Notes on Probability and Measure Theory

August 17, 2021

Contents

1 Overview

1.1 References

The primary reference here is [1]. The book is wonderful for statistical machine learning – it is rigorous, but also accessible (prerequisites are undergrad-level real analysis and mathematical probability). Most importantly, it is structured to build towards the kinds of applications in probability that we care about. (A point of contrast would be a book like that of Stein and Shakarchi, which tends to dwell heavily on things that are of higher interest to pure mathematicians — long existence proofs, Cantor sets and fractals, etc.)

Unless otherwise specified, all references to the "text" refers to this textbook. Likewise the symbol § refers to a Section of that textbook.

1.2 Motivation

Measure theory serves as a critical underpinning for some of the most interesting research in Bayesian statistics and probabilistic machine learning (see work from Stephen G. Walker, Michael Jordan, Tamara Broderick, David Dunson, and so on). Thus, fluency with measure theory opens doors to a higher level of research consumption.

Measure theory is convenient in unifying various kinds of random variables.[1] Lesbesgue integrals have nice limit theorems, and can be seen as the completion of Reimann integrals (in the same way that the real numbers complete the rationals).

2 § 1.1: Some notes on set theory

2.1 Limits of sequences of sets

Definition 2.1.1. The **upper limit** of a sequence of sets is given by

$$\limsup A_n := \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k$$

Alternatively,

$$x \in \limsup A_n \text{ iff } x \in A_n \text{ for infinitely many } n$$

△

Definition 2.1.2. The **lower limit** of a sequence of sets is given by

$$\liminf A_n := \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} A_k$$

Alternatively,

$$x \in \liminf A_n \text{ iff } x \in A_n \text{ eventually (for all but finitely many } n \text{)}$$

△

Discussion 2.1.1. Discuss why the two characterizations of upper limit and lower limit are equivalent. △

Definition 2.1.3. If $\liminf A_n = \limsup A_n = A$, then A is called the **limit** of the sequence $A_1, A_2, \ldots$. △

---

[1]For example, it allows one to work with discrete and absolutely continuous random variables in a unified way. For example, the exponential family includes both types of random variables.
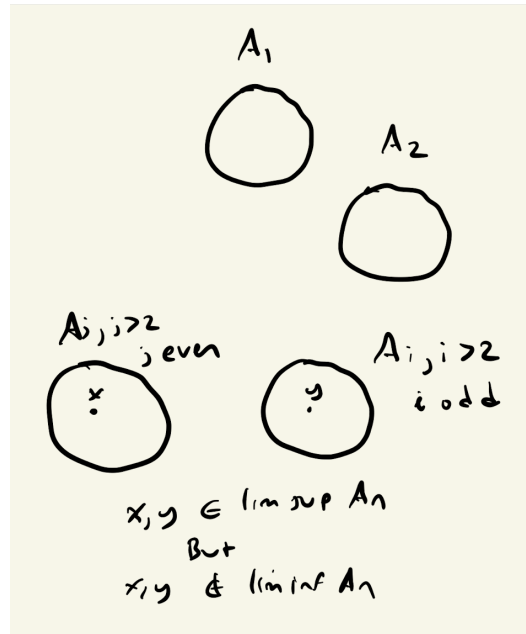
Figure 1: A sequence of sets with empty lower limit and non-empty upper limit.

Now we present a particular kind of limit that will be useful when we discuss continuity of measure.

Definition 2.1.4. If $A_1 \subset A_2 \subset \ldots$ and $\cup_{n=1}^{\infty} A_n = A$, we say that the A_n form a **increasing** sequence of sets with limit A or that the A_n increase to A; we write $A_n \uparrow A$. If $A_1 \supset A_2 \supset \ldots$ and $\cap_{n=1}^{\infty} A_n = A$, we say that the A_n form a **decreasing** sequence of sets with limit A or that the A_n decrease to A; we write $A_n \downarrow A$. $\triangle$

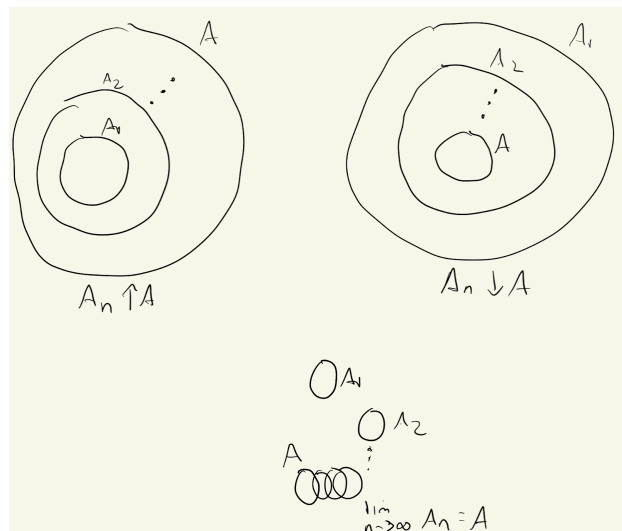Figure 2: An increasing and decreasing sequence of sets, followed by a sequence of sets which is neither, but which has a limit.

One can verify that this definition is consistent with the definition of limits, i.e.

$$\text{If } A_n \uparrow A \text{ or } A_n \downarrow A \text{ then } \liminf A_n = \limsup A_n = A.$$

As shown in Figure 2, limits of increasing and decreasing sequences are very special kinds of limits.

4

2.2 Representing unions as disjoint unions

Remark 2.2.1.   If $A_1, A_2, \ldots$ are subsets of some set Ω, then

$$\bigcup_{n=1}^{\infty} A_n = \dot{\bigcup_{n=1}^{\infty}} \left(A_n \cap A_{n-1}^c \cap \ldots \cap A_1^c \right) \tag{2.2.1}$$

In other words, any union can be re-represented as a disjoint union. This is useful because measures are countably additive on disjoint sets, so we prefer to work with collections of disjoint sets.   △

Remark 2.2.2.   If $A_n \uparrow A$, then (2.2.1) becomes

$$\bigcup_{n=1}^{\infty} A_n = \dot{\bigcup_{n=1}^{\infty}} \left(A_n - A_{n-1} \right) \tag{2.2.2}$$

This is because $A_{n-1} \subset A_n$, so $A_{n-1}^c \supset A_n^c$ by contraposition.   △

3 § 1.2: Fields, σ-fields, measures

3.1 § 1.2.1-1.2.2: Fields and σ-fields

Probability measures, and measures more generally, cannot be defined on all subsets of many spaces that we would like to deal with. For instance, non-measurable sets can be shown to exist even for Lesbesgue measure on the unit interval. Proposition 1.2.6 of [2] shows that there is no definition of $P(A)$ that is defined for all subsets $A \subseteq [0,1]$ satisfying all three conditions below

1. $P([a,b]) = b - a, \quad 0 \le a \le b \le 1.$

2. $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} A_n$ for $A_1, A_2, \ldots$ disjoint subsets of $[0,1]$.

3. $P(A \bigoplus r) = P(A), \quad 0 \le r \le 1$, where $A \bigoplus r$ denotes the *r-shift* of A, i.e.

$$A \bigoplus r := \{a + r : a \in A, a + r \le 1\} \cup \{a + r - 1 : a \in A, a + r > 1\}$$

The solution to this problem is to define measures on a restricted domain, σ-fields.

