Integration

Robert J. Carroll

Last revised: 5 May 2026

---

1 Motivation

The probability handouts to this point have been quietly leaning on integration. The expectation of a continuous random variable was defined as $\mathbb{E}[X] = \int_{\mathbb{R}} x\, f_X(x)\, dx$; the expectation of a general random variable was the Lebesgue integral against the underlying probability measure; the law of large numbers and the central limit theorem both rest on convergence theorems — monotone convergence, dominated convergence — that are integration theorems in disguise. The project assumed calculus from the preface, which covers Riemann integration informally, but the apparatus required to make probability rigorous is Lebesgue integration, and that is genuinely beyond what calculus delivers.

This handout fills the gap. The first piece is to make Riemann integration formally precise, since calculus as taught typically defines $\int_a^b f(x)\, dx$ as "area under the curve" and shifts to the fundamental theorem of calculus before pinning the definition down. The second piece, which is the bulk of the handout, is to develop *Lebesgue integration*: a generalization of the Riemann integral that is defined on a much broader class of functions, behaves well under pointwise limits (in a way Riemann integration does not), and is the right notion for measure-theoretic probability. The third piece is the comparison: when both notions apply, they give the same value, and the integrals one actually computes by hand (polynomials, rationals, exponentials, trig functions) look the same in either framework. The Lebesgue framework matters not for everyday calculation but for what it lets us prove.

The order in the project: this handout sits between the probability-spaces-and-measures handout (which introduces Lebesgue measure on $\mathbb{R}$) and the random-variables-and-expectations handout (which defines expectation as the Lebesgue integral against the probability measure). With these three handouts in sequence, the probability cluster has the foundation it needs without hand-waving.

2 Riemann integration

Almost every applied integral that shows up in undergraduate political-science methods — expected utility under a continuous distribution, probability calculations using normal CDFs, present-value computations $\int_0^T u(c)e^{-\rho t}dt$ in dynamic models, areas under polling-error distributions — is a Riemann integral being computed via the fundamental theorem of calculus. Calculus gave us the Riemann integral informally as "area under the curve," approximated by thinner and thinner rectangles. The rigorous definition makes the approximation precise: an integral exists when fine enough partitions of the interval produce arbitrarily good upper and lower bounds on the area, and the two bounds converge to a common value.

Definition 1. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. A *partition* of $[a, b]$ is a finite sequence $P = (x_0 = a < x_1 < x_2 < \cdots < x_n = b)$. For each subinterval $[x_{i-1}, x_i]$, let $M_i = \sup_{x\in[x_{i-1},x_i]} f(x)$

and $m_i = \inf_{x \in [x_{i-1}, x_i]} f(x)$. The *upper Riemann sum* and *lower Riemann sum* of f over P are

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

The function f is *Riemann integrable* on $[a,b]$ if

$$\inf_P U(f,P) = \sup_P L(f,P),$$

in which case the common value is the *Riemann integral* $\int_a^b f(x)\,dx$.

The standard sufficient conditions for Riemann integrability are easy:

Theorem 2. *Every continuous function on $[a,b]$ is Riemann integrable. Every monotone function on $[a,b]$ is Riemann integrable.*

The continuous case follows from uniform continuity (Heine–Cantor, from the continuity handout): on a compact interval, f is uniformly continuous, so its variation on any sufficiently fine partition is uniformly small, forcing $U(f,P) - L(f,P) \to 0$ as the partition is refined. The monotone case is even simpler: the variation of a monotone function on $[x_{i-1}, x_i]$ is exactly $|f(x_i) - f(x_{i-1})|$, and these telescope. The fundamental theorem of calculus, which we state without proof, ties the integral to differentiation:

Theorem 3 (Fundamental theorem of calculus)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous. Then $F(x) := \int_a^x f(t)\,dt$ is differentiable on (a,b) with $F'(x) = f(x)$. If G is any antiderivative of f on $[a,b]$, then $\int_a^b f(x)\,dx = G(b) - G(a)$.*

This is what makes Riemann integration practically computable: integration becomes "find an antiderivative." For applied work the FTC is the workhorse, and almost every integral one computes in calculus or in undergraduate political-science methods courses is being computed via the FTC.

