Course: Theory of Probability I
Term: Fall 2013
Instructor: Gordan Zitkovic

Lecture 3
The Lebesgue Integral

The construction of the integral

Unless expressly specified otherwise, we pick and fix a measure space $(S, \mathcal{S}, \mu)$ and assume that all functions under consideration are defined there.

Definition 3.1 (Simple functions)**.** A function $f \in \mathcal{L}^0(S, \mathcal{S}, \mu)$ is said to be **simple** if it takes only a finite number of values.

The collection of all simple functions is denoted by $\mathcal{L}^{\text{Simp},0}$ (more precisely by $\mathcal{L}^{\text{Simp},0}(S, \mathcal{S}, \mu)$) and the family of non-negative simple functions by $\mathcal{L}_+^{\text{Simp},0}$. Clearly, a simple function $f : S \to \mathbb{R}$ admits a (not necessarily unique) representation

$$f = \sum_{k=1}^{n} \alpha_k \mathbf{1}_{A_k}, \tag{3.1}$$

for $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and $A_1, \dots, A_n \in \mathcal{S}$. Such a representation is called the **simple-function representation** of f.

When the sets A_k, $k = 1, \dots, n$ are intervals in $\mathbb{R}$, the graph of the simple function f looks like a collection of steps (of heights $\alpha_1, \dots, \alpha_n$). For that reason, the simple functions are sometimes referred to as *step functions*. The Lebesgue integral is very easy to define for non-negative simple functions and this definition allows for further generalizations[1]:

[1] In fact, the progression of events you will see in this section is typical for measure theory: you start with indicator functions, move on to non-negative simple functions, then to general non-negative measurable functions, and finally to (not-necessarily-non-negative) measurable functions. This approach is so common, that it has a name - **the Standard Machine**.

Definition 3.2 (Lebesgue integration for simple functions)**.** For $f \in \mathcal{L}_+^{\text{Simp},0}$ we define the **(Lebesgue) integral** $\int f \, d\mu$ of f with respect to μ by

$$\int f \, d\mu = \sum_{k=1}^{n} \alpha_k \mu(A_k) \in [0, \infty],$$

where $f = \sum_{k=1}^{n} \alpha_k \mathbf{1}_{A_k}$ is a simple-function representation of f,

Problem 3.1. Show that the Lebesgue integral is well-defined for simple functions, i.e., that the value of the expression $\sum_{k=1}^{n} \alpha_k \mu(A_k)$ does not depend on the choice of the simple-function representation of f.

Remark 3.3.

1. It is important to note that $\int f\, d\mu$ can equal $+\infty$ even if f never takes the value $+\infty$. It is enough to pick $f = \mathbf{1}_A$ where $\mu(A) = +\infty$ - indeed, then $\int f\, d\mu = 1\mu(A) = \infty$, but f only takes values in the set $\{0, 1\}$. This is one of the reasons we start with *non-negative* functions. Otherwise, we would need to deal with the (unsolvable) problem of computing $\infty - \infty$. On the other hand, such examples cannot be constructed when μ is a finite measure. Indeed, it is easy to show that when $\mu(S) < \infty$, we have $\int f\, d\mu < \infty$ for all $f \in \mathcal{L}_+^{\text{Simp},0}$.

2. One can think of the (simple) Lebesgue integral as a generalization of the notion of (finite) additivity of measures. Indeed, if the simple-function representation of f is given by $f = \sum_{k=1}^n \mathbf{1}_{A_k}$, for pairwise disjoint $A_1, \dots, A_n$, then the equality of the values of the integrals for two representations $f = \mathbf{1}_{\cup_{k=1}^n A_k}$ and $f = \sum_{k=1}^n \mathbf{1}_{A_k}$ is a simple restatement of finite additivity. When $A_1, \dots, A_n$ are not disjoint, then the finite additivity gives way to finite subadditivity

$$\mu(\cup_{k=1}^n A_k) \leq \sum_{k=1}^n \mu(A_k),$$

but the integral $\int f\, d\mu$ "takes into account" those x which are covered by more than one A_k, $k = 1, \dots, n$. Take, for example, $n = 2$ and $A_1 \cap A_2 = C$. Then

$$f = \mathbf{1}_{A_1} + \mathbf{1}_{A_2} = \mathbf{1}_{A_1 \setminus C} + 2\mathbf{1}_C + \mathbf{1}_{A_2 \setminus C},$$

and so

$$\int f\, d\mu = \mu(A_1 \setminus C) + \mu(A_2 \setminus C) + 2\mu(C) = \mu(A_1) + \mu(A_2) + \mu(C).$$

It is easy to see that $\mathcal{L}_+^{\text{Simp},0}$ is a **convex cone**, i.e., that it is closed under finite linear combinations with non-negative coefficients. The integral map $f \mapsto \int f\, d\mu$ preserves this structure:

Problem 3.2. For $f_1, f_2 \in \mathcal{L}_+^{\text{Simp},0}$ and $\alpha_1, \alpha_2 \geq 0$ we have

1. if $f_1(x) \leq f_2(x)$ for all $x \in S$ then $\int f_1\, d\mu \leq \int f_2\, d\mu$, and
2. $\int (\alpha_1 f_1 + \alpha_2 f_2)\, d\mu = \alpha_1 \int f_1\, d\mu + \alpha_2 \int f_2\, d\mu$.

