

$$P(B|A) \xrightarrow{\text{update}} P(B) = \frac{P(A)P(B|A) + P(A^c)P(B|A^c)}{P(A)P(B|A) + P(A^c)P(B|A^c)}$$

$$= \frac{R}{N} \cdot \frac{R-1}{N-1} + \frac{N-R}{N} \cdot \frac{R}{N-1}$$

$$= \frac{R^2 - R + NR - R^2}{N(N-1)} = \frac{R(N-1)}{N(N-1)} = \frac{R}{N} = P(A)$$

(1st draw, 2nd draw)   
 by ②   
 check ② in LNp.4-4 p. 4-7   
 $P(B|A) < P(B) < P(B|A^c)$    
 Fair   
 $P(A)$

by ③   

$$P(B|A) = P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{(R(R-1))/(N(N-1))}{R/N} = \frac{R-1}{N-1}$$

Notes:

also, same for the 3rd, 4th, ... draw (exercise)

The probabilities are proportional to # of red balls left

$P(A|B) = P(B|A) \Rightarrow$  Symmetry.

Sampling With Replacement:

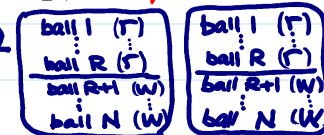
put the ball back into the urn after draw

$$P(A) = \frac{R \cdot N}{N \cdot N} = \frac{R}{N}, \quad P(B) = \frac{N \cdot R}{N \cdot N} = \frac{R}{N}$$

cf. without replacement symmetric outcomes

$$P(A \cap B) = \frac{R \cdot R}{N \cdot N} = \frac{R^2}{N^2}$$

independent (LNp.4-19)



update   

$$P(B|A) = \frac{R^2/N^2}{R/N} = \frac{R}{N} = P(B)$$

cf.  $\Omega$  in LNp.4-6   
 #  $\Omega = N^2$

## Example (Diagnostic Tests)

p. 4-8

The Story. A diagnostic test for a rare disease (e.g., an Xray for lung cancer) is part of a routine physical exam.

Let  $\Omega = \{\text{the whole population of Taiwan}\}$

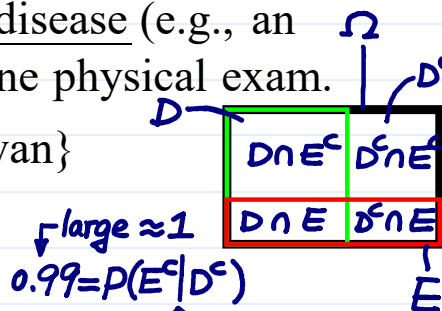
assume symmetric outcomes

usually  $D \neq E$

$D = \{\text{disease is present}\}$

$E = \{\text{test indicates disease present}\}$

i.e.,  $P(D)$  very small.



large  $\approx 1$    
  $0.99 = P(E^c|D^c)$

$P(D^c) = 0.999$

Suppose that  $P(D) = 0.001$ ,  $P(E|D) = 0.98$ ,  $P(E|D^c) = 0.01$

base rate   
 Q: Do you think the test is effective?

large  $\approx 1$

small  $\approx 0$

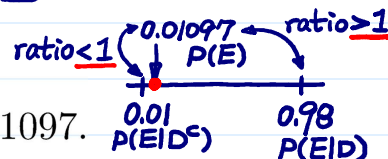
Let us examine it from an alternative viewpoint. Suppose that you are tested as positive. What is the probability that you actually have the disease?

$P(D|E)$

$$P(E) = \frac{P(D)P(E|D) + P(D^c)P(E|D^c)}{P(D)P(E|D) + P(D^c)P(E|D^c)}$$

by ②   

$$= \frac{0.001 \times 0.98 + 0.999 \times 0.01}{0.001 \times 0.98 + 0.999 \times 0.01} = 0.01097$$



update   

$$P(D|E) = \frac{P(D)P(E|D)}{P(D)P(E|D) + P(D^c)P(E|D^c)} = \frac{0.00098}{0.01097} = 0.0893$$

by ③   
 $P(D^c|E) = 1 - P(D|E) = 0.9107$    
 膨脹倍數  $\propto P(E|D)$    
 Q:  $P(E|D)$  large  $\Rightarrow P(D|E)$  large?

