Lecture Notes 9: Measure and Probability
Part A: Measure and Integration

Peter Hammond, loosely based on notes by Andrés Carvajal

Latest revision 2023 September 25th,
typeset from `measure23.tex`.

Recommended textbooks for further reading:
H.L. Royden *Real Analysis*;
R.M. Dudley *Real Analysis and Probability*

Philosophical, Methodological, and Historical Preface

Andrei Nikolayevich Kolmogorov (1933)
Grundbegriffe der Wahrscheinlichkeitsrechnung

This short monograph was the first to set out the fundamental abstract mathematical concept of a probability space.

A probability space is a particular kind of measure space, another abstract concept due to Borel, Lebesgue, and others, in which the probability attached to the whole space is 1.

Like any mathematical model, one based on a probability space "is always wrong, but may be useful".

Indeed, a probability space may, or may not, help us formulate:

- empirical models based on past data;
- predictive models intended to forecast what to expect in data that have not yet been observed.

Our journey starts with measure spaces, before venturing on to probability spaces.

Outline

Power Sets and Indicator Functions

Definition

Given any abstract set S, the power set of S
is the family $\mathcal{P}(S) := \{ T \mid T \subseteq S \}$ of all subsets of S.

Definition

Given any abstract set S and any $T \subseteq S$,
the mapping $S \ni s \mapsto 1_T(s) \in \{0, 1\} \subset \mathbb{R}$
is an indicator function of the set T just in case

$$1_T(s) = 1 \iff s \in T \quad \text{and} \quad 1_T(s) = 0 \iff s \notin T$$

Thus, the function $s \mapsto 1_T(s)$ "indicates" whether $s \in T$.

The Cardinality of a Finite Set

Definition

Given any finite set S, its cardinality, denoted by $\#S$, is the number of its distinct elements.

Remark

Much of mathematical logic has been concerned with extending the concept of cardinality to infinite sets.

Notation

Given any domain set X and any co-domain set Y, let $Y^X := \{\langle y(x)\rangle_{x \in X} \mid \forall x \in X : y(x) \in Y\}$, which is the Cartesian product of copies of Y, one for each element $x \in X$, denote the space of all functions $X \ni x \mapsto f(x) \in Y$.

Counting Finite Power Sets

Theorem

Given any finite set S of n elements,
one has $\#\mathcal{P}(S) = \#\{0,1\}^S = 2^n$.

Proof.

Evidently the mapping $\mathcal{P}(S) \ni T \mapsto 1_T(\cdot) \in \{0,1\}^S$ is a bijection, implying that $\#\mathcal{P}(S) = \#\{0,1\}^S$.

Furthermore $\{0,1\}^S = \{\langle y(s)\rangle_{s \in S} \mid \forall s \in S : y(s) \in \{0,1\}\}$.

When $\#S = n$, this is the Cartesian product of n copies of $\{0,1\}$.

Therefore $\#\{0,1\}^S = 2^n$. $\square$

This result helps explain why the power set $\mathcal{P}(S)$
is often denoted by 2^S, even when S is infinite.

Boolean Algebras, Sigma-Algebras, and Measurable Spaces

Definition

1. The family $\mathcal{A} \subseteq \mathcal{P}(S)$ is a Boolean algebra on S just in case
   - $\emptyset \in \mathcal{A}$;
   - $A \in \mathcal{A}$ implies that the complement $S \setminus A \in \mathcal{A}$;
   - if A, B lie in $\mathcal{A}$, then the union $A \cup B \in \mathcal{A}$.
2. The family $\Sigma \subseteq \mathcal{P}(S)$ is a σ-algebra just in case it is a Boolean algebra with the following stronger property: whenever $\{A_n\}_{n=1}^{\infty}$ is a countably infinite family of sets in Σ, then their union $\cup_{n=1}^{\infty} A_n \in \Sigma$.
3. The pair (S, Σ) is a measurable space just in case Σ is a σ-algebra.

Exercise

Prove that if $\mathcal{A} \subseteq \mathcal{P}(S)$ is a Boolean algebra on S, then $S \in \mathcal{A}$.

Simple Examples

1. Given any set S, the minimal σ-algebra is $\{\emptyset, S\}$.
2. Given any set S, the maximal σ-algebra is 2^S, the power set of all subsets of S.
3. If $\#S = 1$, the only σ-algebras on S are the minimal and the maximal, which coincide.
4. If $\#S = 2$, the only σ-algebras on S are the minimal and the maximal, which differ.
5. If $\#S \geq 3$, then for each $x \in S$ the family $\{\emptyset, \{x\}, S \setminus \{x\}, S\}$ is a σ-algebra on S that is neither minimal nor maximal.
6. In the real line $\mathbb{R}$, the family of all countable and pairwise disjoint collections $\cup_{k \in K} I_k$ of left-open and right-closed intervals $I_k = (a_k, b_k]$ is one particular σ-algebra (which you should verify as an exercise).

   What happens in $\mathbb{Q}$, the set of rational numbers?

Exercise

Exercise

Consider the countable family $\{(\frac{1}{n}, 1] \mid n \in \mathbb{N}\}$
of left-open and right-closed intervals in $\mathbb{Q}$.

The union $\bigcup_{n\in\mathbb{N}}(\frac{1}{n}, 1]$ *includes every member of* $(0, 1] \cap \mathbb{Q}$.