Having defined the integral for $f \in \mathcal{L}_+^{\text{Simp},0}$, we turn to general[2] non-negative measurable functions. In fact, at no extra cost we can consider a slightly larger set consisting of all measurable $[0, \infty]$-valued functions which we denote by $\mathcal{L}_+^0\left([0, \infty]\right)$.

[2] Even though there is no obvious advantage at this point of integrating a function which takes the value $+\infty$, it will become clear soon how convenient it really is.

Definition 3.4 (Lebesgue integral for nonnegative functions)**.** For a function $f \in \mathcal{L}^0_+([0,\infty])$, we define its **Lebesgue integral** $\int f\,d\mu$ by

$$\int f\,d\mu = \sup\left\{\int g\,d\mu \,:\, g \in \mathcal{L}_+^{\text{Simp},0},\ g(x) \le f(x),\ \forall\, x \in S\right\} \in [0,\infty].$$

Note: While there is no question that this definition produces a unique number $\int f\,d\mu$, one can wonder if it matches the previously given definition of the Lebesgue integral for simple functions. A simple argument based on the monotonicity property of part 1. of Problem 3.2 can be used to show that this is, indeed, the case.

Problem 3.3. Show that $\int f\,d\mu = \infty$ if there exists a measurable set A with $\mu(A) > 0$ such that $f(x) = \infty$ for $x \in A$. On the other hand, show that $\int f\,d\mu = 0$ for f of the form

$$f(x) = \infty \mathbf{1}_A(x) = \begin{cases} \infty, & x \in A, \\ 0, & x \ne A, \end{cases}$$

whenever $\mu(A) = 0$.

Note: Relate this to our convention that $\infty \times 0 = 0 \times \infty = 0$.

Finally, we are ready to define the integral for general measurable functions. Each $f \in \mathcal{L}^0$ can be written as a difference of two functions in $\mathcal{L}^0_+$ in many ways. There exists a decomposition which is, in a sense, minimal. We define

$$f^+ = \max(f, 0),\ f^- = \max(-f, 0),$$

so that $f = f^+ - f^-$ (and both f^+ and f^- are measurable). The minimality we mentioned above is reflected in the fact that for each $x \in S$, at most one of f^+ and f^- is non-zero.

Definition 3.5 (Integrable functions)**.** A function $f \in \mathcal{L}^0$ is said to be **integrable** if

$$\int f^+\,d\mu < \infty \text{ and } \int f^-\,d\mu < \infty.$$

The collection of all integrable functions in $\mathcal{L}^0$ is denoted by $\mathcal{L}^1$. The family of integrable functions is tailor-made for the following definition:

Definition 3.6 (The Lebesgue integral)**.** For $f \in \mathcal{L}^1$, we define the **Lebesgue integral** $\int f\,d\mu$ of f by

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu.$$

Remark 3.7.

1. We have seen so far two cases in which an integral for a function $f \in \mathcal{L}^0$ can be defined: when $f \ge 0$ or when $f \in \mathcal{L}^1$. It is possible to combine the two and define the Lebesgue integral for all functions $f \in \mathcal{L}^0$ with $f^- \in \mathcal{L}^1$. The set of all such functions is denoted by

$\mathcal{L}^{0-1}$ and we set

$$\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu \in (-\infty, \infty], \text{ for } f \in \mathcal{L}^{0-1}.$$

Note that no problems of the form $\infty - \infty$ arise here, and also note that, like $\mathcal{L}^0_+$, $\mathcal{L}^{0-1}$ is only a convex cone, and not a vector space. While the notation $\mathcal{L}^0$ and $\mathcal{L}^1$ is quite standard, the one we use for $\mathcal{L}^{0-1}$ is not.

2. For $A \in \mathcal{S}$ and $f \in \mathcal{L}^{0-1}$ we usually write $\int_A f\, d\mu$ for $\int f \mathbf{1}_A\, d\mu$.

Problem 3.4. Show that the Lebesgue integral remains a monotone operation in $\mathcal{L}^{0-1}$. More precisely, show that if $f \in \mathcal{L}^{0-1}$ and $g \in \mathcal{L}^0$ are such that $g(x) \geq f(x)$, for all $x \in S$, then $g \in \mathcal{L}^{0-1}$ and $\int g\, d\mu \geq \int f\, d\mu$.

