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* $\mathcal{A}$ *and* $\mu(A_n) < +\infty$ *for some* n*, then*

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

Note: When taking the limit of an infinite intersection of nested sets, we only get continuity if our measure is finite or one of the nested sets has finite measure.

Example: Take μ to be the counting measure on $\mathbb{N}$ then consider the nested sets

$$A_k := \{k, k+1, k+2, \dots\}$$

then $\cap_k A_k = \emptyset$ but $\lim_{k\to\infty} \mu(A_k) = +\infty$.

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 $\cup_i A_i$.

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

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 $\mathcal{A}$.

□

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.

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

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

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

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

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

□

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. □

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

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

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

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.

□

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.

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

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

□

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. We start be defining an equivalence relation $x \sim y$ if $x - y \in \mathbb{Q}$. This gives us equivalence classes $\mathbb{Q} + x$ for $x \in \mathbb{R}$. By the axiom of choice, we can choose $E \subseteq (0,1)$ such that we pick exactly one element from each class.
Label the rationals in $(-1,1)$ by $r_1, r_2, \ldots$ and define $E_n := E + r_n$ for each rational. Then the E_n are pairwise disjoint as follows. Suppose $E_n \ni e + r_n = e' + r_{n'} \in E_{n'}$ for $n \neq n'$. Then we have two elements of the Vitali set such that $e - e' = r_n - r_{n'} \in \mathbb{Q}$. These two are therefore in the same equivalence class $\mathbb{Q} + x$ so we must have $e = e'$ and hence $n = n'$.
Note also that every real number is a rational distance away from a unique member of the Vitali set and hence

$$\bigcup_{r \in \mathbb{Q}} (E + r) = \mathbb{R} \quad \text{and} \quad (0,1) \subseteq \bigcup_n E_n \quad \text{and} \quad \bigcup_n E_n \subseteq (-1,2)$$

Now we wish to show that E is not Lebesgue measurable. Suppose for contradiction that $E \in \mathcal{L}$ then by translation invariance so too is E_n for all n. Then

$$3 = \lambda((-1,2)) \geq \lambda\left(\cup_n E_n\right) = \sum_n \lambda(E_n) = \sum_n \lambda(E)$$

To avoid the right hand side shooting off to infinity we must have $\lambda(E) = 0$. But then

$$\lambda(0,1) = 1 \leq \sum_n \lambda(E_n) = \sum_n \lambda(E) = 0$$

This is a clear contradiction. □

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 $f_1 \leq f_2 \leq f_3 \leq ...$ and $f = \lim_n f_n$.*

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

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

Then define

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

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 $f_1 \leq f_2 \leq f_3 \leq \ldots$ □

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 Limits of measurable functions

Should show that the liminf and what not of measurable functions are measurable and the sets where they are defined.

4.1 Some Generalisations

Given spaces $(X, \mathcal{A})$ and $(Y, \mathcal{C})$ we can say that $f : X \to Y$ is $(A, \mathcal{C})$-measurable if

$$\forall C \in \mathcal{C} \quad f^{-1}(C) \in \mathcal{A}$$

We can see very clearly that composition of measurable functions yields another measurable function. Checking something is measurable can be quite challenging because we have a lot of sets to check. The following allows us to check a basis of sets rather than there σ-algebra.

Proposition 4.1. *Suppose $\mathcal{C} = \sigma(C_0)$ for some $C_0 \subseteq \mathcal{P}(Y)$ then*

$$f \textit{is measurable} \quad \iff \quad \forall C \in C_0 \; f^{-1}(C) \in \mathcal{A}$$

5 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.

5.1 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 5.1. *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 $g_1 \leq g_2 \leq g_3 \leq \dots$

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

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

5.2 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 5.2. *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

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

$\square$

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

5.3 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.

Note: It might be worth going over the proof that the integral is linear in the arbitrary case.

5.4 Playing with the Integral

One property we will often use to estimate integrals.

Proposition 5.3. *Let $f : X \to \overline{\mathbb{R}}$ be measurable then*

$$f \textit{ integrable} \iff |f| \textit{ integrable}$$

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 5.4. *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. $\square$

Proposition 5.5. *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 $\int f \, d\mu = \int g \, d\mu$. $\square$

5.5 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 5.6 (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.

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

$\square$

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

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

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

Corollary 5.8.

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

Proof. The proof is very similar to the previous corollary. ☐

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 5.9. *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. ☐

5.6 Limit Theorems

Theorem 5.10 (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 $\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$.

Proof. We will suppose that the inequalities hold for every $x \in X$. This leaves us with one inequality left to prove. We approximate each f_n by an increasing sequence of S_+ functions and then select a subsequence of these.
So for each $n \in \mathbb{N}$ we can pick $g_{n,1} \leq g_{n,2} \leq g_{n,3} \leq \ldots$ in S_+ such that $f_n = \lim_{k\to\infty} g_{n,k}$. Then for each $k \in \mathbb{N}$ we define

$$h_k := \max\{g_{1,k}, g_{2,k}, \ldots, g_{k,k}\} \in S_+$$

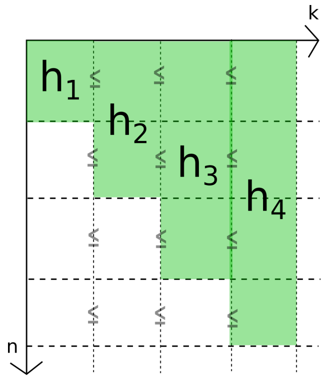

Figure 1: Visualizing the definition h_k. Each square represents a $g_{n,k}$.

