3 Lebesgue integral

In the simplest case, the (Riemann) integral of a non-negative function can be regarded as the area between the graph of that function and the x-axis. Lebesgue integration is a mathematical construction that extends the notion of the integral to a larger class of functions; it also extends the domains on which these functions can be defined. As such, the Lebesgue integral plays an important role in real analysis, probability, and many other areas of mathematics.

3.1 Integration: Riemann vs. Lebesgue

As part of the general movement towards rigour in mathematics in the nineteenth century, attempts were made to put the integral calculus on a firm foundation. The Riemann integral [16] is one of the most widely known examples; its definition starts with the construction of a sequence of easily-calculated integrals which converge to the integral of a given function. This definition is successful in the sense that it gives the expected answer for many already-solved problems, and gives useful results for many other problems.

However, despite the Riemann integral is naturally linear and monotone,[17] it does not interact well with taking limits of sequences of functions, making such limiting functions difficult to analyse (and integrate).[18] The Lebesgue integral is easier to deal with when taking limits under the integral sign; it also allows to calculate integrals for a broader class of functions. For example, the Dirichlet function, which is 0 where its argument is irrational and 1 otherwise, is Lebesgue-integrable, but not Riemann-integrable.

3.1.1 Riemann integral

Recall that a partition of an interval $[a, b]$ is a finite sequence

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b\,.$$

Each $[x_i, x_{i+1}]$ is called a sub-interval of the partition. The mesh of a partition is defined to be the length of the longest sub-interval $[x_i, x_{i+1}]$, that is, it is $\max(x_{i+1} - x_i)$ where $0 \le i \le n-1$.

Let f be a real-valued function defined on the interval $[a, b]$. The Riemann sum of f with respect to the partition $x_0, \ldots, x_n$ is

$$\sum_{i=0}^{n-1} f(t_i)(x_{i+1} - x_i)\,,$$

where each t_i is a fixed point in the sub-interval $[x_i, x_{i+1}]$. Notice that the last expression is the sum of areas of rectangles with heights $f(t_i)$ and lengths $x_{i+1} - x_i$.

[16] proposed by Bernhard Riemann (1826-1866);

[17] see the slides!

[18] This is of prime importance, for instance, in the study of Fourier series, Fourier transforms and other topics.

Loosely speaking, the Riemann integral of f is the limit of the Riemann sums of f as the partitions get finer and finer (ie. the mesh goes to zero), and every function f for which this limit does not depend on the approximating sequence is called integrable.

3.1.2 Lebesgue integral: sketch of the construction

The modern approach to the theory of Lebesgue integration has two distinct parts:
a) a theory of measurable sets and measures on these sets;
b) a theory of measurable functions and integrals on these functions.

Measure theory initially was created to provide a detailed analysis of the notion of length of subsets of the real line and more generally area and volume of subsets of Euclidean spaces. In particular, it provided a systematic answer to the question of which subsets of $\mathbb{R}$ have a length. As was shown by later developments in set theory, it is actually impossible to assign a length to all subsets of $\mathbb{R}$ in a way which preserves some natural additivity and translation invariance properties. This suggests that picking out a suitable class of measurable subsets is an essential prerequisite.

The modern approach to measure and integration is axiomatic. One defines a measure as a mapping μ from a σ-field $\mathcal{A}$ of subsets of a set E, which satisfies a certain list of properties.[19] These properties can be shown to hold in many different cases.

Integration. In the Lebesgue theory, integrals are limited to a class of functions called measurable functions. Let E be a set and let $\mathcal{A}$ be a σ-field of subsets[20] of E. A function $f : E \to \mathbb{R}$ is measurable if the pre-image of any closed interval $[a,b] \subset \mathbb{R}$ is in $\mathcal{A}$, $f^{-1}([a,b]) \in \mathcal{A}$. The set of measurable functions is naturally closed under algebraic operations; in addition (and more importantly) this class is closed under various kinds of point-wise sequential limits, eg., if the sequence $\{f_k\}_{k\in\mathbb{N}}$ consists of measurable functions, then both

$$\liminf_{k\in\mathbb{N}} f_k \qquad \text{and} \qquad \limsup_{k\in\mathbb{N}} f_k$$

are measurable functions.

Let a measure space $(E, \mathcal{A}, \mu)$ be fixed. The Lebesgue integral $\int_E f\,d\mu$ for measurable functions $f : E \to \mathbb{R}$ is constructed in stages:

Indicator functions: If $S \in \mathcal{A}$, ie., the set S is measurable, we define the integral of its indicator function[21] $\mathbb{1}_S$ via

$$\int \mathbb{1}_S \, d\mu = \mu(S)\,.$$

[19] see the slides!
[20] one often calls $(E, \mathcal{A})$ a measurable space, and $(E, \mathcal{A}, \mu)$ a measure space;
[21] recall that $\mathbb{1}_S(x) = 1$ if $x \in S$ and $\mathbb{1}_S(x) = 0$ otherwise

Simple functions: for non-negative simple functions, ie., linear combinations of indicator functions $f = \sum_k a_k \mathbb{1}_{S_k}$ (where the sum is finite and all $a_k \geq 0$), we use linearity to define [22]

$$\mu(f) \equiv \int \Bigl(\sum_k a_k \mathbb{1}_{S_k}\Bigr) d\mu = \sum_k a_k \int \mathbb{1}_{S_k} d\mu = \sum_k a_k \mu(S_k)\,,$$

This construction is obviously linear and monotone.[23] Moreover, even if a simple function can be written as $\sum_k a_k \mathbb{1}_{S_k}$ in many ways, the integral will always be the same.[24]

Non-negative functions: Let $f : E \to [0, +\infty]$ be measurable. We put

$$\int_E f\, d\mu := \sup\Bigl\{ \int_E h\, d\mu : h \leq f, 0 \leq h \text{ simple} \Bigr\}$$

We need to check whether this construction is consistent, ie., if $0 \leq f$ is simple we need to verify whether this definition coincides with the preceding one. Another question is: if f as above is Riemann-integrable, does this definition give the same value of the integral? It is not hard to prove that the answer to both questions is yes.

