1179: Probability
Lecture 2 — Sets, Probability Axioms, and Continuity of Probability Function

Ping-Chun Hsieh (謝秉均)

September 17, 2021

This Lecture

1. Set Operations

2. Probability Axioms

3. Continuity of Probability Functions

- Reading material: Chapter 1.3-1.5

Review: Countable Union/Intersection

- Let $S_1, S_2, S_3 \cdots$ be a sequence of sets

3. $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n = \{x : x \in S_k, \text{ for infinitely many } k\}$
4. $\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} S_n = \{x : x \in S_k, \text{ for all except for finitely many } k\}$

- Show: $\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} S_n = \underbrace{\{x : x \in S_k, \text{ for all except for finitely many } k\}}_{=:A}$
- Proof:

(1) $\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} S_n \subseteq A$

Pick any $x \in \bigcup_{k=1}^{\infty} \underbrace{\bigcap_{n=k}^{\infty} S_n}_{=B_k}$

There must be some k such that

$x \in \bigcap_{n=k}^{\infty} S_n$

$\Rightarrow x \in A$

(2) $A \subseteq \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} S_n$

Pick any $y \in A$

Then, y appears in almost all S_k except for finitely many k

$\Rightarrow$ There must exist M such that $y \in S_k$, for all $k \geq M$

$\Rightarrow y \in \bigcap_{k=M}^{\infty} S_k$

$\Rightarrow y \in \bigcup_{n=k}^{\infty} S_n$

De Morgan's Laws

- Let S_1, S_2 be two sets

1. $\left(S_1 \cup S_2\right)^c = S_1^c \cap S_2^c$

2. $\left(S_1 \cap S_2\right)^c = S_1^c \cup S_2^c$

- Prove this by Venn diagram

De Morgan's Laws (General Case)

- Let $S_1, S_2, S_3 \cdots$ be a sequence of sets

1. $\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c$

2. $\left(\bigcap_n S_n\right)^c = \bigcup S_n^c$

Axioms of Probability

公理

Probability Axioms

⇒ simple fundamentals (definition)

- In a probabilistic model, we assign probability to events (How?)
- Axioms: rules to verify a probabilistic model
- Example: 8 axioms of vector space in linear algebra

Axiom	Meaning
Associativity of addition	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
Commutativity of addition	$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
Identity element of addition	There exists an element $\mathbf{0} \in V$, called the *zero vector*, such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.
Inverse elements of addition	For every $\mathbf{v} \in V$, there exists an element $-\mathbf{v} \in V$, called the *additive inverse* of $\mathbf{v}$, such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.
Compatibility of scalar multiplication with field multiplication	$a(b\mathbf{v}) = (ab)\mathbf{v}$ [nb 2]
Identity element of scalar multiplication	$1\mathbf{v} = \mathbf{v}$, where 1 denotes the multiplicative identity in F.
Distributivity of scalar multiplication with respect to vector addition	$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$
Distributivity of scalar multiplication with respect to field addition	$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

- Why are axioms useful?
- Can we prove axioms?

3 Axioms of Probability

A probability assignment is valid if:

1. $P(A) \geq 0$, for any event A (non-negativity)
2. $P(\Omega) = 1$
3. $A_1, A_2, \cdots$ is an infinite sequence of mutually exclusive events, then
$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

- Can we find $P(\emptyset) = ?$ 0 (By 3)
- $A_1, \cdots, A_n$ are disjoint events, then $P(\bigcup_{i=1}^{n} A_i) = ?$ $\sum_{i=1}^{n} P(A_i)$ (By 3)
- Do we have $P(A) \leq 1$, for any A? (By 1, 2 and 3)

Examples: Probability Assignment

- Example: $\Omega = \{1,2,3,4\}$
  - $P(\{1,2\}) = 3/4$
  - $P(\{1,3,4\}) = 7/8$
  - $P(\{1,3\}) = 1/2$
  - Can this be made a valid probability assignment?

