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 Sum of Countably Many Non-negative Terms

Definition 1.1.1

Let A be a family of countably infinitely many non-negative terms in extended real numbers. Formally,

$$A = \left\{ a_\alpha \in \overline{\mathbb{R}}_{\geq 0} \mid \alpha \in \Lambda \right\}$$

where Λ is a countably infinite index set. Then the sum of terms in A is defined by the sum of the series

$$\sum_{n=1}^{\infty} a_{\sigma(n)}$$

where $\sigma : \mathbb{N}^* \to \Lambda$ is a bijection. ♣

Definition 1.1.2

Let $\{a_n\}$ be a sequence, and $\sigma : \mathbb{N}^* \to \mathbb{N}^*$ a bijection. Let $\hat{a}_n$ be given by

$$\hat{a}_n := a_{\sigma(n)}, \quad n \in \mathbb{N}^*$$

*Then sequence σ is said to be an **rearrangement** of $\{a_n\}$ into sequence $\{\hat{a}_n\}$.* ♣

Theorem 1.1.1

Let $\{a_n\}$ be a sequence of complex numbers, and σ an rearrangement. If the series $\sum a_n$ converges absolutely to sum s, then $\sum a_{\sigma(n)}$ also converges absolutely to s. ♡

Corollary 1.1.1

Let $\{a_n\}$ be a sequence consisting of non-negative terms in extended real numbers, possibly including infinity, and σ be an rearrangement. Then

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = \sum_{n=1}^{\infty} a_n \tag{1.1}$$

♡

⚐ **Note** *Equation* (1.1) *should be understood as a compact expression containing the following two meanings.*

1. *If $\sum a_n$ converges to a (non-negative) sum s, then $\sum a_{\sigma(n)}$ also converges to s.*
2. *If $\sum a_n$ diverges to infinity (either a_n's are all finite numbers and the series diverges to infinity or there exists ∞ among a_n's), then $\sum a_{\sigma(n)}$ also diverges to infinity.*

Definition 1.1.3

*A **double sequence** is in general any function whose domain is $\mathbb{N}^* \times \mathbb{N}^*$. In this book, we are particularly interested in the function whose values are complex numbers, or non-negative*

extended real numbers. The double sequence is denoted by $\{a_{m,n}\}$ where the function variables $m, n \in \mathbb{N}^*$ are referred to as indices, and the function value $a_{m,n}$ is called term. ♣

Note *Concerning the usual notation for a function in above definition, the double sequence should be written as $f(m, n)$ where $m, n \in \mathbb{N}^*$. However, we adopt the notation $a_{m,n}$ to emphasize that the function values, or terms, are indexed by two natural numbers.*

Definition 1.1.4

Let $\{a_{m,n}\}$ be a double sequence, and $\{s_{p,q}\}$ be another double sequence defined by

$$s_{p,q} = \sum_{m=1}^{p} \sum_{n=1}^{q} a_{m,n}$$

*The pair $(\{a_{m,n}\}, \{s_{p,q}\})$ is called the **double series**, and is denoted by $\sum_{m,n} a_{m,n}$. The sequence $\{s_{p,q}\}$ is referred to as the (p, q)-th partial sum of $\sum_{m,n} a_{m,n}$.* ♣

1.2 Classes of Sets

Definition 1.2.1

*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 1.2.2

*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 1.2.1

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 1.2.3

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 1.2.1, any intersections of σ-algebras is also itself a σ-algebra.

Proposition 1.2.2

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 1.2.1 and Proposition 1.2.2, we may introduce the algebra and σ-algebra generated by a class of subsets in Ω.

Definition 1.2.4

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 1.2.1, $\mathcal{A}(\mathcal{C})$ 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 1.2.3**
>
> *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 1.2.1**
>
> *Let $\mathcal{S}$ be a semi-algebra on Ω. Then*
> $$A \in \mathcal{A}(\mathcal{S}) \iff \exists E_1, \ldots, E_n \in \mathcal{S}, \ A = \biguplus_{i=1}^{n} E_i \tag{1.2}$$
> ♡

Note *Note that the integer n in (1.2) is not fixed. That is the choice of n may vary for different set A.*

Proof (Proof of $\Longleftarrow$) The reverse direction of (1.2) 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 (1.2). 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 (1.2). 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 $\mathcal{C}$. ∎

