AI HW2 孫士龍 RE6121011 Q2&3 code ↓

https://colab.research.google.com/drive/1gprSzmVVCNAiR44uP1swC0mjDnMjSRvP?usp=sharing

1、

Probability	Hypothesis		
	$i=1$	$i=2$	$i=3$
$p(H_i)$	0.40	0.35	0.25
$p(E_1\|H_i)$	0.3	0.8	0.5
$p(E_2\|H_i)$	0.9	0.0	0.7
$p(E_3\|H_i)$	0.6	0.7	0.9

approach: iterative

order : E_3, E_1, E_2

solve : $P(h_i|d)$? $P(X|d)$?

t=1 : $P(h_i|E_3) = \frac{P(E_3|h_i)\cdot P(h_i)}{P(E_3)}$

$$P(E_3) = \sum_{i=1}^{3} P(E_3|h_i)P(h_i)$$
$$= P(E_3|h_1)P(h_1) + P(E_3|h_2)P(h_2) + P(E_3|h_3)P(h_3)$$
$$= 0.6\times 0.4 + 0.7\times 0.35 + 0.9\times 0.25$$
$$= 0.24 + 0.245 + 0.225$$
$$= 0.71$$

$P(h_1|E_3) = \frac{P(E_3|h_1)\cdot P(h_1)}{P(E_3)} = \frac{0.6\times 0.4}{0.71} \doteq 0.3380$ *

$P(h_2|E_3) = \frac{P(E_3|h_2)\cdot P(h_2)}{P(E_3)} = \frac{0.9\times 0.35}{0.71} \doteq 0.3951$ *

$P(h_3|E_3) = \frac{P(E_3|h_3)\cdot P(h_3)}{P(E_3)} = \frac{0.9\times 0.25}{0.71} \doteq 0.3169$ *

t=2 : $P(h_i|E_1,E_3) = \frac{P(E_1|h_i)P(h_i|E_3)}{P(E_1|E_3)}$

$$P(E_1|E_3) = \sum_{i=1}^{3} P(E_1|h_i)P(h_i|E_3)$$
$$= P(E_1|h_1)P(h_1|E_3) + P(E_1|h_2)P(h_2|E_3) + P(E_1|h_3)P(h_3|E_3)$$
$$= 0.3\times 0.3380 + 0.8\times 0.3951 + 0.5\times 0.3169$$
$$= 0.1014 + 0.29608 + 0.15845$$
$$= 0.53593$$ *

$P(h_1|E_1,E_3) = \frac{P(E_1|h_1)P(h_1|E_3)}{P(E_1|E_3)} = \frac{0.3\times 0.3380}{0.53593} \doteq 0.1892$ *

$P(h_2|E_1,E_3) = \frac{P(E_1|h_2)P(h_2|E_3)}{P(E_1|E_3)} = \frac{0.8\times 0.3951}{0.53593} \doteq 0.5151$ *

$P(h_3|E_1,E_3) = \frac{P(E_1|h_3)P(h_3|E_3)}{P(E_1|E_3)} = \frac{0.5\times 0.3169}{0.53593} \doteq 0.2957$ *

$t=3$: $P(h_i \mid E_2, E_1, E_3) = \dfrac{P(E_2 \mid h_i)\, P(h_i \mid E_1, E_3)}{P(E_2 \mid E_1, E_3)}$

$$P(E_2 \mid E_1, E_3) = \sum_{i=1}^{3} P(E_2 \mid h_i)\, P(h_i \mid E_1, E_3)$$

$$= P(E_2 \mid h_1)\, P(h_1 \mid E_1, E_3) + P(E_2 \mid h_2)\, P(h_2 \mid E_1, E_3) + P(E_2 \mid h_3)\, P(h_3 \mid E_1, E_3)$$

$$= 0.9 \times 0.1892 + 0 \times 0.5151 + 0.9 \times 0.2959$$

$$= 0.17028 + 0 + 0.26631$$

$$= 0.43659 \ \#$$

$$P(h_1 \mid E_2, E_1, E_3) = \frac{P(E_2 \mid h_1)\, P(h_1 \mid E_1, E_3)}{P(E_2 \mid E_1, E_3)} = \frac{0.9 \times 0.1892}{0.43659} \doteq 0.3900 \ \#$$

$$P(h_2 \mid E_2, E_1, E_3) = \frac{P(E_2 \mid h_2)\, P(h_2 \mid E_1, E_3)}{P(E_2 \mid E_1, E_3)} = \frac{0 \times 0.5151}{0.43659} \doteq 0 \ \#$$

$$P(h_3 \mid E_2, E_1, E_3) = \frac{P(E_2 \mid h_3)\, P(h_3 \mid E_1, E_3)}{P(E_2 \mid E_1, E_3)} = \frac{0.9 \times 0.2959}{0.43659} \doteq 0.6100 \ \#$$

2.

