to the set of all possible outcomes of an experiment.
2. A probability law that assigns nonnegative number $\mathbb{P}(A)$ to each subset $A \subseteq \Omega$ that satisfies Kolmogorov’s axioms.

Let us proceed with a probabilistic model of dice rolls:

Example 1.2.1: Probabilistic Model of Dice Rolls

The sample space of a dice roll may be:

$$\Omega = \{1, 2, \dots, 6\}$$

where the probability law for a fair dice works to be, for example,

$$\mathbb{P}(\{1\}) = \frac{1}{6}$$

All probabilistic models follow the following Kolmogorov’s axioms:

Theorem 1.2.1. Kolmogorov's Axioms

1. $\mathbb{P}(A) \geq 0$
2. Countable-Additivity: If A and B are disjoint events, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
3. $\mathbb{P}(\Omega) = 1$

Among which, we may expand countable-additivity into the following condition:

For a sequence of disjoint, countably infinite events A_1, A_2, \dots ,

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

With these axioms in mind, let us consider an alternative example regarding probabilistic model of dices:

Example 1.2.2: Probabilistic Model of Pairs of Dice Rolls

Let all possible results from one single four-sided dice rolls be represented by:

$$\omega = \{1, 2, 3, 4\}$$

The sample space of a pair of four-sided dice rolls may be represented as:

$$\Omega = \{(i, j) | i \in \omega \wedge j \in \omega\}$$

Notably, this discrete sample space entertains a uniform probability for all possible outcome. In such case, the discrete uniform probability law states that, for some event A :

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

We further derive the following properties of probability laws as the consequence of Kolmogorov's Axioms:

1. $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$
2. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
3. As implied from the previous rule: $\mathbb{P}(A) + \mathbb{P}(B) \geq \mathbb{P}(A \cap B)$
4. $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(A^C \cap B) + \mathbb{P}(A^C \cap B^C \cap C)$

1.3 Conditional Probability

Conditioning is the essence of many probabilistic applications, and primary method of constructing complicated probabilistic measures.

Definition 1.3.1. Conditional Probability

The probability of an event A provided the transpiration of event B is denoted as:

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

We may observe that if the event A is not independent from event B , then the probability we discuss now is fundamentally different from the sole probability of event A 's occurrence.

For demonstration, let us state that A is the event where a fair dice throw presents value 3, and B is where an even value is presented. Then, the event A is dependent upon B such that, with the occurrence of B , event A shall never occur. Therefore, $\mathbb{P}(A|B) = 0$.

On the other hand, let C be the event where value 4 is presented, then $\mathbb{P}(A|C) = \frac{1}{3}$.