(by ⑤)

$P(A_1|B) \stackrel{?}{=} \dfrac{(1/3) \cdot 0}{1/2} = 0.$ ← 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 $C = \{2^{nd}$ 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)

- $P(B \cap C) \stackrel{?}{=} \dfrac{1}{3} \cdot 0 + \dfrac{1}{3} \cdot 0 + \dfrac{1}{3} \cdot 1 = \dfrac{1}{3}.$ $B \neq A_2 \cup A_3$ (check Ω) ← wrong calculation

9/17

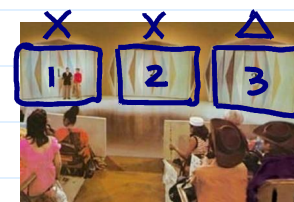$P(C|B) = \dfrac{P(B \cap C)}{P(B)} = \dfrac{1/3}{1/2} = \dfrac{2}{3}$ cf. ↔ $P(A_3|A_2 \cup A_3) = \dfrac{1/3}{2/3} = \dfrac{1}{2}$

➤ 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?

- 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.→ coin example in LNp 4-16

  see intuitive interpretation (next slide)

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

  ← contestant's choice

  $P(A_i|B) = ?$ $\begin{cases} A_i = \{\text{the car is behind the door } i\}, i=1, 2, 3. \\ B = \{\text{door 1 is opened}\} \end{cases}$

  ↳ host's choice → unopened doors: doors 2 & 3.

- $P(A_1) = P(A_2) = P(A_3) = 1/3$

  $P(B)$

  0 1/2 1
  $P(B|A_1)$ $P(B|A_3)$ $P(B|A_2)$

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

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

① ② ③ switch

A	H	win $-\frac{1}{3}$
H	A	win $-\frac{1}{3}$
H	A	lose $-\frac{1}{6}$
H	A	lose $-\frac{1}{6}$

$\} -\frac{1}{3}$
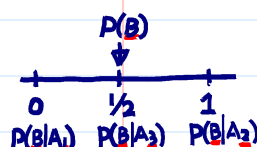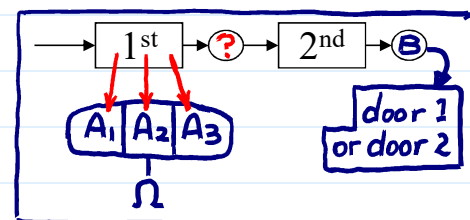
P(switch & win) = $\frac{2}{3}$

— not switch & win

$\square$ $P(A_2|B) = 2/3$

— switch & win

$\square$ Similar result obtained if $B=\{$door 2 is opened$\}$ (exercise)

- Intuitive interpretation. WLOG, assume contestant select door 3.

  Note. not symmetric outcomes $\Omega = \{$ (A-d1, H-d2), (A-d2, H-d1), (A-d3, H-d1), (A-d3, H-d2)$\}$

  #$\{$switch to win$\}$ / #$\Omega = \frac{1}{2}$

➤ **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

  $\boxed{\begin{array}{l} P(B) \Rightarrow B : \Omega \\ o(B) \Rightarrow B : B^c \end{array}}$

  ➤ The odds of an event B:

  check LN$_p$ 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 $P(\cdot|A)$

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

and cf. $\to P(\cdot)$

$O(\cdot) \xrightarrow[\text{after A occurs}]{\text{update}} O(\cdot|A)$ $o(B|A) = o(B) \times \dfrac{P(A|B)}{P(A|B^c)}$ cf. → Bayes' rule $P(\cdot) \xrightarrow{\text{update}} P(\cdot|A)$

❖ **Reading**: textbook, Sec 3.1, 3.2, 3.3, 3.5

Recall with replacement example (LN$_p$. 4-7) ← 獨立 —— **Independence**

$P(A)$

$P(A|B)$ $P(A|B^c)$

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

  for calculation purpose →

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

  cf. → ① in LN$_p$. 4-4

  Otherwise, they are said to be *dependent*.

