- The odd of event B given A:

evaluated under the probability measure $P(\cdot|A)$ → $P(\cdot)$

$$o(B|A) \equiv \frac{P(B|A)}{P(B^c|A)} = \frac{P(A\cap B)/\cancel{P(A)}}{P(A\cap B^c)/\cancel{P(A)}} = \frac{\boxed{P(B)}\cdot P(A|B)}{\boxed{P(B^c)}\cdot P(A|B^c)}$$

and

$$O(\cdot) \xrightarrow[\text{after A occurs}]{\text{update}} O(\cdot|A) \quad o(B|A) = o(B) \times \frac{P(A|\boxed{B})}{P(A|\boxed{B^c})}$$

cf. → Bayes' rule $P(\cdot) \xrightarrow{\text{update}} P(\cdot|A)$

❖ **Reading**: textbook, Sec 3.1, 3.2, 3.3, 3.5 10/5

獨立 —— Independence

→ Recall, with replacement example (LNp.4-7)

- Definition (independence for 2-events case): Two events A and B are said to be *independent* if and only if

It's a property defined on events ↕ cf. The "independent" defined on random variables (future lecture)

for calculation purpose →

$$P(A\cap B) = P(A)P(B).$$

cf. → ① in LNp.4-4

Otherwise, they are said to be *dependent*.

- Notes. If $P(A) > 0$, events A and B are independent if and only if

for interpretation purpose →

$$P(B|\underset{\uparrow\ \text{new information}}{A}) = \frac{P(A\cap B)}{P(A)} = \frac{\cancel{P(A)}P(B)}{\cancel{P(A)}} \,=\, P(B),$$

similarly, if $P(B) > 0$, if and only if $P(A|\underset{\text{new information}}{B}) = P(A)$.

Q: How to interpret the equality? new information

- Example (Sampling 2 balls, LNp.4-6~7). Events A and B were "independent" for sampling with replacement, but "dependent" for sampling without replacement. $P(B|A) \overset{?}{=} P(B)$

- Example (Cards): If a card is selected from a standard deck, let
  - $A = \{\text{ace}\}$ and $B=\{\text{spade}\}$. Then,
  - $P(A) = \frac{4}{52} = \frac{1}{13}$, $P(B) = \frac{13}{52} = \frac{1}{4}$,

    $P(A\cap B) = \frac{1}{52} = P(A)P(B)$
  - Face and Suit are independent

can be changed to any other face

can be changed to any other suit

大小 → 花色 →

Recall ② in LNp.4-4

Note. $P(B|A) > P(B)$, (i.e., $P(A\cap B) > P(A)P(B)$) iff $P(B|A^c) < P(B)$ (i.e., $P(A^c\cap B) < P(A^c)P(B)$)

- Theorem (Independence and Complements, 2-events case).

$P(B|A) = P(B)$ → If A and B are independent, then so are A^c and B. ← $P(B|A^c) = P(B)$

proof. $B = B\cap\Omega = B\cap(A\cup A^c) = \underbrace{(B\cap A)\cup(B\cap A^c)}$ ← mutually exclusive

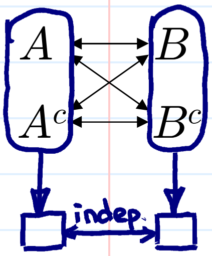
$P(B) = P(B\cap A) + P(B\cap A^c)$

∵ indep. $\overset{*}{\underset{(>)}{=}} P(A)P(B) + P(B\cap A^c)$

$\Rightarrow P(B\cap A^c) \underset{(<)}{=} P(B) - P(A)P(B) = P(B)[1-P(A)]$ $\overset{P(B)P(A^c)}{\Vert}$

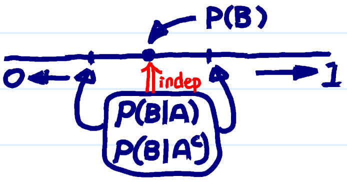
A & B independent

- Corollary: If A and B^c are independent, so are A^c and B^c
- Corollary: If A and B are independent and $0<P(A)<1$, $0<P(B)<1$, then



$$\begin{cases} P(B) = P(B|A) = P(B|A^c), \\ P(B^c) = P(B^c|A) = P(B^c|A^c), \end{cases}$$

$$\begin{cases} P(A) = P(A|B) = P(A|B^c), \\ P(A^c) = P(A^c|B) = P(A^c|B^c), \end{cases}$$



