p. 4-7

$\overleftarrow{P(B|A), P(B|A^c)} \xleftarrow{\text{update}}$

$$P(B) \overset{\text{by } ②}{=} P(A)P(B|A) + P(A^c)P(B|A^c)$$

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

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

($\boxed{r/w}$ 1st draw , $\boxed{r}$ 2nd draw)

check ② in LNp.4-4: $\boxed{P(B|A) < P(B) < P(B|A^c)}$

$$P(B|A) = P(A|B) \overset{\text{by } ③}{=} \frac{P(A\cap B)}{P(B)} = \frac{(R(R-1))/(N(N-1))}{R/N} = \frac{R-1}{N-1}.$$

also, same for the 3rd, 4th, ⋯ draw

- Notes:
  - The probabilities are proportional to # of red balls left
  - $P(A|B) = P(B|A) \Rightarrow$ Symmetry.

9/12

$P(A\cap B)/P(B)$ " $= P(A\cap B)/P(A)$

- Sampling *With Replacement*: put the ball back into the urn after draw

cf. without replacement

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

symmetric outcomes

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

independent (LNp.4-19)

Ω: ball 1 (r), ball R (r), ball R+1 (w), ball N (w) ; ball 1 (r), ball R (r), ball R+1 (w), ball N (w)

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

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

Ω in LNp.4-6 $\xrightarrow{\text{cf.}}$ (□ 1st draw , □ 2nd draw) $\#\Omega = N^2$

p. 4-8

➢ Example (Diagnostic Tests)

- The Story. A diagnostic test for a *rare* disease (e.g., an Xray for lung cancer) is part of a routine physical exam. i.e., P(D) very small

- Let Ω = {the whole population of Taiwan}
  assume symmetric outcomes

  usually $D \neq E$

  D = {disease is present}

  E = {test indicates disease present}

$D \cap E^c$	$D^c \cap E^c$
$D \cap E$	$D^c \cap E$

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

  base rate → $P(D^c) = 0.999$; large ≈ 1 ; small ≈ 0

  Q: Do you think the test is effective?

- 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?

  E occurs ; D occurs ; P(D|E)

$$P(E) \overset{\text{by } ②}{=} P(D)P(E|D) + P(D^c)P(E|D^c)$$

$$= 0.001\times 0.98 + 0.999\times 0.01 = 0.01097.$$

ratio<1, 0.01097 P(E), ratio>1 ; 0.01 $P(E|D^c)$; 0.98 $P(E|D)$

$$P(D|E) \overset{\text{by } ③}{=} \frac{P(D)P(E|D)}{P(D)P(E|D) + P(D^c)P(E|D^c)} = \frac{0.00098}{0.01097} = 0.0893.$$

update ; cf. P(D)

膨脹倍數 ratio = $\frac{P(D|E)}{P(D)} \approx 90$

膨脹倍數 $\propto$ P(E|D) $\xleftrightarrow{\text{cf.}}$ P(D|E)

Q: P(E|D) large $\not\Rightarrow$ P(D|E) large?

$$P(D^c|E) = 1 - P(D|E) = 0.9107$$

車禍&酒駕 (LNp 1-8)

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

- The probability of D increased by a factor of roughly 90 (0.001 → 0.0893) when E occurs, but 0.0893 is still **small**

  new information $P(D|E)$

∵ rare disease ⇒ base rate $P(D)$ is very small

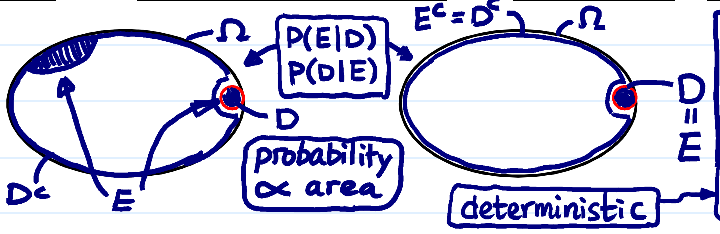
- The $P(E|D^c)$ (=0.01) and $P(E^c|D)$ (=1−$P(E|D)$=0.02) are called the *false positive* and *false negative rates*, respectively.

  假陽性 假陰性

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

•若D則 almost E
•若非D則almost 非E

~~若E則almost D?~~



• 若D則E ⇒ $D\subseteq E$
• 若非D則非E ⇒ $D^c\subseteq E^c$ ⇒ $D\supseteq E$
$D \overset{\text{iff}}{=} E$
$D^c \overset{\text{iff}}{=} E^c$

$1=P(E|D)=P(D|E)$
$0=P(E|D^c)=P(D^c|E)$

若E則D

Ω={all male graduate students}

➢ Example (引自"快思慢想", Kahneman)

- 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。

  S_1 A. 企業管理, S_2 B. 電腦, S_3 C. 工程, S_4 D. 人文與教育, S_5 E. 法律, S_6 F. 醫學, S_7 G. 圖書館學, S_8 H. 物理和生命科學, S_9 I. 社會學和社會工作

$S_1,\dots,S_9$: partition of Ω

$P(S_i)$=?? $i=1,\dots,9$ ← base rate

- 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
- ▫ 湯姆是個很聰明的學生，但缺少真正的創造力。
- ▫ 他喜歡整潔和秩序，他的每一樣東西，不管多少，都有條有理的擺放在恰當位置上。
- ▫ 他的作文有點無趣呆板和機械式，偶爾會出現一些陳舊過時的雙關語和類似科幻想像的句子。
- ▫ 他的好勝心很強，對人冷淡，沒什麼同情心，也不喜歡跟別人來往。
- ▫ 雖然以自我為中心，卻有很強的道德感。

e.g. 敘述謬誤 in LNp.1-7
特質/成功 large
特質^c/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。