3.1.1 σ-fields

Definition 3.1.1.   Let $\mathcal{F}$ be a collection of subsets of a set Ω. Then $\mathcal{F}$ is called a **sigma-field** (or *sigma-algebra*) if it satisfies

a) $\Omega \in \mathcal{F}$

b) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

c) If $A_1, A_2, \ldots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

that is, if $\Omega \in \mathcal{F}$ and $\mathcal{F}$ is closed under complementation and countable unions.   △

Remark 3.1.1.   It follows that σ-fields are closed under countable intersections, since

$$\cap_{i=1}^{\infty} A_i \stackrel{\text{DeMorgan's Law}}{=} \cup_{i=1}^{\infty} A_i^c$$

△

Example 3.1.1. $\mathcal{F} = \{\emptyset, \Omega\}$ is the smallest σ-field on Ω. △

Example 3.1.2. $\mathcal{F} = 2^\Omega$, i.e. the set of all subsets of Ω, is the largest σ-field on Ω. △

Example 3.1.3. If $A \in \Omega$ is non-empty, then $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ is the smallest σ-field containing A. △

Notation 3.1.1. If $\mathcal{C}$ is a class of sets, the smallest σ-field containing the sets of $\mathcal{C}$ is written as $\sigma(\mathcal{C})$. This is sometimes called the *minimal σ-field over C* or the *σ-field generated by C*. △

Exercise 3.1.1. Let $A_1, ..., A_n$ be subsets of Ω. Describe $\mathcal{F} := \sigma(\{A_1, ..., A_n\})$, the smallest σ-field containing $A_1, ..., A_n$. Also describe the number of sets in $\mathcal{F}$. *This is Ash's Problem 1.2.8. We can derive the strict upper bound $|\mathcal{F}| \le 2^{2^n}$. For a complete answer, see GoodNotes.* △

Remark 3.1.2. The gist of exercise 3.1.1 is that the collection $\{A_1, ..., A_n\}$ partitions Ω into up to $M = 2^N$ pieces, and the minimal sigma field contains all possible finite unions of these pieces, so has at most 2^M elements.

△
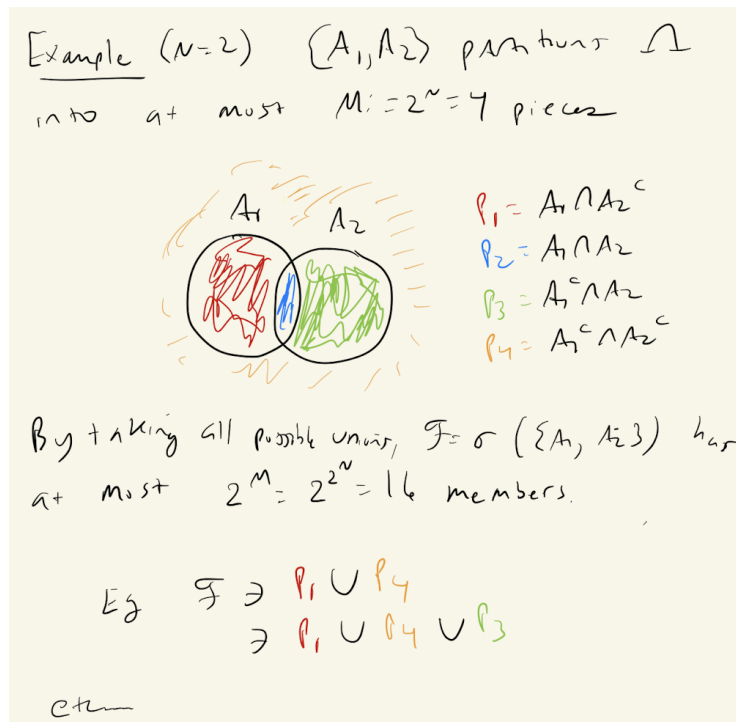
3.1.2 Fields

Fields are more general than σ-fields. Measures are sometimes constructed by being defined on fields, and then extended to σ-fields. Indeed, we will see this strategy with Lesbesgue measure.

Definition 3.1.2. Let $\mathcal{F}$ be a collection of subsets of a set Ω. Then $\mathcal{F}$ is called a **field** (or *algebra*) if satisfies Definition 3.1.1 after replacing condition c) with

c') If $A_1, ...A_n \in \mathcal{F}$ then $\cup_{i=1}^n A_i \in \mathcal{F}$.

that is, if $\Omega \in \mathcal{F}$ and $\mathcal{F}$ is closed under complementation and *finite* unions. △

Example 3.1.4. What is an example of a collection that is a *field*, but not a σ-*field*?

Let $\Omega = \mathbb{R}$ and $\mathcal{F}_0 = \{$finite disjoint unions of right semi-closed intervals $(a, b], a \neq b\}$. Then $\mathcal{F}_0$ is a field, as can be easily verified.[2]
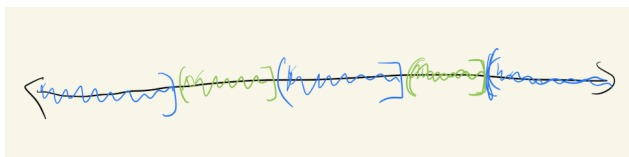
But $\mathcal{F}_0$ is <u>not</u> a σ-field. Note that if $A_n = (-\frac{1}{n}, 0]$, then $\bigcap_{n=1}^{\infty} A_n = \{0\} \notin \mathcal{F}_0$. $\qquad\qquad \triangle$

Remark 3.1.3. A σ-field can also be described as a field that is closed under limits of increasing sequences. For if $A_n \in \mathcal{F}$ and $A_n \uparrow A$, then A is a countable union of sets in $\mathcal{F}$ by definition. Conversely, if $A = \cup_{n=1}^{\infty} A_n$, then set $B_N := \cup_{n=1}^{N} A_n$ and $B_N \uparrow A$. So if $\mathcal{G}$ is the collection of all limits of increasing sequences of sets in some field $\mathcal{F}_0$, we can also describe $\mathcal{G}$ as the collection of all countable unions of sets in $\mathcal{F}_0$. $\qquad\qquad \triangle$

3.1.3 "Good sets" strategy

Ash says that there is a type of reasoning that occurs so often in problems involving σ-fields that it deserves explicit mention. It is called the *good sets strategy*. Suppose you want to show that all members of a σ-algebra $\mathcal{F}$ have some property P. Define "good sets" as those that satisfy the property
$$\mathcal{G} := \{G \in \mathcal{F} : G \text{ has property } P\}$$

The strategy is then to simply

1. Show $\mathcal{G}$ is a σ-algebra

2. Show $\mathcal{G}$ contains some class $\mathcal{C}$ such that $\mathcal{F} = \sigma(\mathcal{C})$

Then you're done!

Why does this work?

$$
\begin{aligned}
&\mathcal{C} \subset \mathcal{G} &&\text{by 2}\\
&\implies \sigma(\mathcal{C}) \subset \sigma(\mathcal{G})\\
&\implies \mathcal{F} \subset \mathcal{G} &&\text{by 1,2}
\end{aligned}
$$
Yet $\mathcal{G} \subset \mathcal{F}$ by definition of $\mathcal{G}$.
So $\mathcal{G} = \mathcal{F}$.
So all sets in $\mathcal{F}$ are good.

Some example applications:

- In the text, Ash uses this strategy (see pp.5) to show that if $\mathcal{C}$ is a class of subsets of Ω, and $A \in \Omega$, then

$$\underbrace{\sigma_\Omega(\mathcal{C}) \cap A}_{\text{take minimal sigma field first, then intersect}} \quad = \quad \underbrace{\sigma_A(\mathcal{C} \cap A)}_{\text{intersect first, then take minimal sigma-field}}$$

---

[2]By convention, we also count (a, ∞) as right semi-closed for $-\infty \leq a < \infty$, which is necessary for the σ-field to be closed under complements.

- See my handwritten homework exercise for § 1.2, Problem 6.

- See the proof of Caratheodory Extension Theorem (Theorem 4.1.3).

