MSTAT 551: Probability and Measure Theory

Gabriel Asare Okyere (PhD)

Department of Statistics and Actuarial Science,
KNUST.

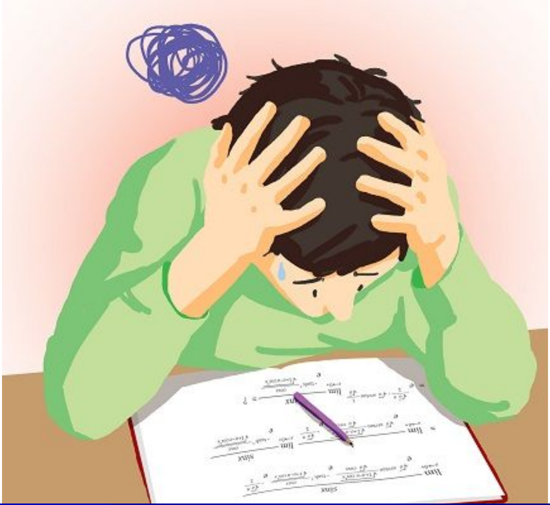
January 2026

Learning objectives

At the end of this section, learners will be able to:

Outer Measure

Suppose you want to assign a ‘length’ to a very complicated set of points on a line, like a set that is infinite or scattered irregularly. How could you systematically estimate its size using only intervals?



Motivation for Outer Measure

1. Measuring Beyond Simple Sets

We know how to measure intervals in $\mathbb{R}$ or rectangles in $\mathbb{R}^n$, but many sets are more complicated.

Example:

Let

$$A = \{0\} \cup \{1/2\} \cup \{2/3\} \cup \cdots \subset \mathbb{R}$$

which is not a single interval.

Idea:

Outer measure allows us to assign a "length" by covering each point with a very small interval and summing their lengths.

Motivation for Outer Measure

2. Approximation from the Outside

Outer measure measures a set by covering it with countably many simple sets and taking the smallest total size.

Example:
The Cantor set in $[0, 1]$ is not a union of intervals.

Idea:
We cover it with a sequence of intervals whose total length can be made arbitrarily small.
Outer measure gives the *infimum* of these total lengths, matching our geometric intuition.

3. Defining Lebesgue Measure Rigorously

Outer measure is the first step in constructing Lebesgue measure.

Example / Idea:

1. Start with outer measure m^* defined on all subsets of $\mathbb{R}$ using intervals
2. Identify sets where σ-additivity holds (measurable sets)
3. Restrict m^* to these sets to get the Lebesgue measure

Outer Measure

Let Ω be a set and $\mathcal{A}$ be an algebra of sets of Ω. Suppose μ is a σ-additive measure on $\mathcal{A}$. From the algebra property: if $A \in \mathcal{A}$, then $A^c := \Omega \setminus A \in \mathcal{A}$.

Definition

For any set $A \subset \Omega$, define its outer measure $\mu^*(A)$ by:

$$\mu^*(A) = \inf \left\{ \sum_{k=1}^{\infty} \mu(A_k) : A_k \in \mathcal{A} \text{ and } A \subset \bigcup_{k=1}^{\infty} A_k \right\} \quad \ldots\ldots\ldots\ldots (1.7)$$

In other words, we consider all coverings $\{A_k\}_{k=1}^{\infty}$ of A by a sequence from the algebra $\mathcal{A}$ and define $\mu^*(A)$ as the infimum of the sum of all $\mu(A_k)$, taken over all such coverings.

Outer Measure

Properties of outer measure

- Null (empty) set: $\mu^*(\emptyset) = 0$
- Monotonicity: if $A \subset B \subset \Omega$, then $\mu^*(A) \leq \mu^*(B)$.
- Countable subadditivity (σ-subadditivity): For any countable collection of sets $\{A_n\}_{n=1}^{\infty} \subset \Omega$:
  $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Outer measure

Lemma

For any set $A \subset \Omega$, $\mu^*(A) < \infty$ and if $A \in \mathcal{A}$, then $\mu^*(A) = \mu(A)$.

Proof

Note that $\emptyset \in \mathcal{A}$ and $\mu(\emptyset) = 0$ because $\mu(\emptyset) = \mu(\emptyset \sqcup \emptyset) = \mu(\emptyset) + \mu(\emptyset)$.

For any set $A \subset \Omega$, consider a covering $\{A_k\} = \{\Omega, \emptyset, \emptyset, \dots\}$ of A.

