

Notes on Probability and Measure Theory

August 3, 2021

Contents

1	Overview	2
1.1	References	2
1.2	Motivation	2
2	§ 1.1: Some notes on set theory	2
3	§ 1.2: Fields, σ-fields, measures	3
3.1	§ 1.2.1-1.2.2: Fields and σ -fields	3
3.1.1	“Good sets” strategy	4
3.2	§ 1.2.3-1.2.4: Measures	4
3.3	§ 1.2.5-1.2.6: Measure-like set functions, and their properties	6
3.4	§ 1.2.7-1.2.8: Continuity of countably additive set functions	7
4	§ 1.3: Extension of measures	9
A	Definitions	9

1 Overview

1.1 References

The primary reference here is [1]. 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.

It is convenient in unifying various kinds of random variables. In particular, 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.

Finally, measure theory is important to have a deeper level of understanding of probabilistic concepts. 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 the uniform distribution. 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. $P([a, b]) = b - a, \quad 0 \leq a \leq b \leq 1.$
2. $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ for A_1, A_2, \dots disjoint subsets of $[0, 1]$.
3. $P(A \oplus r) = P(A), \quad 0 \leq r \leq 1,$ where $A \oplus r$ denotes the r -shift of A , i.e.
$$A \oplus r := \{a + r : a \in A, a + r \leq 1\} \cup \{a + r - 1 : a \in A, a + r > 1\}$$

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

2 § 1.1: Some notes on set theory

Definition 2.0.1. If $A_1 \subset A_2 \subset \dots$ and $\bigcup_{n=1}^{\infty} A_n = A$, 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 $A_1 \supset A_2 \supset \dots$ and $\bigcap_{n=1}^{\infty} A_n = A$, 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

Now some remarks on representing unions as disjoint unions.

Remark 2.0.1. If A_1, A_2, \dots are subsets of some set Ω , then

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \left(A_n \cap A_{n-1}^c \cap \dots \cap A_1^c \right) \quad (2.0.1)$$

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. \triangle

Remark 2.0.2. If $A_n \uparrow A$, then (2.0.1) becomes

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \left(A_n - A_{n-1} \right) \quad (2.0.2)$$

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

3 § 1.2: Fields, σ -fields, measures

3.1 § 1.2.1-1.2.2: Fields and σ -fields

We begin with a discussion of σ -fields, which are the domains of probability measures, and measures more generally. As stated in the motivation (Section 1.2), measures cannot be defined on all subsets of many spaces that we would like to deal with.

σ -fields are important because they are the domain of measures. Fields are important because measures are sometimes constructed by being defined on fields, and then extended to σ -fields.

Here are some definitions.

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

- a) $\Omega \in \mathcal{F}$
- b) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
- c) If $A_1, \dots, A_n \in \mathcal{F}$ 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. \triangle

Definition 3.1.2. 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 Definition 3.1.1 after replacing condition c) with

- c') If $A_1, A_2, \dots \in \mathcal{F}$ 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. \triangle

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

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

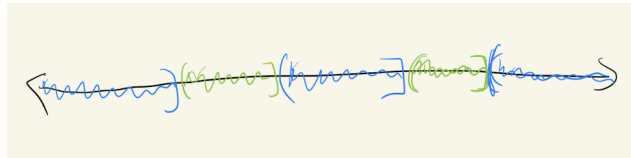
Example 3.1.3. If $A \in \Omega$ is non-empty, then $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ is the smallest σ -field containing A . \triangle

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 \mathcal{C}* or the *σ -field generated by \mathcal{C}* . \triangle

Question 3.1.1. Let A_1, \dots, A_n be subsets of Ω . Describe $\mathcal{F} := \sigma(\{A_1, \dots, A_n\})$, the smallest σ -field containing A_1, \dots, 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 $|\mathcal{F}| \leq 2^{2^n}$. For a complete answer, see GoodNotes.* \triangle

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 $\mathcal{F}_0 = \{\text{finite disjoint unions of right semi-closed intervals } (a, b], a \neq b\}$. Then \mathcal{F}_0 is a field, as can be easily verified.¹



But \mathcal{F}_0 is not a σ -field. Note that if $A_n = (-\frac{1}{n}, 0]$, then $\bigcap_{n=1}^{\infty} A_n = \{0\} \notin \mathcal{F}_0$. \triangle

Remark 3.1.1. 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 $B_N \uparrow A$. \triangle

¹By 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.

3.1.1 “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.

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

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

For another application, see handwritten homework exercises.

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, \dots form a finite or countably infinite collection of disjoint sets in \mathcal{F} , we have countable additivity; that is,

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

△

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

△

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

△

Example 3.2.1. Let $\Omega = \{x_1, x_2, \dots\}$ be a finite or countably infinite set. Let p_1, p_2, \dots 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.

