樣本空間

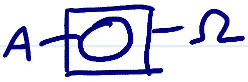
Sample Space and Events

- *Sample Space* Ω: the set of all possible outcomes in a random phenomenon. Examples:

known

outcome is not predictable

  1. Sex of a newborn child: Ω = {girl, boy}
  2. The order of finish in a race among the 7 horses 1, 2, …, 7:
     Ω = {all 7! Permutations of (1, 2, 3, 4, 5, 6, 7)}
  3. Flipping two coins: Ω = {(H, H), (H, T), (T, H), (T, T)}
  4. Lifetime of a transistor: Ω = [0, ∞)



事件

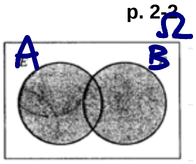
- *Event*: Any (measurable) subset of Ω is an event. Examples:
  1. *A*={girl}: the event - child is a girl.
  2. *A*={all outcomes in Ω starting with a 3}: the event - horse 3 wins the race.

$\Omega=\{w_1,\cdots,w_n\}, 2^{\Omega}=\{\phi,\{w_1\},\cdots,\{w_n\},\{w_1,w_2\},\cdots\}$

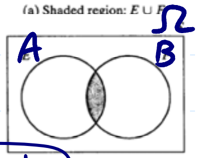
  3. *A*={(H, H), (H, T)}: the event - head appears on the 1st coin.
  4. *A*=[0, 5]: the event - transistor does not last longer than 5 hours.
  - an event occurs: outcome ∈ the event 2^{Ω}: collection of all subsets in Ω.
  - Q: How many different events if #Ω=$n < \infty$? Ans: $2^n = \#2^{\Omega}$

- Set Operations of Events
  - ➢Union. $C = A \cup B \Rightarrow$ C: either A or B occurs
  - ➢Intersection. $C = A \cap B \Rightarrow$ C: both A and B occur
  - ➢Complement. $C = A^c \Rightarrow$ C: A does not occur
  - ➢Mutually Exclusive. $A \cap B = \emptyset \Rightarrow$ A and B have no outcomes in common. disjoint.
  - ➢Definitions of union and intersection for more than two events can be defined in a similar manner

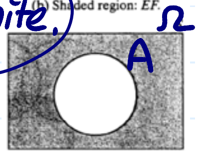
including countable infinite.



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



(b) Shaded region: EF.

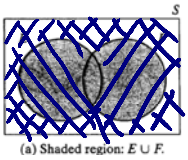


(c) Shaded region: E^c.

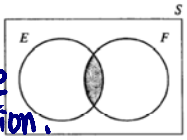
- Some Simple Rules of Set Operations
  - ➢Commutative Laws. $A \cup B = B \cup A$ and $A \cap B = B \cap A$
  - ➢Associative Laws. $(A \cup B) \cup C = A \cup (B \cup C)$
    $(A \cap B) \cap C = A \cap (B \cap C)$.
  - ➢Distributive Laws. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
    $(A \cap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

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

exercise + induction.



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



(b) Shaded region: EF.

❖ **Reading**: textbook, Sec 2.2

Probability Measure 机率測度.

↳ a function: $2^{\Omega} \to [0,1]$

- The Classical Approach

  Sample Space Ω is a *finite* set

  Probability: For an event A,

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

$\#A \to$ # of elements in A

This explains why combinatorial thm play an important role in probability.

read more examples in textbook, Sec.2.5

➢ Example (Roulette):

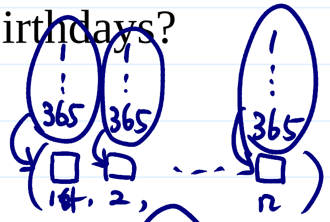
- $\Omega = \{0, 00, 1, 2, 3, 4, \ldots, 35, 36\}$
- $P(\{\text{Red Outcome}\}) = 18/38 = 9/19$.

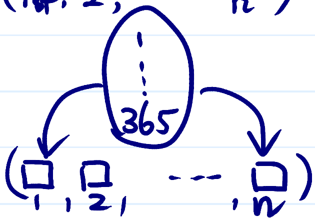
➢ Example (Birthday Problem): n people gather at a party. What is the probability that they all have different birthdays?

