Lecture 1

Set theory, probability law

Bibliography

1. Dimitri P. Bertsekas and John N. Tsitsiklis «**Introduction to Probability»**
2. Seymour Lipschutz, Marc Lipson **«Probability»**

Set theory

Probability makes extensive use of set operations**.**

A **set** is a collection of objects, which are the elements of the set. If S is a set and x is an element of S, we write $x \epsilon S$. If x is not an element of S, we write $x \notin S$.

A set can have no elements, in which case it is called the **empty set** $\emptyset$.
$\Omega - \boldsymbol{universal\ set}$

The complement of a set S, with respect to the universe Ω, is the set $S^{c} = \{x \in \Omega | x \notin S\}$.

The union of two sets S and T is the set of all elements that belong to S or T (or both), and is denoted by $S \cup T = \{x | x \in S \text{ или } x \in T\}$.

The intersection of two sets S and T is the set of all elements that belong to both S and T, and is denoted by $S \cap T = \{x | x \in S \text{ и } x \in T\}$.

The difference of two sets S and T is a set consisting of elements that belong to the set S and do not belong to the set T: $S \backslash T = \{x | x \in S \text{ и } x \notin T\}$

The algebra of sets

$$
\begin{aligned}
S \cup T &= T \cup S. & S \cup (T \cup U) &= (S \cup T) \cup U, \\
S \cap (T \cup U) &= (S \cap T) \cup (S \cap U). & S \cup (T \cap U) &= (S \cup T) \cap (S \cup U). \\
(S^c)^c &= S. & S \cap S^c &= \emptyset. \\
S \cup \Omega &= \Omega. & S \cap \Omega &= S.
\end{aligned}
$$

De Morgan's laws:

$$
\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c, \qquad \left(\bigcap_n S_n\right)^c = \bigcup_n S_n^c.
$$

Examples

1. On the Venn diagram (let all possible intersections be present), select the sets:

-$A \cap (B \cup C)$; -$A^C \cap (B \cup C)$; -$(A \cup C) \cap (B \cup C)$

2. Prove the law of distributivity $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (formally by parsing boolean expressions and with Venn diagram)

3. There are 30 students on list A and 35 students on list B. 20 students are on both list A and list B.

Define:

-number of students only in list A

-number of students only in list B

-number of students in List A or List B

-number of students contained in exactly one list

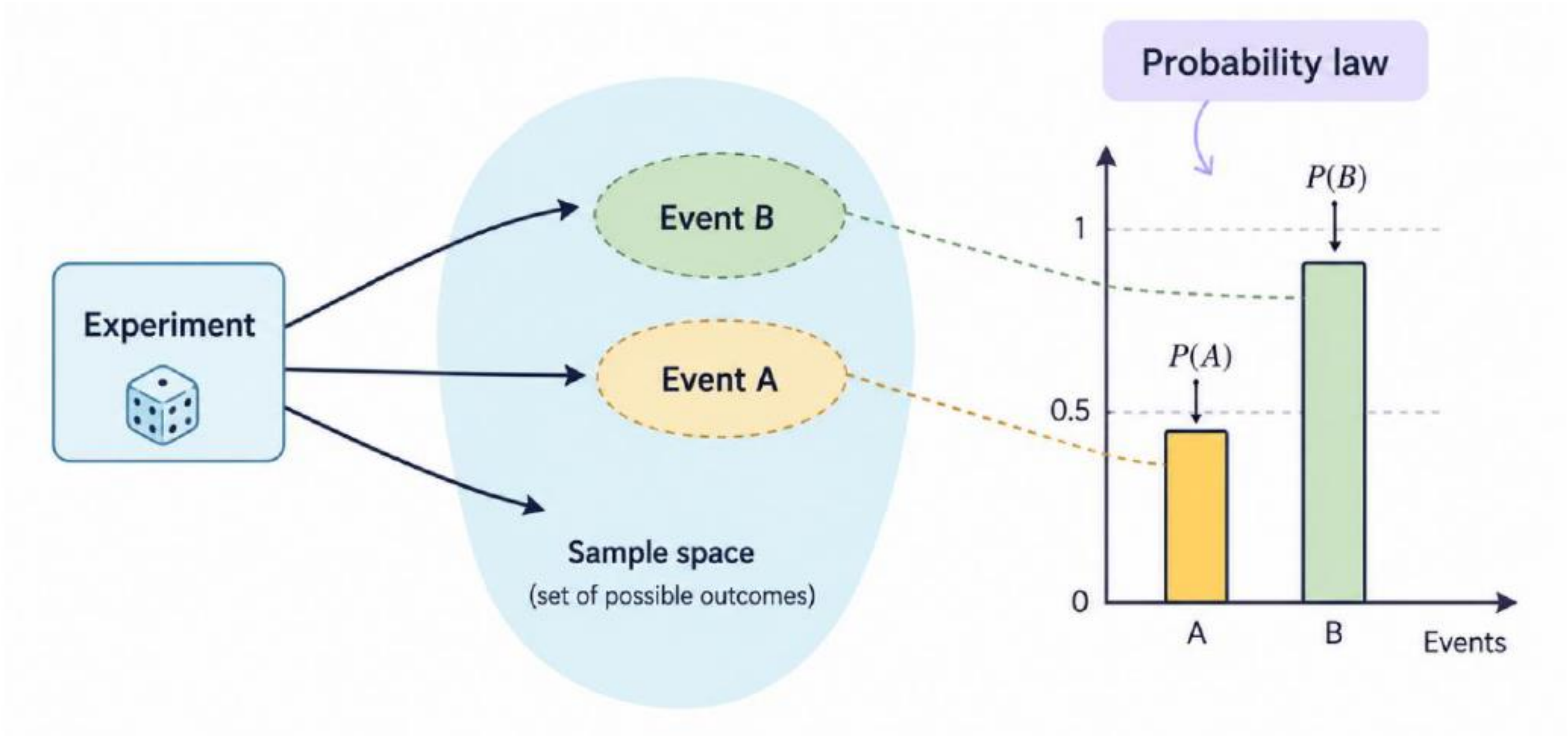
Probabilistic models

A probabilistic model is a mathematical description of an uncertain situation. It must be in accordance with a fundamental framework. Its two main ingredients are listed below and are visualized

Elements of a Probabilistic Model

- The **sample space** Ω, which is the set of all possible **outcomes** of an experiment.
- The **probability law**, which assigns to a set A of possible outcomes (also called an **event**) a nonnegative number $\mathbf{P}(A)$ (called the **probability** of A) that encodes our knowledge or belief about the collective "likelihood" of the elements of A. The probability law must satisfy certain properties to be introduced shortly.

Probabilistic models



The main ingredients of probabilistic model

Probabilistic models

Example. Consider a random experiment in which a six-sided die is rolled once. The result of the experiment is the number that appears on the upper face of the die. The possible results are: 1, 2, 3, 4, 5, 6. These results are called outcomes. The set of all possible outcomes forms the sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Axioms of probability

Let S be a sample space, let Ω be the class of all events, and let P be a real-valued function defined on Ω. Then P is called a *probability function*, and $P(A)$ is called the *probability* of the event A, when the following axioms hold:

- [**P**1] For any event A, we have $P(A) \geq 0$.
- [**P**2] For the certain event S, we have $P(S) = 1$.
- [**P**3] For any two disjoint events A and B, we have $P(A \cup B) = P(A) + P(B)$
- [**P**3] For any infinite sequence of mutually disjoint events $A1$, $A2$, $A3$, . . ., we have
$$P(A1 \cup A2 \cup A3 \cup \cdots) = P(A1) + P(A2) + P(A3) + \cdots$$

Furthermore, when P does satisfy the above axioms, the sample space S will be called a *probability space*.

Theorems of probability spaces

Theorem 1: The impossible event or, in other words, the empty set has probability zero, that is, $P(\emptyset) = 0$

Theorem 2 (Complement Rule): For any event A, we have $P(A^c) = 1 - P(A)$

Theorem 3: For any event A, we have $0 \leq P(A) \leq 1$.

Theorem 4: If $A \subseteq B$, then $P(A) \leq P(B)$.

Theorem 5: For any two events A and B, we have $P(A \backslash \mathrm{B}) = P(A) - P(A \cap B)$

Theorem (Addition Rule) 6: For any two events A and B,

$$\mathrm{P}(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probability axioms. Example 1

Consider an experiment involving a **single roll of a four-sided die**. There are four possible outcomes: 1, 2, 3, and 4. The sample space is $\Omega = \{1, 2, 3, 4\}$. The possible events include $\{1, 2, 3, 4\}$, $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\emptyset$. If the die is fair, all possible outcomes are considered equally likely. Therefore, each elementary outcome is assigned the same probability: $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = 0.25$.

- According to the additivity axiom, the probability of the entire sample space is the sum of the probabilities of all mutually exclusive outcomes: $P(\{1,2,3,4\}) = P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) = 1$.
- This result agrees with the normalization axiom, which states that the probability of the sample space must be equal to one. Therefore, the probability law for this experiment is: $P(\{1,2,3,4\}) = 1$, $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = 0.25$, $P(\emptyset) = 0$.
- This probability law satisfies all three probability axioms: nonnegativity, additivity, and normalization.

Probability axioms. Example 2

Consider an experiment in which two fair six-sided dice are rolled once. The outcome is an ordered pair (i, j), where i is the number shown by the first die and j is the number shown by the second die. The sample space is: $\Omega = \{(1,1), (1,2), ..., (6,5), (6,6)\}$. The sample space contains 36 possible outcomes. We assume that each possible outcome has the same probability $P(\{(i,j)\}) = 1/36$ for every possible outcome in Ω.

Let us construct a probability law that satisfies the three probability axioms. Consider, as an example, the event A = {the sum of the two dice is equal to 7}.

