The Riemann Integral

I know of some universities in England where the Lebesgue integral is *taught* in the first year of a mathematics degree instead of the Riemann integral, but I know of no universities in England where students *learn* the Lebesgue integral in the first year of a mathematics degree. (Approximate quotation attributed to T. W. Körner)

Let $f : [a, b] \to \mathbb{R}$ be a bounded (not necessarily continuous) function on a compact (closed, bounded) interval. We will define what it means for f to be Riemann integrable on [a, b] and, in that case, define its Riemann integral

$$\int_{a}^{b} f(x) \, dx.$$

The integral of f on [a, b] is a real number whose geometrical interpretation is the signed area under the graph y = f(x).

The Riemann integral is the simplest integral to define, and it allows one to integrate every continuous function as well as some not-too-badly discontinuous functions. It isn't the only integral and there are many others, the most important of which is the Lebesgue integral. The Lebesgue integral allows one to integrate unbounded or highly discontinuous functions whose Riemann integrals do not exist, and it has better mathematical properties than the Riemann integral. The definition of the Lebesgue integral, however, requires the use of measure theory, which we will not describe here. In any event, the Riemann integral is adequate for many purposes, and even if one needs the Lebesgue integral, it's better to understand the Riemann integral first.

1.1. Definition of the Riemann integral

We say that two intervals are almost disjoint if they are disjoint or intersect only at a common endpoint. For example, the intervals [0, 1] and [1, 3] are almost disjoint, whereas the intervals [0, 2] and [1, 3] are not.

Definition 1.1. Let *I* be a compact interval of nonzero length. A partition of *I* is a finite collection $\{I_k : k = 1, ..., n\}$ of almost disjoint, compact subintervals of nonzero length whose union is *I*.

A partition of [a, b] with subintervals $I_k = [x_{k-1}, x_k]$ is determined by the set of endpoints of the intervals

$$a = x_0 < x_1 < \dots < x_n = b.$$

Abusing notation, we will denote a partition P either by its intervals

$$P = \{I_1, I_2, \dots, I_n\}$$

or by its endpoints

$$P = \{x_0, x_1, \dots, x_n\}.$$

We'll adopt either notation as convenient; the context should make it clear which one is being used.

Example 1.2. The set of intervals

 $\{[0, 1/5], [1/5, 1/4], [1/4, 1/3], [1/3, 1/2], [1/2, 1]\}$

is a partition of [0, 1]. The corresponding set of endpoints is

$$\{0, 1/5, 1/4, 1/3, 1/2, 1\}.$$

We denote the length of an interval I = [a, b] by |I| = b - a. Note that the sum of the lengths $|I_k| = x_k - x_{k-1}$ of the almost disjoint subintervals in a partition $\{I_k : k = 1, ..., n\}$ of an interval I is equal to length of the whole interval. This is obvious geometrically; algebraically, it follows from the telescoping series

$$\sum_{k=1}^{n} |I_k| = \sum_{k=1}^{n} (x_k - x_{k-1})$$

= $x_n - x_{n-1} + x_{n-1} - x_{n-2} + \dots + x_2 - x_1 + x_1 - x_0$
= $x_n - x_0$
= $|I|$.

Suppose that $f:[a,b]\to \mathbb{R}$ is a bounded function on the compact interval I=[a,b] with

$$M = \sup_{T} f, \qquad m = \inf_{T} f.$$

If $P = \{I_k : k = 1, ..., n\}$ is a partition of I with endpoints $\{x_0, x_1, ..., x_n\}$, let

$$M_k = \sup_{I_k} f, \qquad m_k = \inf_{I_k} f.$$

These suprema and infima are well-defined and finite since f is bounded. Moreover,

$$m \le m_k \le M_k \le M$$

If f is continuous on I, then it is bounded and attains its maximum and minimum values on each interval, but a bounded discontinuous function need not attain its supremum or infimum.

