

Lecture 3: Basics of Measure Theory 2

Lecturer: Yiqiao Zhong

Last update: 16:27:38, 2024-09-11

Disclaimer: materials in this pdf are used for STAT 709 “Mathematical Statistics” at the University of Wisconsin–Madison, but they are not guaranteed to be accurate. Consult textbooks whenever possible.

We continue introducing basic measure theory in this lecture, and we will see important theorems that truly make measure theory useful. Here are some useful abbreviations: w.r.t. means with respect to, a.e. means almost everywhere, DCT means Dominated Convergence Theorem, MCT means Monotone Convergence Theorem, RN means Radon-Nikodym.

Notations. We denote indicator functions using both $\mathbf{1}_A(x)$ and $\mathbf{1}\{x \in A\}$:

$$\mathbf{1}_A(x) = \mathbf{1}\{x \in A\} = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

1 Abstract integration

We generalize Riemann integrals in the following three steps. First, we define the integral of a nonnegative simple function. As usual, let $(\Omega, \mathcal{F}, \nu)$ be a measure space.

Definition 1. Let $k \geq 1$ be an integer, $(a_i)_{i \leq k}$ be any nonnegative scalars, and $A_1, \dots, A_k$ be a partition of Ω , meaning that these sets are disjoint and $\Omega = \cup_{i \leq k} A_i$. The integral of a nonnegative simple function

$\phi(\omega) = \sum_{i=1}^k a_i \mathbf{1}_{A_i}(\omega)$ w.r.t. ν is defined as

$$\int \phi \, d\nu = \sum_{i=1}^k a_i \nu(A_i).$$

Next, we consider a nonnegative measurable function f .

Definition 2. Let f be a nonnegative measurable function and let S_f be the collection of all nonnegative simple functions satisfying $\phi(\omega) \leq f(\omega)$ for any $\omega \in \Omega$. The integral of f w.r.t. ν is defined as

$$\int f \, d\nu = \sup \left\{ \int \phi \, d\nu : \phi \in S_f \right\}.$$

This definition using supreme looks abstract but provides convenience for proving theorems. More intuitively, we can actually choose a sequence of simple functions $(s_n)_{n \geq 1}$ nondecreasing in n such that $f(x) = \lim_{n \rightarrow \infty} s_n(x)$ for every x (exercise).

Finally, for a measurable function f , we first define the positive part of f by

$$f_+(\omega) = \max\{f(\omega), 0\}$$

and the negative part of f by

$$f_-(\omega) = \max\{-f(\omega), 0\}.$$

From the properties of measurable functions in Lecture 2, we see that f_+ and f_- are nonnegative measurable functions. Moreover, $f(\omega) = f_+(\omega) - f_-(\omega)$, and $|f(\omega)| = f_+(\omega) + f_-(\omega)$.

Definition 3. Let f be a measurable function. We say that $\int f d\nu$ exists if and only if at least one of $\int f_+ d\nu$ and $\int f_- d\nu$ is finite, in which case

$$\int f d\nu = \int f_+ d\nu - \int f_- d\nu.$$

When both $\int f_+ d\nu$ and $\int f_- d\nu$ are finite, we say that f is integrable.

Remark 1. The abstract integral unifies both regular (Riemann) integral and infinite series. To see this, suppose that we have a sequence $(a_n)_{n=1}^\infty$ and we are interested in the summation. Consider the counting measure¹ ν defined on $(\mathbb{R}, \mathcal{B})$ satisfying $\nu(A) = \sum_{n=1}^\infty \mathbf{1}\{n \in A\}$, namely counting how many elements are included in a Borel set A .

We can also find any function $a : \mathbb{R} \rightarrow \mathbb{R}$ satisfying $a(n) = a_n$ (the values outside the set of positive integers do not matter in the integral). Then the infinite series can be written into integral with counting measure on positive integers:

$$\sum_{n=1}^\infty a_n = \int a d\nu.$$

You can use Definition 1—3 to verify the above equality (*exercise*).