The Riemann integral has two limitations that motivate the upgrade to Lebesgue integration.

Some bounded functions are not Riemann integrable. The canonical example is the Dirichlet function $\mathbf{1}_{\mathbb{Q}} : [0,1] \to \{0,1\}$, which is 1 at every rational and 0 at every irrational. On any partition of $[0,1]$, every subinterval contains both rationals and irrationals (since both are dense), so $M_i = 1$ and $m_i = 0$ for every i. Thus $U(f,P) = 1$ and $L(f,P) = 0$ for every partition P, and the inf and sup do not match: $\mathbf{1}_{\mathbb{Q}}$ is not Riemann integrable. Yet on cardinality grounds (the rationals are countable, the irrationals are uncountable), one would intuitively want to assign this function the integral 0 — the rationals are a "small" subset of $[0,1]$, and they should not contribute area. Lebesgue integration delivers this.

The pointwise limit of Riemann-integrable functions need not be Riemann-integrable. Enumerate $\mathbb{Q} \cap [0,1]$ as $q_1, q_2, q_3, \ldots$, and define $f_n(x) = 1$ if $x \in \{q_1, \ldots, q_n\}$ and $f_n(x) = 0$ otherwise. Each f_n is Riemann integrable with $\int_0^1 f_n = 0$ (the function is zero except at finitely many points). Pointwise, $f_n \to \mathbf{1}_{\mathbb{Q}}$, which we just saw is not Riemann integrable. So the class of Riemann-integrable functions is not closed under pointwise limits, and the natural identity $\int \lim f_n = \lim \int f_n$ has no straightforward analog. This is a substantial obstacle for proving theorems about expectations, where one constantly takes limits of random variables.

3 The Lebesgue integral

Why bother with a second notion of integration when Riemann is already in hand? Two reasons, both of which directly affect probabilistic modeling in political science. First, Riemann integration runs out of room: the Dirichlet function we just saw is not Riemann integrable, but on cardinality grounds (the rationals are a measure-zero subset of $[0,1]$) it ought to integrate to zero, and Riemann does not deliver. Second, Riemann integration handles limits poorly — the pointwise limit of Riemann-integrable functions need not itself be Riemann-integrable. That is fatal for proving theorems about expectations of limits of random variables, which is much of what probability theory does. The Lebesgue integral, built directly on a measure space rather than on intervals of $\mathbb{R}$, addresses both gaps. The construction is exactly the three-stage one sketched in the random-variables-and-expectations handout, now made rigorous.

Definition 4 (Lebesgue integral, simple functions)**.** A *simple function* on $(\Omega, \mathcal{F})$ is a measurable function taking only finitely many values: $s = \sum_{i=1}^{k} c_i \mathbf{1}_{A_i}$ with $c_i \in \mathbb{R}$ and $A_i \in \mathcal{F}$ pairwise disjoint. For a non-negative simple function ($c_i \geq 0$), the *integral with respect to* μ is

$$\int_\Omega s\, d\mu := \sum_{i=1}^{k} c_i \mu(A_i),$$

with the convention $0 \cdot \infty = 0$ (so a function that is zero on a set of infinite measure contributes nothing).

The well-definedness check (that the integral is independent of the representation as $\sum c_i \mathbf{1}_{A_i}$) is a small but standard exercise. The integral on simple functions is linear and monotone: $\int (as + bt) d\mu = a \int s\, d\mu + b \int t\, d\mu$, and $s \leq t$ implies $\int s\, d\mu \leq \int t\, d\mu$.