We define the upper Riemann sum of f with respect to the partition P by

$$U(f; P) = \sum_{k=1}^{n} M_k |I_k|$$

= $\sum_{k=1}^{n} M_k (x_k - x_{k-1})$

and the lower Riemann sum of f with respect to the partition P by

$$L(f; P) = \sum_{k=1}^{n} m_k |I_k|$$

= $\sum_{k=1}^{n} m_k (x_k - x_{k-1}).$

Geometrically, U(f; P) is the sum of the areas of rectangles based on the intervals I_k that lie above the graph of f, and L(f; P) is the sum of the areas of rectangles that lie below the graph of f. Note that

$$m(b-a) \le L(f;P) \le U(f;P) \le M(b-a).$$

Let $\Pi(a, b)$, or Π for short, denote the collection of all partitions of [a, b]. The set of all upper Riemann sums $\{U(f; P) : P \in \Pi\}$ is bounded from below by m(b-a), and we can therefore define the upper Riemann integral of f on [a, b] by

$$U(f) = \inf_{P \in \Pi} U(f; P).$$

Similarly, the set of all lower Riemann sums $\{L(f; P) : P \in \Pi\}$ is bounded from above by M(b-a), and we define the lower Riemann integral of f on [a, b] by

$$L(f) = \sup_{P \in \Pi} L(f; P).$$

Note the use of "lower-upper" and "upper-lower" approximations for the integrals: we take the infimum of the upper sums and the supremum of the lower sums. We will show in Corollary 1.13 below that we always have $L(f) \leq U(f)$, but they need not be equal.

Definition 1.3. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable on [a, b] if its upper and lower Riemann integrals are equal, meaning that U(f) = L(f). In that case, the Riemann integral of f on [a, b], denoted by

$$\int_{a}^{b} f(x) \, dx, \qquad \int_{a}^{b} f, \qquad \int_{[a,b]} f$$

or similar notations, is the common value of U(f) and L(f).

Let us illustrate the definition of Riemann integrability with a number of examples.

Example 1.4. Define $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1/x & \text{if } 0 < x \le 1, \\ 0 & \text{if } x = 0. \end{cases}$$

Then

$$\int_0^1 \frac{1}{x} \, dx$$

isn't defined as a Riemann integral becuase f is unbounded. In fact, if

$$0 < x_1 < x_2 < \dots < x_{n-1} < 1$$

is a partition of [0, 1], then

$$\sup_{[0,x_1]} f = \infty$$

so the Riemann sums of f are not well-defined.

An integral with an unbounded interval of integration, such as

$$\int_{1}^{\infty} \frac{1}{x} \, dx,$$

also isn't defined as a Riemann integral. In this case, a Riemann sum associated with a partition of $[1, \infty)$ into intervals of finite length (for example, $I_k = [k, k+1]$ with $k \in \mathbb{N}$) is an infinite series rather than a finite sum, leading to questions of convergence.

One can interpret these integrals as limits of Riemann integrals, or improper Riemann integrals,

$$\int_0^1 \frac{1}{x} \, dx = \lim_{a \to 0^+} \int_a^1 \frac{1}{x} \, dx, \qquad \int_1^\infty \frac{1}{x} \, dx = \lim_{b \to \infty} \frac{1}{x} \, dx.$$

Such improper Riemann integrals involve two limits — a limit of Riemann sums to define the Riemann integrals, followed by a limit of Riemann integrals — and they are not proper Riemann integrals in the sense of Definition 1.3. Both of the improper integrals in this example diverge to infinity.

Next, we consider some examples of bounded functions on compact intervals, where Definition 1.3 does apply.

Example 1.5. The constant function f(x) = 1 on [0, 1] is Riemann integrable, and

$$\int_0^1 1 \, dx = 1.$$

To show this, let $P = \{I_1, \ldots, I_n\}$ be any partition of [0, 1] with endpoints

$$\{x_0, x_1, \ldots, x_n\}$$

Since f is constant,

$$M_k = \sup_{I_k} f = 1, \quad m_k = \inf_{I_k} f = 1 \quad \text{for } k = 1, \dots, n,$$

and therefore

$$U(f;P) = L(f;P) = \sum_{k=1}^{n} (x_k - x_{k-1}) = x_n - x_0 = 1.$$

Geometrically, this equation is the obvious fact that the sum of the areas of the rectangles over (or, equivalently, under) the graph of a constant function is always

exactly equal to the area under the graph. Thus, every upper and lower sum of f on [0, 1] is equal to 1, which implies that the upper and lower integrals

$$U(f) = \inf_{P \in \Pi} U(f; P) = \inf\{1\} = 1, \qquad L(f) = \sup_{P \in \Pi} L(f; P) = \sup\{1\} = 1$$

are equal, and the integral of f is 1.

