Warm up

If $P(A|B)=0.2$, $P(C|B)=0.7$, and $P(B)=0.6$

find

$$P(A\cap B) = P(B)\,P(A|B) = 0.6 * 0.2 = 0.12$$

and

$$P(B\cap C) = P(C)\,P(B|C)$$

$$\parallel$$

$$P(C\cap B) = P(B)\,P(C|B) = 0.6 * 0.7 = 0.42$$

17 Apr

- No class on Monday
  ↳ moodle page contain reading assignment

- Class evals are opened

On final { if 70% of class does CE +1
80% +2
90% +4

class evaluations close the morning of our final

- Web Assign can be extended
  - if requested, automatically granted
  - extension lasts 24 hrs
  - ∞ # of extensions

Last class we learned

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

and therefore

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

the probability of A and B occuring

the probability of B occuring * the probability of A occuring give that B occured

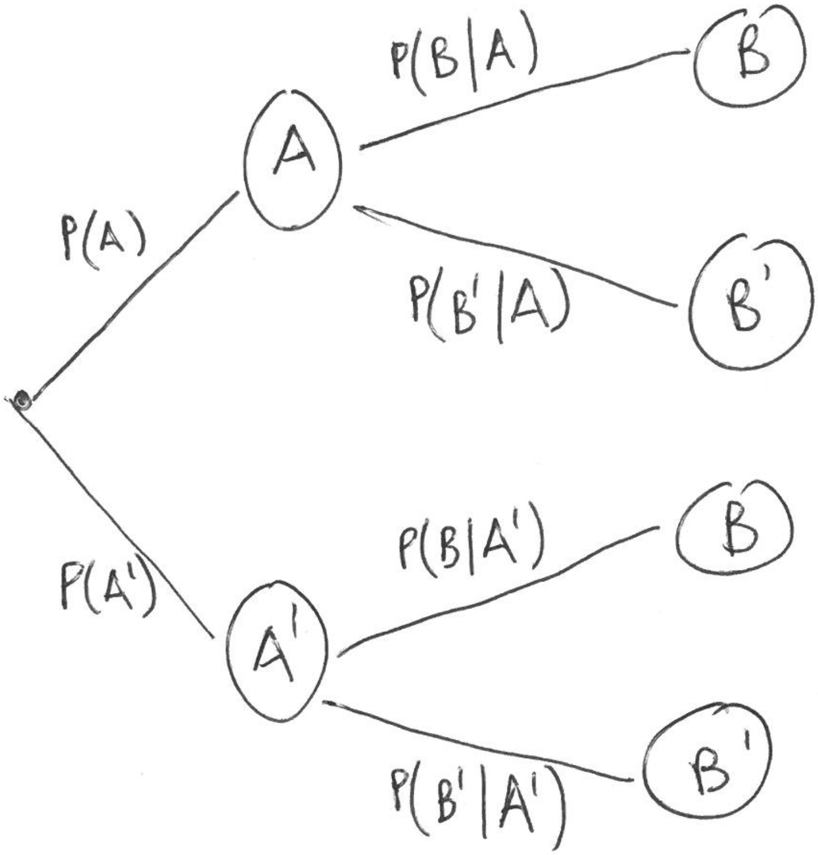
Why did we use $P(B)$?

in one sense we don't need to

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

so it depends on what we know

Reminder about trees...



