Mathematics

Edward Wang

2025

Contents

Part I.

Algebra

Part II.

Analysis

Part III.

Probability

4. Probability

4.1. Basics

Definition 4.1.1. The *sample space*, denoted Ω is the set of all possible outcomes of an experiment.

Each element $\omega \in \Omega$ is called an *outcome*. An *event* E is a subset of Ω, and is hence a set of possible outcomes. Thus Ω itself is an event, the *certain event*, and so is the empty set $\varnothing$, the *impossible event*. An event E *occurs* if the outcome ω is in E.

Definition 4.1.2. If Ω is countable, then we say that it is a *discrete* sample space. Otherwise, if Ω is uncountable, we say that it is a *continuous* sample space.

Since events are sets, we can use the usual properties of sets. In particular, if A and B are events, then $A \cup B$ is an event (that either A or B occur) and $A \cap B$ is an event (that both A and B occur).

Proposition 4.1.1 (Set theory properties)**.** *Let A, B, C be subsets of Ω.*

1. *We have $A \cup \varnothing = A$ and $A \cap \varnothing = \varnothing$.*
2. *We have $A \cup \Omega = \Omega$ and $A \cap \Omega = A$.*
3. *We have $A \cup (\Omega \setminus A) = \Omega$ and $A \cap (\Omega \setminus A) = \varnothing$. In other words, $A \cup A^c = \Omega$ and $A \cap A^c = \varnothing$.*
4. *(Commutativity) We have $A \cup B = B \cup A$ and $A \cap B = B \cap A$.*
5. *(Associativity) We have $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$.*
6. *(Distributivity) We have $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.*
7. *(De Morgan's laws) We have $\Omega \setminus (A \cup B) = (\Omega \setminus A) \cap (\Omega \setminus B)$ and $\Omega \setminus (A \cap B) = (\Omega \setminus A) \cup (\Omega \setminus B)$. In other words, $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$.*

Definition 4.1.3. Two events A and B are *mutually exclusive* or *disjoint* if $A \cap B = \varnothing$. A collection of events $\{E_i\}_{i \in I}$ is mutually exclusive if $E_i \cap E_j = \varnothing$ for all $i \neq j$.

Definition 4.1.4. Let Ω be a sample space and let 2^Ω denote the power set of Ω. Then $\mathbb{P} \colon 2^\Omega \to [0,1]$ is a probability function that assigns to each event a probability. This function $\mathbb{P}$ must satisfy the following axioms formulated by Kolmogorov:

1. For all events E, we have $\mathbb{P}(E) \geq 0$.
2. The probability of the sample space is 1. That is, $\mathbb{P}(\Omega) = 1$.
3. (Countable additivity) Let $E_1, E_2, \ldots$ be a countable sequence of mutually exclusive (disjoint) events. Then,
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

Remark 4.1.1. This differs from the definition of probability space one will encounter later on in a more advanced, measure theoretic course. The more formal definition of a probability space is something like the following: a probability space is a tuple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F} \subseteq 2^\Omega$ is a σ-algebra, and $\mathbb{P}$ is a probability measure $\mathbb{P} \colon \mathcal{F} \to [0,1]$ satisfying $\mathbb{P}(\Omega) = 1$ and countable additivity.

In particular, this our definition does not exactly work for uncountable sample spaces. The restriction for $\mathcal{F}$ to be a σ-algebra is exactly what is needed to 'fix' this issue.

Proposition 4.1.2. *Properties of the probability function:*

1. *We have* $\mathbb{P}(\varnothing) = 0$.
2. *We have* $\mathbb{P}(A^c) = \mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$.
3. *We have* $\mathbb{P}(A) \leq 1$.
4. *We have* $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
5. *If* $A \subseteq B$, *then* $\mathbb{P}(A) \leq \mathbb{P}(B)$.
6. *(Finite additivity) If* $E_1, E_2, \ldots, E_n$ *are mutually exclusive events, then*
$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mathbb{P}(E_i)$$

7. *If* $A_1 \subseteq A_2 \subseteq \ldots$ *and* $B = \bigcup_{i=1}^{\infty} A_i$*, then* $\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(A_n)$.

8. *If* $A_1 \supseteq A_2 \supseteq \ldots$ *and* $B = \bigcap_{i=1}^{\infty} A_i$*, then* $\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(A_n)$.

Proof. We will prove (1), (2), (4), (6), and (7), and leave the rest as exercises.

1. This is a corollary of (2). Since $\varnothing = \Omega^c$, we have $\mathbb{P}(\varnothing) = \mathbb{P}(\Omega^c) = 1 - \mathbb{P}(\Omega) = 1 - 1 = 0$.

2. This is a corollary of finite additivity (6), with $n = 2$. Consider the disjoint events A and A^c. Since $A \cup A^c = \Omega$, we have $1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$, from which the result follows.

3. Exercise.

4. Again, this follows from finite additivity applied to the disjoint sets A and $B \setminus A$. Since $A \cup (B \setminus A) = A \cup B$, we have

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A \cup (B \setminus A)) \\ &= \mathbb{P}(A) + \mathbb{P}(B \setminus A).\end{aligned}$$

Notice that $B = (B \setminus A) \cup (A \cup B)$ is a disjoint union, from which we get $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$, so $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$. This, combined with the previous equation, yields

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(A \cap B).$$

5. Exercise.

6. Consider the mutually exclusive events $E_1, \ldots, E_n$. Let $E_{n+1} = E_{n+2} = \cdots = \varnothing$. Notice that $E_1, \ldots$ still form a mutually exclusive collection of events, so countable additivity applies. Thus

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) &= \sum_{i=1}^{\infty} \mathbb{P}(E_i) \\ &= \sum_{i=1}^{n} \mathbb{P}(E_i) + \sum_{i=n+1}^{\infty} \mathbb{P}(E_i) \\ &= \sum_{i=1}^{n} \mathbb{P}(E_i) + 0,\end{aligned}$$

since $\mathbb{P}(E_i) = \mathbb{P}(\varnothing) = 0$ for all $i > n$. Hence finite additivity holds.

7. Define $B_i = A_i \setminus A_{i-1}$. Clearly all of B_i are mutually exclusive, and moreover, $B = \bigcup_{i=1}^{n} B_i$. Hence by countable additivity,

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(B_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(B_i) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} B_i\right),\end{aligned}$$

   with the last step being due to finite additivity. But $\bigcup_{i=1}^{n} B_i = A_n$, so we have

$$\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

   as required.

8. Exercise. □

In a discrete sample space, all outcomes are disjoint. (This is obviously true, due to the fact that there can be only one outcome of an experiment!) Hence for any event E, we have that

$$\mathbb{P}(E) = \sum_{\omega\in E} \mathbb{P}(\omega).$$

Remark 4.1.2. It is important to note here that the statements $\mathbb{P}(E) = 0 \implies E = \varnothing$ and $\mathbb{P}(E) = 1 \implies E = \Omega$ are **not necessarily true**. For example, let $\Omega = \{H, T\}$ and define $\mathbb{P}(H) = 0$ and $\mathbb{P}(T) = 1$. Notice that $H \neq \varnothing$ and $T \neq \Omega$.

Definition 4.1.5. We say that a collection of events $\{E_i\}_{i\in I}$ is *exhaustive* if

$$\bigcup_{i\in I} E_i = \Omega.$$

In particular, note that A and A^c are exhaustive, and also disjoint.

Definition 4.1.6. We define the *conditional probability* of event A *given* event B as

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)},$$

given that $\mathbb{P}(B) > 0$.