Measures

You can study this material from Stokey and Lucas, Chapter 7; alternatively, in detail from Bass (Real Analysis for Graduate Students). An easy to access reference is Capinski and Kopp (Measure, Integral and Probability). Athreya and Lahiri (Probability Theory) is a detailed treatment.

1 Introduction

Our motivation for studying measure theory is to lay a foundation for modeling probabilities. I want to give a bit of motivation for the structure of measures that has developed by providing a sort of narrative of measurement. This follows from standard treatments of Lebesgue measure that you can find, for example, in Capinski and Kopp, Chapter 2, or in Royden, books that begin by first studying Lebesgue measure on $\Re$. For studying probability, we have to study measures more generally; that will follow the introduction.

People were interested in measuring length, area, volume etc. Let's start with length. How was one to extend the notion of the length of an interval (a, b), $l(a, b) = b - a$ to more general subsets of $\Re$? Given an interval I of any type (closed, open, left-open-right-closed, etc.), let $l(I)$ be its length (the difference between its larger and smaller endpoints). Then the notion of *Lebesgue outer measure* (LOM) of a set $A \in \Re$ was defined as follows. Cover A with a countable collection of intervals, and measure the sum of lengths of these intervals. Take the smallest such sum, over all countable collections of intervals that cover A, to be the LOM of A. That is, $m^*(A) = \inf Z_A$, where

$$Z_A = \{\sum_{n=1}^{\infty} l(I_n) : A \subseteq \cup_{n=1}^{\infty} I_n\}$$

(the I_n referring to intervals). We find that for a singleton set A, $m^*(A) = 0$; for an interval $(a,b), a < b$, $m^*((a,b)) = b - a$, corresponding to our notion of length, and so on.

LOM possesses the property of *(countable) subadditivity*: Suppose $(A_n)_{n=1}^{\infty}$ is a countable collection of subsets of $\Re$. Then $m^*(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} m^*(A_n)$.

We would like the following property for any measure: suppose I break up a set into, say, a countable collection of disjoint sets, then the measure of the mother set, and the sum of the measures of the disjoint sets, ought to be the same. This is *countable additivity*. Unfortunately, LOM on the class of all subsets of $\Re$ **does not** satisfy countable additivity.

People pointed out that the problem is with funny sets, which can be broken up into disjoint subsets in funny ways (e.g. the Banach and Tarski paradox). A popular way out is to restrict attention to a smaller class of subsets of $\Re$ (nevertheless, this is a very large class and is 'all we need') using a definition used by Caratheodory. According to this, we call a set $A \subseteq \Re$ *Lebesgue measurable* if for every set $E \in \Re$, we have

$$m^*(E) = m^*(E \cap A) + m^*(E \cap A^c)$$

Clearly $E \cap A$ and $E \cap A^c$ are disjoint and their union is E; yet, not all sets A have the above property.

Consider the class $\mathcal{M}$ of Lebesgue measurable sets. Restricted to this class, m^* is countable additive: For a countable collection $(E_n)_{n=1}^{\infty}$ of disjoint sets from $\mathcal{M}$, $m^*(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m^*(E_n)$. LOM restricted to $\mathcal{M}$ is simply called Lebesgue measure.

The class $\mathcal{M}$ satisfies (i) $\Re$ and $\emptyset$ belong to $\mathcal{M}$; (ii) $E \in \mathcal{M}$ iff $E^c \in \mathcal{M}$; (iii) if $(E_n)_{n=1}^{\infty}$ all belong to $\mathcal{M}$, then so does their union $\cup_{n=1}^{\infty} E_n$. That is, $\mathcal{M}$ is closed under complementation and countable unions. This got generalized: any class/family of subsets of any universal set that satisfies properties (i)-(iii) is called a σ-algebra.

You can check that the family of all subsets of $\Re$ (and more generally, given any set Ω, the family of all subsets of Ω) is a σ-algebra. But there are lots of σ-algebras that have fewer sets (e.g. the family $\mathcal{M}$ above). You can also check that the intersection of two σ-algebras is a σ-algebra. Now, suppose $\mathcal{C}$ is some family of sets (subsets of some space such as $\Re$, say). Consider all the σ-algebras that contain $\mathcal{C}$; their intersection is thus a σ-algebra, and obviously also contains $\mathcal{C}$. So, this is the smallest σ-algebra that contains $\mathcal{C}$. It is known as the **σ-algebra generated by $\mathcal{C}$**.

Now, Caratheodory's measurability property is hard to check: you must check that the candidate set A and A^c can split each set $E \in \Re$ satisfactorily. On the other hand, we can show that restricting LOM to *any* σ-algebra will yield countable additivity. So are there nice, and large σ-algebras that we can use in $\Re$, instead of the Lebesgue-measurable sets $\mathcal{M}$? A leading example of this is the **Borel σ-algebra** on $\Re$.

We consider $\Re$ as a metric space with the $|\ |$ or d_1 metric. Consider the family of all subsets of $\Re$ that are open sets under this metric. By definition, the Borel σ-algebra $\mathcal{B}$ or $\mathcal{B}(\Re)$ on $\Re$ is the σ-algebra generated by this family of open sets. (More generally, for any topological space (X, τ), the Borel σ-algebra in this space is the σ-algebra generated by the family τ of open sets). Interestingly, if we restrict ourselves to only intervals (even intervals of a particular kind like open intervals), and ask what's the σ-algebra generated by this family of intervals, the answer is the same, the Borel σ-algebra.