- Ω = lists of n from {1, 2, 3, …, 365}
- A = {all permutations}
- $P_n(A) = (365)_n / 365^n$

#Ω

#A





n	8	16	24	32	40
$P_n(A)$	.926	.716	.462 ~0.5	.247	.109

→ smaller

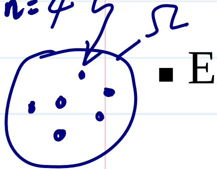
➢ Inadequacy of the Classical Approach

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

- It requires:
  - Finite Ω
  - Symmetric Outcomes

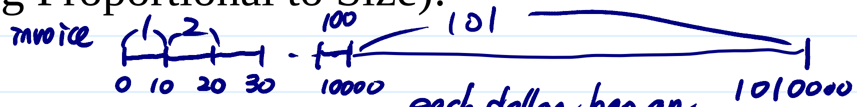
i.e. all outcomes in Ω are equally likely to occur → $\omega \in \Omega$, $P(\omega) = \frac{1}{\#\Omega}$

N=101
n=4
{10,50,78,91}



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

- Example (Sampling Proportional to Size):
  - N invoices.
  - Sample $n < N$.
  - Pick large ones with higher probability.
  - Note: Finite Ω, but *non equally-likely* outcomes.

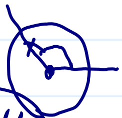


each dollar has an equal chance of being selected, not each invoices.

- Example (Waiting for a success):
  - Play roulette until a win.
  - Ω = {1, 2, 3, …} → infinite, countable set
  - P = ??

discrete sample space
$P: 2^{\Omega} \to [0,1]$

- Example (Uniform Spinner):
  - Random Angle (in radians).
  - Ω = (−π, π] → infinite, uncountable set
  - P = ??



continuous sample space
P: σ-field → [0,1]

- The Modern Approach

2^{Ω}

抽象化

➢ A *probability measure* on Ω is a function *P* from subsets of Ω to the real number [0,1] that satisfies the following axioms 公理

- self-evident truth
- a statement generally acepted as true.

check if the 3 axioms hold in classical approach.

(Ax1) Non-negativity. For any event *A*, *P*(*A*)≥0.

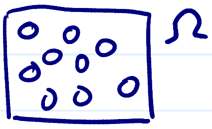
(Ax2) Total one. *P*(Ω)=1

countably 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

disjoint.

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

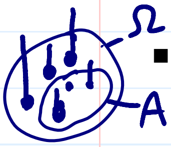


- Notes:
  - These axioms restrict probabilities, but do not define them.
  - Probability is a property of events.

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

9/20 ↓

➢ Define Probability Measures in a Discrete Sample Space.



- **Q**: Is it required to define probabilities on every events? (e.g., *n* possible outcomes in Ω, 2^n-1 possible events) $\because P(\Omega)=1$.
- Suppose Ω = {$\omega_1, \omega_2, \ldots$}, *finite* or *countably infinite*, let $p: \Omega \to R$ satisfy

$P:2^{\Omega}\to\mathbb{R}$ c.f.

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

Let

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

for *A*⊂Ω, then *P* is a probability measure. (exercise) examine Ax1~3

(**Q**: how to define *p*?) → see the arguments about subjective & objective interpretation (LNp.2-15~17)

- Example: In the classical approach, $p(\omega) = 1/\#\Omega$. For example, throw a fair dice, Ω={1, …, 6}, *p*(1)=…=*p*(6)=1/6 and *P*(odd)=*P*({1, 3, 5})=*p*(1)+*p*(3)+*p*(5)=3/6=1/2.

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

- Example (non equally-likely events): Throwing an unfair dice might have *p*(1)=3/8, *p*(2)=*p*(3)=…=*p*(6)=1/8, and *P*(odd)=*P*({1, 3, 5})=*p*(1)+*p*(3)+*p*(5)= 5/8. (c.f., Example in LNp.2-4)
- Example (Waiting for Success – Play Roulette Until a Win):
  - Let *r*=9/19 and *q*=1–*r*=10/19
  - Ω = {1, 2, 3, …} ← infinite, countable.
  - Intuitively, $p(1)=r$, $p(2)=qr$, $p(3)=q^2r$, …, $p(n)=q^{n-1}r$, … >0, and

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

Q: what if we directly define P on 2^{Ω}?