But it does not include 0.

So $\bigcup_{n\in\mathbb{N}}(\frac{1}{n}, 1] = (0.1]$.

Exercise on Boolean Algebras and Sigma-Algebras

Exercise

1. *Let $\mathcal{A}$ be a Boolean algebra on S.*
   Prove that if $A, B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$.
2. *Let Σ be a σ-algebra on S.*
   Prove that if $\{A_n\}_{n=1}^\infty$ is a countably infinite family of sets in Σ, then $\cap_{n=1}^\infty A_n \in \Sigma$.

Hint

1. For part 1, use de Morgan's laws

$$\begin{aligned} S \setminus (A \cap B) &= (S \setminus A) \cup (S \setminus B) \\ S \setminus (A \cup B) &= (S \setminus A) \cap (S \setminus B) \end{aligned}$$

2. For part 2, use the infinite extension of de Morgan's laws:

$$S \setminus (\cap_{n=1}^\infty A_n) = \cup_{n=1}^\infty (S \setminus A_n); \quad S \setminus (\cup_{n=1}^\infty A_n) = \cap_{n=1}^\infty (S \setminus A_n)$$

Finite and Co-finite Sets in a Boolean Algebra

Definition

Given any infinite set S, say that the subset $T \subseteq S$ is co-finite just in case its complement $S \setminus T$ is finite.

Exercise

Let S be any infinite set, and let $\mathcal{F} := \{\{s\} \mid s \in S\}$ denote the family of all singleton subsets of X.

Show that the smallest Boolean algebra $\alpha(\mathcal{F})$ containing all sets in $\mathcal{F}$ consists of all subsets of S that are either finite or co-finite.

Hint *Show that the union of a finite set and a co-finite set is co-finite.*

Generating a Sigma-Algebra

Theorem

Let $\{\Sigma_i \mid i \in I\}$ be any indexed family of σ-algebras.

Then the intersection $\Sigma^\cap := \cap_{i \in I} \Sigma_i$ is also a σ-algebra.

Proof left as an exercise.

Let X be a space, and $\mathcal{F} \subset 2^X$ any family of subsets.

Since 2^X is obviously a σ-algebra,
there exists a non-empty set $\mathcal{S}(\mathcal{F})$ of σ-algebras that include $\mathcal{F}$.

Definition

Let $\sigma(\mathcal{F})$ denote the intersection $\cap\{\Sigma \mid \Sigma \in \mathcal{S}(\mathcal{F})\}$;
it is the smallest σ-algebra that includes $\mathcal{F}$,
otherwise known as the σ-algebra generated by $\mathcal{F}$.

Topological Spaces

Definition

Given a set X, a topology $\mathcal{T}$ on X
is a family of open subsets $U \subseteq X$ satisfying:

1. $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$;
2. if $U, V \in \mathcal{T}$, then $U \cap V \in \mathcal{T}$;
3. if $\{U_\alpha \mid \alpha \in A\}$ is any (possibly uncountable) collection of open sets $U_\alpha \in \mathcal{T}$, then the union $\cup_{\alpha \in A} U_\alpha \in \mathcal{T}$.

A topological space $(X, \mathcal{T})$ is any set X together with a topology $\mathcal{T}$ that consists of all the open subsets of X. $\square$

Parts 2 and 3 of the above definition of topology say that:

- finite intersections of open sets are open;
- arbitrary unions of open sets are open.

Closed Sets, Closures, Interiors, and Boundaries

Definition

Recall that, in a topological space $(X, \mathcal{T})$,
a set S is closed just in case its complement $X \setminus S$ is open. $\square$

Exercise

Prove that if $\{V_\alpha \mid \alpha \in A\}$ *is any (possibly uncountable) collection of closed sets* V_α *in the topological space* $(X, \mathcal{T})$, *then the intersection* $\cap_{\alpha \in A} V_\alpha$ *is closed.*

Definition

Let S be an arbitrary subset of the topological space $(X, \mathcal{T})$.

1. The closure $\text{cl}\, S$ of S
   is the intersection of all the closed sets that are supersets of S.
2. The interior $\text{int}\, S$ of S
   is the union of all the open sets that are subsets of S.
3. The boundary $\text{bd}\, S$ of S, also denoted by ∂S, is $\text{cl}\, S \setminus \text{int}\, S$,
   the complement of the interior in the closure. $\square$

The Metric Topology

Definition

Let (X, d) be any metric space.

The open ball of radius r centred at x is the set

$$B_r(x) := \{y \in X \mid d(x, y) < r\}$$

The metric topology $\mathcal{T}_d$ of (X, d) is the smallest topology that includes the entire family $\{B_r(x) \mid x \in X \text{ and } r > 0\}$ of all open balls in X.

Borel Sets and the Borel Sigma-Algebra

Definition

Let $(X, \mathcal{T})$ be any topological space.

Its Borel σ-algebra is defined as $\sigma(\mathcal{T})$
— i.e., the smallest σ-algebra containing every open set of X.

Each set $B \in \sigma(\mathcal{T})$ is then a Borel set.

Example

Suppose the topological space is a metric space (X, d)
with its metric topology $\mathcal{T}_d$.

