posterior ↓ Prior ↓

$$P(\theta / X=x) = \frac{P(X=x/\theta)\, P(\theta)}{P(X=x)}$$

Lecture 4 : Problems with our model

① What is $p(x)$? $P(X) = <0,0,1,0,1,0> = ?$

doesn't make sense ; can't be calculated w/o θ^{**}

Solution : $\left[p(x) = \sum_{\theta_0 \in \Theta} P(X|\theta_0)\, P(\theta_0) = \int_{\theta_0 \in \Theta} P(X|\theta_0)\, P(\theta_0) \right]$

$P(\theta_0) = 0$ if $\theta_0 \neq \theta$; in $\theta_0 = \theta$ case $P(\theta_0) = 1$

** we solve $P(X;\theta)$ not $P(X)$ when $= \theta^{x}(1-\theta)^{n-x}$

② what is $p(\theta)$? if θ is fixed it is either 0 or 1 } frequentism

No room for uncertainty if fixed.

Solution : treat θ as a random variable to represent uncertainty

θ is given a distribution

• With these adjustments we can solve the "inverse problem"

$P(\theta | X)$ based on $P(X|\theta)$ θ : model (cause)

$P(\text{cause}|\text{effect})$ $P(\text{effect}|\text{cause})$ X : data (effect)

Bayesian ex: $\mathcal{F}$ is a bernoulli where $X = <0,1,1>$ based on frequentism

given w/ full data set

$\Theta = \{0.25, 0.75\}$ $\hat{\theta}_{MLE} = \bar{x} = \frac{2}{3}$

$P(Bern/\theta)) = \theta^{x}(1-\theta)^{3-x}$ 3 outcomes

or $= \sum_{i=1}^{3} \theta^{x_i}(1-\theta)^{1-x_i}$

so $P(X \mid \theta = 0.25) = (0.25)^2(0.75) = \underline{0.047}$

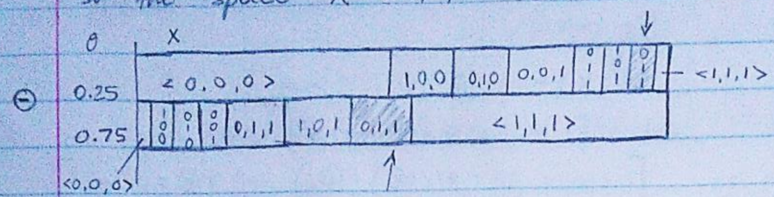
$P(X \mid \theta = 0.75) = (0.25)(0.75)^2 = \underline{0.141}$ so we know <0,1,1> is more likely to occur if $\theta = 0.75$

since θ , X are rv

$\theta \in \Theta = \{0.25, 0.75\}$

$x \in X = Supp[X]^3 = \{0,1\} \times \{0,1\} \times \{0,1\}$ $\left\{ \begin{array}{l} <0,0,0> \ <0,0,1> \ <0,1,1> \\ <1,1,1> \ <1,1,0> \ <1,0,0> \\ <0,1,0> \ <1,0,1> \end{array} \right.$

So the space $X = \langle 0,1,1 \rangle$ exists is $X \times \Theta$



FOR $\theta = 0.25$

$P(X = \langle 0,0,0 \rangle \mid \theta = 0.25) = (0.75)^3 = \underline{0.422}$ (of the space of $\theta = 0.25$)

$P\left(X = \begin{matrix} \langle 1,0,0 \rangle \\ \langle 0,1,0 \rangle \\ \langle 0,0,1 \rangle \end{matrix} \;\middle|\; \theta = 0.25 \right) = (0.25)(0.75)^2 = \underline{0.14}$ (for each; of space of $\theta = 0.25$)

$P\left(X = \begin{matrix} \langle 0,1,1 \rangle \\ \langle 1,0,1 \rangle \\ \langle 1,1,0 \rangle \end{matrix} \;\middle|\; \theta = 0.25 \right) = (0.25)^2 (0.75) = \underline{0.046}$ (for each; of space of $\theta = 0.25$)

$P(X = \langle 1,1,1 \rangle \mid \theta = 0.25) = 0.25^3 = \underline{0.016}$ (of space of $\theta = 0.25$)

since only 2 θ's; probability events is the opposite but in $\theta = 0.75$ space

** Before we can do anything we must know $P(\theta)$
Probability θ is 0.25 or 0.75.
Knowing nothing of the data/system we apply "Principle of indifference"
(that is we equally assign probabilities to each θ in system)

so $P(\Theta) = \begin{cases} 0.25 & \text{w.p } 1/2 \\ 0.75 & \text{w.p } 1/2 \end{cases}$

for example $P(X = \langle 0,0,0 \rangle \mid \theta = 0.25) = 0.422$
but $P(X = \langle 0,0,0 \rangle \cap \theta = 0.25) = 0.211$
since $\theta = 0.25$ is only $1/2$ of the entire space
OR
$$P(X \cap \theta) = P(X \mid \theta)\, P(\theta) = (0.422) \cdot \overset{1/2}{(0.5)}$$
$X = \langle 0,0,0 \rangle$
$\theta = 0.25$

so for $X = \langle 0,1,1 \rangle$ $P(X \cap \theta_1 = 0.25) = P(X \mid \theta_1)\, P(\theta_1) = (0.046)(0.5) = 0.023$

$P(X \cap \theta_2 = 0.75) = P(X \mid \theta_2)\, P(\theta_2) = (0.14)(0.5) = 0.070$

$$\underset{\langle 0,1,1 \rangle}{P(X)} = \sum_{i=1}^{2} P(X \cap \theta_i) = P(X \cap \theta_1) + P(X \cap \theta_2) = 0.023 + 0.070 = 0.093$$

so To find inverse probability

- $$P(\theta_1 = 0.25 \mid X = \langle 0,1,1 \rangle) = \frac{\overset{0.046}{P(X \mid \theta_1)}\ \overset{0.5}{P(\theta_1)}}{P(X)} = \frac{0.023}{0.093} = 0.247$$

And

- $$P(\theta_2 = 0.75 \mid X = \langle 0,1,1 \rangle) = \frac{P(X \mid \theta_2)\, P(\theta_2)}{P(X)} = \frac{0.07}{0.093} = 0.753$$

Note: θ and X are NOT independent

for independence : $P(X \mid \theta) = P(X)$

but $\theta \in \Theta$ affects $P(X = x)$

- Odds ratio for event $X = \langle 0,1,1 \rangle$ allows for comparison w/o finding $P(X)$

$$\frac{P(\theta = 0.25 \mid X)}{P(1-\theta = 0.75 \mid X)} = \frac{P(X \mid \theta_1)}{P(X \mid \underbrace{1-\theta_1}_{\theta_2})} \cdot \frac{P(\theta_1 = 0.25)}{P(1-\theta_1 = \theta_2 = 0.75)}$$

$\frac{0.247}{0.753} \approx 1/3$ <= likelihood ratio $= \frac{0.023}{0.07} = \frac{1}{3}$ $\cdot$ prior odds $= \frac{0.5}{0.5} = 1$

checks out

What if we change Θ ?