- 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

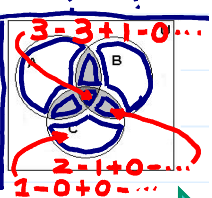
$$P(A_1 \cup \cdots \cup A_n) = \sigma_1 - \sigma_2 + \sigma_3 - \cdots + (-1)^{k+1}\sigma_k + \cdots + (-1)^{n+1}\sigma_n.$$

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

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}

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



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

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

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

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) → $\underline{A_1 \cup \cdots \cup A_{n-1}} \cup \underline{A_n}$

➢ Notes:

- There are $\binom{n}{k}$ summands in σ_k
- In symmetric examples, e.g. $k=2$, $P(A_1 \cap A_2) = P(A_1 \cap A_3) = \cdots = P(A_{n-1} \cap A_n)$

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

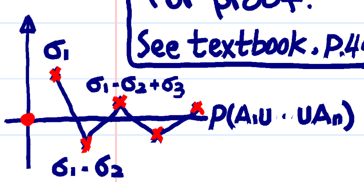
cf. symmetric outcomes (LNp.3-4)

- It can be shown that

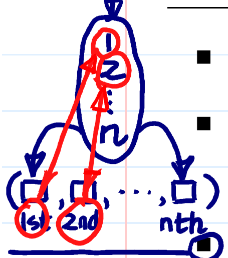
3rd proposition in LNp.3-11

For proof. See textbook. p.44

$$P(A_1 \cup \cdots \cup A_n) \le \sigma_1 \quad \text{加太多}$$
$$P(A_1 \cup \cdots \cup A_n) \ge \sigma_1 - \sigma_2 \quad \text{減太多}$$
$$P(A_1 \cup \cdots \cup A_n) \le \sigma_1 - \sigma_2 + \sigma_3 \quad \text{加太多}$$
$$\cdots \quad \cdots \quad \cdots$$



item → ➢ Example (The Matching Problem). classical prob $= \#A / \#\Omega$ ← symmetric outcomes



- Applications: (a) Taste Testing. (b) Gift Exchange.
- Let Ω be all permutations $\omega = (i_1, \ldots, i_n)$ of $1, 2, \ldots, n$. Thus, $\#\Omega = n!$. e.g. (3, 1, 5, ...) 1st 2nd 3rd

- Let event of interest can be difficult to identify

item

$A_j = \{\omega : i_j = j\}$ and $A = \cup_{i=1}^n A_i$, → at least one person get his/her own gift

A^c → none of them get his/her own gift

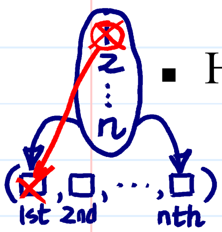
item j: $(\ldots, \boxed{i_j}, \ldots)$ j-th position

Q: $P(A)$=? (What would you expect when n is large?)

P(A)↑? to 1? P(A)↓? to 0? or others?

- By symmetry,

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



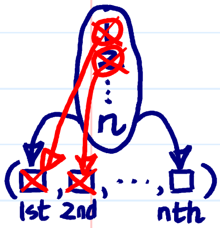
- Here,

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

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

$$\cdots = \cdots,$$

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



for $k = 1, \ldots, n$.

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

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

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

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 $n \to \infty$

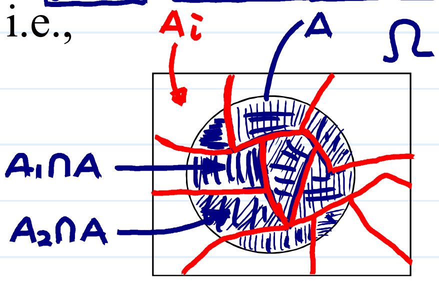
- Note: approximation accurate to 3 decimal places if $n \geq 6$.

countably infinite many or finite ← Let $A_{n+1} = A_{n+2} = \cdots = \phi$

⊙ Proposition: If $A_1, A_2, \ldots,$ is a *partition* of Ω, i.e.,

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

c.f. additivity axiom



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

uncountably infinite Ω

⊙ **Q**: How to define probability in a continuous sample space?

cf. How to define P.M. for discrete Ω (LNp.3-7)

check Examples in LNp.3-1 ~ 3-6

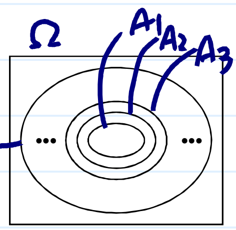
- Monotone Sequences of sets

➢ Definition: A sequence of events $A_1, A_2, \ldots,$ is called *increasing* if

countably infinite many

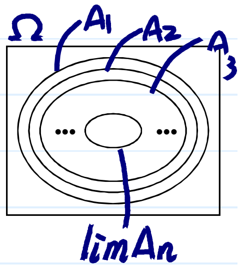
$$A_1 \subset A_2 \subset \cdots \subset A_n \subset A_{n+1} \subset \cdots \subset \Omega$$

