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

generalization of ④

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

$$\sigma_2 = \sum_{1 \le i < j \le n} P(A_i \cap A_j),$$

$\binom{n}{2}$ different $\{i, j\}$

$$\sigma_3 = \sum_{1 \le i < j < k \le n} P(A_i \cap A_j \cap A_k),$$

$\binom{n}{3}$ different $\{i, j, k\}$

$$\cdots = \cdots$$

$$\sigma_k = \sum_{1 \le i_1 < \cdots < i_k \le n} P(A_{i_1} \cap \cdots \cap A_{i_k})$$

$\binom{n}{k}$ different $\{i_1, \cdots, i_k\}$

$$\cdots = \cdots$$

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

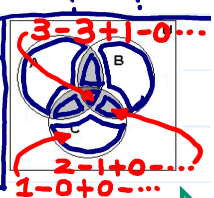
then

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

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

An intuitive proof for discrete Ω:
For $\omega \in \Omega$,
$p(\omega)$ counted m times in σ_1
: : $\binom{m}{2}$: : σ_2
: : $\binom{m}{k}$: : σ_k
: : $\binom{m}{m}$: : σ_m
: : 0 : : σ_{m+1}

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



proof. Prove by induction.
$n=2$, it holds (by ④, LNp.3-11)
$n=3$, $P(A_1 \cup A_2 \cup A_3) = P(A_1 \cup A_2) + P(A_3) - P((A_1 \cup A_2) \cap A_3)$

$(A_1 \cap A_3) \cup (A_2 \cap A_3)$

$= P(A_1) + P(A_2) - P(A_1 \cap A_2)$ (by ④)

$= P(A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)$

$= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$

For $n > 3$, by mathematical induction (exercise) → $A_1 \cup \cdots \cup A_{n-1} \cup A_n$

➢ Notes:

- There are $\binom{n}{k}$ summands in σ_k
- In symmetric examples,

  e.g. $k=2$, $P(A_1 \cap A_2) = P(A_1 \cap A_3) = \cdots = P(A_{n-1} \cap A_n)$

  $$\sigma_k = \binom{n}{k} P(A_1 \cap \cdots \cap A_k)$$

  cf. symmetric outcomes (LNp.3-4)

- It can be shown that

  3rd proposition in LNp.3-11

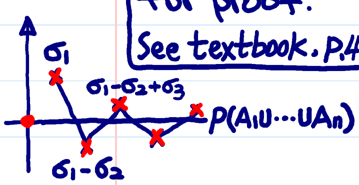
  $$P(A_1 \cup \cdots \cup A_n) \le \sigma_1$$ 加太多

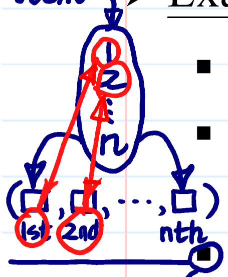
  $$P(A_1 \cup \cdots \cup A_n) \ge \sigma_1 - \sigma_2$$ 減太多

  $$P(A_1 \cup \cdots \cup A_n) \le \sigma_1 - \sigma_2 + \sigma_3$$ 加太多

  $$\cdots \quad \cdots \quad \cdots$$

For proof. See textbook, p.44





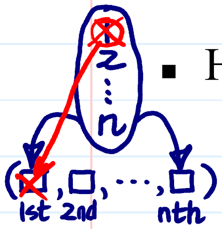
➢ Example (The Matching Problem). classical prob. = #A/#Ω ← symmetric outcomes

- Applications: (a) Taste Testing. (b) Gift Exchange.
- Let Ω be all permutations ω = (i_1, …, i_n) of 1, 2, …, n. Thus, #Ω = n!. e.g. (3, 1, 5, …) 1st 2nd 3rd
- Let A_j = {ω: i_j = j} and $A = \cup_{i=1}^{n} A_i$, (item; event of interest can be difficult to identify; A^c: none of them get his/her own gift; A: at least one person get his/her own gift)

  item i_j: (…, i_j, …) j-th position

  Q: P(A)=? (What would you expect when n is large?) P(A)↑? to 1? P(A)↓? to 0? or others?
- By symmetry,

$$\sigma_k = \binom{n}{k} P(A_1 \cap \cdots \cap A_k),$$



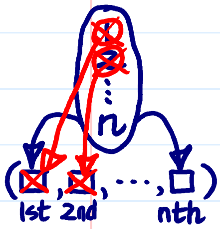
- Here,

$$P(A_1) = \frac{1 \times (n-1)!}{n!} = \frac{1}{n},$$

$$P(A_1 \cap A_2) = \frac{(n-2)!}{n!} = \frac{1}{(n)_2},$$

$$\cdots = \cdots,$$

$$P(A_1 \cap \cdots \cap A_k) = \frac{(n-k)!}{n!} = \frac{1}{(n)_k}, \cdots$$



for k = 1, …, n.

