What is integration?

April 24, 2012

Integration should capture the intuitive ideas of what the area or volume under the graph of a function is. I.e., given a suitable function $f : \mathbb{R}^n \to \mathbb{R}^{\geq 0}$, we get a subspace of $\mathbb{R}^{n+1}$ of points (x, y) where $x \in \mathbb{R}^n$ and $0 \leq y \leq f(x)$. The integral should measure the volume of this subspace and should naturally extend to suitable functions $f : \mathbb{R}^n \to \mathbb{R}$.

Formally, we want an integral to be an operator $\int : \mathcal{F} \to \mathbb{R}$ that is defined on a subspace $\mathcal{F}$ of the space of all functions $\mathbb{R}^n \to \mathbb{R}$. The operator assigns to a function f in $\mathcal{F}$ a real number denoted by $\int f$. Given an integral, we call such a function f integrable with respect to that integral. We want $\int$ to fulfill certain properties:

1. Linearity: If $\int f$ and $\int g$ exists (i.e., f and g are in $\mathcal{F}$), then $\int f + g$ and $\int \lambda f$ exists and
$$\int f + g = \int f + \int g, \int \lambda f = \lambda \int f.$$
2. Monotonicity: If $\int f$ and $\int g$ exists and
$$f \leq g, \qquad \text{then} \qquad \int f \leq \int g.$$
(Here, $f \leq g$ means $f(x) \leq g(x)$ for all $x \in \mathbb{R}^n$. $f \leq g$ induces a partial order on $\mathcal{F}$)
3. Normalized: Let $\chi_{[0,1]^n} : \mathbb{R}^n \to \mathbb{R}$ defined by
$$\chi_{[0,1]^n}(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } 0 \leq x_i \leq 1 \text{ for all } i \\ 0 & \text{otherwise} \end{cases}.$$
Then $\int \chi_{[0,1]^n}$ exists and
$$\int \chi_{[0,1]^n} = 1.$$

4. Translation invariance: Let f defined by $f(x) = g(x + \lambda)$. If g exists, then
$$\int f = \int g.$$
5. Scaling property: Let f defined by $f(x) = g(\lambda x)$ with $\lambda > 0$. If $\int g$ exists, then
$$\int f = 1/\lambda^n \int g.$$
6. Approximations by upper and lower functions: If
$$f_1 \leq f_2 \leq \cdots \leq f_i \leq \cdots \leq f \leq \cdots \leq g_i \leq \cdots \leq g_2 \leq g_1$$
has the property that $\int f_i$ and $\int g_i$ exist and
$$\lim_{i\to\infty} \int f_i = \lim_{i\to\infty} \int g_i,$$
then $\int f$ exists and
$$\int f = \lim_{i\to\infty} \int f_i.$$
7. Countable sums preserved (Optional): If $\int f_i$ and $\int |f_i|$ exists and $\sum \int |f_i| < \infty$, then $\int \sum f_i$ exists and
$$\int \sum f_i = \sum \int f_i.$$
Note: Thinking of two functions f and g as the same if $\int f - g = 0$, this implies that the space $\mathcal{F}$ with the metric $d(f, g) = \int |f - g|$ is a complete metric space.

Definition 1. *Given a subset U of $\mathbb{R}^n$, we can define the characteristic function $\chi_U : \mathbb{R}^n \to \mathbb{R}$ by $\chi_U(x) = 1$ if x in U and $\chi_U(x) = 0$ if x is not in U.*

There are two applications of the characteristic function:

Definition 2. *Given an integral $\int$, a subset U of $\mathbb{R}^n$ is measurable if $\int \chi_U$ exists. We call $\int \chi_U$ the volume or measure of U. If $\int \chi_U$ does not exists, U is not measurable. Thus, an integral $\int$ defines a measure.*

Definition 3. *If U is a measurable subset U of $\mathbb{R}^n$, then $\int_U f$ is defined as $\int \chi_U f$ if the integral exists.*

Definition 4. *A function $f : \mathbb{R}^n \to \mathbb{R}$ has compact support if the set $\{x : f(x) \neq 0\}$ is bounded.*

Note 1. *One can prove (Tarski Paradox) that it is impossible to define an integral that works for all bounded functions with compact support.*

There are two major flavors of integrals and induced measures: Riemann-integral and Lebesgue-integral. The space of all Riemann-integrable functions is the smallest space closed under properties 1 to 6. The Riemann-integral is the unique integral defined on all Riemann-integrable functions that fulfills properties 1 to 6. Similarly, for Lebesgue-integrable functions, the Lebesgue-integral, and properties 1 to 7. In particular, every Riemann-integrable function is Lebesgue-integrable. Of course, one has too show that these operators $\int : \mathcal{F} \to \mathbb{R}$ really exists, are well-defined, and unique.

This definition of the Riemann-integral and Lebesgue-integral and their induced measures is a bit abstract. Luckily, there are more concrete definitions. Let's start with the notion of a set having "no volume", for example, the surface of a ball has no "width", so no volume.

Definition 5. *Let $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$ be a generalized rectangle. The volume of I is defined by $\mathrm{vol}(I) = (b_1 - a_1) \cdots (b_n - a_n)$.*

Definition 6. *A set U has Riemann measure zero (Jordan content zero), if for every $\epsilon > 0$ there is a collection of finitely many generalized rectangles $I_1, \ldots, I_k$ such that they cover U and such that $\sum \mathrm{vol}(I_i) < \epsilon$.*

Definition 7. *A set U has Lebesgue measure zero, if for every $\epsilon > 0$ there is a collection of countably many generalized rectangles $I_1, \ldots$ such that they cover U and such that $\sum \mathrm{vol}(I_i) < \epsilon$.*

Definition 8. *The k-standard partition of $\mathbb{R}^n$ is the partition P by generalized rectangles of the form $[2^{-k}x_1, 2^{-k}(x_i + 1)] \times \cdots \times [2^{-k}x_n, 2^{-k}(x_n + 1)]$ where x_i are integers. Note that the volume of each generalized rectangle is 2^{-kn}.*

Definition 9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a bounded function with compact support. For each generalized rectangle I in the k-standard partition P, pick the infimum, respectively, supremum of f over that generalized rectangle. The sum of all the infimums, respectively, supremums multiplied by the volume I is called lower, respectively, upper Darboux sum.*

Theorem 1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a bounded function with compact support. Then f is Riemann integrable if and only if the lower and upper Darboux sum converges to the same value as $k \to \infty$. In this case, the Riemann integral $\int f$ is equal to the two limits.*

Theorem 2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then f is Riemann integrable if and only if f has compact support, is bounded, and the set where f is not continuous has Lebesgue measure 0.*