Instructor: Dr. Osonde Osoba

Contact info:

office: EEB 420

Hours: MW $17^{00} - 18^{00}$

email: osondeos@usc.edu

class website: blackboard.usc.edu

TA: Yung-Hsuan Chao

email: yunghsuc@usc.edu

office: EE 434 hours: T 2–3pm, F 4–6pm

This week:

- Definition of a Probability Space $(\Omega, \mathcal{A}, P)$
  - Sets (Ω)
    - Operations on sets, logic.
    - mappings and functions between sets
  - Sigma Algebras $(\mathcal{A})$
  - Probability Measures (P)
    - Axioms of Probability

HW 1: L-G: #2-117 (d).

(Due 9-Sept) Gubner: #15 (pg. 50)

Theme:

Probability is about "measuring the relative sizes of sets."
It is a more applied version of measure theory.

The basic structure is as follows:
Suppose an experiment yields random outcomes
① - "Outcomes" are elements of the Sample space (Ω) which is the set of all possible outcomes.

② - "Events" are measurable subsets of the sample space.
③ - The probability measure (P) sizes up events relative to the size of the whole Sample space, Ω.

Sets (or events) are the fundamental entities in probability theory...

I/ Sets :

A set is a collection of elements.

e.g $A = \{a, e, i, o, u\}$

$B = \{x \mid x \text{ is a vowel}\}$

$\mathcal{T} = \{\phi, \{a\}\}$

There is often an implicit assumption of a universal set X i.e. a set that contains all elements in the "universe of discourse".

Operations on sets :

i/ Union : $(\cup)$

$$A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\}$$

or $$\bigcup_{i \in I} E_i = \{x \in X \mid \exists i \in I,\ x \in E_i\}$$

ii/ Intersection : $(\cap)$

$$A \cap B = \{x \in X \mid x \in A \text{ and } x \in B\}$$

or $$\bigcap_{i \in I} E_i = \{x \in X \mid \forall i \in I,\ x \in E_i\}$$

iii) Complement:

$$A' \text{ or } A^c = \{x \in X \mid x \notin A\}$$

iv) Set Difference:

$$A - B = A \cap B^c$$
$$= \{x \in X \mid x \in A \text{ and } x \notin B\}$$

Relations between Sets:

i) Subsethood:

$$A \subset B \text{ iff } \forall x \in A,\ x \in B$$

[or $x \in A \Rightarrow x \in B$]

ii) Equality:

$$A = B \text{ iff } A \subset B \text{ and } B \subset A$$

[or $x \in A \Leftrightarrow x \in B$]

There is a relationship between sets and propositional logic:

Union: $A \cup B \longleftrightarrow p \vee q$

Intersection: $A \cap B \longleftrightarrow p \wedge q$

Subsethood: $A \subset B \longleftrightarrow p \Rightarrow q$

p	q	$p \vee q$	$p \wedge q$	$p \Rightarrow q$	
1	1	1	1	1	
0	1	1	0	1	[e.g $\phi \subset A$]
1	0	1	0	0	
0	0	0	0	1	

Power set (2^X): Every set X has a derived set of all its subsets 2^X

e.g $X = \{a, b\}$

$\Rightarrow 2^X = \{\phi, \{a\}, \{b\}, X\}$.

Ex1: Prove or Disprove

$[A \cup B = B] \Longrightarrow [A \subset B]$

Proof: Assume $[A \cup B = B]$ is true

	Justification
(1) Pick $x \in A$	– Assumption
(2) $\therefore x \in A$ or $x \in B$	– Definition of or
(3) $\therefore x \in A \cup B$	– Def$_n$ of Union ($\cup$)
(4) $\therefore x \in B$	– Since $[A \cup B = B]$
$\therefore \forall x \in A, x \in B$	– Conclusion from (1)–(4)
i.e $[A \subset B]$ // QED	– Def$_n$ of Subsethood ($\subset$)

Ex2: Prove / Disprove:

$$[A \subset B] \Rightarrow [A \cup B = B]$$

Proof: Assume $[A \subset B]$ is true

Need to show: (a) $A \cup B \subset B$

(b) $B \subset A \cup B$

(b) is trivial to show [see step (1) – (3) in Ex1]

To show (a):	justification
(1) pick $x \in A \cup B$	– Assumption
(2) $\therefore$ $x \in A$ or $x \in B$	– Defn of union
(3) Assume $x \in A$	– Assumption to test subcase
(4) $\therefore$ $x \in B$	– Since $[A \subset B]$
(5) Assume $x \in B$	– 2^{nd} subcase
(6) $\therefore$ $x \in B$	– By assumption
(7) $\therefore$ $x \in B$	– results from Combine Subcases
$\therefore$ $\forall x \in A \cup B,\ x \in B$	– Conclusion
i.e $A \cup B \subset B$	– Defn of subsethood ($\subset$)

