Lebesgue Measure

The idea of the Lebesgue integral is to first define a **measure** on subsets of $\mathbb{R}$. That is, we wish to assign a number $m(S)$ to each subset S of $\mathbb{R}$, representing the total length that S takes up on the real number line. For example, the measure $m(I)$ of any interval $I \subseteq \mathbb{R}$ should be equal to its length $\ell(I)$.

Measure should also be **additive**, meaning that the measure of a disjoint union of two sets is the sum of the measures of the sets:

$$m(S \uplus T) \;=\; m(S) + m(T).$$

Indeed, if we want m to be compatible with taking limits, it should be **countably additive**, meaning that

$$m\left(\biguplus_{n\in\mathbb{N}} S_n\right) \;=\; \sum_{n\in\mathbb{N}} m(S_n)$$

for any sequence $\{S_n\}$ of pairwise disjoint subsets of $\mathbb{R}$.

Of course, the measure $m(\mathbb{R})$ of the entire real line should be infinite, as should the measure of any open or closed ray. Thus the measure should be a function

$$m\colon \mathcal{P}(\mathbb{R}) \to [0,\infty]$$

where $\mathcal{P}(\mathbb{R})$ is the power set of $\mathbb{R}$.

> **Question: Measuring Subsets of $\mathbb{R}$**
> Does there exist a function $m\colon \mathcal{P}(\mathbb{R}) \to [0,\infty]$ having the following properties?
>
> 1. $m(I) = \ell(I)$ for every interval $I \subseteq \mathbb{R}$.
>
> 2. For every sequence $S_1, S_2, \ldots$ of pairwise disjoint subsets of $\mathbb{R}$,
>
> $$m\left(\biguplus_{n\in\mathbb{N}} S_n\right) \;=\; \sum_{n\in\mathbb{N}} m(S_n).$$

Surprisingly, the answer to this question is **no**, although it will be a while before we prove this. But it turns out that it is impossible to define a function $m\colon \mathcal{P}(\mathbb{R}) \to [0,\infty]$ satisfying both of the conditions above.

The reason is that there exist certain subsets of $\mathbb{R}$ that really cannot be assigned a measure. In fact, there is a rigorous sense in which *most* subsets of $\mathbb{R}$ cannot be assigned a measure. Interestingly, actual examples of this phenomenon are difficult to construct, with all such constructions requiring the axiom of choice. As a result, such poorly behaved sets are quite rare in practice, and it is possible to define a measure that works well for almost any set that one is likely to encounter.

Thus our plan is to restrict ourselves to a certain collection $\mathcal{M}$ of subsets of $\mathbb{R}$, which we will refer to as the **Lebesgue measurable sets**. We will then define a function

$$m\colon \mathcal{M} \to [0,\infty]$$

called the **Lebesgue measure**, which has all of the desired properties, and can be used to define the Lebesgue integral. The following theorem summarizes what we are planning to prove.

Main Theorem Existence of Lebesgue Measure

There exists a collection $\mathcal{M}$ of subsets of $\mathbb{R}$ (the measurable sets) and a function $m\colon \mathcal{M} \to [0,\infty]$ satisfying the following conditions:

1. *Every interval $I \subseteq \mathbb{R}$ is measurable, with $m(I) = \ell(I)$.*
2. *If $E \subseteq \mathbb{R}$ is a measurable set, then the complement $E^c = \mathbb{R} - E$ is also measurable.*
3. *For each sequence $\{E_n\}$ of measurable sets in $\mathbb{R}$, the union $\bigcup_{n\in\mathbb{N}} E_n$ is also measurable. Moreover, if the sets $\{E_n\}$ are pairwise disjoint, then*

$$m\left(\biguplus_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} m(E_n).$$

Lebesgue Outer Measure

We begin by defining the **Lebesgue outer measure**, which assigns to each subset S of $\mathbb{R}$ an "outer measure" $m^*(S)$. Thus m^* will be a function

$$m^*\colon \mathcal{P}(\mathbb{R}) \to [0,\infty]$$

where $\mathcal{P}(\mathbb{R})$ denotes the power set of $\mathbb{R}$.

