Notes on the Lebesgue Integral

by
Francis J. Narcowich
November, 2013

1 Introduction

In the definition of the Riemann integral of a function $f(x)$, the x-axis is partitioned and the integral is defined in terms of limits of the Riemann sums $\sum_{j=0}^{n-1} f(x_j^*)\Delta_j$, where $\Delta_j = x_{j+1} - x_j$. The basic idea for the Lebesgue integral is to partition the y-axis, which contains the range of f, rather than the x-axis.

This seems like a "dumb" idea at first. Shouldn't the two ways end up giving the same integral? Most of time this is the case, but Lebesgue was after integrating some functions for which the Riemann integral doesn't exist; for example, the Dirichlet function, which is defined on $[0, 1]$:

$$\chi(x) = \begin{cases} 0 & x \text{ rational,} \\ 1 & x \text{ irrational.} \end{cases}$$

Lebesgue's reasoning was that there were uncountably many irrationals versus countably many rationals, so the "area" should be 1. It is is easy to show that the Riemann integral doesn't exist for χ. The integral Lebesgue came up with not only integrates this function but many more. It also has the property that every Riemann integrable function is also Lebesgue integrable.

Many of the common spaces of functions, for example the square integrable functions on an interval, turn out to complete spaces – Hilbert spaces or Banach spaces – if the Riemann integral is replaced by the Lebesgue integral. The idea of splitting the range (y-axis) rather than the domain (x-axis) turns out to be invaluable in dealing with integrating functions over domains that aren't just real numbers. Such integrals arise in many fields, probability theory for instance. Lebesgue's idea turns out to be a *brilliant* "dumb" idea.

We now turn to the technical details involved in the Lebesgue integral, starting with Lebesgue sums. Choose an increasing sequence of points $\mathcal{P} = \{c \le y_0 < y_1 < y_2 < \cdots < y_n \le d\}$, where the range of f is contained in $[c, d]$. As usual, we set $\|\mathcal{P}\| := \max_{0\le j\le n-1}(y_{j+1} - y_j)$. Let $E_j = \{x \in [a, b] : y_j \le f(x) < y_{j+1}\} = f^{-1}([y_j, y_{j+1}))$ and choose a point y_j^* from each interval

$[y_j, y_{j+1}]$. (Note that we can have $y_j^* = y_{j+1}$.) The corresponding Lebesgue sum is

$$L_{\mathcal{P},Y^*}(f) := \sum_{j=0}^{n-1} y_j^* \mu(E_j) \tag{1}$$

where $\mu(E_j)$ denotes the "measure" or "length" of the set E_j and $Y^* = \{y_j^*\}_{j=0}^{n-1}$ For this sum to make sense, we need a concept of measure for more sets than just intervals. For example, $\chi^{-1}([1/2, 3/2))$ is the set of all irrational numbers between 0 and 1. It doesn't contain any intervals at all. This leads to the question of how to extend the concept of measure to subsets of the real line that are much more complicated than simple intervals.

2 Measurable Sets

The ordinary idea of the measure/length of an interval can be thought of as a function that assigns to an interval a nonnegative number. In addition, μ satisfies standard properties; for example, if I and J are disjoint intervals, then $\mu(I \cup J) = \mu(I) + \mu(J)$.

An appropriate generalization of μ would satisfy the same properties, but for a wider class of subsets of $[a, b]$ than just the collection of intervals. Suppose that Σ is a collection of subsets of $[a, b]$ and that we have a function $\mu : \Sigma \to \mathbb{R}$. This is called a *set function*, because its domain consists of subsets of $[a, b]$. To go further and require that μ satisfies the same properties as those of length on intervals, we have to put conditions on Σ. For example, if we want μ to satisfy $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A \cap B = \emptyset$, then $A \cup B$ must be in Σ. Here is a list of conditions that we will require Σ to satisfy:

1. Σ is *non-empty*. It will always contain both $\emptyset$ and $[a, b]$.
2. Σ is *closed under complementation*: $A \in \Sigma$ if and only if $A^{\complement} \in \Sigma$.
3. Σ is *closed under countable unions*: $\{A_j \in \Sigma\}_{j=1}^{\infty}$ then $\cup_{j=1}^{\infty} A_j \in \Sigma$.

A collection of subsets Σ that satisfies that satisfies these conditions is called a *σ-algebra*.

We can now specify the properties that a set function $\mu : \Sigma \to [0, \infty)$ requires so that it mimics length for intervals. We say that μ is a (σ-finite) *measure* if and only if

1. $\mu([a,b]) = b - a$ and $\mu(\emptyset) = 0$.
2. Non-negativity: $\mu(A) \geq 0 \; \forall A \in \Sigma$.
3. Monotonicity: If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.
4. Countable Additivity: If $\{A_j \in \Sigma\}_{j=1}^{\infty}$, with $A_i \cap A_j = \emptyset$, $i \neq j$, then $\mu(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j)$.

