

Lecture 4: Measures

(1)

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 Probability measure on (Ω, \mathcal{F}_0) if

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

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

μ is a measure on (Ω, \mathcal{F}_0) if

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

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 $B_0^{(ori)}$ then \exists an extended prob measure on $\sigma(B_0^{(ori)})$).

Note: Any Prob measure is also a measure since

$$\begin{aligned} \mathcal{F} \text{ is not empty} \Rightarrow & \exists A \in \mathcal{F} \\ \Rightarrow & \Omega = A \cup A^c \in \mathcal{F} \\ \phi = & \Omega^c \in \mathcal{F} \\ \Rightarrow 1 = & P(\Omega) = P(\phi) + \sum_{i=1}^{\infty} P(A_i) \\ \Rightarrow P(\phi) = & 0 \end{aligned}$$

Def: (Ω, \mathcal{F}) is a measurable space if \mathcal{F} is a σ -field over Ω

Def: (Ω, \mathcal{F}, P) is called a Probability Space.

σ -field Prob measure on \mathcal{F}

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

measure on \mathcal{F}

Def:

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

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

μ is infinite if $\mu(\Omega) = \infty$.

μ is σ -finite over $\mathcal{A} \subset \mathcal{F}$ if

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

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

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

μ is σ -finite if μ is σ -finite over \mathcal{F} .

Remark: 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 $\mu(A) = \int_A dP$ where P is a prob (The probabilist's world view of measure theory).

Remark: \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 not a measure if Ω is countable.

Probability

v/s

Measure

Example

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

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

$$P(A) = \sum_{w \in \Omega} I_A(w) = \# \text{ of elements in } A$$

indicator of 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...

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

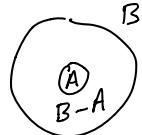
$$\begin{aligned} \text{since } P(\Omega) &= P(A \cup A^c) \\ &= P(A) + P(A^c) \end{aligned}$$

$$2. P(\emptyset) = 0$$

3. If $A \subset B$ then

$$P(B - A) = P(B) - P(A).$$

$$\text{since } B = (B - A) \cup A$$



$$4. A \subset B \implies P(A) \leq P(B)$$

$$\text{by 3 since } 0 \leq P(B - A) \\ = P(B) - P(A)$$

Example

Ω is any set

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

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

μ is called counting measure.

Basic Properties

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

Then...

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

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

$$2. \text{ SAME by definition}$$

$$3. \text{ If } A \subset B \text{ & } \mu(A) < \infty \text{ then}$$

$$\mu(B - A) = \mu(B) - \mu(A).$$

4. SAME, since we still have

$$\mu(B) = \underbrace{\mu(B - A)}_{> 0} + \mu(A)$$

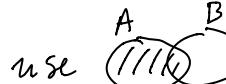
Probability

vrs

Measure

(2)

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

use  = $A \cup (B - A \cap B)$
and 3.

$$6. P\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k)$$

can overlap

use 5 &
induction.

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

Note: property 7 is called continuous
from below

Proof:

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

$$= P\left(\bigcup_{k=1}^n \underbrace{[A_k - A_{k-1}]}_{\substack{\text{Same as } A_n - (A_1 \cup \dots \cup A_{k-1}) \\ \text{so disjoint}}}\right)$$

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

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

$$\text{since } P \text{ is measure}^{\exists} \quad = P\left(\bigcup_{k=1}^{\infty} [A_k - A_{k-1}]\right)$$

$$= P(A)$$

$$\left. \begin{aligned} 5. \mu(A \cup B) &= \mu(A) + \mu(B - A \cap B) \\ &= \mu(A) + \mu(B) - \mu(A \cap B) \\ &\uparrow \\ &\text{if } \mu(A \cap B) < \infty \end{aligned} \right\}$$

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

7. SAME, same proof too

disjointing
technique

Probability

vrs

Measure

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

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

Note: the extra condition is necessary.

e.g. $\Omega = \mathbb{Z}$, $\mathcal{F}_0 = 2^{\mathbb{Z}}$, μ = counting measure
 $A_n = \{n, n+1, \dots\} \downarrow A = \emptyset$ but

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

Proof:

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

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

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

$$\therefore \frac{\mu(A_n \cap A_m)}{\mu(A_m)} \downarrow \frac{\mu(A \cap A_m)}{\mu(A_m)}$$

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

9. SAME. SAME proof too.

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

Proof: $\bigcup_{p=1}^n A_p \uparrow \bigcup_{p=1}^{\infty} A_p$

$$\begin{aligned} \therefore P\left(\bigcup_{p=1}^{\infty} A_p\right) &= \lim_n P\left(\bigcup_{p=1}^n A_p\right) \\ &= \lim_n \sum_{p=1}^n P(A_p - A_1 \cup \dots \cup A_{p-1}) \\ &\leq \lim_n \sum_{p=1}^n P(A_p) \end{aligned}$$

Probability

vrs

Measure

Thm (uniqueness)

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

- (i) $P = Q$ on \mathcal{P}
- (ii) \mathcal{P} is a π -sys

Then $P = Q$ on $\sigma(\mathcal{P})$.