3.2 § 1.2.3-1.2.4: Measures

Definition 3.2.1. A **measure** on a σ-field $\mathcal{F}$ is a non-negative, extended real-valued function μ on $\mathcal{F}$ such that whenever $A_1, A_2, ...$ form a finite or countably infinite collection of disjoint sets in $\mathcal{F}$, we have countable additivity; that is,

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$$

$\triangle$

Definition 3.2.2. A **probability measure** is a measure (Definition 3.2.1) where $\mu(\Omega) = 1$. $\quad \triangle$

Remark 3.2.1. Ash additionally assumes that a measure does not take $\mu(A) = \infty$ for all $A \in \mathcal{F}$.[3] From this, we automatically obtain $\mu(\emptyset) = 0$. For $\mu(A) < \infty$ for some A, and by considering the sequence $A, \emptyset, \emptyset, ...$, we have that $\mu(\emptyset) = 0$ by countable additivity. $\quad \triangle$

Example 3.2.1. Let Ω be any set. Fix $x_0 \in \Omega$. Let $\mathcal{F} = 2^\Omega$. For any $A \in \mathcal{F}$ define $\mu(A) = 1$ if $x_0 \in A$ and $\mu(A) = 0$ if $x_0 \notin A$. Then μ may be called the **unit mass** concentrated at x_0. $\quad \triangle$

Example 3.2.2. Let $\Omega = \{x_1, x_2, ...\}$ be a finite or countably infinite set. Let $p_1, p_2, ...$ be non-negative reals. Let $\mathcal{F} = 2^\Omega$. Define

$$\mu(A) = \sum_{x_i \in A} p_i \quad \text{for all } A \in \mathcal{F}$$

Then μ is a measure on $\mathcal{F}$. We might call it the "point weighting" measure.

- If $p_i \equiv 1 \, \forall \, i$, then μ is called the **counting measure**.

- If $\sum_i p_i = 1$, then μ is a probability measure.

$\triangle$

Example 3.2.3. (*Lesbesgue measure*) Define μ such that

$$\mu(a, b] = b - a \quad \forall \, a, b \in \mathbb{R} : b > a$$

As we will see in Section 4, this requirement determines μ on a large collection of sets, the Borel Sets $\mathcal{B}(\mathbb{R})$, defined as the smallest σ-field of subsets of $\mathbb{R}$ containing all intervals $(a, b] \subset \mathbb{R}$.

We may alternately characterize $\mathcal{B}(\mathbb{R})$ as the smallest σ-field containing

- all intervals $(a, b]$, $a, b \in \mathbb{R}$

- all intervals (a, b), $a, b \in \mathbb{R}$

- all intervals $[a, b)$, $a, b \in \mathbb{R}$

- all intervals $[a, b]$, $a, b \in \mathbb{R}$.

- all intervals (a, ∞), $a \in \mathbb{R}$.

- all intervals $[a, \infty)$, $a \in \mathbb{R}$.

---

[3]Likewise, he assumes that signed measures do not take $\mu(A) = -\infty$ for for all $A \in \mathcal{F}$.

- all intervals $(-\infty, b)$, $b \in \mathbb{R}$.

- all intervals $(-\infty, b]$, $b \in \mathbb{R}$.

- all open sets of $\mathbb{R}$.[4]

- all closed sets of $\mathbb{R}$.[5]

To illustrate these equivalences, let us equate the first two conditions. That is, let us show that a σ-field contains all open intervals (a, b) iff it contains all right semi-closed intervals $(a, b]$. To see this, simply note

$$(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n} \right) \tag{3.2.1a}$$

and

$$(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n} \right] \tag{3.2.1b}$$

$\triangle$

Question 3.2.1. The text gives another description of the Borel sets $\mathcal{B}(\mathbb{R})$ as the smallest σ-field containing $\mathcal{F}_0$, the field of disjoint unions of right semi-closed intervals $(a, b]$. Can we make the same statement about the field of finite disjoint unions of left semi-closed intervals? $\triangle$

3.3 §1.2.5-1.2.6: Properties of measures (and some more general set functions)

The text considers some generalizations of measures that can be obtained

1. by restricting the domain to a field (in other texts, such functions are called *pre-measures*)

2. by only assuming *finite* additivity

3. by allowing the range to be extended reals ($\bar{\mathbb{R}}$) instead of non-negative extended reals ($\bar{\mathbb{R}}_{\geq 0}$).

Remark 3.3.1. With respect to pre-measures, a countably additive function can be defined on a *field* (rather than σ-field) if the condition is taken to hold whenever a countable union *does* happen to still be in the field. Unless otherwise specified, I will assume in these notes by that countably additive functions are always defined on σ-fields, and finitely additive functions are defined on fields. $\triangle$

	Range	
	non-negative extended reals	**extended reals**
countably additive	μ measure	$\tilde{\mu}$ signed measure
finitely additive	μ_0	$\tilde{\mu}_0$

Table 1: Notation for generalizations of measure (For assumed domain in each case, see Remark 3.3.1.)

In Table 1, we introduce some notation to try to clarify more immediately when results hold. Note the relations[6]

$$\{\mu\} \subset \{\mu_0\}, \{\tilde{\mu}\} \subset \{\tilde{\mu}_0\}.$$

---

[4]Recall that an open set is a countable union of open intervals.

[5]Recall that a set is open iff its complement is closed.

[6]So, for example, if something holds for $\tilde{\mu}_0$, it holds for μ. A simple mnemonic is that adding stuff to the notation generalizes the function.

Remark 3.3.2. Being able to work with these generalizations will be important in Section 4 on extension of measures. In particular, it will help us show that we can construct the Lesbesgue measure on the Borel sets. △

Example 3.3.1. Let $\mathcal{F}_0$ be the field of finite disjoint unions of right semi-closed intervals (see Definition A.1.1), and define the set function $\tilde{\mu}_0$ on $\mathcal{F}_0$ as follows[7]:

$$
\begin{aligned}
\tilde{\mu}_0(-\infty, a] &= a, & a &\in \mathbb{R} \\
\tilde{\mu}_0(a, b] &= b - a, & a, b &\in \mathbb{R}, \quad a < b \\
\tilde{\mu}_0(b, \infty) &= -b, & b &\in \mathbb{R} \\
\tilde{\mu}_0(\mathbb{R}) &= 0 \\
\tilde{\mu}_0\left(\bigsqcup_{i=1}^{n} I_i\right) &= \sum_{i=1}^{n} \tilde{\mu}_0(I_i), & &\text{if } I_1, ..., I_n \text{ are right semi-closed intervals}
\end{aligned}
$$

Then $\tilde{\mu}_0$ is finitely additive, but not countably additive on $\mathcal{F}_0$. (Why?) For a proof, see GoodNotes.
△

Measure-like set functions have useful properties. Using the notation in Table 1, we rewrite Theorem 1.2.5 of the text:

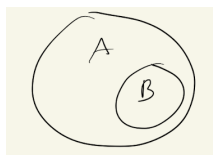**Theorem 3.3.1.** *Let $\tilde{\mu}_0$ be a finitely additive set function on the field $\mathcal{F}_0$. Then*

a) $\tilde{\mu}_0(\emptyset) = 0$

b) $\tilde{\mu}_0(A \cup B) + \tilde{\mu}_0(A \cap B) = \tilde{\mu}_0(A) + \tilde{\mu}_0(B)$ *for all* $A, B \in \mathcal{F}_0$.
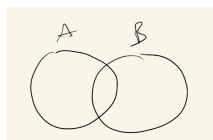
c) *If* $A, B \in \mathcal{F}_0$ *and* $B \subset A$, *then*

$$\tilde{\mu}_0(A) = \tilde{\mu}_0(B) + \tilde{\mu}_0(A - B) \quad \textit{(piece-and-difference decomposition)}$$

[8]*So* $\tilde{\mu}_0(A) \geq \tilde{\mu}_0(B)$ *if* $\tilde{\mu}_0(A - B) \geq 0$. *More generally, for non-negative set functions, we have*

$$\mu_0(A) \geq \mu_0(B) \quad \textit{(monotonicity)}$$

---

[7]This example comes from Problem 4 in Section 1.2 of the text

[8]If the "piece" satisfies $\tilde{\mu}_0(B) < \infty$, we have $\tilde{\mu}_0(A - B) = \tilde{\mu}_0(A) - \tilde{\mu}_0(B)$. One useful takeaway for piece-and-difference decompositions is that : *the finite measure of the difference is the difference of the finite measures*.

d) *Subadditivity holds if $\tilde{\mu}_0$ is non-negative, i.e.*

$$\mu_0(\cup_{i=1}^n A_i) \le \sum_{i=1}^n \mu_0(A_i)$$

$$\mu(\cup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu(A_i)$$

Proof. We prove Theorem 3.3.1 (b). The rest is an exercise for the reader (or see the text).

