CS1800
Discrete Structures
Fall 2017

Lecture 26
11/6/17

Last time

- Finish Entropy
- Start Cond. Prob.
  - Bayes Law

Today

- Finish Cond. Prob. & Bayes Law
- Start Markov Chains & Page Rank

Next time

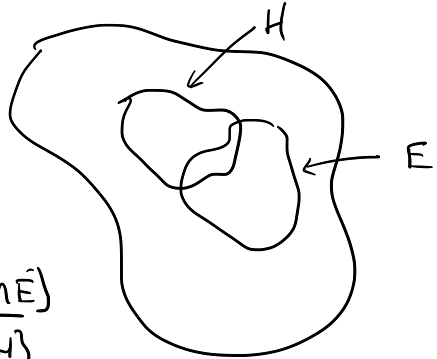
- Finish MC & PR

- Start Module 4: Algorithms & Analysis

Conditional Probability

prob of E given H

$$\Pr\{H \cap E\} = \Pr\{H\} \cdot \underline{\Pr\{E \mid H\}}$$



$$\Rightarrow \Pr\{E \mid H\} = \frac{\Pr\{H \wedge E\}}{\Pr\{H\}}$$

$$= \Pr\{E\} \cdot \Pr\{H \mid E\}$$

$$\Rightarrow \Pr\{H \mid E\} = \frac{\Pr\{H \wedge E\}}{\Pr\{E\}}$$

Independence

$\Pr\{E \mid H\} = \Pr\{E\}$

$\Pr\{H \mid E\} = \Pr\{H\}$

$\Rightarrow \Pr\{H \cap E\} = \Pr\{H\} \cdot \Pr\{E \mid H\} = \Pr\{H\} \cdot \Pr\{E\}$

Bayes Law:

$$\underline{\Pr\{E\}} \cdot \underline{\Pr\{H \mid E\}} = \Pr\{H \wedge E\} = \underline{\Pr\{H\}} \cdot \underline{\Pr\{E \mid H\}}$$

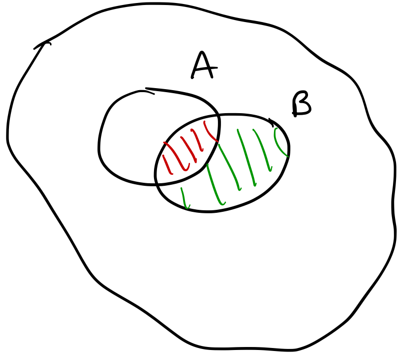
$$\Pr\{H \mid E\} = \frac{\Pr\{E \mid H\} \cdot \Pr\{H\}}{\Pr\{E\}}$$

H = hypothesis: do have Zika?

E = evidence: did blood test come back pos?

Bayes Law

$$\Pr(B|A)\cdot\Pr(A) = \Pr(A\cap B) = \Pr(A|B)\cdot\Pr(B)$$



$$\Pr(A|B) = \frac{\Pr(B|A)\cdot\Pr(A)}{\Pr(B)}$$

$$= \frac{\Pr(B|A)\cdot\Pr(A)}{\Pr(B|A)\cdot\Pr(A) + \Pr(B|\bar{A})\cdot\Pr(\bar{A})}$$

what is $\Pr(B)$?

$$p(B) = \quad + $$

$$= p(B|A)\cdot p(A) + p(B|\bar{A})\cdot p(\bar{A})$$

$$\Pr(H|E) = \frac{\Pr(E|H)\cdot\Pr(H)}{\Pr(E|H)\cdot\Pr(H) + \Pr(E|\bar{H})\cdot\Pr(\bar{H})}$$

Example: Zika in FL 2016

- Prevalence of Zika in S. FL. $p(\text{Zika}) = 10^{-5}$ (1 in 100,000)
- accuracy of blood test is 99 %

i.e. $p(\text{pos. test} \mid \text{Zika}) = 0.99$ ← test subjects who had Zika

$p(\text{pos. test} \mid \text{no Zika}) = 0.01$ ← control group who don't have Zika

- You test <u>positive</u>: what it is the chance that you have Zika?

