Lecture 5: Caratheodory Extension Thm with application to Lebesgue measure.

This is an important Thm to know about.
The proof is less important.
We will skip the details (which are partly covered in the Notes.
That said, understanding the outline of the constructive proof helps one understand where outer/inner measures come from & the nature of non-measurable sets.

Fix a sample space Ω.
Let $\mathcal{F}_0$ be a field over Ω.

Theorem (Carathéodory extension):
Any probability measure P_0 on $(\Omega, \mathcal{F}_0)$ has a unique extension to a probability measure P on $\sigma\langle \mathcal{F}_0 \rangle =: \mathcal{F}$.

Comments on the proof:

Uniqness follows from our "uniqness Thm" since the generators $\mathcal{F}_0$ is a π-sys.

The proof proceeds by adding sets to $\mathcal{F}_0$ in two diffent ways, defining extensions of P_0, then simplifying.

Step 1: $\mathcal{F}^{\uparrow} :=$ closure of $\mathcal{F}_0$ under monotonically increasing set limits

$P^{\uparrow} :=$ extension defined by the monotonic limit $\lim\uparrow P_0(A_n)$.

Now extend $\mathcal{F}^{\uparrow}$ to 2^{Ω} & define

$$P^*(A) := \inf\{P^{\uparrow}(B) : A \subset B \in \mathcal{F}^{\uparrow}\}$$

Step 2: $\mathcal{F}^{\downarrow} :=$ closure of $\mathcal{F}_0$ under monotonically decreasing set limits

$P^{\downarrow} :=$ extension defined by the monotonic limit $\lim\downarrow P_0(A_n)$.

Now extend $\mathcal{F}^{\downarrow}$ to 2^{Ω} & define

$$P_*(A) := \sup\{P^{\downarrow}(B) : B \subset A,\ B \in \mathcal{F}^{\downarrow}\}$$

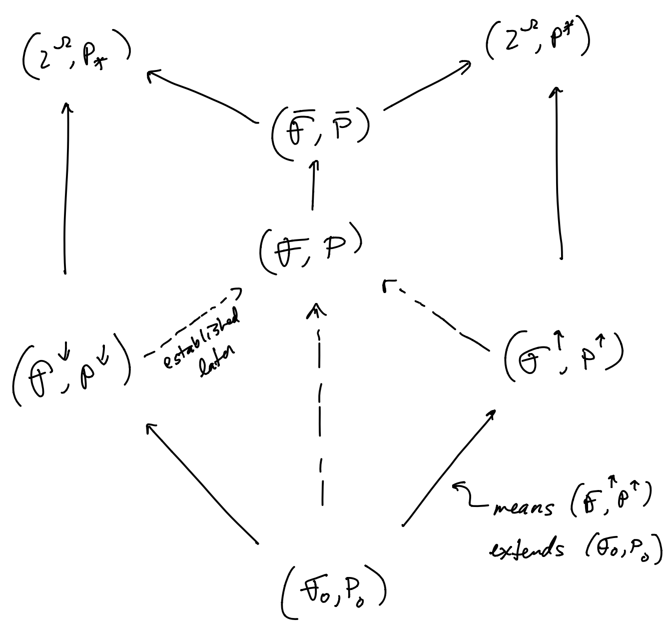
called the inner measure

Step 3: $\bar{\mathcal{F}} = \{A \subset \Omega : P^*(A) = P_*(A)\}$

$\bar{P}(A) :=$ common value of P^* & P_*

show $(\bar{\mathcal{F}}, \bar{P})$ is a prob measure, $\mathcal{F} \subset \bar{\mathcal{F}}$ & $(\mathcal{F}, \bar{P}|_{\mathcal{F}})$ extends $(\mathcal{F}_0, P_0)$.

Here is a visual map



Note: P^* & P_* are not necessarily measures. They can be useful, however, when working with events A that may or may not be in $\mathcal{F} = \sigma\langle \mathcal{F}_0 \rangle$ or $\overline{\mathcal{F}}$. In particular $P_*(A) \leq P^*(A)$ $\forall A \subset \Omega$. (3)

$$\therefore \quad A \subset \Omega \;\&\; P^*(A) = 0 \implies P_*(A) = 0$$

$$\iff A \in \overline{\mathcal{F}} \;\&\; \overline{P}(A) = 0$$

$$\iff A \text{ is } P\text{-neg}$$

e.g. Let $((0,1], \mathcal{B}_0^{(0,1]}, P)$ be the FAP generated by coin flips from Lecture 1. Borel's SLLN shows $\forall \varepsilon > 0$ $\exists B_1, B_2, \ldots \in \mathcal{B}_0^{(0,1]}$ s.t. $N^c \subset \bigcup_{k=1}^{\infty} B_k$ & $\sum_{k=1}^{\infty} P(B_k) \leq \varepsilon$.