Notice that $h_1 \leq h_2 \leq h_3 \leq \dots$ and $f = \lim_{k\to\infty} h_k$. Hence

$$\int f \, d\mu = \lim_{k\to\infty} \int h_k \, d\mu \leq \lim_{n\to\infty} f_n \, d\mu$$

In generality, we can pick $N \in \mathcal{A}$ such that $\mu(N) = 0$ and we have the assumed inequalities $\forall x \in N^{\mathsf{c}}$. We can then apply these previous arguments to N^{c} by considering the functions

$$f\chi_{N^{\mathsf{c}}}, \quad f_1\chi_{N^{\mathsf{c}}} \leq f_2\chi_{N^{\mathsf{c}}} \leq f_3\chi_{N^{\mathsf{c}}} \leq \dots$$

These functions differ on a set contained within a set of measure 0 and hence their integrals must agree with the full integrals. □

Corollary 5.11 (Levi's Theorem)**.** *Given measurable $g_n : X \to [0, +\infty]$ for each $n \in \mathbb{N}$.*

$$\int \left(\sum_{n=1}^{\infty} g_n \right) d\mu = \sum_{n=1}^{\infty} \left(\int g_n \, d\mu \right)$$

Proof. Take $f_n := \sum_{k=1}^{n} g_k$ then the infinite sum is $\lim_{n\to\infty} f_n$ and the f_n are monotone increasing so we can apply the MCT. □

Theorem 5.12 (Fatou's Lemma)**.** *Given a sequence $\{f_n\}$ of functions in $\overline{S_+}$,*

$$\int \left(\liminf_n f_n \right) d\mu \leq \liminf_n \int f_n \, d\mu$$

Proof. For each $k \in \mathbb{N}$ define $g_k := \inf_{n\geq k} f_n \in \overline{S_+}$.

$$g_1 \leq g_2 \leq g_3 \leq \dots \quad \text{and} \quad \liminf_n f_n = \lim_n g_n$$

Apply the Monotone Convergence Theorem to see

$$\int \left(\liminf_n f_n \right) d\mu = \int \left(\lim_n g_n \right) d\mu = \lim_n \int g_n \, d\mu$$

So we need to show $\lim_n \int g_n \, d\mu \leq \liminf_n \int f_n \, d\mu$. Notice for each $n \in \mathbb{N}$ that $g_n \leq f_n \leq f_{n+1} \leq \dots$ and hence

$$\lim_n g_n \, d\mu \leq \liminf_n \int f_n \, d\mu$$

□

Theorem 5.13 (Dominated Convergence Theorem)**.** *Suppose:*

(i) $g : X \to [0, +\infty]$ *is integrable*

(ii) $f, f_1, f_2, \cdots : X \to \overline{\mathbb{R}}$ *are measurable such that for almost every* $x \in X$

$$f(x) = \lim_{n\to\infty} f_n(x) \quad \textit{and} \quad \forall n \in \mathbb{N}\ |f_n(x)| \le g(x)$$

Then:

1. f *and each* f_i *are integrable*
2. $\int f\, d\mu = \lim_{n\to\infty} \int f_n\, d\mu$

Proof. We may assume that *(ii)* holds for every $x \in X$ since this won't change any integrals. Likewise we may assume that $g(x) \neq +\infty$ for all $x \in X$.

1. Given $n \in \mathbb{N}$, $|f_n| \le g \implies \int |f_n| < \int g < +\infty \implies f_n$integrable.

   Then $|f| = \lim_n |f_n| \le \lim_n g = g \implies f$integrable.

2. **Claim:** $\int (g + f)\ d\mu \le \liminf_n \int (g + f_n)\ d\mu$

   This follows by Fatou's Lemma because $g + f_n \ge 0$ is measurable and $g + f = \lim_n (g + f_n)$. Now,

   $$\begin{aligned} \int g\, d\mu + \int f\, d\mu &= \int (g + f)\, d\mu \\ &\le \liminf_n \left(\int (g + f_n)\ d\mu \right) \\ &= \int g\, d\mu + \liminf_n \int f_n\, d\mu \\ \text{and hence} \quad \int f\, d\mu &\le \liminf_n \int f_n\, d\mu \end{aligned}$$

   Now applying the same argument to $-f$ and $\{-f_n\}$ yields

   $$\int (-f)\, d\mu \le \liminf_n \int (-f_n)\, d\mu \implies \int f\, d\mu \ge \limsup_n \int f_n\, d\mu$$

   And hence we have $\int f\, d\mu = \lim_n \int f_n\, d\mu$.

