Probability Theory

Author: Isaac FEI

Preface

Probability theory is a fundamental branch of mathematics with applications in numerous fields, including statistics, computer science, physics, and engineering. While the subject has many practical applications, it is also a beautiful and elegant area of mathematics in its own right. This book is designed to offer readers a thorough and rigorous understanding of the concepts and principles that underlie probability theory.

To that end, this book begins with a detailed treatment of measure theory, which provides the mathematical framework for probability theory. I mainly refer to [1] for this part. From there, the book explores the concepts of random variables and distribution functions, and their properties.

Contents

Chapter 1 Measure Theory

1.1 Classes of Sets

Definition 1.1.1

*A collection of subsets of Ω, $\mathcal{S}$, is a **semi-algebra** if it satisfies the following properties:*

1. $\Omega \in \mathcal{S}$
2. $A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$
3. *The complement of every set in $\mathcal{S}$ can be written as a finite disjoint union of sets in $\mathcal{S}$. Formally, $\forall A \in \mathcal{S},\ \exists E_1, \ldots, E_n \in \mathcal{S}$ such that $A^{\complement} = \biguplus_{i=1}^{n} E_i$*

♣

Example 1.1 Let $\Omega = \mathbb{R}$. Suppose $\mathcal{S}$ is a collection of subsets in $\mathbb{R}$ consisting of intervals of the following forms:

1. $(a, b]$
2. $(-\infty, b]$
3. (a, ∞)
4. $(-\infty, \infty) = \mathbb{R}$

One may check that $\mathcal{S}$ is indeed a semi-algebra by definition. And it is by studying this type of collection of sets that the definition of semi-algebra arises.

Definition 1.1.2

*A collection of subsets of Ω, $\mathcal{A}$, is called an **algebra** if*

1. $\Omega \in \mathcal{A}$
2. $A, B \in \mathcal{A} \implies A \cap B \in \mathcal{A}$
3. $A \in \mathcal{A} \implies A^{\complement} \in \mathcal{A}$

♣

Proposition 1.1.1

Let $\{\mathcal{A}_\alpha\}_{\alpha \in \Lambda}$ be a collection of algebras in Ω where Λ is any index set, then

$$\bigcap_{\alpha \in \Lambda} \mathcal{A}_\alpha$$

is also an algebra.

♠

Proof ■

Definition 1.1.3

*A collection of subsets of Ω, $\mathcal{F}$, is called a σ-**algebra** if*

1. $\Omega \in \mathcal{F}$
2. *$\mathcal{F}$ is closed under countable intersections, i.e,. $A_n \in \mathcal{F}\ \forall n \in \mathbb{N}^* \implies \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$*
3. $A \in \mathcal{F} \implies A^{\complement} \in \mathcal{F}$

♣

Of course, a σ-algebra is also closed under finite intersections (and unions) since we can always write

$$\bigcap_{n=1}^{m} A_n = \bigcap_{n=1}^{m} A_n \cap \bigcap_{n=m+1}^{\infty} \Omega = \bigcap_{n=1}^{\infty} A_n \quad \text{where } A_n = \Omega \text{ for } n \geq m+1$$

Therefore, the σ-algebra is of course an algebra.

Analogous to Proposition 1.1.1, any intersections of σ-algebras is also itself a σ-algebra.

Proposition 1.1.2

Let $\{\mathcal{F}_\alpha\}_{\alpha\in\Lambda}$ be a collection of σ-algebras in Ω where Λ is any index set, then

$$\bigcap_{\alpha\in\Lambda} \mathcal{F}_\alpha$$

is also a σ-algebra.

Having proved Proposition 1.1.1 and Proposition 1.1.2, we may introduce the algebra and σ-algebra generated by a class of subsets in Ω.

