

**PROBABILITY THEORY:
ELEMENTARY RESULTS**

Note: *This work is intended for informative and educational purposes only.*

1. Introduction

The field of Probability Theory is fundamental to the way we live our lives. If you open up the weather app on your phone you might see the phrase *chance of rain*, which is a probability calculated by meteorologists of the chance of it raining within a given time period. Simply put, weather systems are far too complex to know with absolute certainty how things will evolve.

Probability appears in many other traditional areas too. Gambling is perhaps the most ‘in your face’ example. If you’ve ever played roulette you’ll know there are red numbers, black numbers, and the all important green zero (or if you’re reading this in the U.S., there is also the green double zero). These clever additions ensure that there are no strategies that will earn you expected winnings greater than or equal to zero—after all, *the house always wins*.

In this issue I want to build up and present some elementary results from Probability Theory. I would imagine most of you will have covered a large proportion, if not all, of these results before, possibly at school. But as easy as it is to forget them, they are undoubtedly important.

2. The Basics of Set Theory

We begin with a couple of results from Set Theory. When we consider traditional probability examples such as dice rolling and coin flipping we can naturally arrange the outcomes into ‘sets’. In what follows suppose we have two sets, A and B .

- **Union** (\cup)

The union symbol refers to the elements that are in A or B .

- **Intersection** (\cap)

The intersection symbol refers to the elements that are in A and B .

- **Empty or Null Set** (\emptyset)

The empty set contains no elements.

- **Universe or Sample Space** (Ω)

The universe contains all elements.

- **Complement** (c or $'$)

The complement is the negation of its argument, for instance A^c is all elements not in A .

These definitions allow us to introduce some important results known as De Morgan's Laws.

$$(A \cup B)^c = A^c \cap B^c \quad (1)$$

$$(A \cap B)^c = A^c \cup B^c. \quad (2)$$

3. The Foundations of Probability Theory

We now build up some fundamental concepts to establish Probability Theory.

3.1 σ -Algebra

A sigma algebra, denoted \mathcal{F} has the following three properties:

1. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$;
2. $A_1, A_2, \dots, A_n \in \mathcal{F} \implies \bigcup_{i=1}^n A_i \in \mathcal{F}$;
3. $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_i A_i \in \mathcal{F}$.

The difference between points 2 and 3 is that the former states that a *finite* union is an element of the σ -algebra, while the latter states that a *countably infinite* union is an element of the σ -algebra. In fact, the first two points are properties of more general Algebra's while the third is specific to σ -algebra's. A more detailed description of the differences is beyond the scope of this issue.

3.2 General Measure

Combining a universe Ω and a σ -algebra \mathcal{F} we get something called a *measurable space*, denoted (Ω, \mathcal{F}) . This can be extended to include a general *measure* $\mu : \mathcal{F} \rightarrow \mathbb{R}$ which has the following properties.

1. Non-negativity

For all $A \in \mathcal{F}$, we have $\mu(A) \geq 0$;

2. Null Measure

$\mu(\emptyset) = 0$;

3. Countable Additivity

For $A_1, A_2, \dots \in \mathcal{F}$, with $i \neq j$ such that $\mathbb{P}(A_i \cap A_j) = \emptyset$, we have that $\mu(\sum_i A_i) = \sum_i \mu(A_i)$.

Including the measure into our measurable space, gives us the triple $(\Omega, \mathcal{F}, \mu)$ referred to as a *measure space*.

3.3 Probability Measure

It was Kolmogorov who took the ideas presented in measure theory and devised a the fundamental axioms of probability, which are now affectionately known as Kolmogorov's axioms.

Given a sample space Ω , σ -algebra \mathcal{F} , and probability measure \mathbb{P} , we have the triple $(\Omega, \mathcal{F}, \mathbb{P})$. The axioms are:

1. **Non-negativity**

For all $A \in \mathcal{F}$, we have $\mathbb{P}(A) \geq 0$;

2. **Normalisation**

$\mathbb{P}(\Omega) = 1$;

3. **Countable Additivity**

For $A_1, A_2, \dots \in \mathcal{F}$, with $i \neq j$ such that $\mathbb{P}(A_i \cap A_j) = \emptyset$, we have that

$\mathbb{P}(\sum_i A_i) = \sum_i \mathbb{P}(A_i)$.

4. Elementary Results

Now that we have established the basis of Probability Theory we can move on to consider some elementary results. In what follows we will now assume we have two events A and B . Where there are elements which have a subscript, this means we are considering multiple events.

A Single Event

If we have just one event A , then naturally it follows that

$$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1. \quad (3)$$

Disjoint Events

If two events are disjoint, then we say that $\mathbb{P}(A \cap B) = \emptyset$. As a result, we have that

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B). \quad (4)$$

General Union

If two events are not disjoint then we have

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \quad (5)$$

One Event Given Another

We denote $\mathbb{P}(A|B)$ to be the probability of event A happening given that event B has occurred. In this case we have

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad (6)$$

General Intersection

The intersection can be written in either order, by which we mean $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$. We can rearrange the previous result to obtain

$$\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B). \quad (7)$$

Law of Total Probability

Suppose that we split up our universe Ω into partitions such that $\sum_i A_i = \Omega$, with $\mathbb{P}(A_i \cap A_j) = \emptyset$ for $i \neq j$. Then we have

$$\mathbb{P}(B) = \sum_i \mathbb{P}(B|A_i)\mathbb{P}(A_i). \quad (8)$$

Bayes' Theorem

From the General Intersection formula we can exploit the fact that the ordering of A and B are equivalent to obtain the famous *Bayes' Theorem*.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}. \quad (9)$$

Extended Bayes' Theorem

Combing the last two results we can obtain the so called *Extended Bayes' Theorem*.

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_i \mathbb{P}(B|A_i)\mathbb{P}(A_i)} \quad (10)$$

5. Discussion

In this issue we have explored the very building blocks of Probability Theory. Of course, this barely scratches the surface of the world of Probability and the reader will most likely have come across several more advanced features, including the concept of random variables and their extension to stochastic processes.

For those wishing to refresh their memory or perhaps just dig a little deeper into the subject I recommend the books *Probability and Random Processes* by Grimmett and Stirzaker, and *Probability Theory: The Logic of Science* by Jaynes. The latter provides some excellent applications to the real world. If you are feeling particularly adventurous, there is another book *1000 Exercises in Probability* also by Grimmett and Stirzaker.