The event A consists of the following outcomes: $A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$.

Using the additivity axiom, the probability of event A is the sum of the probabilities of its individual outcomes:

$$P(A) = P(\{(1,6)\}) + P(\{(2,5)\}) + P(\{(3,4)\}) + P(\{(4,3)\}) + P(\{(5,2)\}) + P(\{(6,1)\}) = 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 = 6/36 = 1/6$$

Discrete uniform probability law

Discrete Uniform Probability Law

If the sample space consists of n possible outcomes which are equally likely (i.e., all single-element events have the same probability), then the probability of any event A is given by

$$\mathbf{P}(A) = \frac{\text{number of elements of } A}{n}.$$

Discrete uniform probability law. Example 3

Two cards are drawn with replacement from a set of four cards numbered 1, 2, 3, and 4. The outcome is an ordered pair (i, j), where i is the number on the first card drawn and j is the number on the second card drawn. Since each draw has four possible results, the sample space contains $4 \cdot 4 = 16$ equally likely outcomes.

We assume that each card is equally likely to be drawn. Therefore, there are sixteen possible outcomes: $\Omega = \{(1,1), (1,2), (1,3), (1,4), (2,1), \ldots, (4,4)\}$

Each possible outcome has the same probability: $P(\{(i,j)\}) = 1/16$, $i = 1,..,4$, $j = 1,..,4$.

To calculate the probability of an event, we count the number of outcomes belonging to this event and divide by the total number of possible outcomes.

For example:

- *P({the sum of the two numbers is even}) = 8/16 = ½;*
- *P({the sum of the two numbers is greater than 5}) = 6/16 = 3/8;*
- *P({the first number is equal to the second}) = 4/16 = ¼;*
- *P({the first number is smaller than the second}) = 6/16 = 3/8;*
- *P({at least one drawn number is equal to 1}) = 7/16.*

Complement rule. Example 4

A software company has 50 programmers. Among them, 8 programmers are currently working on a critical project. One programmer is selected at random. Find the probability that the selected programmer is not working on the critical project.

Practical lesson

Examples

1. Prove the De Morgan's law: $(A \cup B)^c = A^c \cap B^c$ (formally by parsing boolean expressions and with Venn diagram).
2. For a six-sided die, consider the events A = {the number of points is even}, B = {the number of points is greater than 3}. Find and compare sets in both sides of the equality:

$$-(A \cup B)^c = A^c \cap B^c$$
$$-(A \cap B)^c = A^c \cup B^c$$

3. 2 coins and a dice are tossed.

a) describe a suitable set of all outcomes and find the number of elements in it

b) describe the set of outcomes of events:

-A= {2 heads rolled and even number of points};

-B= {score rolled 2};

-C= {dropped 1 heads and an odd number of points};

-A and B

-only in B

-B and C

-A but not B

Examples

4. Give examples of sets for which $A \cap B = A \cap C, but\ B \neq C; A \cup B = A \cup C, but\ B \neq C.$

5. A two-digit number is conceived. Find the probability that the conceived number will be:

a) a randomly named two-digit number;

b) a randomly named two-digit number, the digits of which are different.

6. A cube, all sides of which are painted, is sawn into a thousand cubes of the same size, which are then thoroughly mixed. Find the probability that a randomly drawn cube has colored faces:

a) one;

b) two;

c) three.

7. Two fair dice are rolled. Find the probabilities of the following events:

- the sum of the numbers rolled is equal to 9;
- the sum of the numbers rolled is equal to 10, and the difference between the numbers on the dice is 2;
- the sum of the numbers rolled is equal to 10, if it is known that the absolute difference between the numbers on the dice is 2;
- the sum of the numbers rolled is equal to 6, and the product of the numbers on the dice is 8.

Examples

4. Give examples of sets for which $A \cap B = A \cap C, but\ B \neq C; A \cup B = A \cup C, but\ B \neq C.$

5. A two-digit number is conceived. Find the probability that the conceived number will be:

a) a randomly named two-digit number;

b) a randomly named two-digit number, the digits of which are different.

6. A cube, all sides of which are painted, is sawn into a thousand cubes of the same size, which are then thoroughly mixed. Find the probability that a randomly drawn cube has colored faces:

a) one;

b) two;

c) three.

7. Two fair dice are rolled. Find the probabilities of the following events:

- the sum of the numbers rolled is equal to 9;
- the sum of the numbers rolled is equal to 10, and the difference between the numbers on the dice is 2;
- the sum of the numbers rolled is equal to 10, if it is known that the absolute difference between the numbers on the dice is 2;
- the sum of the numbers rolled is equal to 6, and the product of the numbers on the dice is 8.

Examples

8. A three-digit number is randomly chosen from all possible three-digit numbers. Find the probability that the chosen number will be:

   a) a number whose last digit is even;

   b) a three-digit number with all digits different.

Homework

Study guide: p.15-17 Tasks 1-9; p. 26 Task 1; p.27 Tasks 5-9