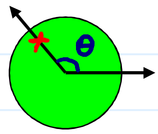
discrete Ω
(finite or countably infinite)
$P: 2^{\Omega} \to [0,1]$

continuous Ω
(uncountably infinite)
P: σ-field $\to [0,1]$
~~2^{Ω}~~

- Example (Waiting for a success):
  - Play roulette until a win.
  - $\Omega = \{1, 2, 3, \ldots\}$ ← infinite countable set
  - $P = ??$
- Example (Uniform Spinner):



  - Random Angle (in radians).
  - $\Omega = (-\pi, \pi]$ ← infinite uncountable set
  - $P= ??$

公理 (Axiom)
1. self-evident truth
2. a statement generally accepted as true

2^{Ω}

- The Modern Approach (抽象化)
  - A *probability measure* on Ω is a function P from subsets of Ω to the real number (or [0, 1]) that satisfies the following axioms:

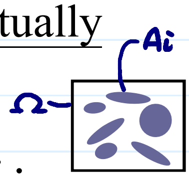
    (Ax1) Non-negativity. For any event A, $P(A)\geq 0$.

    (Ax2) Total one. $P(\Omega)=1$.

    countably infinite many

    (Ax3) Additivity. If $A_1, A_2, \ldots,$ is a sequence of mutually exclusive events, i.e., $A_i \cap A_j = \emptyset$ when $i\neq j$, then

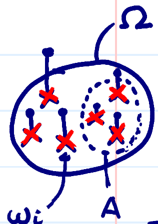
$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots.$$



- Notes: check whether the classical approach satisfies the 3 axioms (exercise)
  - These axioms restrict probabilities, but do not define them.
  - Probability is a property of events. $P: 2^{\Omega} \to [0,1]$

    c.f. $P(A) = \dfrac{\#A}{\#\Omega}$

- Define Probability Measures in a Discrete Sample Space.
  finite or countably infinite Ω
  - **Q**: Is it required to define probabilities directly on every events? (e.g., n possible outcomes in Ω, $2^n - 1$ possible events) $\because P(\Omega)=1$



  - Suppose $\Omega = \{\omega_1, \omega_2, \ldots\}$, *finite* or *countably infinite*, let $p : \Omega \to [0,1]$ satisfy

    no need to be symmetric

    c.f. ① $p(w) \geq 0$ for all $\omega \in \Omega$ and ② $\sum_{\omega\in\Omega} p(w) = 1.$

    probability mass function

    $P: 2^{\Omega} \to [0,1]$ finite or countably infinite many sum

  - Let

    $$P(A) = \sum_{\omega\in A} p(\omega) \Rightarrow P(\{\omega\}) = p(\omega)$$

    by additivity   examine Axiom 1~3   absolutely converge

    for $A\subset\Omega$, then P is a probability measure. (exercise)

    (**Q**: how to define p?) see the argument about subjective vs. objective interpretation (LNp.3-19~21)

- Example: In the classical approach, $p(\omega) = 1/\#\Omega$. For example, throw a fair dice, $\Omega=\{1, \ldots, 6\}$, $p(1)=\ldots=p(6)=1/6$ and $P(\text{odd})=P(\{1, 3, 5\})=p(1)+p(3)+p(5)=3/6=1/2$. $\rightarrow = \frac{\#A}{\#\Omega}$
- Example (non equally-likely events): Throwing an unfair dice might have $p(1)=3/8$, $p(2)=p(3)=\ldots=p(6)=1/8$, and $P(\text{odd})=P(\{1, 3, 5\})=p(1)+p(3)+p(5)= 5/8$. (c.f., Examples in LNp.3-5)
- Example (Waiting for Success – Play Roulette Until a Win): bet on "red"
  - Let $r=9/19$ and $q=1-r=10/19$ ← LNp.3-4
  - $\Omega = \{1, 2, 3, \ldots\}$ ← infinite, countable
  - Intuitively, $p(1)=r$, $p(2)=qr$, $p(3)=q^2r$, …, $p(n)=q^{n-1}r$, … >0, and ∵ independent

$$\sum_{n=1}^{\infty} p(n) = \sum_{n=1}^{\infty} rq^{n-1} = \frac{r}{1-q} = 1.$$

  - For an event $A\subset\Omega$, let

Q: What if we want to directly define P on 2^Ω?