- So, $\sigma_k = \binom{n}{k}\frac{1}{(n)_k} = \frac{1}{k!}$, $= \frac{n!}{k!(n-k)!} \times \frac{(n-k)!}{n!}$

$$P(A) = \sigma_1 - \sigma_2 + \cdots + (-1)^{n+1}\sigma_n = \sum_{k=1}^{n}(-1)^{k+1}\frac{1}{k!},$$

$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$ (−1)

c.f.

$$P(A) = 1 - \sum_{k=0}^{n\to\infty}(-1)^k\frac{1}{k!} \approx 1 - \frac{1}{e} = 0.632 \Rightarrow P(A^c) \approx e^{-1} = 0.368$$

(≈ when $n \to \infty$)

- Note: approximation accurate to 3 decimal places if $n \geq 6$.

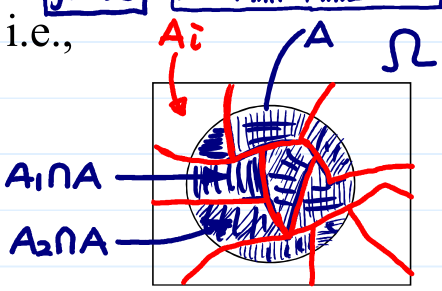
⊙ Proposition: If A_1, A_2, …, is a *partition* of Ω, i.e., (countably infinite many or finite ← Let $A_{n+1} = A_{n+2} = \cdots = \phi$)

cf. additivity axiom

1. $\cup_{i=1}^{\infty} A_i = \Omega$,
2. A_1, A_2, …, are mutually exclusive,

then, for any event $A \subset \Omega$,

$$P(A) = \sum_{i=1}^{\infty} P(A \cap A_i).$$



proof. $A = A \cap \Omega = A \cap (\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty}(A \cap A_i)$ ← mutually exclusive

$P(A) = P(\bigcup_{i=1}^{\infty}(A \cap A_i)) = \sum_{i=1}^{\infty} P(A \cap A_i)$

❖ **Reading**: textbook, Sec 2.4 & 2.5

Probability Measure for Continuous Sample Space

- **Q**: How to define probability in a continuous sample space? (uncountably infinite Ω)
  cf. → How to define P. M. for discrete Ω ↳ check Examples in LNp.3-1 ~ 3-6

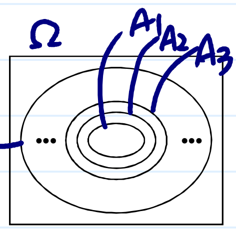
- Monotone Sequences of sets (LNp.3-7)

  ➢ Definition: A sequence of events $A_1, A_2, \ldots,$ is called *increasing* if (countably infinite many)

  $$A_1 \subset A_2 \subset \cdots \subset A_n \subset A_{n+1} \subset \cdots \subset \Omega$$

  and *decreasing* if

  $$A_1 \supset A_2 \supset \cdots \supset A_n \supset A_{n+1} \supset \cdots \supset \emptyset$$

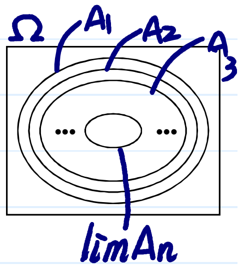


  The limit of an increasing sequence is defined as

  $$\lim_{n\to\infty} A_n = \cup_{i=1}^{\infty} A_i \quad (A_1 \cup \cdots \cup A_n)$$

  and the limit of a decreasing sequence is (countably infinite)