Of course, m^* will not be countably additive. Instead, it will have the weaker property of **countable subadditivity**, meaning that

$$m^*\left(\bigcup_{n\in\mathbb{N}} S_n\right) \leq \sum_{n\in\mathbb{N}} m^*(S_n)$$

for any sequence $\{S_n\}$ of subsets of $\mathbb{R}$.

The outer measure $m*$ should be thought of as our first draft of Lebesgue measure. Indeed, once we determine which subsets of $\mathbb{R}$ are measurable, we will simply restrict m^* to the collection $\mathcal{M}$ of measurable sets to obtain the Lebesgue measure m. Thus, even though m^* is not countably additive in general, it will turn out be countably additive on the collection of measurable sets.

For the following definition, we say that a collection $\mathcal{C}$ of subsets of $\mathbb{R}$ **covers** a set $S \subseteq \mathbb{R}$ if $S \subseteq \bigcup \mathcal{C}$.

Definition: Lebesgue Outer Measure

If $S \subseteq \mathbb{R}$, the **(Lebesgue) outer measure** of S is defined by

$$m^*(S) = \inf\left\{\sum_{I\in\mathcal{C}} \ell(I) \;\middle|\; \mathcal{C} \text{ is a collection of open intervals that covers } S\right\}.$$

It should make intuitive geometric sense that $m^*(J) = \ell(J)$ for any interval J, though we will put off the proof of this for a little while. The difficult part is to show that if we cover an interval J with open intervals, then the sum of the lengths of the open intervals is greater than or equal to the length of J.

Note that $m^*(S)$ may be infinite if $\sum_{I\in\mathcal{C}} \ell(I)$ is infinite for every collection $\mathcal{C}$ of open intervals that covers S. For example, it is not hard to see that $m^*(\mathbb{R})$ must be infinite.

Proposition 1 Properties of m^*

Lebesgue outer measure m^* has the following properties:

1. $m^*(\emptyset) = 0$.

2. *If $S \subseteq T \subseteq \mathbb{R}$, then $m^*(S) \leq m^*(T)$.*

3. *If $\{S_n\}$ is a sequence of subsets of $\mathbb{R}$, then*

$$m^*\left(\bigcup_{n\in\mathbb{N}} S_n\right) \leq \sum_{n\in\mathbb{N}} m^*(S_n)$$

PROOF Statement (1) is obvious from the definition. For (2), let $S \subseteq T \subseteq \mathbb{R}$, and let $\mathcal{C}$ be any collection of open intervals that covers T. Then $\mathcal{C}$ also covers S, so

$$m^*(S) \;\leq\; \sum_{I\in\mathcal{C}} \ell(I).$$

This holds for every cover $\mathcal{C}$ of T by open intervals, and therefore $m^*(S) \leq m^*(T)$.

For (3), let $\{S_n\}$ be a sequence of subsets of $\mathbb{R}$, and let $S = \bigcup_{n\in\mathbb{N}} S_n$. If $m^*(S_n)$ is infinite for some n, then by statement (2) it follows that $m^*(S) = \infty$, and we are done. Suppose then that $m^*(S_n) < \infty$ for all n. For each n, let $\mathcal{C}_n$ be a cover of S_n by open intervals so that

$$\sum_{I\in\mathcal{C}_n} \ell(I) \;\leq\; m^*(S_n) + \frac{\epsilon}{2^n}.$$

Then $\mathcal{C} = \bigcup_{n\in\mathbb{N}} \mathcal{C}_n$ is a cover of S by open intervals, so

$$m^*(S) \leq \sum_{I\in\mathcal{C}} \ell(I) \;\leq\; \sum_{n\in\mathbb{N}} \sum_{I\in\mathcal{C}_n} \ell(I) \;\leq\; \sum_{n\in\mathbb{N}} \left(m^*(S_n) + \frac{\epsilon}{2^n} \right) \;=\; \epsilon + \sum_{n\in\mathbb{N}} m^*(S_n).$$