2^Ω is an uncoutable set

cf.

$$P(A) = \sum_{n\in A} p(n).$$

For example, Odd=$\{1, 3, 5, 7, \ldots\}$

$$\begin{aligned} P(\text{Odd}) &= \sum_{k=0}^{\infty} p(2k+1) = \sum_{k=0}^{\infty} rq^{(2k+1)-1} = r\sum_{k=0}^{\infty} q^{2k} \\ &= r/(1-q^2) = 19/29. \end{aligned}$$

❖ **Reading**: textbook, Sec 2.3 & 2.5 ← read more examples of sample spaces having symmetric outcomes (classical approach)

Some Consequences of the 3 Axioms → 三原色RGB

- Proposition: For any sample space Ω, the probability of the empty set is zero, i.e.,

$$P(\emptyset) = 0. \text{ — (1)}$$

proof. In (Ax3), let $A_1=\Omega$, $A_2=A_3=\cdots=\phi$, $\Rightarrow A_i\cap A_j=\phi\ \forall i,j$.
By (Ax3), $\underset{1}{P(\Omega)} \overset{\text{(Ax2)}}{=} \underset{1}{P(\Omega)} + \sum_{n=2}^{\infty} P(\phi)$, $P(\phi) \geq 0$ (Ax1)

- Proposition: For any finite sequence of mutually exclusive events $A_1, A_2, \ldots, A_n$,

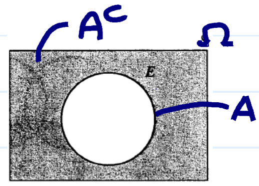
finite version of Ax3

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n). \text{ — (2)}$$

proof. In (Ax3), let $A_{n+1}=A_{n+2}=\cdots=\phi$, then (2) holds ∵ $P(\phi)=0$ by (1)

- Proposition: If A is an event in a sample space Ω and A^c is the complement of A, then

$$P(A^c) = 1 - P(A). \text{ — ③}$$



proof. $A^c \cup A = \Omega$, $A^c \cap A = \phi$

$1 = P(\Omega) = P(A^c \cup A) = P(A^c) + P(A)$ by ②

- Proposition: If A and B are events in a sample space Ω and $A \subset B$, then $P(A) \le P(B)$ and $P(B - A) = P(B \cap A^c) = P(B) - P(A)$.

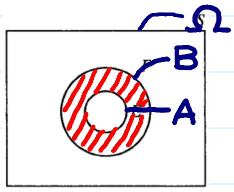
set minus

小機事件 報導 微觀事件

proof. $B = A \cup (B \cap A^c)$ disjoint

$P(B) = P(A) + P(B \cap A^c) \ge P(A)$ by (Ax1)

by ②



Recall. 馬路三寶 example (LNp.1-8)

- Example (摘自 “快思慢想”, Kahneman).

琳達是個三十一歲、未婚、有話直說的聰明女性。她主修哲學，在學生時代非常關心歧視和社會公義的問題，也參與過反核遊行。下面那一個比較可能？

A_1 ▪ 琳達是銀行行員。A_1

$A_1 \cap A_2$ ▪ 琳達是銀行行員，也是活躍的女性主義運動者。A_1 A_2

- Proposition: If A is an event in a sample space Ω, then

This is Axiom 1 in textbook

$$\underbrace{0 \le P(A)}_{\text{Ax1}} \le 1. = P(\Omega) \ \& \ A \subset \Omega$$

