## Contents

1 Probability Space 3

1. Sample space \( \Omega \) ......................................................... 4

2. \( \sigma \)-field \( \mathcal{F} \) .......................................................... 4

3. Probability Measure \( P \) .................................................... 6

   3.1 Measure \( \mu \) ............................................................... 6

   3.2 Probability Measure \( P \) .................................................. 7

4. Learning Objectives ............................................................. 7

5. Appendix ........................................................................ 8
Chapter 1

Probability Space

Let’s consider the experience of throwing a dart on a circular target with radius $r$ (assuming the dart always hits the target), divided in 4 different areas as illustrated in Figure 1.1.

The circles that bound the regions 1, 2, 3, and 4, have radius of, respectively, $\frac{r}{4}$, $\frac{r}{2}$, $\frac{3r}{4}$, and $r$. Therefore, the probability that a dart lands in each region is: $P(1) = \frac{1}{16}$, $P(2) = \frac{3}{16}$, $P(3) = \frac{5}{16}$, $P(4) = \frac{7}{16}$.

For this kind of problems, the theory of discrete probability spaces suffices. However, when it comes to problems involving either an infinitely repeated operation or an infinitely fine operation, this mathematical framework does not apply. This motivates the introduction of a measure-theoretic probability approach to correctly describe those cases. We define the probability space $(\Omega, \mathcal{F}, P)$, where $\Omega$ is the sample space, $\mathcal{F}$ is the event space, and $P$ is the probability
measure. Each of them will be described in the following subsections.

1 Sample space $\Omega$

The sample space $\Omega$ is the set of all the possible results or outcomes $\omega$ of an experiment or observation. For the previously described experience (hereafter referred to as $E_1$), the sample space is $\Omega_1 = \{1, 2, 3, 4\}$. This is an example of a finite sample space.

Considering a variant of $E_1$, an example of an infinitely repeated operation $E_2$ would be throwing the dart until winning the game (hit the region 1). The sample space in this case would be $\Omega_2 = \{1, 21, 31, 41, 221, 231, \ldots\}$, and this is an example of an infinite, yet countable, sample space.

An example of an infinitely fine operation $E_3$ would be throwing the dart and viewing the point of impact as the outcome. For this experience, the sample space is $\Omega_3 = (d, \alpha) : d \in [0, r], \alpha \in [0, 2\pi]$, corresponding to an uncountable sample space.

Thus, concerning their cardinality, we can classify the sample spaces as countable and finite, countable and infinite, and uncountable. A set is countable if there exists a bijective mapping from a subset of the set of natural numbers onto that set. Therefore, all finite sets are countable but not all countable sets are finite. The remaining countable sets are called infinite and countable. For instance, the set of natural numbers itself is countable, but it is not finite. An uncountable set is an infinite set which cannot be put in a one-to-one correspondence with the set of natural numbers $\mathbb{N}$, i.e., one cannot order the elements into a sequence, for example the set real numbers $\mathbb{R}$.

2 $\sigma$-field $\mathcal{F}$

Considering a completely arbitrary non-empty space $\Omega$, a class $\mathcal{F}$ of subsets of $\Omega$ is called a $\sigma$-field $^1$ if it contains the empty set, is closed under the formation of complements, and is closed under the formation of countable unions:

(i) $\emptyset \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$;

(iii) $A_i \in \mathcal{F}$, $i \in I$, $I =$ countable set $\Rightarrow \bigcup_{i \in I} A_i \in \mathcal{F}$.

Since $\Omega$ and the empty set are complementary, by (ii) the first condition can be replaced with $\Omega \in \mathcal{F}$. Being closed under complementation also implies that (iii) is equivalent to say that $\mathcal{F}$ is closed under the formation of countable intersections. This results from DeMorgan’s law, $A \cap B = (A^C \cup B^C)^C$ and $A \cup B = (A^C \cap B^C)^C$. In the more general case where $\mathcal{F}$ is closed

$^1$Some authors also use the terms sigma-field, sigma-algebra or $\sigma$-algebra.
under finite unions/intersections $\mathcal{F}$ is a field. With this result we can concluded that all $\sigma$-fields are also fields, but the opposite is not always true.