First, we break things into disjoint pieces

$$A = \left(A \cap B \right) \biguplus \left(A \cap B^c \right) \qquad \Longrightarrow \quad \tilde{\mu}_0(A) = \tilde{\mu}_0(A \cap B) + \tilde{\mu}_0(A \cap B^c) \qquad (1)$$

$$B = \left(A \cap B \right) \biguplus \left(A^c \cap B \right) \qquad \Longrightarrow \quad \tilde{\mu}_0(B) = \tilde{\mu}_0(A \cap B) + \tilde{\mu}_0(A^c \cap B) \qquad (2)$$

$$A \cup B = \left(A \cap B \right) \biguplus \left(A \cap B^c \right) \biguplus \left(A^c \cap B \right) \quad \Longrightarrow \quad \tilde{\mu}_0(A \cup B) = \tilde{\mu}_0(A \cap B) + \tilde{\mu}_0(A \cap B^c) + \tilde{\mu}_0(A^c \cap B) \quad (3)$$

Summing (1) and (2), we obtain

$$\tilde{\mu}_0(A) + \tilde{\mu}_0(B) = 2\tilde{\mu}_0(A \cap B) + \tilde{\mu}_0(A \cap B^c) + \tilde{\mu}_0(A^c \cap B).$$

We use (3) to simplify the RHS, and the result follows. $\qquad\qquad\square$

Remark 3.3.3. In the proof of Theorem 3.3.1 (b), note that we use a common strategy – breaking sets into disjoint pieces so that we can apply the assumed (finite or countable) additivity of the set function. $\qquad\qquad\triangle$

Remark 3.3.4. Is *finiteness* ($|\mu_g(A)| < \infty \; \forall \, A \in \mathcal{F}_g$) equivalent to *boundedness* ($\sup\{|\mu_g(A)| : A \in \mathcal{F}_g\} < \infty$)?

- $\mu_0, \tilde{\mu}$? ✓

- $\tilde{\mu}_0$? ✗ (too general)

The fact that equivalence holds for signed measures $\tilde{\mu}$ is surprising. Somehow countable additivity compensates for the signedness. See Section 2.1.3 of the text. $\qquad\qquad\triangle$

3.4 §1.2.7-1.2.8: Continuity of countably additive set functions

Countably additive set functions have a basic continuity property. Continuity of measure is a special case.

Theorem 3.4.1. *Let $\tilde{\mu}$ be a countably additive set function on the σ-field $\mathcal{F}$. Then*

a) *(continuity from below) If $A_1, A_2, ... \in \mathcal{F}$ and $A_n \uparrow A$, then $\tilde{\mu}(A_n) \to \tilde{\mu}(A)$ as $n \to \infty$.*

11
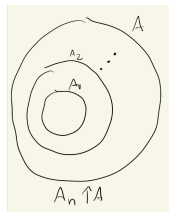
b) *(continuity from above)* If $A_1, A_2, \ldots \in \mathcal{F}$, $A_n \downarrow A$, and $\widetilde{\mu}(A_1)$ is finite, then $\widetilde{\mu}(A_n) \to \widetilde{\mu}(A)$ as $n \to \infty$.

Proof. We prove continuity from below, and leave continuity from above as an exercise to the reader (or see text).

First let us assume that all $\widetilde{\mu}(A_n)$ are finite (*). Then

$$
\begin{aligned}
A &= A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \ldots && \text{by (2.2.2)} \\
\implies \widetilde{\mu}(A) &= \widetilde{\mu}(A_1) + \widetilde{\mu}(A_2 - A_1) + \widetilde{\mu}(A_3 - A_2) + \ldots && \text{(countable additivity)} \\
&= \widetilde{\mu}(A_1) + \widetilde{\mu}(A_2) - \widetilde{\mu}(A_1) + \widetilde{\mu}(A_3) - \widetilde{\mu}(A_2) + \ldots && \text{(Theorem 3.3.1 c), (*)} \\
&= \lim_{n \to \infty} \widetilde{\mu}(A_n) && \text{(telescoping difference)}
\end{aligned}
$$

Now suppose $\widetilde{\mu}(A_n) = \infty$ for some n. So write

$$
\begin{aligned}
A &= A_n \cup A - A_n && \text{(increasing sequence)} \\
\implies \widetilde{\mu}(A) &= \widetilde{\mu}(A_n) + \widetilde{\mu}(A - A_n) && \text{(countable additivity)} \\
&= \infty + \widetilde{\mu}(A - A_n)
\end{aligned}
$$

So $\widetilde{\mu}(A) = \infty$.[9] Replace A by A_k for any $k \geq n$ to also find $\widetilde{\mu}(A_k) = \infty$ for all $k \geq n$ and the result follows.

Finally suppose $\widetilde{\mu}(A_n) = -\infty$ for some n. Then the result follows in the same way as for $\widetilde{\mu}(A_n) = \infty$.

$\square$

Remark 3.4.1. The logic of the proof of Theorem 3.4.1 under the finiteness assumption is as follows. First, we re-represent the union as a disjoint union (the form is particularly simple since the sets are increasing). This allows us to apply countable additivity. Then we apply the piece-and-difference decomposition (and the subtraction is defined under the finiteness assumption). $\triangle$

Remark 3.4.2. In proving Theorem 3.4.1 for the case where $\mu(A_n) = \infty$ for some n, it is tempting to make the simpler argument

$$
\begin{aligned}
\mu(A) &\geq \mu(A_n) && \text{(monotonicity)} \\
\mu(A_k) &\geq \mu(A_n) && \text{(monotonicity)}
\end{aligned}
$$

for $k \geq n$. But recall from Theorem 3.3.1 that monotonicity only holds under non-negativity, and the theorem statement is more general, applying to *signed* set functions as well. $\triangle$

We have the result that finite additivity plus continuity equals countable additivity.

Theorem 3.4.2. *Let $\widetilde{\mu}_0$ be a finitely additive set function on the field $\mathcal{F}_0$. Suppose either*

a) *$\widetilde{\mu}_0$ is continuous from below*

b) *$\widetilde{\mu}_0$ is continuous from above at the empty set.*

Then $\widetilde{\mu}_0$ is countably additive.

Proof. We prove that the conclusion holds under (a) and leave doing the same for (b) as an exercise to the reader (or see text).

---

[9] Note that we cannot have $\widetilde{\mu}(A - A_n) = -\infty$, because that would violate additivity.

Given $A = \bigcup_{n=1}^{\infty} A_n$, we define $P_n := \bigcup_{m \leq n} A_n$ and so $P_n \uparrow A$. So we have

$$\tilde{\mu}_0(P_n) \to \tilde{\mu}_0(A) \qquad \text{(continuity from below)}$$

$$\implies \tilde{\mu}_0\Big(\bigcup_{m \leq n} A_n \Big) \to \tilde{\mu}_0(A) \qquad \text{(definition)}$$

$$\implies \sum_{m=1}^{n} \tilde{\mu}_0(A_n) \to \tilde{\mu}_0(A) \qquad \text{(finite additivity)}$$

Taking $n \to \infty$ gives countable additivity. $\qquad\square$

4 § 1.3: Extension of measures

4.1 Extension and approximation

In Example 3.2.3, we discussed the concept of length of a subset of $\mathbb{R}$; in particular, we mentioned extending the set function given on intervals by $\mu(a, b] = b - a$ to a larger class of subsets of $\mathbb{R}$.

As remarked in Example 3.1.4, if we define $\mathcal{F}_0 = \{$finite disjoint unions of right semi-closed intervals $(a, b], a \neq b\}$, then $\mathcal{F}_0$ is a field, as can be easily verified. And μ can easily be seen to be a finitely additive set function on $\mathcal{F}_0$.