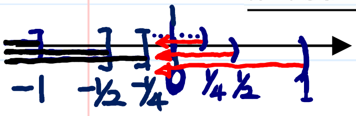
  $$\lim_{n\to\infty} A_n = \cap_{i=1}^{\infty} A_i \quad (A_1 \cap \cdots \cap A_n)$$



  ➢ Example: If $\Omega=\mathbb{R}$ and $A_k = (-\infty, 1/k)$, then A_k's are decreasing and

  $$\lim_{k\to\infty} A_k = \{\omega : \omega < 1/k \ \text{ for all } k \in \mathbb{Z}_+\} = (-\infty, 0].$$

  (exercise) $A_k = (-\infty, -\frac{1}{k}] \uparrow$, $\lim_{k\to\infty} A_k = (-\infty, 0)$



- Proposition: If $A_1, A_2, \ldots,$ is increasing or decreasing, then

  $$\left(\lim_{n\to\infty} A_n\right)^c = \lim_{n\to\infty} A_n^c$$

  cf. LNp.3-3 DeMorgan's Law

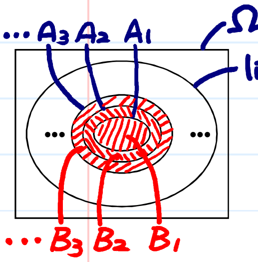
  $(\bigcup_{n=1}^{\infty} A_n)^c = \bigcap_{n=1}^{\infty} A_n^c = \lim_{n\to\infty} A_n^c$ (↑ → ↓)

  $(\bigcap_{n=1}^{\infty} A_n)^c = \bigcup_{n=1}^{\infty} A_n^c = \lim_{n\to\infty} A_n^c$ (↓ → ↑)

- Proposition: If $A_1, A_2, \ldots,$ is increasing or decreasing, then

  $$\lim_{n\to\infty} P(A_n) = P\left(\lim_{n\to\infty} A_n\right). \quad \text{—(5)}$$

  limit of numbers; limit of sets; a P.M.; lim & P exchangable; (A) measurable; What A_n's should be defined first?



proof. (a) A_n increasing.

Let $B_1 = A_1$, $B_2 = A_2 \cap A_1^c, \cdots, B_n = A_n \cap A_{n-1}^c, \cdots$

$\Rightarrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$, $\bigcup_{n=1}^{k} B_n = \bigcup_{n=1}^{k} A_n = A_k$, $B_i \cap B_j = \phi$ for $i \neq j$. (mutually exclusive)

$P(\lim_{n\to\infty} A_n) = P(\bigcup_{n=1}^{\infty} A_n) = P(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} P(B_n) = \lim_{k\to\infty} \sum_{n=1}^{k} P(B_n)$

$= \lim_{k\to\infty} P(\bigcup_{n=1}^{k} B_n) = \lim_{k\to\infty} P(\bigcup_{n=1}^{k} A_n) = \lim_{k\to\infty} P(A_k)$ (by Ax3 Additivity; countable sums)

(b) A_n decreasing $\Rightarrow A_n^c$ increasing

$1 - P(\lim_{n\to\infty} A_n) = P((\lim_{n\to\infty} A_n)^c)$

$= P(\lim_{n\to\infty} A_n^c) = \lim_{n\to\infty} P(A_n^c) = \lim_{n\to\infty} 1 - P(A_n) = 1 - \lim_{n\to\infty} P(A_n)$ (by (a))

LNp.3-6 ⊙ Example (Uniform Spinner): Let $\Omega = (-\pi, \pi]$. Define

$$P((a, b]) = \frac{b-a}{2\pi} \quad (*)$$

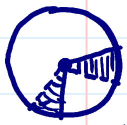
for subintervals $(a, b] \subset \Omega$. Then, extend P to other subsets using the 3 axioms. For example, if $-\pi < a < b < \pi$,

$$P([a,b]) = P\left(\left(\bigcap_{k=1}^{\infty}(a-\tfrac{1}{k}, b]\right)\cap\Omega\right) = P\left(\bigcap_{k=1}^{\infty}\left((a-\tfrac{1}{k}, b]\cap\Omega\right)\right)$$

