Measure Theory - Overview

1 Starting definitions

Start with a set X. $\mathcal{A} \subseteq \mathcal{P}(X)$ is called an algebra if

- $\mathcal{A}$ is non-empty
- $X \in \mathcal{A}$
- $\mathcal{A}$ is closed and under complementation
- $\mathcal{A}$ is closed under finite unions and intersections

We obtain a σ-algebra if we also have closure under countable unions and intersections. We say a set $A \in \mathcal{A}$ is $\mathcal{A}$-measurable.

Note that intersecting σ-algebras obtains a new σ-algebra but taking unions does not necessarily work. A very important σ-algebra is the Borel σ-algebra,

$$\mathcal{B}(\mathbb{R}^d) := \sigma\left(\{\text{open sets in } \mathbb{R}^d\}\right)$$

Note that it can also be formed by all closed sets, closed half-rays or half-open intervals.

Given a σ-algebra $\mathcal{A}$ on a set X, a measure on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ such that

- $\mu(\emptyset) = 0$
- Given disjoint $A_1, A_2, \dots \in \mathcal{A}$,
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

This gives us a measure space $(X, \mathcal{A}, \mu)$.
We call this measure finite if $\mu(X) < \infty$ and σ-finite if we can write X as a union of finite measure sets.

Note measures are always increasingly monotonous and countably sub-additive. We also have the following very important property:

Proposition 1.1 (Continuity of measure)**.** *Given a measure space* $(X, \mathcal{A}, \mu)$*.*

- $A_1 \subseteq A_2 \subseteq \dots$ *in* $\mathcal{A}$ *then*
$$\mu\left(\bigcup_{i} A_i\right) = \lim_{i} \mu(A_i)$$

- $A_1 \supseteq A_2 \supseteq \dots$ *in* $\mathcal{A}$ *then*

$$\mu\left(\bigcap_k A_k\right) = \lim_k \mu(A_k)$$

A very important measure is the Lebesgue measure since it coincides with our natural intuition for the measure of subsets of $\mathbb{R}^d$. First we need to define an outer measure.

An outer measure is a function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ such that

- $\mu^*(\emptyset) = 0$
- $A \subseteq B \subseteq X \implies \mu^*(A) \leq \mu^*(B)$
- Given a countable collection of subsets $A_i \subseteq X$, we have countable sub-additivity

$$\mu^*\left(\bigcup_i A_i\right) \leq \sum_i \mu^*(A_u)$$

Notably we require monotonicity as an axiom and also require that the outer measure is defined on every subset. This is a weaker notion than a measure.
Given $A \subseteq \mathbb{R}$, $\mathcal{C}_A := \{\text{Collections } \{(a_i, b_i)\}_{i=1}^{\infty} \mid -\infty < a_i < b_i < \infty,\ \cup_{i=1}^{\infty}(a_i, b_i) \supseteq A\}$. This is the set of collections of finite open intervals which cover the set A. We can then define the Lebesgue outer measure on A to be

$$\lambda^*(A) := \inf\left\{\sum_i (b_i - a_i) \,\middle|\, \{(a_i, b_i\}_{i=1}^{\infty} \in \mathcal{C}_A\right\}$$

Proposition 1.2. *λ^* is an outer measure on $\mathbb{R}$ and $\lambda^*([a, b]) = b - a$, $\forall a, b \in \mathbb{R}$ such that $a \leq b$.*

Proof. The only difficult things to prove are countable sub-additivity and the desired value for intervals.

(i) Given $A_1, A_2, \dots \subseteq \mathbb{R}$ we may assume that $\lambda^*(A_i) < \infty$ for all i else countable sub-additivity holds trivially. Given $\epsilon > 0$ we can pick $\{(a_{i_n}, b_{i_n})\}_{i=1}^{\infty} \in \mathcal{C}_{A_i}$ such that

$$\sum_{n=1}^{\infty}(b_{i_n} - a_{i_n}) < \lambda^*(A_i) + \frac{\epsilon}{2^i}$$

We can union these countably many countable collections to get another countable collection covering $\cup_i A_i$.

$$\begin{aligned}
\lambda^*\left(\bigcup_i A_i\right) &\leq \sum_j (b_j - a_j) \\
&= \sum_i \left(\sum_n (b_{i_n} - a_{i_n})\right) \\
&\leq \sum_i \left(\lambda^*(A_i) + \frac{\epsilon}{2_i}\right) \\
&\leq \left(\sum_i \lambda^*(A_i)\right) + \epsilon
\end{aligned}$$

Taking $\epsilon \to 0$ yields the result.