First properties of the integral

The wider the generality to which a definition applies, the harder it is to prove theorems about it. Linearity of the integral is a trivial matter for functions in $\mathcal{L}^{\text{Simp},0}_+$, but you will see how much we need to work to get it for $\mathcal{L}^0_+$. In fact, it seems that the easiest route towards linearity is through two important results: an approximation theorem and a convergence theorem. Before that, we need to pick some low-hanging fruit:

Problem 3.5. Show that for $f_1, f_2 \in \mathcal{L}^0_+([0,\infty])$ and $\alpha \in [0,\infty]$ we have

1. if $f_1(x) \leq f_2(x)$ for all $x \in S$ then $\int f_1\, d\mu \leq \int f_2\, d\mu$.
2. $\int \alpha f\, d\mu = \alpha \int f\, d\mu$.

Theorem 3.8 (Monotone convergence theorem)**.** *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^0_+([0,\infty])$ with the property that*

$$f_1(x) \leq f_2(x) \leq \dots \text{ for all } x \in S.$$

Then

$$\lim_n \int f_n\, d\mu = \int f\, d\mu,$$

where $f(x) = \lim_n f_n(x) \in \mathcal{L}^0_+([0,\infty])$, for $x \in S$.

Proof. The (monotonicity) property (1) of Problem 3.5 above implies that the sequence $\int f_n\, d\mu$ is non-decreasing and that $\int f_n\, d\mu \leq \int f\, d\mu$. Therefore, $\lim_n \int f_n\, d\mu \leq \int f\, d\mu$. To show the opposite inequality, we deal with the case $\int f\, d\mu < \infty$ and pick $\varepsilon > 0$ and $g \in \mathcal{L}^{\text{Simp},0}_+$ with $g(x) \leq f(x)$, for all $x \in S$ and $\int g\, d\mu \geq \int f\, d\mu - \varepsilon$ (the case $\int f\, d\mu = \infty$

is similar and left to the reader). For $0 < c < 1$, define the (measurable) sets $\{A_n\}_{n\in\mathbb{N}}$ by

$$A_n = \{f_n \geq cg\}, n \in \mathbb{N}.$$

By the increase of the sequence $\{f_n\}_{n\in\mathbb{N}}$, the sets $\{A_n\}_{n\in\mathbb{N}}$ also increase. Moreover, since the function cg satisfies $cg(x) \leq g(x) \leq f(x)$ for all $x \in S$ and $cg(x) < f(x)$ when $f(x) > 0$, the increasing convergence $f_n \to f$ implies that $\cup_n A_n = S$. By non-negativity of f_n and monotonicity,

$$\int f_n \, d\mu \geq \int f_n \mathbf{1}_{A_n} \, d\mu \geq c \int g \mathbf{1}_{A_n} \, d\mu,$$

and so

$$\sup_n \int f_n \, d\mu \geq c \sup_n \int g \mathbf{1}_{A_n} \, d\mu.$$

Let $g = \sum_{i=1}^k \alpha_i \mathbf{1}_{B_i}$ be a simple-function representation of g. Then

$$\int g \mathbf{1}_{A_n} \, d\mu = \int \sum_{i=1}^k \alpha_i \mathbf{1}_{B_i \cap A_n} \, d\mu = \sum_{i=1}^k \alpha_i \mu(B_i \cap A_n).$$

Since $A_n \nearrow S$, we have $A_n \cap B_i \nearrow B_i$, $i = 1, \dots, k$, and the continuity of measure implies that $\mu(A_n \cap B_i) \nearrow \mu(B_i)$. Therefore,

$$\int g \mathbf{1}_{A_n} \, d\mu \nearrow \sum_{i=1}^k \alpha_i \mu(B_i) = \int g \, d\mu.$$

Consequently,

$$\lim_n \int f_n \, d\mu = \sup_n \int f_n \, d\mu \geq c \int g \, d\mu, \text{ for all } c \in (0,1),$$

and the proof is completed when we let $c \to 1$. □

Remark 3.9.

1. The monotone convergence theorem is really about the robustness of the Lebesgue integral. Its stability with respect to limiting operations is one of the reasons why it is a de-facto “industry standard”.

2. The “monotonicity” condition in the monotone convergence theorem cannot be dropped. Take, for example $S = [0,1]$, $\mathcal{S} = \mathcal{B}([0,1])$, and $\mu = \lambda$ (the Lebesgue measure), and define

$$f_n = n \mathbf{1}_{(0,n^{-1}]}, \text{ for } n \in \mathbb{N}.$$

Then $f_n(0) = 0$ for all $n \in \mathbb{N}$ and $f_n(x) = 0$ for $n > \frac{1}{x}$ and $x > 0$. In either case $f_n(x) \to 0$. On the other hand

$$\int f_n \, d\lambda = n\lambda\left((0, \tfrac{1}{n}]\right) = 1,$$

so that

$$\lim_n \int f_n \, d\lambda = 1 > 0 = \int \lim_n f_n \, d\lambda.$$

We will see later that the while the equality of the limit of the integrals and the integral of the limit will not hold in general, they will always be ordered in a specific way, if the functions $\{f_n\}_{n\in\mathbb{N}}$ are non-negative (that will be the content of Fatou's lemma below).

