Lecture 1

Andrei Antonenko

Jan 24, 2005

1 Introduction

The concept of probability is used to describe and discuss an uncertain situation. The possible approach to this is to talk about the **frequency of the occurrence** of some particular event. For example, when we toss a fair coin many time, the head will appear in approximately 50% of tosses, and the tail will appear in also approximately 50% of tosses. Another possibility is to talk about the **subjective belief** about some situation. While rolling a fair die our belief is that any number is equally likely to appear. Probability is trying to formalize these concepts.

2 Set Theory

The set theory is important part of probability theory, thus I will introduce the basic set operations and definitions, related to it.

Definition 2.1. *A **set** is a collection of objects, which are called the **elements** of the set.*

We will denote sets by capital letters, and their elements by small letters. For example, if S is a set, and x is an element of the set S , we will be using notation $x \in S$ in order to specify that x is an element of S and $x \notin S$ in case x is not an element of S.

Definition 2.2. *A set with no elements is called an **empty set** and is denoted by $\emptyset$.*

There are several ways in which the sets can be specified. If the set contains a finite number of elements, it can be specified by enumerating all it's elements, like

$$S = \{x_1, x_2, \dots, x_n\}.$$

For example, if we roll a die, a set of possible outcomes will be the set $S = \{1, 2, 3, 4, 5, 6\}$. If S contains infinitely many elements, which can be arranged as a sequence $x_1, x_2, \dots,$ then we will write

$$S = \{x_1, x_2, \dots\},$$

and in this case the set S is called **countable**. For example, the set of all natural odd numbers can be written as $S = \{1, 3, 5, \dots\}$.

Another way to write a set of elements which satisfy a given property P is

$$S = \{x | x \text{ satisfies } P\}.$$

For example, the set of all numbers between 0 and 1 can be written as $S = \{x|0 \le x \le 1\}$. Notice, that the elements of this set can not be enumerated and listed as a sequence, so, such sets are called **uncountable**.

Definition 2.3. *If every element of the set T is also an element of the set S, we say that T is a* ***subset*** *of S, and write it as $T \subset S$ of $S \supset T$. If $S \subset T$ and $T \subset S$, then the two sets are called* ***equal****: $S = T$. In this case these sets have the same elements.*

Often it is important to consider the **universal set (universe**), consisting of all elements which are of interest in particular context. This set will be denoted by Ω. For example, when in the given context we talk about all numbers between 0 and 1, we can take $\Omega = [0, 1]$. If we talk about all integer numbers, we can have $\Omega = \{0, -1, 1, -2, 2, \dots\}$. The importance of the proper choice of the universe will be clear from the next section.

3 Set Operations

Definition 3.1. *The* ***complement*** *of the set S with respect to the universe Ω is a set $S^c = \{x|x \notin S\}$.*

Here the universe is important. For example, if $S = \{0, -2, 2, -4, 4, \dots\}$, and $\Omega = \{0, -1, 1, -2, 2, -3, 3, \dots\}$, the complement of S will be the set of all odd numbers: $S^c = \{1, 3, \dots\}$. In case the universe Ω is the set of all real numbers, the complement of S will consist of all odd numbers, and all fractional numbers, i.e. for example 1.5 will be the member of S^c in the latter but not the former case.

Definition 3.2. *The* ***union*** *of two sets S and T is a set $S \cup T$ defined as*

$$S \cup T = \{x|x \in S \text{ or } x \in T \text{ or in both}\}. \tag{1}$$

Definition 3.3. *The* ***intersection*** *of two sets S and T is a set $S \cap T$ defined as*

$$S \cap T = \{x|x \in S \text{ and } x \in T\}. \tag{2}$$

If we would like to consider the union or intersection of several, maybe infinitely many sets, the appropriate definition can be modified as follows:

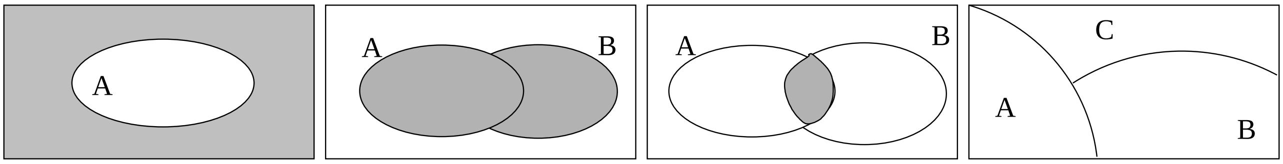
$$\bigcup_{n=1}^{\infty} S_n = \{x|x \in S_n \text{ for some } n\}, \tag{3}$$

$$\bigcap_{n=1}^{\infty} S_n = \{x|x \in S_n \text{ for all } n\}. \tag{4}$$

Definition 3.4. *Two sets are said to be* ***disjoint*** *if their intersection is empty, i.e. they don't have any common elements. More generally, several sets are said to be* ***disjoint*** *if no two of them have common elements.*