lim A_n



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

$A_1 \cup \cdots \cup A_n$

$$\lim_{n\to\infty} A_n = \cup_{i=1}^{\infty} A_i$$

countably infinite

and the limit of an decreasing sequence is

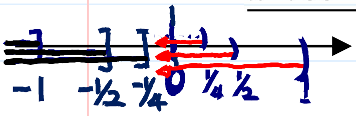
$A_1 \cap \cdots \cap A_n$

$$\lim_{n\to\infty} A_n = \cap_{i=1}^{\infty} A_i$$

➢ 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

cf. LNp.3-3 DeMorgan's Law

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

$\cup$ $\uparrow$ ⟷ $\cap$ $\downarrow$

$\cap$ $\downarrow$ ⟷ $\cup$ $\uparrow$

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

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

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

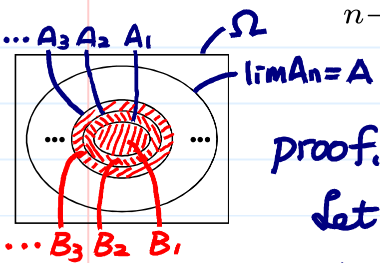
limit of numbers limit of sets

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

a P.M. lim & P exchangable

A measurable

What A_n's should be defined first?



proof. ⓐ A_n increasing.

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

mutually exclusive

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

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

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

by Ax3 Additivity

countable sums

ⓑ 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 ⓐ

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

Actually, it's enough to define P on any $(a,b]$, where a, b are rational numbers ($\because$ rational numbers is a dense set)

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

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)

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

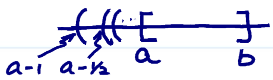
cumulative distribution function

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

Note: $[a,b] \neq (a,b]$ $\Rightarrow$ P not defined on $[a,b]$ under $(*)$

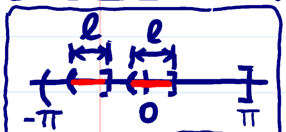
$$\overset{\text{by (5)}}{=} \lim_{k\to\infty} P\left((a-\frac{1}{k}, b] \cap \Omega\right)$$

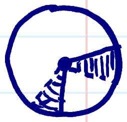
$$\overset{\text{by }(*)}{=} \lim_{k\to\infty} \frac{1}{2\pi}(b-a+\frac{1}{k}) = \frac{b-a}{2\pi}.$$



$A_k \downarrow$, $\lim_{k\to\infty} A_k = [a,b]$

基礎事件
discrete: $\{\omega_i\}$
continuous: $(a,b]$





- Some notes

Note: $P([a,b])$ is derived through $P((a,b])$.

$P(\text{height}=170) = 0$?

- $P(\{a\}) = P([a,b] - (a,b]) = P([a,b]) - P((a,b]) = 0.$

- If $C = \{\omega_1, \omega_2, \ldots\} \subset \Omega$, then

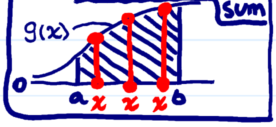
discrete sample space (cf.)

$\bigcup_{i=1}^{\infty}\{\omega_i\}$

$$P(C) \overset{\text{Ax3}}{=} \sum_{i=1}^{\infty} P(\{\omega_i\}) = 0 + 0 + \cdots = 0.$$

$\frac{b-a}{2\pi}$ $\frac{b-a}{2\pi}$

$\int_a^b g(x)\,dx$ ← uncountably infinite sum



In Ax3, countable unions (LNp.3-6)

- 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

obj → 1. This is a fair coin

subj → 2. It's 90% probable that Shakespeare actually wrote Hamlet

- **Q**: What do we mean if we say that the probability of rain tomorrow is 40%? — how to interpret it? — an event A

$P(A) = 0.4$
$o(A) = \frac{0.4}{0.6} = \frac{2}{3}$

  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
random →	N_H	55	493	5143	50329
frequency →	N_H/N	.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 賺 2 : 1 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 $o(A) = \frac{P(A)}{P(A^c)} \geq \frac{2}{1}$, $P(A^c) = 1-P(A)$

$\Rightarrow P(A) \geq 2-2P(A)$

$\Rightarrow P(A) \geq 2/3$

1 : 2 odds on A^c

- Paul accepts the bet. His subjective probability for rain tomorrow is at most 2/3

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

$\Rightarrow P(A) \leq 2/3$

Bayesian approach

➢ Probabilities are simply personal measures of how likely we think it is that a certain event will occur — 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

Summary

probability space (機率空間)

- sample space Ω
- collection of events (2^Ω) — subsets — set operations { $\cup$, $\cap$, A^c }; rules of set operations
- probability measure $P: 2^\Omega \to [0,1]$ — interpretation — objective; subjective

classical approach: $P(A) = \frac{\#A}{\#\Omega}$ — define P

inadequency of classical approach

modern approach: 3 Axioms — restrict P but not define P

consequence — proposition 1: $P(A^c) = 1 - P(A)$
- proposition 2: $P(\phi) = 0$
- ⋮
- proposition: $P(A_1 \cup \cdots \cup A_n) = \mathfrak{S}_1 - \mathfrak{S}_2 + \cdots$

define ← Axiom3 Additivity

discrete Ω — 1. define a small $p: \Omega \to [0,1]$ on any $\omega \in \Omega$
- 2. then extend: $P(A) = \sum_{\omega \in A} p(\omega)$

continuous Ω — 1. define probability on any $(a,b] \subset \Omega$, $a, b \in \mathbb{Q}$
- 2. then extend ← monotone sequence of sets & $P(\lim A_n) = \lim P(A_n)$

❖ **Reading**: textbook, Sec. 2.7