Proposition 3.10 (Approximation by simple functions). *For each $f \in \mathcal{L}^0_+([0,\infty])$ there exists a sequence $\{g_n\}_{n\in\mathbb{N}} \in \mathcal{L}^{\mathrm{Simp},0}_+$ such that*

1. *$g_n(x) \leq g_{n+1}(x)$, for all $n \in \mathbb{N}$ and all $x \in S$,*
2. *$g_n(x) \leq f(x)$ for all $x \in S$,*
3. *$f(x) = \lim_n g_n(x)$, for all $x \in S$, and*
4. *the convergence $g_n \to f$ is uniform on each set of the form $\{f \leq M\}$, $M > 0$, and, in particular, on the whole S if f is bounded.*

Proof. For $n \in \mathbb{N}$, let A^n_k, $k = 1, \dots, n2^n$ be a collection of subsets of S given by

$$A^n_k = \{\tfrac{k-1}{2^n} \leq f < \tfrac{k}{2^n}\} = f^{-1}\Big([\tfrac{k-1}{2^n}, \tfrac{k}{2^n})\Big), \, k = 1, \dots, n2^n.$$

Note that the sets A^n_k, $k = 1, \dots, n2^n$ are disjoint and that the measurability of f implies that $A^n_k \in \mathcal{S}$ for $k = 1, \dots, n2^n$. Define the function $g_n \in \mathcal{L}^{\mathrm{Simp},0}_+$ by

$$g_n = \sum_{k=1}^{n2^n} \tfrac{k-1}{2^n} \mathbf{1}_{A^n_k} + n\mathbf{1}_{\{f \geq n\}}.$$

The statements 2., 3., and 4. follow immediately from the following three simple observations:

- $g_n(x) \leq f(x)$ for all $x \in S$,
- $g_n(x) = n$ if $f(x) = \infty$, and
- $g_n(x) > f(x) - 2^{-n}$ when $f(x) < n$.

Finally, we leave it to the reader to check the simple fact that $\{g_n\}_{n\in\mathbb{N}}$ is non-decreasing. □

Problem 3.6. Show, by means of an example, that the sequence $\{g_n\}_{n\in\mathbb{N}}$ would not necessarily be monotone if we defined it in the following way:

$$g_n = \sum_{k=1}^{n^2} \tfrac{k-1}{n} \mathbf{1}_{\{f \in [\frac{k-1}{n}, \frac{k}{n})\}} + n\mathbf{1}_{\{f \geq n\}}.$$

Proposition 3.11 (Linearity of the integral for non-negative functions). *For $f_1, f_2 \in \mathcal{L}^0_+([0,\infty])$ and $\alpha_1, \alpha_2 \geq 0$ we have*

$$\int (\alpha_1 f_1 + \alpha_2 f_2)\, d\mu = \alpha_1 \int f_1\, d\mu + \alpha_2 \int f_2\, d\mu.$$

Proof. Thanks to Problem 3.5 it is enough to prove the statement for $\alpha_1 = \alpha_2 = 1$. Let $\{g^1_n\}_{n\in\mathbb{N}}$ and $\{g^2_n\}_{n\in\mathbb{N}}$ be sequences in $\mathcal{L}^{\text{Simp},0}_+$ which approximate f^1 and f^2 in the sense of Proposition 3.10. The sequence $\{g_n\}_{n\in\mathbb{N}}$ given by $g_n = g^1_n + g^2_n$, $n \in \mathbb{N}$, has the following properties:

- $g_n \in \mathcal{L}^{\text{Simp},0}_+$ for $n \in \mathbb{N}$,
- $g_n(x)$ is a nondecreasing sequence for each $x \in S$,
- $g_n(x) \to f_1(x) + f_2(x)$, for all $x \in \mathcal{S}$.

Therefore, we can apply the linearity of integration for the simple functions and the monotone convergence theorem (Theorem 3.8) to conclude that

$$\begin{aligned}\int (f_1 + f_2)\, d\mu &= \lim_n \int (g^1_n + g^2_n)\, d\mu = \lim_n \left(\int g^1_n\, d\mu + \int g^2_n\, d\mu \right) \\ &= \int f_1\, d\mu + \int f_2\, d\mu. \end{aligned}$$ □

Corollary 3.12 (Countable additivity of the integral). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^0_+([0,\infty])$. Then*

$$\int \sum_{n\in\mathbb{N}} f_n\, d\mu = \sum_{n\in\mathbb{N}} \int f_n\, d\mu.$$

Proof. Apply the monotone convergence theorem to the partial sums $g_n = f_1 + \dots + f_n$, and use linearity of integration. □

Once we have established a battery of properties for non-negative functions, an extension to $\mathcal{L}^1$ is not hard. We leave it to the reader to prove all the statements in the following problem:

Problem 3.7. The family $\mathcal{L}^1$ of integrable functions has the following properties:

1. $f \in \mathcal{L}^1$ iff $\int |f|\, d\mu < \infty$,
2. $\mathcal{L}^1$ is a vector space,
3. $\left|\int f\, d\mu\right| \leq \int |f|\, d\mu$, for $f \in \mathcal{L}^1$.
4. $\int |f + g|\, d\mu \leq \int |f|\, d\mu + \int |g|\, d\mu$, for all $f, g \in \mathcal{L}^1$.

