Measure Theory: Building up to the Lebesgue Integral

Atharv Chagi

Introduction

One of the most fundamental tools calculus provides is the ability to comprehend sums of an infinite number of arbitrarily small quantities. When it is first taught, definite integration is primarily used to compute areas and volumes. Various ways of calculating area have been around since Ancient Greece, though legitimate integration appeared around the 17th century. The first rigorous definition of an integral was formalized by Bernhard Riemann; this is, most commonly, the first notion of integration students run into. In the 19th century, mathematicians observed limitations in Riemann's integral. As a result, Henri Lebesgue generalized Riemann's integral into a more powerful integral, today known as the Lebesgue integral. In this paper, we examine deficiencies in the Riemann integral, motivate the need for the Lebesgue integral, and go into some depth about Lebesgue measure.

The Riemann Integral

We start with a brief overview of the Riemann Integral.

Suppose f is a bounded function on $[a, b]$. Select a partition of $[a, b]$, $P = \{x_0, x_1, \ldots, x_n\}$ where $a := x_0 < x_1 < \ldots < x_n =: b$.

For each $i \in [1, n]$,

$$M_i := \sup_{x \in [x_{i-1}, x_i]} f(x)$$

For each $i \in [1, n]$,

$$m_i := \inf_{x \in [x_{i-1}, x_i]} f(x)$$

For each $i \in [1, n]$, the **Upper Riemann Sum** $U(f, P, [a, b])$ and **Lower Riemann Sum** $L(f, P, [a, b])$ can be defined as

$$U(f, P, [a, b]) = \sum_{i=1}^{n} (x_i - x_{i-1}) M_i \qquad L(f, P, [a, b]) = \sum_{i=1}^{n} (x_i - x_{i-1}) m_i$$

f is called Riemann Integrable on $[a, b]$ if and only if $L(f, P, [a, b]) = U(f, P, [a, b]) \to L$ as $\|P\| \to 0$ where $\int_a^b f(x)dx = L$.

In the 19th century mathematicians observed several deficiencies in the Riemann Integral.

- Riemann Integrals cannot handle many discontinuities.
- Riemann Integrals cannot handle unbounded functions.
- Riemann Integrals have difficulties with limits.
- Riemann Integrals can't be generalized to abstract spaces due to the need for the partition on the real line.

Limitations of the Riemann Integral

The Dirichlet Function

The Dirichlet Function is often used as a common example to illustrate the deficiences of the Riemann Integral.

Consider a function $D : [0,1] \to \mathbb{R}$.

$$D(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

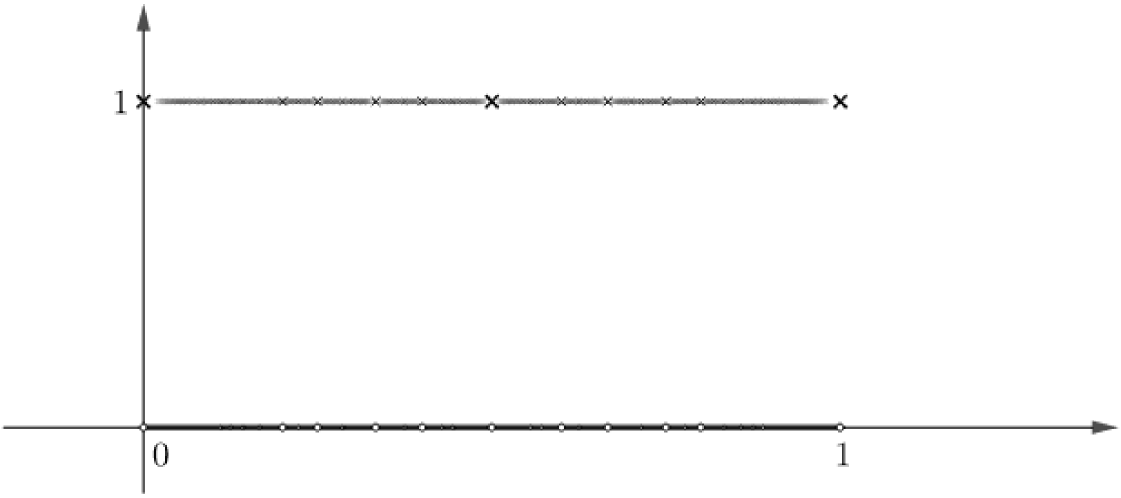

Figure 1: The Dirichlet Function

If $[a,b] \subseteq [0,1]$, then observe that

$$\inf_{x \in [a,b]} f(x) = 0$$

$$\sup_{x \in [a,b]} f(x) = 1$$

for every partition on the real line.

From this result, notice $U(f, P, [a,b]) = 1 \neq L(f, P, [a,b]) = 0$, so we can conclude f is not Riemann integrable. There are much fewer rational numbers than irrational numbers, so in some sense, intuition tells us the integral should be 0. However, there's nothing we can say with the Riemann integral, since f is not Riemann integrable.