Since ϵ was arbitrary, statement (3) follows. ■

Lebesgue Measure

We are now ready to define the measurable subsets of $\mathbb{R}$. There are many possible equivalent definitions of measurable sets, and the following definition is known as **Carethéodory's criterion**. It is not very intuitive, and we shall see equivalent definitions of measurability later on that make much more sense. The advantage of Carethéodory's criterion is that it is relatively easy to use from a theoretical perspective, and also it can be generalized to many other settings.

> **Definition: Lebesgue Measure**
> A subset E of $\mathbb{R}$ is said to be **(Lebesgue) measurable** if
>
> $$m^*(T \cap E) + m^*(T \cap E^c) \;=\; m^*(T).$$
>
> for every subset T of $\mathbb{R}$. In this case, the outer measure $m^*(E)$ of E is called the **(Lebesgue) measure** of E, and is denoted $m(E)$.

The arbitrary subset T of $\mathbb{R}$ that appears in the criterion is known as a **test set**. Note that

$$m^*(T \cap E) + m^*(T \cap E^c) \;\geq\; m^*(T)$$

automatically since m^* is subadditive. Thus a set E is Lebesgue measurable if and only if

$$m^*(T \cap E) + m^*(T \cap E^c) \leq m^*(T)$$

for every test set T.

Note also that Carethéodory's criterion is symmetric between E and E^c. Thus a set E is measurable if and only if its complement E^c is measurable.

Proposition 2 Union of Two Measurable Sets

If E and F are measurable subsets of $\mathbb{R}$, then $E \cup F$ is also measurable.

PROOF Let $T \subseteq \mathbb{R}$ be a test set. Since E is measurable, we know that

$$m^*(T) = m^*(T \cap E) + m^*(T \cap E^c). \tag{1}$$

Also, if we use $T \cap (E \cup F)$ as a test set, we find that

$$m^*\big(T \cap (E \cup F)\big) = m^*(T \cap E) + m^*\big(T \cap E^c \cap F\big). \tag{2}$$

Finally, since F is measurable, we know that

$$m^*(T \cap E^c) = m^*(T \cap E^c \cap F) + m^*(T \cap E^c \cap F^c). \tag{3}$$

Combining equations (1), (2), and (3) together yields

$$m^*(T) = m^*\big(T \cap (E \cup F)\big) + m^*(T \cap E^c \cap F^c).$$

Since $E^c \cap F^c = (E \cup F)^c$, this proves that $E \cup F$ is measurable. ■

Corollary 3 Intersection of Two Measurable Sets

If E and F are measurable subsets of $\mathbb{R}$, then $E \cap F$ is also measurable.

PROOF Since E and F are measurable, their complements E^c and F^c is also measurable. It follows that the union $E^c \cup F^c$ is measurable, and the complement of this is $E \cap F$. ■

Proposition 4 Countable Additivity

Let $\{E_k\}$ be a sequence of pairwise disjoint measurable subsets of $\mathbb{R}$. Then the union $\biguplus_{k\in\mathbb{N}} E_k$ is measurable, and

$$m\left(\biguplus_{k\in\mathbb{N}} E_k\right) = \sum_{k\in\mathbb{N}} m(E_k).$$

PROOF Let $T \subseteq \mathbb{R}$ be a test set, and let $U = \biguplus_{k\in\mathbb{N}} E_k$. We wish to show that

$$m^*(T) \geq m^*(T \cap U) + m^*(T \cap U^c).$$

For each $n \in \mathbb{N}$, let $U_n = \biguplus_{k=1}^{n} E_k$. By the Proposition 2, each U_n is measurable, so

$$m^*(T) = m^*(T \cap U_n) + m^*(T \cap U_n^c).$$

But each $U_n \subseteq U$, so $T \cap U_n^c \supseteq T \cap U^c$, and hence

$$m^*(T) \geq m^*(T \cap U_n) + m^*(T \cap U^c).$$

Thus it suffices to show that $m^*(T \cap U_n) \to m^*(T \cap U)$ as $n \to \infty$.