More generally, the same argument shows that every constant function f(x) = c is integrable and

$$\int_{a}^{b} c \, dx = c(b-a).$$

The following is an example of a discontinuous function that is Riemann integrable.

Example 1.6. The function

$$f(x) = \begin{cases} 0 & \text{if } 0 < x \le 1\\ 1 & \text{if } x = 0 \end{cases}$$

is Riemann integrable, and

$$\int_0^1 f \, dx = 0.$$

To show this, let $P = \{I_1, \ldots, I_n\}$ be a partition of [0, 1]. Then, since f(x) = 0 for x > 0,

$$M_k = \sup_{I_k} f = 0, \quad m_k = \inf_{I_k} f = 0 \quad \text{for } k = 2, \dots, n.$$

The first interval in the partition is $I_1 = [0, x_1]$, where $0 < x_1 \le 1$, and

$$M_1 = 1, \quad m_1 = 0,$$

since f(0) = 1 and f(x) = 0 for $0 < x \le x_1$. It follows that

$$U(f; P) = x_1, \qquad L(f; P) = 0.$$

Thus, L(f) = 0 and

$$U(f) = \inf\{x_1 : 0 < x_1 \le 1\} = 0,$$

so U(f) = L(f) = 0 are equal, and the integral of f is 0. Note that in this example, the infimum of the upper Riemann sums is not attained and U(f; P) > U(f) for every partition P.

A similar argument shows that if a function $f : [a, b] \to \mathbb{R}$ is zero except at finitely many points in [a, b], then it is Riemann integrable and its integral is 0.

The next example shows that the Riemann integral can fail to exist even for bounded functions on a compact interval.

Example 1.7. The Dirichlet function $f: [0,1] \to \mathbb{R}$ defined by

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

is not Riemann integrable. If $P = \{I_k : 1 \le k \le n\}$ is a partition of [0, 1], then

$$M_k = \sup_{I_k} f = 1, \qquad m_k = \inf_{I_k} = 0,$$

since every interval of non-zero length contains both rational and irrational numbers. It follows that

 $U(f; P) = 1, \qquad L(f; P) = 0$

for every partition P of [0, 1], so U(f) = 1 and L(f) = 0 are not equal.

We remark that the Dirichlet function is Lebesgue integrable. Its Lebesgue integral is given by

$$\int_{0}^{1} f = 1 \cdot |A| + 0 \cdot |B|$$

where $A = \mathbb{Q} \cap [0, 1]$ is the set of rational numbers in [0, 1] and $B = [0, 1] \setminus \mathbb{Q}$ is the set of irrational numbers. Here, |E| denotes the Lebesgue measure of the set E, which is a countably additive extension of the length of an interval to more general sets. It turns out that |A| = 0 (as is true for any countable set of real numbers) and |B| = 1, so the Lebesgue integral of the Dirichlet function is 0.

The moral of the previous example is that the Riemann integral of a highly discontinuous function need not exist.

A precise statement about Riemann integrability can be given in terms of Lebesgue measure. One can show that a set $E \subset \mathbb{R}$ has Lebesgue measure zero if and only if for every $\epsilon > 0$ there is a countable collection of open intervals $\{(a_k, b_k) : k \in \mathbb{N}\}$ such that

$$E \subset \bigcup_{k=1}^{\infty} (a_k, b_k), \qquad \sum_{k=1}^{\infty} (b_k - a_k) < \epsilon.$$

A bounded function on a compact interval is Riemann integrable if and only if the set of points at which it is discontinuous has Lebesgue measure zero.

For example, the set of discontinuities of the function in Example 1.6 consists of a single point $\{0\}$, which has Lebesgue measure zero. On the other hand, the Dirichlet function in Example 1.7 is discontinuous at every point of [0, 1], and its set of discontinuities has Lebesgue measure one.