$$= \lim_{k\to\infty} P\left((a-\tfrac{1}{k}, b]\cap\Omega\right)$$

$$= \lim_{k\to\infty} \frac{1}{2\pi}(b-a+\tfrac{1}{k}) = \frac{b-a}{2\pi}.$$

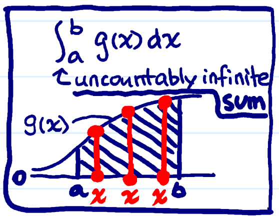
(continuous sample space. Actually, it's enough to define P on any (a,b], where a,b are rational numbers (∵ rational numbers is a dense set). Play a similar role like $\omega_1, \omega_2, \cdots$, in small p for discrete sample space (c.f. how to define P in LNp.3-7). Cumulative distribution function. Note: [a,b] ≠ (a,b] ⇒ P not defined on [a,b] under (*). $A_k = (a-\frac{1}{k}, b]\cap\Omega$, $A_k \downarrow$, $\lim_{k\to\infty} A_k = [a,b]$. by (5). by (*). 基礎事件 discrete: {ωi}, continuous: (a,b]. Note: P([a,b]) is derived through P((a,b]). P(height=170)=0?)



- Some notes
  - □ $P(\{a\}) = P([a,b]-(a,b]) = P([a,b]) - P((a,b]) = 0.$
  - □ If $C = \{\omega_1, \omega_2, \ldots\} \subset \Omega$, then

$$P(C) = \sum_{i=1}^{\infty} P(\{\omega_i\}) = 0 + 0 + \cdots = 0.$$

($\frac{b-a}{2\pi}$; Ax3; $\bigcup_{i=1}^{\infty}\{\omega_i\}$; cf. discrete sample space; In Ax3, countable unions (LNp3-6); $\int_a^b g(x)dx$ — uncountably infinite sum.)



  - □ The probability of all rational outcomes is zero (a countable (but dense) set.)

❖ **Reading**: textbook, Sec. 2.6

Objective vs. Subjective "Interpretation" of Probability

- Evaluate the following statements
  1. This is a fair coin (50%) → tossing the coin ← repeatable case (cf. one-time case) — obj
  2. It's 90% probable that Shakespeare actually wrote Hamlet — subj
- **Q**: What do we mean if we say that the probability of rain tomorrow is 40%? (how to interpret it? rain tomorrow — an event A: P(A) = 0.4, O(A) = 0.4/0.6 = 2/3)

  Objective: Long run relative frequencies

  Subjective: Chosen to reflect opinion

- The Objective (Frequency) Interpretation 頻率
  - ➢ Through Experiment: Imagine the experiment repeated N times. For an event A, let

$$N_A = \#\text{ occurrences of } A.$$

(frequency; random number)

Then,

$$P(A) \equiv \lim_{N\to\infty} \frac{N_A}{N}.$$

(By Law of Large Number (future lecture or course))

➢ Example (Coin Tossing):

N	100	1000	10000	100000
N_H (random)	55	493	5143	50329
N_H/N (frequency)	.550 (.580)	.493 (.512)	.514 (.498)	.503 (.501)

→ 0.5

The result is consistent with $P(H)$=0.5.

- The Subjective Interpretation

★ odds of an event (勝敗率, 賠率)

$o(A) = \frac{P(A)}{P(A^c)} = \frac{P(A)}{1-P(A)} \Rightarrow P(A) = \frac{o(A)}{1+o(A)}$

$o(A) \in [0, \infty)$

$P(A)\uparrow \Leftrightarrow o(A)\uparrow \Leftrightarrow o(A^c)\downarrow$

➢ Strategy: Assess probabilities by imagining bets

➢ Example: 押 2 : 1 賺 odds on A

- Peter is willing to give two to one odds that it will rain tomorrow (event A). His subjective probability for rain tomorrow is at least 2/3

(Peter) true $o(A) = \frac{P(A)}{P(A^c)} \geq \frac{2}{1}$, $P(A^c) = 1-P(A)$ $\Rightarrow P(A) \geq 2-2P(A)$ $\Rightarrow P(A) \geq 2/3$

- Paul accepts the bet (1 : 2 odds on A^c). His subjective probability for rain tomorrow is at most 2/3