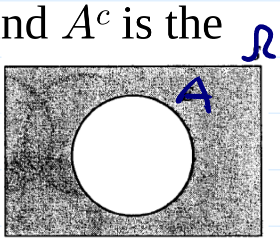
- For an event $A \subset \Omega$, let

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

For example, Odd={1, 3, 5, 7, …}

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

- Some Consequences of the 3 Axioms → 三原色 RGB → any color
  - Proposition: If A is an event in a sample space Ω and A^c is the complement of A, then



$$③ — P(A^c) = 1 - P(A).$$

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

$$1 = P(\Omega) \underset{\text{by}②}{=} P(A^c) + P(A)$$

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

$$① — P(\emptyset) = 0.$$

proof: In (Ax3) Let $A_1 = \Omega$, $A_2 \cdots A_n, \cdots = \phi$

$$\underset{1}{\overset{\shortparallel}{P(\Omega)}} = \underset{1}{\overset{\shortparallel}{P(\Omega)}} + \sum_{n=2}^{\infty} P(\phi)$$

$$P(\phi) = 0$$

finite version of Axiom 3

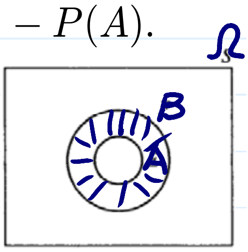
  - Proportion: For any finite sequence of mutually exclusive events $A_1, A_2, \ldots, A_n$,

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

proof: In (Ax3), Let $A_{n+1} = A_{n+2} = \cdots = \phi$. Then, ∵ $P(\phi) = 0$

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

$$P(A) \le P(B) \quad \text{and} \quad P(B - A) = P(B \cap A^c) = P(B) - P(A).$$



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

$$\underbrace{P(B) = P(A) + P(B \cap A^c)}_{\text{by}②} \overset{\text{by}(Ax1)}{\ge} P(A)$$

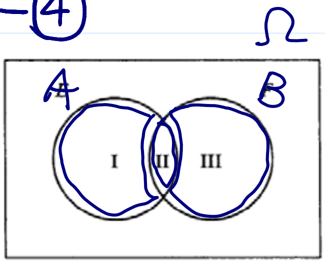
This is Axiom1 in textbook.

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

$$\underset{Ax1}{\boxed{0 \le P(A)}} \le 1. = P(\Omega) \ \& \ A \subset \Omega$$

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

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). — ④$$



proof: $A \cup B = \text{I} \cup \text{II} \cup \text{III}$,& I, II, III mutually exclusive

$$P(A \cup B) \overset{\text{by}②}{=} P(\text{I}) + P(\text{II}) + P(\text{III})$$

$$P(A) = P(\text{I}) + P(\text{II}) \ , \quad P(A \cap B) = P(\text{II})$$

$$P(B) = P(\text{II}) + P(\text{III})$$

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

$$P(A_1\cup\cdots\cup A_n)\le P(A_1)+\cdots+P(A_n). \xleftrightarrow{\text{c.f.}} ② \ \& \ ④$$

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

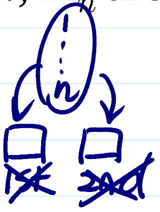
by ④ $\le P(A_1\cup\cdots\cup A_{n-1})+P(A_{n-1})+P(A_n)\le\cdots$

generalization of ④

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

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

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



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

$\to \binom{n}{2}$ different (i, j)

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

$$\cdots = \cdots$$

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

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

$$\cdots = \cdots$$

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

An intuitive proof for discrete sample space.

$p(w)$ counted m times in σ_1

$\vdots \quad \binom{m}{2} \quad \vdots \quad \vdots \quad \sigma_2$

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

$$\boxed{-\binom{m}{0}} + \binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \binom{m}{4} \cdots + (-1)^{k+1}\binom{m}{k} + \cdots + (-1)^{n+1}\binom{m}{m} = 0$$

$-1 = -\binom{m}{0}$

then,

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

---

proof: prove by induction.

$n=2$, it holds ← by ④

$n=3$, $P(\underline{A_1\cup A_2}\cup \underline{A_3}) \overset{④}{=} P(A_1\cup A_2)+P(A_3)-P(\overbrace{(A_1\cup A_2)\cap A_3}^{(A_1\cap A_3)\cup(A_2\cap A_3)})$

$\| \leftarrow$ by ④ $P(A_1)+P(A_2)-P(A_1\cap A_2)$ | $\| \leftarrow$ 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.

- Notes:
  - ▫ There are $\binom{n}{k}$ summands in σ_k
  - ▫ In symmetric examples, e.g. $k=2$, $P(A_1\cap A_2)=P(A_{i_1}\cap A_{i_2})\ \cdots$

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

  - ▫ It can be shown that

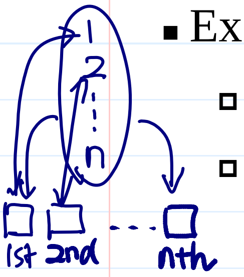
for proof see textbook.

$$P(A_1\cup\cdots\cup A_n) \le \sigma_1$$ 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.

