DS2500 Day 18 Mar 20

Admin:
- submitting data & analysis plan
- April 4th and 7th
    - Project meetings with Prof Higger
        - sign up link available on course schedule
        - 1 student per team sign up (share the link within team please)
    - no classes will be held (to allow time to meet)

Part 1: Bayes Rule
- basic probability (review)
- binary problems (review)
- bayes rule & independence

(Part 2 of today's notes has its own content table of contents too)

PROBABILITY (NOTATION + DEFINITIONS)

LET X BE A RANDOM VARIABLE REPRESENTING OUTCOME OF FAIR 6 SIDED DIE ROLL

"$X = x_1$" $\Longleftrightarrow$ "DIE ROLL WAS 1"

CAPITAL → RANDOM VARIABLE

LOWERCASE W/ SUBSCRIPT IS PARTICULAR OUTCOME

PROBABILITY (NOTATION + DEFINITIONS)

LET X BE A RANDOM VARIABLE REPRESENTING OUTCOME OF FAIR 6 SIDED DIE ROLL

" $P(X = x_1)$ " $\leftrightarrow$ "PROBABILITY THAT DIE OUTCOME IS 1"

SHORTHAND

WE DROP MENTION OF RANDOM VARIABLE

TO SAVE SPACE

$$P(x_0) = P(X = x_0)$$

PROBABILITY AXIOMS

$0 \leq P(x_i) \leq 1$

EACH OUTCOME PROB IS BETWEEN 0 AND 1

$\sum_i P(x_i) = 1$

SUM OF PROBABILITY OF ALL OUTCOMES IS 1

Conditional Probability (motivation)

Let C=1 indicate the event that a person has covid (C=0 otherwise)
Let A=1 indicate the event that an antigen test is positive (A=0 otherwise)

$P(C=1)$

Let us discuss (and express) the following probabilities:

$$P(C=1 \mid A=1) = \frac{P(A=1, C=1)}{P(A=1)}$$

- probability person has a positive antigen test $P(A=1)$

- probability person has covid and a positive antigen test $P(A=1, C=1)$

- probability person has covid given a positive antigen test $P(C=1 \mid A=1)$

→ "GIVEN"

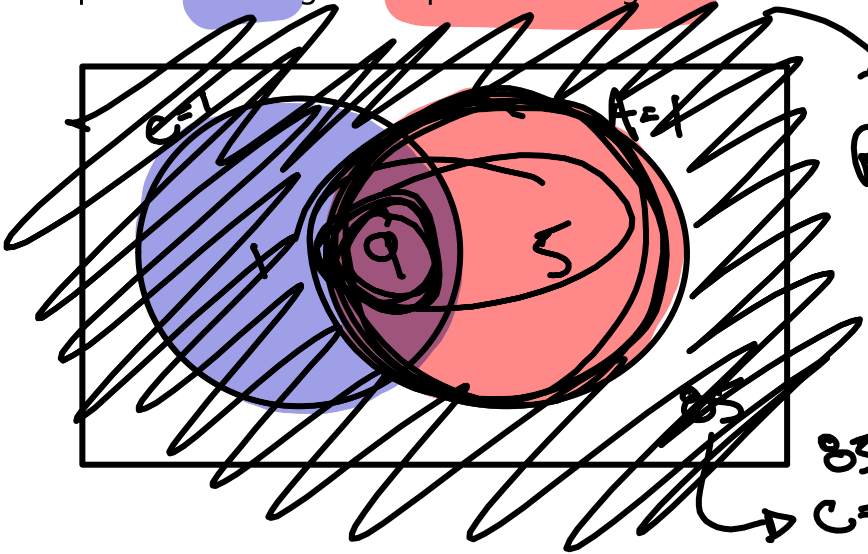
A conditional probability gives the probability of one event given that another has occured.