Proof: Use "good sets".

$$\text{Let } \mathcal{Y} = \{A \in \sigma(\mathcal{P}) : P(A) = Q(A)\}$$

We want to show $\sigma(\mathcal{P}) \subset \mathcal{Y}$.

By Dynkin's π - λ extension of good sets

we need $P \subset \mathcal{Y}$
 \downarrow
 $\pi\text{-system} \quad \lambda\text{-sys.}$

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

To show \mathcal{Y} is a λ -sys just note:

• $\Omega \in \mathcal{Y}$, since $P(\Omega) = Q(\Omega) = 1$

• $A \in \mathcal{Y} \Rightarrow P(A) = Q(A)$

$$\Rightarrow P(A^c) = Q(A^c)$$

$$\Rightarrow A^c \in \mathcal{Y}$$

• $A_1, A_2, \dots \in \mathcal{Y} \Rightarrow P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{\infty} Q(A_k) = Q\left(\bigcup_{k=1}^{\infty} A_k\right)$

disjoint

$$\Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{Y}.$$

$\Omega \in \mathcal{D}$.

Thm (uniqueness)

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

Note: The extra condition is necessary.

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

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

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

$\mu \neq \nu$ both agree on \mathcal{P} but not on $\sigma(\mathcal{P})$

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

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

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

Proof:

Let $C_1, C_2, \dots \in \mathcal{P}$ 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: $\mu(C_k) = 0$

$$\therefore \nu(C_k) = 0$$

$$\therefore P_k = Q_k \text{ on } \sigma(\mathcal{P})$$

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.

$\therefore P_k = Q_k$ on $\sigma(\mathcal{P})$

Now stitch P_k & Q_k together

$$\begin{aligned} \mathcal{D} &= \bigcup_{k=1}^{\infty} C_k = \bigcup_{k=1}^{\infty} C_k - C_1 \cup \dots \cup C_{k-1} \\ &= \bigcup_{k=1}^{\infty} C_k \cap C_1^c \cap \dots \cap C_{k-1}^c \\ &\quad \text{disjoint cover} \\ &\quad \dots \text{but not in } \mathcal{P}. \end{aligned}$$

$$\begin{aligned} \therefore \mu(A) &= \mu\left(\bigcup_{k=1}^{\infty} A \cap C_k \cap C_1^c \cap \dots \cap C_{k-1}^c\right) \\ &= \sum_{k=1}^{\infty} \mu(A \cap C_k \cap C_1^c \cap \dots \cap C_{k-1}^c) \\ &= \sum_{k=1}^{\infty} \nu[C_k] P_k(A \cap C_1^c \cap \dots \cap C_{k-1}^c) \\ &= \sum_{k=1}^{\infty} \nu[C_k] Q_k(A \cap C_1^c \cap \dots \cap C_{k-1}^c) \\ &= \nu(A) \end{aligned}$$

QED

(7)

Continuity is equivalent to countable additivity for P

(8)

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 $\mathcal{F}_0 \subset 2^{\mathbb{N}}$ then the following are equivalent

- (i) P is a prob measure
- (ii) P is continuous from below
- (iii) P is continuous from above

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

Proof:

We already have

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

Show (iv) $\stackrel{(a)}{\Rightarrow}$ (ii) $\stackrel{(b)}{\Rightarrow}$ (i)

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

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

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

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

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

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

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

\therefore (ii) holds

Note: these properties follow by FAP

For (b) suppose (ii) holds (9)

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

$$\begin{aligned} P\left(\bigcup_{k=1}^{\infty} A_k\right) &= \lim_{n \rightarrow \infty} P\left(\bigcup_{k=1}^n A_k\right) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n P(A_k) \\ &= \sum_{k=1}^{\infty} P(A_k) \quad \text{QED} \end{aligned}$$

Null & Negligible sets

Definitions: Let $(\mathbb{R}, \mathcal{F}, \mu)$ 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 .
- $(\mathbb{R}, \mathcal{F}, \mu)$ is complete iff all μ -neg subsets of \mathbb{R} are in \mathcal{F}
- The completion of \mathcal{F} w.r.t μ is

$$\bar{\mathcal{F}} = \sigma(\mathcal{F}, \eta_\mu)$$

where $\eta_\mu := \{N \subset \mathbb{R} : N \text{ is } \mu\text{-neg}\}$.