□

It's also worth knowing that

Theorem 5.14. *Given bounded function* $f : [a, b] \to \mathbb{R}$

(a) f *is Riemann integrable* $\iff$ *for almost every* x*,* f *is continuous at* x

(b) In this case Riemann Integral = Lebesgue Integral

5.7 The Riemann Integral

Given a bounded function $f : [a, b] \to \mathbb{R}$, a partition is $P = \{a_i\}_{i=0}^k$ where

$$a = a_0 < a_1 < \cdots < a_{k-1} < a_k = b$$

We say P' refines P if $P \subseteq P'$ and P' is a partition.
We define the lower sum

$$l(f, P) := \sum_{i=1}^{k} (a_i - a_{i-1}) \inf(f[a_{i-1}, a_i])$$

and the upper sum

$$u(f, P) := \sum_{i=1}^{k} (a_i - a_{i-1}) \sup(f[a_{i-1}, a_i])$$

f is Riemann integrable if

$$\sup_P l(f, P) = \inf_P u(f, P)$$

Then this common value is the Riemann integral $(RI) \int_a^b f(x)dx$.

Theorem 5.15. *Let $a < b$ and $f : [a, b] \to \mathbb{R}$ be bounded*

(a) f is Riemann integrable if and only if

$$\lambda(\{x \in [a, b] \mid f \text{ not continuous at } x\} = 0$$

i.e. the set is λ-measurable and has measure 0.

(b) If one of (a) holds then the f is Lebesgue integrable and

$$(RI) \int_a^b f(x)dx = \int_a^b f d\lambda$$

Proof. Very long, should definitely be read. □

6 Theorems on Measures

$D \subseteq \mathcal{P}(X)$ is a d-system or Dynkin class if

(a) $X \in D$.

(b) $\forall A, B \in D$ such that $B \subseteq A$ we have $A \setminus B \in D$.

(c) D is closed under countable union.

Given any collection of sets $\mathcal{C}$, $d(\mathcal{C})$ is the smallest d-system containing $\mathcal{C}$.
$\mathcal{C} \subseteq \mathcal{P}(X)$ is a π-system if it is closed under finite intersections.

Lemma 6.1. *Let $\mathcal{C}$ be a π-system then $\sigma(\mathcal{C}) = d(\mathcal{C})$.*

Proof. A σ-algebra is a d-system and $d(\mathcal{C})$ is the smallest d-system containing $\mathcal{C}$ and hence we easily see $\sigma(\mathcal{C}) \supseteq d(\mathcal{C})$.
For the opposite direction we want to show that $\mathcal{D} := d(\mathcal{C})$ is a σ-algebra. First we show that it is closed under finite intersections. To this end define

$$D_1 := \{A \in \mathcal{D} \mid \forall C \in \mathcal{C} \quad A \cap C \in \mathcal{D}\}$$

Claim: D_1 is a d-system.
Once this has been shown we can see that $D_1 \supseteq \mathcal{C}$ because $\mathcal{C}$ is closed under intersections. Hence

$$D_1 \supseteq d(\mathcal{C}) \implies d(\mathcal{C}) = \mathcal{D} \supseteq D_1 \supseteq d(\mathcal{C}) \implies \mathcal{D} = D_1$$

Next define

$$D_2 := \{A \in \mathcal{D} \mid \forall C \in \mathcal{D} \quad A \cap C \in \mathcal{D}\}$$

Claim: D_2 is also a d-system.
Then again one can easily see that $D_2 \supseteq \mathcal{C}$ and hence

$$\mathcal{D} \supseteq D_2 \supseteq d(\mathcal{C}) = \mathcal{D} \implies D_2 = \mathcal{D}$$

which shows that $\mathcal{D}$ is closed under finite intersections.
So we have that $\mathcal{D}$ is a $(\pi + d)$-system which means that in fact $\mathcal{D}$ is a σ-algebra and thus yields the opposite inequality because $\sigma(\mathcal{C})$ is the smallest σ-algebra containing $\mathcal{C}$. $\square$

Corollary 6.2. *Given a measurable space $(X\mathcal{A})$ and a π-system $\mathcal{C} \subseteq \mathcal{P}(X)$ such that two measures μ and ν coincide on $\mathcal{C}$. If there exists an increasing sequence of subsets*

$$C_1 \subseteq C_2 \subseteq \dots \quad \mathit{in}\, \mathcal{C}$$

such that $\cup C_n = X$ and $\mu(C_n) < \infty$ then $\mu = \nu$.

Given $f : \mathbb{R} \to \mathbb{R}$ non-decreasing and right continuous at every $x \in \mathbb{R}$, we define the Lebesgue-Stieltfes measure by

$$\lambda_f^* := \inf \left\{ \sum_{i=1}^{\infty} (f(b_i) - f(a_i)) \;\middle|\; A \subseteq \cup_i (a_i, b_i] \right\}$$

Proposition 6.3. *Given a measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mu([a,b]) < \infty$ for all such $a < b$ define*

$$F_\mu(x) := \begin{cases} \mu((0,x]) & \text{for } x \geq 0 \\ -\mu((x,0]) & \text{for } x < 0 \end{cases}$$

Then F_μ is non-decreasing and right continuous and $f(0) = 0$.