(ii) Just think of a nice cover than does the job either exactly or to within ϵ, depending on your philosophy surrounding the set $\mathcal{A}$.

$\square$

By taking d-dimensional 'rectangular' intervals we can use the same procedure to define a Lebesgue measure on $\mathbb{R}^d$ which similarly assigns expected 'volumes' to these rectangles. The proof of this is somewhat more involved. Now for a weird definition.

> Given an outer measure μ^* on X, $B \subseteq X$ is μ^*-measurable if
>
> $$\forall A \in \mathcal{P}(X) \quad \mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^{\mathsf{c}})$$
>
> Intuitively, B is 'nice' if when we want to measure any other set we just measure the part inside and the part outside B and then add the measures together.

It's easy to show that any set with zero outer measure or who's complement has zero outer measure is outer measurable. Define

$$M_{\mu^*} := \{\mu^*\text{-measurable sets}\}$$

Theorem 1.3. *Given an outer measure μ^*, $M = M_{\mu^*}$ is a σ-algebra and μ^* yields a measure when restricted to M_{μ^*}.*

Proof. We certainly have $\sigma, X \in M$ and closure under complementation. First lets prove closure under finite union. Take $B_1, B_2 \in M$ and choose $A \subseteq X$ arbitrary.

$$\begin{aligned}
&\mu^*(A \cap (B_1 \cup B_2)) + \mu^*(A \cap (B_1 \cup B_2)^{\mathsf{c}}) && \\
&= \mu^*\left[(A \cap (B_1 \cup B_2)) \cap B_1\right] + \mu^*\left[(A \cap (B_1 \cup B_2)) \cap B_1^{\mathsf{c}}\right] && \textit{B_1 measurable} \\
&\quad + \mu^*\left[A \cap (B_1 \cap B_2)^{\mathsf{c}}\right] && \\
&= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^{\mathsf{c}} \cap B_2) + \mu * (A \cap B_1^{\mathsf{c}} \cap B_2^{\mathsf{c}}) && \textit{simplify sets} \\
&= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^{\mathsf{c}}) && \textit{B_2 measurable} \\
&= \mu^*(A) && \textit{B_1 measurable}
\end{aligned}$$

So M is certainly an algebra. To obtain countable unions we note the following can be proved by induction. Given $B_1, B_2, \dots \in M$ and any $A \subseteq X$.

$$\mu^*(A) = \sum_{i=1}^{n} \mu^*(A \cap B_i) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{n} B_i\right)^{\mathsf{c}}\right) \quad \forall n \in \mathbb{N}$$

Letting $n \to \infty$, by monotonicity of the outer measure on right term we get

$$\begin{aligned}
\mu^*(A) &\geq \underbrace{\sum_{i=1}^{\infty} \mu^*(A \cap B_i)}_{\text{converges since all terms +ve}} + \mu^*\left(A \cap \left(\textstyle\bigcup_{i=1}^{\infty} B_i\right)^{\mathsf{c}}\right) && \textit{sub-additivity} \\
&\geq \mu^*\left(A \cap \left(\textstyle\bigcup_{i=1}^{\infty} B_i\right)\right) + \mu^*\left(A \cap \left(\textstyle\bigcup_{i=1}^{\infty} B_i\right)^{\mathsf{c}}\right)
\end{aligned}$$

and hence $\cup_i B_i \in M$ because the other inequality is an axiomatic assumption. For arbitrary sets we can just take appropriate complementation to express their union as a union of pairwise disjoint sets.