Clearly, if $f : E \to [0, +\infty]$ is any measurable function, its integral $\int f\, d\mu$ may be infinite.

Signed functions: If $f : E \to [-\infty, +\infty]$ is measurable,[25] we decompose it into the positive and negative parts, $f = f^+ - f^-$, where

$$f^+(x) = \begin{cases} f(x) & \text{if} \quad f(x) > 0\,, \\ 0 & \text{otherwise}\,, \end{cases} \qquad f^-(x) = \begin{cases} -f(x) & \text{if} \quad f(x) < 0\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

Note that the functions $f^+ \geq 0$ and $f^- \geq 0$ satisfy $|f| = f^+ + f^-$. If $\int |f|\, d\mu$ is finite, then f is called Lebesgue integrable. In this case, both integrals $\int f^+\, d\mu$ and $\int f^-\, d\mu$ converge, and it makes sense to define

$$\int f\, d\mu = \int f^+\, d\mu - \int f^- d\mu\,.$$

It turns out that this definition gives the desirable properties of the integral, namely, linearity, monotonicity and regularity when taking limits. The functions, which can be obtained from the above construction, are called Borel functions.[26] The class of Borel functions is very big and sufficient for most practical considerations.[27]

[22] here we always assume that $0 \cdot \infty = \infty \cdot 0 = 0$;

[23] see the slides!

[24] Also, if any two functions f_1 and f_2 coincide almost everywhere, ie., they differ on a set of measure zero, $\mu(x : f_1(x) \neq f_2(x)) = 0$, their integrals are equal, $\mu(f_1) = \mu(f_2)$.

[25] Complex valued functions can be similarly integrated, by considering the real part and the imaginary part separately.

[26] by definition $f : E \to [-\infty, +\infty]$ is Borel, if for every $a \in \mathbb{R}$, $\{x \in E : f(x) \leq a\} \in \mathcal{A}$, ie., is measurable.

[27] it is not easy to construct a non-Borel real-valued function; get in touch, if interested!

3.2 Lebesgue integral: limiting results

The construction described above implies the following limiting property, which is one of the most central in the area:

✍ **Theorem 3.1** (Monotone Convergence Theorem; (MON)). *Let f and $(f_n)_{n\geq1}$ be Borel functions on $(E,\mathcal{A},\mu)$ such that $0\leq f_n\nearrow f$. Then, as $n\to\infty$,*

$$\mu\bigl(f_n\bigr)\nearrow\mu\bigl(f\bigr)\leq\infty\,.$$

The "random variables" version of the result is:

✍ **Theorem 3.2** (Monotone Convergence Theorem; (MON)). *If random variables $X_n\geq0$ are such that $X_n\nearrow X$ as $n\to\infty$, then $\mathsf{E}(X_n)\nearrow\mathsf{E}(X)\leq\infty$ as $n\to\infty$.*

In view of the footnote 24 above, the following result is rather natural:

Corollary 3.3. *Let f and $(f_n)_{n\geq1}$ be non-negative Borel functions on $(E,\mathcal{A},\mu)$ such that, except on a μ-null set N, $0\leq f_n\nearrow f$, i.e.,*

$$^{\forall}x\in E\setminus N,\quad f_n(x)\nearrow f(x)\qquad\text{and}\qquad\mu(N)=0\,.$$

Then $\mu\bigl(f_n\bigr)\nearrow\mu\bigl(f\bigr)\leq\infty$ as $n\to\infty$.

Exercise 3.4. State an analogue of the previous corollary for random variables (using almost sure convergence).

Another important result is

✍ **Theorem 3.5** (Dominated-Convergence Theorem; (DOM)). *Let $(f_n)_{n\geq1}$ and f be Borel functions on $(E,\mathcal{A},\mu)$ such that $f_n(x)$ converges to $f(x)$ for* all *$x\in E$ as $n\to\infty$ and such that the sequence $f_n(x)$ is* dominated *by a non-negative integrable function g, i.e., for all $x\in E$ and $n\in\mathbb{N}$,*

$$f_n(x)\to f(x)\qquad\text{and}\qquad\bigl|f_n(x)\bigr|\leq g(x)\quad\text{with}\quad\mu(g)<\infty\,.\tag{3.1}$$

Then $\mu(f_n)\to\mu(f)$ as $n\to\infty$.

✍ **Theorem 3.6** (Dominated-Convergence Theorem; (DOM)). *Let $(X_n)_{n\geq1}$ and X be random variables such that for* all *$\omega\in\Omega$, we have $X_n(\omega)\to X(\omega)$ as $n\to\infty$. If there is a random variable $Y\geq0$ such that $\mathsf{E}(Y)<\infty$, and for all $\omega\in\Omega$, $\bigl|X_n(\omega)\bigr|\leq Y(\omega)$, then $\mathsf{E}(X_n)\to\mathsf{E}(X)$ as $n\to\infty$.*

Of course, similarly to the corollary above, one can allow the conditions (3.1) to be violated on a set N of measure zero.

Exercise 3.7. State the versions of the last two theorems in the case convergence is violated on a set of measure zero (ie., convergence takes place almost surely).

Various examples of application of these results were discussed in the lectures and tutorials.