FYI. P(B) and P(B|A) not enough to decide $P(B|A^c)$, also need to know P(A). It's because

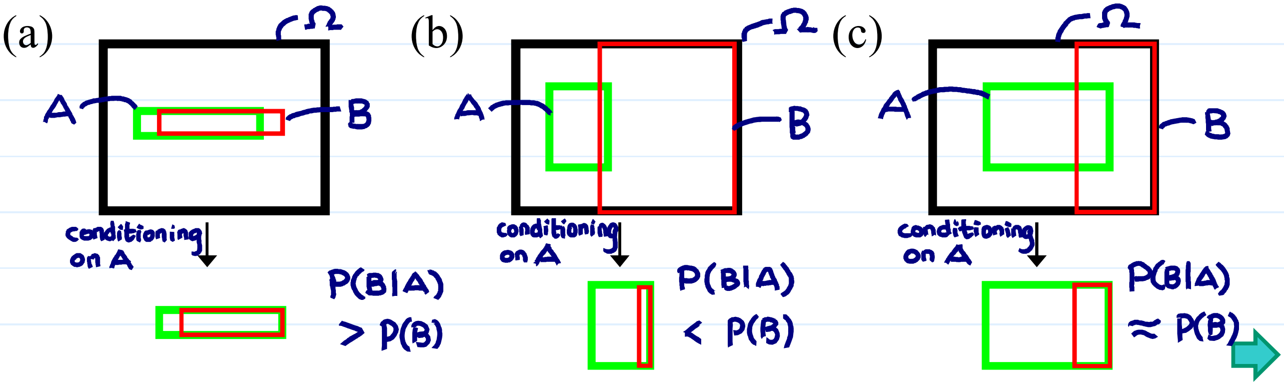
$$P(B|A^c) = \frac{P(B) - P(A)P(B|A)}{1-P(A)}$$

Q: What do these equalities say?

➢ Example. P(A) & P(A^c): weights in ② (LNp.4-4)

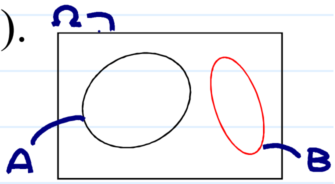
- **Q**: Which of the following graphs represents "green and red events are independent" (assume probability ∝ area)?

Ans.(c)



- **Q**: Let green event={graduate from Tsing-Hua University}, red event={your future dream will come true}. Which of the graphs would you prefer? Ans. usually (a)
- **Q**: What do we prefer? independent? or dependent? ← It depends

➢ Theorem (Independence and Mutually Exclusive). If A and B are mutually exclusive and $P(A)>0$, $P(B)>0$, then A and B are dependent since



$$0 = \frac{P(\phi)}{P(A)} = \frac{P(A\cap B)}{P(A)} = P(B|A) = 0 \neq P(B).$$

What if $B=\phi$?

cf. indep. for 2 events (LNp.4-19)

⊙ Definition (independence for n-events case). Events $A_1, \ldots, A_n$ are said to be *pairwise independent* iff for all $1\le i<j\le n$;

weak

cf.

$$P(A_i \cap A_j) = P(A_i)P(A_j),$$

strong

$A_1, \ldots, A_n$ are said to be *mutually independent* iff for k=2, …, n,

equality holds for any K sets in {A$_1$, …, A$_n$}

$$P(A_{i_1} \cap A_{i_2}) = P(A_{i_1})P(A_{i_2}), \quad \text{for } 1 \le i_1 < i_2 \le n,$$

$$P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) = P(A_{i_1})P(A_{i_2})P(A_{i_3}), \quad \text{for } 1 \le i_1 < i_2 < i_3 \le n,$$

$$\ldots,$$

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}), \quad \text{for } 1 \le i_1 < \cdots < i_k \le n,$$

$$\ldots \quad P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$$

➢ Note:

$A_{t_1}, \cdots, A_{t_k}$ are mutually indep.