Then the Borel σ-algebra is generated
by all the open balls $B_r(x) := \{x' \in X \mid d(x, x') < r\}$ in X.

For the case of the real line when $X = \mathbb{R}$,
its Borel σ-algebra is generated by all the open intervals of $\mathbb{R}$.

Indeed, it is even generated by the countable family
consisting of all the open intervals (q_1, q_2) where $q_1, q_2 \in \mathbb{Q}$.

More Borel Sets

Exercise

Show that every closed subset of a topological space $(X, \mathcal{T})$ is a Borel set.

Definition

A G_δ **set** in any topological space is the intersection of any countable collection of open sets.

Example

In $\mathbb{R}$, the infinite intersection $\bigcap_{n \in \mathbb{N}} (-\frac{1}{n}, \frac{1}{n})$ of open intervals is the G_δ set $\{0\}$, which is not open.

Exercise

Given any topological space $(X, \mathcal{T})$, show that:

1. *the complement of any G_δ subset is the union of a countable collection of closed sets;*
2. *any G_δ subset is a Borel set.*

Outline

Finitely Additive Set Functions

Let $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty + \infty\} = [-\infty, +\infty]$
denote the extended real line which, at each end,
has an endpoint added at infinity.

Let $\bar{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{+\infty\} = [0, +\infty]$ be the non-negative part of $\bar{\mathbb{R}}$.

Any family $\mathcal{F}$ of subsets $A \subseteq X$ is said to be pairwise disjoint
just in case $A \cap B = \emptyset$ whenever $A, B \in \mathcal{F}$ with $A \neq B$.

Definition

Let (X, Σ) be a measurable space.

A mapping $\mu : \Sigma \to \bar{\mathbb{R}}_+$ whose domain is a family of sets
is said to be a set function (but not a set-valued function).

The set function $\mu : \Sigma \to \bar{\mathbb{R}}_+$
is said to be additive (or finitely additive)
just in case, for any pair $\{A, B\}$ of disjoint sets in Σ,
one has $\mu(A \cup B) = \mu(A) + \mu(B)$. □

Implications of Finite Additivity

Lemma

If the set function $\mu : \Sigma \to \bar{\mathbb{R}}_+$ is finitely additive, then $\mu(\emptyset) = 0$.

Proof.

For any non-empty $A \in \Sigma$, the sets A and $\emptyset$ are disjoint.

Additivity implies that $\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)$, so $\mu(\emptyset) = 0$. $\square$

Exercise

For any finite collection $\{A_n\}_{n=1}^k$ of pairwise disjoint sets in Σ, prove by induction on k that finite additivity implies

$$\mu\left(\cup_{n=1}^k A_n\right) = \sum\nolimits_{n=1}^k \mu(A_n)$$

Disjoint Does Not Imply Pairwise Disjoint

Example

Suppose that $S = \{a, b, c\}$ where a, b, c are all different.

Consider the three different pair subsets

$$S_{-a} := S \setminus \{a\} = \{b, c\}$$
$$S_{-b} := S \setminus \{b\} = \{a, c\}$$
$$S_{-c} := S \setminus \{c\} = \{a, b\}$$

These three sets obviously satisfy $S_{-a} \cap S_{-b} \cap S_{-c} = \emptyset$, so are disjoint.

Yet $S_{-a} \cap S_{-b} = \{c\}$, $S_{-a} \cap S_{-c} = \{b\}$, and $S_{-b} \cap S_{-c} = \{a\}$ are all non-empty, so the three sets are not pairwise disjoint.

Additivity for Pairwise Disjoint, but not for Disjoint Sets

Exercise

Let S be any finite set, with power set 2^S.

Show that the only additive function μ on the measurable space $(S, 2^S)$ which satisfies $\mu(\{x\}) = 1$ for all $x \in S$ is the counting measure defined by $\mu(E) = \# E$ for all $E \subseteq S$.

Exercise

Following the previous example, let $S = \{a, b, c\}$ where a, b, c are all different, and let $S_{-x} := S \setminus \{x\}$ for each $x \in S$.

Following the previous exercise, let μ be the counting measure on $(S, 2^S)$.

Verify that, though the sets S_{-a}, S_{-b}, S_{-c} are disjoint, one has

$$\mu(S_{-a} \cup S_{-b} \cup S_{-c}) = \mu(S) = 3$$
$$\neq \mu(S_{-a}) + \mu(S_{-b}) + \mu(S_{-c}) = 3 \cdot 2 = 6$$

Measure as a Countably Additive Set Function

Definition

The set function $\mu : \Sigma \to \bar{\mathbb{R}}_+$ on a measurable space (X, Σ) is said to be σ-additive or countably additive just in case, for any countable collection $\{A_n\}_{n=1}^{\infty}$ of pairwise disjoint sets in Σ, one has

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

A measure on a measurable space (X, Σ) is a countably additive set function $\mu : \Sigma \to \bar{\mathbb{R}}_+$. □

Measure Space

Definition

A measure space is a triple (X, Σ, μ) where

1. Σ is a σ-algebra on X;
2. μ is a measure on the measurable space (X, Σ). $\square$

The Borel Real Line

Example

A prominent example of a measure space
is the Borel real line $(\mathbb{R}, \mathcal{B}, \ell)$ where:

1. $\mathcal{B}$ is the Borel σ-algebra
   generated by the open sets of the real line $\mathbb{R}$;
2. the measure $\ell(J)$ of any interval $J \subset \mathbb{R}$ is its length,
   defined whenever $(a, b) \in \mathbb{R}^2$ with $a \leq b$ by

$$\ell([a, b]) = \ell([a, b)) = \ell((a, b]) = \ell((a, b)) = b - a$$

3. ℓ is extended to all of $\mathcal{B}$ so as to satisfy countable additivity
   (it can be shown that this extension is unique).