$P(\{4\}) = \frac{7}{8} - \frac{1}{2} = \frac{3}{8}$

$P(\{1,2,4\}) = \frac{3}{4} + \frac{3}{8} = \frac{9}{8} > 1$ (contradict to axiom 2)

No

- Example: $\Omega = \{0,1,2,3\cdots\}$
  - $P(\{k\}) = 2^{-k} \cdot |\cos(k\pi + \frac{\pi}{3})|$, for all k
  - Can this be made a valid probability assignment? Yes

$|\cos(k\pi + \frac{\pi}{3})| = \frac{1}{2}$ for all k

$\Rightarrow \sum_{k=0}^{\infty} 2^{-k} = 2$

$\Rightarrow \Omega = 2 \cdot \frac{1}{2} = 1$

Useful Properties

Prove the following properties by the axioms of probability:

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

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

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

- If $A \subseteq B$, then $P(A) \leq P(B)$

Union Bound

- For any events $A_1, A_2, \cdots, A_n$, we have

$$P(\bigcup_{n=1}^{N} A_n) \leq \sum_{n=1}^{N} P(A_n)$$

- Intuition:

- Proof: HW1 problem

Discrete Uniform Probability Law

Theorem: Let Ω be the sample space of an experiment. If Ω has N elements that are equally likely to occur, then for any event A of Ω, we have

$$P(A) = \frac{\text{Number of elements in A}}{N}$$

- How to verify this by using the axioms?