Now, this is the smallest σ-algebra containing the open sets of $\Re$; so, the σ-algebra $\mathcal{M}$ of all Lebesgue-measurable sets is at least as large as $\mathcal{B}$. But in a way that for measurement purposes does not matter. Specifically, for every set $A \in \mathcal{M}$, there exists a set $B \in \mathcal{B}$ that is a subset of A, such that $m^*(A - B) = 0$. So even if $A \notin \mathcal{B}$, there is a set $B \in \mathcal{B}$ that is a subset of A such that what remains after subtracting B is trivial.

2 Classes of Sets

Let Ω be a nonempty set and let $\mathcal{P}(\Omega)$ be its power set; i.e. the collection of all its subsets.

Definition 1 *A collection $\mathcal{F}$ of subsets of Ω is called an* **algebra** *if it contains Ω, and is closed under complementation and pairwise unions.*

Note: By De Morgan's laws, an algebra is closed under pairwise intersections as well. Moreover, by induction, it is closed under *finite* union and intersection. Note also that an algebra must contain the emptyset, $\emptyset$. We can dispense with the explicit requirement in the definition that $\mathcal{F}$ contain Ω if we require explicitly that $\mathcal{F}$ be a *nonempty* collection.

Definition 2 *$\mathcal{F}$ is a* **σ-algebra** *if it's an algebra and is also closed under countable unions.*

Note: A σ-algebra is obviously going to be closed under *monotone* unions as well. But in fact, this is an alternative *definition*, as the following proposition states.

Why define measures on σ-algebras rather than on algebras? First, generally, admitting countable unions and intersections of sets allows carrying out nice limit and continuity operations on a measure on a countably infinite sequence of sets. Second, specifically on the Lebesgue measure: the collection of sets that are finite disjoint unions of half-open intervals is an algebra, and it's easy to extend the notion of length (or area or volume) to them. But we may want to measure more complicated sets that lie outside this algebra.

Proposition 1 *Let $\mathcal{F} \subset \mathcal{P}(\Omega)$. Then $\mathcal{F}$ is a σ-algebra if and only if $\mathcal{F}$ is an algebra and satisfies*

$$A_n \in \mathcal{F}, A_n \subset A_{n+1} \forall n \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{F}$$

Proof. 'Only if' is obvious. For the 'if' part, let $\{B_n\}_{n=1}^{\infty} \subset \mathcal{F}$. Want to show that the countable union is in $\mathcal{F}$. For every n, define $A_n = \cup_{j=1}^{n} B_j$. Since $\mathcal{F}$ is an algebra, $A_n \in \mathcal{F}, \forall n$. Moreover, $A_n \subset A_{n+1}, \forall n$. Since $\mathcal{F}$ is by assumption closed under monotone unions, this implies $\cup_{n\geq 1} A_n \in \mathcal{F}$.

Finally, note that $\cup_{n\geq 1} A_n = \cup_{n\geq 1} B_n$. (Indeed, let $x \in \cup_n A_n$. Then $x \in A_N$, for some $N = 1, 2, ...$ Since $A_N = \cup_{j=1}^{N} B_j$, $x \in B_j$ for some j, and hence $x \in \cup_n B_n$. The converse is as straightforward.)

So, $\cup_{n\geq 1} B_n \in \mathcal{F}$, and so $\mathcal{F}$ is a σ-algebra. ∎

Example 1 *Let Ω be a nonempty set. Then $\mathcal{F}_3 = \mathcal{P}(\Omega)$ and $\mathcal{F}_4 = \{\emptyset, \Omega\}$ are σ-algebras (and so algebras). The latter is called the trivial σ-algebra.*

Let $\Omega = \{a, b, c, d\}$, where these are 4 distinct objects. The $\mathcal{F}_1 = \{\emptyset, \Omega, \{a\}\}$ is not an algebra or σ-algebra, but $\mathcal{F}_2 = \{\emptyset, \Omega, \{a\}, \{b, c, d\}\}$ is a σ-algebra.

Example 2 *Let Ω be an infinite set. Then $\mathcal{F}_5$, the collection of all subsets of Ω that are either finite sets or have complements that are finite sets, is an algebra but not a σ-algebra.*

Indeed, $\Omega \in \mathcal{F}_5$ since its complement is finite. Now suppose $A \in \mathcal{F}_5$. If A is finite, then $A^c \in \mathcal{F}_5$ since it has a finite complement. On the other hand, if $A \in \mathcal{F}_5$ because A^c is finite, then $A^c \in \mathcal{F}_5$ trivially.

Now suppose $A, B \in \mathcal{F}_5$. If one of these sets is finite, then clearly their intersection is finite and so belongs to $\mathcal{F}_5$. On the other hand, suppose neither set is finite. Then A^c, B^c and so $A^c \cup B^c$ are finite. So $(A \cap B)^c = A^c \cup B^c$ is finite, so $A \cap B \in \mathcal{F}_5$. So, $\mathcal{F}_5$ is an algebra.

To see that it's not a σ-algebra, let $\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of distinct objects from Ω. Being finite, each singleton set $\{x_n\} \in \mathcal{F}_5$. But the countable union of the odd-numbered elements, $\cup_{n \text{ odd}}\{x_n\}$ does not belong to $\mathcal{F}_5$, since neither it nor its complement is a finite set.

Note that an intersection of σ-algebras is a σ-algebra. Thus we may define

Definition 3 *If $\mathcal{A}$ is a collection of subsets of Ω, then the σ-algebra* **generated by** $\mathcal{A}$ *is defined as*

$$\sigma\langle\mathcal{A}\rangle = \cap_{\mathcal{F} \in \mathcal{I}(\mathcal{A})} \mathcal{F}$$

where $\mathcal{I}(\mathcal{A})$ is the collection of all σ-algebras containing $\mathcal{A}$.