Since $\Omega, \emptyset \in \mathcal{A}$, it follows from (1) that

$\mu^*(A) \leq \mu(\Omega) + \mu(\emptyset) + \mu(\emptyset) + \cdots = \mu(\Omega) < \infty.$

Assume now $A \in \mathcal{A}$. Considering a covering $\{A_k\} = \{A, \emptyset, \emptyset, \dots\}$

and using that $A, \emptyset \in \mathcal{A}$, we obtain in the same way that

$\mu^*(A) \leq \mu(A) + \mu(\emptyset) + \mu(\emptyset) + \cdots = \mu(A).$

Outer measure

Continuation of Proof

On the other hand, for any sequence $\{A_k\}$ as in (1), we have by the

σ-subadditivity of μ that

$\sum_{k=1}^{\infty} \mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$.

Taking the infimum over all such sequences $\{A_k\}$, we obtain

$\mu(A) \leq \mu^*(A)$, which together with the previous inequality yields

$\mu^*(A) = \mu(A)$. ■

Lemma

The outer measure μ^* is σ-subadditive on 2^{Ω}.

Outer measure

Proof

We need to prove that if $A \subset \bigcup_{k=1}^{\infty} A_k$ where A and A_k are subsets of Ω,

then $\mu^*(A) \leq \sum_{k=1}^{\infty} \mu^*(A_k)$.

By the definition of μ^*, for any set A_k and any $\varepsilon > 0$

there exists a sequence $\{A_{kn}\}_{n=1}^{\infty}$ of sets from R such that

$A_k \subset \bigcup_{n=1}^{\infty} A_{kn}$ and $\mu^*(A_k) \geq \sum_{n=1}^{\infty} \mu(A_{kn}) - \frac{\varepsilon}{2^k}$.

Adding these inequalities over all k, we obtain

$\sum_{k=1}^{\infty} \mu^*(A_k) \geq \sum_{k,n=1}^{\infty} \mu(A_{kn}) - \varepsilon$.

Continuation of proof

On the other hand, by the inclusions $A \subset \bigcup_{k=1}^{\infty} A_k$ and $A_k \subset \bigcup_{n=1}^{\infty} A_{kn}$,

we get $A \subset \bigcup_{k,n=1}^{\infty} A_{kn}$. Since $A_{kn} \in \mathcal{A}$, it follows from (1) that

$\mu^*(A) \leq \sum_{k,n=1}^{\infty} \mu(A_{kn})$.

Comparing with the previous inequality gives

$\mu^*(A) \leq \sum_{k=1}^{\infty} \mu^*(A_k) + \varepsilon$.

Since this holds for any $\varepsilon > 0$, it also holds for $\varepsilon = 0$, which completes the proof. ■

Outer Measure

From Outer Measure to Measurable Sets

Outer measure μ^* assigns a "size" to any set $E \subset \mathbb{R}$.

But not all sets behave nicely under μ^*.

For example, we want additivity: splitting a set shouldn't change total size.

Symmetric Difference

Definition

The symmetric difference of two sets $A, B \subset \Omega$ is the set
$A \triangle B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$

- Clearly, $A \triangle B = B \triangle A$
- Also, $x \in A \triangle B$ if and only if x belongs to exactly one of the sets A, B. That is, either $x \in A$ and $x \notin B$ or $x \notin A$ and $x \in B$.

Symmetric Difference

Lemma

If μ^* is an outer measure on Ω, then $|\mu^*(A) - \mu^*(B)| \leq \mu^*(A \triangle B)$, for arbitrary sets $A, B \subset \Omega$.

Proof of Lemma

Note that

$$A \subset B \cup (A \setminus B) \subset B \cup (A \triangle B)$$

where by the subadditivity of μ^*

$$\mu^*(A) \leq \mu^*(B) + \mu^*(A \triangle B),$$

where

$$\mu^*(A) - \mu^*(B) \leq \mu^*(A \triangle B).$$

Switching A and B, we obtain a similar estimate:

$$\mu^*(B) - \mu^*(A) \leq \mu^*(A \triangle B),$$

hence the inequality in the Lemma follows.

Measurable Sets

We still consider $\mathcal{A}$ to be an algebra on Ω and μ is a σ-additive measure on $\mathcal{A}$, while holding the definition of μ^*.