Since $\bigcup_{k=1}^{\infty} B_k = \lim_n \uparrow \underbrace{\bigcup_{k=1}^{n} B_k}_{\in \mathcal{B}_0^{(0,1]}} \in \mathcal{F}^{\uparrow}$

If we can show P is a measure on $\mathcal{B}_0^{(0,1]}$ then Carathéodory applies &

$P^*(N^c) = 0 \leftarrow$ the inf over $\mathcal{F}^{\uparrow}$ covers contains all $\varepsilon > 0$.

which would imply $P(N^c) = 0$
$\uparrow$ the extension given by Carathéodory.

Another way to see this, since we already know $N \in \mathcal{B}^{(0,1]}$, is

$$P(N^c) \leq P\Big(\bigcup_k B_k^{\varepsilon}\Big) \leq \sum_k P(B_k^{\varepsilon}) \leq \varepsilon, \quad \forall \varepsilon > 0.$$

$\therefore P(N^c) = 0$

Note: P^* & P_* can also be used to show there are sets $A \subset (0,1]$ which are not in $\mathcal{B}^{(0,1]}$ or $\overline{\mathcal{B}}^{(0,1]}$. (4)

This will follow if $\exists$ a prob measure $((0,1], \mathcal{B}_0^{(0,1]}, P)$ & $A \subset (0,1]$ s.t.

$$P_*(A) = 0 \;\&\; P^*(A) = 1$$

... i.e. $A \notin \overline{\mathcal{B}}^{(0,1]}$

Thm (different formulas for P^*, P_*)

Let $(\Omega, \mathcal{F}, P)$ be a measure space. Then $\forall A \subset \Omega$,

$$P^*(A) = \inf\{\overset{\text{since } P = P^{\uparrow}}{P(B)} : A \subset B \in \mathcal{F}^{\uparrow}\}$$

$$= \min\{P(B) : A \subset B \in \mathcal{F}\} \quad (*)$$

$\uparrow$ means the inf is attained

$$P_*(A) = \sup\{P(B) : A \supset B \in \mathcal{F}^{\downarrow}\}$$

$$= \max\{P(B) : A \supset B \in \mathcal{F}\}$$

Part of Proof: Lets prove $(*)$

$$P^*(A) := \inf\{P^{\uparrow}(B) : A \subset B \in \mathcal{F}^{\uparrow}\}$$

$$= \inf\{P^*(B) : A \subset B \in \mathcal{F}^{\uparrow}\}$$

$$\geq \inf\{\underbrace{P^*(B)}_{= P(B)} : A \subset B \in \mathcal{F}\}$$

since P^* extends $P^{\uparrow}$; $\uparrow$ larger set

$$\geq P^*(A)$$

since $A \subset B \Rightarrow P^*(A) \leq P^*(B)$ by the fact that $P^{\uparrow}$ is increasing over increasing $\mathcal{F}^{\uparrow}$ sets (since $P^{\uparrow} = P$ on $\mathcal{F}^{\uparrow}$).

To show the inf is attained

Let $B_1, B_2, B_3, \ldots \in \mathcal{F}$ s.t. $A \subset B$ &

$$P(B_n) \downarrow P^*(A).$$

$\therefore$ Now $A \subset \bigcap_{n=1}^{\infty} B_n \in \mathcal{F}$

$$\therefore P^*(A) \le P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_m P\left(\bigcap_{n=1}^{m} B_n\right)$$

$$\le \lim_m P(B_m)$$

$$= P^*(A).$$

QED

Thm (Approximating P on $\mathcal{F}$ with $\mathcal{F}_0$):

Suppose $A \in \sigma\langle \mathcal{F}_0 \rangle$ & $\varepsilon > 0$. Then

1. $\exists B \in \mathcal{F}_0$ s.t.

$$P(A \Delta B) \le \varepsilon.$$

2. $\exists B_1, B_2, \ldots \in \mathcal{F}_0$ s.t. $A \subset \bigcup_{k=1}^{\infty} B_k$ &

$$P\left(\bigcup_{k=1}^{\infty} B_k - A\right) \le \varepsilon.$$

This thm also holds for σ-finite measures.

Proof: left as an exercise.

Application: Lebesgue measure from coin flips.

