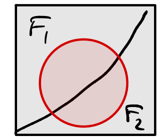
3.3. BAYES FORMULA

- RECAP
  - Def: $P(E|F) = \dfrac{P(E\cap F)}{P(F)}$ $\Rightarrow$ chain rule $P(E\cap F) = P(E|F)\cdot P(F)$
  - LTP: if $S = \cup F_i$ $\Rightarrow$ $P(E) = \sum_i P(E\cap F_i) = \sum_i P(E|F_i)P(F_i)$

    mutually exclusive



- $P(E|F) \neq P(F|E)$ e.g. for E = "has a brain", F = "is a mathematician"
- BF Allows to compute $P(F|E)$ from $P(E|F)$

$$\boxed{P(F|E)} \underset{\text{Def}}{=} \frac{P(E\cap F)}{P(E)} \underset{\text{chain rule}}{=} \boxed{\frac{P(E|F)\,P(F)}{P(E)}} \underset{\text{L.T.P.}}{=} \frac{P(E|F)P(F)}{P(E|F)P(F)+P(E|F^c)\,P(F^c)}$$

BAYES FORMULA

Ex Celiac disease affects 1% of population. ← $P(C)$

A person is tested positive.
What is prob. that s/he has celiac disease?
A blood test is false positive in 5% cases. ← $P(+|C^c)$
(positive result for a healthy person)
A blood test is false negative in 10% cases, ← $P(-|C)$
(negative result for a sick person)

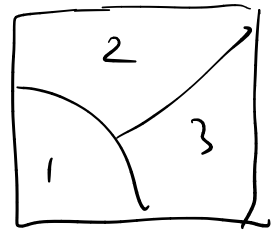
C = "has celiac", $+$ = "test positive".

$$P(C|+) = \frac{\overbrace{P(+|C)}^{1-P(-|C)=0.9}P(C)}{P(+|C)P(C)+P(+|C^c)P(C^c)} = \frac{0.9\times 0.01}{0.9\times 0.01 + 0.05\times 0.99} \approx \boxed{0.18}$$

- So small! Why? Positive tests are dominated by false positives: only a tiny fraction of the population is actually sick.

- Bayes formula generalizes for more than 2 events since L.T.P. does:

Ex (3k) • A hiker went missing in the wilderness. She is one of the three regions with probabilities 0.5, 0.3, 0.2

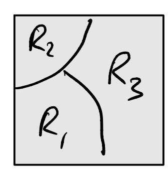


- Whenever anyone is lost in region 1, the search is successful with prob. 0.4 and for regions 2, 3 it is 0.7 and 0.8.
- A search in region 1 was unsuccessful. What is the prob. that the person is in region 1?

R_i = "the hiker is in region i", $i=1,2,3$.

U_i = "the search in region i is unsuccessful"

$$P(R_1|U_1) = \frac{P(U_1|R_1)\,P(R_1)}{P(U_1)}$$

S= 

$$= \frac{P(U_1|R_1)\,P(R_1)}{P(U_1|R_1)P(R_1)+P(U_1|R_2)P(R_2)+P(U_1|R_3)P(R_3)}$$

L.T.P.

$$= \frac{(1-0.4)\,0.5}{(1-0.4)\,0.5+1\times 0.3+1\times 0.2} = 0.375$$

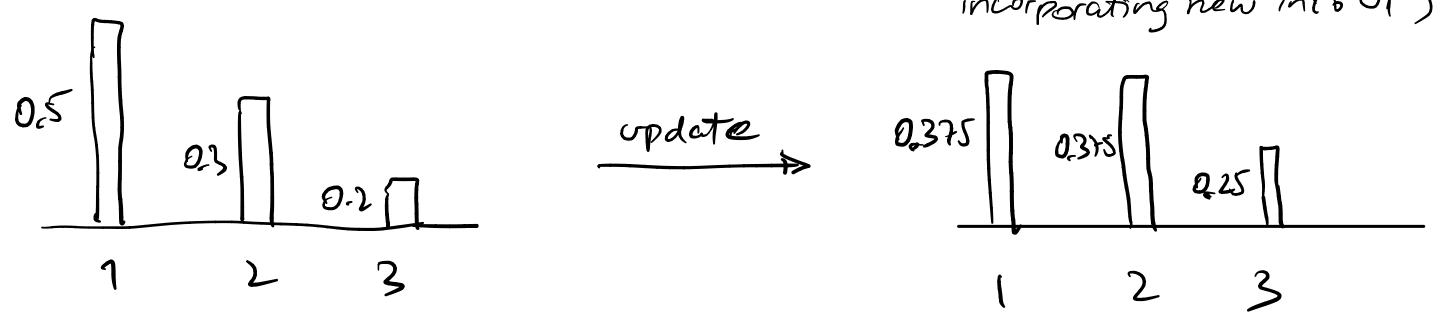
• Similarly, $P(R_2|U_1) = \dfrac{P(U_1|R_2)P(R_2)}{\text{same}} = 0.375$

$P(R_3|U_1) = \dfrac{P(U_1|R_3)P(R_3)}{\text{same}} = 0.25$ } Check!

• Thus, the <u>prior</u> probabilities $P(R_1)$, $P(R_2)$, $P(R_3)$ got updated to the <u>posterior</u> probabilities $P(R_1|U_1)$, $P(R_2|U_1)$, $P(R_3|U_1)$.

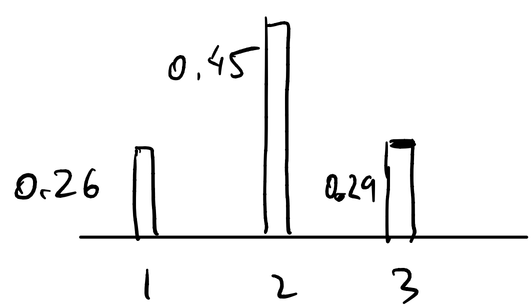
<u>PRIOR</u>
(initial degree of belief)

<u>POSTERIOR</u>
(degree of belief after incorporating new info U_1)



<u>Ex</u> Another search in reg. 1 is unsuccessful

⇒ prob's are further updated to



(Check!)

<u>Remark</u> Bayesian model of <u>learning</u>

(ML, AI, e.g. NLP, ...)

Ex: Instagram meme:

$\underbrace{\text{Sober drivers}}_{S}$ cause 75% of $\underbrace{\text{accidents}}_{A}$ [Instagram]

while $\underbrace{\text{drunk drivers}}_{D}$ cause only 25%

$\Rightarrow$ sobriety increases the odds of an accident by $\times 3$ $\left(= \frac{75}{25}\right)$

?! Paradox.

Assumption is about $P(S|A) = 0.75$, but
$\neq$
Conclusion is about $P(A|S)$. Let's compute by B.F.

$$P(A|S) = \frac{P(S|A)\,P(A)}{P(S)}$$

Compare to

$$P(A|D) = \frac{P(D|A)\,P(A)}{P(D)} \Rightarrow$$

$$\frac{P(A|D)}{P(A|S)} = \underbrace{\frac{P(S)}{P(D)}}_{\geq \frac{9}{1}} \cdot \underbrace{\frac{P(D|A)}{P(S|A)}}_{= \frac{0.25}{0.75} = 3} \geq 3.$$

if we assume that ≤ 1 in 10 drivers is drunk

$\Rightarrow$ drinking (not sobriety) increases the odds of accidents by $\times 3$.