A conditional probability P(a|b) ignores all states which don't meet a condition

Given the following, esitmate

- the prob of covid
- the prob of covid given a positive antigen test

$$\text{Prob Covid} = P(C=1) = \frac{1+9}{1+9+5+85} = 1/10$$



$$P(C=1 \mid A=1) = \frac{9}{9+5} = \frac{9}{14}$$

85 people w/ C=0 A=0

Conditional Prob (Algebraic Definition)

Prob a b happen together

$$P(a|b) = \frac{P(a\,b)}{P(b)}$$

Prob a happens given condition b

Prob b happens

Conditional Prob (Algebraic Definition)

$$P(a|b) = \frac{P(a\,b)}{P(b)} \longleftrightarrow P(a|b)\,P(b) = P(a\,b)$$

Takeaway above:

Multiplying
- a conditional probability
- the probability of condition

Will yield
- prob both outcomes happen together

BAYES RULE (GLORIFIED CONDITIONAL PROBABILITY)

SEE PREVIOUS "TAKEAWAY"

$$P(a|b)\,P(b) = P(ab) = P(b|a)\,P(a)$$

$$\Rightarrow \boxed{P(a|b)} = \frac{P(b|a)\,P(a)}{P(b)}$$

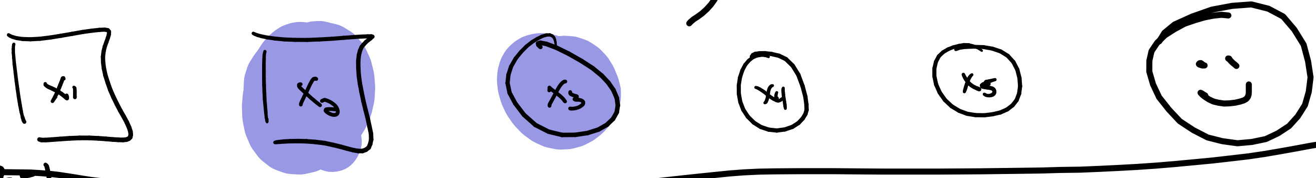
Notice: this formula "swaps" the order of the conditioning: P(A|B) on left P(B|A) on right
Its typical in a Bayes question to be given variables in one order while question asks for other.

Marginalizing

(Adding | Removing a random variable from prob)

B=1 shape is blue

C=1 shape is circle



	B=0	B=1
C=0	x_1 1/5	x_2 1/5
C=1	x_4 x_5 2/5	x_3 1/5

$$P(B=1) = P(B=1, C=0) + P(B=1, C=1)$$
$$= 1/5 + 1/5$$
$$= 2/5$$

Remember: To compute P(B) we can sum P(B, A) for all outcomes in sample space of A

$$P(B=b) = \sum_a P(B=b, A=a)$$

A Helpful Manipulation

Marginalization

$$P(b) = \sum_a P(a\,b)$$

Conditional Prob Definition

$$= \sum_a P(b|a)\,P(a)$$

WHY WAS THAT HELPFUL?

BAYES RULE 1

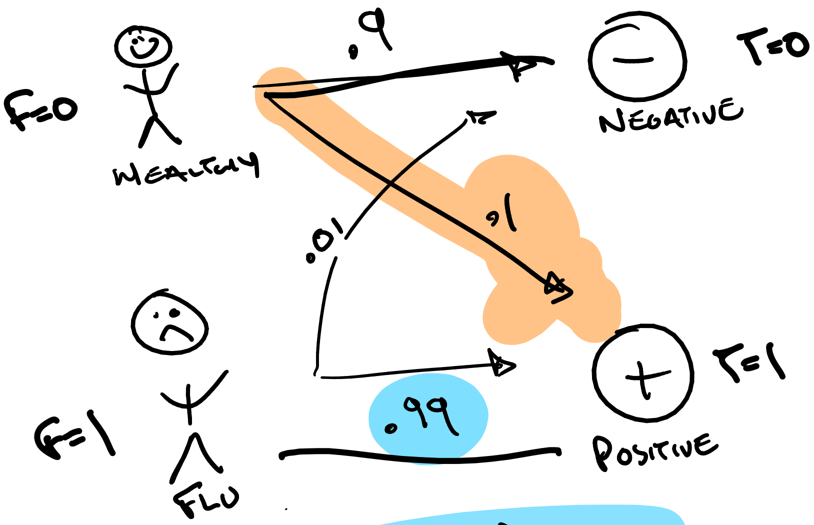
$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