Let $([0,1], \mathcal{B}_0([0,1]), P)$ be the FAP model from lecture 1 i.e. $P(A) = \text{length}(A)$, $\forall A \in \mathcal{B}_0([0,1])$

We will use this to construct Lebesgue measure on $\mathbb{R}^d$ using Caratheodory.

Lemma: If $F_1, F_2, \ldots$ are compact subsets s.t. $\bigcap_{n=1}^{\infty} F_n = \phi$ then $\bigcap_{n=1}^{N} F_n = \phi$ for some N.

Proof:

Suppose not.

$\therefore \forall N \; \exists x_N \in \bigcap_{n=1}^{N} F_n$

$\therefore x_N \in F_n \quad \forall n \le N$

$\therefore \forall n, \; x_N \in F_n \; \forall$ large N

$\therefore \exists$ a subseq $\{x_{N_k}\}$ of $\{x_N\}$ s.t.

$$x_{N_k} \xrightarrow{k \to \infty} x \in F_1$$

To finish notice $x \in F_n \;\; \forall n$ since x_{N_k} is eventually in F_n for large enough k & F_n is closed.

$$\therefore \underbrace{x \in \bigcap_{n=1}^{\infty} F_n}_{\text{contradiction.}}$$

QED.

Thm: $P: \mathcal{B}_0((0,1]) \to [0,1]$ is a probability measure (7)

Proof: It will be sufficient to show P is continuous from above @ $\emptyset$.

Let $A_n \in \mathcal{B}_0((0,1])$ s.t. $A_n \downarrow \emptyset$.

We already know $P(A_n)\downarrow$ so just show $\lim_n P(A_n) = 0$.

Let $A_n^\varepsilon \in \mathcal{B}_0((0,1])$, $F_n^\varepsilon \subset (0,1]$ be closed s.t.

$$A_n^\varepsilon \subset F_n^\varepsilon \subset A_n \quad \&$$

$$P(A_n - A_n^\varepsilon) \le \frac{\varepsilon}{2^n}.$$

e.g. If $A_n = (0, 1/n]$ then $A_n^\varepsilon = \left(\frac{\varepsilon}{2^{n+1}}, \frac{1}{n} - \frac{\varepsilon}{2^{n+1}}\right]$

$$F_n^\varepsilon = \left[\frac{\varepsilon}{2^{n+1}}, \frac{1}{n} - \frac{\varepsilon}{2^{n+1}}\right]$$

Now the lemma gives $\exists N_\varepsilon$ s.t.

$$\bigcap_{n=1}^{N_\varepsilon} F_n^\varepsilon = \emptyset, \quad \text{since } \bigcap F_n^\varepsilon \subset \bigcap A_n = \emptyset.$$

$$\therefore \bigcap_{n=1}^{N_\varepsilon} A_n^\varepsilon = \emptyset$$

$$\therefore P\left(\bigcap_{n=1}^{N_\varepsilon} A_n\right) - \underbrace{P\left(\bigcap_{n=1}^{N_\varepsilon} A_n^\varepsilon\right)}_{=0} \le \sum_{n=1}^{N_\varepsilon} P(A_n - A_n^\varepsilon) \le \varepsilon$$

Since $\bigcap A_n - \bigcap A_n^\varepsilon \subset \bigcup (A_n - A_n^\varepsilon)$

$$\therefore \underbrace{P\left(\bigcap_{n=1}^{N_\varepsilon} A_n\right)}_{\text{this covers } A_k\ \forall \text{ large } k \text{ since } A_1 \supset A_2 \supset \cdots} \le \varepsilon$$

$\therefore P(A_k) \le \varepsilon \quad \forall$ large k

QED

(8)

Now Caratheodory applies so there is a unique extension to a probability measure $((0,1], \mathcal{B}((0,1]), P)$.

This models coin flips in the following way: $X_1, X_2, \ldots$ ← the binary digits.

is a sequence of maps $X_k: (0,1] \to \{0,1\}$ s.t.

$$P(X_k = 1) = 1/2$$
$$P(X_k = 0) = 1/2.$$

We will show the X_k's are indep later.