Definition 1.1.4

Let $\mathcal{C}$ be a class of subsets in Ω. The ***algebra generated by*** *$\mathcal{C}$, is defined by*

$$\mathcal{A}(\mathcal{C}) := \bigcap_{\mathcal{A} \text{ is an algebra, } \mathcal{A}\supseteq\mathcal{C}} \mathcal{A}$$

That is, it is the intersection of all algebras containing $\mathcal{C}$. Similarly, the ***σ-algebra generated by*** *$\mathcal{C}$ is defined by*

$$\sigma(\mathcal{C}) := \bigcap_{\mathcal{F} \text{ is a } \sigma\text{-algebra, } \mathcal{F}\supseteq\mathcal{C}} \mathcal{F}$$

Note that the set

$$\{\mathcal{A} \subset \mathcal{P}(\Omega) \mid \mathcal{A} \text{ is an algebra, } \mathcal{A} \supseteq \mathcal{C}\}$$

is clearly nonempty since the power set, $\mathcal{P}(\Omega)$, itself is one of its element. And then by Proposition 1.1.1, $\mathcal{A}(\mathcal{C})$ is indeed well-defined, and so is $\sigma(\mathcal{C})$.

One important property for $\mathcal{A}(\mathcal{C})$ (resp. $\sigma(\mathcal{C})$) is that it is the *smallest* algebra (resp. σ-algebra) containing $\mathcal{C}$, as stated in the following proposition.

Proposition 1.1.3

Let $\mathcal{C} \subseteq \mathcal{P}(\Omega)$. We have the following:

1. *If $\mathcal{A}$ is an algebra containing $\mathcal{C}$, then $\mathcal{A} \supseteq \mathcal{A}(\mathcal{C})$.*
2. *If $\mathcal{F}$ is a σ-algebra containing $\mathcal{C}$, then $\mathcal{F} \supseteq \sigma(\mathcal{C})$.*

As we see in Example 1.1, it is easy to describe the elements in a semi-algebra directly in terms of subsets of Ω. However, it is difficult in general to describe the elements in an algebra, and it is even more difficult to do the same for the σ-algebra.

It is because there is no explicit form of describing a σ-algebra in terms of subsets in Ω that makes the extension of a measure (or any properties) from the semi-algebra to σ-algebra extremely difficult. As we shall see in many proofs, we will exploit the property of a σ-algebra in an *inductive* way rather than a *constructive* one.

Fortunately, we do have an explicit form for the algebra generated by a semi-algebra $\mathcal{S}$, $\mathcal{A}(\mathcal{S})$, in terms of sets in $\mathcal{S}$. It turns out that it is simply the collection of all finite disjoint unions of sets in $\mathcal{S}$.

Theorem 1.1.1

Let $\mathcal{S}$ be a semi-algebra on Ω. Then

$$A \in \mathcal{A}(\mathcal{S}) \iff \exists E_1, \ldots, E_n \in \mathcal{S},\ A = \biguplus_{i=1}^{n} E_i \tag{1.1}$$

Note *Note that the integer n in* (1.1) *is not fixed. That is the choice of n may vary for different set A.*

Proof (Proof of $\Longleftarrow$) The reverse direction of (1.1) is easy to prove. Suppose $A = \biguplus_{i=1}^{n} E_i$ where each $E_i \in \mathcal{S}$. Then clearly $A \in \mathcal{A}(\mathcal{S})$ since the algebra is closed under finite unions.

(Proof of $\Longrightarrow$) The proof of the forward direction is more interesting.

Note *Note that given $A \in \mathcal{A}(\mathcal{S})$, it is tempting and yet difficult to find a representation of A as the right-hand side of* (1.1)*. Hence, we shall prove this in an indirect way. Instead of finding a specific representation of A, we shall look at all representations described by* (1.1)*. That is, we consider the collection, say $\mathcal{C}$, of all finite disjoint unions of sets in $\mathcal{S}$. And then we show that $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{C}$.*

Define

$$\mathcal{C} = \left\{ E \subseteq \Omega \,\middle|\, \exists E_1, \ldots, E_n \in \mathcal{S},\ E = \biguplus_{i=1}^{n} E_i \right\}$$

The goal is to show that this collection $\mathcal{C}$ is actually an algebra containing $\mathcal{S}$.

It is clear that $\mathcal{C} \supseteq \mathcal{S}$ by the definition of $\mathcal{C}$. What remains to show is that $\mathcal{C}$ is an algebra. We need to check whether it satisfies each of the three properties one after another.

First, note that $\Omega \in \mathcal{C}$ since $\Omega \in \mathcal{S}$ and we have seen that $\mathcal{C} \supseteq \mathcal{S}$.

