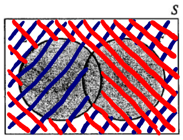
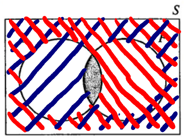
- Some Simple Rules of Set Operations
  - ➢ Commutative Laws. $A \cup B = B \cup A$ and $A \cap B = B \cap A$
  - ➢ Associative Laws. $(A \cup B) \cup C = A \cup (B \cup C)$
    $(A \cap B) \cap C = A \cap (B \cap C)$.
  - ➢ Distributive Laws. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
    $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
  - ➢ DeMorgan's Laws.



(a) Shaded region: $E \cup F$.



(b) Shaded region: EF.

required, not optional

$(\cup_{i=1}^{n} A_i)^c = \cap_{i=1}^{n} A_i^c$ and $(\cap_{i=1}^{n} A_i)^c = \cup_{i=1}^{n} A_i^c$.

by induction (exercise)

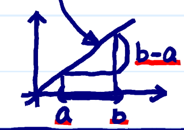
Reading: textbook, Sec 2.2

~~$\Omega \to [0,1]$~~

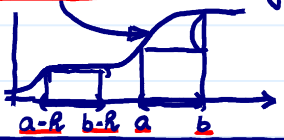
機率測度 → **Probability Measure** → a function: $2^{\Omega} \to [0,1]$

- The Classical Approach $\int g(x)dx$
  - ➢ Sample Space Ω is a *finite* set
  - ➢ Probability: For an event A,

length measure: $y = x$



weight measure $y = F(x)$ non-decreasing



probability space $(\Omega, \mathcal{F}, P)$ collection of events

$$P(A) = \frac{\#A}{\#\Omega}$$

This explains why combinatorial Thm plays an important role in probability.

$\int g(x)\,dF(x)$

- ➢ Example (Roulette):
  - $\Omega = \{0, 00, 1, 2, 3, 4, \ldots, 35, 36\}$
  - $P(\{\text{Red Outcome}\}) = 18/38 = 9/19$.

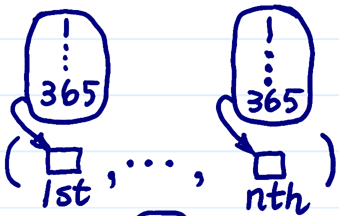
#A #Ω

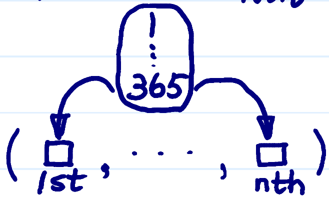
assume birthday is uniform in 365 days

- ➢ Example (Birthday Problem): n people gather at a party. What is the probability that they all have different birthdays?
  - Ω = lists of n from $\{1, 2, 3, \ldots, 365\}$
  - A = {all permutations}
  - $P_n(A) = (365)_n / 365^n$ #A #Ω
  - table:

n	8	16	22	23	32	40
$P_n(A)$	.926	.716	.524	.492	.247	.109

smaller





- Inadequacy of the Classical Approach

$$P(A) = \frac{\#A}{\#\Omega}$$

- ➢ It requires:
  - Finite Ω
  - Symmetric Outcomes

restrict the form of Ω

i.e., all outcomes in Ω are equally likely to occur, for $\omega \in \Omega$ $P(\{\omega\}) = \frac{1}{\#\Omega}$

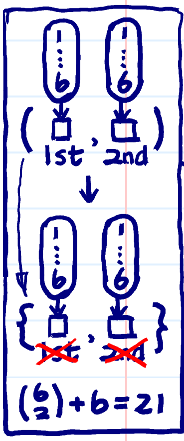
➢ Example (Sum of Two Dice Being 6)

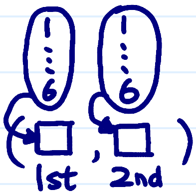
symmetric outcomes

- $\Omega_1 = \{(1,1), (1,2), (2,1), (1,3), (3,1), \ldots, (6,6)\}$, $\#\Omega_1 = 36$,

  {lists} 1/36

  $A = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$, $P(A) = 5/36$.





- $\Omega_2 = \{\{1,1\}, \{1,2\}, \{1,3\}, \ldots, \{6,6\}\}$, $\#\Omega_2 = 21$,

  {sets} 1/36 2/36 2/36

  $A = \{\{1,5\}, \{2,4\}, \{3,3\}\}$, $P(A) = \cancel{3/21}$.

- $\Omega_3 = \{2, 3, 4, \ldots, 12\}$, $\#\Omega_3 = 11$,

  {sums} 1/36 2/36 3/36

  $A = \{6\}$, $P(A) = \cancel{1/11}$.