E occurs

D occurs

膨脹倍數 ratio (check in LNp.4-4)   
  $\frac{P(D|E)}{P(D)} \approx 90$

**Q:** Now, do you still think the test is effective?

- The probability of  $D$  increased by a factor of roughly 90 ( $0.001 \rightarrow 0.0893$ ) when  $E$  occurs, but 0.0893 is still **small**.
- The  $P(E|D^c)$  ( $=0.01$ ) and  $P(E^c|D)$  ( $=1-P(E|D)=0.02$ ) are called the **false positive** and **false negative** rates, respectively.

**Q:** Why the 2 interpretations so different? What causes it?

**Diagram:** Two circles representing the sample space  $\Omega$ . The left circle is labeled "random" and "probability  $\propto$  area". It shows a small red region  $D$  and a larger blue region  $E$ . The right circle is labeled "deterministic" and "D = E". It shows a small red region  $D$  and a larger blue region  $E$ . Arrows indicate the relationship between the two circles.

**Handwritten notes:**

- 車禍 & 酒駕 (LNp.1-7)**
- rare event**
- rare disease  $\Rightarrow$  base rate  $P(D)$  is very small**
- 若D則 almost E**
- 若非D則 almost 非E**
- 若E則 almost D?**
- random**
- probability  $\propto$  area**
- deterministic**
- 若D則 (100%) E  $\Rightarrow D \subseteq E$**
- 若非D則 (100%) 非E  $\Rightarrow D^c \subseteq E^c \Rightarrow D \supseteq E$**
- $D = E$**
- $1 = P(E|D) = P(D|E)$**
- $0 = P(E|D^c) = P(D^c|E)$**
- 若E則D**

**Example (引自“快思慢想”, Kahneman)**

$\Omega = \{\text{all male graduate students}\}$

$S_1, \dots, S_9$ : partition of  $\Omega$

$P(S_i) = ??$   $i=1, \dots, 9$  (base rate)

在1970年代，湯姆是美國某州裡重要大學的研究生，請將下面九個研究所領域排序，標出湯姆就讀這些領域的可能性，1代表可能性最高，9代表最低。

A. 企業管理, B. 電腦, C. 工程, D. 人文與教育, E. 法律, F. 醫學, G. 圖書館學, H. 物理和生命科學, I. 社會學和社會工作

- 下面是湯姆念高三時，心理學家根據心理測驗結果對湯姆的人格素描：

**B**

**eg. 敘述謬誤 in LNp.1-7**

**特質/成功 large**

**特質/失敗 large**

**成功/特質 large**

**other examples in our life?**

把剛剛那幾個領域再排序一下，你認為湯姆最可能是哪個領域的研究生，1代表可能性最高，9代表最低。

在1970年代，受試者的排序大致如下：

- 電腦
- 工程
- 企業管理
- 物理和生命科學
- 圖書館學
- 法律
- 醫學
- 人文與教育
- 社會學和社會工作

**Answer based on  $P(B|S_i)$   $i=1, \dots, 9$ , but, not consider base rate  $P(S_i)$**

**cp. rare event in 1970**

**rare event**

• Extensions of the 3 Useful Formulas (for  $m$ -events case)

1. (Multiplication Law) If  $A_1, \dots, A_m$  are events for which

$$P(A_1 \cap \dots \cap A_{m-1}) > 0, \text{ then}$$

to guarantee the conditional probability well defined

2-events  
(LNp.4-4)

④ —  $P(A_1 \cap \dots \cap A_m)$

meaning

$$= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_m|A_1 \cap \dots \cap A_{m-1}).$$

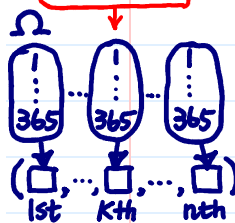
proposition in  
LNp.312 for  
 $P(A_1 \cup \dots \cup A_m)$

proof.

