Lecture 4:

$$P(\theta|x) = \frac{\overbrace{P(x|\theta)}^{\text{likelihood}}\ \overbrace{P(\theta)}^{\text{prior}}}{P(x)}$$

Bayes rule for 2 r.v's x, y

$$P(y|x) = \frac{\overbrace{P(x|y)\, P(y)}^{P(X=x,\, Y=y)}}{\underbrace{P(x)}_{P(X=x)}} = P(Y=y)$$

Assume θ was fixed i.e. $\theta \sim Deg(\theta_0)$

$\rightarrow P(X=x|\Theta=\theta)$ PMF, PDF equal to likelihood.

$$\rightarrow P(\theta) = \begin{cases} 1 & \text{if } \theta = \theta_0 \\ 0 & \text{otherwise } (0|1) \end{cases}$$

$$\Rightarrow P(\theta|x) = \begin{cases} \dfrac{P(x|\theta=\theta_0)}{P(x)} & \text{if } \theta = \theta_0 \\ 0 & \text{if } \theta \neq \theta_0 \end{cases}$$

$$P(x) = \sum_{\theta \in \Theta} P(x|\theta)P(\theta) = P(x|\theta=\theta_0)$$

$$P(x) = \int_{\Theta} P(x|\theta)\, p(\theta)\, d\theta.$$

Assume θ in a non degerate r.v

$$P(\theta|x) = \frac{\overbrace{P(X|\theta)}^{\text{Likelihood}} \; \overbrace{P(\theta)}^{\text{Priore}}}{\underbrace{P(X)}_{\text{Priore predictive distribution}}}$$

$$= \frac{P(X,\theta)}{\sum_{\theta \in \Theta} P(X,\theta)}$$

Priore :- You're thoughts about θ you see any data

Posteriore :- You're thought about θ after you see the data, X

For example; F = iid Bernoulli, But

$$\Theta_0 = \{0.5, 0.75\} \quad \Theta = (0,1)$$

$$X = \langle 0,1,1 \rangle$$

$$P(\theta = 0.75 | X) > P(\theta = 0.51 | X)$$

0.25 X 0.75   0.5

$$= \frac{P(X|\theta = 0.75)\, P(\theta = 0.75)}{\underbrace{P(X|\theta = 0.75)\, P(\theta = 0.75) + P(X|\theta = 0.5)\, P(\theta = 0.5)}_{P(X) = \sum_{\theta \in \Theta} P(X|\theta)P(\theta)}}$$

$$\frac{\overbrace{P(X|\theta = 0.5)}^{0.5^3}\, \overbrace{P(\theta = 0.5)}^{0.5}}{P(X|\theta = 0.75)\, P(\theta = 0.75) + P(X|\theta = 0.5)\, P(\theta = 0.5)}$$

P(x) in the same.

$\Rightarrow P(\theta = 0.75 \mid X) = 0.53$

$\Rightarrow P(\theta = 0.5 \mid X) = 0.47$.

We need $P(\theta = 0.75)$ and $P(\theta = 0.5)$; assume principle of indifference:-

all $\theta \in \Theta$ are equally likely i.e

$$P(\theta) = \frac{1}{\Theta}$$

$$P(\theta) = \begin{cases} 1/2 & \text{if } \theta = 0.5 \\ 1/2 & \text{if } \theta = 0.75 \\ 0 & \text{o/t.} \end{cases}$$

$P(\theta) \xrightarrow{X} P(\theta \mid X)$ Baysian conditionalism.

$x \in X = \{0,1\} \times \{0,1\} \times \{0,1\}$

$= \{\langle 1,1,1\rangle, \langle 1,1,0\rangle, \langle 1,0,1\rangle, \langle 0,1,1\rangle, \langle 0,0,1\rangle, \langle 0,1,0\rangle, \langle 1,0,0\rangle, \langle 0,0,0\rangle\}$

$P(x = \langle 1,1,1\rangle \mid \theta = .75)$

0.922 0.191 0.047 0.016.

Θ								
0.75	⟨1,1,1⟩	⟨1,1,0⟩	⟨1,0,1⟩	⟨0,1,1⟩	⟨0,0,1⟩	⟨0,1,0⟩	⟨1,1,0⟩	⟨0,1,0⟩
0.5	⟨1,1,1⟩	⟨1,1,0⟩	⟨0,0,1⟩	⟨0,1,1⟩	⟨0,0,1⟩	⟨0,1,0⟩	⟨0,1,0⟩	⟨1,0,1⟩

$$P(\theta = 0.75 \mid x = \langle 1,1,0\rangle) = \frac{\boxed{}}{\boxed{} + \boxed{}}$$

$$P(\theta \mid x) = \frac{P(x \mid \theta)\, P(\theta)}{P(x)} = \frac{P(x, \theta)}{\sum_{\theta \in \Theta} P(x, \theta)} = \frac{P(x, \theta)}{P(x)}$$

Another example:-

$\mathcal{F}$ = iid Bernoulli, But.