- Suppose $A_1, \ldots, A_n$ are mutually independent. For $1 \le r < k \le n$, and different $t_1, \ldots, t_r, t_{r+1}, \ldots, t_k \in \{1, 2, \ldots, n\}$, (exercise)

for interpretation purpose →

$$P(A_{t_1} \cap \cdots \cap A_{t_r} \mid A_{t_{r+1}} \cap \cdots \cap A_{t_k}) = P(A_{t_1} \cap \cdots \cap A_{t_r}).$$

cf. Thm in LNp.4-27

- Mutual independence implies pairwise independence; but, the converse statement is usually not true. ← an example in LNp.4-24
- "n events are independent" means "mutually independent"

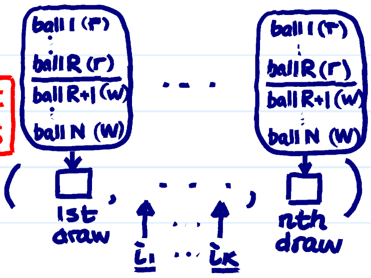
➢ Example (Sampling With Replacement)

- A sample of n balls is drawn with replacement from an urn containing R red and $N-R$ white balls
- Let A_k={red on the k^{th} draw}, then

$\frac{RN^{n-1}}{N^n} = \#A/\#\Omega = P(A_k) = R/N, \quad k=1, \ldots, n.$

symmetric outcomes → Ω

$\#\Omega = N^n$



- For all $1 \le i_1 < \cdots < i_k \le n$, where $k=2, \ldots, n$,

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = \frac{R^k N^{n-k}}{\boxed{N^n}} = \left(\frac{R}{N}\right)^k = P(A_{i_1}) \cdots P(A_{i_k}),$$

\# of lists when balls are labelled

$\Rightarrow A_1, \ldots, A_n$ are mutually independent

➢ Example. Draw one card from a standard deck.

- Let A={Spades or Clubs},
  B={Hearts or Clubs},
  C={Diamonds or Clubs}.
- $P(A) = 26/52 = 1/2$, similarly, $P(B) = P(C) = 1/2$.
- $P(A \cap B) = P(\{\text{Clubs}\}) = \frac{13}{52} = \frac{1}{4} = P(A)P(B)$, similarly,

$$P(A \cap C) = 1/4 = P(A)P(C), \quad P(B \cap C) = 1/4 = P(B)P(C).$$

$\Rightarrow A$, B, and C are pairwise independent

- However,

$$P(A \cap B \cap C) = P(\{\text{Clubs}\}) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C),$$

$P(A|B \cap C) = \frac{P(\{\text{Clubs}\})}{P(\{\text{Clubs}\})} = 1 \neq P(A)$

$\Rightarrow A$, B, and C are *not* mutually independent

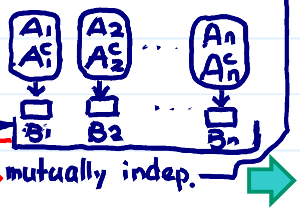
$P(B_{t_1} \cap \cdots \cap B_{t_r} | B_{t_{r+1}} \cap \cdots \cap B_{t_k}) = P(B_{t_1} \cap \cdots \cap B_{t_r})$

➢ Theorem (Independence and Complements, n-events case). ↑cf. Note 1 (LNp.4-23)

$A_1, \ldots, A_n$ are mutually independent if and only if

$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \cdots P(B_n),$$

where B_i is either A_i or A_i^c, for $i=1, \ldots, n$.



mutually indep.

outline of proof.

($\Rightarrow$ only if). Apply Theorem in LNp. 4-20 & by induction

$A_i \xleftrightarrow{\text{indep.}} B$

$A_i^c \xleftrightarrow{\text{indep.}} B$

$P(A_i^c \cap B) = P(A_i^c) \cdot P(B)$

$A_1 \cap \cdots \cap A_n \to \underbrace{A_1 \cap \cdots \cap A_{i_1-1}} \cap A_{i_1}^c \cap \underbrace{A_{i_1+1} \cap \cdots \cap A_n}, \to$ they are mutually indep.

i_1th — true for any k out of n sets — one set ← B — i_2th

$\to \underbrace{A_1 \cap \cdots \cap A_{i_1-1} \cap A_{i_1}^c \cap A_{i_1+1} \cap \cdots \cap A_{i_2-1}} \cap A_{i_2}^c \cap \underbrace{A_{i_2+1} \cap \cdots \cap A_n}, \to \cdots$

one set — i_3th

($\Leftarrow$ if) trivial. $P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$.