$$\cancel{P(A_1)} \times \frac{\cancel{P(A_1 A_2)}}{\cancel{P(A_1)}} \times \frac{\cancel{P(A_1 A_2 A_3)}}{\cancel{P(A_1 A_2)}} \times \dots \times \frac{P(A_1 \cap \dots \cap A_{m-1} \cap A_m)}{\cancel{P(A_1 \cap \dots \cap A_{m-1})}} = P(A_1 \cap \dots \cap A_m).$$

Note: order not matter

symmetric outcomes



Example (Birthday Problem, LNp.3-4). Let  $A_k$  be the event that the  $k^{\text{th}}$  birthday differs from the first  $k-1$ . Then,

$$P(A_1) = 1,$$

might have same birthday

easier to calculate

$$P(A_k|A_1 \cap \dots \cap A_{k-1}) = \frac{365 - k + 1}{365}$$

Q: What benefit does conditional probability bring?

different birthday

$$P(A_1 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap \dots \cap A_{n-1})$$

by ④

$$= \prod_{k=1}^n \frac{365 - k + 1}{365} = \frac{365!}{365^n \cdot (365 - n)!} = \frac{(365)_n}{365^n}.$$

Example (Matching Problem, LNp.3-14). Q: Probability that exactly  $k$  of  $n$  members ( $k \leq n$ ) have matches = ??

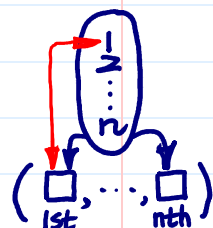
event A

# $\Omega = n!$

Let  $\Omega$  be all permutations  $\omega = (i_1, \dots, i_n)$  of  $1, 2, \dots, n$ .

symmetric outcomes

Let  $A_j = \{\omega: i_j = j\}$  and



$$A = \bigcup_{1 \leq j_1 < \dots < j_k \leq n} (A_1^c \cap \dots \cap A_{j_1-1}^c \cap A_{j_1} \cap A_{j_1+1}^c \cap \dots \cap A_n^c)$$

j<sub>1</sub>th

j<sub>1</sub>th ... j<sub>k</sub>th

By symmetry

mutually exclusive for different  $(j_1, \dots, j_k)$

$$P(A) = \binom{n}{k} \times P(\underbrace{A_1 \cap \dots \cap A_k}_{\subseteq E} \cap \underbrace{A_{k+1}^c \cap \dots \cap A_n^c}_{\subseteq G})$$

Let  $E = A_1 \cap \dots \cap A_k$  and  $G = A_{k+1}^c \cap \dots \cap A_n^c$

Then,  $P(E \cap G) = P(E)P(G|E)$ , where

by ④

$$P(E) = P(A_1)P(A_2|A_1) \cdots P(A_k|A_1 \cap \dots \cap A_{k-1})$$

easier to evaluate

$$\frac{(n-1)!}{n!} = \frac{1}{n} \times \frac{1}{n-1} \times \dots \times \frac{1}{n-k+1} = \frac{(n-k)!}{n!} = \frac{1}{(n)_k}$$

none of the remaining  $n-k$  people get his/her own gifts

$$\text{and, } P(G|E) = \sum_{i=0}^{n-k} (-1)^i \frac{1}{i!} \equiv p_{n-k}$$

check LNp.3-15

Note. In ③,  $A$  &  $A^c$  is a partition

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

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

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

⑤ —  $P(B) = \sum_{i=1}^m P(A_i)P(B|A_i)$

weighted average ( $\because \sum P(A_i)=1$ )

meaning

example

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

e.g.,  $\Omega = \{\omega_1, \dots, \omega_m\}$ , let  $A_i = \{\omega_i\}$

base rate

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

$P(\cdot)$  update after  $B$  occurs

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

check graphs in LNp.4-9

larger  $P(B|A_i)$   
 $\Rightarrow$  larger ratio  $\frac{P(A_i|B)}{P(A_i)}$

But, larger  $P(A_i|B)$ ?

