Lecture 4: Measures

We are going to define measures and probability measures in tandem so we can compare their properties.

Def: Let $\mathcal{F}_0 \subset 2^{\Omega}$ be a field over Ω

P is a <u>Probability measure</u> on $(\Omega, \mathcal{F}_0)$ if

- $P: \mathcal{F}_0 \to [0,1]$
- $P(\Omega) = 1$
- $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$

  when $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_0$ & A_k's are disjoint

μ is a <u>measure</u> on $(\Omega, \mathcal{F}_0)$ if

- $\mu: \mathcal{F}_0 \to [0, \infty]$
- $\mu(\emptyset) = 0$
- $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$

  when $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_0$ & A_k's are disjoint

Note: these definitions are over fields and allow me to state extension thms easily. (i.e. if P is a prob measure on $\mathcal{B}_0^{(0,1]}$ then $\exists$ an extended prob measure on $\sigma\langle \mathcal{B}_0^{(0,1]} \rangle$).

Note: Any Prob measure is also a measure since

$\mathcal{F}$ is not empty $\Rightarrow \exists A \in \mathcal{F}$

$\Rightarrow \Omega = A \cup A^c \in \mathcal{F}$

$\emptyset = \Omega^c \in \mathcal{F}$

$\Rightarrow 1 = P(\Omega) = P(\emptyset) + \underbrace{P(\Omega)}_{=1}$

$\Rightarrow P(\emptyset) = 0$

Def: $(\Omega, \mathcal{F})$ is a <u>measurable space</u> if $\mathcal{F}$ is a σ-field over Ω

Def: $(\Omega, \mathcal{F}, P)$ is called a <u>Probability space</u>.

$\mathcal{F}$: σ-field; P: Prob measure on $\mathcal{F}$

$(\Omega, \mathcal{F}, \mu)$ is called a <u>measure space</u>

μ: measure on $\mathcal{F}$

Def:

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

<u>μ is finite</u> if $\mu(\Omega) < \infty$.

<u>μ is infinite</u> if $\mu(\Omega) = \infty$.

<u>μ is σ-finite over $\mathcal{A} \subset \mathcal{F}$</u> if

$\exists A_1, A_2, \ldots \in \mathcal{A}$ s.t.

$$\Omega = \bigcup_{k=1}^{\infty} A_k \quad \&$$

$$\mu(A_k) < \infty$$

<u>μ is σ-finite</u> if μ is σ-finite over $\mathcal{F}$.

<u>Remark</u>: Most of the measures people work with are σ-finite. In fact, we will show that all non-trivial σ-finite measures can be represented as integrals $\mu(A) = \int_A f\, dP$ where P is a prob (The probabilist's world view of measure theory).

<u>Remark</u>: $\exists$ FAP that are not measures.

e.g. $\mathcal{F}_0$ = finite & co-finite subsets of an infinite sample space.

$$P(A) = \begin{cases} 0 & \text{if } A \text{ is finite} \\ 1 & \text{if } A \text{ is co-finite} \end{cases}$$

Now P is a FAP, P is a measure if Ω is uncountable but P is <u>not</u> a measure if Ω is countable.

Probability vrs Measure

Probability

Example

$\Omega = \{1, 2, \dots, n\}$

$\mathcal{F} = 2^{\Omega}$

$P(A) = \frac{1}{n} \sum_{\omega \in \Omega} \underbrace{I_A(\omega)}_{\text{indicator of } A} = \#$ of elements in A

P is the uniform probability on $\{1, 2, \dots, n\}$

Basic Properties

Suppose P is a probability measure on a field $\mathcal{F}_0$ over Ω

Then...