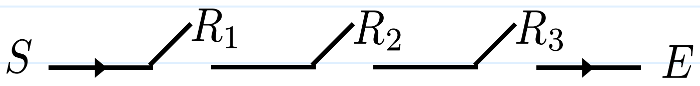
partition of Ω

$$P(A_1 \cap \cdots \cap A_{n-1}) = P(A_1 \cap \cdots \cap A_{n-1} \cap A_n) + P(A_1 \cap \cdots \cap A_{n-1} \cap A_n^c)$$
$$= P(A_1) \cdots P(A_{n-1}) P(A_n) + P(A_1) \cdots P(A_{n-1}) P(A_n^c)$$
$$= P(A_1) \cdots P(A_{n-1}) \left[P(A_n) + P(A_n^c) \right] \quad \boxed{=1}$$

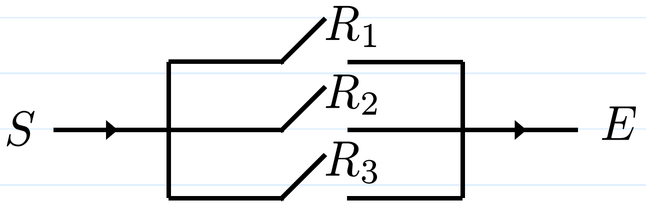
Similarly, $P(A_1 \cap \cdots \cap A_{n-2} \cap A_{n-1}^c) = P(A_1) \cdots P(A_{n-2}) P(A_{n-1}^c)$

- Example (Series and Parallel Connections of Relays).

Series Connection
串聯



Parallel Connection
並聯



- The Story. For n electrical relays $R_1, \ldots, R_n$, let

  $P(A_k) \longleftrightarrow A_k = \{R_k \text{ works properly}\}$, leads to a simpler system

  $k=1,\ldots,n$, and suppose that $A_1, \ldots, A_n$ are independent. meaning?

all events occur →
- Series Connection. The probability that current can flow from S to E (corresponding to the event $A_1 \cap \cdots \cap A_n$) $\subseteq A_i$ is

  $$P(A_1 \cap \cdots \cap A_n) \overset{\because \text{ indep}}{=} P(A_1) \cdots P(A_n) \le P(A_i)$$

  cf. LNp.4-11 → $P(A_1 \cap \cdots \cap A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1})$

at least one event occur →
- Parallel Connection. The probability that current can flow from S to E (corresponding to the event $A_i \subseteq$ → $A_1 \cup \cdots \cup A_n$) is

  $\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c$

  $$P(A_1 \cup \cdots \cup A_n) = 1 - P(A_1^c \cap \cdots \cap A_n^c)$$
  $$\overset{\because \text{ indep.}}{=} 1 - P(A_1^c) \cdots P(A_n^c) = 1 - \prod_{k=1}^{n} [1 - P(A_k)] \quad \ge P(A_i)$$

  cf. LNp.3-12 → $P(A_1 \cup \cdots \cup A_n) = \sigma_1 - \sigma_2 + \sigma_3 - \sigma_4 + \cdots$

★ "Independence" often simplifies the calculation of probability.