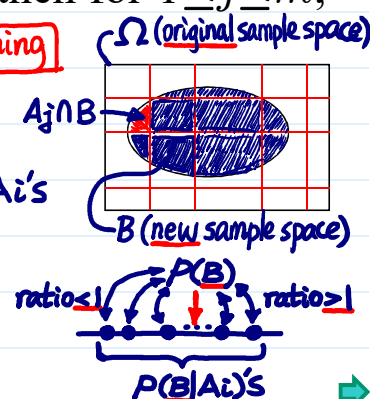
proof. By definition,

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

$\xrightarrow{\text{by ⑤}} \frac{\text{分子}}{\text{分母}}$

$\frac{P(A_j|B)}{P(A_j)} = \frac{P(B|A_j)}{P(B)}$

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



Examples From Bayesians' viewpoint,

data in statistics

$P(A_j)$  = probability of  $A_j$  before  $B$  occurs  $\rightarrow$  prior prob.

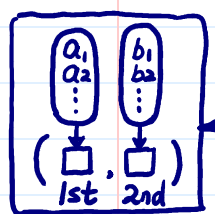
$P(A_j|B)$  = probability of  $A_j$  after  $B$  occurs  $\rightarrow$  posterior prob.

$\Rightarrow$  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

$\frac{P(ANC|B)}{P(C|B)} = P(A|BNC)$

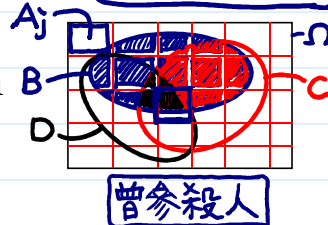


□ The probabilities of the 2<sup>nd</sup>-stage results depend on what happened in the 1<sup>st</sup> stage

□ We never see the result of the 1<sup>st</sup> 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 1<sup>st</sup> stage given the final result



曾參殺人

### ➤ Example (Gold Coins):

# $\Omega = 6$

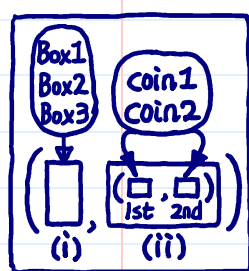
1<sup>st</sup>

?

2<sup>nd</sup>

B

#### ■ The Story.



Box 1 contains 2 silver coins.

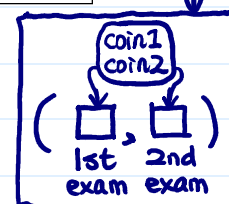
Box 2 contains 1 gold and 1 silver coin.

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)

symmetric outcomes

a partition of  $\Omega$



Q: Given that 1<sup>st</sup> coin is gold, what is the probability that Box  $k$  is selected,  $k=1, 2, 3$ ?

Let  $A_k = \{\text{Box } k \text{ is selected}\}$ ,  $B = \{\text{1<sup>st</sup> coin is gold}\}$ ,

$$P(B|A_k) = \begin{cases} 0, & \text{if } k = 1, \\ 1/2, & \text{if } k = 2, \\ 1, & \text{if } k = 3. \end{cases}$$

subsets of  $\Omega$

easier to evaluate

by ⑤

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

base rate

Note.

$$\text{Sum} = 0 + \frac{1}{2} + 1 \neq 1$$

$$P(A_1|B) = \frac{(1/3) \cdot 0}{1/2} = 0.$$

c.f.  $P(A_1) = P(A_2) = P(A_3) = 1/3$

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

Q: Given that 1<sup>st</sup> coin is gold, what is the probability that 2<sup>nd</sup> coin is gold?

Let  $C = \{\text{2<sup>nd</sup> coin is gold}\}$ .  $P(B \cap C|A_k) = \begin{cases} 0, & \text{if } k = 1, \\ 0, & \text{if } k = 2, \\ 1, & \text{if } k = 3. \end{cases}$

$$P(B \cap C) = \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1 = \frac{1}{3}.$$

$B \neq A_2 \cup A_3$  (check  $\Omega$ )

wrong intuition

9/22

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

$$P(A_3|A_2 \cup A_3) = \frac{1/3}{2/3} = \frac{1}{2}$$

### ➤ 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.