We conclude the present section with two results, which, together with the monotone convergence theorem, play the central role in the Lebesgue integration theory.

Theorem 3.13 (Fatou's lemma). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^0_+([0,\infty])$. Then*

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

Note: The strength of Fatou's lemma comes from the fact that, apart from non-negativity, it requires no special properties for the sequence $\{f_n\}_{n\in\mathbb{N}}$. Its conclusion is not as strong as that of the monotone convergence theorem, but it proves to be very useful in various settings because it gives an upper bound (namely $\liminf_n \int f_n \, d\mu$) on the integral of the non-negative function $\liminf f_n$.

Proof. Set $g_n(x) = \inf_{k\geq n} f_k(x)$, so that $g_n \in \mathcal{L}^0_+([0,\infty])$ and $g_n(x)$ is a non-decreasing sequence for each $x \in S$. The monotone convergence theorem and the fact that $\liminf f_n(x) = \sup_n g_n(x) = \lim_n g_n(x)$, for all $x \in S$, imply that

$$\int g_n \, d\mu \nearrow \int \liminf_n \, d\mu.$$

On the other hand, $g_n(x) \leq f_k(x)$ for all $k \geq n$, and so

$$\int g_n \, d\mu \leq \inf_{k\geq n} \int f_k \, d\mu.$$

Therefore,

$$\lim_n \int g_n \, d\mu \leq \lim_n \inf_{k\geq n} \int f_k \, d\mu = \liminf_n \int f_k \, d\mu.$$

□

Remark 3.14.

1. The inequality in the Fatou's lemma does not have to be equality, even if the limit $\lim_n f_n(x)$ exists for all $x \in S$. You can use the sequence $\{f_n\}_{n\in\mathbb{N}}$ of Remark 3.9 to see that.

2. Like the monotone convergence theorem, Fatou's lemma requires that all function $\{f_n\}_{n\in\mathbb{N}}$ be non-negative. This requirement is necessary - to see that, simply consider the sequence $\{-f_n\}_{n\in\mathbb{N}}$, where $\{f_n\}_{n\in\mathbb{N}}$ is the sequence of Remark 3.9 above.

Theorem 3.15 (Dominated convergence theorem). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^0$ with the property that there exists $g \in \mathcal{L}^1$ such that $|f_n(x)| \leq g(x)$, for all $x \in X$ and all $n \in \mathbb{N}$. If $f(x) = \lim_n f_n(x)$ for all $x \in S$, then $f \in \mathcal{L}^1$ and*

$$\int f \, d\mu = \lim_n \int f_n \, d\mu.$$

Note: The dominated convergence theorem combines the lack of monotonicity requirements of Fatou's lemma and the strong conclusion of the monotone convergence theorem. The price to be paid is the uniform boundedness requirement. There is a way to relax this requirement a little bit (using the concept of *uniform integrability*), but not too much. Still, it is an unexpectedly useful theorem.

Proof. The condition $|f_n(x)| \leq g(x)$, for all $x \in X$ and all $n \in \mathbb{N}$ implies that $g(x) \geq 0$, for all $x \in S$. Since $f_n^+ \leq g$, $f_n^- \leq g$ and $g \in \mathcal{L}^1$, we immediately have $f_n \in \mathcal{L}^1$, for all $n \in \mathbb{N}$. The limiting function

f inherits the same properties $f^+ \leq g$ and $f^- \leq g$ from $\{f_n\}_{n\in\mathbb{N}}$ so $f \in \mathcal{L}^1$, too.

Clearly $g(x) + f_n(x) \geq 0$ for all $n \in \mathbb{N}$ and all $x \in S$, so we can apply Fatou's lemma to get

$$\int g\, d\mu + \liminf_n \int f_n\, d\mu = \liminf_n \int (g + f_n)\, d\mu \geq \int \liminf_n (g + f_n)\, d\mu$$
$$= \int (g + f)\, d\mu = \int g\, d\mu + \int f\, d\mu.$$

In the same way (since $g(x) - f_n(x) \geq 0$, for all $x \in S$, as well), we have

$$\int g\, d\mu - \limsup_n \int f_n\, d\mu = \liminf_n \int (g - f_n)\, d\mu \geq \int \liminf_n (g - f_n)\, d\mu$$
$$= \int (g - f)\, d\mu = \int g\, d\mu - \int f\, d\mu.$$

Therefore

$$\limsup_n \int f_n\, d\mu \leq \int f\, d\mu \leq \liminf_n \int f_n\, d\mu,$$

and, consequently, $\int f\, d\mu = \lim_n \int f_n\, d\mu$. □

Remark 3.16.

Null sets

An important property - inherited directly from the underlying measure - is that it is blind to sets of measure zero. To make this statement precise, we need to introduce some language:

Definition 3.17. Let $(S, \mathcal{S}, \mu)$ be a measure space.