1.3 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 1.3.1**
>
> *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 1.3.2

*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. *$\mu\left(\uplus_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$*

♣

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

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

First, check that $\mathcal{S}$ is a semi-algebra. Consider the set function μ on $\mathcal{S}$ given by

$$\mu(\emptyset) = 0 \quad \text{and} \quad \mu(a, b] = \begin{cases} \infty, & a = 0 \\ b - a, & a > 0 \end{cases}$$

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

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

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 1.3.3

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

Note *In what follows, by saying μ is continuous from below (or above) without specifying a practical set, we mean that μ is continuous from below (or above) at every set in the given class of sets in the context. And we say μ is continuous if it is continuous both from below and above.*

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 1.3.3.

The next lemma shall be useful in many proofs.

Lemma 1.3.1

Let $\mathcal{A}$ be an algebra on Ω. Suppose μ is an additive set function defined on $\mathcal{A}$. Then we have the following:
1. *If μ is σ-additive, then μ is continuous.*
2. *If μ is continuous from below, then μ is σ-additive.*

> *3. If μ is continuous from above at $\emptyset$ and μ is finite, then μ is σ-additive.* ♡

Since the proof of each statement is quite long, and each deserves some additional comments, we shall present the proofs separately.

Proof of 1 of Lemma 1.3.1:

Proof (Continuity from Below) Suppose $E_n, E \in \mathcal{A}$ and $E_n \uparrow E$. Let

$$F_n = E_n \setminus E_{n-1} \quad \forall n \in \mathbb{N}^*$$

where $E_0 = \emptyset$. Note that

1. $E_n = \biguplus_{i=1}^{n} F_i$, and
2. $\biguplus_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n = E$.

Because μ is additive, we have

$$\mu(E_n) = \sum_{i=1}^{n} \mu(F_i) \tag{1.3}$$

And by further exploiting the fact that μ is σ-additive, we have

$$\mu(E) = \mu\left(\biguplus_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mu(F_n) \tag{1.4}$$

Letting $n \to \infty$ on both sides of (1.3), we find

$$\lim_{n \to \infty} \mu(E_n) = \sum_{i=1}^{\infty} \mu(F_i) = \mu(E)$$

where the last equality follows from (1.4). Therefore, μ is continuous from below.

(Continuity from Above) To prove the continuity from below, we are going to exploit the result that μ is continuity from above, which we have just proved. Suppose that $E_n, E \in \mathcal{A}$, $E_n \downarrow E$, and $\mu(E_N) < \infty$ for some $N \in \mathbb{N}^*$.

✎ **Note** *In order to apply the result that μ is continuous from below, we need to construct a sequence of increasing sets out of E_n's.*

Let

$$F_n = E_N \setminus E_n \quad \forall n \in \mathbb{N}^*$$

Note that

1. $F_n \in \mathcal{A}$,
2. $F_n \subseteq F_{n+1}$, and
3. $\bigcup_{n=1}^{\infty} F_n = E_N \setminus \bigcap_{n=1}^{\infty} E_n = E_N \setminus E$.

Hence, indeed $F_n \uparrow E_N \setminus E$. It then follows that $\mu(F_n) \uparrow \mu(E_N \setminus E)$. Since μ is of course increasing, what matters is the following:

$$\lim_{n \to \infty} \mu(F_n) = \mu(E_N \setminus E) \tag{1.5}$$

Note that we have

$$\mu(F_n) + \mu(E_n) = \mu(E_N)$$
$$\mu(E_N \setminus E) + \mu(E) = \mu(E_N)$$

Equating the right-hand sides yields

$$\mu(F_n) + \mu(E_n) = \mu(E_N \setminus E) + \mu(E)$$

Letting $n \to \infty$, we find

$$\lim_{n \to \infty} \mu(F_n) + \lim_{n \to \infty} \mu(E_n) = \mu(E_N \setminus E) + \mu(E)$$

Plugging in the value of $\lim_{n \to \infty} \mu(F_n)$ using (1.5), we obtain

$$\mu(E_N \setminus E) + \lim_{n \to \infty} \mu(E_n) = \mu(E_N \setminus E) + \mu(E) \tag{1.6}$$

Because $\mu(E_N \setminus E)$ is finite since $\mu(E_N)$ is, we are allowed to cancel this term on both sides of (1.6).