It remains to show that we get a measure. Again the only thing to really show is the remaining inequality to get countable additivity. Given disjoint $B_1, B_2, \dots$ in M just take $A = \cup_i B_i$ in the above inequality to get

$$\mu^* \left(\bigcup_{i=1}^{\infty} B_i \right) \geq \sum_{i=1}^{\infty} \mu^*(B_i) + \mu^*(\emptyset)$$

$\square$

1.1 Working with Lebesgue Measures

We are now able to form a measure from the Lebesgue outer measure.

The Lebesgue measurable sets are exactly the λ^*-measurable sets. The resulting σ-algebra is denoted $\mathcal{L}^d$. Restricting λ^* to $\mathcal{L}^d$ yields the Lebesgue measure λ_d.

Proposition 1.4.

$$\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{L}(\mathbb{R}^d)$$

Proof. Take $b \in \mathbb{R}$, we will show that $(-\infty, b] \in \mathcal{L}$ so that we can take σ on either side to obtain the result. Pick any $A \subseteq \mathbb{R}$ such that $\lambda^*(A) < \infty$ and take arbitrary $\epsilon > 0$.
Choose $\{(a_i, b_i)\} \in \mathcal{C}_A$ such that $\sum_i (b_i - a_i) < \lambda^*(A) + \epsilon$. Notice that $(a_i, b_i) \cap B$ and $(a_i, b_i) \cap B^{\mathsf{c}}$ are disjoint intervals whose lengths sum to $b_i - a_i$.

$$\begin{aligned}
&\lambda^*(A \cap B) + \lambda^*(A \cap B^{\mathsf{c}}) && \textit{monotoncity} \\
\leq\ &\lambda^* \left([\cup_i (a_i, b_i)] \cap B \right) + \lambda^* \left([\cup_i (a_i, b_i)] \cap B^{\mathsf{c}} \right) && \textit{countable sub-additivity} \\
\leq\ &\textstyle\sum_i \lambda^*((a_i, b_i) \cap B) + \sum_i \lambda^*((a_i, b_i) \cap B^{\mathsf{c}}) && \textit{rearrange +ve terms} \\
\leq\ &\textstyle\sum_i \left[\text{length}\left((a_i, b_i) \cap B\right) + \text{length}\left((a_i, b_i) \cap B^{\mathsf{c}}\right) \right] \\
=\ &\textstyle\sum_i (b_i - a_i) < \lambda^*(A) + \epsilon
\end{aligned}$$

Now taking $\epsilon \to 0$ we obtain the troublesome inequality. $\square$

It is often useful to be able to approximate the Lebesgue measure from above and from below.

Proposition 1.5 (Regularity of Measure)**.** *Let $A \in \mathcal{L}(\mathbb{R}^d)$ then*

(a) $\lambda(A) = \inf \{\lambda(U) \mid U \textit{ open}, U \supseteq A\}$

(b) $\lambda(A) = \sup \{\lambda(K) \mid K \textit{ compact}, K \subseteq A\}$

(c) $\textit{RHS}(a) = \textit{RHS}(b) \implies A \in \mathcal{L}(\mathbb{R}^d)$

Proof. Every measure is monotonous so we only have one inequality to prove in each case.

(a) Assume $\lambda(A) < \infty$ else we are already done. Given any $\epsilon > 0$, pick $\{R_i\} \in \mathcal{C}_A$ such that $\sum_i \text{vol}(R_i) \leq \lambda^*(A) + \epsilon < \lambda(A) + \epsilon$. Define $U := \cup_i R_i$ which is then an open set such that $A \subseteq U$. Now we have

$$\lambda(U) \leq \sum_i \lambda(R_i) = \sum_i \lambda^*(R_i) = \sum_i \text{vol}(R_i) \leq \lambda(A) + \epsilon$$

Taking $\epsilon \to 0$ yields the result

(b) Again take $\epsilon > 0$ arbitrarily, we split into cases.

Case 1: A is a bounded set.

