

Lecture Notes For: The Fundamentals of Probability

Ali Fele Paranj
alifele@student.ubc.ca

January 18, 2023

This lecture note contains my notes while I was studying the fundamentals of the probability probability. The main content is from the course in term winter2, UBC, 2023. However, I have exanded the content and examples using the material of two books:

- Statistical Modeling and Computation by Kroese
- Probability - Random Variables and Stochastic Processes by Papoulis

1 Fundamentals

The main concept in the field of statistics and probability is the set theory. Basically all we deal with the sets. The whole theory of statistics can be built on that. Let's discuss some fundamental concepts in statistics and then build the theory.

1.1 Random Experiment

To understand the meaning of random experiment, do not over think! The first thing that comes into our minds when we hear the word "random experiment" is its definition! In a nutshell, random experiment is an experiment that its outcome is unknown to us. Like:

- Tossing two coin
- Rolling a dice
- Measuring the number of possible ReadWrite operations on a piece of EEPROM chip

Do not overthink about that. Yes we can go further and discuss stuff like "we can compute the exact movement of dice or coin so it is not random but deterministic" and etc. Here I will not touch the philosophical topics that are very deep and do not necessarily converge to a unified point of view!

The random experiments can be modeled and despite the fact that a random experiment is random, we can deduce many useful information from modeling that. To model a random experiment, we use three important concepts: sample space, events, probability. In the following section, we will discuss each of them in detail.

1.2 Sample Space

Definition: Sample Space

Sample space Ω is simply a set that contains *all possible outcomes* of a random experiment/

For each of random experiments described above, we can define a sample space. For example:

- Ω of Tossing Two Coins:

$$\Omega = \{HH, HT, TH, TT\}$$

- Ω of Rolling a Dice:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Ω of Rolling Two Dices:

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

- Ω of Number of possible ReadWrite operations on a EEPROM chip:

$$\Omega = \mathbb{N}$$

1.3 Events

Definition: Events

Event E is a set of outcomes of a random experiment and is the subset of sample space Ω .

$$E \in \Omega$$

For example for any of the sample spaces specified above, we can define so many possible events. In fact any set that is a subset of the sample space is a valid event of that sample space. For example:

- Tossing Three Coins

- There are at least one Heads:

$$E = \{HHH, HHT, HTH, THH, HTT, THT, TTH\}$$

- There are only two Tails:

$$E = \{TTH, THT, HTT\}$$

- Rolling Two Dices

- The sum of two dices is 4:

$$E = \{(1, 3), (2, 2), (3, 1)\}$$

- there are at least one prime number in the outcome:

$$E = \{(1, 2), (1, 3), (1, 5), (2, 1), (3, 1), (5, 1), (2, 2), (2, 3), (2, 5), \dots, (5, 5)\}$$

Since we have define everything on the basics of set theory, then now we can correspond the everyday concepts to specific operations in the set theory.

Example: The Mapping Between Everyday Language and Sets in the Theory of Probability

- At least one of two events $A, B \in \Omega$ happens: $E = A \cup B$.
- Two events $A, B \in \Omega$ occurs at the same time: $E = A \cap B$.
- Event $A \in \Omega$ does not happen: $E = \bar{A} = \Omega - A$.
- The event A happens but B does not happen: $E = A - B$.

In probability and statistics, we are dealing with three important concepts: sample space Ω , event E , and probability P .

Definition: Disjoint events

If two events has no common elements (i.e. $A \cap B = \emptyset$) then we say that two events are *disjoint*. Basically, if two sets in the venn diagram has nothing is common they are considerent to be disjoint sets.

For example for the random experiment of tossing two coins, the events 1) both coins are heads: $A = \{HH\}$ and 2) both coins are tails: $B = \{TT\}$. Two events A, B are two disjoint events. **Two events being disjoining is NOT the same as being independent.** We will talk about independet events in future.

Note that since the events are basically sets, we can use theorems of set theory to solve the problems.

Theorem: De Morgan's Laws

If A, B are two sets then:

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

Proof. the proof is left as an excercise!

□

1.4 Probability

The last fundamental ingeridient in modeling a random experiment, is to define a probability for each event. The probability should intuitively reflect how likely an event is probable to happen. This probability should satisfy some fundamental properties which are explained as follows.

Axiom: Axioms of probability (Kolmogorov axioms)

Suppose that $A, B \in \Omega$ is an event and \mathbb{P} is a probability function. Then \mathbb{P} should satisfy the following properties:

1. $0 \leq \mathbb{P}(A) \leq 1$
2. $\mathbb{P}(\Omega) = 1$
3. For the events $E_1, E_2, \dots, E_n \in \Omega$ that are mutually exclusive (i.e. disjoint events):

$$\mathbb{P}\left(\bigcup_i E_i\right) = \sum_i \mathbb{P}(E_i)$$

These axioms are called the fundamental axioms of probability and also the Kolmogorov axioms. We are free to define any kind of probability function that we want but it is important that 1) It should align with our common sense, 2) It should satisfy the Kolmogorov axioms.

Using the axioms above, we can observe and prove several interesting properties of the probability function. In the following box we have expressed some of them.

Theorem: Basic Properties of the Probability Function

Suppose that \mathbb{P} is a probability function and $A, B \in \Omega$ are events of the sample space Ω . We can show that the probability function has the following properties:

1. $\mathbb{P}(\emptyset) = 0$
2. If $A \subset B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$.
3. $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Proof. The properties can be proved using the basic set theory theorems.

1. Since \emptyset is the complement of Ω , so these two sets are disjoint (i.e. $\emptyset \cap \Omega = \emptyset$). On the other hand from the set theory we know that $\emptyset \cup \Omega = \Omega$. So $\mathbb{P}(\emptyset \cup \Omega) = \mathbb{P}(\Omega)$. On the other hand, using the third axiom we can write: $\mathbb{P}(\emptyset \cup \Omega) = \mathbb{P}(\emptyset) + \mathbb{P}(\Omega)$. Comparing the two recent equations we can conclude that $\mathbb{P}(\emptyset) = 0$.

The proofs for 2,3,4 are left as an exercise. However, the solutions can be found in the book "Statistical Modeling and Computation by Kroese" chapter 1.

□

Example: Defining a simple probability function

Let's define a probability function for the rolling n dice experiment that is both aligned with our common sense and also satisfy the Kolmogorov equations. Suppose that the Ω is the sample space and $E \in \Omega$ is an event. Then let's define:

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

in which the $|E|$ means the cardinality (number of elements) of the set E .

1.5 Conditional Probability and Independentness

1.6 Law of Total Probability

1.7 Baye's Rule

Example: Tossing Coin

Choose one number in $\{1, 2, \dots, 10\}$ randomly. Then toss that many dices. What is the probability that you get the sum equal to 3.

Example: Medical Test

Null Hypothesis (H_0): The person is *not* healthy (so if the results of test is positive it means the null hypothesis is accepted and the person is infected and is not healthy).

In a nutshell: positive test \rightsquigarrow person is infected!

Question: If the test is positive what is the probability that person is infected.

Example: Monty Hall Problem

Suppose that I have chose the box #1 and the Monty Hall has opened the box #2. What is the chance that the prize is in box #1 and what is the probability that the prize is in box #2.