This gives us a nice bijection

$$\{f : \mathbb{R} \to \mathbb{R} \mid \text{non-decreasing, right continuous, } f(0) = 0\} \leftrightarrow \{\text{measure } \mu \mid \mu((a,b]) < \infty \; \forall a < b\}$$

6.1 Product Measures

Given measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{C})$ a rectangle is any set $A \times C$ with $A \in \mathcal{A}$ and $C \in \mathcal{C}$. Define $\mathcal{R} := \{\text{rectangles}\}$ then the product σ-algebra is

$$\mathcal{A} \times \mathcal{C} := \sigma(\mathcal{R})$$

Given any subset $E \subseteq X \times Y$ and $f : X \times Y \to Z$, for $x \in X$ we define the section

$$E_x := \{y \in X \mid (x, y) \in E\}$$

and then

$$f_x : Y \to Z \quad \text{by} \quad y \mapsto f(x, y)$$

i.e. we restrict f to the vertical line E_x. We likewise define E^y and $f^y : X \to Z$ by restricting f to the horizontal line E^y.

Example:

$$\mathcal{B}(\mathbb{R}^2) = \sigma(\text{2D intervals}) \subseteq \sigma(\text{rectangles}) = \mathcal{B}(R) \times \mathcal{B}(R)$$

Given a rectangle $A \times C \in \mathcal{R}$ we can write $A \times C = A \times \mathbb{R} \cap \mathbb{R} \times C$. If we define projection to the first coordinate π_1 then we see

$$A \times \mathbb{R} = \pi_1^{-1}(A) \in \mathcal{B}(\mathbb{R}^2)$$

since $A \in \mathcal{B}(\mathbb{R})$ and projection is a continuous function. Likewise $\mathbb{R} \times C \in \mathcal{B}(\mathbb{R}^2)$ and hence $A \times C \in \mathcal{B}(\mathbb{R}^2)$. We may conclude that

$$\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$$

Lemma 6.4. *(a)* $E \in \mathcal{A} \times \mathcal{C} \implies \forall x \;\; E_x \in \mathcal{C}$ *and* $\forall y \;\; E^y \in \mathcal{A}$.

(b) If $f : X \times Y \to \mathbb{R}$ *is* $(\mathcal{A} \times \mathcal{C}, \mathcal{B}(\mathbb{R}))$*-measurable then* f_x *is* $\mathcal{C}$*-measurable and* f^y *is* $\mathcal{A}$*-measurable* $\forall x, y$.

Proof. This is done by the standard procedure:

(i) Prove that $\{E \subseteq X \times Y \mid E_x \in \mathcal{C}\}$ is a σ-algebra.

(ii) Prove that all rectangles belong to this set.

□

Proposition 6.5. *Given* σ*-finite measure spaces* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{C}, \nu)$ *and* $E \in \mathcal{A} \times \mathcal{C}$*, the function*

$$I_E : X \to [0, +\infty], \quad x \mapsto \nu(E_x)$$

is $\mathcal{A}$*-measurable for all* $x \in X$.

Proof. Show that I_E is measurable for all rectangles and then that the set

$$\mathcal{F} := \{E \in \mathcal{A} \times \mathcal{C} \mid I_E \text{ is measurable}\}$$

is a d-system. Then since the rectangles form a π-system we get

$$\mathcal{F} \supseteq d(R) = \sigma(R) = \mathcal{A} \times \mathcal{C}$$

□

Theorem 6.6. *Given σ-finite measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{C}, \nu)$ and $E \in \mathcal{A} \times \mathcal{C}$, there is a unique measure $\mu \times \nu$ on $(X \times Y, \mathcal{A} \times \mathcal{C})$ such that for all $A \times C \in \mathcal{R}$*

$$(\mu \times \nu)(A \times C) = \mu(A) \cdot \nu(C)$$

and moreover given any $E \in \mathcal{A} \times \mathcal{C}$

$$(\mu \times \nu)(E) = \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y)$$

Proof. We get uniqueness from the rectangle equality and our previous result about measure uniqueness. We then show that the last formula defines a measure with the desired properties. ☐

6.2 Fubini's Theorem

Proposition 6.7 (Tonelli's Theorem)**.** *Given σ-finite measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{C}, \nu)$ and some function $f : X \times Y \to [0, +\infty]$ which is $(\mathcal{A} \times \mathcal{C})$-measurable, the following holds*

(a) $x \mapsto \int_Y f_x d\nu$ is measurable.

(b) $\int_{X \times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f_x d\nu\right) d\mu$.

Proof. This proof follows the standard format of proving the result for simple functions and then extending it to measurable function by the monotone convergence theorem. ☐

Theorem 6.8 (Fubini's Theorem)**.** *Given σ-finite measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{C}, \nu)$ and $E \in \mathcal{A} \times \mathcal{C}$ and some function $f : X \times Y \to \overline{\mathbb{R}}$ which is $(\mu \times \nu)$-integrable then*

(a) For almost every $x \in X$, f_x is ν-integrable and for almost every $y \in Y$, f^y is μ-integrable.