BAYES RULE 2

$$P(a|b) = \frac{P(b|a)P(a)}{\sum_i P(b|a_i)P(a_i)}$$

Bayes Rule Ex

Given flu occurs in .04 of population, what is the probability one has flu given they test positive?



$$P(F=1\mid T=1) = \frac{P(T=1\mid F=1)\,P(F=1)}{P(T=1)}$$

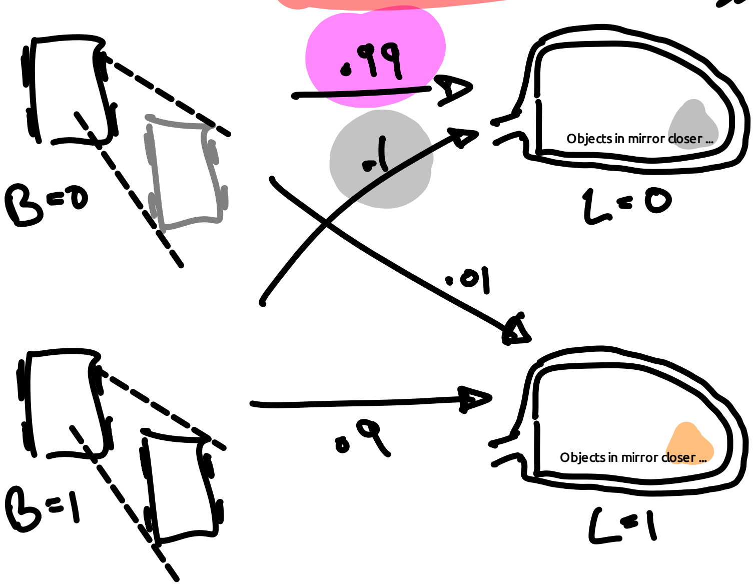
$$= \frac{P(T=1\mid F=1)\,P(F=1)}{P(T=1\mid F=1)\,P(F=1) + P(T=1\mid F=0)\,P(F=0)}$$

$$P(T=1\mid F=1) = .99$$

$$= \frac{.99 \cdot .04}{.99 \cdot .04 + .1 \cdot .96} \approx .29$$

In Class Assignment 1

A blind spot monitor produces a warning light (L=1) when it estimates that a car is in one's blind spot (B=1). Given that the light is off, whats the probability that a car is one's blind spot? (Assume that a car is in your blindspot .02 [redacted] of the time while driving.)