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



jth position ⇔ item j.

- Example (The Matching Problem).
  - Applications: (a) Taste Testing. (b) Gift Exchange.
  - Let Ω be all permutations ω = (i_1, …, i_n) of 1, 2, …, n. Thus, #Ω = n!.

    (3, 1, 5, ⋯) 1st 2nd 3rd.

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

    item position

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

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

- By symmetry,

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

- Here,

$$P(A_1) = \frac{1\times(n-1)!}{n!} = \frac{1}{n},$$
$$P(A_1\cap A_2) = \frac{(n-2)!}{n!} = \frac{1}{(n)_2},$$
$$\cdots = \cdots,$$
$$P(A_1\cap\cdots\cap A_k) = \frac{(n-k)!}{n!} = \frac{1}{(n)_k}.$$

for k = 1, …, n.

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

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

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

when $n\to\infty$

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

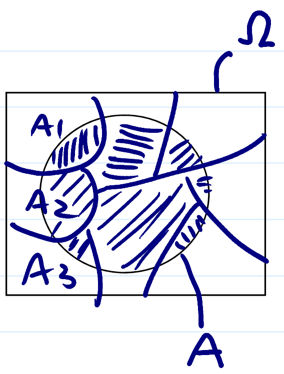
- Note: approximation accurate to 3 decimal places if $n \geq 6$.

➢ Proportion: If A_1, A_2, …, is a *partition* of Ω, i.e.,

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: $A = A\cap\Omega = A\cap\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty}\underbrace{(A\cap A_i)}_{\text{mutually exclusive}}$

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

- Monotone Sequences

→check example in LNp.2-4

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

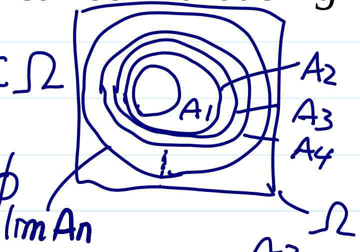
C.f. define P.M. for discrete sample space LNp.2-6

- Definition: A sequence of events A_1, A_2, …, is called *increasing* if

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

and *decreasing* if

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



The limit of an increasing sequence is defined as

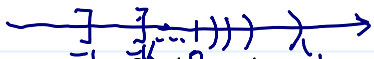
$$\lim_{n\to\infty} A_n = \cup_{i=1}^{\infty} A_i$$

and the limit of an decreasing sequence is

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

- Example: If Ω=R and $A_k = (-\infty, 1/k)$, then A_k's are decreasing and

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



(exercise) $A_k = (-\infty, -\frac{1}{k}] \uparrow$. Ans $\lim A_n = (-\infty, 0)$

- Proportion: If A_1, A_2, …, is increasing or decreasing, then

LNp.2-2 DeMorgan's Law.

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

$(\bigcup_{i=1}^{\infty} A_i)^c = \bigcap_{i=1}^{\infty} A_i^c = \lim A_i^c$

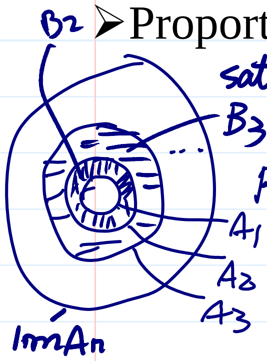
$(\bigcap A_i)^c = \bigcup A_i^c = \lim A_i^c$

- Proportion: If A_1, A_2, …, is increasing or decreasing, then

satisfy (Ax1)~(Ax3)

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

exchangable. ⑤ $B_i \cap B_j = \phi$



proof: ① An increasing

Let $B_n = A_n \cap A_{n-1}^c$, $B_1 = A_1$ $\Rightarrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$, $\bigcup_{n=1}^{k} B_n = \bigcup_{n=1}^{k} A_n$

$P(\lim A_n) = P(\bigcup_{n=1}^{\infty} A_n) = P(\bigcup_{n=1}^{\infty} B_n) \underset{(Ax3)}{=} \sum_{n=1}^{\infty} P(B_n) = \lim_{k\to\infty} \sum_{n=1}^{k} P(B_n)$

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

$= \lim_{k\to\infty} P(A_k)$

② An decreasing, ⇒ A_n^c increasing.