Borel Sigma Algebras

A *topological space* is a pair (S, τ) where S is a nonempty set, and τ is a collection of subsets of S that contains S and is closed under pairwise intersections and arbitrary unions.

The sets belonging to τ are called *open sets.*

If (S, d) is a metric space, say under a d_p metric, the collection of sets open under this metric forms a topology.

Definition 4 *The* **Borel σ-algebra** *on a topological space (S, τ) is the σ-algebra $\sigma\langle\tau\rangle$ generated by the collection of open subsets of S.*

Let $\Re^k$ be k-dimensional Euclidean space with a d_p metric. $\mathcal{B}(\Re^k)$ denotes the Borel σ-algebra on $\Re^k$.

Proposition 2 $\mathcal{B}(\Re^k)$ *is also generated by each of the following classes of sets.*

$$\mathcal{O}_1 = \{(a_1, b_1) \times ... \times (a_k, b_k) : -\infty \leq a_i < b_i \leq \infty, i = 1, ..., k\}$$

$$\mathcal{O}_2 = \{(-\infty, x_1) \times ... \times (-\infty, x_k) : x_1, ..., x_k \in \Re\}$$

$$\mathcal{O}_3 = \{(a_1, b_1) \times ... \times (a_k, b_k) : a_i, b_i \in Q, a_i < b_i, i = 1, ..., k\}$$

$$\mathcal{O}_4 = \{(-\infty, x_1) \times ... \times (-\infty, x_k) : x_1, ..., x_k \in Q\}$$

Before sketching the proof, note that every open set in $\Re$ is a countable union of open intervals (in particular intervals with rational endpoints); and this can be generalized to $\Re^k$. Indeed, let A be an open set in $\Re$. So every point $x \in A$ is at the center of an open ball (open interval) $I_x = (a_x, b_x)$ that is entirely contained in A. So $A = \cup_{x \in A} I_x$. But between any 2 real numbers we can find a rational number, so there exist rational numbers $a_{x'}, b_{x'}$ s.t. $a_x < a_{x'} < x < b_{x'} < b_x$, so that setting $I_{x'} = (a_{x'}, b_{x'})$, we have $A = \cup_{x \in A} I_{x'}$. Moreover, since the rationals are countable, the number of distinct endpoints of the collection of $I_{x'}$'s is countable and therefore so is the number of these intervals.

Also note a couple of simple facts. If $\mathcal{C}$, $\mathcal{D}$ are collections of subsets of Ω, then we have

Fact(i) $\mathcal{C} \subset \sigma\langle\mathcal{D}\rangle$ implies $\sigma\langle\mathcal{C}\rangle \subset \sigma\langle\mathcal{D}\rangle$.

That is: $\mathcal{C}$ is contained in the σ-algebra $\sigma\langle\mathcal{D}\rangle$, which is therefore at least as large as the smallest σ-algebra containing $\mathcal{C}$.

Fact (ii) $\mathcal{C} \subset \mathcal{D}$ implies $\sigma\langle\mathcal{C}\rangle \subset \sigma\langle\mathcal{D}\rangle$.

This follows from Fact (i), since the antecedent implies the antecedent of (i).

Now for a sketch of the proof.

Proof. All the collections of sets above are open subsets of $\Re^k$. So for $i = 1, ..., 4$, $\sigma\langle\mathcal{O}_i\rangle \subset \mathcal{B}(\Re^k)$. We first prove the converse for classes 1 and 3 together.

Take any open set $A \subset \Re^k$. Then $A = \cup_{n=1}^{\infty} B_n$, where the B_n's are from class $\mathcal{O}_3$. So the countable union must be in the σ-algebra $\sigma\langle\mathcal{O}_3\rangle$. Thus the collection of *all* open sets is a subset of this σ-algebra; hence the σ-algebra generated by these, $\mathcal{B}(\Re^k)$ is a subset of it as well (by Fact (i)). The proof for $\mathcal{O}_1$ also follows, because $\mathcal{O}_3 \subset \mathcal{O}_1$. So the σ-algebra generated by the former is a subset of that generated by the latter (by Fact (ii)).

The result holds for classes 2 and 4 as well. To start simply, note that any open set in $\Re$ is a countable union of open intervals of type (a, b). Now, every (a, b) is itself a countable union of the form $\cup_n[(-\infty, b) - (-\infty, a + (1/n))]$. Each set in the union belongs to the σ-algebra generated by $\mathcal{O}_2$. A Countable union of countable unions is countable, so every open set in $\Re$ is a countable union of sets from the σ-algebra generated by $\mathcal{O}_2$. This can be generalized to $\Re^k$ and the proof then follows along the lines above. ∎

Dynkin Systems etc.

Definition 5 *A class $\mathcal{C}$ of subsets of Ω is a* **π-system** *if it is closed under pairwise intersections.*

So that's a nice easy class (easier than say a σ-algebra).

Definition 6 *A class $\mathcal{L}$ of subsets of Ω is a* **Dynkin system** *(or λ-system) if (i) $\Omega \in \mathcal{L}$, (ii) $A, B \in \mathcal{L}, A \subset B \Rightarrow (B - A) \in \mathcal{L}$, (i.e. $\mathcal{L}$ is closed* *under monotone relative complementation, and (iii)* $A_n \in \mathcal{L}, A_n \subset A_{n+1} \forall n = 1, 2, ... \Rightarrow \cup_{n\geq 1} A_n \in \mathcal{L}$. *(i.e.* $\mathcal{L}$ *is closed under monotone countable unions).*

Note that every σ-algebra $\mathcal{S}$ is a λ-system. Indeed, if $A, B \in \mathcal{S}, A \subset B$, then $B - A = B \cap A^c$ and this obviously belongs to $\mathcal{S}$. Note also that the intersection of λ-systems is a λ-system, so the notion of the λ-system generated by a class of sets is well-defined as the intersection of all λ-systems containing this class of sets.