1. $N \in \mathcal{S}$ is said to be a **null set** if $\mu(N) = 0$.

2. A function $f : S \to \bar{\mathbb{R}}$ is called a **null function** if there exists a null set N such that $f(x) = 0$ for $x \in N^c$.

3. Two functions f, g are said to be **equal almost everywhere** - denoted by $f = g$, a.e. - if $f - g$ is a null function, i.e., if there exists a null set N such that $f(x) = g(x)$ for all $x \in N^c$.

Remark 3.18.

1. In addition to almost-everywhere equality, one can talk about the almost-everywhere version of any relation between functions which can be defined on points. For example, we write $f \leq g$, a.e. if $f(x) \leq g(x)$ for all $x \in S$, except, maybe, for x in some null set N.

2. One can also define the a.e. equality of sets: we say that $A = B$, a.e., for $A, B \in \mathcal{S}$ if $\mathbf{1}_A = \mathbf{1}_B$, a.e. It is not hard to show (do it!) that $A = B$ a.e., if and only if $\mu(A \triangle B) = 0$ (Remember that $\triangle$ denotes the symmetric difference: $A \triangle B = (A \setminus B) \cup (B \setminus A)$).

3. When a property (equality of functions, e.g.) holds almost everywhere, the set where it fails to hold is not necessarily null. Indeed, there is no guarantee that it is measurable at all. What is true is that it is *contained* in a measurable (and null) set. Any such (measurable) null set is often referred to as the **exceptional set**.

Problem 3.8. Prove the following statements:

1. The almost-everywhere equality is an equivalence relation between functions.

2. The family $\{A \in \mathcal{S} : \mu(A) = 0 \text{ or } \mu(A^c) = 0\}$ is a σ-algebra (the so-called **μ-trivial σ-algebra**).

The "blindness" property of the Lebesgue integral we referred to above can now be stated formally:

Proposition 3.19. *Suppose that $f = g$, a.e,. for some $f, g \in \mathcal{L}^0_+$. Then*

$$\int f\, d\mu = \int g\, d\mu.$$

Proof. Let N be an exceptional set for $f = g$, a.e., i.e., $f = g$ on N^c and $\mu(N) = 0$. Then $f\mathbf{1}_{N^c} = g\mathbf{1}_{N^c}$, and so $\int f\mathbf{1}_{N^c}\, d\mu = \int g\mathbf{1}_{N^c}\, d\mu$. On the other hand $f\mathbf{1}_N \leq \infty\mathbf{1}_N$ and $\int \infty\mathbf{1}_N\, d\mu = 0$, so, by monotonicity, $\int f\mathbf{1}_N\, d\mu = 0$. Similarly $\int g\mathbf{1}_N\, d\mu = 0$. It remains to use the additivity of integration to conclude that

$$\begin{aligned}\int f\, d\mu &= \int f\mathbf{1}_{N^c}\, d\mu + \int f\mathbf{1}_N\, d\mu \\ &= \int g\mathbf{1}_{N^c}\, d\mu + \int g\mathbf{1}_N\, d\mu = \int g\, d\mu.\end{aligned}$$ □

A statement which can be seen as a converse of Proposition 3.19 also holds:

Problem 3.9. If $f \in \mathcal{L}^0_+$ and $\int f\, d\mu = 0$, show that $f = 0$, a.e.

Hint: What is the negation of the statement "$f = 0$. a.e." for $f \in \mathcal{L}^0_+$?

The monotone convergence theorem and the dominated convergence theorem both require the sequence $\{f_n\}_{n\in\mathbb{N}}$ functions to converge for each $x \in S$. A slightly weaker notion of convergence is required, though:

Definition 3.20. A sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ is said to **converge almost everywhere** to the function f, if there exists a null set N such that

$$f_n(x) \to f(x) \text{ for all } x \in N^c.$$

Remark 3.21. If we want to emphasize that $f_n(x) \to f(x)$ for *all* $x \in S$, we say that $\{f_n\}_{n\in\mathbb{N}}$ converges to f **everywhere**.

Proposition 3.22 (Monotone (almost-everywhere) convergence theorem)**.** *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^0_+([0,\infty])$ with the property that*

$$f_n \le f_{n+1} \text{ a.e., for all } n \in \mathbb{N}.$$

Then

$$\lim_n \int f_n \, d\mu = \int f \, d\mu,$$

if $f \in \mathcal{L}^0_+$ and $f_n \to f$, a.e.