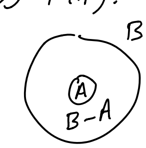
1. $P(A^c) = 1 - P(A)$

   since $P(\Omega) = P(A \cup A^c)$
   $= P(A) + P(A^c)$

2. $P(\emptyset) = 0$

3. If $A \subset B$ then
   $P(B - A) = P(B) - P(A)$.

   since $B = (B - A) \cup A$



4. $A \subset B \Rightarrow P(A) \le P(B)$

   by 3 since $0 \le P(B - A)$
   $= P(B) - P(A)$

Measure

Example

Ω is any set

$\mathcal{F} = 2^{\Omega}$

$\mu(A) = \sum_{\omega \in \Omega} I_A(\omega)$

defined as the sup over finite sums. Can prove a version of fubini to prove the measure axioms hold

μ is called counting measure.

Basic Properties

Suppose μ is a measure on a field $\mathcal{F}_0$ over Ω

Then...

1. $\mu(A^c) = \mu(\Omega) - \mu(A)$ if $\mu(A) < \infty$

   since $\mu(\Omega) = \mu(A) + \mu(A^c)$
   but can only move $\mu(A)$ over if finite

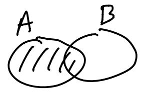
2. SAME by definition

3. If $A \subset B$ & $\mu(A) < \infty$ then
   $\mu(B - A) = \mu(B) - \mu(A)$.

4. SAME, since we still have
   $\mu(B) = \underbrace{\mu(B - A)}_{\ge 0} + \mu(A)$

Probability vrs Measure

5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

use  $= A \cup (B - A \cap B)$

and 3.

6. $P\left(\underbrace{\bigcup_{k=1}^{n} A_k}_{\text{can overlap}}\right) \le \sum_{k=1}^{n} P(A_k)$

use 5 & induction.

7. If $A_n \uparrow A \in \mathcal{F}_0$ then $P(A_n) \uparrow P(A)$

Note: property 7 is called continuous from below

Proof:

$P(A_n) = P\left(\bigcup_{k=1}^{n} A_k\right)$ since $A_1 \subset A_2 \subset \dots$

$= P\left(\bigcup_{k=1}^{n} [A_k - \underbrace{A_{k-1}}]\right)$

Same as $A_k - (A_1 \cup \dots \cup A_k)$.
So dissoint

$= \sum_{k=1}^{n} P(A_k - A_{k-1})$, $A_0 := \phi$

$\uparrow \sum_{k=1}^{\infty} P(A_k - A_{k-1})$

since P is a Prop measure

$= P\left(\bigcup_{k=1}^{\infty} [A_k - A_{k-1}]\right)$

$= P(A)$

5. $\mu(A \cup B) = \mu(A) + \mu(B - A \cap B)$

$= \mu(A) + \mu(B) - \mu(A \cap B)$

if $\mu(A \cap B) < \infty$

6. SAME. For proof use the first part of 5 & induction

7. SAME, same proof too

disjointizing technique

Probability vrs Measure

8. If $A_n \downarrow A \in \mathcal{F}_0$ then $P(A_n) \downarrow P(A)$

Called <u>Continuity from above</u>

<u>Proof</u>

First Note that $A_n \downarrow A \iff A_n^c \uparrow A^c$

By 7. we have

$$\underbrace{P(A_n^c)}_{=1-P(A_n)} \uparrow \underbrace{P(A^c)}_{=1-P(A)}$$

$\therefore \; P(A_n) \downarrow P(A)$.

Called countable subadditivity

9. $P\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} P(A_k)$.

<u>Proof</u>: $\bigcup_{k=1}^{n} A_k \uparrow \bigcup_{k=1}^{\infty} A_k$

$$\therefore \; P\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_n P\left(\bigcup_{k=1}^{n} A_k\right)$$

$$= \lim_n \sum_{k=1}^{n} P(A_k - A_1 \cup \cdots \cup A_{k-1})$$

$$\leq \lim_n \sum_{k=1}^{n} P(A_k)$$

8. If $A_n \downarrow A \in \mathcal{F}_0$ and $\mu(A_m) < \infty$ for some m

then $\mu(A_n) \downarrow \mu(A)$.

