Date

Page

Introduction to Bayesian Inference

⇒ Types of Inferencing problems:

1. **Model building** Vs **Inferning unobserved Variables**

Given ground truth data (input & output) we want to estimate what is the model

Given the mode we want to Infer an Unobserved Variable

2. **Hypothesis testing** Vs **Estimation**

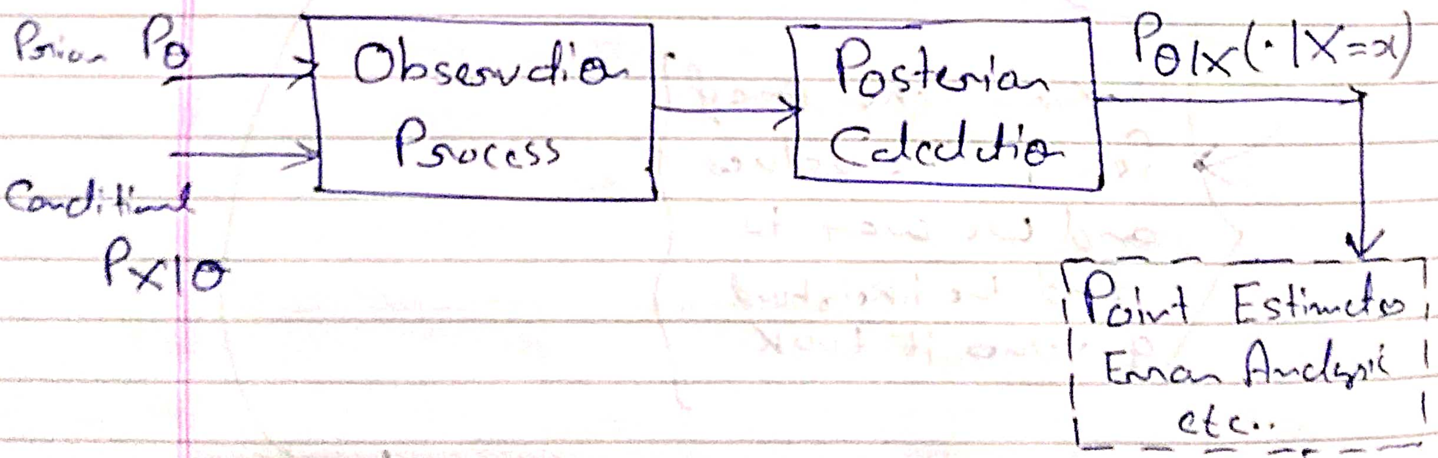
Unknown takes one of few possible values and we want to find the likelyhood of values it took

estimating a continuous value as close as possible

★ The bayesian Inference framework

- Unknown Θ
  - → Treated as a random variable
  - → Prior distribution p_Θ or f_Θ
- Observation X
  - → Observation model $p_{X|\Theta}$ or $f_{X|\Theta}$
- Use appropriate versions of the Bayes rule to find $p_{\Theta|X}(\cdot|X=x)$ or $f_{\Theta|X}(\cdot|X=x)$

  (Posterior)



Date
Page

★ Point Estimates in Bayesian Inference

1. Maximum a posteriori probability (MAP)

$$p_{\Theta|X}(\theta^* | x) = \max_{\theta} p_{\Theta|X}(\theta | x)$$

$$f_{\Theta|X}(\theta^* | x) = \max_{\theta} f_{\Theta|X}(\theta | x)$$

2. Least mean square (LMS)

↳ Conditional expectation $E[\Theta | X = x]$

★ Inferring the unknown bias of a coin and the Beta distribution

→ Coin with bias Θ; prior $f_\Theta(\cdot)$

→ fix n; K = number of heads

→ Assume $f_\Theta(\cdot)$ is uniform in $[0,1]$

$$f_{\Theta|K}(\theta | k) = \frac{f_\Theta(\theta)\, p_{K|\Theta}(k|\theta)}{p_K(k)}$$

$$= \frac{1 \cdot \binom{n}{k} \theta^k (1-\theta)^{n-k}}{p_K(k)}$$

$$= \frac{1}{d(n,k)} \theta^k (1-\theta)^{n-k}$$

$$f_{\Theta|K}(\theta|k) = \frac{1}{d(n,k)} \theta^k (1-\theta)^{n-k}$$

{ This is called **Beta distribution** with parameters $(k+1, n-k+1)$. }

$\Rightarrow$ If prior is Beta $f_\Theta(\theta) = \frac{1}{c} \theta^\alpha (1-\theta)^\beta$

$$f_{\Theta|K}(\theta|k) = \frac{\frac{1}{c} \theta^\alpha (1-\theta)^\beta \binom{n}{k} \theta^k (1-\theta)^{n-k}}{P_K(k)}$$

$$= d\, \theta^{\alpha+k} (1-\theta)^{\beta+n-k}$$

$\Rightarrow$ Posterior is also a beta distribution.

$\Rightarrow$ This property can be exploited to update posterior in a recursive manner as we get new data.

Point estimate

① MAP estimate

$$\hat{\theta}_{MAP} = \max_\theta \left[\theta^k (1-\theta)^{n-k}\right]$$

$$= \max_\theta \left[k \log\theta + (n-k)\log(1-\theta)\right]$$

$\frac{k}{\theta} - \frac{n-k}{1-\theta} = 0$ { Setting derivative to 0 }

$\Rightarrow \hat{\Theta}_{Map} = k/n$

② LMS estimate

$$E[\Theta | K=k] = \int_0^1 \theta f_{\Theta|K}(\theta|k)\, d\theta$$

$$= \frac{1}{d(n,k)} \int_0^1 \theta^{k+1} (1-\theta)^{n-k}\, d\theta$$

$$\boxed{\int_0^1 \theta^{\alpha}(1-\theta)^{\beta} d\theta = \frac{\alpha!\,\beta!}{(\alpha+\beta+1)!}}$$ {from calculus}

$$d(n,k) = \int_0^1 \theta^{k}(1-\theta)^{n-k} d\theta$$

$$= \frac{k!\,(n-k)!}{(n+1)!}$$

$$\int_0^1 \theta^{k+1}(1-\theta)^{n-k} d\theta = \frac{(k+1)!\,(n-k)!}{(n+2)!}$$

So, $E[\Theta | K=k] = \dfrac{k+1}{n+2}$