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

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

for interpretation purpose →

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

└ new information

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

Q: How to interpret the equality? new information

p. 4-20

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

➢ Example (Cards): If a card is selected from a standard deck, let

can be changed to any other face (牌面)

$A = \{ace\}$ and $B = \{spade\}$. Then,

can be changed to any other suit (花②)

$P(A) = \dfrac{4}{52} = \dfrac{1}{13}$, $P(B) = \dfrac{13}{52} = \dfrac{1}{4}$,

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., $P(A^c \cap B) < P(A^c)P(B)$)

$P(A \cap B) = \dfrac{1}{52} = P(A)P(B)$ *Recall ② in LNp.4-4*

■ Face and Suit are independent

$P(B|A) = P(B)$ ➢ Theorem (Independence and Complements, 2-events case).

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

proof. $B = B \cap \Omega = B \cap (A \cup A^c) = \underline{(B \cap A)} \cup \underline{(B \cap A^c)}$

mutually exclusive

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

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

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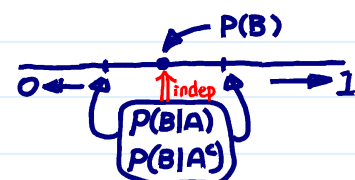
---

p. 4-21

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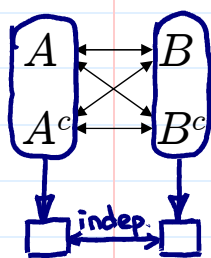
$\begin{bmatrix} P(B) = P(B|A) = P(B|A^c), \\ P(B^c) = P(B^c|A) = P(B^c|A^c), \end{bmatrix}$

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

P(B)

$0 \longleftarrow \uparrow_{indep} \longrightarrow 1$　P(B|A) P(B|A^c)

Q: What do these equalities say?

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

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

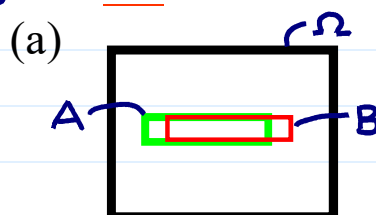➢ Example.　*P(A) & P(A^c): weights in ② (LNp.4-4) →*

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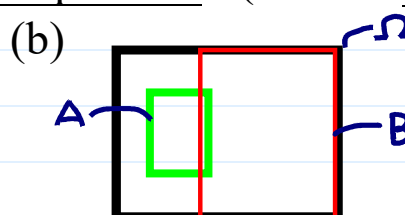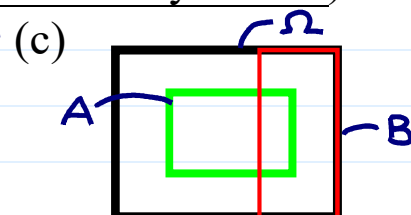
Ans. (c)

(a)　　　　　(b)　　　　　(c)

conditioning on A　　　conditioning on A　　　conditioning on A

P(B|A) > P(B)　　　P(B|A) < P(B)　　　P(B|A) ≈ P(B)

- **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)]

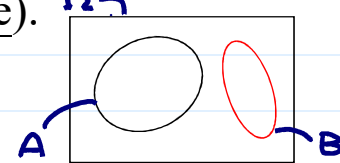- **Q**: What do we prefer? independent? or dependent? [It depends]

➤ 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

[What if $B = \phi$?]

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

[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 $1 \leq i < j \leq n$; [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 $k = 2, \ldots, n$,

[equality holds for any K sets in $\{A_1, \ldots, A_n\}$]

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

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

$$\cdots,$$

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

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

➤ Note: [$A_{t_1}, \ldots, A_{t_k}$ are mutually indep.]

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

[for interpretation purpose →]

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

[Thm in LNp.4-27] [cf.]

- Mutual independence implies pairwise independence; but, the converse statement is usually not true. [— an example in LNp.4-24]

[9/19]

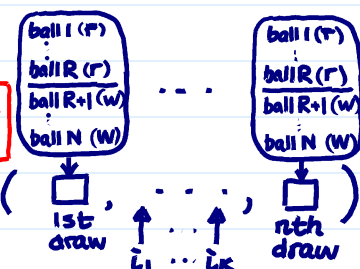- "n events are independent" means "mutually independent"

➤ 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 kth draw}, then [symmetric outcomes]

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

$$\#\Omega = N^n$$

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

[# of lists when balls are labelled.]

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

$$\Rightarrow A_1, \ldots, A_n \text{ are mutually independent}$$