Proof. There are "$\infty + 1$ a.e.-statements" we need to deal with: one for each $n \in \mathbb{N}$ in $f_n \le f_{n+1}$, a.e., and an extra one when we assume that $f_n \to f$, a.e. Each of them comes with an exceptional set; more precisely, let $\{A_n\}_{n\in\mathbb{N}}$ be such that $f_n(x) \le f_{n+1}(x)$ for $x \in A_n^c$ and let B be such that $f_n(x) \to f(x)$ for $x \in B^c$. Define $A \in \mathcal{S}$ by $A = (\cup_n A_n) \cup B$ and note that A is a null set. Moreover, consider the functions $\tilde{f}$, $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ defined by $\tilde{f} = f\mathbf{1}_{A^c}$, $\tilde{f}_n = f_n\mathbf{1}_{A^c}$. Thanks to the definition of the set A, $\tilde{f}_n(x) \le \tilde{f}_{n+1}(x)$, for all $n \in \mathbb{N}$ and $x \in S$; hence $\tilde{f}_n \to \tilde{f}$, everywhere. Therefore, the monotone convergence theorem (Theorem 3.8) can be used to conclude that $\int \tilde{f}_n \, d\mu \to \int \tilde{f} \, d\mu$. Finally, Proposition 3.19 implies that $\int \tilde{f}_n \, d\mu = \int f_n \, d\mu$ for $n \in \mathbb{N}$ and $\int \tilde{f} \, d\mu = \int f \, d\mu$. □

Problem 3.10. State and prove a version of the dominated convergence theorem where the almost-everywhere convergence is used. Is it necessary for all $\{f_n\}_{n\in\mathbb{N}}$ to be dominated by g for all $x \in S$, or only almost everywhere?

Remark 3.23. There is a subtlety that needs to be pointed out. If a sequence $\{f_n\}_{n\in\mathbb{N}}$ of measurable functions converges to the function f *everywhere*, then f is necessarily a measurable function (see Proposition 1.23). However, if $f_n \to f$ only almost everywhere, there is no guarantee that f is measurable. There is, however, always a measurable function which is equal to f almost everywhere; you can take $\liminf_n f_n$, for example.

Additional Problems

Problem 3.11 (The monotone-class theorem)**.** Prove the following result, known as the *monotone-class theorem* (remember that $a_n \nearrow a$ means that a_n is a non-decreasing sequence and $a_n \to a$)

Hint: Use Theorems 3.10 and 2.12

> *Let $\mathcal{H}$ be a class of bounded functions from S into $\mathbb{R}$ satisfying the following conditions*
>
> 1. *$\mathcal{H}$ is a vector space,*
> 2. *the constant function 1 is in $\mathcal{H}$, and*
> 3. *if $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of non-negative functions in $\mathcal{H}$ such that $f_n(x) \nearrow f(x)$, for all $x \in S$ and f is bounded, then $f \in \mathcal{H}$.*
>
> *Then, if $\mathcal{H}$ contains the indicator $\mathbf{1}_A$ of every set A in some π-system $\mathcal{P}$, then $\mathcal{H}$ necessarily contains every bounded $\sigma(\mathcal{P})$-measurable function on $\mathcal{S}$.*

Problem 3.12 (A form of continuity for Lebesgue integration)**.** Let $(S, \mathcal{S}, \mu)$ be a measure space, and suppose that $f \in \mathcal{L}^1$. Show that for each $\varepsilon > 0$ there exists $\delta > 0$ such that if $A \in \mathcal{S}$ and $\mu(A) < \delta$, then $\left|\int_A f\, d\mu\right| < \varepsilon$.

Problem 3.13 (Sums as integrals)**.** In the measure space $(\mathbb{N}, 2^{\mathbb{N}}, \mu)$, let μ be the counting measure.

1. For a function $f : \mathbb{N} \to [0, \infty]$, show that
$$\int f\, d\mu = \sum_{n=1}^{\infty} f(n).$$
2. Use the monotone convergence theorem to show the following special case of Fubini's theorem
$$\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} a_{kn} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} a_{kn},$$
whenever $\{a_{kn} : k, n \in \mathbb{N}\}$ is a double sequence in $[0, \infty]$.
3. Show that $f : \mathbb{N} \to \mathbb{R}$ is in $\mathcal{L}^1$ if and only if the series
$$\sum_{n=1}^{\infty} f(n),$$
converges absolutely.

Problem 3.14 (A criterion for integrability)**.** Let $(S, \mathcal{S}, \mu)$ be a finite measure space. For $f \in \mathcal{L}^0_+$, show that $f \in \mathcal{L}^1$ if and only if

$$\sum_{n \in \mathbb{N}} \mu(\{f \geq n\}) < \infty.$$

Problem 3.15 (A limit of integrals)**.** Let $(S, \mathcal{S}, \mu)$ be a measure space, and suppose $f \in \mathcal{L}^1_+$ is such that $\int f \, d\mu = c > 0$. Show that the limit

$$\lim_n \int n \log\Big(1 + (f/n)^\alpha\Big) \, d\mu$$

exists in $[0, \infty]$ for each $\alpha > 0$ and compute its value.

Hint: Prove and use the inequality $\log(1 + x^\alpha) \leq \alpha x$, valid for $x \geq 0$ and $\alpha \geq 1$.

Problem 3.16 (Integrals converge but the functions don't ...)**.** Construct a sequence $\{f_n\}_{n \in \mathbb{N}}$ of continuous functions $f_n : [0,1] \to [0,1]$ such that $\int f_n \, d\mu \to 0$, but the sequence $\{f_n(x)\}_{n \in \mathbb{N}}$ is divergent for each $x \in [0,1]$.