Note *As we can see, it is necessary to assume $\mu(E_n)$ in the definition of continuity from above for we need to subtract a finite number on both side of* (1.6) *to obtain the desired result.*

Therefore,

$$\lim_{n \to \infty} \mu(E_n) = \mu(E)$$

which shows μ is indeed continuous from above. ∎

Note that in the above proof, we always consider the addition of two set functions instead of subtraction. For example, we write

$$\mu(E_N \setminus E_n) + \mu(E_n) = \mu(E_N)$$

where $E_N \setminus E_n = F_n$, instead of

$$\mu(E_N \setminus E_n) = \mu(E_N) - \mu(E_n) \tag{1.7}$$

though it is more tempting to write (1.7). The reason is that (1.7) may not be valid in general for chances are that $\mu(E_n) = \infty$.

Therefore, we must be very careful and first ensure that $\mu(B) < \infty$ before writing down

$$\mu(A \setminus B) = \mu(A) - \mu(B)$$

Proof of 2 of Lemma 1.3.1:

Proof It is given that μ is additive, continuous from below, we want to show that it is σ-additive. Suppose $E_n, E \in \mathcal{A}$ and $E = \biguplus_{n=1}^{\infty} E_n$. Let

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

It is clear that $F_n \uparrow E$. Hence, $\mu(F_n) \uparrow \mu(E)$ since μ is continuous from below. Exploiting the fact that μ is additive, we have

$$\mu(F_n) = \mu\left(\biguplus_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mu(E_i)$$

Therefore,

$$\sum_{i=1}^{\infty} \mu(E_i) = \mu(E)$$

since we have already known the limit of $\mu(F_n)$ is $\mu(E)$. This shows that μ is indeed σ-additive. ∎

Proof of 3 of Lemma 1.3.1:

Proof Suppose that μ is additive, continuous from above at $\emptyset$, and μ is finite. We want to show μ is

σ-additive. Suppose $E_n, E \in \mathcal{A}$ and $E = \biguplus_{n=1}^{\infty} E_n$. Construct a sequence of sets F_n as follows:

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

Because $F_n \downarrow \emptyset$ and $\mu(F_1) < \infty$ (actually $\mu(F_n) < \infty$ for all n), we have $\mu(F_n) \downarrow 0$ by the definition of continuity from above. Since μ is finite and additive, we can write

$$\mu(F_n) = \mu(E) - \mu(\biguplus_{i=1}^{n-1} E_i) = \mu(E) - \sum_{i=1}^{n-1} \mu(E_i)$$

Therefore, letting $n \to \infty$ yields

$$\sum_{i=1}^{\infty} \mu(E_i) = \mu(E)$$

This implies μ is σ-additive. ∎

As the reader might notice, μ is assumed finite in the third statement of Lemma 1.3.1. We shall use an example (see Example 1.4) to address that we really need this assumption. But before showing this example, we must first study how to extend a set function from a semi-algebra to the algebra generated by it.

1.4 Extension of Measures

The next theorem shows that we can extend a set function on a semi-algebra *uniquely* to a set function on the algebra generated by it.

> **Theorem 1.4.1**
>
> *Let $\mathcal{S}$ be a semi-algebra on Ω, and μ an additive set function. Then there exists an additive function ν on the algebra generated by $\mathcal{S}$, $\mathcal{A}(\mathcal{S})$, such that*
>
> 1. *ν is additive,*
> 2. *ν is an extension of μ, i.e., $\nu(E) = \mu(E) \; \forall E \in \mathcal{S}$, and*
> 3. *ν is a unique extension, that is, if there are two additive set functions, ν_1 and ν_2, on $\mathcal{A}(\mathcal{S})$ such that $\nu_1(E) = \nu_2(E) \; \forall E \in \mathcal{S}$, then $\nu_1(A) = \nu_2(A) \; \forall A \in \mathcal{A}(\mathcal{S})$.* ♡

In the proofs of existence and uniqueness, showing the existence is often way difficult than that of uniqueness. As in the proof of this theorem, the first thing we need to do is to construct such an extension ν. Fortunately, the construction is easy and natural to think of thanks to the well-explained structure of the algebra generated by semi-algebra (Theorem 1.2.1).