Now, we show $\mathcal{C}$ is closed under intersections. Let $E, F \in \mathcal{C}$. By definition, E and F can be written as

$$E = \biguplus_{i=1}^{n} E_i \quad \text{and} \quad F = \biguplus_{j=1}^{m} F_j$$

where $E_i, F_j \in \mathcal{S}$. It then follows that

$$E \cap F = \biguplus_{i=1}^{n} \left(E_i \cap \biguplus_{j=1}^{m} F_j \right) = \biguplus_{i=1}^{n} \biguplus_{j=1}^{m} (E_i \cap F_j)$$

Note that $E_i \cap F_j \in \mathcal{S}$. Hence, we have $E \cap F \in \mathcal{C}$ since the above equation says $E \cap F$ is a finite disjoint union of sets, $(E_i \cap F_j)$'s, in $\mathcal{S}$.

We also need to show $\mathcal{C}$ is closed under complements. Let $E \in \mathcal{C}$. The set E can be written as

$$E = \biguplus_{i=1}^{n} E_i$$

where $E_i \in \mathcal{S}$. Then, the complement of E,

$$E^{\complement} = \bigcap_{i=1}^{n} E_i^{\complement}$$

By the property of semi-algebra, each $E_i^{\complement}$ can be written as

$$E_i^{\complement} = \biguplus_{j=1}^{m_i} F_{i,j}$$

where $F_{i,j} \in \mathcal{S}$. Therefore, $E_i^{\complement} \in \mathcal{C}$ by the definition of $\mathcal{C}$. And then $E^{\complement} \in \mathcal{C}$ since we have already shown that $\mathcal{C}$ is closed under finite intersections.

In summary, we have proved so far that $\mathcal{C}$ is indeed an algebra containing $\mathcal{S}$. It then follows that $\mathcal{A}(\mathcal{S}) \subseteq \mathcal{C}$ since $\mathcal{A}(\mathcal{S})$ is the smallest algebra containing $\mathcal{S}$. Therefore, for each $A \in \mathcal{A}(\mathcal{S})$, A can always be written as a finite disjoint union of sets in $\mathcal{S}$ by the definition of $\mathcal{C}$. ■

1.2 Set Functions

A **set function** is a function defined on collection of sets whose range is usually the union of non-negative real numbers and the positive infinity, $\mathbb{R}_{\geq 0} \cup \{\infty\}$.

Definition 1.2.1

*A set function, μ, defined on $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ is said to be **additive** or **finitely additive** if*

1. $\mu(\emptyset) = 0$, *and*
2. $\mu\left(\biguplus_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$.

♣

By the second condition in above definition, the natural class of sets to consider is the algebra since we would want that the class is closed under finite unions of sets.

Note *The condition $\mu(\emptyset) = 0$ is actually redundant in all cases except the one where $\mu(A) = \infty$ for all $A \in \mathcal{C}$. To see this, suppose that $\mu(A) < \infty$ for some $A \in \mathcal{C}$. We have $A = A \uplus \emptyset$. It then follows from the second condition that $\mu(A) = \mu(A) + \mu(\emptyset)$, which further implies $\mu(\emptyset) = 0$ since $\mu(A) < \infty$.*

If $A \subseteq B$, then it is natural to expect that $\mu(A) \leq \mu(B)$. It is indeed the case. But the proof is not completely trivial. We can always write $B = A \uplus (B \setminus A)$. Then, we have

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$$

Note that the above inequality holds no matter whether $\mu(B \setminus A)$ is finite or not.

Definition 1.2.2

*A set function, μ, defined on $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ is said to be **σ-additive** if*

1. $\mu(\emptyset) = 0$, *and*
2. $\mu\left(\biguplus_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$

♣

Example 1.2 Let $\Omega = (0, 1)$ and

$$\mathcal{C} = \{(a, b] \mid 0 \leq a < b < 1\}$$

Consider the set function μ on $\mathcal{C}$ given by

$$\mu(a, b] = \begin{cases} \infty, & a = 0 \\ b - a, & a > 0 \end{cases}$$

We claim that μ is additive on $\mathcal{C}$. Suppose that

$$(a,b] = \biguplus_{i=1}^{n} (a_i, b_i]$$

There are two cases.