However, $\mathcal{F}_0$ is not a σ-field. So how can we extend this function to a measure on a larger class of subsets? The challenges are:

- *We need to show that μ is countably additive.* We will do this in Section 5. Moreover, in that section, we will generalize our problem to set functions given by $\mu(a, b] = F(b) - F(a)$, where F is an increasing right-continuous function from $\mathbb{R}$ to $\mathbb{R}$.

- *We need to extend μ to $\sigma(\mathcal{F}_0)$, the minimal σ-field containing $\mathcal{F}_0$.* In other words, we need to extend μ to the Borel sets. We will handle the problem in this section more generally. In this section, we will deal with the problem of extending a measure on $\mathcal{F}_0$ to a measure on $\sigma(\mathcal{F}_0)$. We do so using Carathéodory's Theorem (Theorem 4.1.3). Along the way, we will use Theorem 4.1.1 and Theorem 4.1.2 to prove Theorem 4.1.3.

Theorem 4.1.1. *(Theorem 1.3.6 [1]) A finite measure on a field $\mathcal{F}_0$ can be extended to a measure on $\sigma(\mathcal{F}_0)$.*

Proof. See pp. 12-17 of [1]. $\qquad\square$

Theorem 4.1.2. *(Monotone Class Theorem) Let $\mathcal{F}_0$ be a field of subsets of Ω and $\mathcal{C}$ be a class of subsets of Ω that is monotone (if $A_n \in \mathcal{C}$ and $A_n \uparrow A$ or $A_n \downarrow A$, then $A \in \mathcal{C}$). If $\mathcal{C} \supset \mathcal{F}_0$ then $\mathcal{C} \supset \sigma(\mathcal{F}_0)$, then minimal σ-field over $\mathcal{F}_0$.*

Proof. See pp. 18-19 of [1]. $\qquad\square$

Remark 4.1.1. During the proof of Theorem 4.1.2, some key observations are made about the relationship between monotone classes and σ-fields:

a) A monotone class that is also field is a sigma-field. (See Remark 3.1.3.)

b) The smallest monotone class and smallest sigma-field over a field coincide.

$$\triangle$$

Theorem 4.1.3. *(Carathéodory Extension Theorem) Let μ be a measure on the field $\mathcal{F}_0$ of subsets of Ω, and assume that μ is σ-finite on $\mathcal{F}_0$, so that Ω can be decomposed as $\cup_{n=1}^{\infty} A_n$ where $A_n \in \mathcal{F}_0$ and $\mu(A_n) < \infty$ for all n. Then μ has a unique extension to a measure on $\mathcal{F} := \sigma(\mathcal{F}_0)$, the minimal σ-field over $\mathcal{F}_0$.*

Proof. (We follow the argument of [1], but add some detail.) First we prove existence. [Without loss of generality, we assume the A_n are disjoint. This is possible because we can use (2.2.1) to re-express the countable union as a disjoint countable union: $\Omega = \cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$, where $B_i := A_i \cap A_{i-1}^c \ldots \cap A_1^c$.]

If we define $\mu_n(A) = \mu(A \cap A_n)$ for each $A \in \mathcal{F}_0$, then we can decompose μ into a countable sum of finite measures:

- μ_n is a measure on $\mathcal{F}_0$. [Its countable additivity is inherited from μ. If $\cup_{i=1}^{\infty} A_i$ is a disjoint union, then so is $\cup_{i=1}^{\infty} (A_i \cap A_n)$, and $\mu(\cup_{i=1}^{\infty} (A_i \cap A_n)) = \sum_{i=1}^{\infty} \mu(A_i \cap A_n)$ since $A_i \cap A_n$ are in $\mathcal{F}_0$.]

- μ_n is finite. [True because $\mu_n(A) = \mu(A \cap A_n) \overset{\text{monotonicity}}{\leq} \mu(A_n) < \infty$.]

- $\mu = \sum_{n=1}^{\infty} \mu_n$. [True because $\mu(A) = \mu(A \cap \Omega) = \mu(A \cap (\cup_{n=1}^{\infty} A_n)) = \mu(\cup_{n=1}^{\infty} (A \cap A_n)) = \sum_{n=1}^{\infty} \mu(A \cap A_n) = \mu_n(A).$]

Now by Theorem 4.1.1, we can extend each μ_n to a measure μ_n^* on $\mathcal{F}$. Thus $\mu^* := \sum_{n=1}^{\infty} \mu_n^*$ extends μ to $\mathcal{F}$. Moreover, μ^* is still a measure since the order of summation in a double series of nonnegative terms can be reversed. [Countable additivity still holds since:

$$
\begin{aligned}
\mu^*(\cup_{i=1}^{\infty} A_i) &= \sum_{n=1}^{\infty} \mu_n^*(\cup_{i=1}^{\infty} A_i) \\
&= \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} \mu_n^*(A_i) \qquad \text{\small μ_n^* is measure, so countably additive} \\
&= \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \mu_n^*(A_i) \qquad \text{\small reverse order of summation for double series with non-negative terms} \\
&= \sum_{i=1}^{\infty} \mu^*(A_i) \qquad \text{\small def. of μ^*}
\end{aligned}
$$

].

Now we prove uniqueness. That is, we prove that if λ is a measure on $\mathcal{F}$ and $\lambda = \mu^*$ on $\mathcal{F}_0$, then $\lambda = \mu^*$ on $\mathcal{F}$. To see this, as before, we decompose the measure into a sum of finite measures: $\lambda = \sum_{n=1}^{\infty} \lambda_n$ where $\lambda_n := \lambda(A_n \cap A)$. Now by assumption $\lambda_n = \mu_n^*$ on $\mathcal{F}_0$. Where are they equal on $\mathcal{F}$? Let us define the "good sets" (recall Section 3.1.3)

$$\mathcal{G} := \{A \in \mathcal{F} : \lambda_n(A) = \mu_n^*(A)\}$$

Now we can show $\mathcal{G} = \mathcal{F}$ – that is, *all* sets in the σ-field are good sets – by observing

- $\mathcal{G}$ is a monotone class. [This is true by continuity from below (see Theorem 3.4.1). In particular, a countable union can be considered the limit of an increasing sequence of partial unions (See Remark 3.1.3.) As a result, the measure of the limiting set is determined, as the limit of the the measure of the sets in that sequence.]

- $\mathcal{G} \supset \mathcal{F}_0$. [This is true by construction.]

And so by Monotone Class Theorem (Theorem 4.1.2), we have $\mathcal{G} \supset \mathcal{F}$. But by construction $\mathcal{G} \subset \mathcal{F}$, and so $\mathcal{G} = \mathcal{F}$. Therefore $\lambda_n = \mu_n^*$ for each n.

So

$$\lambda \overset{\text{decomposition}}{=} \sum_n \lambda_n = \sum_n \mu_n^* \overset{\text{recomposition}}{=} \mu^*,$$

proving uniqueness. $\qquad\square$

Remark 4.1.2. The proof of Theorem 4.1.3 reveals the appeal of σ-finite measures – they can be decomposed as the countable sum of finite measures (and the order of summation of double series can be reversed for nonnegative series, so countable additivity still holds). $\qquad\triangle$

In Remark 4.1.1 (b), we observed that minimal σ-fields over a field can be characterized as the minimal monotone classes over a field – so we merely need to close the field over increasing and decreasing sequences of sets. This idea suggests that if $\mathcal{F}_0$ is a field and $\mathcal{F} = \sigma(\mathcal{F}_0)$, sets in $\mathcal{F}$ can be approximated in some sense by sets in $\mathcal{F}_0$. The following result formalizes this notion.