Theorem 1 *(***The** $\pi - \lambda$ **theorem***). If* $\mathcal{C}$ *is a* π*-system, then* $\lambda\langle\mathcal{C}\rangle = \sigma\langle\mathcal{C}\rangle$*.*

Proof. $\sigma\langle\mathcal{C}\rangle$ is a λ-system containing $\mathcal{C}$, and therefore contains the λ-system generated by $\mathcal{C}$ as a subset. To show the converse, we show that $\lambda\langle\mathcal{C}\rangle$ is a σ-algebra if $\mathcal{C}$ is a π-system.

Note first that a λ-system $\mathcal{L}$ is closed under complementation, since $A, \Omega, \in \mathcal{L}, A \subset \Omega$, and $A^c = \Omega - A$. We show that if $\mathcal{C}$ is a π-system, then $\lambda\langle\mathcal{C}\rangle$ is closed under intersection as well, and so is an algebra; since it is also closed under monotone countable unions, we can appeal to Proposition 1 and say it's a σ-algebra.

Consider $\lambda_2(\mathcal{C}) = \{A \in \lambda\langle\mathcal{C}\rangle : A \cap B \in \lambda\langle\mathcal{C}\rangle \forall B \in \lambda\langle\mathcal{C}\rangle\}$. Clearly, $\lambda_2(\mathcal{C}) \subset \lambda\langle\mathcal{C}\rangle$. We are done if we show the converse. To show it, we show that $\mathcal{C} \subset \lambda_2(\mathcal{C})$ and that the latter is a λ-system.

Consider $\lambda_1(\mathcal{C}) = \{A \in \lambda\langle\mathcal{C}\rangle : A \cap B \in \lambda\langle\mathcal{C}\rangle \forall B \in \mathcal{C}\}$. Clearly, $\lambda_1(\mathcal{C}) \subset \lambda_2(\mathcal{C})$. Moreover, since $\mathcal{C}$ is a π-system, it is a subset of $\lambda_1(\mathcal{C})$. So, $\mathcal{C} \subset \lambda_2(\mathcal{C})$.

Finally, $\lambda_2(\mathcal{C})$ is a λ-system. (i) $\Omega \in \lambda_2(\mathcal{C})$ is trivial. (ii) If $A_1, A_2 \in \lambda_2(\mathcal{C}), A_1 \subset A_2$, then for every $B \in \lambda_2(\mathcal{C})$, $(A_1 \cap B) \subset (A_2 \cap B)$, these sets are in $\lambda\langle\mathcal{C}\rangle$, and so is there difference because of closure w.r.t. relative complementation. But $(A_2 \cap B) - (A_1 \cap B) = (A_2 - A_1) \cap B$, so $(A_2 - A_1) \in \lambda_2(\mathcal{C})$. The third property is as easy to show. ∎

One of the applications of the above result has to do with the result that if 2 finite measures agree on a π-system, then they agree on the σ-algebra generated by it.

3 Measures

Definition 7 *Let $\mathcal{F}$ be an algebra on Ω. A* **measure** *is a function defined on $\mathcal{F}$ that satisfies*

(a) $\mu(A) \in [0, \infty] \forall A \in \mathcal{F}$.

(b) $\mu(\emptyset) = 0$.

(c) For any disjoint collection of sets $A_1, A_2, ..., \in \mathcal{F}$ with $\cup_{n\geq 1} A_n \in \mathcal{F}$,

$$\mu(\cup_{n\geq 1} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

Measures have the following properties.

Proposition 3 *Let $\mathcal{F}$ be an algebra on Ω and μ a measure defined on it. Then*

(i) (finite additivity) $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$, for all mutually disjoint $A_1, A_2 \in \mathcal{F}$.

(ii) (monotonicity) $\mu(A) \leq \mu(B), \forall A, B \in \mathcal{F} s.t.\ A \subset B$.

(iii) (monotone continuity from below (mcfb)) For any collection $A_1, A_2, ... \in \mathcal{F}$ with $A_n \subset A_{n+1}, \forall n$ and $\cup A_n \in \mathcal{F}$,

$$\mu(\cup_{j\geq 1} A_j) = \lim_{n\to\infty} \mu(A_n)$$

(iv) (countable subadditivity) For any collection $A_1, A_2, ... \in \mathcal{F}$ with $\cup A_n \in \mathcal{F}$,

$$\mu(\cup A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

(v) (mcfa) For any collection $A_1, A_2, ... \in \mathcal{F}$ with $A_n \supset A_{n+1}, \forall n$, $A = \cap_n A_n$, $A \cap A_n \in \mathcal{F}$ and $\mu(A_k) < \infty$ for some k,

$$\mu(A) = \lim_{n\to\infty} \mu(A_n)$$

(vi) (inclusion-exclusion formula) $A_i \in \mathcal{F}, i = 1, ..., k, \mu(A_i) < \infty, \forall i$ implies

$$\mu(A_1 \cup ... \cup A_k) = \sum_{i=1}^{k} \mu(A_i) - \sum_{1\le i<j<k} \mu(A_i \cap A_j) + ... + (-1)^{k-1}\mu(A_1 \cap ... \cap A_k)$$

Proof of (i). Set $A_n = \emptyset$ for $n = 3, 4, ...$ and use countable additivity.