The largest $\sigma$-field containing all possible subsets of $\Omega$, including the empty set, is called power set. It contains $2^n$ elements, where $n$ is the number of elements of $\Omega$, and is usually denoted by $P(\Omega)$. Being $\mathcal{A}$ a collection of subsets of $\Omega$, there is a unique $\sigma$-field generated by $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$, which is the smallest $\sigma$-field containing $\mathcal{A}$. In cases where $\sigma(\mathcal{A})$ is more difficult to obtain than $P(\Omega)$, we can use $\mathcal{F} = P(\Omega)$, as $\sigma(\mathcal{A}) \subseteq P(\Omega)$.

**Example 1.** Suppose $\Omega = \{a, b, c\}$. Then, $P(\Omega) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{ab\}, \{bc\}, \{ac\}, \Omega\}$. We can see that $|P(\Omega)| = 2^3 = 8$ elements. Considering $\mathcal{A} = \{a, b\}$, $\sigma(\mathcal{A}) = \{\emptyset, \{a\}, \{b\}, \Omega\}$. Therefore, it is clear that $\sigma(\mathcal{A}) = P(\mathcal{A}) \subset P(\Omega)$.

An interesting curiosity is the analogy between the elements of the power set and the sequence of binary numbers, where "1" means that the corresponding element is in the subset, and "0" means the opposite, as shown in Table 1.1.

<table>
<thead>
<tr>
<th>$abc$</th>
<th>Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 000</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>1 001</td>
<td>${c}$</td>
</tr>
<tr>
<td>2 010</td>
<td>${b}$</td>
</tr>
<tr>
<td>3 011</td>
<td>${cb}$</td>
</tr>
<tr>
<td>4 100</td>
<td>${a}$</td>
</tr>
<tr>
<td>5 101</td>
<td>${ac}$</td>
</tr>
<tr>
<td>6 110</td>
<td>${ab}$</td>
</tr>
<tr>
<td>7 111</td>
<td>${abc}$</td>
</tr>
</tbody>
</table>

Table 1.1: Power set construction example

The power set is useful when we are dealing with countable sample spaces. However, when this is not the case, one needs a different way to define the event space. Thus, when the sample space is uncountable, the Borel $\sigma$-field is used.

The Borel $\sigma$-field $\mathcal{B}$ is the smallest $\sigma$-field in a topological space $\mathcal{X}$ such that every open set in $\mathcal{X}$ belongs to $\mathcal{B}$. The members of $\mathcal{B}$ are called Borel sets. If $\mathcal{X} = \mathbb{R}$, $\mathcal{B}$ can be generated by any of the intervals $(a, b), (a, b], [a, b), [a, b]$ with $-\infty \leq a \leq b \leq +\infty$, i.e.:

\[ \mathcal{B} = \sigma((a, b)) = \sigma((a, b]) = \sigma([a, b)) = \sigma([a, b]) = \sigma(\text{open subsets of } \mathbb{R}). \]

**Example 2.** Considering $\Omega = [0, 1]$, and $\mathcal{A} = [0, \frac{1}{2}), [\frac{1}{2}, 1], \sigma(\mathcal{A}) = \{\emptyset, [0, \frac{1}{2}), [\frac{1}{2}, 1], \Omega\}$. Being $\mathcal{B}([0, 1])$ the collection of all the Borel sets generated by $[0, 1]$, it is clear that $\sigma(\mathcal{A}) \subseteq \mathcal{B}([0, 1])$ and, therefore, both can be used as a valid event space $\mathcal{F}$ associated with the sample space $\Omega$. 
3 Probability Measure $P$

3.1 Measure $\mu$

A measure $\mu$ is a function defined on a $\sigma$-field $\mathcal{M}$ over a sample space $\Omega$ and taking values in $\mathbb{R}_0^+$ such that the following properties are satisfied:

(i) The empty set has measure zero: $\mu(\emptyset) = 0$.