Theorem 4.1.4. *(Approximation Theorem) Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $\mathcal{F} = \sigma(\mathcal{F}_0)$ where $\mathcal{F}_0$ is a field of subsets of Ω. Let μ be σ-finite on $\mathcal{F}_0$. Then for every $A \in \mathcal{F}$ and fixed $\epsilon > 0$, there is a set $B \in \mathcal{F}_0$ such that $\mu(A \triangle B) < \epsilon$.*

Example 4.1.1. This interesting example (from [1] pp. 20) provides a counterexample to the theorems when $\mathcal{F}_0$ is not σ-finite.

$\triangle$

1.3.12 Example. Let Ω be the rationals, $\mathscr{F}_0$ the field of finite disjoint unions of right-semiclosed intervals $(a, b] = \{\omega \in \Omega: a < \omega \leq b\}, a, b$ rational [counting (a, ∞) and Ω itself as right-semiclosed; see 1.2.2]. Let $\mathscr{F} = \sigma(\mathscr{F}_0)$. Then:

(a) $\mathscr{F}$ consists of all subsets of Ω.

(b) If $\mu(A)$ is the number of points in A (μ is counting measure), then μ is σ-finite on $\mathscr{F}$ but not on $\mathscr{F}_0$.

(c) There are sets $A \in \mathscr{F}$ of finite measure that cannot be approximated by sets in $\mathscr{F}_0$, that is, there is no sequence $A_n \in \mathscr{F}_0$ with $\mu(A \bigtriangleup A_n) \to 0$.

(d) If $\lambda = 2\mu$, then $\lambda = \mu$ on $\mathscr{F}_0$ but not on $\mathscr{F}$.

Thus both the approximation theorem and the Carathéodory extension theorem fail in this case.

PROOF.   (a)   We have $\{x\} = \bigcap_{n=1}^{\infty} (x - (1/n), x]$, and therefore all singletons are in $\mathscr{F}$. But then all sets are in $\mathscr{F}$ since Ω is countable.

(b)   Since Ω is a countable union of singletons, μ is σ-finite on $\mathscr{F}$. But every nonempty set in $\mathscr{F}_0$ has infinite measure, so μ is not σ-finite on $\mathscr{F}_0$.

(c)   If A is any finite nonempty subset of Ω, then $\mu(A \bigtriangleup B) = \infty$ for all nonempty $B \in \mathscr{F}_0$, because any nonempty set in $\mathscr{F}_0$ must contain infinitely many points not in A.

(d)   Since $\lambda\{x\} = 2$ and $\mu\{x\} = 1, \lambda \neq \mu$ on $\mathscr{F}$. But $\lambda(A) = \mu(A) = \infty, A \in \mathscr{F}_0$ (except for $A = \emptyset$). $\square$

4.2   Completion of measure spaces

Definition 4.2.1.   A measure μ on a σ-field $\mathcal{F}$ is said to be *complete* iff whenever $A \in F$ and $\mu(A) = 0$, we have $B \in F$ for all $B \subset A$. $\triangle$

Definition 4.2.2.   The *completion* of a measure space $(\Omega, \mathcal{F}, \mu)$ is given by $(\Omega, \mathcal{F}_\mu, \mu)$, where

$$\mathcal{F}_\mu := \{A \cup S : A \in \mathcal{F}, S \subset N \text{ for some } N \in \mathcal{F} \text{ with } \mu(N) = 0\}$$

and where μ is extended to $\mathcal{F}_\mu$ by setting $\mu(A \cup S) = \mu(A)$. $\triangle$

Remark 4.2.1.   Let us show that Definition 4.2.2 is a valid definition by showing that

1. $\mathcal{F}_\mu$ *is a σ-field.*

2. μ *is a measure on* $\mathcal{F}_\mu$.

3. *The completion is complete.*

We justify these in turn:

1. $\mathcal{F}_\mu$ is closed under countable unions, since

$$\cup_{i=1}^\infty (A_i \cup S_i) = \underbrace{(\cup_{i=1}^\infty A_i)}_{\in \mathcal{F}} \cup \underbrace{(\cup_{i=1}^\infty S_i)}_{\text{has measure 0}}$$

where the term on the right has measure 0 because $\cup_{i=1}^\infty S_i \subset \cup_{i=1}^\infty N_i \in \mathcal{F}$, and $\mu(\cup_{i=1}^\infty N_i) = \sum_{i=1}^\infty \mu(N_i) = 0$.

F_μ is also closed under complements, since $S \subset N \implies N^c \subset S^c$, and so

$$(A \cup S)^c = (A^c \cap S^c) = \underbrace{(A^c \cap N^c)}_{\in \mathcal{F}} \cup \underbrace{(A^c \cap S^c - N^c)}_{\text{has measure 0}}$$

where the term on the right has measure 0 by monotonicity, because $A^c \cap S^c - N^c \subset S^c - N^c = S^c \cap (M^c)^c = S^c \cap N \subset N$.

2. First, we show that countable additivity holds in $\mathcal{F}_\mu$.

$$\mu(\cup_{i=1}^\infty (A_i \cup S_i)) \stackrel{\text{see below}}{=} \mu(\cup_{i=1}^\infty A_i) \stackrel{\mu \text{ countably additive on } \mathcal{F}}{=} \sum_{i=1}^\infty \mu(A_i) \stackrel{\text{construction of extension}}{=} \sum_{i=1}^\infty \mu(A_i \cup S_i)$$

The first equality holds because we can re-represent a disjoint union $\cup_{i=1}^\infty (A_i \cup S_i) = (\cup_{i=1}^\infty A_i) \cup (\cup_{i=1}^\infty S_i)$. Since $\cup_{i=1}^\infty S_i \subset \underbrace{\cup_{i=1}^\infty N_i}_{\text{has measure 0 in } \mathcal{F}}$, we have that $\mu((\cup_{i=1}^\infty A_i) \cup (\cup_{i=1}^\infty S_i)) = \mu(\cup_{i=1}^\infty A_i)$.

Next, we show that μ is invariant to decompositions: if $A_1 \cup S_1 = A_2 \cup S_2$, then $\mu(A_1 \cup S_1) = \mu(A_2 \cup S_2)$, or more simply $\mu(A_1) = \mu(A_2)$.
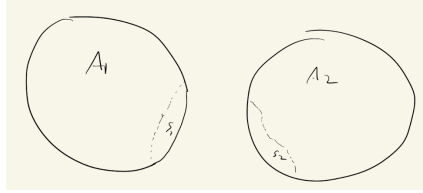
We have

$$\mu(A_1) \stackrel{\text{countable additivity}}{=} \mu(A_1 \cap A_2) + \mu(A_1 \cap A_2^c) \stackrel{\text{see below}}{=} \mu(A_1 \cap A_2) \stackrel{\text{monotonicity}}{\leq} \mu(A_2)$$

where the second equality holds since $A_1 \cap A_2^c \subset S_2$ (which, in turn, holds since $x \in A_1 \implies x \in A_2$ or $x \in S_2$, so $x \in A_1$ and $x \notin A_2 \implies x \in S_2$).

By symmetry, $\mu(A_2) \leq \mu(A_1)$, so $\mu(A_1) = \mu(A_2)$.

3. By the definition of a complete measure, we need to show that if $B \in \mathcal{F}_\mu$ and $\mu(B) = 0$ then $C \in \mathcal{F}_\mu$ for all $C \subset B$.

Now $B \in \mathcal{F}_\mu \implies B = \underbrace{A}_{\in \mathcal{F}} \cup \underbrace{S}_{\subset N \in \mathcal{F}\,:\,\mu(N) = 0}$.

So our assumption $\mu(B) = 0$ gives us $\mu(A) = 0$, since $\mu(B) = \mu(A \cup S) \stackrel{\text{choice of extension}}{=} \mu(A) = 0$.

Now since we have assumed $C \subset B$ we have

$$\mu(C) \stackrel{\text{monotonicity}}{\leq} \mu(B) \stackrel{B \in \mathcal{F}_\mu}{=} \mu(A \cup S) \stackrel{\text{subadditivity}}{\leq} \mu(A) + \mu(S) \stackrel{\text{see above}}{=} 0 + \mu(S) = 0 + 0 = 0$$

Since μ is non-negative, this implies that $\mu(C) = 0$.