To prove this claim, observe first that

$$\begin{aligned} m^*(T \cap U_k) &= m^*(T \cap U_k \cap E_k) + m^*(T \cap U_k \cap E_k^c) \\ &= m^*(T \cap E_k) + m^*(T \cap U_{k-1}). \end{aligned}$$

for each k. By induction, it follows that

$$m^*(T \cap U_n) = \sum_{k=1}^{n} m^*(T \cap E_k)$$

for each n. Then

$$\sum_{k=1}^{n} m^*(T \cap E_k) = m^*(T \cap U_n) \leq m^*(T \cap U) \leq \sum_{k\in\mathbb{N}} m^*(T \cap E_k),$$

where the last inequality follows from the countable subadditivity of m^*. By the squeeze theorem, we conclude that

$$\lim_{n\to\infty} m^*(T \cap U_n) = m^*(T \cap U) = \sum_{k\in\mathbb{N}} m^*(T \cap E_k),$$

which proves that U is measurable. Moreover, in the case where $T = \mathbb{R}$, the last equation gives

$$m(U) = \sum_{k\in\mathbb{N}} m(E_k).$$

■

Corollary 5 Countable Union of Measurable Sets

If $\{E_k\}$ is any sequence of measurable subsets of $\mathbb{R}$, then the union $\bigcup_{k\in\mathbb{N}} E_k$ is measurable.

PROOF Let $U_n = \bigcup_{k=1}^{n} E_k$ for each k, and let $F_n = U_n - U_{n-1}$, with $F_1 = U_1$. By Proposition 2, we know that each U_n is measurable, and thus $F_n = U_n \cap U_{n-1}^c$ is measurable by Corollary 3. But the sets $\{F_n\}$ are disjoint, and

$$\biguplus_{n\in\mathbb{N}} F_n = \bigcup_{k\in\mathbb{N}} E_k$$

so $\bigcup_{k\in\mathbb{N}} E_k$ is measurable. ■

The Geometry of Intervals

All that remains in proving the desired properties of Lebesgue measure is to show that intervals in $\mathbb{R}$ are measurable, with $m(I) = \ell(I)$ for any interval I. Unlike all of the work so far, proving this requires exploiting the geometry of intervals in a significant way.

We begin with the following proposition.

Proposition 6 Intervals are Measurable

Every interval J in $\mathbb{R}$ is Lebesgue measurable.

PROOF Since each interval in $\mathbb{R}$ is the intersection of two rays, it suffices to prove that each ray in $\mathbb{R}$ is measurable.

Let R be a ray in $\mathbb{R}$, and let $T \subseteq \mathbb{R}$ be a test set. We wish to prove that

$$m^*(T) \geq m^*(T \cap R) + m^*(T \cap R^c)$$

If $m^*(T) = \infty$ then we are done, so suppose that $m^*(T) < \infty$. Let $\epsilon > 0$, and let $\mathcal{C}$ be a cover of T by open intervals so that

$$\sum_{I\in\mathcal{C}} \ell(I) \leq m^*(T) + \frac{\epsilon}{2}.$$

Since the sum $\sum_{I\in\mathcal{C}} \ell(I)$ is finite, $\mathcal{C}$ must be countable (see the appendix on sums). Let $\{I_1, I_2, \ldots\}$ be an enumeration of the elements of $\mathcal{C}$, where we set $I_n = \emptyset$ for

$n > |\mathcal{C}|$ if $\mathcal{C}$ is finite. Then each of the intersections $I_n \cap R$ and $I_n \cap R^c$ is an interval, with
$$\ell(I_n \cap R) + \ell(I_n \cap R^c) \;=\; \ell(I_n).$$