These are fairly general conditions and they form the basis of the field of measure theory. Our aim here is simply to construct the Lebesgue measure. We note that the most natural choice for the class Σ would be the one comprising *all* subsets of $[a,b]$. Unfortunately, a theorem whose proof employs the axiom of choice shows that there is no measure for this class. Instead, we will use a procedure that simultaneously constructs μ and Σ.

We start with the outer measure of a set, which can be defined for arbitrary subsets of $[a,b]$. Every open set G in $[a,b]$ is the disjoint union of intervals, $G = \cup_i (a_i, b_i)$ and we define the outer measure to be $\mu^*(G) = \sum_i (b_i - a_i)$. This is of course the natural generalization of length to open sets.

We now turn to the general case. Let $A \subseteq [a,b]$ for some finite $a \neq b \in \mathbb{R}$. We define the *outer measure* $\mu^*(A)$ by

$$\mu^*(A) = \inf\{\mu^*(G) : A \subset G \, , G \text{ is open in } [a,b]\}.$$

We can also define the inner measure $\mu_*(A)$:

$$\mu_*(A) = b - a - \mu^*([a,b] \backslash A).$$

If $\mu_*(A) = \mu^*(A)$, then we say that A is *Lebesgue measurable*. The class Σ is then just defined as all the collection of all Lebesgue measurable sets and the Lebesgue measure of $A \in \Sigma$ is defined as $\mu(A) := \mu_*(A) = \mu^*(A)$. It is not hard to show that all open sets and all closed sets are measurable, and that if A is measurable, so is its complement $A^{\complement}$. With more work one can show that the measurable sets form a σ-algebra and the Lebesgue measure is a non-negative measure defined on Σ. When we are dealing only with intervals, the Lebesgue measure coincides with the usual concept of length.

It is always true that both the inner and outer measure are positive and that $\mu^*(A) \geq \mu_*(A)$, so that to verify A is measurable only requires showing that $\mu^*(A) \leq \mu_*(A)$. For instance, a set has measure 0 if and only if, for every $\epsilon > 0$, there is an open set $G \supset A$ such that $\mu(G) < \epsilon$. Here is a specific case.

Example 1. Every countable set has measure 0.

Proof. Let $A = \{x_1, x_2, \ldots, x_n, \ldots\}$. Let $I_j = (x_j - \frac{\epsilon}{2^{j+1}}, x_j + \frac{\epsilon}{2^{j+1}})$. Then, we see that $A \subset \cup I_j$, and hence

$$\begin{aligned}\mu^*(A) &= \inf\{\mu^*(G) : A \subset G\,, G \text{ is open}\} \\ &\leq \mu^*(\cup_j I_j) \\ &= \sum_j (\frac{\epsilon}{2^{j+1}} + x_j - x_j + \frac{\epsilon}{2^{j+1}}) \\ &= \sum_j \frac{\epsilon}{2^j} = \epsilon\end{aligned}$$

Since this holds for arbitrary $\epsilon > 0$, we see that $\mu^*(A) = 0$. □

3 The Lebesgue Integral

We briefly introduce the concept of the Lebesgue integral in this section and discuss some important theorems for convergence. The Lebesgue integral on a measurable set A will be denoted by either $\int_A f(x)\,dx$ or $\int_A f(x)\,d\mu(x)$. We say that a function $f : A \to \mathbb{R}$ is *measurable* if for every measurable $E \subset \mathbb{R}$, $f^{-1}(E)$ is a measurable subset of A. This is analogous to the concept of continuity, where a function is said to be continuous if the pre-image of any open set is open. Let A be a measurable set. We define the *characteristic* function of A (or *indicator* function of A) by

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

We say a function is simple if it is a finite linear combination of characteristic functions. We define the Lebesgue integral for simple functions by

$$\int_E s\,d\mu = \int_E \sum_{j=1}^{N} c_j \chi_{A_j}\,d\mu = \sum_{j=1}^{n} c_j \mu(A \cap E_j).$$

For non-negative measurable functions, the Lebesgue integral on a measurable set A can be defined by

$$\int_A f\,d\mu = \sup\{\int_A s\,d\mu : 0 \leq s \leq f \text{ and } s \text{ is simple}\}.$$

We say that a function f is integrable on A if $\int_A |f| d\mu < \infty$. Let A, B denote measurable sets of finite measure.

