STAT 205A/MATH 218A: Probability Theory

Professor Shirshendu Ganguly
Notes by Albert Zhang

Fall 2020

Contents

1 Measure Theory

To develop probability at a higher level, measure theory is required to lay the foundations. Indeed, much of probability theory is just measure theory restricted to the case where the measure of the entire space is equal to 1. We start with basic measure theory constructions, and then move on to develop concepts such as random variables and their expectation in this context.

1.1 The Problem of Measure

We need to have some notion of measure that is well-defined and satisfies some desired "axioms" of how volume should behave. For example, perhaps we would like to have translation invariance:

$$\mu(E) = \mu(E + x).$$

Another desirable axiom could be disjoint countable additivity. That is, if we have disjoint sets $(E_n)_{n=1}^{\infty}$, with $E = \cup_n E_n$, then

$$\mu(E) = \sum_{n=1+}^{\infty} \mu(E_n).$$

But it turns out that there exists sets for which a measure would be "hard" to define given these axioms. We go over the construction of such a set known as the Vitali set. In particular, consider the base set $I = [-1, 2]$. Define an equivalence relation on I given by

$$x \sim y \iff x - y \in \mathbb{Q}.$$

Consider the restriction of the classes in $[0, 1]$. Note that the size of each class is countable. For each equivalence class B, pick $x_B \in B \cap [0, 1]$ (check that there always exists an x to be picked, note that this has to do with axiom of choice). Now consider

$$E = \{x_B : B \text{ is an equivalence class}\}.$$

Can we determine the size of E? To answer this, we first represent $[-1, 2]$ in terms of E. Note that

$$[0, 1] \subseteq \bigcup_{q \in [-1,1] \cap \mathbb{Q}} (E + q) \subseteq [-1, 2]$$

Since the $(E + q)$ disjoint for each $q \in [-1, 1] \cap \mathbb{Q}$, then we'd have

$$\infty \cdot \mu(E) = \sum_{q \in [-1,1] \cap \mathbb{Q}} \mu(E + q) = \mu\left(\bigcup_{q \in [-1,1] \cap \mathbb{Q}} (E + q)\right) \leq \mu([-1, 2]) < \infty.$$

This implies $\mu(E) = 0$, but then

$$0 = \sum_{q \in [-1,1] \cap \mathbb{Q}} \mu(E + q) = \mu\left(\bigcup_{q \in [-1,1] \cap \mathbb{Q}} (E + q)\right) \geq \mu([0, 1]) > 0,$$

clearly a contradiction. It follows that we cannot define a measure on every set that satisfies both translational invariance and disjoint countable additivity as axioms!

1.2 Probability Spaces

Recall from measure theory that a *measure space* is a triple

$$(\Omega, \mathcal{F}, \mu),$$

where Ω is the base set, $\mathcal{F}$ is the σ-algebra, and $\mu : \mathcal{F} \to [0, \infty]$ is our measure. A *probability measure* $\mu = \mathbb{P}$ is just the special case where $\mathbb{P}(\Omega) = 1$. The σ-algebra $\mathcal{F}$ by definition is a collection of subsets of Ω that satisfies

(i) $\Omega \in \mathcal{F}$.
(ii) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
(iii) If $(A_n) \subset \mathcal{F}$ is a countable sequence of sets, then $\cup_n A_n \in \mathcal{F}$.

Note that (ii) and (iii) imply (i).

Example 1.1. Verify that each of the following are valid σ-algebras.

- $\mathcal{F} = \{\varnothing, \Omega\}$
- $\mathcal{F} = 2^\Omega$
- $\mathcal{F} = \{A \subset \Omega : A \text{ is countable or co-countable}\}$

For the definition of a measure, we will only assume the axioms of nonnegativity and countable disjoint additivity. In particular, $\mu : \mathcal{F} \to [0, \infty]$ is a measure if it satisfies

(i) $\mu(A) \geq \mu(\varnothing) = 0$ for every $A \in \mathcal{F}$.
(ii) If A_n is a countable sequence of disjoint sets and $A = \sqcup_n A_n$, then

$$\mu(A) = \sum_n \mu(A_n).$$

Note: here and henceforth we may use $\sqcup_n A_n$ to denote disjoint union.

Example 1.2. Let $\Omega = \{1, 2, \ldots, n\}$ and $\mathcal{F} = 2^\Omega$. Then all possible ways to define a measure on $(\Omega, \mathcal{F})$ can be obtained by assigning a measure to each singleton. That is,

$$\mu(\{k\}) = p_k$$

for $k \in \Omega$.

Surprisingly, many familiar properties follow from this set of axioms alone.