For each n, let J_n and K_n be open intervals containing $I_n \cap R$ and $I_n \cap R^c$, respectively, such that
$$\ell(J_n) \;\leq\; \ell(I_n \cap R) + \frac{\epsilon}{2^{n+2}} \qquad \text{and} \qquad \ell(K_n) \;\leq\; \ell(I_n \cap R^c) + \frac{\epsilon}{2^{n+2}}.$$
Then $\{J_n\}_{n\in\mathbb{N}}$ is a cover of $T \cap R$ by open intervals, and $\{K_n\}_{n\in\mathbb{N}}$ is a cover of $T \cap R^c$ by open intervals, so
$$\begin{aligned} m^*(T \cap R) + m^*(T \cap R^c) \;&\leq\; \sum_{n\in\mathbb{N}} \ell(J_n) + \sum_{n\in\mathbb{N}} \ell(K_n) \\ &\leq\; \sum_{n\in\mathbb{N}} \left(\ell(I_n \cap R) + \frac{\epsilon}{2^{n+2}}\right) + \sum_{n\in\mathbb{N}} \left(\ell(I_n \cap R^c) + \frac{\epsilon}{2^{n+2}}\right) \\ &=\; \frac{\epsilon}{2} + \sum_{n\in\mathbb{N}} \ell(I_n) \;\leq\; m^*(T) + \epsilon. \end{aligned}$$
Since ϵ was arbitrary, this proves the desired inequality. ■

All that remains is to prove that the measure of any interval is equal to its length. For this we need the famous Heine-Borel theorem, which we will state and prove next. Those familiar with point-set topology should recognize this theorem as a special case of the statement that closed intervals in $\mathbb{R}$ are compact. In fact, the notion of compactness in point-set topology arose as a generalization of this theorem.

Theorem 7 Heine-Borel Theorem

Let $[a, b]$ be a closed interval in $\mathbb{R}$, and let $\mathcal{C}$ be a family of open intervals that covers $[a, b]$. Then there exists a finite subcollection of $\mathcal{C}$ that covers $[a, b]$.

PROOF Let S be the set of all points $s \in [a, b]$ for which the interval $[a, s]$ can be covered by some finite subcollection of $\mathcal{C}$. Note that $a \in S$, since the interval $[a, a]$ is just a single point. Our goal is to prove that $b \in S$.

Let $x = \sup(S)$. Since $S \subseteq [a, b]$, we know that $x \in [a, b]$. Therefore, there exists an interval $(c, d) \in \mathcal{C}$ that contains x. Since $c < x$, there is some point $s \in S$ that lies between c and x. Let $\{(c_1, d_1), \ldots, (c_n, d_n)\}$ be a finite subcollection of $\mathcal{C}$ that covers $[a, x]$. Then the collection $\{(c_1, d_1), \ldots, (c_n, d_n), (c, d)\}$ covers $[a, x]$, which proves that $x \in S$.

Moreover, if $x < b$, then there exists an $\epsilon > 0$ such that $x + \epsilon \in [a, b]$ and $x + \epsilon \in (c, d)$. Then the collection $\{(c_1, d_1), \ldots, (c_n, d_n), (c, d)\}$ covers $[a, x + \epsilon]$, which proves that $x + \epsilon \in S$, a contradiction since x is the supremum of S. We conclude that $x = b$, and therefore $b \in S$. ■

In addition to the Heine-Borel theorem, the following proof will use the Riemann integral and characteristic functions. If S is any subset of $\mathbb{R}$, the **characteristic function** (or **indicator function**) for S is the function $\chi_S \colon \mathbb{R} \to \mathbb{R}$ defined by

$$\chi_S(x) = \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases}$$

Note that if I is an interval then

$$\int_{-\infty}^{\infty} \chi_I(x)\, dx = \ell(I).$$

Proposition 8 Measure of an Interval

If J is any interval in $\mathbb{R}$, then $m(J) = \ell(J)$.