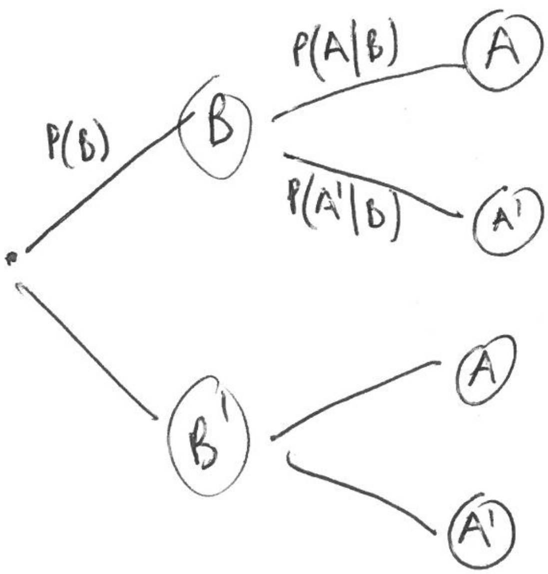
$P(A \cap B) = P(A)P(B|A)$

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

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

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

maybe our problem tells us $P(B)$ not $P(A)$

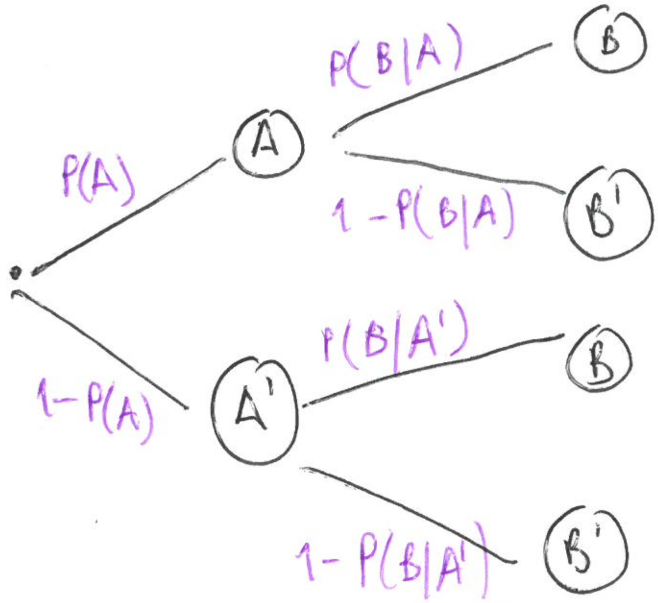


$P(B \cap A) = P(B)P(A|B)$

⋮

If we knew $P(A)$, $P(B|A)$, $P(B|A')$

Can we find $P(B)$?



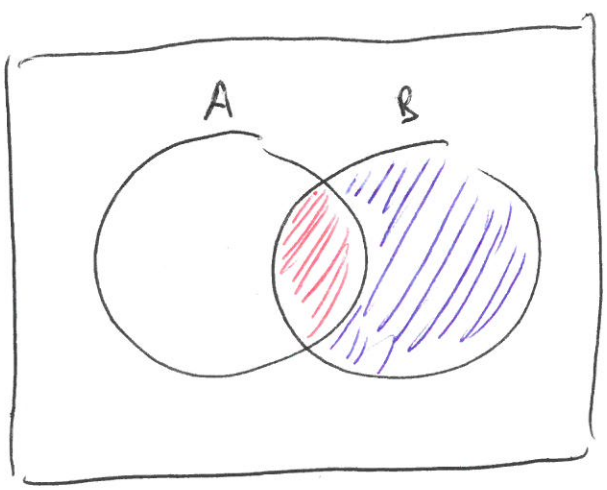
In which scenarios does B occur?

$A \cap B$ and $A' \cap B$

these are disjoint

$$P(B) = P(A \cap B) + P(A' \cap B)$$
$$= P(A)P(B|A) + (1 - P(A))P(B|A')$$
$$= P(A)P(B|A) + P(A')P(B|A')$$

does this make sense?



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

$P(\;) = P(A' \cap B) = P(A')P(B|A')$

Imagine a particular disease exists in

.1 % of the population.