The abstract integral inherits most natural properties such as linearity of the Riemann integral; see [1][Sec. 1.4] for example. Note that properties that hold a.e. does not have any impact on the integral values. For instance, if $f = g$, a.e. then $\int f d\nu = \int g d\nu$.

1.1 Connection between Lebesgue integrals and Riemann integrals

If ν is a Lebesgue measure on $\mathbb{R}$, then we have the following nice properties.

1. If f is Riemann integrable on $[a, b]$, then f is Lebesgue integrable on $[a, b]$, and the two integrals have the same value.
2. If f is bounded on $[a, b]$. Then f is Riemann integrable iff f is continuous a.e. on $[a, b]$.
3. If f is Lebesgue integrable on $[a, b]$ and

$$F(x) = \int_a^x f(t) dt,$$

then $F'(x) = f(x)$ a.e. on $[a, b]$.

4. If F is differentiable on $[a, b]$ and F' is Lebesgue integrable on $[a, b]$, then

$$F(b) - F(a) = \int_a^b F'(t) dt.$$

Property 2 says that for bounded functions, f is Lebesgue integrable if it is modification of a continuous function on a zero-measure set. Riemann integrable functions include piecewise continuous functions, but $\mathbf{1}_{\mathbb{Q}}$ is unfortunately not Riemann integrable—this issue is now resolved with Lebesgue integrability.

¹We may use consider the counting measure directly on the discrete set $\{1, 2, \dots\}$, which makes little difference in this example.

Some conventions. The integral with respect to the Lebesgue measure in $(\mathbb{R}^d, \mathcal{B}^d)$ is denoted by

$$\int f \, dx \quad \text{or} \quad \int f(x) \, dx \quad \text{or} \quad \int_{\mathbb{R}^d} f(x) \, dx$$

which is consistent with the Riemann integral. We may or may not use double/multiple integral notation such as $\int \int$ if there is a need.

When we say a property or equality holds **almost everywhere** or a.e. with respect to a measure ν if the set such that property or equality does not hold has zero measure. By default, ν is the Lebesgue measure. If we consider a probability measure instead of the Lebesgue measure, then we say **almost surely** or a.s. For example, $f(x) \geq 0$ a.e. means that the set $\{x : f(x) < 0\}$ has zero Lebesgue measure.

In almost every scenario, there is little difference of a property holding a.e. versus always, since the union of countably many zero-measure sets still has zero measure. For this reason, sometimes people don't even bother putting a.e. in their statement.

2 Convergence Theorems

The following result gives sufficient conditions under which the limits and integrals are interchangeable.

Theorem 1. Let $f_1, f_2, \dots$ be a sequence of Borel functions on $(\Omega, \mathcal{F}, \nu)$.

(i)(Fatou's lemma) If $f_n \geq 0$, then

$$\int \liminf_n f_n \, d\nu \leq \liminf_n \int f_n \, d\nu.$$

(ii)(Dominated convergence theorem) If $\lim_{n \rightarrow \infty} f_n = f$ a.e. and there exists an integrable function g such that $|f_n| \leq g$ a.e., then

$$\int \lim_{n \rightarrow \infty} f_n \, d\nu = \lim_{n \rightarrow \infty} \int f_n \, d\nu.$$

(iii)(Monotone convergence theorem) If $0 \leq f_1 \leq f_2 \leq \dots$ and $\lim_{n \rightarrow \infty} f_n = f$ a.e., then

$$\int \lim_{n \rightarrow \infty} f_n \, d\nu = \lim_{n \rightarrow \infty} \int f_n \, d\nu.$$

Remark 2. Intuitively, why is Fatou's lemma true? Recall that we defined integrals as supreme over simple functions...see [the min-max perspective](#) at this link.