(b) The function

$$I_f(x) := \begin{cases} \int f_x d\nu & \text{if } f_x \text{ is integrable} \\ 0 & \text{otherwise} \end{cases}$$

is μ-integrable and likewise $I^f(y)$ is ν-integrable.

(c) $\int_{X \times Y} f d(\mu \times \nu) = \int_X I_f d\mu = \int_Y I^f d\nu$.

Example: Algorithm for Fubini's Theorem
Given some measurable $f : X \times Y \to \overline{R}$

1. Write $f = f^+ - f^-$ which are both measurable.
2. Apply Tonelli's tells us $x \mapsto \int f_x^+ d\nu$ and $x \mapsto \int f_x^- d\nu$ are both measurable.
3. Compute

$$A^+ := \int_X \left(\int_Y f_x^+ d\nu\right) d\mu$$
$$A^- := \int_X \left(\int_Y f_x^- d\nu\right) d\mu$$

4. If both $A^+, A^- < \infty$ then Tonelli tells us that

$$\int_{X\times Y} f^+ d(\mu\times\nu) = A^+ < +\infty$$
$$\int_{X\times Y} f^- d(\mu\times\nu) = A^- < +\infty$$

5. Hence f is $(\mu\times\nu)$-integrable and Fubini tells us

$$\int_{X\times Y} f d(\mu\times\nu) = A^+ - A^-$$

6.3 Signed measures

For a measurable space $(X,\mathcal{A})$ and a function $\mu : \mathcal{A} \to [-\infty, +\infty]$ is called a signed measure if

(a) $\mu(\emptyset) = 0$

(b) Given any measurable disjoint sets $A_1, A_2, \dots$ we have countable additivity

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

Note:

- Since the left hand side of *(b)* is defined so to the right hand side must be defined. Hence there is no disjoint $A, B \in \mathcal{A}$ such that $\mu(A) = \infty$ and $\mu(B) = -\infty$ otherwise their union would not have a well-defined measure.
- Even more strongly, if $\mu(A) = \infty$ and $\mu(B) = -\infty$ for some $A, B \in \mathcal{A}$ then one of the following occurs:
  - $\mu(A\cap B) \neq \mu(B) \implies \mu(A\cap B) = \mu(B) - \mu(A^{\mathsf{c}}\cap B) \implies \mu(B\setminus A) = -\infty$
  - $\mu(A\cap B) \neq \mu(A) \implies \mu(A\setminus B) = +\infty$

  These are both contradictions and so we can assume that one of $\pm\infty$ never occurs.

For a signed measure μ on $(X,\mathcal{A})$, a set $A \subseteq X$ is called a positive set (resp. negative set) if:

(i) $A \in \mathcal{A}$.

(ii) $\forall B \subseteq A$ such that B is measurable we have $\mu(B) \geq 0$ (resp. $\mu(B) \leq 0$).

Lemma 6.9. *Given a signed measure μ and $A \in \mathcal{A}$,*

$$-\infty < \mu(A) < 0 \implies \exists \text{ negative set } B \subseteq A \text{ such that } \mu(B) \leq \mu(A)$$

Proof. We proceed by induction on n, contracting a measurable set A_n each time. For each $n \in \mathbb{N}$ define $\delta_n := \sup\left\{\mu(E) \,\middle|\, E \text{ measurable }, E \subseteq A \setminus \left(\bigcup_{i=1}^{n-1} A_i\right)\right\}$. Note that $\delta_n \geq 0$ since we may always take the empty set.

Now pick any measurable $A_n \subseteq A \setminus \left(\bigcup_{i=1}^{n-1} A_i\right)$ such that $\mu(A_n) \geq \min(\frac{\delta_n}{2}, 1)$.
Having done this process let $A_\infty = \cup_n A_n$ and then let $B := A \setminus A_\infty$.
Then $\mu(A_\infty) = \sum_n \mu(A_n) \geq 0$. So by finite additivity $\mu(A) = \mu(A_\infty) + \mu(B) \geq \mu(B)$.
We claim that B is a negative set. Since $\mu(A) > -\infty$ and $\sum_n \mu(A_n) = \mu(A_\infty) < +\infty$. Then since the sum converges we must have $\mu(A_n) \to 0$ and hence $\delta_n \to 0$.
Now take any measurable $E \subseteq B$, we must have that $\mu(E) \leq \delta_n$ for all n and hence $\mu(E) \leq 0$. □

Theorem 6.10 (Kahn Decomposition Theorem)**.** *Given any signed measure μ on $(X\mathcal{A})$ there is a partition $X = P \sqcup N$ such that P is a positive set and N is a negative set.*

Proof. WLOG we can assume that $\mu : \mathcal{A} \to (-\infty, +\infty]$ then we define

$$L := \inf\{\mu(A) \mid A \text{ is a negative set}\}$$

then choose any negative sets A_n such that $\mu(A_n) \to L$ (note we don't know yet whether they are disjoint).
Define $N := \cup_n A_n$ which is negative since for all measurable $B \subseteq N$

$$\mu(B) = \underbrace{\mu(B \cap A_1)}_{\subseteq A_1} + \underbrace{\mu(B \cap (A_2 \setminus A_1))}_{\subseteq A_2} + \cdots \leq 0$$

and hence $L \leq \mu(N)$. Now for every n we have that

$$\mu(N) = \mu(A_n) + \underbrace{\mu(N \setminus A_n)}_{\leq 0} \leq \mu(A_n) \implies \mu(N) \leq L$$

by taking the limit. Hence we have $\mu(N) = L > -\infty$. Now let $P = X \setminus n$, this is a positive set. □