In other instances, we can have a sequence of continous bounded functions which are Riemann integrable, but the limit of the sequence converges to a function that is not Riemann integrable.

Let $r_1, r_2, ...$ be a sequence containing each rational number in [0,1]. Define $f_k : [0,1] \to \mathbb{R}$ where $k \in \mathbb{Z}^+$. Consider the sequence of function

$$f_k(x) = \begin{cases} 1 & \text{if } x \in \{r_1, ..., r_k\} \\ 0 & otherwise \end{cases}$$

Notice $f_k(x)$ is Riemann integrable and that $\int_0^1 f_k(x)dx = 0$. This is because $f_k(x)$ is bounded, has a finite number of discontinuities, and is continuous almost everywhere. To actually compute the integral, the integral can be split into its continuous regions and limits can be taken to handle discontinuities. Though, observe the limit of f_k. Define $f : [0, 1] \to \mathbb{R}$.

$$\lim_{k\to\infty} f_k(x) = f$$

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

The sequence of functions converges pointwise to f as the number of rational numbers in the sequence tends to infinity. Despite f being the limit of Riemann integrable functions, f itself is not Riemann integrable. f is exactly the Dirichlet function from the previous example, and is not Riemann integrable.

One way we could potentially go around this fault is by using uniform convergence. Uniform convergence gives that sequence is guaranteed to converge to a Riemann integrable function. The main issue is uniform convergence is restrictive in that we can only work with a certain class of functions. There should be a more suitable method for being able to integrate more functions.

The abilities of the Riemann integral are restricted by its dependence on the real line. We have also seen how the Riemann integral doesn't play well with limits. A lot of analysis relies on the interplay between limits and integrals, so a good theory of integration should permit these interactions (at least with suitable functions).

The Lebesgue Integral

"I have to pay a certain sum, which I have collected in my pocket. I take the bills and coins out of my pocket and give them to the creditor in the order I find them until I have reached the total sum. This is the Riemann integral. But I can proceed differently. After I have taken all the money out of my pocket I order the bills and coins according to identical values and then I pay the several heaps one after the other to the creditor. This is my integral." - Henri Lebesgue

Above is a quote from Lebesgue about his theory of integration in relation to Riemann. What Lebesgue is saying will become clear after an overview of the Lebesgue Integral. Riemann integrals depend on partitions of the real line, but Lebesgue theory permits integration over a set.

For example, we can define a function $f : \mathbb{N} \to \mathbb{R}$, whose Riemann integral doesn't seem to make sense, but the Lebesgue integral can be considered. f can be considered as a sequence: $\{x_n = f(n)\}_{n\in\mathbb{N}}$.

In Riemann theory, we start by defining a partition on the domain. In contrast, Lebesgue integration defines a partition on the y-axis. Suppose f is bounded on $[a, b] \subseteq \mathbb{R}$. On the y axis, create a partition P on $[m, M]$ where

$$m = \inf_{x\in[a,b]} f(x) \qquad\qquad M = \sup_{x\in[a,b]} f(x)$$

$$P = \{y_0, y_1, \ldots, y_n\} \text{ where } m := y_0 < y_1 < \ldots < y_n =: M.$$

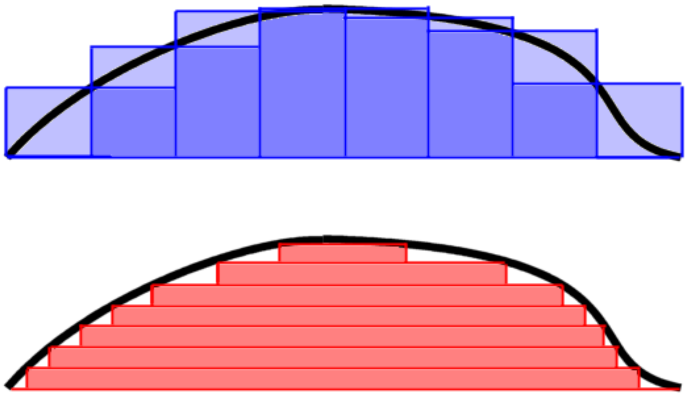

Figure 2: In blue: Riemann partition on the domain. In red: Lebesgue partition on the range.

Riemann integration looks at the lower and upper Riemann sums and takes limits. The way Lebesgue integration equivalently obtains this value is by computing the heights, y_i multiplied by some notion of length of the corresponding set $f^{-1}[y_{i-1}, y_i]$ on the domain. With a properly defined notion of length (measure), the corresponding set does not have to be an interval.

The pressing question now becomes: how is measure of a set defined?

Lebesgue Measure

We want to construct some function, say μ, that acts as notion of length for subsets of $\mathbb{R}$. Intuitively, there's a few things that jump to mind. Suppose A and B are disjoint subsets of $\mathbb{R}$.