Take $C \supseteq A$ which is compact. Now by (a) there is U open with $U \supseteq C \setminus A$ such that

$$\lambda(U) \leq \lambda(C \setminus A) + \epsilon$$

Now define $K := C \setminus U$. Then C is closed and U is open so K is closed and K lives within C so is bounded. Hence K is bounded. Also note $K \subseteq A$.

$$\begin{aligned} \lambda(C) &\leq \lambda(K) + \lambda(U) \\ &\leq \lambda(K) + \lambda(C \setminus A) + \epsilon \end{aligned}$$

Hence

$$\lambda(K) \geq \lambda(C) - \lambda(C \setminus A) - \epsilon = \lambda(A) - \epsilon$$

Taking $\epsilon \to 0$ yields $\sup \geq \lambda(A)$.

Case 2: A is an unbounded set.

The issue this time is we can't really choose that C. This time define $A_i := A \cap [-i, i]^d$ and set $A := \cup_i A_i$ to see

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$$

Continuity of measure tells us that $\lim_{n\to\infty} \lambda(A_i) = \lambda(A)$. So given $\epsilon > 0$ there is $n \in \mathbb{N}$ such that $\lambda(A_n) \geq \lambda(A) - \frac{\epsilon}{2}$. Case 1 tells us that there is compact K such that $K \subseteq A_n \subseteq A$ with the property that

$$\lambda(K) \geq \lambda(A_n) - \frac{\epsilon}{2} \geq \lambda(A) - \epsilon$$

Taking $\epsilon \to 0$ yields our result.

$\square$

One nice property of the Lebesgue measure is translation invariance.

Proposition 1.6 (Translation invariance). *Fix $x \in \mathbb{R}^d$ then*

(a) $\forall A \in \mathcal{P}(\mathbb{R}^d) \quad \lambda^*(A) = \lambda^*(A + x)$

(b) $A \in \mathcal{L} \implies A + x \in \mathcal{L}, \quad \lambda(A + x) = \lambda(A)$

Proof. (a) Given a covering collection $\{(a_i, b_i)\} \in \mathcal{C}_A$ we can just translate these intervals and the volume is preserved.

(b) We first show that $A + x$ is λ^*-measurable.

$$\begin{aligned} &\lambda^*(B \cap (A + x)) + \lambda^*(B \cap (A + x)^\mathsf{c}) && \textit{using (a)} \\ &= \lambda^*((B - x) \cap A) + \lambda^*((B - x) \cap A^\mathsf{c}) && A \in \mathcal{L} \\ &= \lambda^*(B - x) && \textit{using (a)} \\ &= \lambda^*(B) \end{aligned}$$

and hence $A + x \in \mathcal{L}$. Also $\lambda(A + x) = \lambda^*(A + x) = \lambda^*(A) = \lambda(A)$.

$\square$

The question arises whether there exists a set which cannot be measured by the omnipotent Lebesgue. This depends on your view of the Axiom of Choice.

Theorem 1.7 (Vitali Set). *Assuming the axiom of choice, $\exists E \subseteq (0, 1)$ such that $E \notin \mathcal{L}(\mathbb{R})$.*

Proof. Can't be bothered to do this atm. $\square$

2 Extended Real Line

We can extend the real line by adding two points

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$$

and abiding by the following conventions

- $+\infty + x = +\infty \qquad \forall x \in (-\infty, +\infty]$
- $-\infty + x = -\infty \qquad \forall x \in [-\infty, +\infty)$
- $$x \cdot (+\infty) = (+\infty) \cdot x = \begin{cases} -\infty & x \in [-\infty, 0) \\ 0 & x = 0 \\ +\infty & x \in (0, +\infty] \end{cases}$$
- $$x \cdot (-\infty) = (-\infty) \cdot x = \begin{cases} +\infty & x \in [-\infty, 0) \\ 0 & x = 0 \\ -\infty & x \in (0, +\infty] \end{cases}$$

We need a topology for this by giving a base of open sets

$$\{(a, b) \mid a, b \in \mathbb{R}\} \cup \{[-\infty, a) \mid a \in \mathbb{R}\} \cup \{(a, \infty] \mid a \in \mathbb{R}\}$$

Then a set is closed if and only if all sequences contain their limits (including limits at infinity). Under this topology $\overline{\mathbb{R}}$ is compact.

3 Measurable Functions

Before we can define measurable functions we need to note a few equivalences.