Once we prove the existence of Lebesgue measure \mathcal{L} on \mathbb{R} (10)

$$\bar{\mathcal{B}}^\mathbb{R} := \sigma(B^\mathbb{R}, \eta_\mathcal{L})$$

= Lebesgue measurable sets of \mathbb{R}

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

The following thm shows how to extend $(\mathbb{R}, \mathcal{F}, \mu)$ to $(\mathbb{R}, \bar{\mathcal{F}})$.

claim: If $(\mathbb{R}, \mathcal{F}, \mu)$ is a measure space then $\sigma(\mathcal{F}, \eta_\mu) = \{F \cup N : F \in \mathcal{F}, N \in \eta_\mu\}$

Proof:

The inclusion " \supset " is trivial. "RHS".

For " \subset " use good sets.

- Clearly $\mathcal{F}, \eta_\mu \subset \text{RHS}$ since $\emptyset \in \mathcal{F}$ & $\emptyset \in \eta_\mu$
- $\emptyset \in \text{RHS}$.
- RHS is closed under countable unions since $\bigcup_k (F_k \cup N_k) = (\bigcup_k F_k) \cup (\bigcup_k N_k) \in \text{RHS}$

- $A \in \text{RHS} \Rightarrow A = F \cup N$ for a μ -null $C \supset N$

$$A^c = F^c \cap N^c$$

$$= (F^c \cap N^c) \cap (C \cup C^c)$$

$$\begin{aligned} &\text{since } C^c \subset N^c \\ &\therefore C^c \cap N^c = C^c \\ &= (F^c \cap N^c \cap C^c) \cup (F^c \cap N^c \cap C) \end{aligned}$$

$\in \eta_\mu$ since covered by C

$$A^c \in \text{RHS} \in \bar{\mathcal{F}}$$

QED

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

(11)

$$\text{by } \bar{\mu}(F_{UN}) = \mu(F)$$

$\uparrow \quad \downarrow$
 $\epsilon F \quad \epsilon \mathcal{N}_\mu$

Claim:

- (i) $\bar{\mu}$ is well defined
- (ii) $\bar{\mu}$ is a measure on (\mathcal{R}, \bar{F}) extending μ on (\mathcal{R}, F) .
- (iii) $\bar{\mu}$ is unique
- (iv) $(\mathcal{R}, \bar{F}, \bar{\mu})$ is complete.

Proof:

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

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

For suppose $w \in F_1 - F_2$

$$\therefore w \in F_1 \quad \& \quad w \notin F_2$$

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

$$\therefore w \in N_2$$

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

$$\therefore \mu(F_1) \leq \mu(F_1 \cup F_2)$$

$$= \mu(F_2 \cup (F_1 - F_2))$$

$$= \mu(F_2) + \mu(F_1 - F_2)$$

$\underbrace{\quad}_{\sim}$

$$\& \mu(F_2) \leq \mu(F_1) \text{ similarly}$$

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

(12)

$$\therefore \bar{\mu}(F_{UN}) = \bar{\mu}(F_2 \cup N_2).$$

For (ii): only need to show

$$\bar{\mu}\left(\bigcup_k (F_k \cup N_k)\right) = \bar{\mu}\left(\left[\bigcup_k F_k\right] \cup \left[\bigcup_k N_k\right]\right)$$

disjoint over k ϵF $\epsilon \mathcal{N}_\mu$

$$:= \mu\left(\bigcup_k F_k\right)$$

$$\begin{aligned} &\overbrace{F_k \cup N_k \text{ disjoint}}^{\text{implies } F_k \text{'s disjoint}} = \sum_k \mu(F_k) \\ &= \sum_k \bar{\mu}(F_k \cup N_k) \end{aligned}$$

For (iii):

Let ν be a measure on (\mathcal{R}, \bar{F})

s.t. $\nu = \mu$ on F .

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

$$\therefore \nu(F_{UN}) = \nu(F) = \mu(F) = \bar{\mu}(F_{UN})$$

\uparrow

\geq by "increasing"

\leq by sub-additivity

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

For (iv): let $\bar{N} \subset \mathcal{R}$ be $\bar{\mu}$ -neg

$$\therefore \bar{N} \subset F_{UN} \subset F \cup A$$

\nearrow for $F \in F \& N \in \mathcal{N}_\mu$ \nwarrow for $A \in F$ s.t. $\mu(A) = 0$

$\underbrace{0 = \bar{\mu}(F_{UN})}_{\text{s.t.}} = \mu(F)$

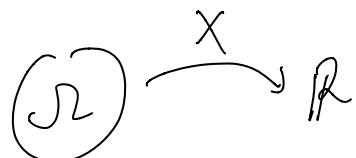
$\therefore \bar{N}$ is μ -neg so \bar{F} contains all $\bar{\mu}$ -neg sets

QED

What are we going to use
measures for

(13)

Random variables will be maps



where (Ω, \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
 (Ω, \mathcal{F}) the Radon-Nikodym
derivative $\frac{dP|X^{-1}}{dQ|X^{-1}}$ will be

The likelihood ratio of P to Q
for X .

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

(14)

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
w.e. Ω . 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 $(\Omega, \mathcal{F}, \mu)$
or the physical random
draws it models.