Note. In ⑤⑥, A & A^c is a partition

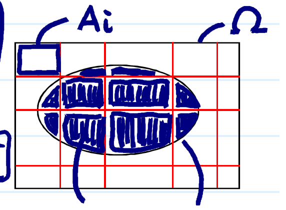
- $P(A) = \binom{n}{k}\frac{1}{(n)_k}p_{n-k} = \frac{p_{n-k}}{k!} \approx \frac{e^{-1}}{k!}$, when n is large

$\bigcup_{i=1}^{m} A_i = \Omega$, $A_i \cap A_j = \phi$, $i \neq j$

2. (Law of Total Probability) Let $A_1, \ldots, A_m$ be a partition of Ω and $P(A_i)>0$, $i=1, \ldots, m$, then for any event $B \subset \Omega$,

meaning → example

⑤ — $P(B) = \sum_{i=1}^{m} P(A_i)P(B|A_i)$. weighted average ($\because \sum P(A_i)=1$)



proof. $B = B \cap \Omega$
$= B \cap (\bigcup_{i=1}^{m} A_i)$
$= \bigcup_{i=1}^{m} (B \cap A_i)$ mutually exclusive

$P(B) = \sum_{i=1}^{m} P(B \cap A_i)$ LNp.3-15
$= \sum_{i=1}^{m} P(A_i) \cdot P(B|A_i)$ by ①

10/3

e.g., $\Omega = \{\omega_1, \cdots, \omega_m\}$, let $A_i = \{\omega_i\}$ like base rate

3. (Bayes' Rule) Let $A_1, \ldots, A_m$ be a partition of Ω and $P(A_i)>0$, $i=1, \ldots, m$. If B is an event such that $P(B)>0$, then for $1 \le j \le m$,

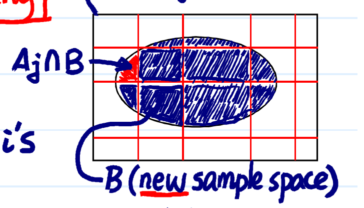
$P(\cdot) \xrightarrow[\text{after } B \text{ occurs}]{\text{update}} P(\cdot|B)$

check graphs in LNp 4-9

⑥ — $P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{m} P(A_i)P(B|A_i)}$. meaning example

not a P.M. w.r.t A_i's

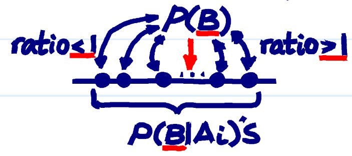
Ω (original sample space)



larger $P(B|A_i)$
⇒ larger $\frac{P(A_i|B)}{P(A_i)}$
~~⇒~~ But, larger $P(A_i|B)$?

proof By definition,

$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)}$ $\frac{\xrightarrow{\text{by ①}} \text{分子}}{\xrightarrow{\text{by ⑤}} \text{分母}}$



Examples

- From Bayesians' viewpoint, data in statistics

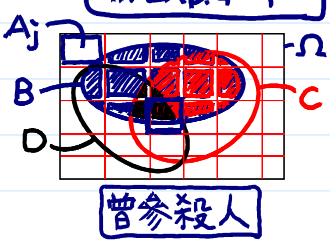
$P(A_j)$=probability of A_j *before* B occurs → *prior* prob. 事前機率

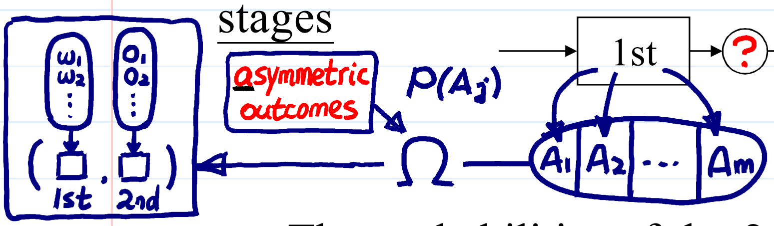
$P(\cdot)$ update ↓ B occurs $P(\cdot|B)$ update ↓ C occurs $P(\cdot|B \cap C)$ ↓ ⋮

$P(A_j|B)$=probability of A_j *after* B occurs → *posterior* prob. 事後機率

⇒ The Bayes' rule tells how to update the probabilities of A_j in light of the new information (i.e., B occurs) check examples in LNp.4-1

- An Application of Bayes' Rule. Suppose that a random experiment consists of two random stages





⇒ $A_1, \cdots, A_m$: a partition of Ω.

- The probabilities of the 2nd-stage results depend on what happened in the 1st stage → key

$P(B|A_j)$ easier to evaluate

- We never see the result of the 1st stage, only the final result → B occurs i.e., it's a random outcome to you

- We may be interested in finding the probability for outcomes in the 1st stage given the final result → $P(A_j|B)$

➢ Example (Gold Coins): $\#\Omega = 6$

1st → ? → 2nd → B

- The Story.