Prior probabilities:	Likelihood
P(swim) = 0.3	P(hot | swim) = 0.8
P(golf) = 0.25	P(sunny | swim) = 0.7
P(hiking) = 0.45	P(dry | swim) = 0.2
	P(hot | golf) = 0.4
	P(sunny | golf) = 0.9
	P(dry | golf) = 0.3
	P(hot | hiking) = 0.4
	P(unny | hiking) = 0.8
	P(dry | hiking) = 0.7

(1) order: HSD → observed value q = HSD

$P(\text{swim} \mid q) \doteq 0.2082$

$P(\text{golf} \mid q) \doteq 0.1673$

$P(\text{hiking} \mid q) \doteq 0.6245$ → The next suggested activity is hiking. #

$P(\text{hot} \mid q) \doteq 0.4833$

$P(\text{sunny} \mid q) \doteq 0.7959$

$P(\text{dry} \mid q) \doteq 0.5290$ → The next evidence will be sunny. #

(2) order: SSSSS → observed value q = SSSSS

$P(\text{swim} \mid q) \doteq 0.1459$

$P(\text{golf} \mid q) \doteq 0.4293$

P(hiking|q) ≒ 0.4268 ⇒ The next suggested activity is golf. ※

P(hot|q) ≒ 0.4584

P(sunny|q) ≒ 0.8281

P(dry|q) ≒ 0.4561 ⇒ The next evidence will be sunny. ※

(3) order: HHHHH ⇒ observed value q = HHHHH

P(swim | q) ≒ 0.9320

P(golf | q) ≒ 0.0293

P(hiking|q) ≒ 0.0439 ⇒ The next suggested activity is swimming. ※

P(hot|q) ≒ 0.9928

P(sunny|q) ≒ 0.7092

P(dry|q) ≒ 0.2243 ⇒ The next evidence will be hot ※

(4) order: DDDDD ⇒ observed value q = DDDDD

P(swim | q) ≒ 0.0012

P(golf | q) ≒ 0.0080

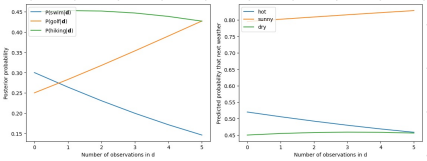
P(hiking|q) ≒ 0.9908 ⇒ The next suggested activity is hiking. ※

P(hot|q) ≒ 0.4005

P(sunny|q) ≒ 0.8007

P(dry|q) ≒ 0.6962 ⇒ The next evidence will be sunny. ※

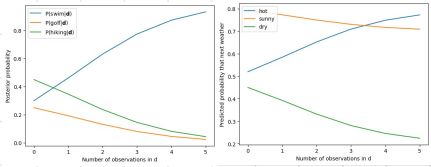
(5) order: SSSSS



若 order 是 SSSSS，從 likelihood 可看到 sunny 在給定 golf 的情況下最高，swim下最低，所以 golf 的後驗機率會上升，swim會下降，如圖所示。

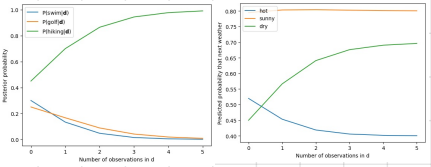
又若 golf 的後驗機率上升，且因給定 golf 情況下 sunny 的 likelihood 最高，所以下一次預測是 sunny 的機率也會上升。

order: HHHHH



若order是HHHHH，從likelihood可看到hot在給定swim的情況下最高，
所以swim的後驗机率會上升，而其他則會降低，如圖所示。
又若swim的後驗机率上升，且因給定swim情況下hot的likelihood最高，
所以下一个預測是hot的机率也會上升。

order: DDDDD



若order是DDDDD，從likelihood可看到dry在給定hiking的情況下最高，
所以hiking的後驗机率會上升，而其他則會降低，如圖所示。
又若hiking的後驗机率上升，且因給定hiking情況下sunny和dry的likelihood皆很高，
所以下一个預測是sunny和dry的机率可能會上升。
但因sunny在給定是golf情況下的likelihood最高，
所以golf的後驗机率之下降導致sunny可能跟著下降，
因此sunny在這些權衡下持平，只有dry上升。