Proposition 3.1. *Given a measurable space $(X, \mathcal{A})$. If $Y = \mathbb{R}$ or $\overline{\mathbb{R}}$, $A \in \mathcal{A}$ and $f : A \to Y$. The following are equivalent:*

(a) $\forall t \in \mathbb{R} \quad f^{-1}([-\infty, t]) \in \mathcal{A}$

(b) $\forall t \in \mathbb{R} \quad f^{-1}((t, +\infty]) \in \mathcal{A}$

(c) $\forall t \in \mathbb{R} \quad f^{-1}([-\infty, t)) \in \mathcal{A}$

(d) $\forall t \in \mathbb{R} \quad f^{-1}([t, +\infty]) \in \mathcal{A}$

(e) $\forall$ *open* $U \subseteq Y \quad f^{-1}(U) \in \mathcal{A}$

(f) $\forall$ *closed* $B \subseteq Y \quad f^{-1}(B) \in \mathcal{A}$

(g) $\forall B \in \mathcal{B}(Y) \quad f^{-1}(B) \in \mathcal{A}$

Proof. There's an awful lot to prove here. $\square$

A function $f : A \to \overline{\mathbb{R}}$ or $\mathbb{R}$ is $\mathcal{A}$-measurable if $\forall t \in \mathbb{R}$,

$$\{f < t\} := \{x \in A \mid f(x) < t\} \in \mathcal{A}$$

A function f is simple if $f(A)$ is finite.

One consequence is that all Borel-measurable functions are Lebesgue-measurable since $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{L}(\mathbb{R}^d)$.

Proposition 3.2. *If $f, g : A \to \overline{\mathbb{R}}$ are measurable then $\{f < g\}$, $\{f \leq g\}$ and $\{f = g\}$ are in $\mathcal{A}$.*

Proof. Notice that we only really need to show the first is in $\mathcal{A}$. We express this as a countable combination of measurable sets:

$$B = \bigcup_{r \in \mathbb{Q}} \{f < r \text{ and } g > r\} = \bigcup_{r \in \mathbb{Q}} (\{f < r\} \cap \{g > r\})$$

□

We can define maximum and minimum functions which by this last proposition are measurable functions themselves.

$$(f \vee g)(x) = \max\{f(x), g(x)\}$$
$$(f \wedge g)(x) = \min\{f(x), g(x)\}$$

Pointwise sup, inf, lim sup, lim inf and lim of sequences of measurable functions also define measurable functions.

Note: Given $f_n : A \to \overline{\mathbb{R}}$ measurable we can show that $B := \{x \in A \mid lim_{n\to\infty} f_n(x) \text{ exists }\}$ is measurable and is the domain we use to define the pointwise limit function $\lim_n f_n$.
Also, given functions

$$(X, \mathcal{A}) \xrightarrow[\text{measurable}]{f} (\mathbb{R}, \mathcal{B}) \xrightarrow[\text{Borel measurable}]{g} (\mathbb{R}, \mathcal{B})$$

their composition $f \circ g$ is also measurable.

We can also see that the set of measurable functions forms a vector space under appropriate pointwise operations. We can see that f^2 is measurable because

$$\{f^2 < t\} = \left\{f < \sqrt{t}\right\} \cap \left\{f > -\sqrt{t}\right\}$$

Define the following two very important functions:

$$f^+ := f \vee 0$$
$$f^- := -(f \wedge 0)$$

We will come to use the following technical proposition very often:

Proposition 3.3. *Given $f : A \to [0, +\infty]$ measurable, there exist measurable simple functions $f_n : [0, +\infty)$ such that $f_1 \leq f_2 \leq f_3 \leq \ldots$ and $f = \lim_n f_n$.*

Proof. Given $n \in \mathbb{N}$, for every $k \in (1, \dots, n \cdot 2^n)$ define the set

$$A_{n,k} := \left\{ \frac{k-1}{2^n} \leq f < \frac{k}{2^n} \right\} \in \mathcal{A}$$

Then define

$$f_n(x) := \begin{cases} \frac{k-1}{2^n} & \text{if } \exists k \in \{1, \dots, n \cdot 2^n\} \text{ such that } x \in A_{k,n} \\ n & \text{otherwise} \end{cases}$$

Where f has a finite value, the maximum error is $\frac{1}{2^n} \to 0$ as $n \to \infty$. Where f has infinite value $f_n(x) = n \to \infty$ as $n \to \infty$. Certainly $f_1 \leq f_2 \leq f_3 \leq \dots$ □

By applying this proposition to f^+ and f^- separately and combining the results we can see that any measurable f is the limit of measurable simple functions.