Atoms and Non-Atomic Measure Spaces

Definition

An atom in a measure space (X, Σ, μ) is a set $A \in \Sigma$ such that $\mu(A) > 0$ and, for all $B \in \Sigma$ with $B \subset A$, one has $\mu(B) \in \{0, \mu(A)\}$.

Equivalently, there is no $\alpha \in (0, 1)$ and set $B \in \Sigma$ with $B \subset A$ such that $\mu(B) = \alpha\mu(A)$.

The measure space (X, Σ, μ) is non-atomic just in case no set $A \in \Sigma$ is an atom. $\square$

Exercise

Given any measure space (X, Σ, μ), prove that:

1. *if $x \in X$ satisfies $\mu(\{x\}) > 0$, then $\{x\}$ is an atom;*
2. *if (X, Σ, μ) is non-atomic and $S \in \Sigma$ is a countable set, then $\mu(S) = 0$.*

Prove too that the Borel real line is non-atomic as a measure space.

Probability Measure and Probability Space

Definition

Consider a measure space (X, Σ, μ).

The measure μ is a probability measure just in case $\mu(X) = 1$.

Then (X, Σ, μ) is a probability space. $\square$

Often one writes $(\Omega, \mathcal{F}, \mathbb{P})$ in this case, where:

1. Ω is the sample space;
2. $\mathcal{F}$ is the σ-algebra (or σ-field) of measurable events;
3. for each event $E \in \mathcal{F}$, the probability that E occurs is $\mathbb{P}(E)$.

Then, because $\mathbb{P}$ is a measure satisfying $\mathbb{P}(\Omega) = 1$,
one has $0 \leq \mathbb{P}(E) \leq 1$ for all $E \in \mathcal{F}$.

Probability as Normalized Measure

Definition

A measure space (X, Σ, μ) is:

1. finite just in case $\mu(X) < +\infty$;
2. σ-finite just in case there is a countable collection $\{S_n\}_{n \in \mathbb{N}}$ of measurable sets $S_n \in \Sigma$ with $\mu(S_n) < +\infty$ for all $n \in \mathbb{N}$ such that $X = \cup_{n \in \mathbb{N}} S_n$. □

Obviously any finite measure space (X, Σ, μ) can be given a normalized measure defined for all $E \in \Sigma$ by $\mathbb{P}(E) = \mu(E)/\mu(X)$.

This normalization makes $\mathbb{P}(X) = 1$, so $(X, \Sigma, \mathbb{P})$ is a probability space.

Exercise

Verify that the Borel real line $(\mathbb{R}, \mathcal{B}, \ell)$ *is not a finite measure space, but it is* σ*-finite.*

Lebesgue Measurable Subsets of the Real Line

Definition

In the Borel real line $(\mathbb{R}, \mathcal{B}, \ell)$ a subset $N \subset \mathbb{R}$, even if it is not a Borel set, is null just in case there exists a Borel subset $B \in \mathcal{B}$ with $\ell(B) = 0$ such that $N \subseteq B$. □

Let $\mathcal{N}$ denote the family of all null subsets of $\mathbb{R}$.

These null sets can be used to generate the Lebesgue σ-algebra of Lebesgue measurable sets, which is $\sigma(\mathcal{B} \cup \mathcal{N})$.

The symmetric difference of any two sets S and B is defined as the set

$$S \triangle B := (S \setminus B) \cup (B \setminus S) = (S \cup B) \setminus (S \cap B)$$

of elements s that belong to one of the two sets, but not to both.

One can show that $S \in \sigma(\mathcal{B} \cup \mathcal{N})$ if and only if there exists a Borel set $B \in \mathcal{B}$ such that $S \triangle B \in \mathcal{N}$ — i.e., S differs from a Borel set only by a null set.

The Lebesgue Real Line

There is a well-defined function $\lambda : \sigma(\mathcal{B} \cup \mathcal{N}) \to \bar{R}_+$
that satisfies $\lambda(S) := \ell(B)$ whenever $S \triangle B \in \mathcal{N}$.

Moreover, one can prove
that the function $S \mapsto \lambda(S)$ is countably additive.

This makes λ a measure, called the Lebesgue measure.

The associated measure space $(\mathbb{R}, \sigma(\mathcal{B} \cup \mathcal{N}), \lambda)$
is called the Lebesgue real line.

Because $\lambda(\mathbb{R}) = +\infty$, the Lebesgue real line cannot be normalized
to form a probability space.

Outline

Measurable Functions and Measurable Partitions

Definition

Let (X, Σ, μ) be a measure space,
and $(\mathbb{R}, \sigma(\mathcal{B} \cup \mathcal{N}), \lambda)$ the Lebesgue real line.

