Lecture 4

Instructor: Quan Zhou

For more details about the materials covered in this note, see Chapters 5.1 and 5.2 of Resnick [3] and Chapter 1.4 of Durrett [2].

4.1 Simple functions

We first give a formal definition of simple functions. Note that we require them to be real-valued and measurable in the definition.

Definition 4.1. Given a measurable space $(\Omega, \mathcal{F})$, a simple function f is a measurable mapping from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with a finite range.

Remark 4.1. Given a simple function f, let $\{a_1, \dots, a_k\}$ denote its range, where $a_1, \dots, a_k$ are distinct real numbers, and $A_i = f^{-1}(a_i)$. Then, $A_1, \dots, A_k$ are disjoint and measurable, and we can express f by $f = \sum_{i=1}^k a_i \mathbb{1}_{A_i}$.

Lemma 4.1. *Let $f, g \colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be simple functions. Then, $f + g$ is also a simple function. For any $a \in \mathbb{R}$, af is also a simple function.*

Proof. Try it yourself. $\square$

Lemma 4.2. *Let $f \colon (\Omega, \mathcal{F}) \to ([0, \infty], \mathcal{B}([0, \infty]))$ be a non-negative extended real-valued Borel function.*[1] *There exists a sequence of simple functions $\{\varphi_n\}$ such that $0 \leq \varphi_n \uparrow f$ pointwisely; i.e. $0 \leq \varphi_1 \leq \varphi_2 \leq \cdots \leq f$ and $\lim_{n\to\infty} \varphi_n(x) = f(x)$ for every $x \in \Omega$. Further, if f is bounded, then $\varphi_n \uparrow f$ uniformly.*

Proof. Here we only prove the first claim. The idea is that we construct φ_n by (1) truncating f at height 2^n, and (2) rounding $f(x)$ to the greatest integer multiple of 2^{-n} that does not exceed $f(x)$. Explicitly, we define the simple function φ_n by

$$\varphi_n(x) = \sup \left\{ \frac{k}{2^n} \colon k = 0, 1, 2, \dots, \frac{k}{2^n} \leq \min\{f(x), 2^n\} \right\}.$$

It is not difficult to show that (1) $\varphi_n \in [0, \infty)$, (2) φ_n is measurable, and (3) $\varphi_n \leq \varphi_{n+1}$ for every n. To prove $\varphi_n(x) \to f(x)$, first assume that $f(x) < \infty$. Then the approximation error $|\varphi_n(x) - f(x)|$ vanishes at rate 2^{-n}. If $f(x) = \infty$, then $\varphi_n(x) = 2^n \to \infty$. $\square$

[1] It can be shown that $\mathcal{B}(\bar{\mathbb{R}}) = \{A \subset \bar{\mathbb{R}} \colon A \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\}$.

4.2 Lebesgue integral

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f\colon (\Omega, \mathcal{F}) \to (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$. Now we show how to construct the Lebesgue integral $\int_\Omega f d\mu$. Other notation often used for denoting Lebesgue integral includes

$$\int_\Omega f(\omega)\mu(d\omega), \int_\Omega f(\omega)d\mu(\omega), \int f d\mu, \int f.$$

It is better to use the first two for clarity. There are a few ways to construct $\int_\Omega f d\mu$. Here we present the construction that is commonly used in real analysis books, which is slightly different from the approaches in Resnick's and Durrett's books.

Step 1. Consider a non-negative simple function $\varphi(\omega) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(\omega)$ where $a_1, \dots, a_n \in [0, \infty)$ and $A_1, \dots, A_n$ are disjoint measurable subsets of Ω that partition Ω. Define

$$\int_\Omega \varphi d\mu = \sum_{i=1}^n a_i \mu(A_i). \tag{1}$$

Note that the right-hand side makes sense since φ is assumed measurable (and thus $A_i \in \mathcal{F}$). The definition immediately implies that, for any measurable set A, $\int_\Omega \mathbb{1}_A d\mu = \mu(A)$.

Step 2. For $f\colon (\Omega, \mathcal{F}) \to ([0, \infty], \mathcal{B}([0, \infty]))$, define

$$\int_\Omega f d\mu = \sup\left\{ \int_\Omega \varphi d\mu : 0 \le \varphi \le f, \text{ and } \varphi \text{ is simple} \right\}. \tag{2}$$

We need to prove that when f is a non-negative simple function, the two definitions in Step 1 and Step 2 give the same value (see Lemma 4.3).

Equivalently, we can define $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega \varphi_n d\mu$ using any sequence of simple functions φ_n such that $0 \le \varphi_n \uparrow f$. This approach requires us to check that (1) $\int_\Omega \varphi_n d\mu$ does converge, and (2) the limit does not depend on the sequence $\{\varphi_n\}$ we pick [3, §5.2.2]. In the next lecture, we will prove the monotone convergence theorem (MCT) using the definition given in (2). The equivalence of the two definitions then follows from MCT.