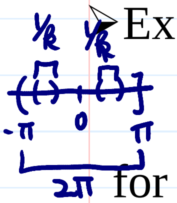
$1 - P(\lim A_n) = 1 - P(\bigcap_{n=1}^{\infty} A_n) = P((\bigcap_{n=1}^{\infty} A_n)^c)$

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

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

- Example (Uniform Spinner): Let Ω = (–π, π]. Define

continuous sample space



C.f. P.M. for discrete sample space

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

play a role similar to $w_1, w_2, \cdots$ in small p for discrete sample space

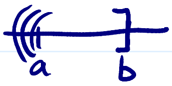
a, b 有理數

for subintervals (a, b]⊂Ω. Then, extend P to other subsets using the 3 axioms. For example, if –π < a < b < π,

$$\begin{aligned} P([a,b]) &= P\left(\left(\bigcap_{k=1}^{\infty}(a-\frac{1}{k},b]\right)\cap\Omega\right) = P\left(\bigcap_{k=1}^{\infty}\left((a-\frac{1}{k},b]\cap\Omega\right)\right) \\ &= \lim_{k\to\infty} P\left((a-\frac{1}{k},b]\cap\Omega\right) \\ &= \lim_{k\to\infty}\frac{1}{2\pi}(b-a+\frac{1}{k}) = \frac{b-a}{2\pi}. \end{aligned}$$

Note $[a,b] \neq (a,b]$ ⇒ P not defined in previous slide.

by ⑤



$A_k=(a-\frac{1}{k},b] \downarrow, \lim_{k\to\infty} A_k=[a,b]$

$=P([a,b])$ ← Note: $P([a,b])$ is defined through $P((a,b])$

- 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
    $$P(C) = \sum_{i=1}^{\infty} P(\{\omega_i\}) = 0 + 0 + \cdots = 0.$$
    $\overset{\shortparallel}{\cup\{w_i\}}$ a countable set.
  - ☐ The probability of a rational outcome is zero

c.f. discrete sample space

- Objective vs. Subjective Interpretations of Probability
  - **Q**: What do we mean if we say that the probability of rain tomorrow is 40%? ← how to interpret it?

    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

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

    Then,
    $$P(A) \equiv \lim_{N\to\infty}\frac{N_A}{N}.$$
    ← frequency
  - Example (Coin Tossing):

N	100	1000	10000	100000
N_H	55	493	5143	50329
N_H/N	.550	.493	.514	.503

→0.5

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

※ odds of an event → 勝敗率

$O(A) = \frac{P(A)}{P(A^c)} = \frac{P(A)}{1-P(A)}$

$P(A) = \frac{O(A)}{1+O(A)}$

- The Subjective Interpretation
  - Strategy: Assess probabilities by imagining bets
  - Examples:
    - ☐ Peter is willing to give two to one odds that it will rain tomorrow. His subjective probability for rain tomorrow is at least 2/3

2 : 1

A

$O(A) = \frac{P(A)}{P(A^c)} \geq \frac{2}{1}$

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

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

$P(A) \geq \frac{2}{3}$

$O(A^c) = \frac{P(A^c)}{P(A)} \geq \frac{1}{2}$, $P(A^c) = 1 - P(A)$

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

- Paul accepts the bet. His subjective probability for rain tomorrow is at most ~~1~~2/3

present personal degree of belief.

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

e.g. how likely you think Shakespeare wrote Hamlet?

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

Summary

- probability space
  - sample space Ω
  - collection of events (2^Ω) — subsets
    - set operations — $\cup$, $\cap$, A^c
    - rule of set operations.
  - probability measure $P: 2^\Omega \to [0,1]$ — interpretation
    - subjective
    - objective
  - classical approach: $P(A) = \frac{\#A}{\#\Omega}$ — define
  - modern approach: 3 axioms — consequence
    - proposition 1: $P(A^c) = 1 - P(A)$
    - $P(\phi) = 0$
    - ⋮
    - proposition: $P(A_1 \cup \cdots \cup A_n) = \sigma_1 - \sigma_2 + \cdots$
  - restrict P but not define.
    - discrete Ω → 1st, define a small p on $\omega_1, \omega_2, \omega_3 \cdots$
      - then extend ← $P(A) = \sum_{\omega \in A} p(\omega)$
    - continuous Ω
      - → 1st, define prob. on $(a,b]$, $a, b \in \mathbb{Q}$
      - then extend ← monotone sequence of sets.

❖ **Reading**: textbook, Sec. 2.3~2.7