The function $X \ni x \mapsto f(x) \in \mathbb{R}$ is measurable
(with respect to the σ-algebras Σ on X and $\sigma(\mathcal{B} \cup \mathcal{N})$ on $\mathbb{R}$)
just in case the set $f^{-1}(B) = \{x \in X \mid f(x) \in B\}$ is Σ-measurable
for every Lebesgue measurable set $B \in \sigma(\mathcal{B} \cup \mathcal{N})$. □

Example

Let X and Y be topological spaces.

The function $X \ni x \mapsto f(x) \in Y$ is continuous
just in case the set $f^{-1}(B)$ is open in X whenever B is open in Y.

Then any continuous function $f : X \to Y$ is measurable
provided that X and Y are each given their Borel σ-algebra.

Step Functions

Recall that for any set $E \subseteq X$, the indicator function of E satisfies

$$X \ni x \mapsto 1_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Definition

A real-valued mapping $X \ni x \mapsto f(x) \in \mathbb{R}$ is a step function
just in case there is a collection $\{I_k\}_{k \in K_m}$
of m pairwise disjoint open intervals $I_k = (a_k, b_k) \subset \mathbb{R}$
together with a corresponding collection $\{c_k\}_{k \in K_m}$
of m constants $c_k \in \mathbb{R}$ such that $f(x) \equiv \sum_{k=1}^{m} c_k \, 1_{I_k}(x)$. □

Graphs of Step Functions

Exercise

Show that the graph in $\mathbb{R}^2$
of the non-trivial step function $f(x) \equiv \sum_{k=1}^{m} c_k \, 1_{I_k}(x)$
consists of:

1. *one finite collection* $\{I_k \times \{c_k\}\}_{k=1}^{m}$
   of m *finitely long horizontal line segments*
   on which y *belongs to the range* $\cup_{k=1}^{m}\{c_k\}$ *of* f*;*
2. *a complementary finite collection of line segments*
   along the horizontal axis $y = 0$*,*
   of which the two "at the ends" are infinitely long.

Integrating a Step Function

Definition

The integral of any step function

$$\mathbb{R} \ni x \mapsto f(x) = \sum_{k=1}^{m} c_k \, 1_{I_k}(x) \in \mathbb{R}$$

is defined as $\sum_{k=1}^{m} c_k \, \ell(I_k)$ where, for each $k \in \mathbb{N}_m$, the finite length of the interval I_k is $\ell(I_k)$.

Simple Functions

Definition

Given any measurable space (X, Σ),
the finite collection $\{E_k | k \in \mathbb{N}_m\}$
of m pairwise disjoint measurable sets $E_k \in \Sigma$
is a measurable partition of X just in case $\cup_{k=1}^m E_k = X$. □

Definition

A real-valued mapping $X \ni x \mapsto f(x) \in \mathbb{R}$ is a simple function
just in case there exist a measurable partition $\{E_k | k \in \mathbb{N}_m\}$ of X
together with a corresponding collection $(c_k)_{k=1}^m$
of m different real constants such that $f(x) \equiv \sum_{k=1}^m c_k 1_{E_k}(x)$. □

Note that the range $f(X) := \{y \in \mathbb{R} \mid \exists x \in X : y = f(x)\}$
of the simple function $f(x) = \sum_{k=1}^m c_k 1_{E_k}(x)$
is precisely the finite set $\{0\} \cup \{c_k | k \in \mathbb{N}_m\}$ of m real constants,
together with 0.

Step Functions Are Simple

Lemma

Any step function $\mathbb{R} \ni x \mapsto f(x) \equiv \sum_{k=1}^{m} c_k \, 1_{I_k}(x) \in \mathbb{R}$

where the sets $\{I_k\}_{k \in K_m}$ *are* m *pairwise disjoint intervals* $I_k \subset \mathbb{R}$

is identical to a simple function $\mathbb{R} \ni x \mapsto \tilde{f}(x) \equiv \sum_{k=1}^{m+1} \tilde{c}_k \, 1_{E_k}(x)$
where:

1. *for each* $k \in \mathbb{N}_m$ *one has* $\tilde{c}_k = c_k$ *and* $E_k = I_k$*;*
2. $E_{m+1} = \mathbb{R} \setminus \cup_{k \in K_m} I_k$ *and* $\tilde{c}_{m+1} = 0$*.*

Proof.

By obvious and routine checking of a few details. □

Let $\mathcal{F}_0$ denote the set of all real-valued step functions defined on $\mathbb{R}$.

Let $\mathcal{F}(X, \Sigma)$ denote the set of all real-valued simple functions defined on the measurable space (X, Σ).

It is easy to see that both $\mathcal{F}_0$ and $\mathcal{F}(X, \Sigma)$ are real vector spaces.

Integrable Simple Functions

We have seen how to integrate step functions defined on $\mathbb{R}$.

What about simple functions which are defined on a general measure space (X, Σ, μ)?

For as many functions $f : X \mapsto \mathbb{R}$ as possible, we want to define the integral $\int_X f(x) \, \mathrm{d}\,\mu = \int_X f(x) \, \mu(\mathrm{d}\,x)$.

Definition

The simple function $f(x) = \sum_{k=1}^m c_k \, 1_{E_k}(x)$ on (X, Σ, μ)

is μ-integrable just in case one has $\mu(E_k) < +\infty$ for all $k \in \mathbb{N}_m$.