Proof of (ii). $B = A \cup (B - A)$ and the latter two sets are disjoint. Using finite additivity, $\mu(B) = \mu(A) + \mu(B - A) \ge \mu(A)$.

Proof of (iii). If $\mu(A_k) = \infty$ for some k, then by monotonicity this is true for all $n \ge k$ so both LHS and RHS equal ∞. Now suppose $\mu(A_n) < \infty \forall n$. Let $\{B_n\}$ be a sequence of sets defined as follows: $B_1 = A_1$ and for all $n > 1$, $B_n = A_n - A_{n-1}$. By finite additivity, we have $\mu(B_n) = \mu(A_n) - \mu(A_{n-1}), \forall n$ (with $A_0 = \emptyset$). Moreover, the B_n's form a collection of disjoint sets, and their union equals $\cup A_n$. So,

$$\mu(\cup A_n) = \mu(\cup B_n) = \sum \mu(B_n) = \lim_{N\to\infty} \sum_{n=1}^{N} [\mu(A_n) - \mu(A_{n-1})]$$

$$= \lim_{N\to\infty} \mu(A_N)$$

Proof of (iv). First, note that finite subadditivity holds. Indeed, $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2 - A_1) \le \mu(A_1) + \mu(A_2)$. The first equality follows from finite additivity, and the second from monotonicity. This result can be extended to any finite number of sets by induction.

Now, for $n \geq 1$, let $D_n = \cup_{i=1}^n A_i$. By finite subadditivity, $\mu(D_n) \leq \sum_{i=1}^n \mu(A_i)$, for all n. Since $\{D_n\}$ is a nested increasing sequence of sets with union $\cup_{i\geq 1} A_i$, by mcfb we have the first equality below.

$$\mu(\cup_{i\geq 1} A_i) = \lim_{n\to\infty} \mu(D_n) \leq \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A_i) = \sum_{n=1}^{\infty} \mu(A_n)$$

Proof of (v). WLOG let $\mu(A_1) < \infty$. Let $C_n = A_1 - A_n \forall n$ and $C_\infty = A_1 - A$. The C_n's are monotone nested increasing to C_∞ so by mcfb, $\mu(C_n) \uparrow \mu(C_\infty)$. Note that by finite additivity, $\mu(A_1) = \mu(C_\infty) + \mu(A)$, so by finiteness of $\mu(A_1)$, we get

$$\mu(C_\infty) = \mu(A_1) - \mu(A) \quad \dagger$$

.

Now, also by finite additivity, we have $\mu(A_1) = \mu(A_n) + \mu(C_n)$, so by finiteness of $\mu(A_1)$, we have $\mu(C_n) = \mu(A_1) - \mu(A_n)$. Now take limits on both sides. We get $\mu(C_\infty) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n)$. Compare with †. ■

Some Terms:

$(\Omega, \mathcal{F})$, where Ω is a nonempty set and $\mathcal{F}$ is a σ-algebra is called a **measurable space**.

A measure μ is called **finite** if $\mu(\Omega) < \infty$, and σ-finite if there is a countable collection of sets from the relevant σ-algebra whose union is Ω, and all sets of the collection are of finite measure. For example, the Lebesgue measure of $\Re$ is infinity, but the collection of all intervals of the form $(-n, n), n = 1, 2, ...$ has countable union $\Re$, and all of these intervals have positive measure equal to $2n$. So Lebesgue measure on say the Borel σ-algebra on $\Re$ is σ-finite.

Theorem 2 *(Uniqueness of Measures). Let μ_1, μ_2 be 2 finite measures on $(\Omega, \mathcal{F})$ where $\mathcal{F}$ is the σ-algebra generated by a π-system $\mathcal{C}$. If $\mu_1(C) = \mu_2(C) \forall C \in \mathcal{C}$ and $\mu_1(\Omega) = \mu_2(\Omega)$, then $\mu_1(A) = \mu_2(A) \forall A \in \mathcal{F}$.*

Proof. Let $\mathcal{L} \equiv \{A \in \mathcal{F} : \mu_1(A) = \mu_2(A)\}$. It is routine to check that $\mathcal{L}$ is a Dynkin system, and since it contains the π-system $\mathcal{C}$ and is a subset of $\mathcal{F} = \sigma\langle\mathcal{C}\rangle$, it must be equal to $\mathcal{F}$. ∎

3.1 Extension Theorems and Lebesgue-Stieltjes Measures

The idea is to start by assigning measures to a simple class of sets (such as assigning lengths to intervals on $\Re$), and then to use an approximation procedure to *extend* this measure to the more complicated sets in a σ-algebra. The simple class we will use is called a semialgebra or elementary family.

Definition 8 *A class $\mathcal{C} \subset \mathcal{P}(\Omega)$ is called a* **semialgebra** *if (i) it is closed under pairwise intersections, and (ii) if for every $A \in \mathcal{C}$, A^c is a disjoint union of a finite collection of sets from $\mathcal{C}$. We will also require a third thing: that $\emptyset \in \mathcal{C}$.*

Note that if a semialgebra has 2 sets at least, then it has a third (their intersection), and then, taking intersections, we see that the semialgebra must contain $\emptyset$.

Example 3 $\Omega = \Re$. $\mathcal{C} = \{(a, b] | -\infty \leq a, b < \infty\} \cup \{(a, \infty) | -\infty \leq a < \infty\}$.

Example 4 *(i)* $\Omega = \Re$. $\mathcal{C} = \{I \subset \Re | a, b \in I, a < b, \Rightarrow (a, b) \subset I\}$.

That is, I is an interval.

(ii) $\Omega = \Re^k$. $\mathcal{C} = \{I_1 \times ... \times I_k | I_j \text{ is an interval in } \Re\}$.