QED (a)

$$(a), (b) \Leftrightarrow [A \cup B = B]$$

Eg 3: [De Morgan's Law for Sets]

logic : $(p \vee q)^c = p^c \wedge q^c$

sets : $(A \cup B)^c = A^c \cap B^c$

Proof:

to show: i) $(A \cup B)^c \subset A^c \cap B^c$	justification
Pick $x \in (A \cup B)^c$	Assumption
$\therefore\ x \notin A \cup B$	defn: complement
$\therefore\ x \notin A$ and $x \notin B$	defn: $\cup$
$\therefore\ x \in A^c$ and $x \in B^c$	defn: complement
$\therefore\ x \in A^c \cap B^c$	defn: $\cap$

i.e $\therefore\ (A \cup B)^c \subset A^c \cap B^c$

ii) to show: $A^c \cap B^c \subset (A \cup B)^c$	justification
Pick $x \in A^c \cap B^c$	Assumption
$\therefore\ x \in A^c$ and $x \in B^c$	defn: $\cap$
$\therefore\ x \notin A$ and $x \notin B$	defn: complement
$\therefore\ x \notin (A \cup B)$	defn: $\cup$
$\therefore\ x \in (A \cup B)^c$ //	defn: complement

OR ii) **Proof by Contradiction**

Suppose $A^c \cap B^c \not\subset (A \cup B)^c$

$\Rightarrow \exists\, x \in A^c \cap B^c \ni x \notin (A \cup B)^c$

$\Rightarrow x \in A \cup B \Rightarrow [x \in A$ or $x \in B]$ ⟸ contradiction //

But $x \in A^c \cap B^c \Rightarrow [x \notin$ and $x \notin B]$

Function and Mappings between sets:

$$f : X \longrightarrow Y$$

Domain — Codomain/Range

A function (f) is a mapping between sets (X and Y) such that each element of the domain $x \in X$ maps to a unique element in the codomain $f(x) \in Y$.

A function is one-to-one (injective) if distinct inputs map to distinct outputs

i.e $\forall x_1, x_2 \in X, \; [x_1 \neq x_2] \Rightarrow [f(x_1) \neq f(x_2)]$

A function is onto (surjective) if every element in the range is the image of an element in the domain.

i.e $\forall y \in Y, \; \exists \; x \in X$ such that $y = f(x)$

A function is a one-to-one correspondence (bijective) if it is both one-to-one and onto.

A bijective function f has a well-defined inverse f^{-1}. $f(x) = y$, $f^{-1}(y) = x$ $\forall x \in X, y \in Y$

We can make the distinction between the familiar "point-to-point" action of functions and the behavior of set functions f and f^{-1} as defined below.

The <u>image</u> set function f:

$$f : 2^X \longrightarrow 2^Y$$

$$\boxed{f(A) = \{f(x) \mid x \in A\}}$$

$[A \subset X, \ f(A) \subset Y]$

The <u>preimage</u> (or pullback) set function:

$$f^{-1} : 2^Y \longrightarrow 2^X$$

$$\boxed{f^{-1}(B) = \{x \in X \mid f(x) \in B\}}$$

$[f^{-1}(B) \subset X, \ B \subset Y]$

The set functions above differ from the regular point-to-point function in that the preimage set function $f^{-1}(B)$ <u>always exists</u> even if f is not a bijection in the point-to-point sense.

Ex 4: $f : x \longmapsto x^2$

$A = [-1, 2] \qquad B = \{4, 9\}$

$\Rightarrow \; f(A) = [0, 4]$

$f^{-1}(B) = \{-2, 2, -3, 3\}$

Ex 5: Prove / Disprove:

$$[A \subset B] \Rightarrow [f^{-1}(A) \subset f^{-1}(B)]$$

Proof: Assume $[A \subset B]$ is true	justification
(1) pick $x \in f^{-1}(A)$	Assumption
(2) $\therefore$ $f(x) \in A$	defn of f^{-1}
(3) $\therefore$ $f(x) \in B$	Since $A \subset B$
(4) $\therefore$ $x \in f^{-1}(B)$	defn of f^{-1}
$\therefore$ $f^{-1}(A) \subset f^{-1}(B)$ // QED	defn of $\subset$

Ex 6: Prove / Disprove:

$$f(A \cap B) = f(A) \cap f(B)$$

Disprove: [by counterexample]