Box1 | Box2 | Box3 — a partition of Ω

Box1, Box2, Box3 — coin1, coin2 — (i) — (ii) 1st, 2nd — symmetric outcomes

coin1, coin2 — (1st exam, 2nd exam)

  - Box 1 contains 2 silver coins.
  - Box 2 contains 1 gold and 1 silver coin.
  - Box 3 contains 2 gold coins.
  - Experiment: (i) Select a box at random and, (ii) Examine the 2 coins in order (assuming all choices are equally likely at each stage)

  1st stage → never see the result. 2nd stage

- **Q**: Given that 1st coin is gold, what is the probability that Box k is selected, k=1, 2, 3? $P(A_k|B)$ ← event of interest

  - Let A_k = {Box k is selected}, B = {1st coin is gold}, — subsets of Ω
  - $$P(B|A_k) = \begin{cases} 0, & \text{if } k=1, \\ 1/2, & \text{if } k=2, \\ 1, & \text{if } k=3. \end{cases}$$ ← easier to evaluate ← cf.

  Note. sum $= 0 + \frac{1}{2} + 1 \neq 1$

  by ⑤ $$P(B) = \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 1 = \frac{1}{2}.$$

  - by ⑥ $$P(A_1|B) = \frac{(1/3)\cdot 0}{1/2} = 0.$$ cf. $P(A_1)=P(A_2)=P(A_3)=1/3$

    Similarly, $P(A_2|B) = 1/3$, $P(A_3|B) = 2/3$.

- **Q**: Given that 1st coin is gold, what is the probability that 2nd coin is gold? (B) — $P(A_3|B)$, $A_3\cap B = C\cap B$

  - Let C = {2nd coin is gold}. $$P(B\cap C|A_k) = \begin{cases} 0, & \text{if } k=1, \\ 0, & \text{if } k=2, \\ 1, & \text{if } k=3. \end{cases}$$
  - by ⑤ $$P(B\cap C) = \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 1 = \frac{1}{3}$$ $B \neq A_2\cup A_3$ (check Ω) ← wrong calculation

    $$P(C|B) = \frac{P(B\cap C)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}$$ cf. $P(A_3|A_2\cup A_3) = \frac{1/3}{2/3} = \frac{1}{2}$

➢ Example (TV Game Show: Let's Make A Deal)

- The story.
  1. The contestant is given an opportunity to select one of three doors.
  2. Behind one of the doors is a great prize (say, a car) and there is nothing behind the other two doors.
  3. The host knows which door contains the car, but the contestant does not.

4. After the contestant select a door, the host opens an *empty door* that the contest *did not* pick.

key

5. After opening an empty door, the host always offers the contestant the opportunity to switch to the other remaining *unopened* door.

- **Q**: Should the contestant switch to the other door or not?

see intuitive interpretation (next slide)

- Argument 1 (*The Drunkar's Walk* by L. Mlodinow): "Two doors are available --- open one and you win; open the other and you lose …, your chances of wining are 50/50." cf. Example in LNp.4-6

- Argument 2. Without loss of generality, assume that the contestant select door 3. Let

contestant's choice

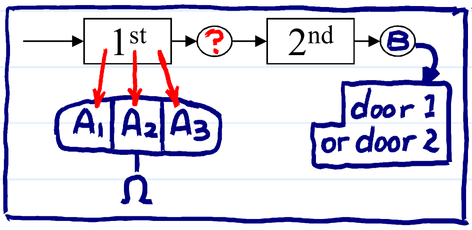
$P(A_3|B)$ $=?$

A_i = {the car is behind the door i}, i=1, 2, 3.

B = {door 1 is opened} → unopened doors: doors 2&3.

host's choice

- $P(A_1) = P(A_2) = P(A_3) = 1/3$
- $P(B|A_1) = 0, \quad P(B|A_2) = 1,$ $P(B|A_3) = 1/2$



1	2	3 (c)	switch
A	H		win - 1/3
H	A		win - 1/3
H		A	lose - 1/6
	H	A	lose - 1/6

(lose: 1/3)

P(switch & win) = 2/3

- $P(A_3|B) = \dfrac{(1/3)\times(1/2)}{(1/3)\times 0 + (1/3)\times 1 + (1/3)\times(1/2)} = 1/3$

not switch & win

- $P(A_2|B) = 2/3$

switch & win

- Similar result obtained if B={door 2 is opened} (exercise)

- Intuitive interpretation. WLOG, assume contestant select door 3. Ω = { (A-d1, H-d2), (A-d2, H-d1), (A-d3, H-d1), (A-d3, H-d2) }

Note. not symmetric outcomes #{switch to win}/#Ω = 1/2

➢ **Q**: Why are the 3 formulas useful in calculating probabilities? (Note: They all benefit from conditional probabilities.)

Ans: (i)繁→簡&簡&⋯&簡; $P(\cdot) \to P(\cdot|A) \Rightarrow \Omega \xrightarrow[\text{sample space}]{\text{reduced}} A$

(ii)簡=conditioning because the sample space is reduced from Ω to a smaller set. (e.g., in many previous examples, $P(B|A)$'s are known or easier to evaluate)

- Odds and Conditional Odds

$P(B) \Rightarrow B : \Omega$
$o(B) \Rightarrow B : B^c$

➢ The odds of an event B:

check LNp 3-20

$$o(B) \equiv \frac{P(B)}{P(B^c)} = \frac{P(B)}{1-P(B)} \iff P(B) = \frac{o(B)}{1+o(B)}$$

➢ The odd of event B given A:

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

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

and $\hookrightarrow P(\cdot)$

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

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

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

new information

similarly, if $P(B) > 0$, if and only if $P(A|B) = P(A)$.

Q: How to interpret the equality? new information