We first assess the measure of sets in a semialgebra.

Definition 9 *A function μ defined on a semialgebra $\mathcal{C}$ is a* **measure** *(or protomeasure) if (i) $\mu(C) \in [0, \infty] \forall C \in \mathcal{C}$, (ii) $\mu(\emptyset) = 0$, and (iii) $\mu(\cup_{n \geq 1} A_n) = \sum_1^n \mu(A_n)$, if the A_n's are a countable collection of sets from $\mathcal{C}$ that are disjoint and have union in $\mathcal{C}$.*

Next, we'll extend μ to the algebra generated by $\mathcal{C}$. Note first that for any semialgebra $\mathcal{E}$, the collection of finite disjoint unions of sets from this collection constitute the algebra generated by $\mathcal{E}$.

Lemma 1 *If $\mathcal{E}$ is a semialgebra, the collection $\mathcal{A}$ of finite disjoint unions of members of $\mathcal{E}$ is an algebra.*

Proof. Suppose $A, B \in \mathcal{A}$, so $A = \cup_1^n A_j$ for disjoint $A_j \in \mathcal{E}, | = \infty, .., \backslash$ and $B = \cup_1^m B_j$ for disjoint $B_j's \in \mathcal{E}$. To show that $A \cap B \in \mathcal{A}$, note that $A \cap B = \cup_{j,k}(A_j \cap B_k)$, a finite union of disjoint sets of $\mathcal{E}$. $A_j, B_k \in \mathcal{E} =>$ $\mathcal{A}_| \cap \mathcal{B}_{||} \in \mathcal{E}$, and $A_j \cap B_k$ and $A_i \cap B_s$ are disjoint, if $j \neq i, k \neq s$.

To show that $\mathcal{A}$ is closed under complements, let $A = \cup_1^n A_j \in \mathcal{A}$ where the $A_j's$ are disjoint sets in $\mathcal{E}$. So $A^c = \cap_1^n A_j^c$. Each A_j^c is a finite disjoint union of members of $\mathcal{E}$; for simplicity, let $A_j^c = B_j^1 \cup B_j^2$. Thus $A^c = \cap_1^n (B_j^1 \cup B_j^2)$ $= \cup\{B_1^{k_1} \cap ... \cap B_n^{k_n} : k_1, ..., k_n = 1, 2\}$. The sets in curly brackets are finite intersections of sets in $\mathcal{E}$ and hence are in $\mathcal{E}$; the union is finite and of disjoint sets. So, $A^c \in \mathcal{A}$. ∎

In fact, the collection of finite disjoint unions of the members of a semialgebra $\mathcal{E}$ forms the algebra $\mathcal{A}(\mathcal{E})$ generated by the semialgebra.

Now for the extension of μ to this algebra.

Proposition 4 *Let μ be a measure on a semialgebra $\mathcal{C}$. For each $A \in \mathcal{A} \equiv \mathcal{A}(\mathcal{C})$, set*

$$\bar{\mu}(A) = \sum_{i=1}^{k} \mu(B_i)$$

,

if $A = \cup_{i=1}^{k} B_i$ *for some finite collection* $B_i, i = 1, ..., k$ *of disjoint sets from* $\mathcal{C}$. *Then,*

(i) $\bar{\mu}(A)$ *is independent of the representation of* A *as* $A = \cup_{i=1}^{k} B_i$.

(ii) $\bar{\mu}$ *is countably additive on* $\mathcal{A}$. *If* $A_1, A_2, ...$ *is a countable collection of disjoint sets in* $\mathcal{A}$ *with union in* $\mathcal{A}$, *then*

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

.

Proof. (i) Let $A = \cup_{i=1}^{k} B_i = \cup_{j=1}^{n} C_j$ be 2 different representations of A in terms of finite collections of disjoint sets from $\mathcal{C}$. We show $\sum_{i=1}^{k} \mu(B_i) = \sum_{j=1}^{n} \mu(C_j)$.

For any i, since $B_i \subset A$, we have $B_i = \cup_j (B_i \cap C_j)$, a disjoint union. So by additivity, $\sum_{i=1}^{k} \mu(B_i) = \sum_{i=1}^{k} \sum_{j=1}^{n} \mu(B_i \cap C_j)$. By a similar argument, $\sum_{j=1}^{n} \mu(C_j)$ also equals this RHS.

(ii). It is straightforward to verify that $\bar{\mu}$ is finitely additive. Now, suppose $A_n \in \mathcal{A}, \setminus = \infty, \in, ...$ is a countable collection of disjoint sets, with $\cup_n A_n \in \mathcal{A}$. So for every n, there exists a finite collection of disjoint sets $\{B_{nj}\}$ in $\mathcal{C}$ whose union is A_n; and there exists a finite collection $\{B_i\}$ of disjoint sets in $\mathcal{C}$ whose union is $\cup_n A_n$. So,

$$\cup_i B_i = \cup_n A_n = \cup_n \cup_j B_{nj}$$

. So,

$$\bar{\mu}(\cup_n A_n) = \sum_{i=1}^{k} \mu(B_i) = \mu(\cup_i B_i) = \mu(\cup_n \cup_j B_{nj}) = \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} \mu(B_{nj})$$

The first equality follows by definition of $\bar{\mu}$, the second and fourth by countable (hence finite) additivity of μ.