1. If f is a bounded and measurable function, then the integral exists.
2. Integration is a linear operation. That is, if f and g are both integrable and $a, b \in \mathbb{R}$, then $\int_A (af + bg)\, d\mu = a \int_A f\, d\mu + b \int_A g\, d\mu$.
3. If $\int_A f(x)^2\, dx$ and $\int_A g(x)^2\, dx$ both exist, then $\int_A f(x)g(x)\, dx$ and $\int_A (f(x)+g(x))^2\, dx$ exist.
4. $\int_{A\cup B} f\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu$, where $\mu(A \cap B) = 0$.
5. If $f = g$ almost everywhere, then $\int_A f\, d\mu = \int_A g\, d\mu$.
6. If the Riemann integral of f exists, then the Lebesgue integral exists and the integrals are equal.

Theorem 1. *(Monotone Convergence Theorem): Let $\{f_j\}$ be a collection of measurable functions on A that satisfy $0 \le f_1 \le f_2 \le \cdots \le f_n \le \ldots$ almost everywhere. Define f as the pointwise limit $f(x) = \lim f_j(x)$. Then, f is measurable and*

$$\lim \int_A f_j\, d\mu = \int_A f\, d\mu.$$

Theorem 2. *(Dominated Convergence Theorem): Let $\{f_j\}$ be a set of measurable functions that converges pointwise to a function f and assume there exists an integrable function g such that $|f_j(x)| \le g(x)$ almost everywhere. Then, f is integrable and*

$$\lim \int_A f_j\, d\mu = \int_A f\, d\mu.$$

Theorem 3. *(Fubini's Theorem): Let f be measurable on $A \times B$. If*

$$\int_{A\times B} |f(x,y)|\, d(x,y) < \infty$$

then $\int_A \int_B f(x,y)\, d\mu(x) d\mu'(y)$ exists and the order of integration may be switched.

We define the space $L^p([a,b])$ to be the space of p integrable functions for $1 \leq p < \infty$. That is,

$$L^p([a,b]) = \{f : \int_a^b |f(x)|^p < \infty\}.$$

The space $L^p([a,b])$ is a normed space when it is equipped with the norm

$$\|f\|_p = (\int_a^b |f(x)|^p d\mu(x))^{\frac{1}{p}}.$$

For the case $p = \infty$, we define

$$\|f\|_\infty = \text{ ess sup}|f| = \inf\{a \in \mathbb{R} : \mu(\{x : |f|(x) > a\}) = 0\}$$

and we define

$$L^\infty([a,b]) = \{f : \|f\|_\infty < \infty\}.$$

Example 2. Just because a function is in L^∞ doesn't necessarily mean that it is bounded in the usual sense. Here is an example of this.

$$f(t) = \begin{cases} 1 & t \in \mathbb{R}\backslash\mathbb{Q} \\ q & t = \frac{p}{q} \end{cases} \tag{2}$$

where $\frac{p}{q}$ is reduced to lowest terms and $q > 0$. This function is unbounded, but

$$\|f\|_\infty = \text{ ess sup}|f(x)| = 1$$

because the set of rational numbers has measure zero. This implies that, although f is unbounded, $f \in L^\infty$.

Theorem 4. *The space $L^p([a,b])$ is a complete space for $1 \leq p \leq \infty$.*

This theorem implies that $L^p([a,b])$ is a Banach space. In the special case of $p = 2$, $L^2([a,b])$ is a Hilbert space.

4 Examples

Example 3. Suppose $f_n(x) = \frac{1}{\sqrt{x+\frac{1}{n}}}$ for $x \in [0,1]$ and $n \geq 1$. Let f denote the pointwise limit of f_n. Show that f is integrable.

Proof. We see that $f_{n+1}(x) = \frac{1}{\sqrt{x+\frac{1}{n+1}}} \geq \frac{1}{\sqrt{x+\frac{1}{n}}} = f_n(x)$ holds for all n and $\lim_{n\to\infty} f_n(x) = \frac{1}{\sqrt{x}}$ almost everywhere. Then, by the monotone convergence theorem, $\frac{1}{\sqrt{x}}$ is integrable and

$$\lim_{n\to\infty} \int_0^1 \frac{dx}{\sqrt{x+\frac{1}{n}}} = \int_0^1 \frac{dx}{\sqrt{x}}.$$

□

Example 4. Define

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\, dt \tag{3}$$

where $f \in L^1(\mathbb{R})$. Then, $\hat{f}$ is a continuous function.

Proof. By a straightforward computation, we see that

$$\begin{aligned}\hat{f}(\omega+h) - \hat{f}(\omega) &= \int_{-\infty}^{\infty} f(t)(e^{-i(\omega+h)t} - e^{-i\omega t})\, dt \\ &= \int_{-\infty}^{\infty} f(t)e^{-i\omega t}(e^{-iht} - 1)dt\end{aligned}$$

The integrand is bounded above by $2|f(t)|$ almost everywhere, and since $f \in L^1$, $2|f(t)|$ is integrable. Then, an application of the dominated convergence theorem implies that $\lim_{h\to 0} \hat{f}(\omega+h) = \hat{f}(\omega)$ which implies that $\hat{f}$ is continuous. □