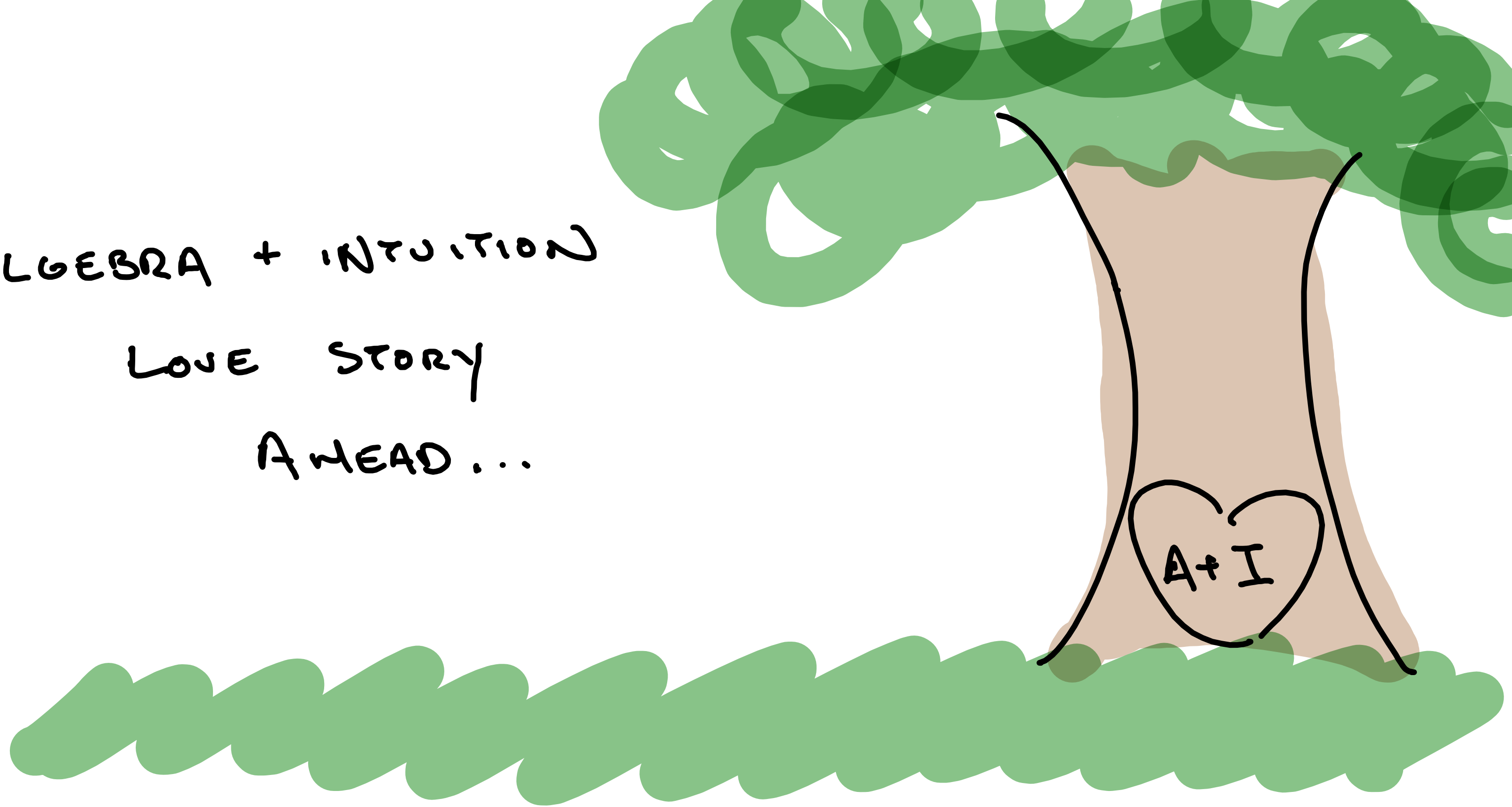


$$P(B=1|L=0) = \frac{P(L=0|B=1)\,P(B=1)}{P(L=0|B=1)\,P(B=1) + P(L=0|B=0)\,P(B=0)}$$

$$= \frac{.1 \cdot .02}{.1 \cdot .02 + .99 \cdot (1-.02)}$$

$$\approx .00205$$

ALGEBRA + INTUITION
LOVE STORY
AHEAD...



INDEPENDENCE + CONDITIONAL PROB

INDEPENDENCE

INTUITION:

Random variables x, y are independent if observing any outcome of one doesn't impact our beliefs about the other.

ALGEBRA:

FOR EACH OUTCOME PAIR x, y

$$P(X=x, Y=y) = P(X=x)\,P(Y=y)$$

Bayes Rule shows the equivilence of the algebraic and intuitive definitions above!

INDEPENDENCE + CONDITIONAL PROB

INDEPENDENCE

INTUITION:

Random variables x, y are independent if observing any outcome of one doesn't impact our beliefs about the other.

ALGEBRA:

FOR EACH OUTCOME PAIR x, y

$$P(X=x\ Y=y) = P(X=x)\,P(Y=y)$$

$$P(X|Y) = \frac{P(X\,Y)}{P(Y)} = \frac{P(X)P(Y)}{P(Y)} = P(X)$$

Notice that P(X|Y) = P(X). Observing Y has no impact on the prob of X!