On the other hand, $\bar{\mu}(A_n) = \sum_{j=1}^{k_n} \mu(B_{nj})$, so

$$\sum_{n=1}^{\infty} \bar{\mu}(A_n) = \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} \mu(B_{nj})$$

$\blacksquare$

Now we wish to extend μ to $\sigma(\mathcal{A}) = \sigma(\mathcal{C})$. We can do this by using μ to define an **outer measure** on all sets in $\mathcal{P}(\Omega)$ and then look at the induced measure on only the sets in $\sigma(\mathcal{C})$. We do this using Caratheodory's method. This introduces the notion of sets that are measurable according to an *outer* measure, and this in fact typically gives us a larger σ-algebra than $\sigma(\mathcal{C})$.

Definition 10 *A function $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ is called an* **outer measure** *on Ω if*

(i) $\mu^*(\emptyset) = 0$.

(ii) monotonicity: $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$.

(iii) Countable subadditivity: For every countable collection $\{A_n\}_{n\geq 1}$ of sets in $\mathcal{P}(\Omega)$,

$$\mu^*\left(\cup_{n\geq 1} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$$

Definition 11 *(Caratheodory). A set A is called* **μ^*-measurable** *if for every $E \in \Omega$,*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

That is, if we use a μ^*- measurable set and its complement to partition any set, we get a finite additivity result for that set.

Definition 12 *A measure space $(\Omega, \mathcal{F}, \nu)$ is* **complete** *if for every $A \in \mathcal{F}$ s.t. $\nu(A) = 0$, $\mathcal{P}(A) \subset \mathcal{F}$.*

Theorem 3 *(Caratheodory). Let* μ^* *be an outer measure on* Ω*. The collection* $\mathcal{M} \equiv \mathcal{M}_{\mu^*}$ *of all* μ^**-measurable sets of* Ω *is a* σ*-algebra;* μ^* *restricted to* $\mathcal{M}$ *is a measure; and* $(\Omega, \mathcal{M}, \mu^*)$ *is a complete measure space.*

The proof can be found in most textbooks on Measure Theory, including **A-L**.

Theorem 4 *(Caratheodory's Extension Theorem). Let* μ *be a measure on a semialgebra* $\mathcal{C}$ *with* $\mu(\emptyset) = 0$*. Define* μ^* *on all sets* A *in* $\mathcal{P}(\Omega)$ *as follows:*

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) | \{A_n\}_{n \geq 1} \subset \mathcal{C}, A \subset \cup_{n \geq 1} A_n \right\}$$

(If there is no such cover for A, let $\mu^*(A) = \infty$*, the convention being the one for minimizing over an empty feasible set). Then,*

(i) μ^* *is an outer measure.*

(ii) $\mathcal{C} \subset \mathcal{M}_{\mu^*}$*.*

(iii) $\mu^* = \mu$ *on* $\mathcal{C}$*.*

Proof. (i) (a). $\mu^*(\emptyset) = 0$. To see this, just cover the set with a countable collection of empty sets, and note that $\mu(\emptyset) = 0$. (b) $A \subseteq B$ implies $\mu^*(A) \leq \mu^*(B)$. Indeed, all collections that cover B cover A as well, so the infimum for A is pulled out of a larger set. (c) For countable subaddivity, suppose $A_n \in \Omega, n = 1, 2, \ldots.$ Fix any $\epsilon > 0$. For each of the A_n's , there exists a collection $(C_{nj}), j = 1, 2, \ldots$ of sets from $\mathcal{C}$ whose union covers A_n and $\mu^*(A_n) + \frac{\epsilon}{2^n} \geq \sum_{j=1]\infty \mu(C_{nj})}$. So, $\cup_{n=1}^{\infty} A_n \subseteq \cup_{n=1}^{\infty} \left(\cup_{j=1}^{\infty} C_{nj} \right)$, so again by the infimum property of $\mu^*()$ we have:

$$\begin{aligned} \mu^*(\cup_{n=1}^{\infty} A_n) &\leq \textstyle\sum_{n=1}^{\infty} \left(\sum_{j=1}^{\infty} \mu(C_{nj}) \right) \\ &\leq \textstyle\sum_{n=1}^{\infty} \left(\mu^*(A_n) + \frac{\epsilon}{2^n} \right) \\ &= \textstyle\sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, countable subadditivity follows.

(ii) Let $A \in \mathcal{C}$. We want to show that then A is μ^*-measurable. For any $E \subset \Omega$, $E = (E \cap A) \cup (E \cap A^c)$. So by subadditivity of outer measures, $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$. So we just need to show that the inequality can reverse its direction.

Let $\{A_n\}_{n\geq 1}$ be a countable cover of E, pulled out of $\mathcal{C}$. So, $\{(A_n \cap A)\}_n$ and $\{A_n \cap A^c\}_n$ are countable covers of $(E \cap A)$ and $(E \cap A^c)$. Moreover, since $A \in \mathcal{C}$, $A^c = \cup_{i=1}^k B_i$, for some finite collection of disjoint sets $\{B_i\}$ from $\mathcal{C}$. So, $\{(A_n \cap B_i) | i = 1, ..., k, n = 1, 2, ...\}$ is a countable cover of $(E \cap A^c)$. So,

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \sum_{i=1}^{k} \mu(A_n \cap B_i) \quad \dagger$$

On the other hand, for any of the sets A_n, we have that it equals the disjoint union $A_n = (A_n \cap A) \cup \cup_{i=1}^k (A_n \cap B_i)$, so by finite additivity of μ on $\mathcal{C}$,

$$\mu(A_n) = \mu(A_n \cap A) + \sum_{i=1}^{k} \mu(A_n \cap B_i)$$

. So,

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \sum_{i=1}^{k} \mu(A_n \cap B_i)$$

Comparing with $\dagger$, we have

$$\sum_{n=1}^{\infty} \mu(A_n) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Since this is true for every cover of E, the above inequality will hold if the LHS equals the infimum, $\mu^*(E)$.