Note: It is possible to construct a set this is Lebesgue measurable but not Borel measurable. Its rather long winded but worth a read.

4 Integration

The aim of this section is to define the integral on a measure space $(X, \mathcal{A}, \mu)$. We define this function iteratively on an increasingly large subset of functions.

4.1 Simple Functions

Define

$$S_+ := \{f : X \to [0, +\infty \mid f \text{ simple and } \mathcal{A}\text{-measurable}\}$$

So given $f \in S_+$ we can write $f = \sum_i a_i \chi_{A_i}$ for some $a_i \in [0, +\infty)$ and $A_1, \dots, A_m$ disjoint and measurable. The a_i are not distinct and so this is not a unique presentation.

We can now define the integral to be

$$\int f \, d\mu := \sum_{i=1}^{m} a_i \, \mu(A_i) = \sum_{a \in f(X)} a \, \mu(f^{-1}(a))$$

It can be shown with some ease that this is a linear, increasing function. We also get the desirable property that we can swap limit and integral in certain circumstances.

Proposition 4.1. *Let f and $f_1 \leq f_2 \leq f_3 \leq \dots$ in S_+ with $f = \lim_n f_n \in S_+$, then*

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

Proof. By monotonicity we certainly have

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

For the opposite inequality, write $f = \sum_i a_i \, \chi_{A_i}$. Take some arbitrary $\epsilon > 0$. Define the following sets

$$A_{n,i} := \{x \in A_i \mid f_n(x) \geq (1 - \epsilon) a_i\} \in \mathcal{A}$$

and notice these are nested sets satisfy

$$A_{1,i} \subseteq A_{2,i} \subseteq A_{3,i} \subseteq \dots \quad \text{such that} \quad \cup_n A_{n,i} = A_i$$

Define $g_n := \sum_{i=1}^k (1-\epsilon) a_i \chi_{A_{n,i}} \leq f_n$ which also satisfies $g_1 \leq g_2 \leq g_3 \leq \dots$

$$\begin{aligned} lim_n \int f_n \, d\mu &\geq \lim_n \int g_n \, d\mu \\ &= \textstyle\sum_{i=1}^k (1-\epsilon) a_i \mu(A_{n,i}) \\ &= (1-\epsilon) \textstyle\sum_{i=1}^k a_i \lim_n \mu(A_{n,i}) \\ &= (1-\epsilon) \textstyle\sum_{i=1}^k a_i \mu(A_i) && \text{measure cty} \\ &= (1-\epsilon) \int f \, d\mu \end{aligned}$$

Taking $\epsilon \to 0$ yields the remaining inequality. □

4.2 Non-negative measurable functions

Define

$$\overline{S_+} := \{\text{measurable } f : X \to [0, +\infty]\}$$

Given $f \in \overline{S_+}$ we can define the integral by

$$\int f \, d\mu := \sup \left\{ \int g \, d\mu \,\middle|\, g \in S_+, \; g \leq f \right\}$$

Note that this is certainly consistent with our original definition for S_+

Proposition 4.2. *Given $f_1 \leq f_2 \leq \dots$ in S_+, and $d := \lim_n f_n$ then $f \in \overline{S_+}$. Moreover, $\int f \, d\mu = \lim_n \int f_n \, d\mu$.*