Theorem 6.11 (Jordan Decomposition Theorem)**.** *For every Hahn decomposition theorem of $X = P \sqcup N$ of a signed measure μ on $(X, \mathcal{A})$ then there are measures μ^+, μ^- such that at least one one is finite such that $\mu = \mu^+ - \mu^-$ and $\mu^+(N) = 0$ and $\mu^-(P) = 0$.*
Moreover, such measures are unique and do not depend on the choice of N and P.

Proof. **Existence:** Define $\mu^+(A) = \mu(A \cap P)$ and $\mu^-(A) = -\mu(A \cap N)$ for all $A \in \mathcal{A}$. Then since A is measurable we have that $\mu(A) = \mu(A \cap P) + \mu(A \cap N) = \mu^+(A) - \mu^-(A)$ and $\mu(N \cap P) = \mu(\emptyset) = 0$.
At least one of the holds $\mu(N) = -\infty$, $\mu(P) = -\infty$ since they are disjoint sets. Hence one of the new measures is finite.
Independence on decomposition: Given any $A \in \mathcal{A}$ we would like to show that

$$\mu^+(A) = \sup\{\mu(B) \mid B \subseteq A \text{ measurable}\}$$
$$\mu^-(A) = \sup\{-\mu(B) \mid B \subseteq A \text{ measurable}\}$$

These do not depend on N or P so we get our uniqueness, we will just prove the first identity. Given any $B \subseteq A$ we can notice

$$\mu(B) = \mu^+(B) - \mu^-(B) \leq \mu^+(B) \leq \mu^+(A)$$

and hence we have $\geq$. Then we need to find a measurable B such that $\mu(B) \geq \mu^+(A)$. Well notice

$$\mu^+(A) = \mu^+(A \cap P) + \mu^+(A \cap N) = \mu^+(A \cap P) = \mu^+(A \cap P) - \mu - (A \cap P) = \mu(A \cap P)$$

and so we can just take take $B = A \cap P$. □

6.4 Absolute continuity

Given two measures μ, ν on a $(X, \mathcal{A})$ we say that ν is absolutely continuous with respect to μ if

$$\forall A \in \mathcal{A} \quad \mu(A) = 0 \implies \nu(A) = 0$$

and we write $\nu << \mu$.

Lemma 6.12. *Suppose that μ and ν are measures on $(X, \mathcal{A})$ and that $\nu(X) < +\infty$, Then*

$$\nu << \mu \iff \forall \epsilon > 0 \ \exists \delta > 0 \ \ s.t. \ \ \mu(A) < \delta \implies \nu(A) < \epsilon$$

Proof. "$\Longleftarrow$" : Let $\mu(A) = 0$ then given any $\epsilon > 0$ there is a $\delta > 0$ such that $\mu(A) < \delta \implies \nu(A) < \epsilon$. But $\mu(A) < \delta$ for any such delta and hence $\nu(A) < \epsilon$ for any such ϵ and hence $\nu(A) = 0$.
"$\Longrightarrow$" : Suppose not then there exists an $\epsilon > 0$ and a sequence of sets $A_k \in \mathcal{A}$ such that

$$\mu(A_k) < \frac{1}{2^k} \quad \text{but} \quad \nu(A_k) \geq \epsilon$$

Now let $B_n :== \bigcup_{k-n}^{\infty} A_k$. Notice that $\mu(B_n) \leq \sum_{k=n}^{\infty} \mu(A_k) < \sum_{k=n}^{\infty} \frac{1}{2^k} = \frac{1}{2^{n-1}}$. Moreover, $\nu(B_n) \geq \nu(A_n) \geq \epsilon$. Let $B := \bigcap_{n=1}^{\infty} B_n$ so that

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots \supseteq B$$

Since we assumed that ν was finite, Borel-Cantelli tells us that

$$\nu(B) = \lim_{n \to \infty} \nu(B_n) \geq \epsilon$$

But by assumption $\mu(B) = 0$. This contradicts absolute continuity. □

Theorem 6.13 (Radon-Nikodyn Theorem)**.** *Suppose μ and ν are σ-finite measures on $(X, \mathcal{A})$ such that $\nu << \mu$. Then $\exists$ a measurable function $f : x \to [0, +\infty)$ such that*

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

Moreover, for all such functions h, we have that $h = f$ μ-almost everywhere. We call such function the Radon-Nikodyn derivative which we denote $\frac{d\nu}{d\mu}$.