(iii). Let $A \in \mathcal{C}$. Since A covers itself, clearly $\mu^*(A) \leq \mu(A)$ by definition. To prove the inequality in the opposite direction, note that that will hold if $\mu^*(A) = \infty$. So suppose $\mu^*(A) < \infty$. Then, by definition there exists a countable cover $\{A_n\}_{n\geq 1}$ of A, pulled out from $\mathcal{C}$, such that for every $\epsilon > 0$,

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(A) + \epsilon$$

(This is due to $\mu^*(A)$ being an infimum). WLOG let the sets $\{A_n\}_{n\geq 1}$ be disjoint (for if not, then we can use them to get a countable collection of disjoint sets with the same union). Then A equals the disjoint union $A = \cup_{n\geq 1}(A \cap A_n)$. So,

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A \cap A_n) \leq \sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(A) + \epsilon$$

The first equality above follows from countable additivity of μ on $\mathcal{C}$, and the first inequality from monotonicity. Since the above is true for every $\epsilon > 0$, we have $\mu(A) \leq \mu^*(A)$. ∎

Given a measure μ on a semialgebra $\mathcal{C}$, the measure space $(\Omega, \mathcal{M}_{\mu^*}, \mu^*)$ is called the **Caratheodory extension of** μ. Since $\mathcal{M}_{\mu^*}$ is a σ-algebra containing $\mathcal{C}$, it contains $\sigma\langle\mathcal{C}\rangle$, the σ-algebra generated by $\mathcal{C}$. The latter σ-algebra need not be complete, as we see in an example below.

Lebesgue-Stieltjes Measures on $\Re$

Let $F : \Re \to \Re$ be nondecreasing. Let $F(x+) \equiv \lim_{y\downarrow x} F(y)$, $F(x-) \equiv \lim_{y\uparrow x} F(y)$, $(\forall x \in \Re)$, $F(\infty) = \lim_{y\uparrow\infty} F(y), F(-\infty) = \lim_{y\downarrow -\infty} F(y)$. Let

$$\mathcal{C} = \{(a, b] | -\infty \leq a \leq b < \infty\} \cup \{(a, \infty) | -\infty \leq a < \infty\}$$

$\mathcal{C}$ is a semialgebra. Define

$$\mu_F((a, b]) = F(b+) - F(a+)$$

,

$$\mu_F((a, \infty)) = F(\infty) - F(a+)$$

.

It is easy to show that μ_F as defined is a measure on $\mathcal{C}$ (see problems 1.22-23 of **A-L**). F-like functions are models for distribution functions.

Definition 13 *The Caratheodory extension* $(\Re, \mathcal{M}_{\mu_F^*}, \mu_F^*)$ *of* μ_F *is called a* **Lebesgue-Stieltjes measure space***, and* μ_F^* *the Lebesgue-Stieltjes measure generated by* F.

Since $\sigma\langle\mathcal{C}\rangle = \mathcal{B}(\Re)$, the class of Borel sets of $\Re$, every Lebesgue-Stieltjes measure μ_F^* is also a measure on the measurable space $(\Re, \mathcal{B}(\Re))$.

Now notice that μ_F^* is finite on bounded intervals. A measure that is finite on bounded intervals is called a **Radon measure**. Now, given any Radon measure μ on $(\Re, \mathcal{B}(\Re))$, define

$$F(x) = \begin{cases} \mu((0, x]) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\mu((x, 0]) & \text{if } x \leq 0 \end{cases}$$

Then $\mu_F = \mu$ on $\mathcal{C}$. By the uniqueness of the extension, $\mu_F^* = \mu$ on $\mathcal{B}(\Re)$. So every Radon measure is a Lebesgue-Stieltjes measure.

Definition 14 *When* $F(x) = x, x \in \Re$*, the measure* μ_F^* *is called the* **Lebesgue measure** *and the* σ*-algebra* $\mathcal{M}_{\mu_F^*}$ *is the class of Lebesgue measurable sets.*

I need to put in stuff about completion and Lebesgue versus Borel σ-algebra, skipped for now due to time constraint. Better still, fill in worked out details of various properties of Lebesgue measure.

4 Probability Measures

Definition 15 $(\Omega, \mathcal{F}, \mathcal{P})$ *is a probability space if* $\mathcal{F}$ *is a* σ*-algebra on* Ω *and* P *is a measure s.t.* $P(\Omega) = 1$.

Probabilists and especially statisticians often call $\mathcal{F}$ a σ-field. The elements $\mathcal{F}$ (subsets of Ω) are called events.

Suppose $P(B) > 0$. Then the number $P(A|B) = \frac{P(A \cap B)}{P(B)}$ is called the conditional probability of A given B.

Now, for this set B, consider the family of all sets $(A \cap B)$ s.t. $A \in \mathcal{F}$. These are all subsets of B; in fact, this family of sets forms a σ-algebra on B; it's called $\mathcal{F}_{\mathcal{B}}$. It can be shown that the mapping $A \mapsto P(A|B)$ is countably additive on $\mathcal{F}_{\mathcal{B}}$.

We say two events A and B are independent if $P(A \cap B) = P(A)P(B)$. Two σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent if for every $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$ we have $P(F_1 \cap F_2) = P(F_1)P(F_2)$.

Extending this to a finite number of events or σ-algebras needs care, because pairwise independence as above neither implies nor is implied by a single similar relationship for a larger number of events. So for example we define the events $A_1, \ldots, A_n$ to be independent if for every $k \in \{2, \ldots, n\}$, for all choices of distinct indices $i_1, \ldots, i_k$ from $\{1, \ldots, n\}$, we have

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$$

.