Definition 3.5. *A collection of sets is said to form a* ***partition*** *of a set S, if they are disjoint, and they union is equal to S.*

All of these sets can be seen on **Venn's diagrams**. The following four pictures show the complement of A, union of A and B, intersection of A and B, and in the 4th picture sets A, B, and C for a partition of the universe.



4 Algebra of Sets

The set operations have the following properties:

(1) $S \cap T = T \cap S$, $S \cup T = T \cup S$;

(2) $S \cap (T \cap U) = (S \cap T) \cap U$, $S \cup (T \cup U) = (S \cup T) \cup U$;

(3) $S \cap (T \cup U) = (S \cap T) \cup (S \cap U)$, $S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$;

(4) $(S^c)^c = S$;

(5) $S \cap S^c = \emptyset$, $S \cup S^c = \Omega$;

(6) $S \cap \Omega = S$, $S \cup \Omega = \Omega$.

Another two particularly important properties are **de Morgan's laws:**

$$\left(\bigcap_n S_n\right)^c = \bigcup_n S_n^c \tag{5}$$

$$\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c \tag{6}$$

In case of two sets, they say that

$$(S \cap T)^c = S^c \cup T^c \tag{7}$$

and

$$(S \cup T)^c = S^c \cap T^c \tag{8}$$

Proof. To prove the first law let's assume that $x \in (\cup_n S_n)^c$. Then, $x \notin \cup_n S_n$ and thus $x \notin S_n$ for any n. Therefore, $x \in \cap_n S_n^c$. That shows that $(\cup_n S_n)^c \subset \cap_n S_n^c$. By reversing this argument, we establish the converse inclusion. The second law is proven in the similar way. $\square$

5 Probabilistic Models

Every random situation involves an underlying process, which we will be calling an **experiment**. The experiment can be tossing a coin, rolling a die, observing the weather, or a behavior of a stock market. The experiment produces one of the several outcomes. The set of all possible outcomes is called a **sample space**, which will be denoted by Ω. Subsets of the sample space (i.e. collections of possible outcomes) are called **events**. The correct choice of the sample space is important for specifying a probabilistic model. The sample space should be exhaustive, i.e. all possible outcomes should belong to it, and it should have enough details to distinguish between all outcomes, which are of an interest in particular context.

The second part of a probabilistic model is a **probability law**. The probability law is a number $P(A)$ (which is called a **probability of** A), assigned to any event A from Ω, which is intuitively a likelihood of the event A happening. The probability law should satisfy the following axioms:

1. **(Nonnegativity)** $P(A) \geq 0$ for every event A.

2. **(Additivity)** For any two disjoint events A and B,

$$P(A \cup B) = P(A) + P(B). \tag{9}$$

   This axiom is generalized in the following way: if there is an infinite sequence of disjoint events $A_1, A_2, \ldots$, then

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots. \tag{10}$$

3. **(Normalization)** The probability of the universe Ω is equal to 1: $P(\Omega) = 1$.

The probability law can be visualized as a unit mass spread over the sample space. Then $P(A)$ is the total mass assigned to all elements of A.

From these axioms it follows that a probability of the empty event $\emptyset$ is equal to 0:

$$1 = P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) = 1 + P(\emptyset),$$

and therefore, $P(\emptyset) = 0$.

From the additivity axiom it follows, that for any, say, three disjoint events A_1, A_2, A_3,

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3).$$

More properties of the probability law will be considered in the following lecture.

So, we can see, that in order to state a probabilistic model, we need to specify two things:

1. The sample space Ω – set of all possible outcomes of an experiment;

2. The probability law, which assigns a nonnegative number $P(A)$ to all sets of possible outcomes (to subsets of Ω). The probability law must satisfy the axioms, given above.

We will consider several examples of probabilistic model in the next lecture.