ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

# The Countable Additivity in the Axiomatic Definition of Probability

**Ailing Shi***

*School of Education, Lanzhou University of Arts and Science, Lanzhou, China*
*Sailing317@163.com*
**Corresponding author*

***Abstract:*** *This paper focuses on the understanding of countable additivity in the axiomatic definition of probability. Through concept decomposition, example verification and theoretical deduction, the core cognition is refined. The results show that the essence of countable additivity is the rule to solve the probability calculation of an infinite number of mutually exclusive events; as one of the three properties of axiomatic probability, it is based on non-negativity and normalization, and is incomplete without it. In the specific case verification process, this paper follows the logic of "defining event, calculation of a single event's probability, and finding the target probability". The conclusions show that to understand countable additivity, it is necessary to closely follow the four-layer framework of 'essence-theory-application-reality', in order to grasp its core value and application logic.*

***Keywords:*** *Countable additivity; Cases; Definition decomposition*

## 1. Introduction

In the axiomatic definition of probability, a real-valued function defined on a sample space denoted by $\Omega$, must satisfy three conditions to be called a probability [1]: non-negativity, normalization, and countable additivity. The first two conditions are more intuitive and have lower comprehension barriers, while the third condition, which involves logic at the infinite level, is often the hardest to understand. First, let's look at non-negativity, which clarifies the lower limit of probability values, meaning that the probability of any random event occurring cannot be negative. In other words, for any event $A$, we have $P(A) \geq 0$. This aligns completely with our understanding of 'likelihood'. For example, it is absolutely invalid logically to say 'the probability of something happening is -0.5'. Now, let's look at normalization, which sets the upper limit of probability values, meaning that the probability of an event that is certain to happen (i.e., the sample space itself) is defined as 1. This is equivalent to setting a standardized measurement benchmark for probability, allowing the probabilities of all events to be compared within the range of 0 to 1.

In comparison, the third condition, additivity, is more abstract. It addresses how to calculate the probability of the union of an infinite number of mutually exclusive events. If there exists a sequence of pairwise mutually exclusive events, denoted as $A_1, A_2, \cdots, A_n, \cdots$, then the total probability that at least one of these infinite events occurs is equal to the sum of the probabilities of each event occurring individually, i.e., $P(A) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots$. This property extends the concept from 'a finite number of events' to 'an infinite number of events', ensuring logical consistency when dealing with infinite sample spaces (such as recording the number of vehicles on a certain road segment during a specific time period). It also lays the foundation for deriving properties or important formulas of probability (such as the total probability formula and Bayes' theorem).

Let's start by understanding countable additivity of probability from four aspects.

## 2. Countable, Mutually Exclusive, and Additive

This section will begin by breaking down the axiomatic definition of probability, focusing on the three core concepts of 'countable', 'mutually exclusive', and 'additive'. Through step-by-step analysis, it aims to lay the foundation for understanding countable additivity.

Firstly, focus on the concept of 'countable'. In probability theory, 'countable' and 'enumerable' are

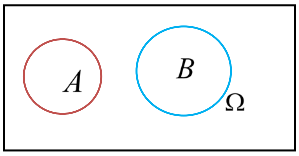
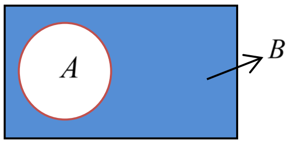
ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

completely equivalent terms, both originating from the definition of countable sets [2]. A countable set (or 'enumerable set') essentially refers to an infinite set that can establish a one-to-one correspondence with the set of positive integers (1, 2, 3, ...). This means that the elements in the set can be arranged in a specific order, like the positive integers, and listed one by one, such as the set of natural numbers (0, 1, 2, ...), the set of integers (...,-2,-1,0,1,2,...), the set of rational numbers (all numbers that can be expressed as fractions), and so on. These all belong to countable sets. From the perspective of set classification, sets can be divided into finite sets and infinite sets; infinite sets can be further subdivided into countably infinite sets (abbreviated as countable sets) and uncountably infinite sets (abbreviated as uncountable sets). This means that not all infinite sets are countable. For example, all real numbers within the interval (0,1) (including finite decimals like 0.1, 0.01, and irrational numbers like $\pi/4$, $\sqrt{2}$ ) cannot establish a one-to-one correspondence with the set of positive integers, making them a typical uncountable set.