We can therefore write $C = \underbrace{\emptyset}_{\in \mathcal{F}} \cup \underbrace{C}_{\text{has measure 0}}$, so $C \in \mathcal{F}_\mu$.

Thus, μ on $\mathcal{F}_\mu$ is complete, since any subset of measure 0 is contained in $\mathcal{F}_\mu$.

$\triangle$

5 § 1.4: Lesbesgue-Stieltjes Measures and Distribution Functions

Definition 5.0.1. A *Lesbesgue-Stieltjes measure* on $\mathbb{R}$ is a measure μ on $\mathcal{B}(\mathbb{R})$ such that $\mu(I) < \infty$ for each bounded interval I. $\triangle$

Definition 5.0.2. A *distribution function* on $\mathbb{R}$ is a map $F : \mathbb{R} \to \mathbb{R}$ that is increasing [$a < b$ implies $F(a) \leq F(b)$] and right continuous [$\lim_{x \to x_0^+} F(x) = F(x_0)$]. $\triangle$

Here we show that the formula $\mu(a, b] = F(b) - F(a)$ sets up a one-to-one correspondence between distribution functions and Lesbesgue-Stieltjes measures. Figure 3 helps to illuminate how working with distribution functions allows us to cover absolutely continuous, discrete, and mixed random variables in a single paradigm.
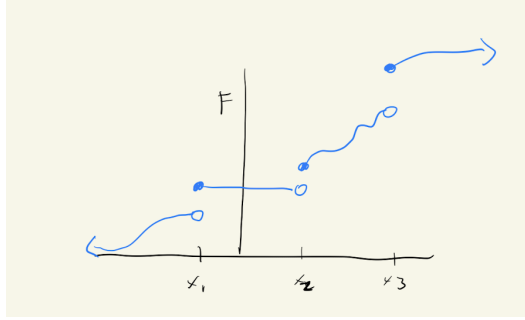

Figure 3: A distribution function with positive mass on points that is not concentrated on a countable set

5.1 § 1.4.2 Each Lesbesgue-Stietljes measure uniquely determines a distribution function (up to an additive constant)

First, we show that to every Lesbesgue-Stieltjes measure, there is a unique distribution function (up to an additive constant). This is the easy part.

Theorem 5.1.1. *Let μ be a Lesbesgue-Stietljes measure on $\mathbb{R}$. Let $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ be defined (up to additive constant) by $F(b) - F(a) = \mu(a, b]$ for $a < b$. Then F is a distribution function.*

Proof. First we show that F is increasing. We have $F(b) - F(a) = \mu(a, b] \geq 0$, since μ is non-negative.

Next we show that F is right continuous. By the continuity (from above) of measure,

$$\lim_{b' \to b+} [F(b') - F(a)] = \lim_{b' \to b+} \mu(a, b'] = \mu(a, b]$$

and continuity from above applies since Lesbesgue-Stietljes measures are finite on any interval.

Thus, rearranging,

$$\lim_{b' \to b+} F(b') = \mu(a, b] + F(a) = \left(F(b) - F(a) \right) + F(a) = F(a)$$

$\square$

5.2 § 1.4.3-1.4.4 Each distribution function (identified up to additive constant) uniquely determines a Lesbesgue-Stietljes measure

Now the harder part. We need to show that every distribution function F (identified up to additive constant) uniquely determines a Lesbesgue-Stieltjes measure.

To start, by a similar reasoning as we've seen before (e.g. see Section 4.1), it is straightforward to show that the formula $\mu(a, b] = F(a) - F(b), a, b \in \mathbb{R}, a < b$ defines a finitely additive set function on $\mathcal{F}_0(\overline{\mathbb{R}})$, the field of disjoint unions of right semi-closed intervals of the extended reals.

The challenge will be to show that this set function is countably additive. If we can do that, then we can apply Carathéodory's Extension Theorem to extend this function to $\mathcal{B}(\mathbb{R})$.

Lemma 5.2.1. The set function μ is countably additive on $\mathcal{F}_0(\overline{\mathbb{R}})$.   △

Proof. We assume $F(\infty) - F(-\infty) < \infty$, so that μ is finite. (We leave the case where $F(\infty) - F(-\infty) = \infty$ to the reader, or see the text.) Our strategy will be to show that μ is continuous from above, in which case we can apply Theorem 3.4.2 (b) to show that the set function is countably additive.

Let A_n be a sequence of sets in $\mathcal{F}_0(\overline{\mathbb{R}})$ such that $A_n \downarrow \emptyset$. Now each A_n is a finite union of disjoint r.s.c. intervals.
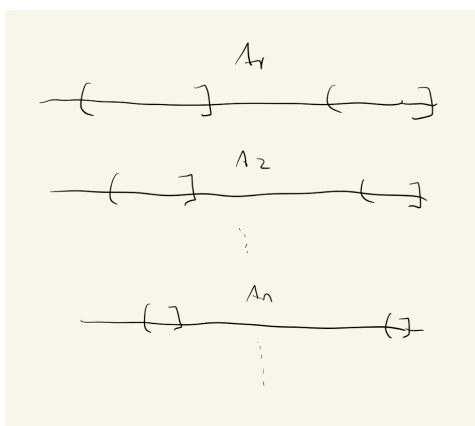
Suppose one such interval is $(a, b]$. By the right continuity of F, we can find intervals $(a', b]$ that approximate $(a, b]$ from the inside arbitrarily well, since

$$\mu(a', b] = F(b) - F(a') \to \mu(a, b] = F(b) - F(a) \text{ as } a' \to a \text{ from the right}$$

Thus, we can find sets $B_n \in \mathcal{F}_0(\overline{\mathbb{R}})$ whose closures $\overline{B}_n \subset A_n$ and where $\mu(B_n)$ approximates $\mu(A_n)$ to any desired $\epsilon > 0$.
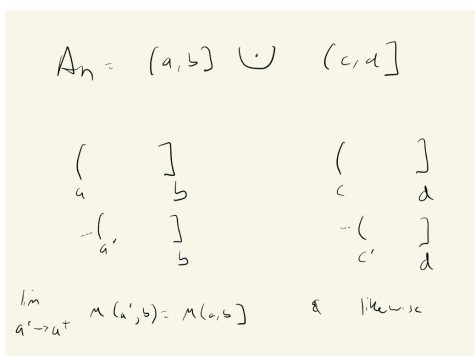
For any fixed $\epsilon > 0$, we can choose such a sequence B_n such that $\mu(A_n) - \mu(B_n) < \epsilon 2^{-n}$ $(*)$, and we find

a) $\cap_{n=1}^{\infty} \overline{B}_n = \emptyset$.   [True because each $\overline{B}_n \subset A_n$, so $\cap_{n=1}^{\infty} \overline{B}_n \subset \cap_{n=1}^{\infty} A_n = \emptyset$.]

19

b) $\cap_{k=1}^{n}\overline{B}_k = \emptyset$ for sufficiently large n. [We have $\overline{\mathbb{R}} \overset{\text{item a)}}{=} (\overline{\mathbb{R}} - \cap_{n=1}^{\infty}\overline{B}_n) \overset{\text{DeMorgan (A.2.1)}}{=} \cup_{n=1}^{\infty}(\overline{\mathbb{R}} - \overline{B}_n)$. So $\{\overline{\mathbb{R}} - \overline{B}_n\}$ is an open cover of $\overline{\mathbb{R}}$. By the Heine-Borel theorem, there must be a finite subcover. So for sufficiently large n, we have $\cup_{k=1}^{n}(\overline{\mathbb{R}} - \overline{B}_k) = \overline{\mathbb{R}}$. Taking complements of both sides, and once again applying DeMorgan's law (A.2.1) to the relative complement, we find $\cap_{k=1}^{n}\overline{B}_k = \emptyset$.]