Definition 5 (Lebesgue integral, non-negative functions)**.** For a non-negative measurable function $f : \Omega \to [0, \infty]$, the *integral with respect to* μ is

$$\int_\Omega f\, d\mu := \sup \left\{ \int_\Omega s\, d\mu : 0 \leq s \leq f,\ s \text{ simple} \right\} \in [0, \infty].$$

The supremum exists in $[0, \infty]$, but may be infinite: not every non-negative measurable function has a finite integral. The integral is monotone (larger functions have larger integrals) and additive ($\int (f+g) d\mu = \int f\, d\mu + \int g\, d\mu$), though linearity in scalar multiples requires a small argument.

Definition 6 (Lebesgue integral, general functions)**.** For a measurable function $f : \Omega \to \mathbb{R}$, write $f = f^+ - f^-$ where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$ (both non-negative measurable). The *Lebesgue integral* is

$$\int_\Omega f\, d\mu := \int_\Omega f^+\, d\mu - \int_\Omega f^-\, d\mu,$$

provided at least one of $\int f^+\, d\mu$ and $\int f^-\, d\mu$ is finite. The function f is *Lebesgue integrable* (or just *integrable*) when $\int |f|\, d\mu < \infty$, which is equivalent to both $\int f^+, \int f^- < \infty$.

For a measurable subset $A \in \mathcal{F}$, we write $\int_A f\, d\mu := \int_\Omega f \mathbf{1}_A\, d\mu$. The standard properties — linearity, monotonicity, the triangle inequality $|\int f\, d\mu| \leq \int |f| d\mu$ — all extend to the general case from the non-negative case.

Example 7 (The Dirichlet function, redeemed)**.** The function $\mathbf{1}_{\mathbb{Q}}$ on $[0,1]$ with Lebesgue measure λ is a simple function: $\mathbf{1}_{\mathbb{Q}} = 1 \cdot \mathbf{1}_{\mathbb{Q}\cap[0,1]} + 0 \cdot \mathbf{1}_{[0,1]\setminus\mathbb{Q}}$. By Definition 4, $\int_{[0,1]} \mathbf{1}_{\mathbb{Q}}\, d\lambda = 1 \cdot \lambda(\mathbb{Q} \cap [0,1]) + 0 \cdot \lambda([0,1] \setminus \mathbb{Q}) = 1 \cdot 0 + 0 \cdot 1 = 0$. So the Dirichlet function is Lebesgue integrable with integral 0, just as the cardinality intuition suggests — the rationals are a measure-zero subset of $[0,1]$, and integration weights them as such.

Example 8 (The Lebesgue integral on a probability space)**.** When $(\Omega, \mathcal{F}, \mu)$ is a probability space (so $\mu = \mathbb{P}$), the Lebesgue integral $\int_\Omega f\, d\mathbb{P}$ is the *expectation* $\mathbb{E}[f]$ from the random-variables-and-expectations handout. The three-stage construction here is the rigorous version of the construction sketched there. With this in hand, every claim about expectations made in the probability handouts has a precise integration-theoretic backing.

4 Convergence theorems

When can we swap a limit and an integral? In probability terms, when does $\mathbb{E}[\lim X_n] = \lim \mathbb{E}[X_n]$? The question comes up everywhere in applied work: posterior expectations as fresh data arrive, expected utility as the time horizon grows, sample averages as the sample size grows, expected-value computations under sequences of approximating distributions. The Riemann theory handles this poorly; pointwise limits of Riemann-integrable functions need not even be Riemann-integrable. The Lebesgue theory delivers two clean theorems with mild hypotheses — monotone convergence and dominated convergence — and they are the workhorse limit-versus-integral tools in measure-theoretic probability.