- A diagnostic to detect this disease that always says 'you don't have it' is correct 99.9% of the time

↳ this isn't useful though because

$$P\left(\begin{matrix}\text{test}\\\text{positive}\end{matrix}\;\middle|\;\begin{matrix}\text{have}\\\text{disease}\end{matrix}\right) = 0$$

- Change it to always say 'You do have it'

here $P\left(\begin{matrix}\text{test}\\\text{positive}\end{matrix}\;\middle|\;\begin{matrix}\text{have}\\\text{disease}\end{matrix}\right) = 1$

~~but~~ still, no good, why

$$P\left(\begin{matrix}\text{test}\\\text{negative}\end{matrix}\;\middle|\;\begin{matrix}\text{don't}\\\text{have it}\end{matrix}\right) = 0$$

- Somehow this real life example very intuitively relies on condition probabilities

A steroid testing company says

96% of steroid users will test positive

$$P(TP \mid US) = .96$$

and only 6% of non steroid users test positive

$$P(TP \mid US') = 0.06$$

In a particular sport it's believed only 5% of athletes use steroids.

Imagine an athlete tests positive but claims it's a mistake.

How likely is this?

We might think 6% <u>BUT</u>

What are we asking?

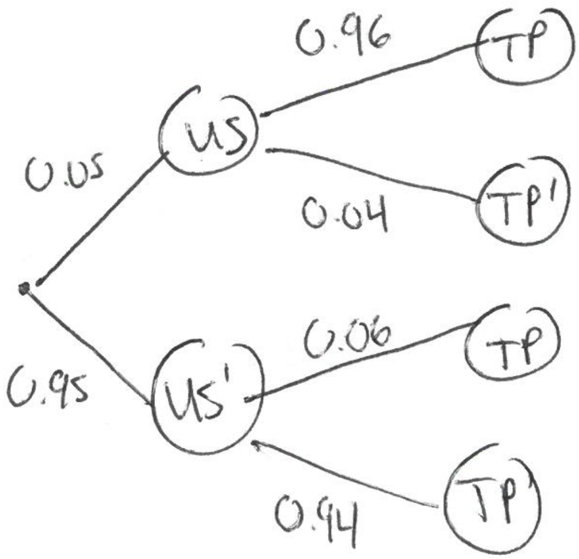
$P(US' \mid TP)$ and this <u>isn't</u> $P(TP \mid US')$

So what's $P(US' | TP)$?

$$P(US' | TP) = \frac{P(US' \cap TP)}{P(TP)}$$

we don't know these

BUT we can find them



$P(\quad) = 0.05 * 0.96$

$P(US' \cap TP) = 0.95 * 0.06$

$$P(TP) = P(US' \cap TP) + P(US \cap TP)$$

so

$$P(US' | TP) = \frac{0.95 * 0.06}{0.95 * 0.06 + 0.05 * 0.96} = 0.\overline{54}$$

More 50% of those that test positive aren't using

* this is a bad diagnostic *

How did we get this

$$P(US' \mid TP) = \frac{P(US' \cap TP)}{P(TP)}$$

$$= \frac{P(US')\,P(TP \mid US')}{P(US)\,P(TP \mid US) + P(US')\,P(TP \mid US')}$$

In general

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

equivalently

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

This is called

Bayes' Theorem

In Greenland it snows on average once every 30 days and when it does glaciers have a 27% chance of growing. If it doesn't snow, there's still a 7% of glacial growth.

What's the probability it's snowing ~~when~~ if I know glaciers are growing?

$$P(S|G) = \frac{P(S \cap G)}{P(G)}$$

$$= \frac{P(S)P(G|S)}{P(S)P(G|S) + P(S')P(G|S')}$$

$$= \frac{(1/30)(.27)}{(1/30)(.27) + (29/30)(0.07)}$$

$$= 0.09375$$

An AI vehicle identifier can correctly a

sedan 85% = $P(DS|S)$

truck 60% = $P(DT|T)$

In Italy only 5% of vehicles are pickup trucks

$$P(T|DT) = \frac{P(T)P(DT|T)}{P(T)P(DT|T) + P(T')P(DT|T')}$$

$$= \frac{(0.05)(0.60)}{(0.05)(0.60) + (0.95)(0.15)}$$

$$= 0.174$$

if our identifier tells us a truck passed, it was probably a sedan

if $$\frac{(0.40)(0.6)}{(0.40)(0.6) + (0.6)(0.15)} = 0.\overline{72}$$