AMAT 362 - Prob. for Stats - Lecture 6

Lec 5: Conditional Probability $P(A|B)$

Law of Total Probability

$\hookrightarrow P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$

If $B_1, \dots, B_n$ partitions Ω, $P(B_i) > 0$ $\Rightarrow$ $\boxed{P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)}$ (Pitman)

RULE OF AVERAGED CONDITIONALS

Lec 6: BAYES' RULE

$$P(A|B) \neq P(B|A)$$

INDEPENDENCE e.g. when $P(A|B) = P(A)$ NOT ALWAYS TRUE

PROSECUTOR'S FALLACY

- City w/ 1 Million people
- A horrible crime is committed
- A sketch of the suspect is made FTD

only $\frac{1}{10{,}000}$ people fit the description

Prosecutor says of an arrested suspect

1) Probability of FTD is low
2) It is unlikely that an innocent person FTD.
3) Therefore it is unlikely the defendant is innocent.

What is being said in Step 2?

$$P(FTD \mid I) = \text{prob. of fitting the description assuming } \underbrace{\text{innocent}}_{I} = \frac{P(FTD \cap I)}{P(I)}$$

N.B. $\frac{1}{10{,}000} \cdot 1 \text{ mill} = \frac{10^6}{10^4} = 10^2 = 100$ people FTD

$$= \frac{\frac{99}{10^6}}{\frac{999{,}999}{10^6}} = \frac{99}{999{,}999} \sim \frac{1}{10^4}$$

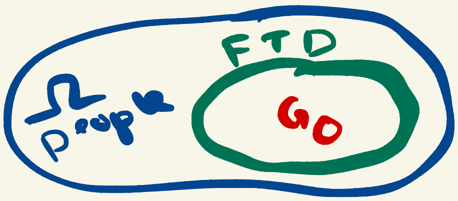
Q: Is it true that

$$P(FTD \mid I^c) \overset{?}{\neq} 1 - P(FTD \mid I)$$

A: No!

$$P(FTD \mid I^c) = \frac{P(FTD \text{ AND Guilty})}{P(\text{Guilty})} = \frac{P(G)}{P(G)} = 1$$

$P(I^c) = P(\text{Guilty})$



N.B. What is true $P(A \mid B) = 1 - P(A^c \mid B)$

Q: What should we be calculating?

$$P(G \mid FTD) = 1 - \underbrace{P(I \mid FTD)}$$

$$P(I \mid FTD) = \frac{P(I \cap FTD)}{P(FTD)} = \frac{\frac{99}{\cancel{10^6}}}{\frac{100}{\cancel{10^6}}} = \frac{99}{100}$$

$P(G \mid FTD) = 1\%$

3 FORMS of BAYES' RULE

1) $P(B|A) = \dfrac{P(A|B)P(B)}{P(A)}$ simplest

2) If $P(A), P(B), P(B^c) > 0$

then $P(B|A) = \dfrac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$

3) If $B_1, \ldots, B_n$ partition the space of possibilities Ω

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$

BAYESIAN LANGUAGE

In form 3 of Bayes' Rule

$P(B_i)$ = The "prior" probability of B_i occuring before observing A

$P(B_i|A)$ = The "posterior" probability of B_i knowing A occurred.

THE FALSE POSITIVE "PARADOX"

HIV ⇝ AIDS

In 2016 in NYS $\frac{760}{100,000}$ had HIV

Let B_1 = has HIV ↪ $P(B_1) = \frac{760}{100,000}$

B_2 = no HIV → $P(B_2) = \frac{99,240}{100,000}$

3rd Gen antigen test

↳ 3 months after exposure

$P(+/B_1) = 99.7\% = \frac{997}{1000}$

$P(+/B_2) = 1.5\%$ ← FALSE POSITIVE RATE

N.B. $P(+/B_1) + P(+/B_2) \neq 100\%$

b/c $P(A/B) + P(A/B^c) \neq 1$

BIG QUESTION: $P(B_1 | +) = ?$

BAYES' RULE $P(B_1|+) = \dfrac{P(+|B_1)P(B_1)}{P(+)}$

LAW OF TOTAL PROBABILITY → $P(+)$

$$P(+) = P(+|B_1)P(B_1) + P(+|B_2)P(B_2)$$
$$= (.997)\left(\frac{760}{100,000}\right) + (.015)\left(\frac{99,240}{100,000}\right)$$

758 out of 760 HIV + people will test positive

BUT 1,488 of the 99,240 of HIV − people will test positive

$P(+) \sim 2.24\%$

$$\Rightarrow P(B_1|+) = \frac{758}{758 + 2{,}246} \sim 33\%$$

If you test positive don't panic! only ~33% chance of having HIV

N.B. 4th Antigen test

$P(+|B_1) \sim 95\%$ $\quad P(+|B_2) \sim \frac{40}{10{,}000}$

$\Rightarrow P(HIV|+) \sim 64.5\%$

INDEPENDENCE!

Def A & B are <u>independent</u> if

$$P(A \cap B) = P(A)P(B)$$

OR $P(A|B) = P(A)$ OR $P(B|A) = P(B)$

Ex

Dealt 2 cards from 52 card deck

E_1 = 1st card is Black

E_2 = 2nd card is Black

Q: Are E_1 & E_2 independent?

A: NOT INDEP B/C

$$P(E_2 | E_1) = \left(\frac{25}{51}\right)$$

NOT = SO NOT INDEP

BUT $P(E_2) = \frac{26}{52} = \frac{1}{2}$

Why? Sampling without replacement → NON-INDEP EVENTS

Q: Deck w/ 2 cards 1R, 1B

DEALT ONE, Then replace it and shuffle, deal again

(Sampling w/ replacement)

E_1 & E_2 as before

$$P(E_2 / E_1) = P(E_2) = \frac{1}{2}$$