Borel's SLLN is now translated to say

$$P\Big(\underbrace{\lim_n \frac{X_1 + \cdots + X_n}{n} = 1/2}_{\text{This is the set } N \in \mathcal{B}((0,1])}\Big) = 1$$

The extension to $\mathcal{B}^{(0,1]}$

Notice that $((0,1], \mathcal{B}((0,1]), P)$ also models the uniform measure on $(0,1]$ since P restricts to "length" on $\mathcal{B}_0((0,1])$

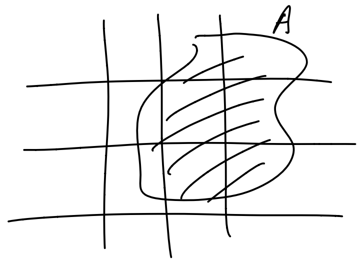
To get Lebesgue measure on $\mathbb{R}^d$ use a similar construction to derive the uniform probability measure $((i,i+1], \mathcal{B}((i,i+1]), P_i)$ on $(i,i+1] := (0,1]^d + i$ where $i \in \mathbb{Z}^d$ &

$$P(A) = \mathrm{vol}(A)$$

$\forall A \in \mathcal{B}_0((i,i+1])$

Now define $\mathcal{L}^d(A) := \sum_{i\in\mathbb{Z}^d} P_i(A\cap(i,i+1])$

$\forall A\in\mathcal{B}(\mathbb{R}^d)$



Notice that $A\cap(i,i+1] \in \mathcal{B}((i,i+1])$

so $P(A\cap(i,i+1])$ is defined.

$\mathcal{B}(\Omega)\cap\Omega_0 = \mathcal{B}(\Omega_0)$

Thm (Properties of $\mathcal{L}^d$)

(1) $\mathcal{L}^d$ is a measure on $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$.

(2) $\mathcal{L}^d(A) = vol(A)$ when $A\in\mathcal{B}_0(\mathbb{R}^d)$ & is the unique measure, σ-finite over $\mathcal{B}_0(\mathbb{R}^d)$, with this property.

(3) $\mathcal{L}^d(A+x) = \mathcal{L}^d(A)$, $\forall x\in\mathbb{R}^d$, $\forall A\in\mathcal{B}(\mathbb{R}^d)$

(4) $\mathcal{L}^d(A)=0$ if A is a k-dim hyperplane of $\mathbb{R}^d$ where $k<d$

(5) $\mathcal{L}^d(TA) = |\det T|\,\mathcal{L}^d(A)$, $\forall A\in\mathcal{B}(\mathbb{R}^d)$ if $T:\mathbb{R}^d\to\mathbb{R}^d$ is linear & non-singular.

Note: $TA\in\mathcal{B}(\mathbb{R}^d)$ in this case.... but be careful since $\exists$ singular linear maps s.t. $TA\notin\mathcal{B}(\mathbb{R}^d)$ (see Billingsley.)

See the Notes or Billingsley for proof details.

__Remarks:__

- For (2) use uniqueness of measures and the fact that $\mathcal{L}^d$ is σ-finite over $\mathcal{A} = \{(a,b] : -\infty < a_i < b_i < \infty\}\cup\{\emptyset\}$ & $\mathcal{B}(\mathbb{R}^d) = \sigma\langle\mathcal{A}\rangle$ & $\mathcal{A}$ is a π-sys.
- For (3) use good sets to show $A+x\in\mathcal{B}(\mathbb{R}^d)$ & uniqueness Thm.
- For (4) use the following Thm which is proved in the Notes & will come up later.

Thm: If $(\Omega,\mathcal{F},\mu)$ is a σ-finite measure space then __NO__ $A\in\mathcal{F}$ has the form

$$A=\bigcup_{i\in I} B_i$$

where $B_i\in\mathcal{F}$ are disjoint, $\mu(B_i)>0$ & I is uncountable.

Remark: This is a good time to recall the fact that an uncountable union of measure zero sets may result in a set with non-zero measure.

Def: Let $A \subset \mathbb{R}^d$. Then

- A is Borel measurable iff $A \in \mathcal{B}(\mathbb{R}^d)$
- A is Lebesgue measurable iff $A \in \overline{\mathcal{B}(\mathbb{R}^d)}$

where $(\mathbb{R}^d, \overline{\mathcal{B}(\mathbb{R}^d)}, \bar{\mathcal{L}}^d)$ be the completion of $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ w.r.t $\mathcal{L}^d$.

Thm (why we need σ-fields)

- $\mathcal{B}(\mathbb{R}) \subsetneq \overline{\mathcal{B}(\mathbb{R})} \subsetneq 2^{\mathbb{R}}$
- There exists no measure μ on $2^{\mathbb{R}}$ s.t. $\mu(A+x) = \mu(A)$, $\forall x \in \mathbb{R}\ \forall A \in \mathcal{B}(\mathbb{R})$ & $\mu(A) = \text{length}(A)$, $\forall A \in \mathcal{B}_0(\mathbb{R})$.