In case $f(x) = \sum_{k=1}^m c_k \, 1_{E_k}(x)$ is μ-integrable,

we define $\int_X f(x) \, \mathrm{d}\,\mu := \sum_{k=1}^m c_k \, \mu(E_k)$, which does converge. $\square$

In particular, integrability requires that the support of f defined by $\operatorname{supp} f := \{x \in X \mid f(x) \neq 0\}$ satisfies $\mu(\operatorname{supp} f) < +\infty$.

The Heaviside and Dirichlet Functions

Example

The Heaviside step function $\mathbb{R} \ni x \mapsto H(x) \in \{0, 1\}$ is defined by $H(x) := 1_{[0,\infty)}(x)$.

It is not λ-integrable, where λ is the Lebesgue measure on $\mathbb{R}$.

Exercise

The Dirichlet simple function $\mathbb{R} \ni x \mapsto D(x) \in \{0, 1\}$ *is defined by* $D(x) := 1_{\mathbb{Q}}(x)$.

Explain why it is not a step function.

Measurable Functions

Definition

Given the measure space (X, Σ, μ),
the function $X \ni x \mapsto f(x) \in \mathbb{R}$ is measurable just in case
the inverse image $f^{-1}(B) := \{x \in X \mid f(x) \in B\}$
of each Borel set $B \subset \mathbb{R}$ satisfies $f^{-1}(B) \in \Sigma$. $\square$

Note that we have defined a simple function to be measurable.

Outline

Upper and Lower Bounds

In this subsection, we consider the case of a finite measure satisfying $\mu(X) < +\infty$.

In case $X \subseteq \mathbb{R}$ and μ is Lebesgue measure, this implies that X must be bounded — for example, $X = [a, b]$.

In case μ is a probability measure satisfying $\mu(X) = 1$, it is automatically a finite measure.

Upper and Lower Step Functions

Recall that $\mathcal{F}_0$ denotes the family
of step functions $\mathbb{R} \ni x \mapsto \sum_{k=1}^{m} c_k 1_{I_k}(x)$.

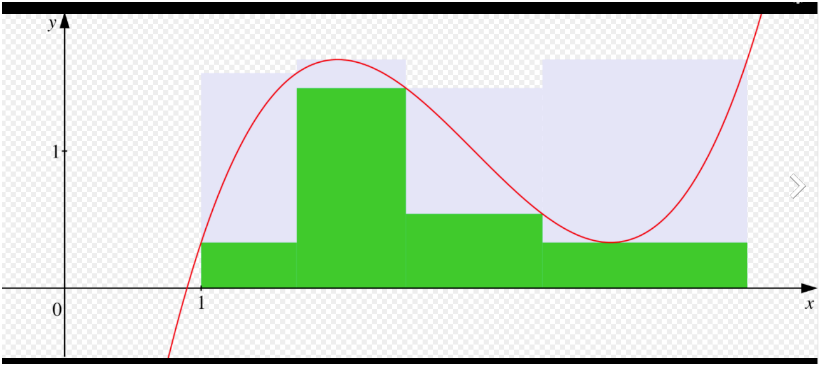
Definition

Given any function $\mathbb{R} \ni x \mapsto f(x) \in \mathbb{R}$, define the two sets

$$\begin{aligned} \mathcal{F}_0^{+}(f) &:= \{f^{+} \in \mathcal{F}_0 \mid \forall x \in X : f^{+}(x) \geq f(x)\} \\ \mathcal{F}_0^{-}(f) &:= \{f^{-} \in \mathcal{F}_0 \mid \forall x \in X : f^{-}(x) \leq f(x)\} \end{aligned}$$

of step functions whose graph lies respectively above or below
that of the function f.

Upper and Lower Step Functions Illustrated



When trying to find the integral of the red curve,
a lower approximation is the sum of the four green rectangles,
and an upper approximation adds the sum of the grey rectangles.

Source: `https://en.wikipedia.org/wiki/Darboux_integral`.
This also illustrates decreasing error as you add more steps.

Upper and Lower Integrals of Step Functions

The integral $\int_X f^+(x)\,\mu(\mathrm{d}\,x)$ of each step function $f^+ \in \mathcal{F}_0^+(f)$ is an over-estimate of the true integral $\int_X f(x)\,\mu(\mathrm{d}\,x)$ of f.

But the integral $\int_X f^-(x)\,\mu(\mathrm{d}\,x)$ of each step function $f^- \in \mathcal{F}_0^-(f)$ is an under-estimate of the true integral $\int_X f(x)\,\mu(\mathrm{d}\,x)$ of f.

Definition

The upper integral and lower integral of f are, respectively:

$$\begin{array}{rrcl} & I^+(f) & := & \inf_{f^+ \in \mathcal{F}_0^+(f)} \int_X f^+(x)\,\mu(\mathrm{d}\,x) \\ \text{and} & I^-(f) & := & \sup_{f^- \in \mathcal{F}_0^-(f)} \int_X f^-(x)\,\mu(\mathrm{d}\,x) \quad \square \end{array}$$

These are respectively the smallest possible over-estimate and greatest possible under-estimate of the integral.

Of course, in case f is itself a step function, one has $I^+(f) = I^-(f) = \int_X f(x)\,\mu(\mathrm{d}\,x)$.