$f(x) = x^2$	$f(A) = [0, 9]$	$f(A) \cap f(B) = [0, 4]$
$A = [-3, 1]$	$f(B) = [0, 4]$	$\neq [0, 1]$
$B = [-1, 2]$	$f(A \cap B) = [0, 1]$	$\therefore f(A) \cap f(B) \neq f(A \cap B)$ //
$A \cap B = [-1, 1]$		

Ex 7 : [Probability by Intuition]

The experiment is 2 coin flips

$\Rightarrow$ The set of all possible outcomes is

$$\Omega = \{HH, HT, TH, TT\}$$

i) The event (E) of observing 1 tail ??

$$E = \{HT, TH\} \subset 2^{\Omega}$$

ii) The "probability" or relative size of E with respect to Ω ??

$$P(E) = \#E / \#\Omega = 2/4$$

$$= 1/2$$

The power set of the sample space 2^{Ω} is often "too big". So we often work with a smaller collection of event, $\mathcal{A}$. This collection of events must satisfy certain rules.

Defn: A sigma-algebra ($\mathcal{Q}$) is a subcollection of the power set 2^{Ω} that satisfies the following conditions: [CUT]

i) C: Closed under Complements

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

ii) U: Closed under Countable Unions

$$A_i \in \mathcal{Q} \;\; \forall i \;\; \Rightarrow \;\; \left(\bigcup_i A_i\right) \in \mathcal{Q}$$

iii) T: Total set

$$\Omega \in \mathcal{Q}$$

An event $A \subset \Omega$ is any element of a sigma algebra on Ω. Events are also called measurable sets.

$(\Omega, \mathcal{Q})$ is a measurable space iff $\mathcal{Q} \subset 2^{\Omega}$ is a sigma-algebra.

Examples of sigma-algebras on a sample space Ω include:

$\mathcal{Q} = \{\phi, \Omega\} \subset 2^{\Omega}$ [the trivial σ-alg]

$\mathcal{Q} = 2^{\Omega} \subseteq 2^{\Omega}$ [the discrete σ-alg]

Ex 8: [sigma-algebras on the 2-coin flip Ω]

$\Omega = \{TT, HH, TH, HT\}$

Examples of a σ-alg on Ω includes:

$\mathcal{Q}_1 = \{\phi, \{TT\}, \{HH, TH, HT\}, \Omega\}$

$\mathcal{Q}_2 = \{\phi, \{HH, TT\}, \{TH, HT\}, \Omega\}$

Define the probability measure (P)

Suppose $(\Omega, \mathcal{Q})$ is a measurable space then

$$P : \mathcal{Q} \longrightarrow [0, 1]$$

is a probability measure iff:

P is: [CAT]

i) CA: Countably Additive

given $\{A_k\}_k$ such that $A_i \cap A_j = \phi \quad \forall i \neq j$

then $$\boxed{P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)}$$

ii) T: Total set has unit mass

$$\boxed{P(\Omega) = 1}$$

Definition:

$(\Omega, \mathcal{Q}, P)$ is a **Probability Space**

iff $\begin{cases} \mathcal{Q} \text{ is a sigma algebra on } \Omega \\ P \text{ is a probability measure} \end{cases}$

Ex 9: [Probability Space on the 2-coin flip case]

$\Omega = \{HH, HT, TH, TT\}$

$\mathcal{Q} = \{\phi, \{HT, TH, TT\}, \{TT\}, \Omega\}$

$P: E \in \mathcal{Q} \longmapsto \#E / \#\Omega$

is $(\Omega, \mathcal{Q}, P)$ a probability space?

Ans: Need to check:

i) $\mathcal{Q}$ is a σ-algebra (CUT)

ii) P is a probability measure (CAT)

i) $\{HT, TH, TT\}^c = \{HH\} \notin \mathcal{Q}$

$\therefore$ $\mathcal{Q}$ is not a sigma-algebra

$\therefore$ $(\Omega, \mathcal{Q}, P)$ is not a probability space.

Qu: How do we generate a σ-algebra from an arbitrary set collection?

e.g. Generate a σ-algebra from the collection

$$\mathcal{Q} = \{\phi, \{TT\}, \{TH, HT, TT\}, \Omega\}.$$

Ans: for countable finite σ-alg, just add the missing subsets.