▫ Combination of Series and Parallel Connections

$B_1=\{T_1 \text{ works}\}$

$P(B_1)=1-[(1-P(A_1))(1-P(A_2))]$

$B_2=\{T_2 \text{ works}\}$

$P(B_2)=P(B_1)\cdot P(A_3)$

$B_3=\{T_3 \text{ works}\}$

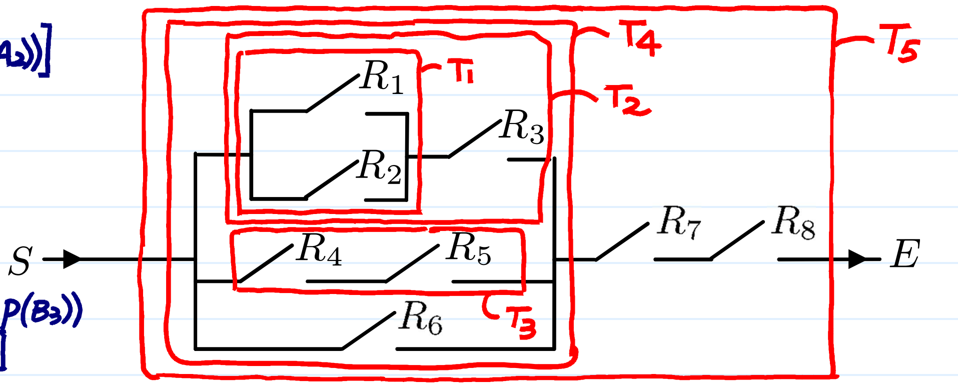
$P(B_3)=P(A_4)P(A_5)$

$B_4=\{T_4 \text{ works}\}$

$P(B_4)=1-[(1-P(B_2))(1-P(B_3))(1-P(A_6))]$

$B_5=\{T_5 \text{ works}\}$

$P(B_5)=P(B_4)\cdot P(A_7)\cdot P(A_8)$



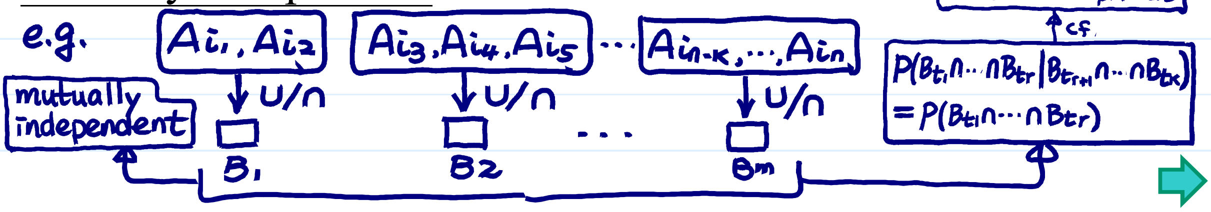
$$\underbrace{\Big(\underbrace{\big(\underbrace{(\underbrace{A_1\cup A_2}_{B_1})\cap A_3}_{B_2}\big)\cup(\underbrace{A_4\cap A_5}_{B_3})\cup A_6}_{B_4}\Big)\cap A_7\cap A_8}_{B_5}$$

Q: Why B_1, A_3 indep.?
B_2, B_3, A_6 indep.?
B_4, A_7, A_8 indep.?

- Theorem. If $A_1, \ldots, A_n$ are mutually independent and $B_1, \ldots, B_m$, $m\leq n$, are formed by taking unions or intersections of mutually exclusive subgroups of $A_1, \ldots, A_n$, then $B_1, \ldots, B_m$ are mutually independent.

$(i_1, i_2, \cdots, i_n)$: a permutation of $(1,2,\cdots,n)$

Note 1 in LNp.4-23



sketch of proof.

(i) For $m=2$, WLOG, suppose that

∵ indep

$B_1=A_1\cap\cdots\cap A_j$, $B_2=A_k\cap\cdots\cap A_n$, $1\leq j<k\leq n$ ($k = j+1$)

$\Rightarrow P(B_1)=P(A_1)\cdots P(A_j)$ & $P(B_2)=P(A_k)\cdots P(A_n)$ &

$P(B_1\cap B_2)=P(A_1\cap\cdots\cap A_j\cap A_k\cap\cdots\cap A_n)$

$=P(A_1)\cdots P(A_j)\cdot P(A_k)\cdots P(A_n)=P(B_1)P(B_2)$

(ii) Next, if $B_1=A_1\cap\cdots\cap A_j$, $1\leq j<k\leq n$

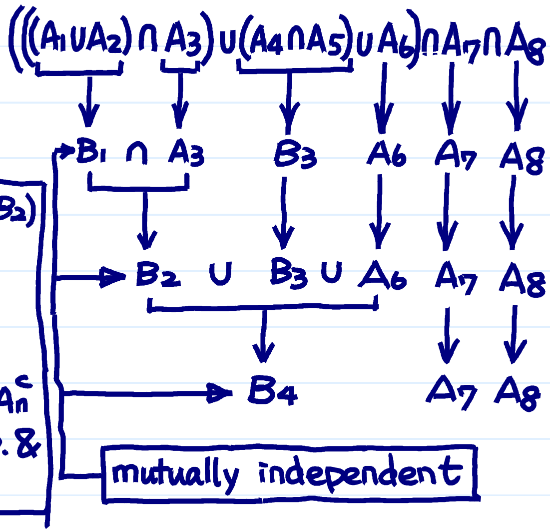
$B_2=A_k\cup\cdots\cup A_n$

$\Rightarrow B_2^c=A_k^c\cap\cdots\cap A_n^c$

Then, B_1 & B_2^c are indep. (∵ $A_1,\cdots,A_j,A_k^c,\cdots,A_n^c$ mutually indep. & by (i))

$\Rightarrow B_1$ & B_2 are indep.