The Darboux Integral

Definition

The function $\mathbb{R} \ni x \mapsto f(x) \in \mathbb{R}$ is Darboux integrable
just in case its upper and lower integrals $I^+(f)$ and $I^-(f)$
are both well defined and equal,
in which case its Darboux integral is the common value
of its upper and lower integrals. $\square$

Theorem

*The function $\mathbb{R} \ni x \mapsto f(x) \in \mathbb{R}$ is Darboux integrable
if and only if it is Riemann integrable,
in which case its Darboux and Riemann integrals are equal.*

Outline

Upper and Lower Simple Functions

Let (X, Σ, μ) be any measure space.

Let $f(x) = \sum_{k=1}^{m} c_k \, 1_{E_k}(x)$ be any simple function on (X, Σ, μ).

Recall that, by definition, that simple function f is μ-integrable just in case one has $\mu(E_k) < +\infty$ for all $k \in \mathbb{N}_m$.

Let $\mathcal{F}_S(X, \Sigma, \mu)$ denote the set of μ-integrable simple functions on the measure space (X, Σ, μ).

Given an arbitrary function $f : X \to \mathbb{R}$, define the two sets

$$\begin{aligned} \mathcal{F}^*(f; X, \Sigma, \mu) &:= \{f^* \in \mathcal{F}_S(X, \Sigma, \mu) \mid \forall x \in X : f^*(x) \geq f(x)\} \\ \mathcal{F}_*(f; X, \Sigma, \mu) &:= \{f_* \in \mathcal{F}_S(X, \Sigma, \mu) \mid \forall x \in X : f_*(x) \leq f(x)\} \end{aligned}$$

of μ-integrable simple functions
that are respectively upper or lower bounds for the function f.

Upper and Lower Bounds on an Integral

Given an arbitrary function $f : X \to \mathbb{R}$,
suppose there exists a "meaningful definition"
of the integral $J = \int_X f(x)\, \mu(\mathrm{d}\, x)$.

Then the well-defined integral $\int_X f^*(x)\, \mu(\mathrm{d}\, x)$
of each μ-integrable simple function $f^* \in \mathcal{F}^*(f; X, \Sigma, \mu)$,
should be an over-estimate of the true integral J of f.

Similarly, the integral $\int_X f_*(x)\, \mu(\mathrm{d}\, x)$
of each μ-integrable simple function $f_* \in \mathcal{F}_*(f; X, \Sigma, \mu)$,
is an under-estimate of the true integral J of f.

Upper and Lower Integrals

Inspired by the previous definition of the Darboux integral,
we define the upper integral and lower integral of f as, respectively

$$\begin{array}{rrcl} & I^*(f) & := & \inf_{f^* \in \mathcal{F}^*(f;X,\Sigma,\mu)} \int_X f^*(x)\,\mu(\mathrm{d}\,x) \\ \text{and} & I_*(f) & := & \sup_{f_* \in \mathcal{F}_*(f;X,\Sigma,\mu)} \int_X f_*(x)\,\mu(\mathrm{d}\,x) \end{array}$$

These are respectively the smallest possible over-estimate
and greatest possible under-estimate
of the integral $J = \int_X f(x)\,\mu(\mathrm{d}\,x)$.

Example

Of course, in case f is itself a μ-integrable simple function,
one has $I^*(f) = I_*(f) = \int_X f(x)\,\mu(\mathrm{d}\,x)$

Integrability and the Lebesgue Integral

Definition

Let $X \ni x \mapsto f(x) \in \mathbb{R}$ be defined on the measure space (X, Σ, μ).

1. The function f is integrably bounded just in case the mapping $X \ni x \mapsto |f(x)| \in \mathbb{R}_+$ is bounded above by a μ-integrable simple function.
2. The function f is Lebesgue integrable just in case its upper and lower integrals $I^*(f)$ and $I_*(f)$ are equal.
3. In case f is integrable, its Lebesgue integral $\int_X f(x)\,\mu(\mathrm{d}\,x)$ is defined as the common value of its upper integral $I^*(f)$ and its lower integral $I_*(f)$.

Main Theorem

Theorem

The function $X \ni x \mapsto f(x) \in \mathbb{R}$ on (X, Σ, μ) is Lebesgue integrable if and only if it is both measurable and integrably bounded.

Proof.

See, for example, the cited text by Royden. $\square$

Integration over an Interval or Other Measurable Set

Let $X \ni x \mapsto f(x) \in \mathbb{R}$ be a function
defined on the measure space (X, Σ, μ)
that is measurable and integrably bounded.

Let $E \in \Sigma$ be any measurable set, with indicator function $1_E(x)$.

Then the function $X \ni x \mapsto 1_E(x) f(x) \in \{f(x), 0\} \subset \mathbb{R}$
is also measurable and integrably bounded.

So we can define the integral of f over E by

$$\int_E f(x)\, \mu(\mathrm{d}\, x) := \int_X 1_E(x) f(x)\, \mu(\mathrm{d}\, x)$$

In case (X, Σ, μ) is the Lebesgue real line,
and E is the interval $[a, b]$,
one usually writes $\int_a^b f(x)\, \mathrm{d}\, x$ instead of $\int_{[a,b]} f(x)\, \mu(\mathrm{d}\, x)$.