Definition of Measurable sets

A set $A \subset \Omega$ is called measurable (with respect to the algebra $\mathcal{A}$ and the measure μ) if, for any $\varepsilon > 0$, there exist $B \in \mathcal{A}$ such that $\mu^*(A \triangle B) < \varepsilon$.

In other words, set A is measurable if it can be approximated by sets from $\mathcal{A}$ arbitrarily closely.

Measurable Sets

Alternative Definition of Measurable Sets

Let μ^* be an outer measure on a set Ω. A subset $E \subset \Omega$ is called **measurable** (with respect to μ^*) if for every subset $A \subseteq \Omega$:

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

where $E^c = \Omega \setminus E$ is the complement of E.

It essentially says that E splits any set A 'nicely' with respect to outer measure.

Measurable sets

Examples of measurable sets

1. Intervals in the real line. i.e. $(a, b), [a, b), [a, b], (-\infty, a), \text{or}(b, +\infty)$.
2. Finite sets and countable sets. i.e. $\{a, b, c\}, \mathbb{N}, \mathbb{Q}$, etc.
3. Complements of measurable sets. i.e. $\mathbb{R} \setminus \mathbb{Q}$, etc

Measurable sets

Are all sets Measurable? Why?

Stefan Banach – Polish Mathematician

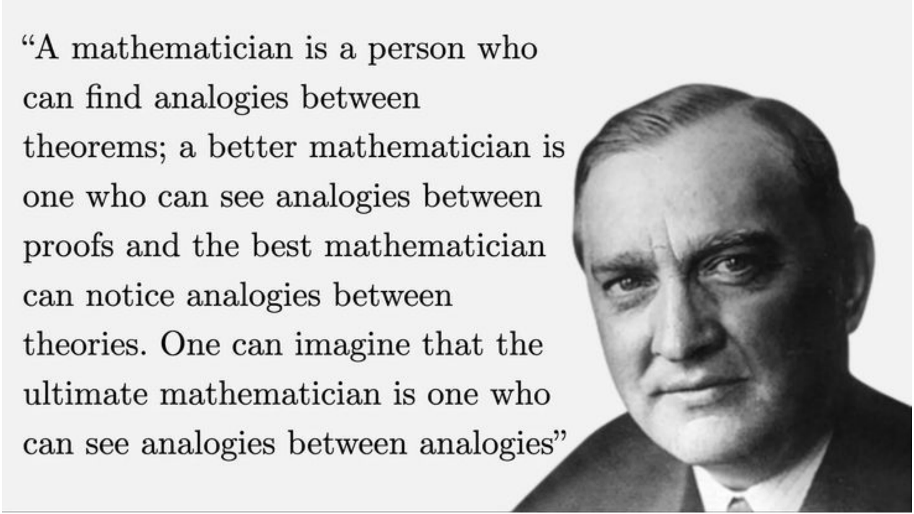


Figure: Stefan Banach

Alfred Tarski – Polish-American Mathematician

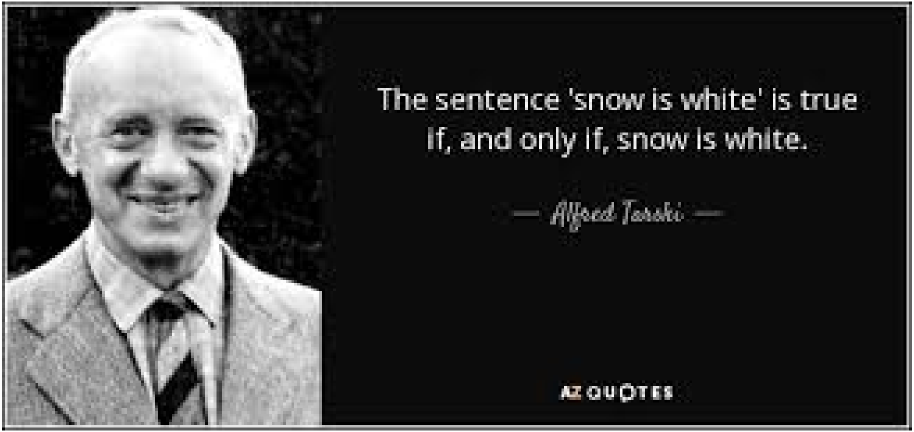


Figure: Alfred Tarski

Giuseppe Vitali – Italian Mathematician

Non-Measurable sets

Why All Sets Are Not Measurable

- **Length, area, volume in real life**
  - When measuring objects (a rope, a piece of land, a container), we only deal with well-behaved sets: intervals, rectangles, polygons, or unions of them
  - These are all measurable because their size can be consistently determined
- **Constructed 'pathological' sets**
  - Some sets are created using the axiom of choice and are so irregular that you cannot assign a meaningful size
  - Example: Vitali set—choosing one number from each equivalence class modulo rationals between 0 and 1. No consistent length can be assigned
- **3D paradoxical sets**
  - Banach–Tarski paradox shows you can partition a ball into strange pieces that can be reassembled into two identical ctc

Figure: examples of non-measurable sets.