For every set $A \in \mathcal{A}(\mathcal{S})$, we know that it can written as a finite disjoint union of sets in $\mathcal{S}$. Namely, $\exists E_1, \ldots, E_n \in \mathcal{S}$ such that

$$A = \biguplus_{i=1}^{n} E_i \tag{1.8}$$

Therefore, it is natural to define

$$\nu(A) = \sum_{i=1}^{n} \mu(E_i) \tag{1.9}$$

But note that the representation of A in (1.8) is not unique. Hence, in the proof we are about to present, we need to first show ν is well-defined by (1.9), regardless of different representations of A.

Proof (Well-Definedness of ν) Let $A \in \mathcal{A}$. Suppose

$$A = \biguplus_{i=1}^{n} E_i = \biguplus_{j=1}^{m} F_j$$

where $E_i, F_j \in \mathcal{S}$. We need to show $\sum_{i=1}^{n} \mu(E_i) = \sum_{j=1}^{m} \mu(F_j)$. Since $E_i \subseteq A$, we have

$$E_i = E_i \cap A = \biguplus_{j=1}^{m} E_i \cap F_j$$

It then follows that

$$\mu(E_i) = \sum_{j=1}^{m} \mu(E_i \cap F_j) \tag{1.10}$$

since μ is additive. Similarly, we can also show

$$\mu(F_j) = \sum_{i=1}^{n} \mu(E_i \cap F_j) \tag{1.11}$$

Summing up (1.10) over i and summing up (1.11) over j, we find

$$\sum_{i=1}^{n} \mu(E_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \mu(E_i \cap F_j) = \sum_{j=1}^{m} \mu(F_j)$$

Therefore, ν is well-defined by (1.9).

(Additivity) Next, we show ν is additive. It suffices to show

$$\mu(A \uplus B) = \mu(A) + \mu(B) \quad \forall A, B \in \mathcal{A}(\mathcal{S}) \tag{1.12}$$

Suppose

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

where $E_i, F_j \in \mathcal{S}$. By the definition of ν, we have

$$\nu(A) = \sum_{i=1}^{n} \mu(E_i) \quad \text{and} \quad \nu(B) = \sum_{j=1}^{m} \mu(F_j)$$

On the other hand, since A and B are disjoint, it is clear that $A \uplus B$ is a disjoint union of all E_i's and F_j's, which implies

$$\nu(A) = \sum_{i=1}^{n} \mu(E_i) + \sum_{j=1}^{m} \mu(F_j) = \nu(A) + \nu(B)$$

This is exactly (1.12).

(ν Being An Extension of μ) This part is rather obvious since set E in $\mathcal{S}$ can be regarded as a union of itself. Then, by (1.9), we have

$$\nu(E) = \mu(E)$$

(Uniqueness) Finally, we need to show the extension of μ is unique. Suppose that there are two additive set function, ν_1 and ν_2, on $\mathcal{A}(\mathcal{S})$ such that

$$\nu_1(E) = \nu_2(E) \quad \forall E \in \mathcal{S}$$

Let $A \in \mathcal{A}(\mathcal{S})$. Then there exists $E_1, \ldots, E_n \in \mathcal{S}$ such that $A = \biguplus_{i=1}^{n} E_i$ (Theorem 1.2.1). Because ν_1 and ν_2 are additive, we have

$$\nu_1(A) = \sum_{i=1}^{n} \nu_1(E_i) \quad \text{and} \quad \nu_2(A) = \sum_{i=1}^{n} \nu_2(E_i)$$

Therefore, we see that indeed $\nu_1(A) = \nu_2(A)$. ∎

Example 1.4 Recall Example 1.2 where $\Omega = (0, 1]$,

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

is a semi-algebra, and μ is a set function on $\mathcal{S}$ given by

$$\mu(\emptyset) = 0 \quad \text{and} \quad \mu(a, b] = \begin{cases} \infty, & a = 0 \\ b - a, & a > 0 \end{cases}$$

We have shown in Example 1.2 that μ is additive but not σ-additive.

Now, consider the algebra generated by $\mathcal{S}$, $\mathcal{A}(\mathcal{S})$. By Theorem 1.4.1, we have a unique extension ν of μ on $\mathcal{A}(\mathcal{S})$ such that ν is also additive.

> **Theorem 1.4.2 (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