<u>Note</u>: the extra condition is necessary.

e.g. $\Omega = \mathbb{Z}$, $\mathcal{F}_0 = 2^\Omega$, μ = counting measure

$A_n = \{n, n+1, \ldots\} \downarrow A = \emptyset$ but

$$\underbrace{\mu(A_n)}_{\infty} \not\downarrow \underbrace{\mu(A)}_{=0}$$

<u>Proof</u>:

Let $n \geq m$ & note $A_m \supset A_n \supset A$.

$\therefore \; \mu(A) \leq \mu(A_n) \leq \mu(A_m) < \infty$

Define $P(B) := \dfrac{\mu(B \cap A_m)}{\mu(A_m)}$

P is a prob measure on $(\Omega, \mathcal{F}_0)$.

$\therefore \; P(A_n) \downarrow P(A)$ by 8 for Probs

$\therefore \; \dfrac{\mu(A_n \cap A_m)}{\cancel{\mu(A_m)}} \downarrow \dfrac{\mu(A \cap A_m)}{\cancel{\mu(A_m)}}$

$\therefore \; \mu(A_n) \downarrow \mu(A)$ since $A_n \cap A_m = A_n$ & $A \cap A_m = A$.

9. SAME. SAME proof too.

Probability vrs Measure

Thm (uniqueness)

Let P & Q be two prob measures on $(\Omega, \sigma\langle \mathcal{P} \rangle)$. If

(i) $P = Q$ on $\mathcal{P}$

(ii) $\mathcal{P}$ is a π-sys

Then $P = Q$ on $\sigma\langle \mathcal{P}\rangle$.

Proof: Use "good sets".

Let $\mathcal{G} = \{A \in \sigma\langle\mathcal{P}\rangle : P(A) = Q(A)\}$

we want to show $\sigma\langle \mathcal{P}\rangle \subset \mathcal{G}$.

By Dynkin's π-λ extension of goodsets we need $\underset{\pi\text{-system}}{\mathcal{P}} \subset \underset{\lambda\text{-sys.}}{\mathcal{G}}$

Clearly $\mathcal{P} \subset \mathcal{G}$ & $\mathcal{P}$ is a π-sys.

To show $\mathcal{G}$ is a λ-sys just Note:

- $\Omega \in \mathcal{G}$, since $P(\Omega) = Q(\Omega) = 1$
- $A \in \mathcal{G} \Rightarrow P(A) = Q(A)$
  $\Rightarrow P(A^c) = Q(A^c)$
  $\Rightarrow A^c \in \mathcal{G}$
- $\underbrace{A_1, A_2 \ldots}_{\text{disjoint}} \in \mathcal{G} \Rightarrow P\left(\bigcup_k A_k\right) = \sum_{k=1}^{\infty} P(A_k)$
  $= \sum_{k=1}^{\infty} Q(A_k)$
  $= Q\left(\bigcup_k A_k\right)$
  $\Rightarrow \bigcup_k A_k \in \mathcal{G}$.

QED.

Thm (uniqueness)

Let μ & ν be two measures on $(\Omega, \sigma\langle\mathcal{P}\rangle)$. If

(i) $\mu = \nu$ on $\mathcal{P}$

(ii) $\mathcal{P}$ is a π-sys

(iii) μ, ν are σ-finite over $\mathcal{P}$

Then $\mu = \nu$ on $\sigma\langle\mathcal{P}\rangle$.

Note: The extra condition is necessary.