Lemma 4.3. *For the Lebesgue integral of a non-negative simple function f, the two definitions given in* (1) *and* (2) *are the same.*

Proof. Since f is simple, f can be written as $f(\omega) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(\omega)$ where $a_1, \dots, a_n \in [0, \infty)$ and $A_1, \dots, A_n$ are disjoint subsets of Ω. Let

$$I_1 = \sum_{i=1}^n a_i \mu(A_i).$$

We need to show that $I_1 = I_2$, where

$$I_2 = \sup \left\{ \int_\Omega \varphi d\mu : \ 0 \leq \varphi \leq f, \text{ and } \varphi \text{ is simple} \right\}.$$

Since f is simple, we immediately have that $I_1 \leq I_2$ by choosing $\varphi = f$. Next, for any two non-negative simple functions f, φ such that $f \geq \varphi$, one can show that $\int f d\mu \geq \int \varphi d\mu$ using the definition of the Lebesgue integral given in Step 1; see Proposition 4.1(iv). Taking sup on both sides we get $I_1 \geq I_2$. Hence, $I_1 = I_2$. □

Step 3. For $f \colon (\Omega, \mathcal{F}) \to (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$, define the positive part and negative part of f as follows:

$$f^+(\omega) = f(\omega) \vee 0, \qquad f^-(\omega) = (-f(\omega)) \vee 0.$$

Hence, $f = f^+ - f^-$ and $|f| = f^+ + f^-$. If at least one of $\int_\Omega f^+ d\mu$ and $\int_\Omega f^- d\mu$ is finite, we say f is quasi-integrable and the integral exists (or is defined). If both are finite, we say f is integrable, which is equivalent to requiring that $\int_\Omega |f| d\mu < \infty$. When the integral exists, we define it by

$$\int_\Omega f d\mu = \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu.$$

Definition 4.2. When μ is a probability measure, we define $E[X] = \int_\Omega X d\mu$, which is called the expectation of X.

Definition 4.3. For $A \in \mathcal{F}$, define

$$\int_A f d\mu = \int_\Omega \mathbb{1}_A f d\mu,$$

provided that the right-hand side is defined.

Example 4.1. Consider the probability space $(\Omega, \mathcal{F}, \delta_x)$ where δ_x denotes a unit point mass on x. Then, for any measurable f, $\int_\Omega f d\delta_x = f(x)$. To prove this, first check that it holds true for non-negative simple functions. Next, for $f \geq 0$ and any simple function $0 \leq \phi \leq f$, we have

$$\int_\Omega \phi(\omega)\delta_x(d\omega) = \phi(x) \leq f(x).$$

Now if we choose $\phi(\omega) = f(x)\mathbb{1}_{\{x\}}(\omega)$, the supremum is achieved and thus $\int f d\delta_x = f(x)$. Finally, for a general measurable f, we have $\int f d\delta_x = f^+(x) - f^-(x) = f(x)$.

Example 4.2. Consider the measure space $(\Omega, \mathcal{P}(\Omega), \#)$ where $\Omega = \{a_1, a_2, \dots\}$ and $\#$ is the counting measure. Then, for any measurable f, it can be shown that $\int_\Omega f d\# = \sum_i f(a_i)$.

4.3 Properties of Lebesgue integral

Proposition 4.1. *Let f, g be Lebesgue integrable functions defined on $(\Omega, \mathcal{F}, \mu)$.*

(i) If $f \geq 0$ a.e., then $\int f d\mu \geq 0$.

(ii) $\forall\, a \in \mathbb{R}$, $\int a f d\mu = a \int f d\mu$.

(iii) $\int (f+g) d\mu = \int f d\mu + \int g d\mu$.

(iv) If $g \leq f$ a.e., then $\int f d\mu \geq \int g d\mu$.

(v) If $g = f$ a.e., then $\int f d\mu = \int g d\mu$.

(vi) $|\int f d\mu| \leq \int |f| d\mu$.

Remark 4.2. "a.e." means almost everywhere. When we say some property holds a.e., it means that there exists a set $N \in \mathcal{F}$ with $\mu(N) = 0$ such that the property holds on N^c. When μ is a probability measure, we often say almost surely, which is abbreviated as "a.s.".

Proof of part (iii). We need to prove the claim using the definition of the Lebesgue integral, i.e. the three-step construction. So let's start from non-negative simple functions and prove the property using the corresponding definition of the Lebesgue integral.