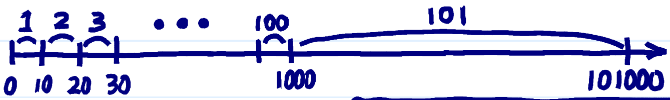
non equally-likely outcomes

can still use Ω_2, Ω_3 to calculate correct probability, but not through $P(A) = \#A/\#\Omega$

➢ Example (Sampling Proportional to Size):

$N=101$
$n=4$
$\{10, 21, 39, 85\}$

- N invoices.
- Sample $n < N$.
- Pick large ones with higher probability.
- Note: Finite Ω, but *non equally-likely* outcomes.



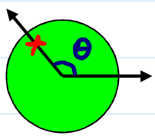
each dollar should have equal chance of being selected, not each invoice

$\#\Omega = \binom{N}{n}$

➢ 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
- $P = ??$

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

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

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:

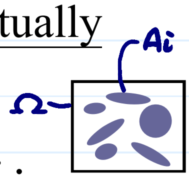
(Ax1) Non-negativity. For any event A, $P(A) \geq 0$.

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

(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

countably infinite many

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



- Notes:

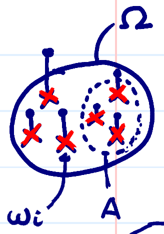
check whether the classical approach satisfies the 3 axioms (exercise)

  - These axioms restrict probabilities, but do not define them.
  - Probability is a property of events. $P: 2^\Omega \to [0,1]$

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

➢ 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. probability mass function

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

$P: 2^\Omega \to [0,1]$ 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

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

absolutely converge

(**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 $P(\text{odd})=P(\{1, 3, 5\})=p(1)+p(3)+p(5)=3/6=1/2$. $= \frac{\#A}{\#\Omega}$

➢ 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 $P(\text{odd})=P(\{1, 3, 5\})=p(1)+p(3)+p(5)= 5/8$. (cf., Examples in LNp.3-5)

bet on "red"

➢ Example (Waiting for Success – Play Roulette Until a Win):

- 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

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

∵ independent

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

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

cf.

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

2^Ω is an uncoutable set

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

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

❖ **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), $\underbrace{P(\Omega) = P(\Omega)}_{(Ax2)} + \sum_{n=2}^{\infty} \underbrace{P(\phi)}_{\geq 0\ (Ax1)}$

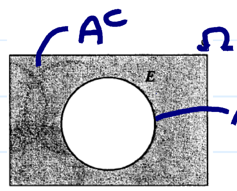
- Proposition: For any finite sequence of mutually exclusive events $A_1, A_2, \ldots, A_n$,

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 $\because P(\phi)=0$ by ①

- Proposition: If A is an event in a sample space Ω and A^c is the complement of A, then $P(A^c) = 1 - P(A)$. — ③

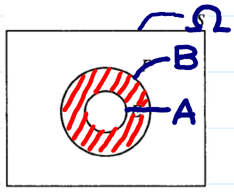


proof. $A^c \cup A = \Omega$, $A^c \cap A = \phi$
$1 = P(\Omega) = P(A^c \cup A) = P(A^c) + P(A)$ by ②

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

少見事件 報導 常見事件

proof. $B = A \cup (B \cap A^c)$ (disjoint)
$P(B) = P(A) + P(B\cap A^c) \geq P(A)$ by (Ax1); $=$ by ②



- Example (摘自 "快思慢想", Kahneman).

Recall. 馬路三寶 example (LNp.1-8)

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，(A_1) 也是活躍的女性主義運動者。(A_2)

This is Axiom 1 in textbook

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

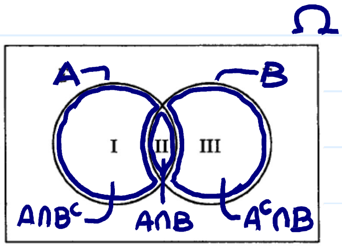
$$\underbrace{0 \le P(A)}_{\text{Ax1}} \le 1. = P(\Omega) \ \& \ A \subset \Omega$$

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

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

remove disjoint codition

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



$$\begin{aligned}
P(A\cup B) &\overset{\text{Ax3}}{=} P(\mathrm{I}) + P(\mathrm{II}) + P(\mathrm{III}) \\
P(A) &\overset{\text{Ax3}}{=} P(\mathrm{I}) + P(\mathrm{II}) \\
P(B) &\overset{\text{Ax3}}{=} P(\mathrm{II}) + P(\mathrm{III}) \\
P(A \cap B) &= P(\mathrm{II})
\end{aligned} \quad \text{by (2)}$$

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

$$P(A_1 \cup \cdots \cup A_n) \le P(A_1) + \cdots + P(A_n). \overset{\text{cf}}{\longleftrightarrow} (2) \ \& \ (4)$$

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)

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

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