For example, $\Omega = (0,1]$, $\mathcal{P} = \{(0,x] : 0 < x \leq 1\}$,

$\mu(A) = \sum_{\omega \in \Omega} I_A(\omega)$ i.e. counting measure

$\nu(A) = \sum_{\omega \in \Omega} I_{A \cap \text{rational}}(\omega)$

μ & ν both agree on $\mathcal{P}$ but not on $\sigma\langle\mathcal{P}\rangle$

since $\mu(\text{irrationals}) = \infty$

$\nu(\text{irrationals}) = 0$

Notice this set is in $\mathcal{B}^{(0,1]}$

Proof:

Let $C_1, C_2, \ldots \in \mathcal{P}$ which cover Ω and $\mu(C_k), \nu(C_k) < \infty$. Define

$$P_k(A) = \begin{cases} \mu(A \cap C_k)/\mu(C_k) & \text{if } \mu(C_k) > 0 \\ 0 & \text{o.w.} \end{cases}$$

$$Q_k(A) = \begin{cases} \nu(A \cap C_k)/\nu(C_k) & \text{if } \nu(C_k) > 0 \\ 0 & \text{o.w.} \end{cases}$$

case 1: $\mu(C_k) = 0$

$\therefore \nu(C_k) = 0$

$\therefore P_k = Q_k$ on $\sigma\langle\mathcal{P}\rangle$

Case 2: $\mu(C_k) > 0$

$\therefore \nu(C_k) > 0$ & both P_k, Q_k are prob measures.

Since $A \in \mathcal{P} \Rightarrow A \cap C_k \in \mathcal{P}$ we have $P_k = Q_k$ on $\mathcal{P}$.

Both P_k & Q_k are prob measures.

$\therefore \quad P_k = Q_k$ on $\sigma\langle \mathcal{P} \rangle$

Now stitch P_k & Q_k together

$$\Omega = \bigcup_{k=1}^{\infty} C_k = \bigcup_{k=1}^{\infty} C_k - C_1 \cup \cdots \cup C_{k-1}$$

$$= \bigcup_{k=1}^{\infty} \underbrace{C_k \cap C_1^c \cap \cdots \cap C_{k-1}^c}_{\text{disjoint cover ... but not in } \mathcal{P}.}$$

$$\therefore \mu(A) = \mu\left(\bigcup_{k=1}^{\infty} A \cap C_k \cap C_1^c \cap \cdots \cap C_{k-1}^c\right)$$

$$= \sum_{k=1}^{\infty} \mu(A \cap C_k \cap C_1^c \cap \cdots \cap C_{k-1}^c)$$

$$= \sum_{k=1}^{\infty} \mu[C_k] P_k(A \cap C_1^c \cap \cdots \cap C_{k-1}^c)$$

$$= \sum_{k=1}^{\infty} \nu[C_k] Q_k(A \cap C_1^c \cap \cdots \cap C_{k-1}^c)$$

$$= \nu(A)$$

QED

Continuity is equivalent to countable additivity for P

The next thm is only needed for probability & is useful for showing a P is measurable

Thm:

If P is a finitely additive prob on a field $\mathcal{F}_0 \subset 2^{\Omega}$ then the following are equivalent

(i) P is a prob measure

(ii) P is continuous from below

(iii) P is continuous from below

(iv) $\forall A_1, A_2, \ldots \in \mathcal{F}_0$ s.t. $A_n \downarrow \emptyset$ one has $\underbrace{P(A_n) \downarrow 0}_{\text{continuous from above at } \emptyset}$

Proof:

we already have

(i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv).

show (iv) $\overset{(a)}{\Rightarrow}$ (ii) $\overset{(b)}{\Rightarrow}$ (i)

For (a) suppose (iv) & let $\underbrace{A_n}_{\text{all in } \mathcal{F}_0} \uparrow A$

$\therefore \quad A_n^c \downarrow A^c$

$\therefore \quad A_n^c - A^c \downarrow \emptyset$

$\therefore \quad P(A_n^c - A^c) \downarrow 0$ by (iv)

$\therefore \quad P(A_n^c) - P(A^c) \downarrow 0$, since $A_n \subset A$ so that $A^c \subset A_n^c$

$\therefore \quad -P(A_n) + P(A) \downarrow 0$

$\therefore \quad P(A_n) \uparrow P(A)$

$\therefore$ (ii) holds

Note: these properties follow by FAP

(9)

For (b) suppose (ii) holds

Let $\underbrace{A_1, A_2, \dots}_{\text{disjoint}} \in \mathcal{F}_0$ & $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_0$

$$\therefore P\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_n \uparrow P\left(\bigcup_{k=1}^{n} A_k\right)$$
$$= \lim_n \sum_{k=1}^{n} P(A_k)$$
$$= \sum_{k=1}^{\infty} P(A_k)$$

QED.