Proof. We have already seen that $f \in \overline{S_+}$ because it is the limit of a sequence of measurable functions. By our new definition of the integral we have

$$\int f_1 \, d\mu \leq \int f_2 \, d\mu \leq \dots \leq \int f \, d\mu$$

and hence certainly $\lim_n \int f_n \, d\mu \leq \int f \, d\mu$. So if the limit is an upper bound, it is certainly the least such upper bound.
So for the converse inequality it suffices to show that given $g \in S_+$ such that $g \leq f$ we have $\int g \, d\mu \leq lim_n \int f_n \, d\mu$. Well consider

$$g \wedge f_1 \leq g \wedge f_2 \leq \dots \in S_+$$

We have that $f_n \to f \geq g$ and hence $\lim_{n\to\infty}(g \wedge f_n) = g$. So the previous proposition tells use that

$$\int g \, d\mu = \lim_{n\to\infty} \int (g \wedge f_n) \, d\mu \leq \lim_{n\to\infty} \int f_n \, d\mu$$

Again we can show that this new integral is still a linear, increasing operator on $\overline{S_+}$. □

4.3 Arbitrary Measurable Functions

Finally given any $f : X \to \overline{\mathbb{R}}$ define the integral to be

$$\int f \, d\mu := \begin{cases} \text{UNDEFINED} & \text{if } \int f^+ \, d\mu = \int f^- \, d\mu = +\infty \\ \int f^+ \, d\mu - \int f^- \, d\mu & \text{otherwise} \end{cases}$$

f is called μ-integrable if $\int f^+ \, d\mu < +\infty$ and $\int f^- \, d\mu \;\; +\infty$.

In the case $f \in \overline{S_+}$, then $f^- = 0$ and hence the definitions coincide.

4.4 Playing with the Integral

One property we will often use to estimate integrals.

Proposition 4.3. *Let $f : X \to \overline{\mathbb{R}}$ be measurable then*

$$f \text{ integrable} \quad \iff \quad |f| \text{ integrable}$$

Moreover, $|\int f \, d\mu| \leq \int |f| \, d\mu$.

We say that a measure space $(X, \mathcal{A}, \mu)$is complete if

$$\forall A \in \mathcal{A} \text{ such that } \mu(A) = 0 \quad \forall B \subseteq A \quad B \in \mathcal{A}$$

i.e. every subset of a 0-measure set is measurable.
The completion of $(X, \mathcal{A}\mu)$ is $(X, \mathcal{A}_\mu, \overline{\mu})$ where

$$\begin{aligned} \mathcal{A}_\mu &:= \{A \subseteq X \mid \exists E, F \in \mathcal{A} \text{ such that } E \subseteq A \subseteq F, \; \mu(F \setminus E) = 0\} \supseteq \mathcal{A} \\ \overline{\mu}(A) &:= \mu(F) = \mu(E) \end{aligned}$$

The proof that the completion of a measure space is in fact a complete measure space is omitted and non-examinable.

A property $P : X \to \{\text{true}, \text{false}\}$ holds almost everywhere if

$$\exists N \in \mathcal{A} \text{ such that } \mu(N) = 0, \; N \supseteq P^{-1}(\text{false})$$

Proposition 4.4. *Suppose $(X, \mathcal{A}, \mu)$is complete and $f, g : X \to \overline{\mathbb{R}}$ such that $f(x) = g(x)$ for almost every x. Then f is measurable $\iff$ g is measurable.*

Proof. Suppose that f is measurable and $\exists N \in \mathcal{A}$ such that $\mu(N) = 0$ and $\{f \neq g\} \subseteq N$.

$$\{g \leq t\} = (\{f \leq t\} \cap N^{\mathsf{c}}) \cup (\{g \leq t\} \cap N)$$

Note $\{f \leq t\} \in \mathcal{A}$ since f is measurable and certainly $N^{\mathsf{c}} \in \mathcal{A}$. The second set is a subset of N and N has 0 measure and hence the second set is measurable by completeness. So $\{g \leq t\} \in \mathcal{A}$ and so g is measurable. ☐

Proposition 4.5. *Suppose $f, g : X \to \overline{\mathbb{R}}$ are measurable such that $f = g$ almost everywhere. If f is integrable then g is integrable. Moreover $\int f \, d\mu = \int g \, d\mu$.*

Proof. Pick $N \in \mathcal{A}$ with $\mu(N) = 0$ such that $\{f \neq g\} \subseteq N$. Define

$$h(x) := \begin{cases} +\infty & x \in N \\ 0 & x \notin N \end{cases}$$

Consider the following sequence of simple measurable, non-negative functions.