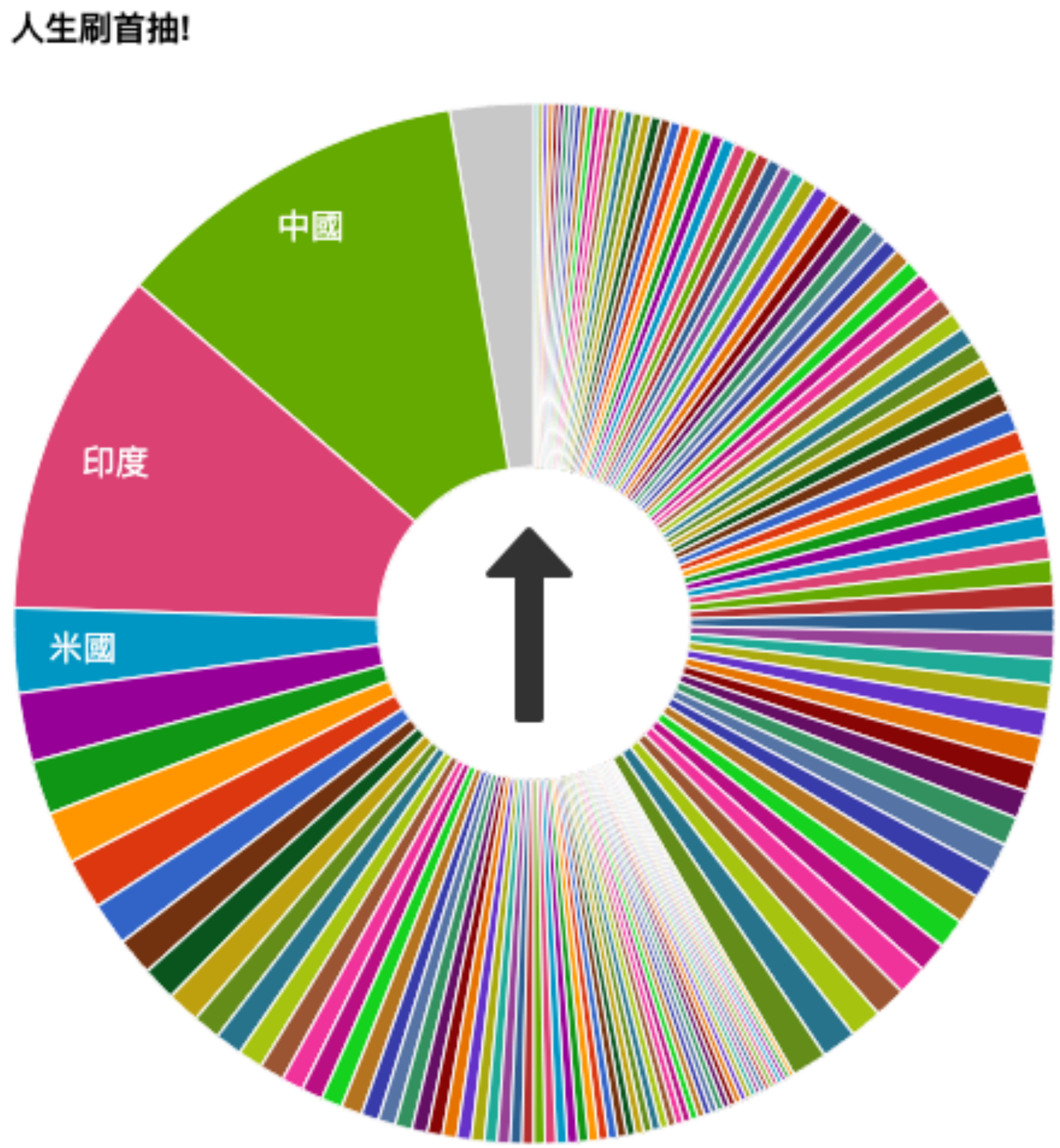
Example: Probability of Rain

- Experiment for probability of rain forecast:

例	降水量
1	0.1mm
2	0.0mm
3	4.8mm
4	0.3mm
5	0.0mm
6	1.2mm
7	0.0mm
8	2.4mm
9	0.9mm
10	0.5mm

- Procedure: Collect all historical data points of similar weather condition
- Model: All data points are equally likely to occur
- The rainy event = {rainfall $\geq$ 1mm}
- P(rainy event) = ?

Example: The Lottery of Birth



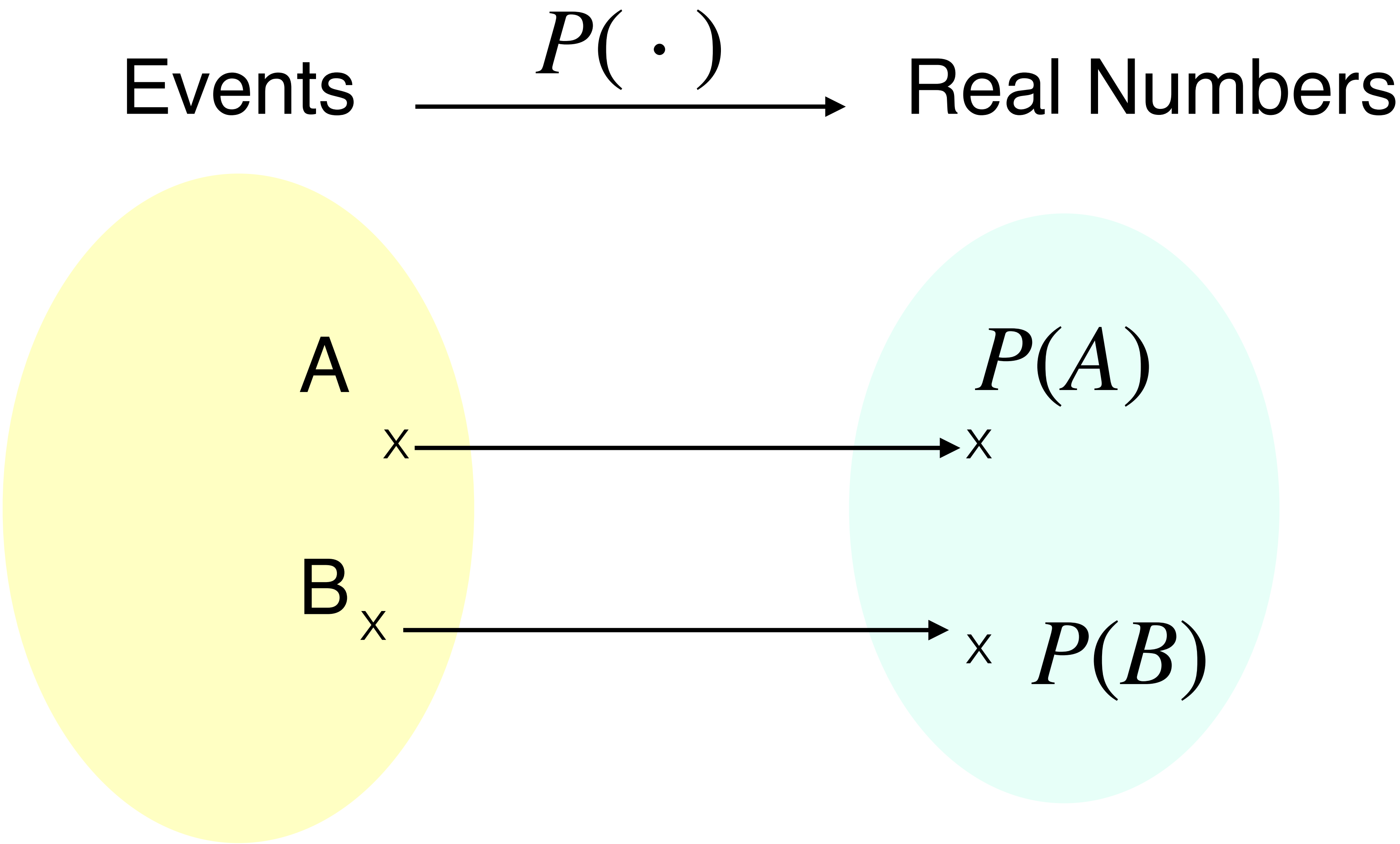
- Sample space = ?
- Probability assignment?
- P(born in Taiwan) = ?
- Veil of ignorance