Proof. Look over this! □

Note: We do need the σ-finite assumption
Let μ be the counting measure and λ the Lebesgue measure on $([0, 1], \mathcal{B})$. Then $\lambda << \mu$ since μ is only 0 on the empty set.
Can we have $\lambda(A) = \int_A g \, d\mu$? No:
If g is non-zero on at least one point then look at $A = \{x\}$ then $\lambda(A) = 0$ but $\int_A g \, d\mu = g(x) \neq 0$.
So g must be identically zero which easily leads to a contradiction.

6.5 $\mathcal{L}^p$ spaces

Fix some $(X, \mathcal{A})$ measure space and real number $p \in [1, +\infty)$ then

$$\mathcal{L}^p := \{\text{measurable } f : X \to \mathbb{R} \mid |f|^p \text{ is integrable}\}$$

We can also define

$$\mathcal{L}^\infty := \{\text{bounded measurable } f : X \to \mathbb{R}\}$$

We can give this space a norm by

$$||f||_\infty := \inf \{M \geq 0 \mid \{|f| > M\} \text{ is locally } \mu\text{-null}\} \in [0, +\infty)$$

$A \subseteq X$ is called μ-null if $\exists N \in \mathcal{A}$ such that $\mu(N) = 0$ and $N \supseteq A$.
A is called locally μ-null if $\forall B \in \mathcal{A}$ with $\mu(B) < +\infty$, $A \cap B$ is null.
$p, q \in (1, +\infty)$ are conjugate exponents if $\frac{1}{p} + \frac{1}{q} = 1$ or $\{p, q\} = \{1, \infty\}$.

Lemma 6.14 (Young's Inequality)**.** *Given conjugate exponents $p, q \in (1, \infty)$ and $x, y \geq 0$*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

Proposition 6.15 (Holder's Inequality)**.** *Given conjugate exponents $p, q \in [1, +\infty]$. If $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$ then $fg \in \mathcal{L}^1$ and*

$$\int |fg| \, d\mu \leq ||f||_p ||g||_q$$

Proposition 6.16 (Minkowski's Inequality)**.** *Given any $p \in [1, +\infty]$,*

$$f, g \in \mathcal{L}^p \implies ||f + g||_p \leq ||f||_p + ||g||_p$$

Corollary 6.17. *Given any $p \in [1, +\infty]$, then $\mathcal{L}^p(X, \mathcal{A}, \mu)$ is a vector space and $||\cdot||_p$ is a semi-norm.*

Let $\mathcal{N}^p = \mathcal{N}^p(X, \mathcal{A}, \mu) := \{f \in \mathcal{L}^p \mid ||f||_p = 0\}$
Then we can define $L^p := \frac{\mathcal{L}^p}{\mathcal{N}^p}$. This can be seen to be a normed space.

Theorem 6.18. *Given any $p \in [1, +\infty]$, $(L^p, ||\cdot||_p)$ is complete.*

Proof. It is enough to show that for all $\{f_n\}$ in $\mathcal{L}^p$

$$\sum_{k=1}^{\infty} ||f_k||_p < +\infty \implies \exists f \in \mathcal{L}^p \quad s.t. \quad ||\sum_{k=1}^{n} f_k - f||_p \to 0$$

Define $g_n : X \to [0, +\infty]$ by $g_n(x) = \sum_{k=1}^{n} |f_k(x)|$ and then $g(x) := \lim_{n\to\infty} g_n(x)^p$.
By the monotone convergence theorem we can see that $\int |g| \, d\mu < +\infty$.
Then define

$$f(x) := \begin{cases} \sum_{k=1}^{\infty} & \text{if } g(x) \neq +\infty \\ 0 & \text{otherwise} \end{cases}$$

Then by the dominated convergence theorem wee see that

$$\int |f|^p \, d\mu \leq \lim_{n\to\infty} \int \sum < +\infty$$

so on and so forth... $\square$

7 Modes of Convergence

Given a measure space $(X, \mathcal{A}, \mu)$and measurable functions $f, f_1, f_2, f_3, \cdots : X \to \mathbb{R}$ we say

- (f_n) converges to f almost everywhere if

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for almost every } x \in X$$

- (f_n) converges to f in measure if

$$\forall \epsilon > 0 \qquad \lim_{n\to\infty} \mu\left(\{x \in X \mid |f_n(x) - f(x)| > \epsilon\}\right) = 0$$

Proposition 7.1. *Given a finite measure space, almost everywhere convergence $\implies$ convergence in measure.*

Proof. Given $\epsilon > 0$ define $A_n := \{|f_n - f| > \epsilon\}$. Let $B_n := \cup_{i\geq n} A_i$ be the set of points where f_n and f differ by more than ϵ at some point in the future. Notice that $B_1 \supseteq B_2 \supseteq \dots$ and all the B_i are measurable. Now define $B = \cap_n B_n$ and note that $B \subseteq \{x \in X \mid f_n(x) \not\to f(x)\}$ since B is the set of points where no matter how far in the sequence you go there will always be a time in the future where f_n and f disagree significantly. Therefore

$$0 = \mu(B) = \mu(\cap_n B_n) = \lim_{n\to\infty} \mu(B_n)$$

where the last inequality is thanks to μ being a finite measure. But note that $A_n \subseteq B_n$ for all n and hence $\mu(A_n) \to 0$ as required. $\square$

Lemma 7.2 (Borel-Cantelli)**.** *Given a measure space $(X, \mathcal{A}, \mu)$and $A_1, A_2, \dots \in \mathcal{A}$ such that $\sum_n \mu(A_n) < \infty$, let*

$$A := \{x \in X \mid x \in A_n \textit{ for infinitely many } n\} = \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right)$$

then $\mu(A) = 0$.