So now we use a piece-and-difference decomposition (Theorem 3.3.1 (b)):

$$A_n = \left(\bigcap_{k=1}^{n} B_k\right) \biguplus \left(A_n - \bigcap_{k=1}^{n} B_k\right)$$

$$\implies \mu(A_n) = \underbrace{\mu(\cap_{k=1}^{n}B_k)}_{0} + \mu(A_n - \cap_{k=1}^{n}B_k) \qquad \text{countable additivity; item b) above}$$

$$\overset{1}{\leq} \mu(\cup_{k=1}^{n}(A_k - B_k)) \qquad \text{monotonicity}$$

$$\leq \sum_{k=1}^{n} \mu(A_k - B_k) \qquad \text{finite subadditivity}$$

$$= \sum_{k=1}^{n} \mu(A_k) - \mu(B_k) \qquad \text{piece-and-difference decomposition; also uses finiteness}$$

$$\leq \epsilon \sum_{k=1}^{n} 2^{-k}$$

$$< \epsilon.$$

where the monotonicity property in (1) applies because $A_n - \cap_{k=1}^{n}B_k \overset{\text{DeMorgan}}{=} \cup_{k=1}^{n}(A_n - B_k) \subset \cup_{k=1}^{n}(A_k - B_k)$.

In summary, we have seen that for any fixed $\epsilon > 0$, we have $\mu(A_n) < \epsilon$ for sufficiently large n. Thus, $\mu(A_n) \to 0$, and so μ is continuous from below. So by Theorem 3.4.2 (b), μ is countably additive. $\square$

Remark 5.2.1. The proof of Lemma 5.2.1 is a very cool application of Heine-Borel! In trying to show continuity from above, we started out with an *infinite* intersection of sets. But in showing that the measure of the sequence converged, we needed to work with *finite* collection so that we could apply *finite* subadditivity, since that's all we had to use, by assumption. $\triangle$

Theorem 5.2.1. *Let F be a distribution function on $\mathbb{R}$, and let $\mu(a, b] = F(b) - F(a), a < b$. Then there is a unique extension of μ to a Lesbesgue-Stietljes measure on $\mathbb{R}$.*

Proof. See text. $\square$

Remark 5.2.2. The proof of Theorem 5.2.1 essentially directly applies Caratheódory's Extension Theorem, since we know from Lemma 5.2.1 that μ is countably additive on $\mathcal{F}_0(\mathbb{R})$, a field from which the Borel sets are generated. The only real additional work is a tedious technical detail to identify a μ-preserving correspondence between sets in $\mathcal{F}_0(\overline{(\mathbb{R})})$ (over which we proved countable additivity) and sets in $\mathcal{F}_0(\mathbb{R})$ (which is the field we actually want to extend). $\triangle$

5.3 § 1.4.5 Properties of Lesbesgue-Stietljes measure

Before extension, we had $\mu(a, b] = F(b) - F(a)$ for $a < b$ where F is a distribution function. The set function μ was defined only on $\mathcal{F}_0(\mathbb{R})$, the field of disjoint unions of r.s.c interval. But after extension, μ is defined on $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{F}_0)(\mathbb{R}))$, which allows us to measure other types of intervals as well (by expressing those intervals as countable unions or intersections of r.s.c intervals; recall (3.2.3).)

Proposition 5.3.1. *Let μ be a Lesbesgue-Stieltjes measure, and let F be its associated distribution function. Let $F(x^-) = \lim_{y\to x^-} F(y)$. Then*

a) $\mu(a, b] = F(b) - F(a)$

b) $\mu(a, b) = F(b^-) - F(a)$

c) $\mu[a, b] = F(b) - F(a^-)$

d) $\mu[a, b) = F(b^-) - F(a^-)$

e) $\mu\{x\} = F(x) - F(x^-)$

f) $\mu(-\infty, x] = F(x) - F(-\infty)$

g) $\mu(-\infty, x) = F(x^-) - F(-\infty)$

h) $\mu(x, \infty) = F(\infty) - F(x)$

i) $\mu[x, \infty) = F(\infty) - F(x^-)$

j) $\mu(\mathbb{R}) = F(\infty) - F(-\infty)$

Proof. We prove some of these statements and leave the rest to the reader.

For (b), note that $(a, b) = \cup_{n=1}^{\infty}(a, b - \frac{1}{n}]$. So let $A_n = (a, b - \frac{1}{n}]$. Then by continuity from below,

$$\mu(a, b) = \lim_{n \to \infty} \mu(A_n) = \lim_{n \to \infty} \left[F(b - \frac{1}{n}) - F(a)\right] = F(b^-) - F(a)$$

For (c), note that $[a, b] = \cap_{n=1}^{\infty}(a - \frac{1}{n}, b]$. So by continuity from above (which applies since the sets in the intersection have finite measure),

$$\mu(a, b] = \lim_{n \to \infty} \left[F(b) - F(a - \frac{1}{n})\right] = F(b) - F(a^-)$$

For (e), note that $\{x\} = \cap_{n=1}^{\infty}(x - \frac{1}{n}, x]$. So the statement follows by the same argument as used in (c).

For (i), we can write $[x, \infty) = \cup_{n=1}^{\infty}[x, x + n)$. So by continuity from below,

$$\mu[x, \infty) = \lim_{n \to \infty} \mu[x, x + n) \overset{(d)}{=} \lim_{n \to \infty} \left[F\big((x + n)^-\big) - F(x^-)\right] = F(\infty) - F(x^-)$$

For (j), we can write $\mathbb{R} = \cup_{n=1}^{\infty}[-n, n]$. So by continuity from below,

$$\mu(\mathbb{R}) = \lim_{n \to \infty} \mu[-n, n] \overset{(c)}{=} \lim_{n \to \infty} \left[F(n) - F(-n)\right] = F(\infty) - F(-\infty)$$

$\square$

Remark 5.3.1. (*Continuity at a point iffi measure zero at a point*)

1. Note that

$$\mu\{x\} = 0 \quad \Leftrightarrow \quad \text{F is continuous at } x \tag{5.3.1}$$

   which holds by Proposition 5.3.1 part e) and the fact that F is already right-continuous by definition.

2. The magnitude of the discontinuity corresponds with the measure of $\{x\}$.

$\triangle$

Remark 5.3.2. The characterization of continuity in Remark 5.3.1 in terms of measure zero can be an interesting way to prove continuity, or prove the existence of functions with interesting properties. For instance, take a countable set $S = \{x_1, x_2, ...\}$ and non-negative weights $\{w_1, w_2, ...\}$. such that $\sum_i w_i < \infty$. Then define $\mu(A) = \sum_i \{w_i : x_i \in A\}$. Now μ is a Lesbesgue-Stietljes measure (and is in fact a finite measure), since $\mu(I) < \infty$ for each bounded interval I. By taking S to be the rationals, we have proven the existence of an increasing function $F : \mathbb{R} \to \mathbb{R}$ that is continuous on the irrationals and discontinuous on the rationals [since each Lesbesgue-Stieltjes measure determines a distribution function F (up to additive constant), and the set of continuities is given by (5.3.1)]. $\triangle$

References

[1] Robert B Ash, B Robert, Catherine A Doleans-Dade, and A Catherine. *Probability and measure theory*. Academic Press, 2000.

[2] Jeffrey S Rosenthal. *First Look At Rigorous Probability Theory, A*. World Scientific Publishing Company, 2006.

A Appendix

A.1 Right semi-closed intervals

Definition A.1.1. A **right semi-closed interval** is a set of the form $(a, b] = \{x : a < x \leq b\}, -\infty \leq a < b < \infty$. By convention, we also count (a, ∞) as right semi-closed for $-\infty \leq a < \infty$. $\triangle$

A.2 DeMorgan's Law applies to relative complements

Remark A.2.1. DeMorgan's Law also holds for relative complements. That is, given a sequence of sets $A_1, A_2, ...$ that are subsets of another set X, we have:

$$X - \bigcap_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (X - A_n) \tag{A.2.1}$$

$\triangle$