Therefore, it can be clearly stated that the countable additivity in the axiomatic definition of probability can also be equivalently expressed as enumerable additivity. The core limitation of this property is that the collection of events for which probabilities can be added must be countably infinite, meaning these events can be listed one by one like natural numbers or integers, and cannot be infinitely uncountable.

Next, let's analyze the concept of mutually exclusive. In probability theory, mutually exclusive is often used interchangeably with mutually incompatible as equivalent terms, serving as a core concept to describe the relationship between events. Intuitively, mutually exclusive can be simply understood as two events are like water and oil, having no overlapping parts. From a set theory perspective (an important theoretical foundation of probability theory), it is more rigorously defined as two events corresponding to sets that have no common elements, i.e., their intersection is an empty set (denoted as $A\cap B=\phi$, where $A$ and $B$ represent the two events). Returning to the context of probability theory itself, the essence of mutually exclusive is that two events cannot occur simultaneously, such as 'flipping a coin and getting head' and 'flipping this coin and getting tails', which are typical mutually exclusive events, and it is impossible for both to occur in a single flip.

*Table 1 Comparison between mutually exclusive events and opposite events*

| Comparison dimension | Mutually exclusive events | Opposite events |
|---|---|---|
| **Defintion** | $A\cap B=\phi$ | $A\cap B=\phi$ and $A\cup B=\Omega$ |
| **Venn diagram** |   |   |
| **The relationship between the two** | Mutually exclusive events are not necessarily opposite events | Opposite events are always mutually exclusive events. |
| **The probabilistic relationship** | $P(A\cap B)=0$, but $P(A)+P(B)$ does not necessarily equal 1 | $P(A\cap B)=0$, and $P(A)+P(B)=1$. |
| **The implied probability meaning** | In an experiment, events $A$ and $B$ cannot happen at the same time, but they can both not happen at the same time. | In an experiment, either $A$ or $B$ occurs, but not both. |
| **Example (rolling a die once)** | $A$ = ``appear '1' '', $B$ =``appear '5' ''. | $A$ ='appear an odd number of dots', $B$ ='appear an even number of dots'. |

It is particularly important to note that the concept of mutually exclusive (mutually incompatible) events in probability theory is easily confused with the concept of 'opposite (mutually inverse) events'. Although both describe a rejection relationship between events, they have essential differences, which can be clearly distinguished through the table 1.

Finally, let's analyze the concept of additivity, which is a core component of countable additivity. It is not merely the accumulation of probabilities, but rather it requires combining the concepts of 'countable' and 'mutually exclusive' from previous discussions to form a complete logical loop. Additivity does not refer to the sum of probabilities of a finite number of events, but rather to the summation of probabilities of countably infinite events. Mathematically, this is the sum of a series. When

ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

we accumulate the probabilities of countably many mutually exclusive events, we are essentially calculating the convergence value of an infinite series. This convergence value represents the total probability of the combined event. This characteristic further aligns with the definition of the 'countable' concept discussed earlier. Because the events are 'countable' (able to be listed one by one like natural numbers), they can form an ordered infinite series, ensuring the logical rigor and computational feasibility of probability summation. If the events are uncountable, an ordered series cannot be formed, and additivity cannot be discussed.

Essentially, the core rule of additivity is that under the precondition that events satisfy 'countable' and 'pairwise mutually exclusive', the total probability of the sum event (that is, the event where 'at least one event occurs') formed by these events is equal to the sum of the probabilities of each event occurring independently. Here, the 'sum event' corresponds to the union of a countable number of mutually exclusive events in set theory (denoted as $A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots$), and its probability calculation follows the rule $P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots$.

## 3. Cases analysis

This section uses two cases to intuitively demonstrate the value of establishing countable additivity in the axiomatic definition of probability: it is not an abstract mathematical rule, but a core tool for solving the probability of the sum of infinitely many mutually exclusive events, aiming to ensure that probability analysis can cover a wider range of application scenarios.

**Case 1:** In the experiment of tossing a die infinitely many times, use countable additivity to derive the probability that "a '1' appears for the first time on the $n$-th toss".

To solve this problem, we can combine the countable additivity in the axiomatic definition of probability and follow the logic of 'define events → calculate the probability of individual events → find the target probability'.