We would like if:

- The measure of a union of sets is the sum of measure of individual sets.
  Countable Additivity: $\mu(A \cup B) = \mu(A) + \mu(B)$
- The measure of a point is zero.
  $\mu(\{x\}) = 0$ where x is a point in a set.
- The measure of an interval is its length.
  $\mu([a, b)) = b - a$
- Moving a set around doesn't change it's measure.
  Translation Invariance: $\mu(t + A) = \mu(A)\ \forall t \in \mathbb{R}$

These qualities of μ would be most desirable for all subsets of $\mathbb{R}(\mathcal{P}(\mathbb{R}))$. Unfortunately, μ cannot have all of these properties. In 1905, Vitali constructed a subset of $\mathbb{R}$ which violated these rules. You will often see the theorem written as such:

Theorem. There is no function μ with all of the following properties:

- $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$

- $\mu(I) = \text{length}(I)$
- $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$ where A_i are disjoint subsets of $\mathbb{R}$
- $\mu(t + A) = \mu(A)$ for all $t \in \mathbb{R}$, for all $A \subseteq \mathbb{R}$

It's not possible to define μ on $\mathcal{P}(\mathbb{R})$. It turns out we need to restrict the definition of μ to a particular kind of set, called a σ-algebra.

A **σ-algebra** S is defined as a subset of $\mathcal{P}(\mathbb{R})$, such that :

- $\varnothing \in S$
- If $A \in S$, then $A^c \in S$
- If $A_1, ..., A_n \in S$, then their union is in S.

All of the tools are present now to define μ properly.

If X is a set and S is a σ-algebra, a measurable space is a pair (X, S). The **measure** on a measurable space is the function $\mu : S \to [0, \infty)$ where:

- $\mu(\varnothing) = 0$
- μ exhibits countable additivity for disjoint sets in S.

The Banach Tarski Paradox

Like other paradoxes, the Banach Tarksi paradox is nothing short of mind bending. The Banach Tarski paradox states that it is possible to take a 3D ball, break it into pieces, and reassemble it into 2 identical 3D balls.

In summary, the Banach Tarski construction works as follows:

- Begin with a ball B in 3D space.
- Apply a decomposition to the ball B into a finite number of disjoint subsets.
- Apply transformations and rigid motions to these pieces to obtain two copies of B.

While it seems impossible to consider, the result relies upon the existence of non measurable sets. As iterated previously, Vitali was able to construct a set that broke the properties of μ. There are some subsets of $\mathbb{R}^3$ that are unmeasurable and the heart of the illuision lies in these subsets.

The decomposition breaks B into unmeasurable sets. Upon rearranging these pieces, volume is not necesarrily conserved. Our intuition is useless against unmeasurable sets.

What's important to note is that these abstract constructions in mathematics aren't really reflective of physical manipulations in the real world. Our intuition can only take us so far.

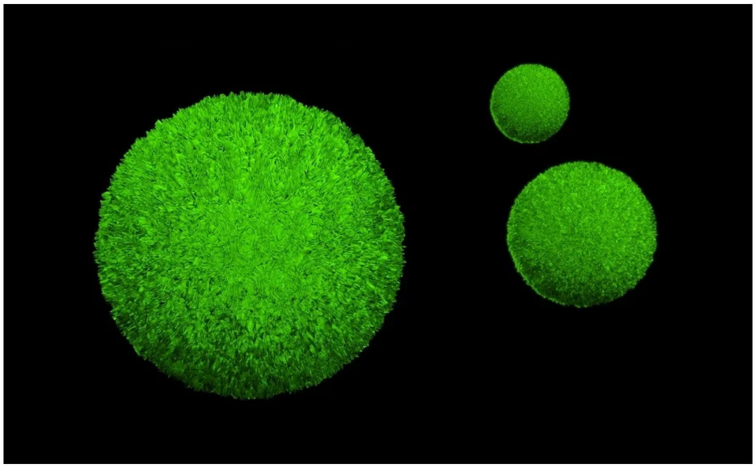

Figure 3: Banach Tarski Paradox Balls.

Conclusion

Measure theory is one of many topics that teaches to let go of intuition momentarily. In terms of applications, measure theory provides an extremely rigorous framework for probability theory when it comes to dealing with random variables and calculating expected values.

The Lebesgue integral is really powerful; it harbors extremely elegant generalizations of integration and allows for the continued growth of fields like probability theory, economics, and machine learning.

The story of Riemann and Lebesgue puts into perspective the importance of accumulated knowledge. Progress is created on the shoulders of giants, and it is important that students keep finding awe in the vast knowledge pool created by history.

References

- Axler, S. (2004). Measure and Integration. Springer Science & Business Media.
- Holmes I. (2019). Limitations of the Riemann Integral. AMUSE Seminar Texas A&M University.
- Hunter, J. K. (2011). Measure Theory . University of California, Davis, Department of Mathematics.
- Kaseorg A. (2007) The Banach-Tarski Paradox. Massachusetts Institute of Technology 18.504 Seminar in Logic.
- Tao, T. (2011). Introduction to Measure Theory. American Mathematical Society.