(Case 1: $a = 0$) If $a = 0$, then $\mu(a,b] = \infty$. And there exists i such that $a_i = 0$, which implies $\mu(a_i, b_i] = \infty$. In this case, we have

$$\infty = \mu(a,b] = \sum_{i=1}^{n} \mu(a_i, b_i] = \infty$$

(Case 2: $a > 0$) Without loss of generality, we may assume $a_i < a_{i+1}$ for all i. Then it is clear that $a_1 = a$, $b_n = b$, and

$$b_i = a_{i+1} \quad \forall i = 1, \dots, n-1$$

It follows that

$$\sum_{i=1}^{n} \mu(a_i, b_i] = \sum_{i=1}^{n} b_i - a_1 = b - a = \mu(a,b]$$

Therefore, we conclude that μ is indeed additive. However, μ is not σ-additive. To see this, we note that

$$(0, \frac{1}{2}] = \biguplus_{i=2}^{\infty} (\frac{1}{i+1}, \frac{1}{i}]$$

We have

$$\sum_{i=2}^{\infty} \mu(\frac{1}{i+1}, \frac{1}{i}] = \sum_{i=2}^{\infty} \left(-\frac{1}{i+1} + \frac{1}{i} \right) = \lim_{n \to \infty} \left(\frac{1}{2} - \frac{1}{n+1} \right) = \frac{1}{2}$$

whereas

$$\mu(0, \frac{1}{2}] = \infty$$

Therefore, μ is not σ-additive.

Definition 1.2.3

Let μ be a set function defined on $\mathcal{C} \subseteq \mathcal{P}(\Omega)$. Then

1. *μ is said to be **continuous from below** at set $E \in \mathcal{C}$ if*

$$E_n \in \mathcal{C}, \; E_n \uparrow E \implies \mu(E_n) \uparrow \mu(E)$$

2. *μ is said to be **continuous from above** at set $E \in \mathcal{C}$ if*

$$E_n \in \mathcal{C}, \; E_n \downarrow E \quad \textit{and} \quad \exists n_0 \in \mathbb{N}^*, \; \mu(E_{n_0}) < \infty \implies \mu(E_n) \downarrow \mu(E)$$

♣

The symbol $E_n \uparrow E$ means that

$$E_n \subseteq E_{n+1} \quad \forall n \in \mathbb{N}^* \quad \text{and} \quad \bigcup_{n=1}^{\infty} E_n = E$$

And by writing $\mu(E_n) \uparrow \mu(E)$, we mean that

1. $\{\mu(E_n)\}$ is an increasing sequence (of extended real numbers), and
2. $\lim_{n \to \infty} \mu(E_n) = \mu(E)$.

In fact, in the above definition, we only need to assume the limit of $\mu(E_n)$ is $\mu(E)$ since we have already noticed that

$$A \subseteq B \implies \mu(A) \leq \mu(B)$$

Now, we pay attention to the condition in the second statement of the above definition, i.e., the definition of the continuity from above. Note that comparing to the continuity from below, we impose an additional requirement that $\mu(E_{n_0})$ is finite for some index n_0. The reason why we do this is explained through the following example.

Example 1.3 Let $\Omega = \mathbb{R}$ and $E_n = [n, \infty)$. We would want the length of E_n, $\lambda(E_n)$, to be infinity. Actually, λ is a measure, called Lebesgue measure, on the Borel σ-algebra of the real line, which is a generalization of the concept of length.

Note that $E_n \downarrow \emptyset$. We have $\lambda(E_n) = \infty$ whereas $\lambda(\emptyset) = 0$. In this case, $\lambda(E_n)$ does not tend to $\lambda(\emptyset)$ as $n \to \infty$. If we do not assume $\lambda(E_n)$ is finite for some $n = n_0$, then we cannot say λ is continuous from above at $\emptyset$ in this case. However, we really want that the measure to be continuous from above at all sets. Therefore, it is necessary to impose an additional condition as in Definition 1.2.3.

1.3 Extension of Measures

Theorem 1.3.1 (Carathéodory's Extension Theorem)

Chapter 2 Sets and Events

2.1 Probability Spaces

2.2 Distributions

References

[1] S. J. Taylor and J. F. C. Kingman. *Introduction to Measure and Integration*. Cambridge [Eng.]: University Press, 1973. ISBN: 978-0-521-09804-5.

Index

Symbols

A

C

S