Null & Negligible sets

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

- A is μ-null iff $A \in \mathcal{F}$ & $\mu(A) = 0$.
- A is μ-negligible iff $\exists$ a μ-null cover of A.
- $(\Omega, \mathcal{F}, \mu)$ is complete iff all μ-neg subsets of Ω are in $\mathcal{F}$
- The completion of $\mathcal{F}$ w.r.t μ is
$$\overline{\mathcal{F}} = \sigma\langle \mathcal{F}, \mathcal{N}_\mu \rangle$$
where $\mathcal{N}_\mu := \{N \subset \Omega : N \text{ is } \mu\text{-neg}\}$.

(10)

Once we proove the existance of Lebesgue measure $\mathcal{L}$ on $\mathbb{R}$

$$\overline{\mathcal{B}^{\mathbb{R}}} := \sigma\langle \mathcal{B}^{\mathbb{R}}, \mathcal{N}_{\mathcal{L}} \rangle$$
$=$ Lebesgue measurable sets of $\mathbb{R}$

where $\mathcal{B}^{\mathbb{R}}$ = Borel measurable sets of $\mathbb{R}$.

The following thm shows how to extend $(\Omega, \mathcal{F}, \mu)$ to $(\Omega, \overline{\mathcal{F}})$.

claim: If $(\Omega, \mathcal{F}, \mu)$ is a measure space then $\sigma\langle \mathcal{F}, \mathcal{N}_\mu \rangle = \underbrace{\{F \cup N : F \in \mathcal{F}, N \in \mathcal{N}_\mu\}}_{\text{"RHS"}}$

Proof:

The inclusion "$\supset$" is trivial.

For "$\subset$" use good sets.

- Clearly $\mathcal{F}, \mathcal{N}_\mu \subset$ RHS since $\emptyset \in \mathcal{F}$ & $\emptyset \in \mathcal{N}_\mu$
- $\emptyset \in$ RHS.
- RHS is closed under countable unions since $\bigcup_k (F_k \cup N_k) = \underbrace{\left(\bigcup_k F_k\right)}_{\in \mathcal{F}} \cup \underbrace{\left(\bigcup_k N_k\right)}_{\in \mathcal{N}_\mu}$
- $A \in$ RHS $\Rightarrow A = F \cup N$ for a μ-null $C \supset N$
$$\Rightarrow A^c = F^c \cap N^c$$
$$= (F^c \cap N^c) \cap (C \cup C^c)$$
$$= \underbrace{(F^c \cap N^c \cap C^c)}_{= C^c} \cup \underbrace{(F^c \cap N^c \cap C)}_{\in \mathcal{N}_\mu \text{ since covered by } C}$$

since $C^c \subset N^c$
$\therefore C^c \cap N^c = C^c$

$\Rightarrow A^c \in$ RHS $\in \mathcal{F}$

QED

Now we can define $\bar\mu$ on $\bar{\mathcal{F}}$

by $$\bar\mu(\underset{\in\mathcal{F}}{F}\cup\underset{\in\mathcal{N}_\mu}{N})=\mu(F)$$

Claim:

(i) $\bar\mu$ is well defined

(ii) $\bar\mu$ is a measure on $(\Omega,\bar{\mathcal{F}})$ extending μ on $(\Omega,\mathcal{F})$.

(iii) $\bar\mu$ is unique

(iv) $(\Omega,\bar{\mathcal{F}},\bar\mu)$ is complete.