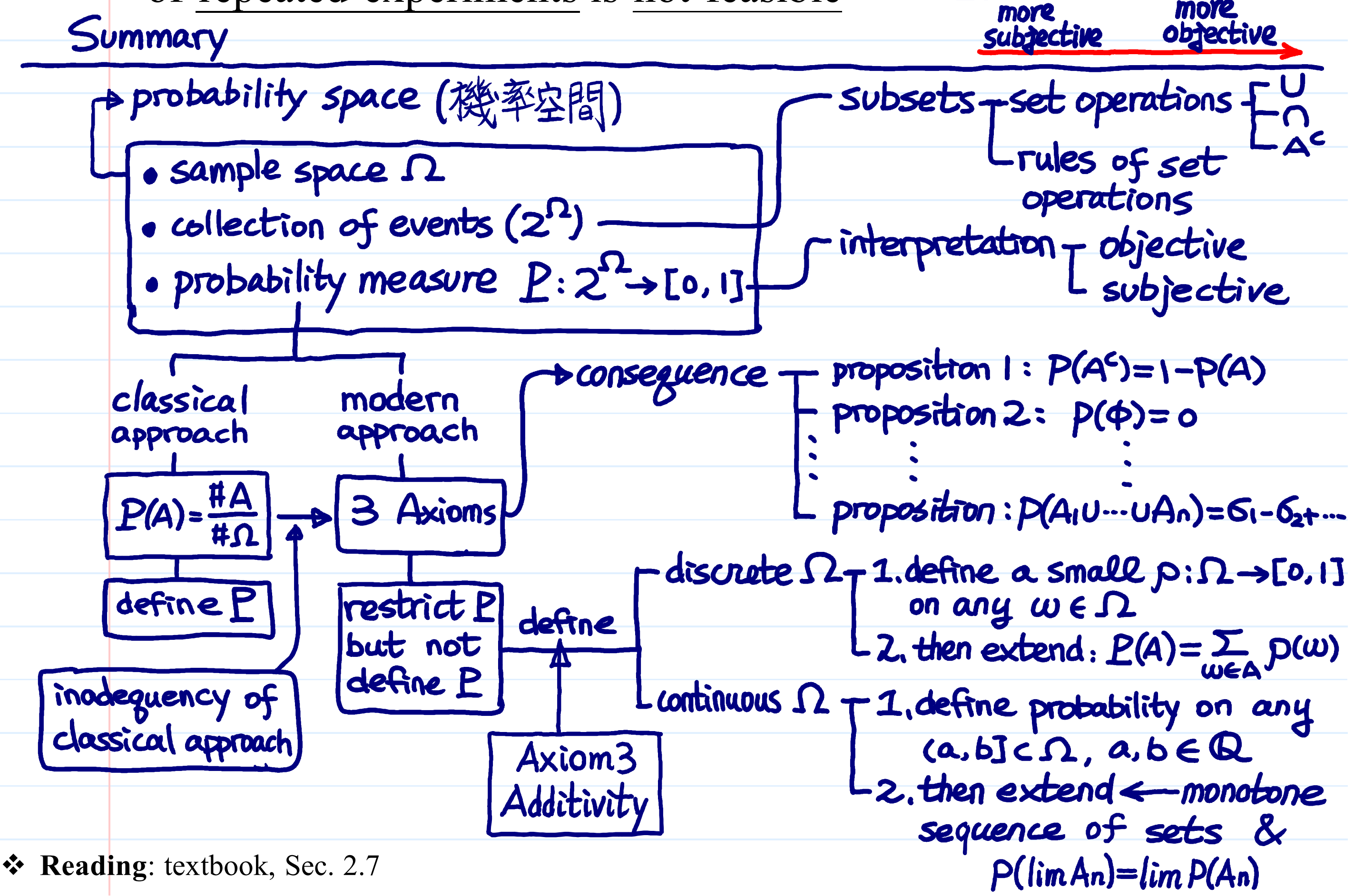
(Paul) true $o(A^c) = \frac{P(A^c)}{P(A)} \geq \frac{1}{2}$ $\Rightarrow P(A) \leq 2/3$

➢ Probabilities are simply personal measures of how likely we think it is that a certain event will occur (Bayesian approach; degree of personal belief)

(e.g. 2 in LNp.3-19.)

➢ This can be applied even when the idea of repeated experiments is not feasible

repeatable case: A A^c, A A^c, …, A A^c … → time; 1st, 2nd, …, nth …; more subjective → more objective



❖ **Reading**: textbook, Sec. 2.7