From now on, we'll only consider the Riemann integral; "integrable" will mean "Riemann integrable, and "integral" will mean "Riemann integral."

1.2. Refinements of partitions

As the previous examples illustrate, a direct verification of integrability from Definition 1.3 is unwieldy even for the simplest functions because we have to consider all possible partitions of the interval of integration. We can obtain simpler and more easily verified conditions for Riemann integrability by considering refinements of partitions.

Definition 1.8. A partition $Q = \{J_{\ell} : \ell = 1, ..., m\}$ is a refinement of a partition $P = \{I_k : k = 1, ..., n\}$ if every interval I_k in P is an almost disjoint union of one or more intervals J_{ℓ} in Q

Equivalently, if we represent partitions by their endpoints, then Q is a refinement of P if $Q \supset P$, meaning that every endpoint of P is an endpoint of Q.

Example 1.9. Consider the partitions of [0, 1] with endpoints

 $P = \{0, 1/2, 1\}, \qquad Q = \{0, 1/3, 2/3, 1\}, \qquad R = \{0, 1/4, 1/2, 3/4, 1\}.$

Then Q is not a refinement of P but R is a refinement of P.

Given two partitions, neither need be a refinement of the the other. However, two partitions P, Q always have a common refinement; the smallest one is $R = P \cup Q$, meaning that the endpoints of R are exactly the endpoints of P or Q (or both).

Example 1.10. Let $P = \{0, 1/2, 1\}$ and $Q = \{0, 1/3, 2/3, 1\}$, as in Example 1.9. Then Q isn't a refinement of P and P isn't a refinement of Q. The partition $S = P \cup Q$, or

$$S = \{0, 1/3, 1/2, 2/3, 1\},\$$

is a refinement of both P and Q. The partition S is not a refinement of R, but $T = R \cup S$, or

$$T = \{0, 1/4, 1/3, 1/2, 2/3, 3/4, 1\},\$$

is a common refinement of all of the partitions $\{P, Q, R, S\}$.

As we show next, refining partitions decreases upper sums and increases lower sums. (The proof is easier to understand than it is to write out — draw a picture!)

Proposition 1.11. If $f : [a,b] \to \mathbb{R}$ is bounded and P, Q are partitions of [a,b] such that Q is refinement of P, then

$$U(f;Q) \le U(f;P), \qquad L(f;P) \le L(f;Q).$$

Proof. Let

$$P = \{I_k : k = 1, \dots, n\}, \qquad Q = \{J_\ell : \ell = 1, \dots, m\}$$

be partitions of [a, b], where the intervals are listed in increasing order of their endpoints. Define

$$M_k = \sup_{I_k} f, \quad m_k = \inf_{I_k} f, \qquad M'_\ell = \sup_{J_\ell} f, \quad m'_\ell = \inf_{J_\ell} f.$$

Since Q is a refinement of P, each interval I_k in P is an almost disjoint union of intervals $\{J_{\ell} : p_k \leq \ell \leq q_k\}$ in Q for some p_k, q_k , where

$$1 = p_1 \le q_1 < p_2 \le q_2 < p_3 \le q_3 < \dots < p_n \le q_n = n.$$

In particular, $J_{\ell} \subset I_k$ if $p_k \leq \ell \leq q_k$, so

$$M'_{\ell} \le M_k, \qquad m_k \le m'_{\ell} \qquad \text{for } p_k \le \ell \le q_k.$$

Using the fact that the sum of the lengths of the J-intervals is the length of the I-interval, we get that

$$\sum_{\ell=p_k}^{q_k} M'_{\ell} |J_{\ell}| \le \sum_{\ell=p_k}^{q_k} M_k |J_{\ell}| = M_k \sum_{\ell=p_k}^{q_k} |J_{\ell}| = M_k |I_k|.$$

It follows that

$$U(f;Q) = \sum_{\ell=1}^{m} M_{\ell}' |J_{\ell}| = \sum_{k=1}^{n} \sum_{\ell=p_k}^{q_k} M_{\ell}' |J_{\ell}| \le \sum_{k=1}^{n} M_k |I_k| \le U(f;P)$$