(ii) Countable additivity or $\sigma$-additivity: if $A_1, A_2, \ldots$ is a countable collection of pairwise disjoint sets in $\mathcal{F}$, then:

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

A measure space $(\Omega, \mathcal{F}, \mu)$ is said to be complete if $A \subset B, B \in \mathcal{F}$, and $\mu(B) = 0$, together imply that $A \in \mathcal{F}$ and hence that $\mu(A) = 0$. In appendix there can be found some interesting properties and their corresponding proofs related with measure spaces.

The Borel measure, defined on the Borel $\sigma$-field $\mathcal{B}$, gives to the interval $(a, b]$, with $a < b$, the measure $b - a$. However, the Borel measure is not complete, which is why in practice the complete Lebesgue measure is preferred.

The Lebesgue measure, defined on the Lebesgue $\sigma$-field $\mathcal{L}$, is a completion of the Borel measure and $\mathcal{L}$ is such that $\mathcal{B} \subset \mathcal{L}$. Actually, $\mathcal{L}$ is obtained by adding to $\mathcal{B}$ all the subsets of sets of measure zero that belong to the power set of $\mathbb{R}$. This finds explanation in the fact that a subset of a set of measure zero might not be Borel measurable.

Therefore every Borel measurable set is also Lebesgue measurable, and both measures agree. Note that the elements of the Lebesgue $\sigma$-field are called Lebesgue measurable sets.

Formally the Lebesgue measure is constructed in the following way:

- For any half-open set $]a, b]$, with $a < b$, define its length as $L(]a, b]) = b - a$.
- For any subset of $\mathbb{R}$ define $\mu^*(B)$ as
  $$
  \mu^*(B) = \inf \{ L(E) : \forall E \supseteq B \text{ where } E \text{ is a countable union of disjoint intervals } ]a_i, b_i] \}.
  $$
- A set $A \subseteq \mathbb{R}$ is Lebesgue measurable, i.e., it belongs to the Lebesgue $\sigma$-algebra, if
  $$
  \mu^*(B) = \mu^*(B \cap A) + \mu^*(B - A)
  $$
  for all $B \subseteq \mathbb{R}$.
- The Lebesgue measure is defined by $\mu(A) = \mu^*(A)$ for all $A$ in the Lebesgue $\sigma$-algebra $\mathcal{L}$.

However, there are some sets that are not measurable. One example of such sets are the Vitali sets, named after Vitali who proved their existence. For more details on this subject please refer to [6].
3.2 Probability Measure $P$

Given a $\sigma$-field $\mathcal{F}$ of subsets of $\Omega$, any measure $P : \mathcal{F} \to [0,1]$ such that $P(\Omega) = 1$ is called a probability measure on $\mathcal{F}$. In addition, if $\mathcal{F}$ is complete with respect to $P$, we say that the triple $(\Omega, \mathcal{F}, P)$ forms a probability space.

Therefore we can define a probability measure $P : \mathcal{F} \to [0,1]$ according to the following properties:

(i) $P(\Omega) = 1$.

(ii) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ with $A_i \in \mathcal{F}, A_i \cap A_j = \emptyset, i \neq j$.

Similarly to what happens with basic probabilities, probability measures also follow the properties specified below, for $A, B \in \mathcal{F}$:

(i) $A \subseteq B \Rightarrow P(A) \leq P(B)$

(ii) $P(A) = 1 - P(A^C)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P(B) = \sum_{i \in I} P(B \cap A_i)$, being $A_i$ a partition of $\Omega$. A partition of $\Omega$ is a set $\{A_i : i \in I\}$ of pairwise disjoint events such that $\bigcup_{i} A_i = \Omega$.