Problem 3.17 (...or they do, but are not dominated)**.** Construct an sequence $\{f_n\}_{n \in \mathbb{N}}$ of continuous functions $f_n : [0,1] \to [0, \infty)$ such that $\int f_n \, d\mu \to 0$, and $f_n(x) \to 0$ for all x, but $f \notin \mathcal{L}^1$, where $f(x) = \sup_n f_n(x)$.

Problem 3.18 (Functions measurable in the generated σ-algebra)**.** Let $S \neq \emptyset$ be a set and let $f : S \to \mathbb{R}$ be a function. Prove that a function $g : S \to \mathbb{R}$ is measurable with respect to the pair $(\sigma(f), \mathcal{B}(\mathbb{R}))$ if and only if there exists a Borel function $h : \mathbb{R} \to \mathbb{R}$ such that $g = h \circ f$.

Problem 3.19 (A change-of-variables formula)**.** Let $(S, \mathcal{S}, \mu)$ and $(T, \mathcal{T}, \nu)$ be two measurable spaces, and let $F : S \to T$ be a measurable function with the property that $\nu = F_* \mu$ (i.e., ν is the push-forward of μ through F). Show that for every $f \in \mathcal{L}^0_+(T, \mathcal{T})$ or $f \in \mathcal{L}^1(T, \mathcal{T})$, we have

$$\int f \, d\nu = \int (f \circ F) \, d\mu.$$

Problem 3.20 (The Riemann Integral)**.** A finite collection $\Delta = \{t_0, \dots, t_n\}$, where $a = t_0 < t_1 < \dots < t_n = b$ and $n \in \mathbb{N}$, is called a **partition** of the interval $[a, b]$. The set of all partitions of $[a, b]$ is denoted by $P([a, b])$.

For a bounded function $f : [a, b] \to \mathbb{R}$ and $\Delta = \{t_0, \dots, t_n\} \in P([a, b])$, we define its **upper and lower Darboux sums** $U(f, \Delta)$ and $L(f, \Delta)$ by

$$U(f, \Delta) = \sum_{k=1}^n \left(\sup_{t \in (t_{k-1}, t_k]} f(t) \right) (t_k - t_{k-1})$$

and

$$L(f,\Delta) = \sum_{k=1}^{n} \left(\inf_{t\in(t_{k-1},t_k]} f(t) \right) (t_k - t_{k-1}).$$

A function $f : [a,b] \to \mathbb{R}$ is said to be **Riemann integrable** if it is bounded and

$$\sup_{\Delta\in P([a,b])} L(f,\Delta) = \inf_{\Delta\in P([a,b])} U(f,\Delta).$$

In that case the common value of the supremum and the infimum above is called the **Riemann integral** of the function f - denoted by $(R)\int_a^b f(x)\,dx$.

1. Suppose that a bounded Borel-measurable function $f : [a,b] \to \mathbb{R}$ is Riemann-integrable. Show that

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

2. Find an example of a bounded an Borel-measurable function $f : [a,b] \to \mathbb{R}$ which is not Riemann-integrable.

3. Show that every continuous function is Riemann integrable.

4. It can be shown that for a bounded Borel-measurable function $f : [a,b] \to \mathbb{R}$ the following criterion holds (and you can use it without proof):

   f is Riemann integrable if and only if there exists a Borel set $D \subseteq [a,b]$ with $\lambda(D) = 0$ such that f is continuous at x, for each $x \in [a,b] \setminus D$. Show that

   - all monotone functions are Riemann-integrable,
   - $f \circ g$ is Riemann integrable if $f : [c,d] \to \mathbb{R}$ is Riemann integrable and $g : [a,b] \to [c,d]$ is continuous,
   - products of Riemann-integrable functions are Riemann-integrable.

5. Let $([a,b], \mathcal{B}([a,b])^*, \lambda^*)$ denote the completion of $([a,b], \mathcal{B}([a,b]), \lambda)$. Show that any Riemann-integrable function on $[a,b]$ is $\mathcal{B}([a,b])^*$-measurable.

   Hint: Pick a sequence $\{\Delta_n\}_{n\in\mathbb{N}}$ in $P([a,b])$ so that $\Delta_n \subseteq \Delta_{n+1}$ and $U(f,\Delta_n) - L(f,\Delta_n) \to 0$. Using those partitions and the function f, define two sequences of Borel-measurable functions $\{\overline{f}_n\}_{n\in\mathbb{N}}$ and $\{\underline{f}_n\}_{n\in\mathbb{N}}$ so that $\underline{f}_n \nearrow \underline{f}$, $\overline{f}_n \searrow \overline{f}$, $\underline{f} \leq f \leq \overline{f}$, and $\int(\overline{f} - \underline{f})\,d\lambda = 0$. Conclude that f agrees with a Borel measurable function on a complement of a subset of the set $\{\underline{f} \neq \overline{f}\}$ which has Lebesgue measure 0.