Example 1. Recall the following example. Let $\mathbb{Q}$ be the set of rational numbers in $[0, 1]$, and let $A_1, A_2, \dots$ be the sets such that $A_1 \subsetneq A_2 \subsetneq \dots \subsetneq \mathbb{Q}$, $|A_n| = n$ and $\bigcup_{n=1}^{\infty} A_n = \mathbb{Q}$. Define $f_n(x)$ as follows:

$$f_n(x) = \begin{cases} 1, & x \in A_n \\ 0, & \text{otherwise.} \end{cases}$$

Use DCT (choosing $g(x) = 1$) or MCT (since $0 \leq f_1(x) \leq f_2(x) \leq \dots$) to derive

$$\int_0^1 \lim f_n(x)^2 \, dx = \lim \int_0^1 f_n(x) \, dx = 0.$$

Remark 3. This example shows that our newly defined Lebesgue integral overcomes the previous non-integrability issue in Riemann integrals, since we have enlarged the space of integrable functions under the new definition. More importantly, it is easy to check exchangeability is valid.

²Note: Prop. 1.4 on Jun Shao's book indicates that $\lim f_n(x)$ is Lebesgue measurable.

Example 2. Recall the following example. Let $f_n(x) = \begin{cases} 1, & x \in [n, n+1] \\ 0, & \text{otherwise} \end{cases}$, we have

$$\int \liminf f_n dx = 0 < \liminf \int f_n dx = 1.$$

Remark 4. This example shows when we have strict inequality in Fatou's lemma. Also, note that we cannot apply DCT, since the dominating function is $\sup_n f_n(x) = 1$ for all $x \geq 0$, which is not integrable. Sometimes, when DCT cannot be applied, it is a good indication that integral and limit cannot be exchanged.

Example 3. (Interchangeability of differentiation and integration). Let $(\Omega, \mathcal{F}, \nu)$ be a measure space and, for any fixed $\theta \in \mathbb{R}$, let $f(\omega, \theta)$ be a Borel function on Ω . Suppose that $\partial f(\omega, \theta)/\partial \theta$ exists a.e. for $\theta \in (a, b) \in \mathbb{R}$ and that $|\partial f(\omega, \theta)/\partial \theta| \leq g(\omega)$ a.e., where g is an integrable function on Ω . Then, for each $\theta \in (a, b)$, $\partial f(\omega, \theta)/\partial \theta$ is integrable and, by Theorem 1(ii),

$$\frac{d}{d\theta} \int f(\omega, \theta) d\nu = \int \frac{\partial f(\omega, \theta)}{\partial \theta} d\nu.$$

3 Fubini's theorem

The following theorem states how to evaluate an integral w.r.t. a product measure via iterated integration.

Theorem 2. (Fubini's theorem). Let ν_i be a σ -finite measure on $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$, and let f be a Borel function on $\prod_{i=1}^2 (\Omega_i, \mathcal{F}_i)$. Suppose that either $f \geq 0$ or f is integrable w.r.t. $\nu_1 \times \nu_2$. Then

$$g(\omega_2) = \int_{\Omega_1} f(\omega_1, \omega_2) d\nu_1$$

exists a.e. ν_2 and defines a Borel function on Ω_2 whose integral w.r.t. ν_2 exists, and

$$\int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2) d\nu_1 \times \nu_2 = \int_{\Omega_2} \left[\int_{\Omega_1} f(\omega_1, \omega_2) d\nu_1 \right] d\nu_2.$$

Remark 5. When ν_1, ν_2 are counting measures, then Fubini's theorem says we can exchange summations of infinite series under some conditions. For example, for nonnegative array $(a_{ij})_{i,j \geq 1}$,

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}.$$

If one of ν_1, ν_2 is a counting measure, then we have guarantees for exchanging $\int$ with $\sum$ under stated conditions.

4 Radon-Nikodym derivative

In elementary probability/statistics, we define probability density function *pdf* for some continuous random variables. How to define density functions in the general setting? We need RN derivative in measure theory.

Given two measures λ and ν defined on the same measurable space $(\Omega, \mathcal{F})$, we say λ is *absolutely continuous* w.r.t. ν and write $\lambda \ll \nu$ if the following holds

$$\nu(A) = 0 \quad \text{implies} \quad \lambda(A) = 0 \quad \text{for all } A \in \mathcal{F}.$$

Check: $\lambda(A) := \int_A f d\nu$, $A \in \mathcal{F}$ satisfies $\lambda \ll \nu$.