Proof:

For (i): Suppose $F_1\cup N_1=F_2\cup N_2$

$\therefore\ F_1-F_2\subset N_2$ & $F_2-F_1\subset N_1$

For suppose $\omega\in F_1-F_2$

$\therefore\ \omega\in F_1$ & $\omega\notin F_2$

$\therefore\ \omega\in\underbrace{F_1\cup N_1}_{=F_2\cup N_2}$ & $\omega\notin F_2$

$\therefore\ \omega\in N_2$

$\therefore\ \mu(F_1-F_2)=\mu(F_2-F_1)=0$

$$\begin{aligned}\therefore\ \mu(F_1)&\le\mu(F_1\cup F_2)\\&=\mu(F_2\cup(F_1-F_2))\\&=\mu(F_2)+\underbrace{\mu(F_1-F_2)}_{=0}\end{aligned}$$

& $\mu(F_2)\le\mu(F_1)$ similarly

$\therefore\ \mu(F_1)=\mu(F_2)$

$\therefore\ \bar\mu(F_1\cup N_1)=\bar\mu(F_2\cup N_2)$.

For (ii): only need to show

$$\bar\mu\Big(\underbrace{\bigcup_k(F_k\cup N_k)}_{\text{disjoint over }k}\Big)=\bar\mu\Big(\Big[\underbrace{\bigcup_k F_k}_{\in\mathcal{F}}\Big]\cup\Big[\underbrace{\bigcup_k N_k}_{\in\mathcal{N}_\mu}\Big]\Big)$$

$$:=\mu\Big(\bigcup_k F_k\Big)$$

$F_k\cup N_k$ disjoint implies F_k's disjoint

$$=\sum_k\mu(F_k)$$

$$=\sum_k\bar\mu(F_k\cup N_k)$$

For (iii):

Let ν be a measure on $(\Omega,\bar{\mathcal{F}})$ s.t. $\nu=\mu$ on $\mathcal{F}$.

$\therefore\ \nu=\mu$ on $\mathcal{N}_\mu$

$\therefore\ \nu(F\cup N)=\nu(F)=\mu(F)=\bar\mu(F\cup N)$

$\ge$ by "increasing"

$\le$ by sub-additivity

$\therefore\ \nu=\mu$ on $\bar{\mathcal{F}}$.

For (iv) let $\bar N\subset\Omega$ be $\bar\mu$-neg

$\therefore\ \bar N\subset F\cup N\subset F\cup A$

for $F\in\mathcal{F}$ & $N\in\mathcal{N}_\mu$ s.t. $0=\bar\mu(F\cup N)=\mu(F)$

for $A\in\mathcal{F}$ s.t. $\mu(A)=0$

$\therefore\ \bar N$ is μ-neg so $\bar{\mathcal{F}}$ contains all $\bar\mu$-neg sets

QED

What are we going to use measures for

Random variables will be maps

$$\Omega \xrightarrow{X} \mathbb{R}$$

where $(\Omega, \mathcal{F}, P)$ is a prob space & X is $\mathcal{F}$-measurable.

$$EX := \int_\Omega X(\omega)\, dP(\omega)$$

integration w.r.t. measure P

For two measures P & Q on $(\Omega, \mathcal{F})$ the Radon-Nikodym derivative $\dfrac{dPX^{-1}}{dQX^{-1}}$ will be the likelihood ratio of P to Q for X.

Radon-Nikodym derivatives will also give use conditional expected values, etc.

why does probability seem so rich a subject when its essentially just measure theory under the constraint $\mu(\Omega) = 1$

I think the answer is that once $\mu(\Omega) = 1$ you can understand μ as modeling a random draw $\omega \in \Omega$. So the extra assumption essentially gives an isomorphism from the abstract $(\Omega, \mathcal{F}, \mu)$ to a physical random procedure.

So a solution can be made on the abstract structure $(\Omega, \mathcal{F}, \mu)$ <u>or</u> the physical random draws it models.