△

Example 3.2.2. (*Lebesgue 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, ∞) , $a \in \mathbb{R}$.
- all intervals $[a, \infty)$, $a \in \mathbb{R}$.
- all intervals $(-\infty, b)$, $b \in \mathbb{R}$.
- all intervals $(-\infty, b]$, $b \in \mathbb{R}$.
- all intervals $[a, b]$, $a, b \in \mathbb{R}$.
- all open sets of \mathbb{R} .²
- all closed sets of \mathbb{R} .³

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

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

and

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

△

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? △

²Recall that an open set is a countable union of open intervals.

³Recall that a set is open iff its complement is closed.

3.3 § 1.2.5-1.2.6: Measure-like set functions, and their properties

The text considers some generalizations of measures that can be obtained by restricting the domain to a field, by assuming merely finite additivity, or 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. The first two relaxations above often go together. However, 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. In my notes, I will simply things by assuming that countably additive functions are always defined on σ -fields. \triangle

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

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 relations⁴

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

Example 3.3.1. Let \mathcal{F}_0 be the field of finite disjoint unions of right semi-closed intervals (see Definition A.0.1), and define the set function $\tilde{\mu}_0$ on \mathcal{F}_0 as follows⁵:

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

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

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

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

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

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

⁴So, 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.

⁵This example comes from Problem 4 in Section 1.2 of the text

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

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

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

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

$$A = (A \cap B) \cup (A \cap B^c) \implies \tilde{\mu}_0(A) = \tilde{\mu}_0(A \cap B) + \tilde{\mu}_0(A \cap B^c) \quad (1)$$

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

$$A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B) \implies \tilde{\mu}_0(A \cup B) = \tilde{\mu}_0(A \cap B) + \tilde{\mu}_0(A \cap B^c) + \tilde{\mu}_0(A^c \cap B) \quad (3)$$

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. \square

Remark 3.3.2. 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. \triangle

Remark 3.3.3. Is *finiteness* ($|\mu_g(A)| < \infty \forall A \in \mathcal{F}_g$) equivalent to *boundedness* ($\sup\{|\mu_g(A)| : A \in \mathcal{F}_g\} < \infty$)?

- $\mu_0, \tilde{\mu}$? \checkmark
- $\tilde{\mu}_0$? \times (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. \triangle

Remark 3.3.4. 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 Lebesgue measure on the Borel sets. \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 $A_1, A_2, \dots \in \mathcal{F}$ and $A_n \uparrow A$, then $\tilde{\mu}(A_n) \rightarrow \tilde{\mu}(A)$ as $n \rightarrow \infty$.
- b) (continuity from above) If $A_1, A_2, \dots \in \mathcal{F}$, $A_n \downarrow A$, and $\tilde{\mu}(A_n)$ is finite, then $\tilde{\mu}(A_n) \rightarrow \tilde{\mu}(A)$ as $n \rightarrow \infty$.

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 $\tilde{\mu}(A_n)$ are finite (*). Then

$$\begin{aligned}
A &= A_1 \cup A_2 - A_1 \cup A_3 - A_2 \cup \dots && \text{by (2.0.2)} \\
\implies \tilde{\mu}(A) &= \tilde{\mu}(A_1) + \tilde{\mu}(A_2 - A_1) + \tilde{\mu}(A_3 - A_2) + \dots && \text{(countable additivity)} \\
&= \tilde{\mu}(A_1) + \tilde{\mu}(A_2) - \tilde{\mu}(A_1) + \tilde{\mu}(A_3) - \tilde{\mu}(A_2) + \dots && \text{(Theorem 3.3.1 c), (*)} \\
&= \lim_{n \rightarrow \infty} \tilde{\mu}(A_n)
\end{aligned}$$

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

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

So $\tilde{\mu}(A) = \infty$.⁶ Replace A by A_k for any $k \geq n$ to also find $\tilde{\mu}(A_k) = \infty$ for all $k \geq n$ and the result follows.

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

□

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 $\tilde{\mu}_0$ be a finitely additive set function on the field \mathcal{F}_0 . Then $\tilde{\mu}_0$ is countably additive if either

- a) $\tilde{\mu}_0$ is continuous from below
- b) $\tilde{\mu}_0$ is continuous from above at the empty set.

Proof. We prove (a) and leave (b) as an exercise to the reader (or see text).

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

$$\begin{aligned}
&\tilde{\mu}_0(P_n) \rightarrow \tilde{\mu}_0(A) && \text{(continuity from below)} \\
\implies \tilde{\mu}_0\left(\bigcup_{m \leq n} A_m\right) &\rightarrow \tilde{\mu}_0(A) && \text{(definition)} \\
\implies \sum_{m=1}^n \tilde{\mu}_0(A_m) &\rightarrow \tilde{\mu}_0(A) && \text{(finite additivity)}
\end{aligned}$$

Taking $n \rightarrow \infty$ gives countable additivity. □

⁶Note that we cannot have $\tilde{\mu}(A - A_n) = -\infty$, because that would violate additivity.

4 § 1.3: Extension of measures

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 Definitions

Definition A.0.1. 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