(iii) The other cases are similar.



- Definition (conditional independence): Events $B_1, \ldots, B_n$ are (*pairwise* or *mutually*) *independent* under the probability measure $P(\cdot|A)$. ← Recall. $P(\cdot)$ & $P(\cdot|A)$ in LNp.4-3

income & gender | professional job

  - e.g., B_1 and B_2 are conditionally independent given A iff

$$P(B_1\cap B_2|A)=P(B_1|A)P(B_2|A),$$

or, equivalently, $\frac{P(B_1\cap B_2|A)}{P(B_2|A)}=\frac{P(B_1\cap B_2\cap A)/\cancel{P(A)}}{P(B_2\cap A)/\cancel{P(A)}}=P(B_1|B_2\cap A).$

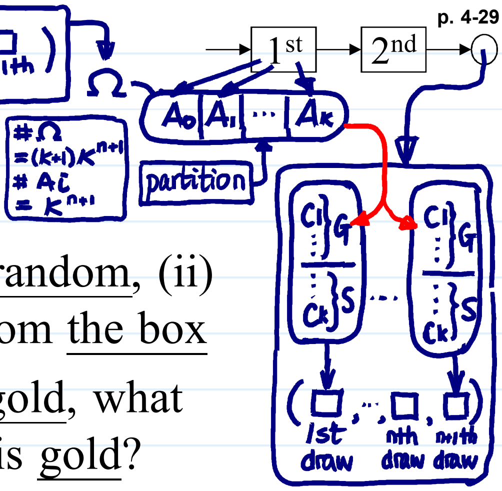
$$P(B_1|B_2\cap A)=P(B_1|A) \text{ or } P(B_2|B_1\cap A)=P(B_2|A)$$

$P(\cdot)$
↓ update
$P(\cdot|A)$
↓ update
$P(\cdot|B_2\cap A)$

cf. Example in LNp.4-15

Example (Gold Coins):

$\omega = (\underset{\text{Box}}{\triangle}, \underset{\text{1st}}{\square}, \cdots, \underset{\text{n+1th}}{\square})$



- The Story.
  - ☐ Box i contains i gold coins and $k-i$ silver coins, i=0,1,…,k.
  - ☐ Experiment: (i) Select a box at random, (ii) Draw coins *with replacement* from the box
- **Q**: Given that first n draws are all gold, what is the probability that $(n+1)^{\text{st}}$ draw is gold?
  - ☐ Let A_i = {Box i is selected}, B = {first n draws are gold}, C = {$(n+1)^{\text{st}}$ draw is gold}
  - ☐ By applying law of total probability on $P(\cdot|B)$,

By ⑤ under P(•|B)

$$P(C|B) \overset{\downarrow}{=} \sum_{i=0}^{k} P(A_i|B)P(C|A_i \cap B)$$

evaluated under P(•|Ai)

  - ☐ Because B and C are *conditionally independent* given A_i,

By ⑥ under P(•)

$$P(C|A_i \cap B) \overset{\downarrow}{=} P(C|A_i) = i/k$$

B & C independent under P(•|Ai)

  - ☐ By Bayes' rule,

$$\frac{1}{k+1} = P(A_i) \xrightarrow{\text{update}} P(A_i|B) \overset{\downarrow}{=} \frac{P(A_i)P(B|A_i)}{\sum_{j=0}^{k} P(A_j)P(B|A_j)} = \frac{[\cancel{1/(k+1)}](i/k)^n}{\sum_{j=0}^{k}[\cancel{1/(k+1)}](j/k)^n}$$

what happens if n↑∞?