Similarly,

$$\sum_{\ell=p_k}^{q_k} m'_{\ell} |J_{\ell}| \ge \sum_{\ell=p_k}^{q_k} m_k |J_{\ell}| = m_k |I_k|,$$

and

$$L(f;Q) == \sum_{k=1}^{n} \sum_{\ell=p_k}^{q_k} m'_{\ell} |J_{\ell}| \ge \sum_{k=1}^{n} m_k |I_k| \ge L(f;P).$$

Using this result, we can prove that all lower sums are less than or equal to all upper sums, not just the lower and upper sums associated with the same partition.

Proposition 1.12. If $f : [a,b] \to \mathbb{R}$ is bounded and P, Q are any partitions of [a,b], then

$$L(f;P) \le U(f;Q).$$

Proof. Let R be a refinement of both P and Q. Then, by Proposition 1.11,

$$L(f; P) \le L(f; R), \qquad U(f; R) \le U(f; Q).$$

It follows that

$$L(f;P) \le L(f;R) \le U(f;R) \le U(f;Q).$$

An immediate consequence of this result that the lower integral is always less than or equal to the upper integral.

Corollary 1.13. If $f : [a, b] \to \mathbb{R}$ is bounded, then

$$L(f) \le U(f).$$

Here we use the following lemma, whose conclusion is obvious but requires proof.

Lemma 1.14. Suppose that A, B are nonempty sets of real numbers such that $a \leq b$ for all $a \in A$ and $b \in B$. Then $\sup A \leq \inf B$.

Proof. The condition implies that every $b \in B$ is an upper bound of A, so $\sup A \leq b$. Hence, $\sup A$ is a lower bound of B, so $\sup A \leq \inf B$.

1.3. Existence of the Riemann integral

The following theorem gives a simple criterion for integrability.

Theorem 1.15. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if for every $\epsilon > 0$ there exists a partition P of [a, b] (depending on ϵ) such that

$$U(f;P) - L(f;P) < \epsilon.$$

Proof. First, suppose that the condition holds. Let $\epsilon > 0$ and choose a partition P that satisfies the condition. Then, since $U(f) \leq U(f; P)$ and $L(f; P) \leq L(f)$, we have

$$0 \le U(f) - L(f) \le U(f; P) - L(f; P) < \epsilon.$$

Since this inequality holds for all $\epsilon > 0$, we must have U(f) = L(f), and f is integrable.

Conversely, suppose that f is integrable, meaning that U(f) = L(f). Given any $\epsilon > 0$, there are partitions Q, R such that

$$U(f;Q) < U(f) + \frac{\epsilon}{2}, \qquad L(f;R) > L(f) - \frac{\epsilon}{2}.$$

Let P be a common refinement of Q and R. Then, by Proposition 1.11,

$$U(f;P) - L(f;P) \le U(f;Q) - L(f;R) < U(f) - L(f) + \epsilon.$$

Since U(f) = L(f), the condition follows.

Furthermore, it is sufficient to find a sequence of partitions whose upper and lower sums approach each other.

Theorem 1.16. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if there is a sequence (P_n) of partitions such that

$$\lim_{n \to \infty} \left[U(f; P_n) - L(f; P_n) \right] = 0.$$

In that case,

$$\int_{a}^{b} f = \lim_{n \to \infty} U(f; P_n) = \lim_{n \to \infty} L(f; P_n).$$

Proof. First, suppose that the condition holds. Then, given $\epsilon > 0$, there is an $n \in \mathbb{N}$ such that $U(f; P_n) - L(f; P_n) < \epsilon$, so Theorem 1.15 implies that f is integrable and U(f) = L(f). Moreover, since $U(f) \leq U(f; P_n)$ and $L(f; P_n) \leq L(f)$,

$$0 \le U(f; P_n) - U(f) = U(f; P_n) - L(f) \le U(f; P_n) - L(f; P_n).$$

Since the limit of the right-hand side is zero, the 'squeeze' theorem implies that

$$\lim_{n \to \infty} U(f; P_n) = U(f).$$

It then also follows that

$$\lim_{n \to \infty} L(f; P_n) = \lim_{n \to \infty} U(f; P_n) - \lim_{n \to \infty} \left[U(f; P_n) - L(f; P_n) \right] = U(f).$$

Conversely, if f is integrable then, by Theorem 1.15, for every $n \in \mathbb{N}$ there exists a partition P_n such that

$$0 \le U(f; P_n) - L(f; P_n) < \frac{1}{n}$$

and then $[U(f; P_n) - L(f; P_n)] \to 0$ as $n \to \infty$.

