Discussion Section 3

Case study: HIV & AIDS diagnosis

AMS 131, 6 Aug 19

1

$0 \quad\longrightarrow\quad C \quad\longrightarrow$ HIV antibody concentration

ELISA says HIV $\ominus$ (below C) — ELISA says HIV $\oplus$ (above C)

Decision: you have to act (based on science)

Science: search for knowledge without action

A = (really is HIV $\oplus$)

not A = (really is HIV $\ominus$)

$\oplus$ = (ELISA says HIV $\oplus$)

$\ominus$ = (ELISA says HIV $\ominus$)

prevalence $P(A) = 0.004$

sensitivity $P(\underset{\text{data}}{\oplus} \mid \underset{\text{truth}}{A}) = 0.96$

specifizity $P(\ominus \mid \text{not } A) = 0.97$

when blood arrives, compare ELISA says $\oplus$

$$P(A \mid \oplus) = ?$$

truth — data

Q: how does $P(A|B)$ relate to $P(B|A)$?

almost always not the same:

$P(\text{clouds overhead} \mid \text{raining}) = \text{high}$

$P(\text{raining} \mid \text{clouds overhead}) = \text{low}$

first A: Bayes (1760)

village: some people dying (effect; data); why? (cause?)

$P(\text{truth} \mid \text{data})$

$P(\text{cause} \mid \text{effect})$ harder

$P(\text{effect} \mid \text{cause})$ easier

$P(\text{data} \mid \text{truth})$

possible causes

bad food

bad water

bad air

disease

truth

$$P(A \text{ and } B) = \boxed{\begin{matrix} P(A) \cdot P(B \mid A) \\ \shortparallel \\ P(B) \cdot P(A \mid B) \end{matrix}}$$

So

$$P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$$

$$\boxed{P(B \mid A) = \frac{P(B)\,P(A \mid B)}{P(A)}}$$

Bayes's theorem for T/F propositions (sets)

$$P(\text{cause} \mid \text{effect}) = \frac{P(\text{cause})\,P(\text{effect} \mid \text{cause})}{P(\text{effect})}$$

posterior info — prior info — data info

$$\boxed{P(\text{truth} \mid \text{data}) = \frac{P(\text{truth})\,\boxed{P(\text{data} \mid \text{truth})}}{P(\text{data})}}$$

a priori — a posteriori

before data — new data arrives — after data → time

$$\begin{pmatrix}\text{posterior}\\ \text{info.}\\ \text{about truth}\\ \text{given data}\end{pmatrix} = \frac{\begin{pmatrix}\text{prior}\\ \text{information}\\ \text{about truth}\end{pmatrix}\begin{pmatrix}\text{data}\\ \text{info.}\\ \text{about}\\ \text{truth}\end{pmatrix}}{\begin{pmatrix}\text{normalizing}\\ \text{constant}\end{pmatrix}}$$

[annoying]

truth data

prevalence sensitivity

$$P(A \mid \oplus) = \frac{P(A)\cdot P(\oplus \mid A)}{P(\oplus)}$$

partitions

LTP

3 methods to get

(1)

dichotomous ↔ having 2 possible values

truth

sensitivity

data

		A	not A	total
ELISA says	$\oplus$	384	2,988	3,372
	$\ominus$	16	96,612	96,628
total		400	99,600	100,000

prevalence

$$P(A \mid \oplus) = \frac{384}{3372} \doteq 11\% \ (!)$$

$$P(\text{not } A \mid \ominus) = \frac{96612}{96628} \doteq 0.9998$$

truth

data

	A	not A
$\oplus$	true ✓ positive	false × positive
$\ominus$	false × negative	true ✓ negative

ELISA says $\oplus$ but really is $\ominus$

blood bank: lose 1 unit blood ($850) bad

ELISA says $\ominus$ but really is $\oplus$

blood bank: somebody gets HIV terrible

western blot: a lot more (WB) accurate than ELISA but more expensive

adaptive strategy:

run ELISA on $\frac{1}{2}$ blood;

if $\ominus$ announce $\ominus$; if $\oplus$

run (WB) on other $\frac{1}{2}$ & announce WB result method 2 (creative laziness)

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

$$\overline{P(\text{not }B|A)} = \frac{P(\text{not }B)\,P(A|\text{not }B)}{\cancel{P(A)}}$$

$$\left(\frac{P(B|A)}{P(\text{not }B|A)}\right) = \left(\frac{P(B)}{P(\text{not }B)}\right) \cdot \left(\frac{P(A|B)}{P(A|\text{not }B)}\right)$$

$$\begin{pmatrix}\text{posterior}\\ \text{odds ratio}\\ \text{in favor}\\ \text{of } B\\ \text{given } A\end{pmatrix} = \begin{pmatrix}\text{prior}\\ \text{odds}\\ \text{ratio}\\ \text{in}\\ \text{favor}\\ \text{of } B\end{pmatrix}\begin{pmatrix}\text{Bayes}\\ \text{factor}\end{pmatrix}$$

Bayes factor $\updownarrow$ likelihood ratio $\updownarrow$ data odds in favor of B