Theorem 9 (Monotone convergence, Beppo Levi)**.** *Let (f_n) be a sequence of non-negative measurable functions on $(\Omega, \mathcal{F}, \mu)$ with $f_n \uparrow f$ pointwise (i.e., $f_n(\omega) \le f_{n+1}(\omega)$ for every n and every ω, and $f_n(\omega) \to f(\omega)$). Then*

$$\int_\Omega f_n\, d\mu \to \int_\Omega f\, d\mu \in [0, \infty].$$

The non-negativity and monotonicity hypotheses are essential: without monotonicity, the sequence $f_n = n\mathbf{1}_{[0,1/n]}$ on $[0,1]$ has $\int f_n\, d\lambda = 1$ for every n but $f_n \to 0$ pointwise, with $\int 0\, d\lambda = 0$. The integrals do not converge to the integral of the limit. Monotone convergence rules out this kind of escape-to-infinity by demanding that the sequence is increasing pointwise.

Lemma 10 (Fatou's lemma)**.** *For any sequence (f_n) of non-negative measurable functions on $(\Omega, \mathcal{F}, \mu)$,*

$$\int_\Omega \liminf_n f_n\, d\mu \le \liminf_n \int_\Omega f_n\, d\mu.$$

Fatou's lemma is the workhorse one-sided inequality of integration theory and is what underwrites several almost-sure convergence arguments in probability. The other direction need not hold without further hypotheses, as the example above shows.

Theorem 11 (Dominated convergence, Lebesgue)**.** *Let (f_n) be a sequence of measurable functions on $(\Omega, \mathcal{F}, \mu)$ with $f_n \to f$ pointwise (or μ-almost everywhere). Suppose there is a non-negative integrable function g such that $|f_n| \le g$ pointwise (or a.e.) for every n. Then f is integrable and*

$$\int_\Omega f_n\, d\mu \to \int_\Omega f\, d\mu, \qquad \int_\Omega |f_n - f|\, d\mu \to 0.$$

Dominated convergence is the most-applied of the three theorems and is the one most likely to come up in everyday probabilistic work. It is what justifies pulling expectations through limits when you have a uniform integrable bound: $\mathbb{E}[\lim X_n] = \lim \mathbb{E}[X_n]$ when $|X_n| \leq Y$ a.s. for some integrable Y. The bounded-convergence theorem (special case where g is a constant) is a frequently-cited corollary; the version most useful in probability is the dominated form.[1]

5 Riemann versus Lebesgue

Should a political-economy modeler think of integrals as Riemann or Lebesgue? The practical answer is reassuring: it almost never matters for the integrals one actually computes. Continuous functions on compact intervals are integrable in both senses with the same value, so the calculus-style computations of expected utility, present values, and confidence-interval probabilities are unchanged whichever framework one keeps in mind. The Lebesgue framework matters when the modeler wants to *prove* something — a swap of limit and integral, an existence result for an estimator, an asymptotic theorem — where the convergence theorems above do work that the Riemann theory cannot. This section makes the relationship between the two notions precise.

Theorem 12. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. Then f is Lebesgue integrable on $[a, b]$ (with respect to Lebesgue measure λ), and the two integrals agree:*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f\,d\lambda.$$

The proof is direct: a Riemann-integrable function is approximated above and below by step functions (simple functions whose level sets are finite unions of intervals), the inf and sup of step-function integrals match the Riemann integral by definition, and these are also bounds on the Lebesgue integral, which is squeezed between them.

A sharper characterization. The set of Riemann-integrable functions on $[a, b]$ is not just smaller than the set of Lebesgue-integrable functions; it is exactly characterized by where the function is continuous.

Theorem 13 (Lebesgue's characterization of Riemann integrability)**.** *A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if the set of points at which f is discontinuous has Lebesgue measure zero.*

The Dirichlet function fails on this criterion: it is discontinuous everywhere on $[0, 1]$, and the discontinuity set $[0, 1]$ has Lebesgue measure $1 \neq 0$, so the function is not Riemann integrable. A function with at most countably many discontinuities, by contrast, is Riemann integrable (countable sets have Lebesgue measure zero), recovering the standard sufficient conditions including all continuous functions and all monotone functions on a compact interval.