Step (1). We claim that for two non-negative simple functions f, g, we have $\int (f+g) d\mu = \int f d\mu + \int g d\mu$. By definition, we can express f, g as

$$f = \sum_{i=1}^{m} a_i \mathbb{1}_{A_i}, \qquad g = \sum_{i=1}^{n} b_i \mathbb{1}_{B_i},$$

where a_i, b_i are non-negative real numbers and $\{A_i\}$ (or $\{B_i\}$) is a partition of Ω. Hence, we have

$$(f+g)(\omega) = \sum_{i=1}^{m} \sum_{j=1}^{n} (a_i + b_j) \mathbb{1}_{A_i \cap B_j}(\omega),$$

and clearly $f + g$ is also a non-negative simple function. Further, $\{A_i \cap B_j : 1 \leq i \leq m, 1 \leq j \leq n\}$ forms a new partition of Ω. Therefore, by the definition of Lebesgue integral for non-negative simple functions, we have

$$\begin{aligned}
\int (f+g) d\mu &= \sum_{i=1}^{m} \sum_{j=1}^{n} (a_i + b_j) \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{m} a_i \sum_{j=1}^{n} \mu(A_i \cap B_j) + \sum_{j=1}^{n} b_j \sum_{i=1}^{m} \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{m} a_i \mu(A_i) + \sum_{j=1}^{n} b_j \mu(B_j) \\
&= \int f d\mu + \int g d\mu.
\end{aligned}$$

Step (2). We prove that for any two non-negative measurable functions f and g, we also have $\int (f+g) d\mu = \int f d\mu + \int g d\mu$. We do not use the definition given in (2). Instead, we use the alternative definition discussed in Step 2 in Section 4.2. (Again, it follows from the monotone convergence theorem, which will be proven in the next lecture.) Choose sequences of non-negative simple functions $\{f_n\}$ and $\{g_n\}$ such that $f_n \uparrow f$ and $g_n \uparrow g$. They always exist by

Lemma 4.2. Since $(f_n + g_n) \uparrow (f + g)$, we find that

$$\begin{aligned}\int (f+g)d\mu &= \lim_{n\to\infty} \int (f_n + g_n)d\mu \\ &= \lim_{n\to\infty} \left(\int f_n d\mu + \int g_n d\mu \right) \\ &= \lim_{n\to\infty} \int f_n d\mu + \lim_{n\to\infty} \int g_n d\mu \\ &= \int f d\mu + \int g d\mu.\end{aligned}$$

Step (3). We first prove that for non-negative integrable functions f_1, f_2, we have $\int (f_1 - f_2)d\mu = \int f_1 d\mu - \int f_2 d\mu$. Let $f = f_1 - f_2$. Since $f = f^+ - f^-$, we have $f_1 + f^- = f_2 + f^+$, which gives

$$\int f_1 d\mu + \int f^- d\mu = \int f_2 d\mu + \int f^+ d\mu.$$

By definition, $\int f d\mu = \int f^+ d\mu - \int f^- d\mu = \int f_1 d\mu - \int f_2 d\mu$.

Now we consider arbitrary *integrable* functions f, g. Notice that

$$f + g = (f^+ - f^-) + (g^+ - g^-) = (f^+ + g^+) - (f^- + g^-).$$

Hence, $\int (f+g)d\mu = \int (f^+ + g^+)d\mu - \int (f^- + g^-)d\mu = \int f d\mu + \int g d\mu$.

The proof is complete. □

Proof of the remaining part(s). Try it yourself. □

Proposition 4.2. *Assume X, Y are random variables defined on the same probability space such that $E[X]$ and $E[Y]$ exist (may be equal to infinity).*

(i) If $E[X^+] < \infty$ and $E[Y^+] < \infty$, then $E[X + Y] = E[X] + E[Y]$. The condition can also be replaced by $E[X^-] < \infty$ and $E[Y^-] < \infty$.

(ii) $\forall\, a, b \in \mathbb{R}$, $E[aX + b] = aE[X] + b$.

(iii) If $X \geq Y$ a.e., then $E[X] \geq E[Y]$.

Proof. Try it yourself. □

Theorem 4.1. *Let $f\colon [a,b] \to \mathbb{R}$ be Riemann integrable for some $-\infty < a \le b < \infty$. Then, f is Lebesgue integrable on $[a,b]$ and the two integrals coincide.*

Proof. See the textbook. $\square$

Remark 4.3. See a real analysis textbook for the construction of Riemann integrals. A function $f\colon [a,b] \to \mathbb{R}$, is Riemann integrable if and only if f is bounded and the set of discontinuity points of f has Lebesgue measure zero (the latter is known as "Lebesgue's integrability criterion"). Note that an improperly Riemann integrable function f on $[a,\infty)$ or $(a,b]$ may not be Lebesgue integrable.

References

[1] Dennis D. Cox. *The Theory of Statistics and Its Applications*. Unpublished.

[2] Rick Durrett. *Probability: Theory and Examples*, volume 49. Cambridge university press, 2019.

[3] Sidney Resnick. *A Probability Path*. Springer, 2019.

[4] Halsey Royden and Patrick Michael Fitzpatrick. *Real analysis*. China Machine Press, 2010.

[5] Terence Tao. *Analysis*, volume 1. Springer, 2006.