STAT 206

7 Jan 24

lecture

1

this time: Bayes's Theorem for propositions

next time: model uncertainty

(*) ~~applied math paradigm~~

$\boxed{1}$ real-world problem

$\boxed{\mathbb{P} = (Q, C)}$

natural language

$\boxed{2}$ math restatement (translation) (**)

honest

$= (\ \underline{\ ?\ } \ )$

math

$\boxed{3}$ solve math problem ([redacted])

$\boxed{4}$ translate math solution back to natural language

(fairly hard)

(*) specialized to prob. & stat.

(**) consistency: (no logical contradictions) all math statements true

(**) completeness: nothing relevant & important left out quality of human activity ← how measure

process: what humans do — harder

outcomes: results of process — easier

problem

$\mathbb{P} = (\mathbb{Q}, \mathbb{C}) = (\theta, \mathcal{D}, \mathcal{B})$

real world questions

real-world context

unknown

usually in 206 $\Theta = \mathbb{R}$ or $\mathbb{R}^k$

k positive finite integer

data

$\{B_1, \dots, B_b\}$

consistency

completeness

uncertainty

as information ↑, good usually uncertainty ↓

but if mistake in prob. formulation, unc. ↑

"Bayesian prob." = quantification of your uncertainty about something (θ) of interest to you"

~~in applied math. modeling~~ we need to be constantly asking ourselves reality-check questions (often qualitative)

ex. if blood test is good,

$P(\theta = 1 \mid B)$ should go up when computing $P(\theta = 1 \mid y_1 = 1, B)$

0.01

A C B

$P(A)$ $P(A \mid B)$ $P(A \mid \underline{BC}) = P(A \mid CB)$

B and C

B, C

BC

B arrives

C arrives

Q how correctly update from $P(A|B)$ 4

to $P(A|BC)$?

T/F

A Bayes's Theorem

with +/- could get prob > 1 or < 0

$$P(A|BC) \stackrel{?}{=} \cancel{P(A|B) \pm ?}$$

$$\stackrel{?}{=} P(A|B) \cdot \boxed{?}$$

def: A, B T/F prop.

$$P(A|B) \triangleq \begin{cases} \dfrac{P(AB)}{P(B)} & \text{if } P(B)>0 \\ \text{undefined} & P(B)=0 \end{cases}$$

is defined to be

mathematical ~~dungeon~~ timeout zone

division by 0

If $1=2$ then any prop. B can be "shown" true

5.1

$$P(A|BC) = P(A|B) \cdot \boxed{?}$$

expand out conditional probability definitions

$$= \boxed{\frac{P(ABC)}{P(BC)}} \circledast = \frac{P(AB)}{\circledast \cdot P(B)} \cdot \boxed{\frac{P(B)}{P(AB)} \; \frac{P(ABC)}{P(BC)}}$$

get rid of | insert ⊛

call this box ⊛⊛

---

in playing-around mode, you discover that there are (several) 2 valid ways to reassemble the probabilities in box ⊛⊛ in conditional probability terms:

I

$$\frac{P(B)}{P(AB)} \; \frac{P(ABC)}{P(BC)} = \frac{1}{P(A|B)} \cdot P(A|BC)$$

this is

this is

---

but this is just a 1=1 tautology:

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

true but useless

switch order of denominator multiplication

II $$\frac{P(B)\,P(ABC)}{P(AB)\,P(BC)} = \boxed{\frac{P(B)}{P(BC)}} \cdot \boxed{\frac{P(ABC)}{P(AB)}}$$

this is $\frac{1}{P(C|B)}$

this is $P(C|AB)$

this yields (go back to (1) above

$$\boxed{P(A|BC) = \frac{P(A|B)\,P(C|AB)}{P(C|B)}}$$

this works

$\boxed{CS1}$

recall that

$A = (\theta = 1) =$ unknown

$B = \mathcal{B} =$ background

$C = (y_1 = 1) =$ data D

in $\boxed{CS1}$

$$P(\theta = 1 | y_1 = 1, \mathcal{B}) = \frac{P(\theta = 1 | \mathcal{B})\,P(y_1 = 1 | \theta = 1, \mathcal{B})}{P(y_1 = 1 | \mathcal{B})}$$

$P(A|BC) = \dfrac{\boxed{P(A|B)}\,\boxed{P(C|AB)}}{\boxed{P(C|B)}}$ ← test (6) sensitivity (0.999)

θ → A, B → B, D → C

0.01 → $P(A|B)$

? → $P(C|B)$

$P(\theta|D\,B) = \boxed{P(\theta|B)} \cdot \dfrac{P(D|\theta\,B)}{P(D|B)}$

pos → $P(\theta|B)$

LHS: $P(\theta|D\,B)$ (after data D arrives) (a posteriori)

posterior info about θ given D and B

RHS: $P(\theta|B)$: prior info about θ given B; a priori (before data arrives)

$P(D|\theta\,B)$: a bit mysteriously people call this likelihood info about θ given D and B (even though it's currently written $P(D|\theta\,B)$)

$P(D|B)$: (annoying) normalizing constant

LHS (left hand side) is a function of θ for fixed D so RHS (right-hand side) also has to have the same structure; $P(D|B)$ is constant in θ & plays the role of a normalizing constant

$$p(A|BC) = \boxed{p(A|C)} \frac{p(B|AC)}{p(B|C)}$$

1 2 3 6 4

0.01 ← 0.999 ?

θ D B

p_0 a priori — likelihood info about θ given D

$$p(\theta_{=1}|DB) = \frac{p(\theta|B)\, p(D|\theta_{=1} B)}{p(D|B)}$$

a posteriori

prior info about θ given B

(annoying) normalizing constant

posterior info about θ given D and B

LHS fn of θ for fixed D and B

	truth (θ) HIV(+)	HIV(−)	
test says (+) (y_i)	999		
(−)	1		
	1,000	99,000	100,000

sens.

prev.

$$p(y_i = 1 \mid \theta = 1, B) = \text{sens.} = .999$$

expected counts

	truth (θ)	
	($\theta=1$) HIV+	($\theta=0$) HIV−
($y_1=1$) test says +	999	
($y_1=0$) test says −	1	
	1,000	99,000

100,000

1,000 ← $(0.01) \cdot 100{,}000$

99,000 ← (by subtraction)

step 1 prevalence is $p(\theta \mid B) = 0.01$ (for people similar to Bob in all relevant ways)

step 2 sensitivity is $p(y_1 = 1 \mid \theta = 1, B) = 0.999$

since $\theta = 1$ we're in the 1^{st} column of the 2x2 contingency table above; $(0.999) \cdot 1{,}000 = 999$ expected people in the upper left cell, leaving 1 expected in the lower left cell