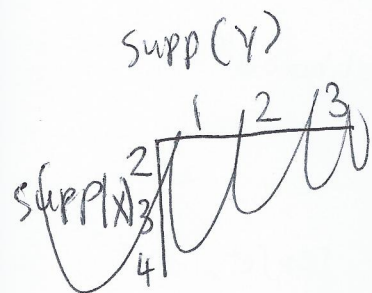


3-1)

$$P(A_i | B) = \frac{P(B|A_i) P(A_i)}{\sum_{i=1}^n P(B|A_i) P(A_i)} \quad \xrightarrow{390-3} P(A_i, B)$$

$P(B)$

Baye's Theorem

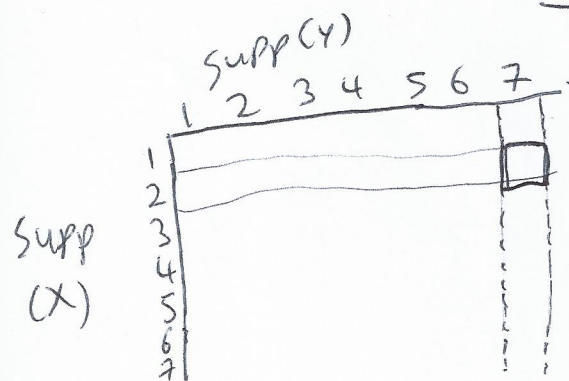


$$P(X=2 | Y=7) = \frac{P(X=2, Y=7)}{P(Y=7)}$$

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

$$\rightarrow \sum P(X, 7)$$

$$X + \text{Supp}(7)$$



dis joint

$$P(Y=7) = P(X=1, Y=7) + P(X=2, Y=7) + P(X=3, Y=7) + P(X=4, Y=7) + P(X=5, Y=7) + P(X=6, Y=7) + P(X=7, Y=7)$$

$$f(x, y) = \frac{f(x, y)}{f(y)} = \frac{f(y|x) f(x)}{f(y)}$$

$$P(Y) = \sum_{x \in \text{Supp}(X)} P(X, Y)$$

marginalization

$$f(y) = \int f(x, y) dx$$

$x \in \text{Supp}(X)$

$f(x, y)$

$$P(\theta|x) = \frac{P(x|\theta) p(\theta)}{P(x)} \leftarrow \begin{array}{l} \text{Posterior} \quad \text{likelihood of } x \quad \text{Prior for } \theta \\ \text{Prior for } (x) \\ \text{"Marginal likelihood"} \end{array}$$

Bayesian

Frequentist

$$P(\theta|x) = \frac{P(x;\theta) P(\theta)}{P(x)} \leftarrow \begin{array}{l} \text{makes no sense,} \\ \text{either 0, or 1} \end{array}$$

θ is set.

X big
no, no

Makes no sense because you cannot
calculate $P(x)$ without knowing what θ is

Coin flip $\in \{0, 1\}$

$$P(X|\theta=0.75) = .25 \cdot .75 \cdot .75 = .141$$

$$P(X|\theta=\text{~~0.000~~ } 0.5) = .5 \cdot .5 \cdot .5 = .125$$

$$P(\theta=0.25) ?? \quad P(\theta=0.75) ??$$

Principle of indifference

All prior hypothesis about θ are equally likely.

$$\underbrace{P(\theta=0.25)}_{\text{idea 1}} = \underbrace{P(\theta=0.75)}_{\text{idea 2}} = \frac{1}{2}$$

$$\Omega = X \times \mathbb{H}_0 \xrightarrow{\quad} \mathbb{H}_0 \subseteq \mathbb{H} \quad 3-2)$$

$$\xrightarrow{\quad} \text{Supp}[X]^n$$

$\langle 0,0,0 \rangle$				$\langle 1,0,0 \rangle$	$\langle 0,1,0 \rangle$	$\langle 0,0,1 \rangle$	110	011	101	111	$\theta = 0.25$
000	100	010	001	110	101	011	$\langle 0,0,0 \rangle$				$\theta = 0.75$

$$\text{Sample space} = \text{Supp}(X)^3$$

$$X = \{(0,0,0), (0,0,1), (0,1,0), (1,0,0), (0,1,1), (1,0,1), (1,1,0), (1,1,1)\}$$

$$P(\theta = 0.25) \quad P(0,0,0 | \theta = 0.25) = .75^3 = .421$$

$$\text{is } \langle 0,0,0 \rangle, \theta = 0.25 \cap \langle 0,0,0 \rangle, \theta = 0.75 = \emptyset?$$