[1] The historical motivation for Lebesgue integration was precisely the limit-of-integrals problem. Riemann's work on Fourier series in the 1850s pushed the question of when $\int \lim f_n = \lim \int f_n$ holds, and the answer for Riemann integration was unsatisfying — the conditions are restrictive and proofs are clumsy. Lebesgue's 1902 thesis introduced the integral that now bears his name, with the convergence theorems as the headline payoff. The shift from Riemann to Lebesgue integration was one of the central technical developments of early-twentieth-century analysis, and it is what made measure-theoretic probability (Kolmogorov 1933) possible. Folland (1999) chapter 2 gives the standard development of the integral and the convergence theorems; Billingsley (1995) chapter 3 gives the probability-flavored version.

For applied work, the practical takeaway is small. The integrals one writes down in calculus, in econometrics, and in everyday political-economy modeling — $\int x f_X(x)dx$, $\int e^{-x^2/2}dx$, $\int_0^\infty u(c)e^{-\rho t}dt$ — are integrals of continuous functions, where Riemann and Lebesgue agree. The Lebesgue formulation is what one needs to prove things rigorously, especially in the convergence-theorem step where one wants to swap a limit and an integral.

6 What's next

This handout closes the immediate gap in the project's foundations. With Riemann and Lebesgue integration in place, the random-variables-and-expectations handout's definition of expectation is no longer hand-waving, and the convergence-and-limit-theorems handout's invocations of monotone and dominated convergence are now backed by the corresponding theorems above.

A few directions one could extend, none developed here:

- *Product measures and Fubini's theorem.* For a product measure space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$, Fubini's theorem says $\int f\, d(\mu_1 \otimes \mu_2) = \int (\int f\, d\mu_1)\, d\mu_2$ when f is integrable on the product. This is the formal version of "compute a multiple integral by iterating one-dimensional integrals," and it is the foundation of joint distributions in probability.
- *L^p spaces.* The set of μ-integrable functions with $\int |f|^p\, d\mu < \infty$ forms a Banach space $L^p(\Omega, \mathcal{F}, \mu)$ under the norm $\|f\|_p := (\int |f|^p\, d\mu)^{1/p}$, modulo functions equal μ-a.e. The cases $p = 1, 2, \infty$ are the standard ones. Convergence in L^p from the convergence-and-limit-theorems handout is convergence in this norm.
- *The Radon–Nikodym theorem*, which characterizes when a measure has a density with respect to another, was invoked in the random-variables handout to define the PDF and conditional expectation. The proof is a substantial application of the integration theory developed here.

For broader treatments at this level, see Folland (1999) chapter 2 (the canonical careful presentation), Billingsley (1995) chapters 1–3, or Rudin (1976) chapter 11 (briefer, less probability-flavored).

7 Exercises

Exercise 14. Compute $\int_0^1 x^2\, dx$ directly from Definition 1, using uniform partitions P_n with $x_i = i/n$ for $i = 0, 1, \ldots, n$. Identify M_i and m_i on each subinterval, write out $U(f, P_n)$ and $L(f, P_n)$, and let $n \to \infty$ to get the value $1/3$.

Exercise 15. Show that the Dirichlet function $\mathbf{1}_{\mathbb{Q}} : [0,1] \to \{0,1\}$ is not Riemann integrable, by explicitly computing $U(f, P) = 1$ and $L(f, P) = 0$ for any partition P of $[0, 1]$.

Exercise 16. Verify Example 7: $\int_{[0,1]} \mathbf{1}_{\mathbb{Q}}\, d\lambda = 0$. (*Hint*: this uses only that $\mathbb{Q} \cap [0, 1]$ has Lebesgue measure zero, which we showed in the probability-spaces-and-measures handout.)