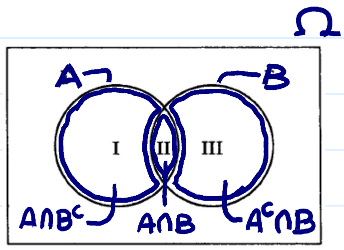
- Proposition: If A and B are two events in a sample space Ω, then

$P(A \cup B) \le P(A) + P(B) \Leftarrow$ $$P(A \cup B) = P(A) + P(B) - P(A \cap B). \text{ — ④}$$

remove disjoint codition

proof. $A \cup B = \mathrm{I} \cup \mathrm{II} \cup \mathrm{III}$ and

$\mathrm{I}, \mathrm{II}, \mathrm{III}$ mutually exclusive



$P(A \cup B) = P(\mathrm{I}) + P(\mathrm{II}) + P(\mathrm{III})$

$P(A) = P(\mathrm{I}) + P(\mathrm{II})$

$P(B) = P(\mathrm{II}) + P(\mathrm{III})$ by ②

$P(A \cap B) = P(\mathrm{II})$

- Proposition: If $A_1, A_2, \ldots, A_n$ are events in a sample space Ω, then

$$P(A_1 \cup \cdots \cup A_n) \le P(A_1) + \cdots + P(A_n). \overset{\text{c.f.}}{\longleftrightarrow} ② \ \& \ ④ \quad \text{LNp.3-9}$$

proof. $P((A_1 \cup \cdots \cup A_{n-1}) \cup A_n) \le P(A_1 \cup \cdots \cup A_{n-1}) + P(A_n)$

by ④ $\le P(A_1 \cup \cdots \cup A_{n-2}) + P(A_{n-1}) + P(A_n)$

$\le \cdots \le P(A_1) + \cdots + P(A_n)$

- Proposition (inclusion-exclusion identity): If $A_1, A_2, \ldots, A_n$ are any n events, let

generalization of ④

Q: For an outcome w contained in m out of the n events, how many times is its probability $p(w)$ repetitively counted in $\sigma_1, \ldots, \sigma_n$?

$$\sigma_1 = \sum_{i=1}^{n} P(A_i),$$

$\{\square, \square\}$ 1st, 2nd

$$\sigma_2 = \underbrace{\sum_{1\le i<j\le n}}_{\binom{n}{2}\text{ different }\{i,j\}} P(A_i \cap A_j),$$

$$\sigma_3 = \underbrace{\sum_{1\le i<j<k\le n}}_{\binom{n}{3}\text{ different }\{i,j,k\}} P(A_i \cap A_j \cap A_k),$$

$$\cdots = \cdots$$

$$\sigma_k = \underbrace{\sum_{1\le i_1<\cdots<i_k\le n}}_{\binom{n}{k}\text{ different }\{i_1,\cdots,i_k\}} P(A_{i_1} \cap \cdots \cap A_{i_k})$$

$$\cdots = \cdots$$

$$\sigma_n = P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

An intuitive proof for discrete Ω:

For $\omega \in \Omega$,

$p(\omega)$ counted m times in σ_1

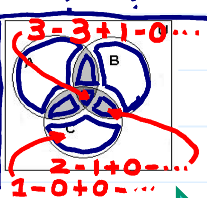
$\vdots$ $\quad\binom{m}{2}\quad$ $\vdots$ σ_2

$\vdots$ $\quad\binom{m}{k}\quad$ $\vdots$ σ_k

$\vdots$ $\quad\binom{m}{m}\quad$ $\vdots$ σ_m

$\vdots$ $\quad 0 \quad$ $\vdots$ σ_{m+1}

$\boxed{-\binom{m}{0}} + \binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \cdots + (-1)^{k+1}\binom{m}{k} + \cdots + (-1)^{m+1}\binom{m}{m} = 0$ (LNp.2-7)



then

$$P(A_1 \cup \cdots \cup A_n) = \sigma_1 - \sigma_2 + \sigma_3 - \cdots + (-1)^{k+1}\sigma_k + \cdots + (-1)^{n+1}\sigma_n.$$