Theorem 3. (RN derivative) Let ν and λ be two measures on $(\Omega, \mathcal{F})$ and ν be σ -finite. If $\lambda \ll \nu$, then there exists a nonnegative Borel function f on Ω such that the following holds

$$\lambda(A) = \int_A f d\nu, \quad \forall A \in \mathcal{F}.$$

Furthermore, f^3 is unique a.e. ν , i.e. if $\lambda(A) = \int_A g d\nu$ for any $A \in \mathcal{F}$, then $f = g$ a.e. ν .

Remark 6. In statistics, usually we have a measure (e.g. the measure induced by random variables), and

$$f \text{ is } \begin{cases} \text{pdf} & \text{if } \nu \text{ is Lebesgue measure} \\ \text{pmf} & \text{if } \nu \text{ is counting measure.} \end{cases}$$

RN derivatives satisfy linearity, chain rule, etc.

Example 4. (i) Let $\Omega = \{0, 1, 2, 3, \dots\}$, $\nu_1(\{n\}) = P_n$.

(ii) Let $\Omega = \mathbb{R}$, ν_2 be the Gaussian measure.

(iii) (Mixed) Let $X_1 \sim \bar{\nu}_1$, $X_2 \sim \nu_2$, where $\bar{\nu}_1$ is defined on $(\mathbb{R}, \mathcal{B})$,

$$\nu_3 = \frac{1}{2}\bar{\nu}_1 + \frac{1}{2}\nu_2$$

is not a pdf or pmf in the classical setting, but $\nu_3 \ll (\text{Lebesgue measure} + \text{counting measure})$.

Example 5. Recall the example from Lecture 2: suppose that $U \sim \text{Unif}([0, 2\pi])$, let $\begin{cases} X = \cos U \\ Y = \sin U \end{cases}$ and $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ be the unit circle. The induced measure of (X, Y) is not absolutely continuous w.r.t. the Lebesgue measure m in $\mathbb{R}^2$, since $m(S) = 0$ but the induced measure of (X, Y) is supported on S .

Proposition 1. (change of variables) Let X be a random k -vector with a Lebesgue pdf f_X and let $Y = g(X)$, where g is a Borel function from $(\mathbb{R}^k, \mathcal{B}^k)$ to $(\mathbb{R}^k, \mathcal{B}^k)$. Let $A_1, \dots, A_m$ be disjoint sets in $\mathcal{B}^k$ such that $\mathbb{R}^k - (A_1 \cup \dots \cup A_m)$ has Lebesgue measure 0 and g on A_j is one-to-one with a nonvanishing Jacobian, i.e., the determinant $\text{Det}(\partial g(x)/\partial x) \neq 0$ on $A_j, j = 1, \dots, m$. Then Y has the following Lebesgue pdf:

$$f_Y(x) = \sum_{j=1}^m |\text{Det}(\partial g_j^{-1}(x)/\partial x)| f_X(g_j^{-1}(x)).$$

5 Further reading

The main references for measure theory in this lecture note are [3, 4, 1]. Most of the materials are standard. The connection between the Riemann integral and the Lebesgue integral is taken from [3][Thm. 11.33].

For more explanation about almost everywhere and the role of zero measure sets, a reference is “The role played by sets of measure zero” Chap. 1 of [2]. For an example where the condition for Fubini’s theorem is necessary, see [1][Example 1.7.5].

References

- [1] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 5th edition, 2019.

³Note: f is called RN derivative or RN density.

- [2] W. Rudin. *Real and Complex Analysis*. Higher Mathematics Series. McGraw-Hill Education, 1987.
- [3] Walter Rudin. *Principles of mathematical analysis*, volume 3. McGraw-hill New York.
- [4] Jun Shao. *Mathematical Statistics*. Springer Science & Business Media, 2nd edition, 2013.