Exercise 17 (Expected loss for a candidate)**.** Voter ideal points are uniformly distributed on $[0, 1]$ (Lebesgue measure). A candidate at position $c \in [0, 1]$ generates per-voter loss $u(x) = -(x - c)^2$ when the voter has ideal point x. Compute the expected loss $\mathbb{E}[u] = \int_{[0,1]} -(x - c)^2\, d\lambda(x)$ as a Riemann integral. (Recall that since the integrand is continuous, Riemann and Lebesgue agree.)

Exercise 18 (Median voter, optimal candidate)**.** Continuing Exercise 17: minimize the expected loss over c. (Take the derivative with respect to c inside the integral — a step justified by the dominated convergence theorem, but here easy to do directly — and set to zero.) The minimizer is the *mean* of the voter ideal-point distribution, which for uniform-on-$[0,1]$ is $c^* = 1/2$. Note: under quadratic loss, the optimal candidate position is the mean of the distribution; under absolute-value loss $|x - c|$, the optimum is the median. The two coincide for any symmetric distribution, including uniform.

Exercise 19. Apply the monotone convergence theorem to compute

$$\lim_{n\to\infty} \int_{[0,1]} (1 - (1-x)^n)\, d\lambda(x).$$

(*Hint*: the integrands form an increasing sequence with pointwise limit 1 on $(0, 1]$.)

Exercise 20 (Polling estimator expectation)**.** The polling estimator $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$ has $\mathbb{E}[\bar{X}_n] = (1/n)\sum_{i=1}^n \mathbb{E}[X_i] = p$ when $\mathbb{E}[X_i] = p$ for each i. Spell out the steps using the linearity of the Lebesgue integral. Argue that this argument applies in identical form to (i) discrete X_i (Bernoulli, where the integral reduces to a sum), and (ii) absolutely continuous X_i (where the integral is against a PDF). The unified Lebesgue framework is what lets the same line of reasoning handle both.

Exercise 21. Apply the dominated convergence theorem to evaluate

$$\lim_{n\to\infty} \int_{[0,1]} \frac{nx}{1 + n^2x^2}\, d\lambda(x).$$

The pointwise limit of the integrands is 0 except at $x = 0$. Find an integrable function dominating $|nx/(1 + n^2x^2)|$ uniformly in n. (*Hint*: AM-GM gives $n^2x^2 + 1 \geq 2nx$, so $nx/(1 + n^2x^2) \leq 1/2$.) Conclude the limit equals 0.

Exercise 22 (Conditional expectation as a restricted integral)**.** Continuing Exercise 17 with $c = 1/2$: compute $\mathbb{E}[u \mid x \in [0.4, 0.6]] = \int_{[0.4,0.6]} -(x - 0.5)^2\, d\lambda(x)/\lambda([0.4, 0.6])$. Compare with $\mathbb{E}[u]$ under the unconditional uniform distribution. Interpret: conditioning on a narrower voter type-range produces a smaller expected loss for the optimal candidate; this is what makes targeted-segment polling more informative than aggregate polling for a candidate's positioning decisions, in the simplest version of the calculation.

Exercise 23. Show every continuous function $f : [a, b] \to \mathbb{R}$ is both Riemann and Lebesgue integrable, with the same value. Then exhibit a function on $[0, 1]$ that is Lebesgue integrable but not Riemann integrable, distinct from the Dirichlet function. (*Hint*: Lebesgue's characterization (Theorem 13) says you need a function whose discontinuity set has positive Lebesgue measure but is still bounded, so the Lebesgue integral exists.)

References

Billingsley, Patrick (1995). *Probability and Measure*. 3rd ed. New York: Wiley.

Folland, Gerald B. (1999). *Real Analysis: Modern Techniques and Their Applications*. 2nd ed. New York: Wiley.

Rudin, Walter (1976). *Principles of Mathematical Analysis*. 3rd ed. New York: McGraw-Hill.