Theorem 1.3 (Properties of Measure)

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space.

(i) *monotonicity.* If $A \subset B$, then $\mu(A) \leq \mu(B)$.
(ii) *subadditivity.* If $A = \cup_n A_n$, then $\mu(A) \leq \sum_n \mu(A_n)$.
(iii) *continuity from below.* If $A_n \uparrow A$, then $\mu(A_n) \uparrow \mu(A)$.
(iv) *continuity from above.* If $A_n \downarrow A$ and $\mu(A_k) < \infty$ for some finite k, then $\mu(A_n) \downarrow \mu(A)$.

Proof. Monotonicity. Since A and $B \setminus A$ are disjoint, we have

$$\mu(A) \leq \mu(A) + \mu(B \setminus A) = \mu(B).$$

Subadditivity. Disjointify the sets as follows:

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 \setminus A_1 \\ &\vdots \\ B_n &= A_n \setminus (\cup_{k=1}^{n-1} A_k) \end{aligned}$$

Then $A = \cup_n B_n$, so we have

$$\mu(A) = \mu(\cup_n B_n) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Continuity from below. Disjointify the set using (B_n) as from before. We have $A = \cup_n B_n$, and so

$$\mu(A) \leq \sum_n \mu(B_n). \tag{1}$$

By monotonicity,

$$\mu(A_1) \leq \mu(A_2) \leq \cdots \leq \mu(A_n) \leq \cdots \leq \mu(\cup_n A_n),$$

so $\mu(A_n)$ converges and is bounded by $\mu(A)$. Since $\mu(A_n) = \sum_{k=1}^{n} \mu(B_k)$, it follows that $\mu(A_n)$ converges to $\sum_n \mu(B_n)$. So (1) becomes

$$\mu(A) \leq \lim_n \mu(A_n).$$

But we also have that

$$\lim_n \mu(A_n) \leq \mu(A),$$

so we deduce that $\mu(A_n) \uparrow \mu(A)$.
Continuity from above. The proof follows from (iii) by taking complement (*exercise:* check that the assumption $\mu(A_k) < \infty$ is necessary by coming up with an example where the property fails without it). □

We now talk about some important σ-algebras. For any $\mathcal{A} \subset 2^{\Omega}$, we define the σ-algebra *generated* by $\mathcal{A}$ as the smallest σ-algebra which contains $\mathcal{A}$. Explicitly,

$$\sigma(\mathcal{A}) := \bigcap_{\mathcal{F}:\mathcal{F} \text{ is a } \sigma\text{-algebra containing } \mathcal{A}} \mathcal{F}.$$

Exercise 1.4. Check that this definition is well-defined, and that $\sigma(\mathcal{A})$ is a σ-algebra.

A common tactic in proving that two σ-algebras are the same is through their generating sets. In particular, suppose $X, Y \subset 2^{\Omega}$, and we want to show that

$$\sigma(X) = \sigma(Y).$$

Then it suffices to show that

$$X \subseteq \sigma(Y), \quad Y \subseteq \sigma(X).$$

The *Borel* σ-algebra is one generated by the collection of open sets in any topological space. The most common case is in $\Omega = \mathbb{R}$, and we denote $\mathcal{B} = \mathcal{B}(\mathbb{R})$ the Borel σ-algebra of $\mathbb{R}$ under the usual topology.

Exercise 1.5. Show that the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is generated by the collection of open intervals. Likewise show that it can be alternatively generated by the collection of closed intervals (hint: $\mathbb{Q}$ is dense in $\mathbb{R}$).

1.2.1 Constructing the Uniform Measure

We will now attempt to define the uniform measure on $\mathcal{B}(\mathbb{R})$. Before we begin, a rough outline:

1. First define it on "intervals" for a "semi-algebra".
2. Extend the definition to an "algebra".
3. Extend to a σ-algebra, namely $\mathcal{B}(\mathbb{R})$.

A *semi-algebra* $\mathcal{S}$ has the following properties:

- If $S_1, S_2 \in \mathcal{S}$, then $S_1 \cap S_2 \in \mathcal{S}$.
- If $S \in \mathcal{S}$, then S^c is a finite disjoint union of sets in $\mathcal{S}$.