John Rawls

http://z358z358.github.io/Roulette/#-LOw_CgDNn8VxaIzdtBE

Continuity of Probability Functions

Probability Assignment is a Function of Events



- The function $P(\cdot)$ needs to satisfy the 3 axioms

Review: Continuity of Functions

- What is a continuous function?
- Example: $f(x) = \sin(x)$
- Example: $f(x) = \lfloor x \rfloor$

Definition: A function $f : \mathbb{R} \to \mathbb{R}$ is **continuous** on $\mathbb{R}$ if and only if, for every convergent sequence $\{x_n\}_{n=1}^{\infty}$ with limit $\lim_{n\to\infty} x_n = x$, we have:

$$\lim_{n\to\infty} f(x_n) = f(x)$$

Continuity of Probability Function

- A sequence of events $E_1, E_2, \cdots$ is **increasing** if

$$E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \subseteq E_{n+1} \subseteq \cdots$$

Theorem: For any increasing sequence of events $E_1, E_2, \cdots$, we have

$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$$

- Is this trivial? Do we need a proof?

- Issue: Interchange of limiting operations

Interchange of Limiting Operations

- Example: $f_n(x) = (\sin nx)/\sqrt{n},\ n = 1,2,3,\cdots$

  Do we have $\lim_{n\to\infty} \frac{d}{dx} f_n(x) = \frac{d}{dx}\left(\lim_{n\to\infty} f_n(x)\right)$?

Interchange of Limiting Operations (Cont.)

- Example: $f_n(x) = \begin{cases} n, & \text{if } x \in (0, \frac{1}{n}) \\ 0, & \text{otherwise} \end{cases}$

  Do we have $\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 \left(\lim_{n\to\infty} f_n(x) \right) dx$?

Proof: Continuity of Probability Function

Theorem: For any increasing sequence of events $E_1, E_2, \cdots$, we have

$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$$

- Proof:

1-Minute Summary

1. Set operations

- Countable union / intersection and De Morgan's laws

2. Probability Axioms

- 3 axioms
- Valid probability assignments

3. Continuity of Probability Functions

- Increasing sequence of events
- Interchange of limiting operations