Yes. different piece of the universe.

$$\text{let } X = (0,1,1) \quad \text{what is } P(\theta = 0.25)$$

$$P(\theta = 0.25 | X = \langle 0,1,1 \rangle)$$

$$\langle 0,1,1 \rangle = \Omega_0 \text{ @ Universe}$$

Therefore

$$= \frac{P(0,1,1 | \theta = 0.25)}{P(0,1,1 | \theta = 0.25) + P(0,1,1 | \theta = 0.75)}$$

all $P(\langle 0,1,1 \rangle)$

$$= \frac{P(0,1,1 | \theta = 0.25) P(\theta = 0.25)}{[P(0,1,1 | \theta = 0.25) (P(\theta = 0.25))] + [P(0,1,1 | \theta = 0.75) P(\theta = 0.75)]}$$

$$= \frac{.047 * .5}{.047 * .5 + .141 * .5} = .25$$

$$P(\theta = -.75 | (0, 1, 1)) = .75$$

$$= P(\theta = -.25)^c | (0, 1, 1) = 1 - .25 = .75$$

$$P(X) = \sum_{i=1}^n P(X_i | \theta_i) \text{ or } \int_{\Theta} P(X, \theta) d\theta$$

$$= \sum_{i=1}^n P(X | \theta_i) P(\theta_i) \text{ or } \int_{\Theta} P(X | \theta) P(\theta) d\theta$$

$$P(X, \theta) \neq P(X) P(\theta)$$

$$\text{Ex } P(0, 1, 1, -.25) \neq P(X) P(\theta = -.25)$$

$$.047 * .5 \neq .094 * .5$$

Imagine $\Theta = \{0.1, 0.25, 0.5, 0.75, 0.9\}$ 3-3)

$P(\theta_i) = \frac{1}{5} \forall_i$ each θ are $\frac{1}{5}$ likely to be real θ .

data = $X = (0, 1, 1)$ what is $P(\theta|X)$?

$$P(X|\theta) = \cancel{0.1} \cdot \cancel{0.1} = \cancel{0.01}$$

$$\square \quad \theta = 0.1 \Rightarrow 0.009$$

$$\square \quad \theta = 0.25 \Rightarrow 0.0469$$

$$\square \quad \theta = 0.5 \Rightarrow 0.125$$

$$\square \quad \theta = 0.75 \Rightarrow 0.1406$$

$$\square \quad \theta = 0.9 \Rightarrow 0.081$$

$$P(\theta|X) = \frac{P(X|\theta) P(\theta)}{P(X)} \propto P(X|\theta) P(\theta)$$

→ not a function of θ

$$\frac{P(\theta = 0.75|X)}{P(\theta = 0.9|X)} = \frac{P(X|\theta = 0.75) P(\theta = 0.75)}{P(X|\theta = 0.9) P(\theta = 0.9)}$$

$$\hat{\theta}_{\text{map}}^{\text{Bayes}} = \underset{\theta \in \mathcal{H}_0}{\operatorname{argmax}} \{P(\theta|x)\} = \underset{\theta \in \mathcal{H}_0}{\operatorname{argmax}} \left\{ \frac{P(x|\theta) P(\theta)}{P(x)} \right\}$$

$$= \underset{\theta \in \mathcal{H}_0}{\operatorname{argmax}} \{P(x|\theta) P(\theta)\}$$

maximum a posteriori (map)

In $(0,1,1)$

$$\hat{\theta}_{\text{map}}^{\text{Bayes}} = .75$$

$$\hat{\theta}_{\text{MLE}} = .66$$

) not equal

because $\mathcal{H}_0 \neq \mathcal{H} = (0,1)$

we were only limited to 5 choices of θ .

Bayesian conditionalism

$$P(\theta = 0.25) = 0.5 \rightarrow P(\theta = 0.25 | \langle 0,1,1 \rangle) = 0.25$$

Data comes in and updates our prior beliefs.