**Step 1:** Define events and anchor a group of mutually incompatible countable events. Let event $A_n \, (n = 1,2,3\cdots)$ represent "toss a die infinitely many times, and a '1' appears for the first time at the $n$-th toss". When $n = 1$, $A_1$ corresponds to tossing a '1' on the first toss; when $n \geq 2$, $A_n$ means a '1' appears for the first time at the $n$-th toss, and no '1' appears in the first $n-1$ tosses. From the relationship between events, events $A_1, A_2, \cdots, A_n, \cdots$ have the following two properties:

(1) Pairwise mutually exclusive, that is, $A_i A_j = \phi, \, i \neq j, i, j = 1,2,\cdots$. This is because the position where '1' first appears is unique. In other words, "first appearing '1' at the $i$-th toss" and "first appearing '1' at the $j(j \neq i)$-th toss" cannot occur simultaneously, so their intersection is an impossible event.

(2) There are infinitely many countable events in this group. Obviously, the value of $n$ can be all positive integers, which can be listed one by one and thus belongs to the category of 'countable'.

**Step 2:** Calculate the probability of individual events. Usually, we calculate the probability of event $A_n \, (n = 1,2,\cdots)$. The occurrence of event $A_n$ means "a '1' appears for the first time at the $n$-th toss", that is, no '1' appears in the first $(n-1)$-th tosses, and the outcomes of those tosses are the other 5 points except '1' (i.e., points 2–6). Let event, $B_i$, represent "a '1' is tossed at the $i$-th time", then event, $\overline{B}_i$, represents "no '1' is tossed at the $i$-th time". Obviously, $B_1, B_2, \cdots, B_n, \cdots$ are pairwise independent (whether a '1' is tossed in the previous time does not affect the probability of tossing a '1' in the next time) and

$$P(B_i) = \frac{1}{6}, P(\overline{B}_i) = \frac{5}{6}, \; i = 1,2,\cdots. \tag{1}$$

Also, since event $A_n$ and the events $B_1, B_2, \cdots, B_n$ satisfy $A_n = \overline{B}_1 \cap \overline{B}_2 \cap \cdots \cap \overline{B}_{n-1} \cap B_n$, from the independence of events, we can obtain

ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

$$\begin{aligned} P(A_n) &= P(\overline{B}_1 \cap \overline{B}_2 \cap \cdots \cap \overline{B}_{n-1} \cap B_n) \\ &= P(\overline{B}_1)P(\overline{B}_2)\cdots P(\overline{B}_{n-1})P(B_n) \\ &= \left(\frac{5}{6}\right)^{n-1}\frac{1}{6}. \end{aligned} \tag{2}$$

**Step 3:** Apply countable additivity to find the probability. Let event $A_n$ represent "toss a die infinitely many times and find that a'1'appears for the first time at the $n$-th toss". Combining the previous analysis, it is easy to obtain $A = A_1 \cup A_2 \cup \cdots$. According to the countable additivity in the axiomatic definition of probability, since the sequence of events $A_1, A_2, \cdots, A_n, \cdots$ are pairwise mutually exclusive, then

$$\begin{aligned} P(A) &= P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) \\ &= P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots \\ &= \sum_{n=1}^{\infty} \left(\frac{5}{6}\right)^{n-1} \frac{1}{6} = \frac{1}{6} \frac{1}{1 - \frac{5}{6}} = 1, \end{aligned} \tag{3}$$

The third-to-last equation here uses the formula for the sum of a geometric sequence. The calculation result, $P(A) = 1$, shows that although the probability of getting a '1' each time is $\frac{1}{6}$, as long as the number of trials is sufficient, a '1' will eventually appear, which is completely consistent with our understanding.

**Case 2:** Deriving the probability of 'at least one occurrence in time $[0,T]$'using countable Additivity in the context of Poisson process

The Poisson process is a classic model in probability theory, which describes the law of random events occurring within continuous time, such as the number of times a customer service staff serves customers within a unit time, the traffic flow of a certain road section within a certain period of time, and so on. In this process, events have the characteristics of stationarity, independence, and sparsity[3]. Below, we use countable additivity to derive the probability of 'at least one event occurring within the time'. This example needs to combine the probability of the sum of infinitely many events and the definition of the Poisson distribution, and its difficulty is higher than the discrete application scenario of Case 1. We still follow the logic of 'define events → calculate the probability of individual events → find the target probability'.

**Step 1:** Clarify the core parameters in the Poisson process and define events.

Let the positive constant $\lambda$ represent the event occurrence rate of the Poisson process, that is, the average number of events occurring per unit time. Let $A$ represent 'at least one event occurs within time $[0,T]$', and $B_k\ (k = 0,1,2,\cdots)$ represent 'exactly $k$ events occur within time $[0,T]$', where $k = 0$ represents 'exactly 0 events occur within time $[0,T]$', that is, no event occurs within time $[0,T]$. By analyzing events $A$ and $B_k\ (k = 0,1,2,\cdots)$, the following conclusions can be easily drawn.

