Probability Triples

Yicong Lin

February, 2020

Probability is a term that seems straightforward; people use it all the time. You may wonder the reason of learning the concepts such as σ-field, the axioms of probability, etc., in the first week. Although modern probability theory is abstract, the fundamental ideas are not very different from the Analysis courses. Consider the real function

$$f(x) = \frac{1}{x},$$

which obviously is not defined at $x = 0$. Therefore, the domain of the function is $D_f = \mathbb{R}\backslash\{0\}$. **A probability function $\mathbb{P}$ is essentially a function with a σ-field as its domain**. Just as ruling out the point $x = 0$ at which the function f cannot be defined, σ-field excludes the event sets to which probabilities cannot be assigned. These exceptional sets are also called non-measurable sets; see "Extension: existence of non-measurable subsets of $[0, 1]$" in course book. The main source of difficulty is the (uncountably) infiniteness .

The capability of dealing with "infiniteness" sets the modern probability theory apart from the classical one. More specifically, the classical probability is established based on outcomes which are defined as the elements of a sample space S; see pages 21 - 27 in lecture slides. Recall that subsets of S are called events. That is, for an event set $A = \bigcup_{i=1}^{n}\{a_i\} \subseteq S$, the classical probability of A is defined as $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(\{a_i\})$, where $a_i \in A$ is an outcome, and $n \in \mathbb{N}$. However, we are in trouble when there are uncountably [1] many outcomes. We cannot assign a probability to an uncountable set A by simply adding the probabilities of every outcome; see the example on page 11 in lecture slides.

For this reason, we start thinking whether it is possible to assign a probability on the event $A \subseteq S$ *directly*. Unfortunately, we *cannot* define probabilities on **all** the subsets; see the conclusion on page 16 in lecture slides. This is similar to being unable to define the function $f(x) = 1/x$ on every point in $\mathbb{R}$. To define probabilities coherently, in the first step, it is natural to eliminate the troublesome/non-measurable subsets. Since it is extremely difficult to search out the subsets that are strange, we thereby consider which subsets (events) are interesting in real life. We try to find out the subsets, which can be assigned with probabilities, as many as possible. For convenient, we will collect all these events into a σ-field (or called σ-algebra) $\mathcal{A}$. Define A and B as two subsets of S. Before continuing, I should mention that we interpret $A \cup B$ as at least A or B happens, similarly, $A \cap B$ as A and B happen simultaneously. Now let's consider which sets should be included in $\mathcal{A}$. We would like to study the impossible event $\emptyset$ and the certain event S.

[1] Read 'Extension: countability of infinite sets' in course book. For example, the set of natural numbers $\mathbb{N}$ is a countable set. The real number space $\mathbb{R}$ is an uncountable set. Both countable and uncountable sets have infinite sizes. However, we say the size of $\mathbb{R}$ is larger than $\mathbb{N}$.

Thus $\emptyset$ and S are put into $\mathcal{A}$. If an event A happens, i.e. $A \in \mathcal{A}$, we also want to study the case when A does not happen, i.e. $A^c \in \mathcal{A}$. If A and B happen, i.e. $A \in \mathcal{A}$ and $B \in \mathcal{A}$, it is natural to consider $A \cup B \in \mathcal{A}$ and $A \cap B \in \mathcal{A}$ as well.

For a modern probability theorist, it is important to take (countably/uncountably) infiniteness into considerations. However, the most we can do is to require $\mathcal{A}$ to be closed under countable operations. That is, if $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ and $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$. What is the reason why it cannot be closed under uncountable operations? It is because we will get contradictory results. We note that any subset of S (event) can be always written as a set of outcomes that are finitely/countably/uncountably many. For example, a subset $(0, 1/2) \subseteq S = [0, 1]$ can be written as $(0, 1/2) = \bigcup_{x \in (0,1/2)} \{x\}$, where $\{x\} \subseteq S$. If $\mathcal{A}$ is closed under uncountable operations, then all the subsets of S are automatically included in $\mathcal{A}$. However, as mentioned, it is impossible to define probabilities on all of the events/subsets. For this reason, $\mathcal{A}$ cannot take uncountable unions and intersections into consideration.

Fortunately, if $\mathcal{A}$ satisfies all the requirements above, it is general enough to cover almost all the interesting events in real life. It can be also shown that a probability function $\mathbb{P}$ is well-defined on $\mathcal{A}$. There is some redundancy above. For example, if $\mathcal{A}$ is closed under complement and countable unions, it is automatically closed under countable intersections by de Morgan's laws; see page 14 in lecture slides. Formally, we call $(S, \mathcal{A}, \mathbb{P})$ is a probability triple, where

- i S is any non-empty set of possible outcomes;
- ii $\mathcal{A}$ is a σ-field (σ-algebra); see page 17 in lecture slides. It is a collection of events satisfying the following conditions:
  - (a) $S \in \mathcal{A}$;
  - (b) if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$;
  - (c) if $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$;
- iii $\mathbb{P} : \mathcal{A} \to [0, 1]$ is a mapping, see page 20 in lecture slides, satisfying the Axioms of Probability:
  - (a) $\mathbb{P}(S) = 1$;
  - (b) $\mathbb{P}(A) \geq 0$, $A \in \mathcal{A}$;
  - (c) countable additivity: if $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, are disjoint, then $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

There are some remarks: first, the symbol σ is usually considered as a countable union. $\mathcal{A}$ has some algebraic structure in the sense that it is closed under countable operations. Therefore, we call $\mathcal{A}$ a σ-algebra. Second, since $S \in \mathcal{A}$ and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, where $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, the probability function $\mathbb{P}$ makes sense for Points 3(a) and 3(c) above. Third, the Axioms of Probability adhere to our intuitions. **Any function $\mathbb{P}$ that satisfies these conditions is called a probability function.** Both the frequentist and Bayesian interpretations of probability are sufficient to these conditions. Finally, we again do not extend 3(c) to uncountable additivity. Otherwise, the probability can be simply defined on outcomes, for instance, $\mathbb{P}([0, 1]) = \sum_{x \in [0,1]} \mathbb{P}(\{x\})$.