賭徒謬誤(LNp.1-8) Bayesian ⟷ Frequentist

  - ☐ Hence, $P(C|B) = \sum_{i=0}^{k}(i/k)^{n+1} \Big/ \sum_{j=0}^{k}(j/k)^n \overset{?}{\neq} P(C)$ (exercise)

by ⑤ under P(•)

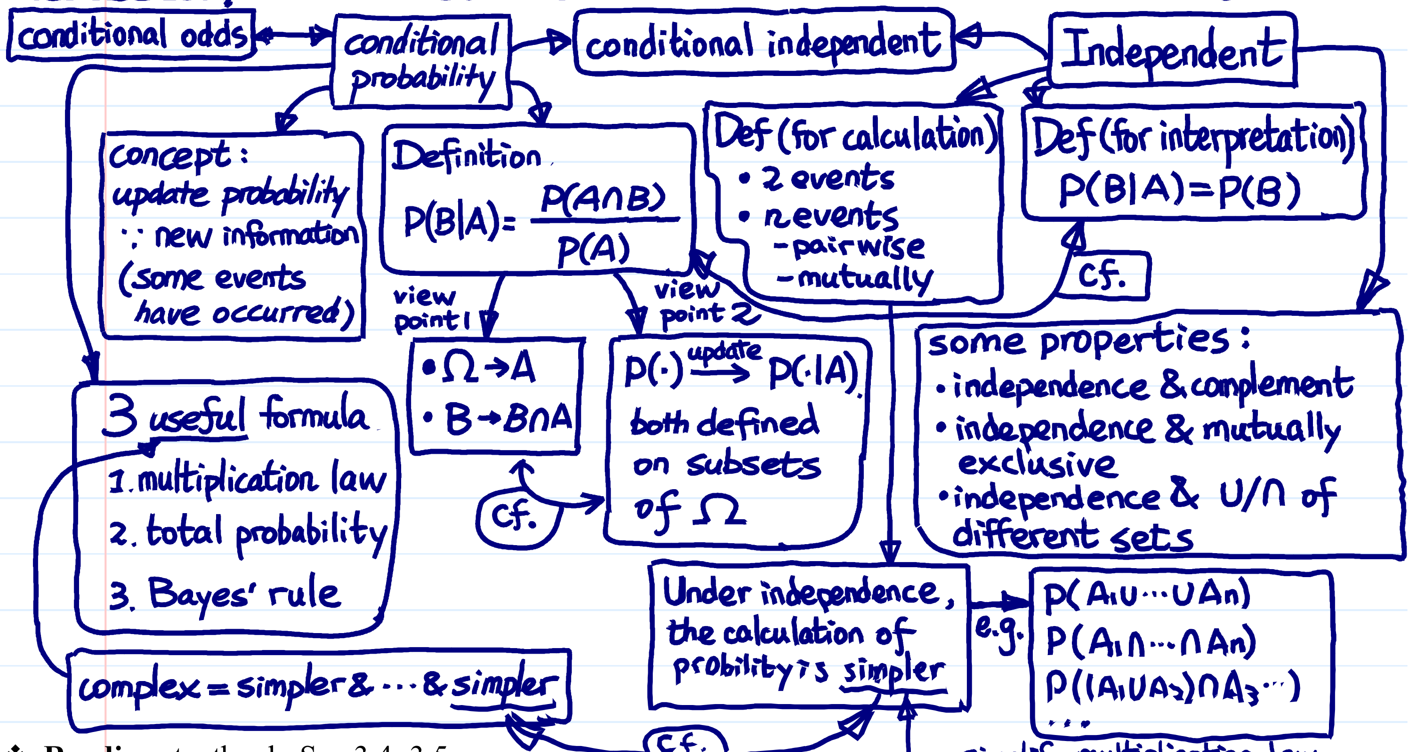
increases to 1, as n ↑ ∞

cf.

- **Q**: Are the events B and C independent under $P(\cdot)$? Ans. No.

Intuition is ?

Summary


conditional odds

conditional probability

conditional independent

Independent

concept: update probability ∵ new information (some events have occurred)

Definition $P(B|A) = \frac{P(A \cap B)}{P(A)}$

view point 1
- Ω→A
- B→B∩A

view point 2
$P(\cdot) \xrightarrow{\text{update}} P(\cdot|A)$ both defined on subsets of Ω

cf.

Def (for calculation)
- 2 events
- n events
  - pairwise
  - mutually

Def (for interpretation) P(B|A)=P(B)

cf.

3 useful formula
1. multiplication law
2. total probability
3. Bayes' rule

some properties:
- independence & complement
- independence & mutually exclusive
- independence & ∪/∩ of different sets

Under independence, the calculation of probability is simpler

e.g. P(A1∪⋯∪An)
P(A1∩⋯∩An)
P((A1∪A2)∩A3⋯)
⋯

complex = simpler & ⋯ & simpler

cf.

∵ simplify multiplication law

❖ **Reading**: textbook, Sec 3.4, 3.5