PROOF Note first that, for every $\epsilon > 0$, there exists an open interval J' containing J so that $\ell(J') \leq \ell(J) + \epsilon$. Then the singleton collection $\{J'\}$ of open intervals covers J, so

$$m(J) \leq \ell(J') \leq \ell(J) + \epsilon.$$

Since ϵ was arbitrary, it follows that $m(J) \leq \ell(J)$.

Now let $\mathcal{C}$ be any collection of open intervals that covers J. Let $\epsilon > 0$, and let K be a closed subinterval of J such that $\ell(K) \geq \ell(J) - \epsilon$. By the Heine-Borel theorem, there exists a finite subcollection $\{I_1, \ldots, I_n\}$ of $\mathcal{C}$ that covers K. Then

$$\chi_{I_1} + \cdots + \chi_{I_n} \geq \chi_K$$

so

$$\begin{aligned} \sum_{I \in \mathcal{C}} \ell(I) &\geq \ell(I_1) + \cdots + \ell(I_n) = \int_{-\infty}^{\infty} \chi_{I_1}(x)\, dx + \cdots + \int_{-\infty}^{\infty} \chi_{I_n}(x)\, dx \\ &= \int_{-\infty}^{\infty} \big(\chi_{I_1}(x) + \cdots + \chi_{I_n}(x)\big)\, dx \geq \int_{-\infty}^{\infty} \chi_K(x)\, dx = \ell(K) \geq \ell(J) - \epsilon. \end{aligned}$$

Since ϵ was arbitrary, it follows that

$$\sum_{I \in \mathcal{C}} \ell(I) \;\geq\; \ell(J)$$

which proves that $m(J) \geq \ell(J)$. ■

Exercises

1. If $\{E_n\}$ is a sequence of measurable sets, prove that the intersection $\bigcap_{n\in\mathbb{N}} E_n$ is measurable.

2. Prove that if $S \subseteq \mathbb{R}$ and $m^*(S) = 0$, then S is measurable.

3. a) If $E \subseteq F$ are measurable sets, prove that $F - E$ is measurable.
   b) Prove that if $m(E) < \infty$ then $m(F - E) = m(F) - m(E)$.

4. If E and F are measurable sets with finite measure, prove that
$$m(E \cup F) = m(E) + m(F) - m(E \cap F).$$

5. Suppose that $E \subseteq S \subseteq F$, where E and F are measurable. Prove that if $m(E) = m(F)$ and this measure is finite, then S is measurable as well.

6. Prove that every countable subset of $\mathbb{R}$ is measurable and has measure zero.

7. Given a nested sequence $E_1 \subseteq E_2 \subseteq \cdots$ of measurable sets, prove that
$$m\Biggl(\bigcup_{n\in\mathbb{N}} E_n\Biggr) \;=\; \sup_{n\in\mathbb{N}} m(E_n).$$

8. a) Let $E_1 \supseteq E_2 \supseteq \cdots$ be a nested sequence of measurable sets with
$$\bigcap_{n\in\mathbb{N}} E_n = \emptyset.$$
   Prove that if $m(E_1) < \infty$, then $m(E_n) \to 0$ as $n \to \infty$.
   b) Let $E_1 \supseteq E_2 \supseteq \cdots$ be a nested sequence of measurable sets, and suppose that $m(E_1) < \infty$. Prove that
$$m\Biggl(\bigcap_{n\in\mathbb{N}} E_n\Biggr) \;=\; \inf_{n\in\mathbb{N}} m(E_n).$$
   c) Give an example of a nested sequence $E_1 \supseteq E_2 \supseteq \cdots$ of measurable sets such that $m(E_n) = \infty$ for all n but
$$m\Biggl(\bigcap_{n\in\mathbb{N}} E_n\Biggr) \;<\; \infty.$$