Upper and Lower Bounds on an Integral

Exercise

Let $X \ni x \mapsto f(x) \in \mathbb{R}$ be a function defined on the measure space (X, Σ, μ) that is measurable and integrably bounded.

Let $E \in \Sigma$ be any measurable set, with indicator function $1_E(x)$.

Suppose that $a \leq f(x) \leq b$ for all $x \in E$.

1. *For any $f^* \in \mathcal{F}^*(f; X, \Sigma, \mu)$ and $f_* \in \mathcal{F}_*(f; X, \Sigma, \mu)$, show that for all $x \in E$ one has*

$$1_E(x) f^*(x) \geq 1_E(x) a \quad \text{and} \quad 1_E(x) f_*(x) \leq 1_E(x) b$$

2. *Show that $\mu(E) a \leq \int_E f(x)\, \mu(\mathrm{d}\, x) \leq \mu(E) b$.*

The Integral of a Nonnegative Function is a Measure

Exercise

Prove the following:

1. *If E and E' are subsets of X, then the indicator functions satisfy $1_{E \cup E'} = 1_E + 1_{E'}$ if and only if E and E' are disjoint.*
2. *If E and E' are disjoint measurable subsets of the measure space (X, Σ, μ), and $X \ni x \mapsto f(x) \in \mathbb{R}$ is integrable w.r.t. μ, then $\int_{E \cup E'} f(x)\mu(\mathrm{d}\,x) = \int_E f(x)\mu(\mathrm{d}\,x) + \int_{E'} f(x)\mu(\mathrm{d}\,x)$.*
3. *If $(E_n)_{n \in \mathbb{N}}$ is an infinite sequence of pairwise disjoint subsets of X, then:*
   - $1_{\cup_{n=1}^k E_n} = \sum_{n=1}^k 1_{E_n}$ *for each* $k \in \mathbb{N}$;
   - $1_{\cup_{n=1}^\infty E_n} = \sup_k 1_{\cup_{n=1}^k E_n} = \sup_k \sum_{n=1}^k 1_{E_n} = \sum_{n=1}^\infty 1_{E_n}$.
4. *If $X \ni x \mapsto f(x) \in \mathbb{R}_+$ is integrable w.r.t. μ, then $\Sigma \ni E \mapsto \int_E f(x)\mu(\mathrm{d}\,x) \in \mathbb{R}_+$ is a measure on (X, Σ).*

The Integral of a General Function is a Signed Measure

A general function $X \ni x \mapsto f(x) \in \mathbb{R}$
may have negative values at some points $x \in X$

Then the mapping $\Sigma \ni E \mapsto \int_E f(x)\mu(\mathrm{d}\,x) \in \mathbb{R}$
will generally have negative values for some measurable sets $E \in \Sigma$.

So $E \mapsto \int_E f(x)\mu(\mathrm{d}\,x)$ will generally not be a measure,
whose values must be nonnegative.

Instead, the mapping is a signed measure
on the measurable space (X, Σ).

That is, it is a σ-additive set function
whose values are allowed to be negative.

Outline

The Integral as a Function of Its Upper Limit

Let $I \ni x \mapsto f(x) \in \mathbb{R}$ be any continuous function defined on a finite interval $I = [A, B) \subset \mathbb{R}$.

Then the function $\mathbb{R} \ni x \mapsto 1_I(x) f(x)$ is measurable and also integrably bounded, because its range $f(I)$ is a finite interval in $\mathbb{R}$.

Given any fixed $a \in [A, B)$ and the Lebesgue measure λ on $\mathbb{R}$, we can define the integral function of the upper limit b by

$$[a, B) \ni b \mapsto J(b) := \int_a^b f(x) \lambda(\mathrm{d}\, x)$$

Theorem (Leibniz's Formula for the Lebesgue Integral)

At any point $b \in [A, B)$ where $[A, B) \ni x \mapsto f(x)$ is continuous, the integral function $J(b)$ is differentiable, with $J'(b) = f(b)$.

Proof of Leibniz's Formula

Proof.

For fixed $a, b \in [A, B)$ with $a < b$, for all small $h > 0$,
define $\phi_*(h)$ and $\phi^*(h)$ respectively
as the infimum and supremum of the set $\{f(x) \mid x \in (b, b+h)\}$.

These definitions imply that $h\phi_*(h) \leq \int_b^{b+h} f(x)\,\lambda(\mathrm{d}\,x) \leq h\phi^*(h)$.
But the Newton quotient of J at b is

$$q(h) = \frac{1}{h}[J(b+h) - J(b)] = \frac{1}{h}\int_b^{b+h} f(x)\,\lambda(\mathrm{d}\,x)$$

It follows that $\phi_*(h) \leq q(h) \leq \phi^*(h)$ for all small h.

Then continuity of f at b implies that, for all small $\epsilon > 0$,
there exists $\delta > 0$ such that $|x - b| < \delta$ implies $|f(x) - f(b)| < \epsilon$.

Hence $|h| < \delta$ implies that $f(b) - \epsilon < \phi_*(h) \leq \phi^*(h) < f(b) + \epsilon$.

This proves that as $h \to 0$,
so $\phi_*(h)$, $\phi^*(h)$ and therefore $q(h)$ all converge to $f(b)$. $\square$