Axiom of Choice (AC)

Formal Statement

The **axiom of choice (AC)** asserts that for any family of nonempty sets $\{S_i\}_{i \in I}$, there exists a function f (called a *choice function*) such that

$$f(i) \in S_i \quad \text{for each } i \in I.$$

Simpler Terms

You can always pick one element from each set, even if there are infinitely many sets and no explicit rule for choosing.

It is needed because it guarantees selections from infinite families of sets. However, it is controversial and thus leads to paradoxical results like the Banach–Tarski paradox (a ball can be split and reassembled into two identical balls).

Illustrative Examples of Non-Measurable Sets

Vitali set ($\mathbb{R}$):

- Consider all real numbers between 0 and 1.
- Partition them into equivalence classes where numbers differ by a rational number.
- Pick one number from each class.
- The resulting set is the Vitali set.
- You cannot assign a consistent length to it, even though each piece "looks like a number."

Banach–Tarski paradox (3D):

- Start with a solid 3D ball (like a basketball).
- Using very strange, non-physical pieces, it is possible to cut it into finitely many pieces and reassemble them into two balls of the same size.
- Each piece is non-measurable, meaning no volume can be consistently assigned.

Assignment

Illustrate the non-measurable Bernstein set of real numbers.

Theorem

Carathéodory's extension theorem

Let $\mathcal{A}$ be an algebra on a set Ω and μ be a σ-additive measure on $\mathcal{A}$. Denote by $\mathcal{M}$ the family of all measurable subsets of Ω. Then the following is true:

1. $\mathcal{M}$ is a σ-algebra containing $\mathcal{A}$.
2. The restriction of μ^* on $\mathcal{M}$ is a σ-additive measure (that extends measure μ from $\mathcal{A}$ to $\mathcal{M}$).
3. If $\tilde{\mu}$ is a σ-additive measure defined on a σ-algebra $\mathcal{F}$ such that $\mathcal{A} \subset \mathcal{F} \subset \mathcal{M}$, then $\tilde{\mu} = \mu^*$ on $\mathcal{M}$.

From this theorem, we can make the following claims.

Claim 1: *The family Ω of all measurable sets is an algebra containing $\mathcal{A}$.*

- If $A \in \mathcal{A}$, then A is measurable because:

$$\mu^*(A \triangle A) = \mu^*(\emptyset) = \mu(\emptyset) = 0,$$

  where $\mu^*(\emptyset) = \mu(\emptyset)$.

- Hence, $\mathcal{A} \subset \Omega$ and the entire space Ω is a measurable set.

To verify $\mathcal{M}$ is an algebra: Show that for $A_1, A_2 \in \mathcal{M}$, both $A_1 \cup A_2$ and $A_1 \setminus A_2$ are measurable.

- Let $A = A_1 \cup A_2$. For any $\varepsilon > 0$, there exist $B_1, B_2 \in \mathcal{A}$ such that

$$\mu^*(A_1 \triangle B_1) < \varepsilon \quad \text{and} \quad \mu^*(A_2 \triangle B_2) < \varepsilon \ldots (1.19)$$

- Set $B = B_1 \cup B_2 \in \mathcal{A}$.

Proof

- Then by the symmetric difference Lemma and the subadditivity of μ^* respectively, we have

$$A \triangle B \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$$

and

$$\mu^*(A \triangle B) \leq \mu^*(A_1 \triangle B_1) + \mu^*(A_2 \triangle B_2) < 2\varepsilon \ldots (1.20)$$

- Since $\varepsilon > 0$ is arbitrary and $B \in \mathcal{A}$, then A is measurable. Similarly, $A_1 \setminus A_2 \in \mathcal{M}$.

Claim 2: μ^* is σ-additive on $\mathcal{M}$.