Note that, since

$$\lim_{n \to \infty} \left[U(f; P_n) - L(f; P_n) \right] = \lim_{n \to \infty} U(f; P_n) - \lim_{n \to \infty} L(f; P_n),$$

the conditions of the theorem are satisfied if the limits $U(f; P_n)$ and $L(f; P_n)$ exist and are equal. Conversely, the proof of the theorem shows that if the limit of

 $U(f; P_n) - L(f; P_n)$ is zero, then the limits of $U(f; P_n)$ and $L(f; P_n)$ exist and are equal. This isn't true for general sequences, where one may have $\lim (a_n - b_n) = 0$ even though $\lim a_n$ and $\lim b_n$ do not exist.

Example 1.17. Consider $f(x) = x^2$ on [0,1]. Let P_n be a partial of [0,1] into *n*-intervals of equal length 1/n, with $x_k = k/n$ for $0 \le k \le n$. Then, using the formula for the sum of squares

$$\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1),$$

we get

$$U(f; P_n) = \sum_{k=1}^n x_k^2 \cdot \frac{1}{n} = \frac{1}{n^3} \sum_{k=1}^n k^2 = \frac{1}{6} \left(1 + \frac{1}{n} \right) \left(2 + \frac{1}{n} \right)$$

and

$$L(f; P_n) = \sum_{k=1}^n x_{k-1}^2 \cdot \frac{1}{n} = \frac{1}{n^3} \sum_{k=1}^{n-1} k^2 = \frac{1}{6} \left(1 - \frac{1}{n} \right) \left(2 - \frac{1}{n} \right)$$

It follows that

$$\lim_{n \to \infty} U(f; P_n) = \lim_{n \to \infty} L(f; P_n) = \frac{1}{3},$$

and Theorem 1.16 implies that so x^2 is integrable on [0, 1] with

$$\int_{0}^{1} x^{2} \, dx = \frac{1}{3}$$

The main application of Theorems 1.15 is the following fundamental result that every continuous function is Riemann integrable.

Theorem 1.18. A continuous function $f : [a, b] \to \mathbb{R}$ on a compact interval is Riemann integrable.

Proof. A continuous function on a compact set is bounded, so its upper and lower sums are well-defined, and we just need to verify the condition in Theorems 1.15.

Let $\epsilon>0.$ A continuous function on a compact set is uniformly continuous, so there exists $\delta>0$ such that

$$|f(x) - f(y)| < \frac{\epsilon}{b-a}$$
 for all $x, y \in [a, b]$ such that $|x - y| < \delta$.

Choose a partition $P = \{I_k : k = 1, ..., n\}$ of [a, b] such that $|I_k| < \delta$ for every k; for example, we can take n intervals of equal length (b-a)/n with $n > (b-a)/\delta$.

Since f is continuous, it attain its maximum and minimum values M_k and m_k on the compact interval I_k at points x_k and y_k in I_k , which satisfy $|x_k - y_k| < \delta$. It follows that

$$M_k - m_k = f(x_k) - f(y_k) < \frac{\epsilon}{b-a}.$$

The upper and lower sums of f therefore satisfy

$$U(f;P) - L(f;P) = \sum_{k=1}^{n} M_k |I_k| - \sum_{k=1}^{n} m_k |I_k|$$
$$= \sum_{k=1}^{n} (M_k - m_k) |I_k|$$
$$< \frac{\epsilon}{b-a} \sum_{k=1}^{n} |I_k|$$
$$< \epsilon,$$

and Theorems 1.15 implies that f is integrable.