It's not too hard to check that the collection of all clopen intervals forms a semi-algebra, i.e.

$$\mathcal{S} = \{(a, b] : a, b \in \mathbb{R}\}.$$

Now let $\mu((a, b]) := b - a$. We claim that μ is countably additive. Let $A = \cup_n B_n$ be a disjoint union, where all sets are in $\mathcal{S}$. Clearly

$$\mu(A) \geq \sum_n \mu(B_n).$$

For the other direction, we use compactness. Since

$$A' = \left[a + \frac{1}{m}, b\right] \subset A$$

is compact, and

$$B_n = (c_n, d_n] \subset (c_n, d_n + \epsilon/2^n) = B_n',$$

we have that

$$A' \subset \cup_{j=1}^k B_{n_j}'.$$

Then we have a finite subcovering, so we may write

$$b - \left(a + \frac{1}{m}\right) \leq \sum_{j=1}^{k} [d'_{n_j} - c'_{n_j}] \leq \epsilon + \sum_{j=1}^{k} [d_{n_j} - c_{n_j}]$$

Now, since ϵ and m are arbitrary, we have

$$\mu(A) \leq \sum_n \mu(B_n).$$

We continue by extending our measure μ from a semi-algebra to an algebra. An *algebra* $\mathcal{A}$ satisfies

- $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$.
- $A_1, A_2 \in \mathcal{A}$ implies $A_1 \cap A_2 \in \mathcal{A}$.

We claim that the collection of all finite disjoint unions of clopen intervals,

$$\mathcal{A} = \{\sqcup_{i=1}^n A_i : A_i \in \mathcal{S}\}$$

is an algebra. For finite intersections, note that

$$(\cup_{i=1}^n A_i) \cap \left(\cup_{j=1}^m B_j\right) = \sqcup_{i,j}(A_i \cap B_j)$$

is a finite disjoint union of clopen intervals, and thus belongs to $\mathcal{A}$.

For complements, we wish to show $\cap_{i=1}^n A_i^c$ belongs to $\mathcal{A}$. Since each $A_i \in \mathcal{S}$, we know that A_i^c is a disjoint union of clopen intervals, and thus belongs to $\mathcal{A}$. Since we've shown that finite intersections belong to $\mathcal{A}$, we are done.

We now must extend our measure μ to $\mathcal{A}$. By a *measure on an algebra*, we mean a set function μ which satisfies

(i) $\mu(A) \geq \mu(\varnothing) = 0$ for all $A \in \mathcal{A}$.

(ii) If $A_n \in \mathcal{A}$ and $\sqcup_{n=1}^\infty A_n = A \in \mathcal{A}$, then

$$\mu(\sqcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n).$$

The first thing we need to check that the measures assigned to sets are well-defined. Suppose

$$A = \cup_{i=1}^k B_i = \cup_{j=1}^l B_j,$$

where $B_i, C_j \in \mathcal{S}$. We wish to prove

$$\mu(A) = \sum_{i=1}^k \mu(B_i) \stackrel{?}{=} \sum_{j=1}^l \mu(C_j).$$

Note that

$$B_i = \sqcup_{j=1}^l [B_i \cap C_j],$$

and similarly,

$$C_j = \sqcup_{i=1}^k [C_j \cap B_i].$$

Then by countable additivity of the measure defined on $\mathcal{S}$, we have

$$\sum_{i=1}^k \mu(B_i) = \sum_{i=1}^k \sum_{j=1}^l \mu(B_i \cap C_j) = \sum_{j=1}^k \mu(C_j),$$

verifying that μ is well-defined.

We will next prove that μ is countably additive on $\mathcal{A}$. Note that finite additivity follows immediately by decomposing each set into a finite disjoint union as in the definition of $\mathcal{A}$. Then, this immediately implies monotonicity, which we'll need to show countable additivity.

Let (A_i) be a sequence in $\mathcal{A}$ such that $\sqcup_i A_i = A \in \mathcal{A}$. By monotonicity, we have

$$\mu(A) \geq \sum_{i=1}^n \mu(A_i).$$

Taking the limit $n \to \infty$, we deduce that

$$\mu(A) \geq \sum_{i=1}^\infty \mu(A_i).$$

To prove the other direction, write

$$A = \cup_{j=1}^k C_j, \quad C_j \in \mathcal{S}$$
$$A_i = \cup_{l=1}^{m_i} C_l^{(i)}, \quad C_l^{(i)} \in \mathcal{S}.$$

Then it suffices to show, for each j,

$$\mu(C_j) \leq \sum_{i=1}^\infty \mu(C_j \cap A_i).$$

since we can just use finite disjoint additivity while summing over each j. We may write

$$C_j \cap A_i = \cup_{l=1}^{m_i} [C_j \cap C_l^{(i)}] = \cup_{i=1}^\infty C_j \cap A_i.$$

The rest of the proof is just an exercise in set manipulation, and left as an exercise.

Now, we've extended μ from $\mathcal{S}$ to $\mathcal{A}$. Our next step is to extend μ to $\sigma(\mathcal{A})$ using Caratheodory's extension theorem, whose proof is left as optional reading in the next section.