4 Learning Objectives

At this point, the reader should be able to:

- Define sample spaces and identify their cardinality

- Understand what a $\sigma$-field is and be able to generate it from a sample space:
  - Definition
  - Power set
  - Borel set

- Understand the concept of measure and the different examples of measure presented

- Define probability measure and be able to apply its properties

- Identify and determine probability spaces in different contexts
5 Appendix

In the previous section, we introduced the definition of measure and its associated properties. Here, we present the proofs of some interesting results. In what follows, consider \(\mathcal{M}\) a \(\sigma\)-field of the space \(X\). Recalling from the definition of measure, the properties below must be satisfied:

(i) \(P(\Omega) = 1\).

(ii) \(P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)\) with \(A_i \in \mathcal{F}, A_i \cap A_j = 0, i \neq j\).

Result 1 \(\mu(\emptyset) = 0\)

We know that \(X, \emptyset \in \mathcal{M}\).

We also know, from the definition of measure \(\mu\), that if \(A_i \in \mathcal{M}, i = 1, \ldots, \infty\) and \(A_i \cap A_j = 0\) if \(i \neq j\) then \(\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)\).

As \(X \cap \emptyset = \emptyset\), we are in the conditions described above so \(\mu(X \cup \emptyset) = \mu(X) + \mu(\emptyset)\).

Also \(X \cup \emptyset = X\) so \(\mu(X \cup \emptyset) = \mu(X) = \mu(X) + \mu(\emptyset)\), then \(\mu(\emptyset) = 0\).

Result 2 \(\mu(A_1 \cup A_2 \cup \ldots \cup A_k) = \sum_{i=1}^{k} \mu(A_i)\) if \(A_i \cap A_j = 0, i \neq j\)

We know from the definition of measure \(\mu\) that if \(A_i \in \mathcal{M}, i = 1, \ldots, \infty\) and \(A_i \cap A_j = 0\) if \(i \neq j\) then \(\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)\).

If we have \(A = (A_1 \cup A_2 \cup \ldots \cup A_k)\) and \(A_i \cap A_j = 0\) when \(i \neq j\) its the same as \(A = \bigcup_{i=1}^{\infty} A_i, i = 1, \ldots, \infty\) as long as \(A_i = \emptyset, i = k + 1, \ldots, \infty\).

Then, if \(A_i = \emptyset, i = k + 1, \ldots, \infty\), \(\mu(A_1 \cup A_2 \cup \ldots \cup A_k) = \sum_{i=1}^{k} \mu(A_i)\) if \(A_i \cap A_j = 0\) when \(i \neq j\).

Result 3 \(A \subset B \Rightarrow \mu(A) \leq \mu(B), A, B \in \mathcal{M}\),

\(A, B \in \mathcal{M}\) and \(A \subset B\).

\(A \cap (B \setminus A) = \emptyset\) and \(A \cup (B \setminus A) = B\) as \(A \subset B\).

Then \(\mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) = \mu(B)\) so \(\mu(A) \leq \mu(B)\).

Result 4 \(A = \bigcup_{i=1}^{\infty} A_i, A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots \Rightarrow \lim_{n \to \infty} \mu(A_i) = \mu(A)\)

If we have \(A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots\), the union of those sets is equal to the last set, so \(A = \bigcup_{i=1}^{\infty} A_i = A_\infty\).

Then \(\lim_{n \to \infty} \mu(A_i) = \mu(A_\infty) = \mu(A)\).

Result 5 \(A = \bigcap_{i=1}^{\infty} A_i, A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots \Rightarrow \lim_{n \to \infty} \mu(A_i) = \mu(A)\)

If we have \(A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots\), the intersection of those sets is equal to the last set so \(A = \bigcap_{i=1}^{\infty} A_i = A_\infty\).

Then \(\lim_{n \to \infty} \mu(A_i) = \mu(A_\infty) = \mu(A)\).
Bibliography