Since $\mathcal{M}$ is an algebra and μ^* is σ-subadditive, it suffices to prove that μ^* is finitely additive on $\mathcal{M}$.
Let us prove that, for any two disjoint measurable sets A_1 and A_2, we have

$$\mu^*(A) = \mu^*(A_1) + \mu^*(A_2)$$

where $A = A_1 \cup A_2$. Then we have the inequality

$$\mu^*(A) \leq \mu^*(A_1) + \mu^*(A_2)$$

so that we are left to prove the opposite inequality

$$\mu^*(A) \geq \mu^*(A_1) + \mu^*(A_2).$$

Proof

For any $\varepsilon > 0$, there are sets $B_1, B_2 \in \mathcal{A}$ such that

$$\mu^*(A_1 \triangle B_1) < \varepsilon \quad \text{and} \quad \mu^*(A_2 \triangle B_2) < \varepsilon.$$

Set $B = B_1 \cup B_2 \in \mathcal{A}$. Then

$$\mu^*(A \triangle B) \le \mu^*(A_1 \triangle B_1) + \mu^*(A_2 \triangle B_2) < 2\varepsilon,$$

where $A = A_1 \cup A_2$. In particular,

$$|\mu^*(A) - \mu^*(B)| \le \mu^*(A \triangle B) < 2\varepsilon.$$

On the other hand, since $B \in \mathcal{A}$, we have

$$\mu^*(B) = \mu(B) = \mu(B_1 \cup B_2) = \mu(B_1) + \mu(B_2) - \mu(B_1 \cap B_2).$$

Proof

Next, we estimate $\mu(B_i)$ from below via $\mu^*(A_i)$ and show that $\mu(B_1 \cap B_2)$ is small enough.
Indeed, for any $i = 1, 2$, we have

$$|\mu^*(A_i) - \mu^*(B_i)| \leq \mu^*(A_i \triangle B_i) < \varepsilon,$$

whence

$$\mu(B_1) \geq \mu^*(A_1) - \varepsilon \quad \text{and} \quad \mu(B_2) \geq \mu^*(A_2) - \varepsilon.$$

On the other hand, if $A_1 \cap A_2 = \emptyset$, then

$$B_1 \cap B_2 \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2),$$

so

$$\mu(B_1 \cap B_2) \leq \mu^*(A_1 \triangle B_1) + \mu^*(A_2 \triangle B_2) < 2\varepsilon.$$

Proof

It follows that

$$\mu^*(A) \geq (\mu^*(A_1) - \varepsilon) + (\mu^*(A_2) - \varepsilon) - 2\varepsilon - 2\varepsilon = \mu^*(A_1) + \mu^*(A_2) - 6\varepsilon.$$

Letting $\varepsilon \to 0$, we finish the proof.

Applicability of Carathéodory's Extension Theorem

Constructing Lebesgue Measure

- Start with the length function defined on intervals of $\mathbb{R}$.
- Apply Carathéodory's theorem to extend this pre-measure to the σ-algebra of Lebesgue measurable sets.
- This gives us the rigorous foundation of integration.

Probability Measures

- Define probabilities on an algebra of "simple events" (like finite unions of intervals).
- Extend to the σ-algebra of Borel sets, ensuring random variables and distributions are well-defined.
- Example: Extending the uniform distribution from intervals in $[0, 1]$ to all Borel subsets.

Reading Assignment

Read on the following applicability of Caratheordory's Extension Theorem

1. Product Measures
2. Outer Measure and Measurability
3. Abstract Measure Spaces

Group Assignment

Prove the following claims:

1. $\mathcal{M}$ is σ-algebra

2. If $\mathcal{F}$ is a σ-algebra such that

$$\mathcal{A} \subset \mathcal{F} \subset \mathcal{M},$$

and $\tilde{\mu}$ is a σ-additive measure on Σ such that $\tilde{\mu} = \mu$ on $\mathcal{A}$. Then $\tilde{\mu} = \mu^*$ on $\mathcal{F}$.
Prove that

$$\tilde{\mu}(A) = \mu^*(A) \quad \text{for any } A \in \mathcal{F}.$$

References

A. Grigoryan, *Measure Theory and Probability*, University of Bielefeld, Lecture Notes.

K. B. Athreya and S. N. Lahiri, *Measure and Probability Theory*, Springer, 2006.

P. Billingsley, *Probability and Measure*, 3rd ed., Wiley, 1995.

A Probability Space, *Measure Theory and Probability*, YouTube lecture series, 2023.
Link: `https://youtu.be/swa1VRYms3Q`