$\Longrightarrow$ <u>Not</u> $p(\text{pos. test} \mid \text{Zika}) = 0.99$

$\Longrightarrow$ You want $p(\text{Zika} \mid \text{pos. test})$!

$$p(\text{zika} \mid \text{pos test}) = \frac{p(\text{pos. test} \mid \text{zika}) \cdot p(\text{zika})}{p(\text{pos test})}$$

$$= \frac{p(\text{pos. test} \mid \text{zika}) \cdot p(\text{zika})}{p(\text{pos test} \mid \text{zika}) \cdot p(\text{zika}) + p(\text{pos test} \mid \text{not zika}) \cdot p(\text{not zika})}$$

$$= \frac{0.99 \times 10^{-5}}{0.99 \times 10^{-5} + 0.01 \times (1 - 10^{-5})}$$

$$= \frac{0.0000099}{0.0000099 + 0.0099999}$$

$$\approx 0.00099$$

$$\approx 0.1\%$$ i.e. only 1 in 1,000 !

Seems wildly counter intuitive, but...

10,000,000 people in FL. 10^7

w/ zika : $10^7 \cdot 10^{-5} = 10^2 = 100$

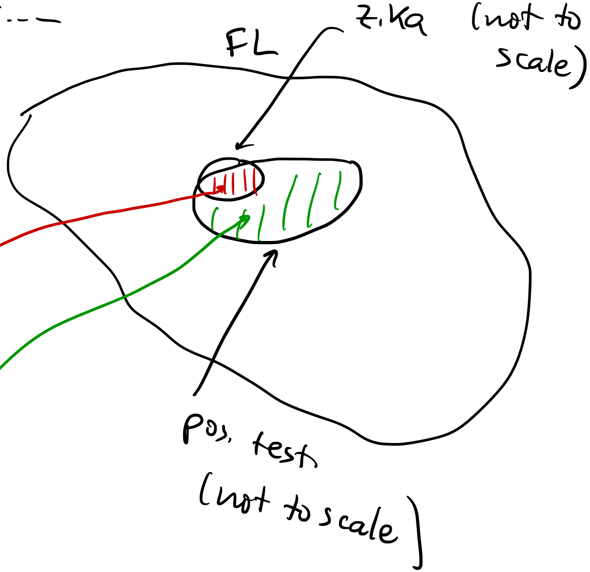
w/o zika : 9,999,900

---

test pos w/ zika : $100 \cdot 0.99 = 99$

test pos. w/o zika : $9{,}999{,}900 \times 0.01$

$= 99{,}999$



$\therefore$ If you test pos.,

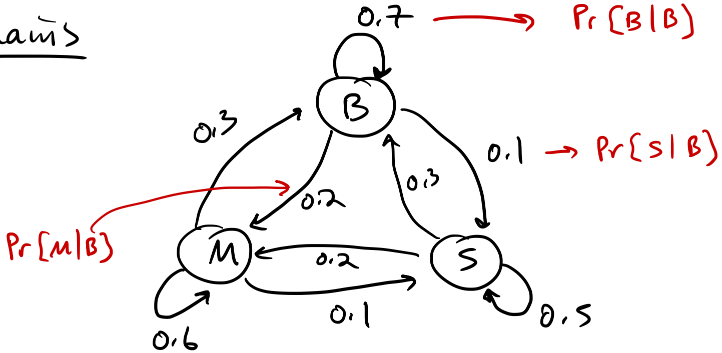
you're about 1,000 times more likely to be among 99,999 who <u>don't</u> have zika but test pos. than among the 99 who do have zika and test pos.

Markov chains

B: Bertucci's

M: Margaritas

S: Sato



3 reds ⇒ must add to 1