$P(S_i|B)$=?? $i=1,\dots,9$

$P(S_i) \xrightarrow{\text{update}} P(S_i|B)$

- 在1970年代，受試者的排序大致如下：

  1. 電腦, 2. 工程, 3. 企業管理, 4. 物理和生命科學, 5. 圖書館學, 6. 法律, 7. 醫學, 8. 人文與教育, 9. 社會學和社會工作

cf

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

- Extensions of the 3 Useful Formulas (for m-events case) cf. 2-events (LNp.4-4)

1. (Multiplication Law) If $A_1, \ldots, A_m$ are events for which $P(A_1 \cap \cdots \cap A_{m-1}) > 0$, then

   to garantee the conditional probability well defined

   ④ — $P(A_1 \cap \cdots \cap A_m)$ meaning

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

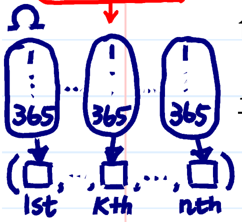
   cf. proposition in LNp.3-12 for $P(A_1 \cup \cdots \cup A_m)$

   proof. $\cancel{P(A_1)} \times \dfrac{\cancel{P(A_1 \cap A_2)}}{\cancel{P(A_1)}} \times \dfrac{\cancel{P(A_1 \cap A_2 \cap A_3)}}{\cancel{P(A_1 \cap A_2)}} \times \cdots \times \dfrac{P(A_1 \cap \cdots \cap A_{m-1} \cap A_m)}{\cancel{P(A_1 \cap \cdots \cap A_{m-1})}}$

   $= P(A_1 \cap \cdots \cap A_m)$. Note order not matter

- Example (Birthday Problem, LNp.3-4). Let A_k be the event that the k^{th} birthday differs from the first $k-1$. Then, $P(A_1)=1$,

symmetric outcomes



might have same birthday

easier to calculate

Q: What benefit does conditional probability bring ?

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

different birthday

- $P(A_1 \cap \cdots \cap A_n) \overset{\text{by ④}}{=} P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1})$

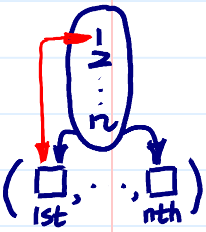
$$= \prod_{k=1}^{n} \frac{365-k+1}{365} = \frac{365!}{365^n \cdot (365-n)!} = \frac{(365)_n}{365^n}.$$

- Example (Matching Problem, LNp.3-14). **Q**: Probability that exactly k of n members ($k \le n$) have matches = ?? ← $k=0$

event A

- Let Ω be all permutations $\omega = (i_1, \ldots, i_n)$ of $1, 2, \ldots, n$. #Ω=n!

  symmetric outcomes



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

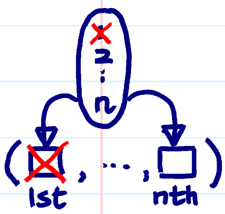
  j_1th j_2th ⋯ j_kth

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

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

- By symmetry

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



- Let $E = A_1 \cap \cdots \cap A_k$ and $G = A_{k+1}^c \cap \cdots \cap A_n^c$
- Then, $P(E \cap G) \overset{\text{by ①}}{=} P(E)P(G|E)$, where

by④ $P(E) = P(A_1)P(A_2|A_1) \cdots P(A_k|A_1 \cap \cdots \cap A_{k-1})$ easier to evaluate

$\dfrac{(n-1)!}{n!}$ $\quad \dfrac{(n-2)!}{(n-1)!}$

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

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

and, $P(G|E) = \displaystyle\sum_{i=0}^{n-k} (-1)^i \frac{1}{i!} \equiv p_{n-k}$ check LNp.3-15

Note. In ② ③, 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 ($\because \Sigma P(A_i)=1$)

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

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

like base rate

A_i, Ω, $B \cap A_i$, B

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 $1 \leq j \leq m$,

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

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)

$A_j \cap B$

B (new 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,

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

ratio≤1, ratio≥1

$P(B|A_i)$'s

Examples

- From Bayesians' viewpoint,

data in statistics

事前机率

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

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

事後机率

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

⇒ 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_j, B, D, Ω, C

曾参殺人

$(\omega_1, \omega_2, \vdots)$ $(o_1, o_2, \vdots)$ $(\square, \square)$ 1st 2nd

asymmetric outcomes

$P(A_j)$

→ 1st → ? → 2nd → B

Ω — $A_1 | A_2 | \cdots | A_m$ ⇒ $A_1, \cdots, A_m$: a partition of Ω.

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

$P(B|A_j)$

easier to evaluate

key

- We never see the result of the 1st stage, only the final result → B occurs

i.e., it's a random outcome to you

$P(A_j|B)$

- We may be interested in finding the probability for outcomes in the 1st stage given the final result

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

1st → ? → 2nd → B

- The Story. Ω

Box 1 | Box 2 | Box 3 — a partition of Ω

(coin 1, coin 2) — 1st exam, 2nd exam

Box1, Box2, Box3; coin1, coin2; (i), (ii) — symmetric outcomes

 - 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? $P(A_3|B)$ — 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. $\leftrightarrow P(A_3|A_2\cup A_3) = \frac{1/3}{2/3} = \frac{1}{2}$

   9/17

➢ 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.