Corollary 7.3. *Given a measure space $(X, \mathcal{A}, \mu)$and $f_1, f_2, f_3, \dots : X \to \mathbb{R}$ all measurable, if $f_n \to f$ in measure then $\exists n_1 < n_2 < \dots$ such that $f_{n_i} \to f$ almost everywhere.*

Theorem 7.4 (Egorov's Theorem)**.** *$(X, \mathcal{A}, \mu)$a measure space and $f, f_1, f_2, \dots : X \to \mathbb{R}$ measurable such that $f_n \to f$ almost everywhere. If $\mu(X) < \infty$ then*

$$\forall \epsilon > 0\ \exists B \in \mathcal{A} \text{ s.t. } \mu(B^{\mathsf{c}}) < \epsilon \ \text{ and } \ f_n \to f \text{ uniformly on } B$$

Proof. Take arbitrary $\epsilon > 0$. For each $n \in \mathbb{N}$ define

$$g_n(x) := \sup_{j\geq n} |f_j(x) - f(x)|$$

Then $g_n : X \to [0, +\infty]$ is measurable and since $f_n \to f$ almost everywhere, so too $g_n \to 0$ almost everywhere. Now these g_n may have some infinite points so we define

$$g'_n(x) := \begin{cases} g_n(x) & \text{if } g_n(x) < \infty \\ 0 & \text{otherwise} \end{cases}$$

Note we still have $g'_n \to 0$ almost everywhere and hence $g'_n \to 0$ in measure. Therefore, for every $k \in \mathbb{N}$ we have n_k such that

$$\mu\left\{g'_{n_k} > \frac{1}{k}\right\} < \frac{\epsilon}{2^k}$$

Define $B_k := \left\{g_{n_k} \leq \frac{1}{k}\right\}$ and subsequently $B := \cap_k B_k$. Note also that $\mu(B_k^{\mathsf{c}}) < \frac{\epsilon}{2^k}$ and hence

$$\mu(B^{\mathsf{c}}) \leq \sum_k \mu(B_k^{\mathsf{c}}) < \sum_k \frac{\epsilon}{2^k} < \epsilon$$

Finally, to prove uniform convergence, given any $\delta > 0$ take $k > \frac{1}{\delta}$ then for any $x \in B$ and $n \geq n_k$ we have

$$|f_n(x) - f(x)| \leq g_{n_k}(x) \underbrace{\leq}_{\text{since}} \frac{1}{k} < \delta$$

□

For $f, f_1, f_2, \dots \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ we say (f_n) converges to f in mean if

$$\int |f_n(x) - f(x)| d\mu(x) \to 0 \quad \text{as } n \to \infty$$

Lemma 7.5. *Convergence in mean $\implies$ convergence in measure.*

Proof. Using Markov's inequality, given any $\epsilon > 0$,

$$\mu\{|f_n - f| > \epsilon\} \leq \frac{1}{\epsilon}\int |f_n - f|\, d\mu \to 0$$

□

Proposition 7.6. *$(X, \mathcal{A}, \mu)$ a measure space and $f, f_1, f_2, \dots \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ If $f_n \to f$ almost everywhere or in measure and there is an integrable $g : X \to [0, +\infty]$ such that for almost every x, $|f| \leq g$ and for all $n \in \mathbb{N}$ $|f_n| \leq g$ then $f_n \to f$ in mean.*

Proof. Suppose $f_n \to f$ almost everywhere then almost everywhere we have $|f_n - f| \leq 2g$. We then apply dominated convergence theorem to get

$$\lim_{n\to\infty} \int |f_n - f|\, d\mu = \int \lim_{n\to\infty} \underbrace{|f_n - f|}_{0 \text{ a.e.}}\, d\mu = 0$$

Now suppose $f_n \to f$ in measure but, for contradiction, not in mean. So there is an $\epsilon > 0$ and a sequence $n_1 < n_2 < \dots$ such that for all k

$$\int |f_{n_k} - f|\, d\mu > \epsilon \tag{1}$$

Convergence in measure implies the existence of an almost everywhere convergent subsequence so we have $k_1 < k_2 < \dots$ such that $f_{n_{k_i}} \to f$ almost everywhere as $i \to \infty$. By the first part of the proof $f_{n_{k_i}} \to f$ in mean which contradicts (1). □

Theorem 7.7 (Lusin's Theorem)**.** *Let $A \in \mathcal{L}(\mathbb{R}^d)$ with $\lambda(A) < \infty$ and $f : A \to \mathbb{R}$ Lebesgue measurable. Then $\forall \epsilon > 0$ $\exists$compact $K \subseteq A$ such that $\lambda(A \setminus K) < \epsilon$ and $f\big|_K$ is continuous.*