1.2.2 Caratheodory's Extension Theorem

We say a measure μ on space $(\Omega, \mathcal{F})$ is *σ-finite* if there exists a countable covering of Ω by finite measure sets in $\mathcal{F}$. We will need this condition to prove uniqueness of extension.

Example 1.6. Consider the semi-algebra $\mathcal{S} = \{(a, b] \cap \mathbb{Q} : a, b \in \mathbb{R}\}$ on the space $\Omega = \mathbb{Q}$. Then $\sigma(\mathcal{S}) = 2^{\mathbb{Q}}$. Note that the cardinality of each element of $\mathcal{S}$ is either ∞ or 0, so we can define a measure μ on $\mathcal{S}$ by

$$\mu(A) = \begin{cases} \infty & \text{if } |A| = \infty \\ 0 & \text{o.w.} \end{cases}$$

Now, we want to construct two distinct extensions $\mu_1 \neq \mu_2$ that agree with μ on $\mathcal{S}$. One possible example is

- μ_1 is the cardinality of a set (counting measure).
- $\mu_2 = 2\mu_1$.

Note now that μ is not σ-finite.

Theorem 1.7

Given a countably additive measure on an algebra $\mathcal{A}$, it can be extended to a measure on $\sigma(\mathcal{A})$. If μ is σ-finite on $\mathcal{A}$, then the extension is unique.

Proof of Uniqueness. We use the good set principle. This is a general strategy for showing some desired property is true for a sigma algebra. We show that the class of sets satisfying the property is closed under σ-algebra operations. But then any σ-algebra containing the σ-algebra generated by some class $\mathcal{A}$ must contain the whole of $\sigma(\mathcal{A})$. A result of this type that we will use is Dynkin's $\pi - \lambda$ theorem.

A *π-system* is a collection of sets closed under finite intersection. A *λ-system* is a collection G of sets satisfying

(i) $\Omega \in G$.
(ii) $A \subset B$ and $A, B \in G$ implies $B \setminus A \in G$.
(iii) $A_i \in G$ and $A_i \uparrow A$ implies $A \in G$.

Lemma 1.8

Let $\mathcal{P}$ be a π-system that is contained in a λ-system $\mathcal{L}$. Then

$$\sigma(\mathcal{P}) \subset \mathcal{L}.$$

We will first prove the case where $\mu(X) < \infty$. With this tool, note that a semi-algebra is a π-system, and let

$$\mathcal{L} = \{A : \mu_1(A) = \mu_2(A)\},$$

where μ_1 and μ_2 are the two extensions that we wish to show are equal. We know that $\mathcal{S} \subset \mathcal{L}$, so it suffices to show that $\mathcal{L}$ is a λ-system. A σ-algebra is also a λ-system, so given any π-system $\mathcal{P}$, we know that $\sigma(\mathcal{P})$ is the smallest λ-system conttaining $\mathcal{P}$. To verify $\mathcal{L}$ is a λ-system:

(i) $\Omega \in \mathcal{L}$ because $\Omega \in \mathcal{A}$.

(ii) Let $A \subset B$ and $A, B \in \mathcal{L}$. Then $\mu_1(A) = \mu_2(A)$ and $\mu_1(B) = \mu_2(B)$. Then, since $\mu_1(\Omega) = \mu(\Omega) = \mu_2(\Omega) < \infty$, we have

$$\mu_1(B \setminus A) = \mu_1(B) - \mu_1(A) = \mu_2(B) - \mu_2(A) = \mu_2(B \setminus A).$$

(iii) If $A_i \in \mathcal{L}$ and $A_i \uparrow A$, then we use continuity from below to see that

$$\begin{aligned} \mu_1(A_i) &\to \mu_1(A) \\ \mu_2(A_i) &\to \mu_2(A), \end{aligned}$$

for $A \in \mathcal{L}$.

Exercise: Modify the proof slightly to include the σ-finite case. □

Proof (Sketch) of Existence. Let $B \subset \mathbb{R}$. How do we define $\mu(B)$? We could try to approximate B by a union of intervals. This is known as *outer measure*. In particular, let

$$\mu_*(B) := \inf_{(A_i) \subset \mathcal{A},\ B \subset \cup_i A_i} \sum_{i=1}^{\infty} \mu(A_i)$$

We wish the outer measure to satisfy

(i) Monotonicity.

(ii) $\mu_*(\varnothing) = 0$.

(iii) Countable subadditivity.

Then we just need to upgrade countable subadditivity to countable additivity. Check the appendix in Durrett for the whole proof. □