We need something more robust for larger sets (e.g infinite and/or uncountable sets)

Denote $\sigma(\mathcal{Q})$ as the σ-algebra generated by the subcollection $\mathcal{Q} \subset 2^{\Omega}$

$$\boxed{\sigma(\mathcal{Q}) = \bigcap_{\alpha} \mathcal{Q}_{\alpha}}$$

where $\mathcal{Q}_\alpha$ is a σ-algebra such that

$$\mathcal{Q} \subset \mathcal{Q}_\alpha \quad \forall \alpha.$$

C: Containment $\mathcal{Q} \subset \mathcal{Q}_\alpha \ \forall \alpha$

I: Intersection $\bigcap_\alpha \mathcal{Q}_\alpha$

A: σ-Algebras : $\mathcal{Q}_\alpha$ is a σ-alg $\forall \alpha$

[unions of σ-algs may not be σ-alg]

This gives us a means for describing probability spaces for continuous-valued sample spaces like $\Omega = \mathbb{R}$ or $\Omega = [0,1]$

The <u>Borel sigma-algebra</u> $\mathcal{B}(\mathbb{R})$ is the sigma-algebra generated by the collection of of intervals on $\mathbb{R}$:

$$\boxed{\mathcal{B}(\mathbb{R}) = \sigma\left(\{(-\infty, a] : a \in \mathbb{R}\}\right)}$$

<u>EX 10</u>: [a simple probability space on [0,1]]

$$\Omega = [0,1]$$
$$\mathcal{Q} = \mathcal{B}([0,1])$$
$$P(E) = \int_E 1 \, d\omega$$

- $\mathcal{Q}$ is a σ-algebra (CUT) by construction of the Borel sigma algebra on $[0,1]$
- $P(E)$ is CAT since:

<u>CA</u>: $\int_{\cup_i E_i} 1 \, d\omega = \sum_i \int_{E_i} 1 \, d\omega$

<u>T</u>: $\int_\Omega 1 \, d\omega = 1$ //

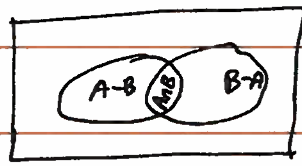
"Axioms" of Probability:

i) Addition Theorem:

$$P(A) + P(B) = P(A \cap B) + P(A \cup B)$$

proof:

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



All disjoint.

Also: $A = (A-B) \cup (A \cap B)$

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

$\therefore$ by CA:

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

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

$$\therefore \; P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ii) $P(A \cup B) \leq P(A) + P(B)$

since $P(A \cap B) \geq 0$

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

proof:

$A \cup A^c = \Omega$

$A \cap A^c = \phi$

$\therefore$ CA $\Rightarrow P(A \cup A^c) = P(A) + P(A^c)$

I $\Rightarrow P(A \cup A^c) = P(\Omega) = 1$

$\therefore \; P(A) + P(A^c) = 1$

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

QED

ii/: $P(\phi) = 0$

since $\phi = \Omega^c$ and $P(\Omega) = 1$

iii/: $A \subseteq B \Rightarrow P(A) \leq P(B)$

Proof:

$A \subseteq B \Rightarrow B = (B-A) \cup A$ [disjoint decomp.]

$\Rightarrow P(B) \stackrel{C.A.}{=} P(B-A) + P(A)$

$\therefore P(B) \geq P(A)$ since $P(B-A) \geq 0$

etc...

Closing Remarks:

i/ Countability:

A set Ω is countable if the same cardinality as the set of integers $\mathbb{Z}$ or a subset thereof. Examples of countable sets include:

$\mathbb{N} = \{0, 1, 2, 3, \cdots\}$ [natural numbers]

$\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z}\}$ [rational numbers]

$\Omega = \{a, b, c, d\}$ or any other discrete set of finite size

We use $|\Omega|$ to denote the cardinality of Ω.
Then $|\Omega| = |\mathbb{Z}|$ if there is a bijection (1-1, onto function) between Ω and $\mathbb{Z}$

Set cardinality obeys the following rules

(a) $A \subseteq B \implies |A| \leq |B|$

(b) $|A| < |2^A|$

Example of uncountable sets include:

$\mathbb{R} = \{\text{real numbers}\}$, $[0, 1]$

$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$, etc...

ii/ <u>Banach – Tarski Paradox</u> : (B-T)

This is an example of what could go wrong if we try to do measure theory (or probability, in particular) on the whole power set 2^Ω instead of using σ-algebras. This example only works in $\mathbb{R}^d$ $d \geq 3$

Ω

unit sphere in $\mathbb{R}^d$

$P(\Omega) = 1$

$\longrightarrow$ i/ Use <u>axiom of choice</u> to split Ω into special subsets.
ii/ Glue subsets back together without stretching ["isometric transformation"] $\longrightarrow$

we get

Ω_1 and Ω_2

$P(\Omega_1) = 1 = P(\Omega_2)$

i.e. $P(\Omega_1) + P(\Omega_2) = 2 \geq P(\Omega)$

$\boxed{\text{Moral: Use } \sigma\text{-algebras to avoid Pathologies like this.}}$