(1) $B_k\ (k = 0,1,2,\cdots)$ form a countable mutually exclusive event group. On the one hand, because all possible values of $k$ are non-negative integers, which can be listed one by one, satisfying the definition of 'countable'. On the other hand, because 'exactly $i$ events occur within time $[0,T]$' and 'exactly $j (j \neq i)$ events occur within time $[0,T]$' cannot happen at the same time, the event group $B_k\ (k = 0,1,2,\cdots)$ is pairwise mutually exclusive.

(2) The relationship between event $A$ and the event group $B_k\ (k = 0,1,2,\cdots)$ is $A = B_1 \cup B_2 \cup \cdots \cup B_n \cup \cdots$. This is because the event 'at least one occurrence' is equivalent to the event 'exactly 1 occurrence, or exactly 2 occurrences, or exactly 3 occurrences, ...', and event $A$ and event $B_0$

ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

are mutually opposite events (that is, the event 'no occurrence at all' and the event 'at least 1 occurrence' are mutually opposite events).

**Step 2:** Introduce the probability formula of the Poisson distribution.

According to the property of the Poisson process [2], the probability that 'exactly $j(j \neq i)$ events occur within time $[0,T]$' follows the Poisson distribution, and the formula is

$$P(B_k) = \frac{(\lambda T)^k \mathrm{e}^{-\lambda T}}{k!}, \quad k = 0,1,2,\cdots, \tag{4}$$

where 'e' is the natural constant (approximately 2.71828). In particular, the probability of 'exactly 0 occurrences' is

$$P(B_0) = \mathrm{e}^{-\lambda T}. \tag{5}$$

**Step 3:** Derive $P(A)$ using countable additivity.

Methodologically, it is entirely feasible to calculate $P(A)$ using the inverse probability formula. However, it should be noted that the focus of the following derivation is not merely on calculating the result, but on clarifying the practical meaning of countable additivity and making it more perceptible through process decomposition.

Since $B_0, B_1, B_2, \cdots, B_n, \cdots$ exhaust all possible results of the number of events occurring within the time $[0,T]$, according to the normativity of probability (the probability of a certain event is 1), we have

$$P(B_0 \cup B_1 \cup B_2 \cup \cdots) = 1. \tag{6}$$

Since the event group $B_k$ $(k = 0,1,2,\cdots)$ is pairwise mutually exclusive, according to countable additivity, we have

$$P(B_0) + P(B_1) + P(B_2) + \cdots = 1. \tag{7}$$

Also, because $A = B_1 \cup B_2 \cup \cdots \cup B_n \cup \cdots$ ,by using countable additivity again, we obtain

$$P(A) = P(B_1) + P(B_2) + \cdots. \tag{8}$$

Combining Eqs. (7) and (8) gets

$$P(A) = 1 - P(B_0) = 1 - \mathrm{e}^{-\lambda T}. \tag{9}$$

In this case, countable additivity is the link connecting the probability of individual events $P(B_k)$ and the target probability $P(A)$ .Without this property, it would be impossible to establish a practical calculation model for the Poisson process.

## 4. Analysis of Difficulties in Cases

The difficulty of the first example is reflected in the following three aspects.

(1) In-depth understanding of event mutual exclusivity. In relatively complex practical problems, how to accurately judge whether multiple events are mutually exclusive, especially when the description of events is rather implicit, judgment errors may occur. In this die-tossing example, it is necessary to deeply understand that the position where '1' first appears in each die toss is uniquely determined, and there is no situation where '1' first appears at different times simultaneously. This is the key to understanding the mutual exclusivity of events.

(2) Understanding of the concept of countable infinity. A set being countable means that the elements in the set can establish a one-to-one correspondence with the set of positive integers. This description may be rather abstract for some readers who have less contact with set theory and the concept of infinity. In fact, being countable means that the elements can be arranged into an infinite sequence. In the die-tossing example, since the first occurrence of '1' happens at the 1st time, the 2nd time, the 3rd time, ..., and so on infinitely, this infinite and sortable characteristic is a manifestation of countable infinity.

ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

(3) Application of independence in event probability calculation. In the die-tossing experiment, it is mentioned that $B_1, B_2, \cdots, B_n, \cdots$ are mutually independent. However, accurately grasping and applying the probability concept of independence is a difficulty. Beginners may confuse the independence and mutual exclusivity of events, and in practical problems, judging whether events are mutually independent is not an easy task. In the die-tossing scenario, the result of each toss does not affect each other. But in other complex scenarios, how to correctly judge and apply independence needs to be focused on.

The difficulty of the second example focuses on ‘conversion of event dimensions’ and ‘implicit verification of infinite series’, which is specifically reflected in the following two aspects.

(1) Mapping from ‘continuous time’ to ‘discrete events’. The time axis of the Poisson process is continuous (for example, the interval $[0,T]$ is continuous), but the number of event occurrences is discrete. It is necessary to first define countable discrete events ( $B_0, B_1, B_2, \cdots, B_n, \cdots$) in the continuous-time scenario according to the target requirements, so as to satisfy the ‘countable’ premise of countable additivity. This is the key difference from discrete scenarios like the first example.

(2) Implicit guarantee of infinite series convergence. From the perspective of series convergence, countable additivity requires the sum of the probabilities of infinitely many events to converge, and the probability sum of the Poisson distribution exactly satisfies this condition, that is,

$$\sum_{k=0}^{\infty} P(B_k) = \sum_{k=0}^{\infty} \frac{(\lambda T)^k}{k!} \mathrm{e}^{-\lambda T} = \mathrm{e}^{-\lambda T} \mathrm{e}^{\lambda T} = 1 . \tag{10}$$

The second-to-last equation in the above formula uses the Taylor expansion of $\mathrm{e}^x$ , that is, $\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Formula (10) shows that the sum of $P(B_k)$ is 1, which ensures the rigor of the application of countable additivity. Without this convergence verification, directly using countable additivity will lead to logical loopholes.

## 5. Conclusion

This paper, through concept decomposition, case verification, and theoretical derivation, sorts out the core cognition for understanding countable additivity in the axiomatic definition of probability, which is specifically reflected in the following four aspects.

First, the essence of countable additivity is a rule established to handle the probabilities of infinitely many mutually exclusive events. When events satisfy countability (having a one-to-one correspondence with the set of positive integers) and ‘pairwise mutual exclusivity’ (their intersection is empty), the probability of the sum event is equal to the sum of the probabilities of infinitely many events. Thus, the additivity of finitely many mutually exclusive events is extended to infinite scenarios.

Second, countable additivity is the link of the axiomatic system of probability theory. As one of the three major properties of the axiomatization of probability, it takes non-negativity and normativity as premises to ensure the logical self-consistency of infinite summation. At the same time, it provides a basis for the construction of the law of total probability and Poisson processes. If it is missing, scenarios such as infinite trials cannot be handled, and the axiomatic system will not be complete.

Third, follow the application logic of the three-step method of ‘premise-decomposition-summation’, that is, first verify the mutual exclusivity and countability of the event group; then decompose the target event into the sum event of countable mutually exclusive events; finally, calculate the total probability through infinite series summation (convergence needs to be verified). This is a practical guide for countable additivity and also a standard for judging scenario adaptability.

Fourth, countable additivity is a bridge from theory to practice. Countable additivity can be used to explain the law that a small-probability event is almost certain to occur in infinite trials, and can also convert continuous-time scenarios (such as Poisson processes) into the summation of discrete events, providing operable models for fault assessment, traffic prediction, etc. It is the core for relevant probability theories to move from abstraction to practicality.

ISSN 2522-6398 Vol. 9, Issue 1: 1-7, DOI: 10.25236/FER.2026.090101

To sum up, understanding countable additivity needs to focus on the four levels of ‘essence-theory-application-reality’. It is a rule for infinite summation in probability, a cornerstone for the construction of the axiomatic system, and a tool for practical application. Only by grasping this framework can one understand its core value and application logic.

## Acknowledgements

This research is supported by the Doctoral Special Project of Yanyuan Science and Technology Innovation Fund (No.2023BSZX05)

## References

*[1] Mao S S, Cheng Y M, Pu X L. Probability Theory and Mathematical Statistics. 2nd Edition. Beijing: Higher Education Press, 2010.*
*[2] Cheng Q X, et al. Fundamentals of Real Variable Function and Functional Analysis. 4th Edition. Beijing: Higher Education Press, 2019.*
*[3] Sheldon M. Ross. Introduction to Probability Models. Tenth Edition. Singapore: Elsevier (Singapore) Pte Ltd., 2010.*