$$\chi_N \leq 2\chi_N \leq 3\chi_N \leq \cdots \leq \lim_n (n\chi_N) = h$$

Hence

$$\int h \, d\mu = \lim_{n\to\infty} \int n\chi_N \, d\mu = \lim_{n\to\infty} n\mu(N) = \lim_{n\to\infty} 0 = 0$$

Certainly $g^+ \leq f^+ + h$ and hence $\int g^+ \, d\mu \leq \int f^+ \, d\mu + \int h \, d\mu \leq \int f^+ \, d\mu < +\infty$. Similarly we can show that $\int g^- \, d\mu \leq \int f^- \, d\mu < +\infty$ and so g is integrable. We can repeat this whole proof in the opposite direction to get the opposite inequalities and hence $\int f \, d\mu = \int g \, d\mu$. $\square$

4.5 Application to Probability Theory

Suppose we have a random variable Y. We need a measure space with the following structure.

- $X = \{\text{elementary outcomes}\}$
- $\mathcal{A} = \{\text{events}\}$
- $\mu(A) = \mathbb{P}(A)$
- $\mu(X) = 1$ so that this is a probability space.

Then $Y : X \to \overline{\mathbb{R}}$ is a measurable function. We define the expectation of Y to be

$$\mathbb{E}(Y) := \int Y \, d\mu$$

Proposition 4.6 (Markov's Inequality)**.** *Given $f : X \to [0, +\infty]$ measurable and $t \in (0, +\infty)$. Let $A := \{f \geq t\}$. Then*

$$\mu(A) \leq \frac{1}{t} \int_A f \, d\mu \leq \frac{1}{t} \int f \, d\mu$$

Proof.

$$t\chi_A \leq f\chi_A \leq f \underbrace{\Longrightarrow}_{\text{integrate}} t\mu(A) \leq \int_A f \, d\mu \leq \int f \, d\mu$$

$\square$

Phrasing this in terms of random variables we see that given a random variable $Y \geq 0$ then

$$\mathbb{P}(Y \geq t) \leq \frac{1}{t}\mathbb{E}(Y) \qquad \forall t \in (0, +\infty)$$

Corollary 4.7. *Suppose $f : X \to \overline{\mathbb{R}}$ is a measurable function. Then*

$$\int |f| \, d\mu = 0 \implies f = 0 \text{ almost everywhere}$$

Proof. Given any $n \in \mathbb{N}$

$$\mu\left\{|f| \geq \frac{1}{n}\right\} \leq n \int |f| \, d\mu = 0$$

Now $\{f \neq 0\} = \cup_{n \in \mathbb{N}} \left\{|f| \geq \frac{1}{n}\right\}$ and $\mu\left(\cup_{n \in \mathbb{N}} \left\{|f| \geq \frac{1}{n}\right\}\right) = 0$. $\square$

Corollary 4.8.

$$f : X \to \overline{\mathbb{R}} \quad \textit{integrable} \quad \Longrightarrow \quad |f| < +\infty \quad \textit{a.e.}$$

Proof. The proof is very similar to the previous corollary. $\square$

The following space will be of vital importance

$$\mathcal{L}^1(X, \mathcal{A}, \mu, \mathbb{R}) := \{f : X \to \mathbb{R} \mid \text{integrable}\}$$

We will often just refer to this as $\mathcal{L}^1$.

Corollary 4.9. *Let $f : X \to \overline{\mathbb{R}}$ be a measurable function. Then*

$$f \quad \textit{integrable} \iff \exists g \in \mathcal{L}^1 \quad \textit{s.t.} \quad g = f \textit{ a.e.}$$

Proof. Just set g to be the same as f except on a set of 0-measure where f is ∞ where we define g to be 0. $\square$

4.6 Limit Theorems

Theorem 4.10 (Monotone Convergence Theorem)**.** *Let f and $f_1, f_2, \ldots$ be measurable functions $X \to [0, +\infty]$ such that for almost every x*

$$f_1(x) \leq f_2(x) \leq \ldots \quad \textit{and} \quad f(x) = \lim_{n \to \infty} f_n(x)$$

then $\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu$.