$\Theta = (0,1)$ subset

$\Theta_0 = [0.1, 0.25, 0.5, 0.75, 0.9]$

$x = \langle 0,1,1 \rangle$

$P(\theta) : \frac{1}{|\Theta_0|}$

prior? $P(\theta) = \begin{cases} 0.2 \ (1/5) & \text{if } \theta \in \Theta_0 \\ 0 & \text{o/t} \end{cases}$

Principle of indifference.

What if I want the most likely value of θ given x?

$$P(\theta \mid x) = \underset{\theta \in \Theta_0}{\text{avgmax}} \{ P(\theta \mid x) \}$$

$$= \text{avgmax} \left\{ \frac{P(x \mid \theta) P(\theta)}{P(x)} \right\}$$

MAP : Maximum of Posterior

$$\hat{\theta}_{MAP} = \text{avgmax} [P(x \mid \theta) p(\theta)]$$

under the principle of indifference

$$= \underset{\theta \in \Theta}{\text{avgmax}} [P(x \mid \theta)].$$

$$\Downarrow$$

$$\mathcal{L}(\theta, x) = \hat{\theta}_{MLE}$$

PMF of Bernoulli $= p^x (1-p)^x$.

p if $x=1$

$1-p$ if $x=0$.

$P(x=0 \mid \theta=0.1) P(x=1 \mid \theta=0.1) P(x=1 \mid \theta=0.1)$

$P(x \mid \theta = 0.1) = 0.1^2 \times 0.9 = 0.009$ standard PMF of Bernoulli

$P(x \mid \theta = 0.25) = .25^2 \times .75 = 0.047$

$P(x \mid \theta = 0.5) = .5^2 \times .5 = 0.125$

$P(x \mid \theta = 0.75) = .75^2 \times .25 = 0.141$

$P(x \mid \theta = 0.9) = 0.9^2 \times 0.01 = 0.081$

$\hat{\theta}_{MAP} = 0.75$

PMF $\qquad$ likelihood $\sum_{x \in \mathcal{X}} P(x|\theta) = 1$

$$\sum_{x \in \mathcal{X}} P(x) = 1, \quad \sum_{\theta \in \Theta} P(x|\theta)$$

PMF $\quad \sum_{\theta \in \Theta} P(\theta) = 1, \quad \sum_{\theta \in \Theta} P(\theta|x) = 1$

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{\sum_{\theta \in \Theta} P(x|\theta)P(\theta)}$$

under principle of indifference

$$= \frac{P(x|\theta)\,P(\theta)}{P(\theta)\sum P(x|\theta)}$$

$$= \frac{P(x|\theta)}{\sum P(x|\theta)}$$

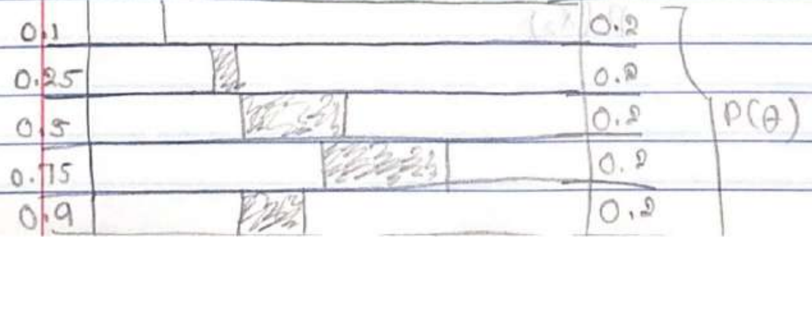
$$P(\theta = 0.75|x) = \frac{0.141}{0.009 + 0.47 + 0.125 + 0.141 + 0.081}$$

$$= \frac{0.141}{0.403} = 0.35$$

$$P(\theta = 0.75) \longrightarrow P(\theta = \underset{.35}{0.75}|x)$$

$$0.2 \xrightarrow{x} 0.35$$

$P(\theta = 0.75)$ without data $\qquad$ $P(\theta = 0.75|x)$ given data



For example,

$\tilde{F}$ = iid Bernoulli Bul

$c\Theta = (0,1)$

$\Theta_0 = \{0.5, 0.75\}$

$x = \langle 0, 1, 1 \rangle$

Principle of indifference

$P(\theta) = 0.5$

$P(\theta = 0.75 \mid x_1, x_2, x_3)$

After seeing x_1

$P(\theta = 0.75 \mid x_1 = 0)$

$$= \frac{P(x_1 = 0 \mid \theta = 0.75)\, P(\theta = 0.75)}{P(x_1 = 0 \mid \theta = 0.75)\, P(\theta = 0.75) + P(x_i = 0 \mid \theta = 0.5)\, P(\theta = 0.5)}$$

$$= \frac{0.25}{0.25 + 0.5} = \frac{1}{3}$$

$P(\theta = 0.5 \mid x = 0) = 2/3$

New my prior changes;

$$P(\theta) = \begin{cases} 1/3 & \text{if } \theta = 0.75 \\ 2/3 & \text{if } \theta = 0.25 \end{cases}$$

$$P(\theta \mid x_2) = \frac{P(x_2 \mid \theta)\, P(\theta)}{P(x_2)}$$