3、 (1) order: HSD ⇒ observed value d = HSD

P(swim | d) ≒ 0.2523

P(golf | d) ≒ 0.2432

P(hiking|d) ≒ 0.5045 ⇒ The next suggested activity is hiking. ✌

P(hot|d) ≒ 0.5009

P(sunny|d) ≒ 0.7990

P(dry|d) ≒ 0.4766 ⇒ The next evidence will be sunny. ✌

(2) order: SSSSS → observed value d = SSSSS

P(swim | d) ≒ 0.1547

P(golf | d) ≒ 0.5436

P(hiking | d) ≒ 0.3017 → The next suggested activity is golf. ✌

P(hot | d) ≒ 0.4619

P(sunny | d) ≒ 0.8389

P(dry | d) ≒ 0.4052 → The next evidence will be sunny. ✌

(3) order: HHHHH → observed value d = HHHHH

P(swim | d) ≒ 0.9912

P(golf | d) ≒ 0.0294

P(hiking | d) ≒ 0.0294 → The next suggested activity is swimming. ✌

P(hot | d) ≒ 0.7965

P(sunny | d) ≒ 0.7088

P(dry | d) ≒ 0.2196 → The next evidence will be hot ✌

(4) order: DDDDD → observed value d = DDDDD

P(swim | d) ≒ 0.0019

P(golf | d) ≒ 0.0142

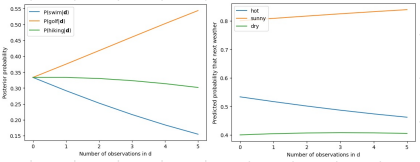
P(hiking | d) ≒ 0.9839 → The next suggested activity is hiking. ✌

P(hot | d) ≒ 0.4007

P(sunny | d) ≒ 0.8012

P(dry | d) ≒ 0.6934 → The next evidence will be sunny. ✌

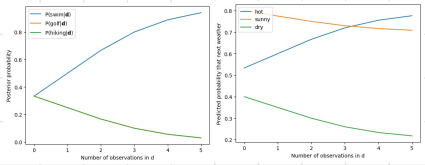
(5) order: SSSSS



從 likelihood 可看到 sunny 在給定 golf 的情況下最高，swim 下最低，所以 golf 的後驗機率會上升，swim 會下降，如圖所示。

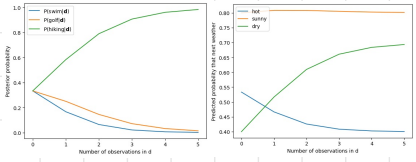
又若 golf 的後驗機率上升，且因給定 golf 情況下 sunny 的 likelihood 最高，所以下一步預測則是 sunny 的機率也會上升。

order: HHHHH



若 order 是 HHHHH，從 likelihood 可看到 hot 在給定 swim 的情況下最高，其他兩個一樣低
所以 swim 的後驗機率會上升，而其他則會一起降低，圖中 golf 和 hiking 的曲線重疊。
又若 swim 的後驗機率上升，且因給定 swim 情況下 hot 的 likelihood 最高，
所以下一次預測是 hot 的機率也會上升。

order: DDDDD



若 order 是 DDDDD，從 likelihood 可看到 dry 在給定 hiking 的情況下最高，
所以 hiking 的後驗機率會上升，而其他則會降低，如圖所示。
又若 hiking 的後驗機率上升，且因給定 hiking 情況下 sunny 和 dry 的 likelihood 皆很高，
所以下一次預測是 sunny 和 dry 的機率可能會上升。
但因 sunny 在給定是 golf 情況下的 likelihood 最高，
所以 golf 的後驗機率之下降 導致 sunny 可能跟著下降，
因此 sunny 在這些權衡下持平，只有 dry 上升。

Comparative Analysis

比較使用之种不同的 prior probabilities，觀察第 1 ~ 4 題，可以發現答案皆相同。從圖表來看，除了初始值不同，其他上升或下降的趨勢皆相同，
因此可以得出 2 個結論：

1. 當證據少時，prior probability 的差異可能會影響結果，
   當證據增多時，posterior probability 主要依賴於證據，
   prior probability 的影響則減少。
2. 即使起始的 prior probability 不同，但最終可依靠觀察到的證據
   做出可靠的結果，證明了 Bayesian inference 的穩健性。