EE559 Week 14 HW 113909387?4 Pin-Hsuan Lee

1. $p(x|\theta) \sim U(0,\theta) = \begin{cases} 1/\theta, & 0 \le x \le \theta \\ 0, & \text{otherwise} \end{cases}$

$\hat{\theta}_{ML} = \underset{\theta}{\arg\max} \{ p(D|\theta) \}$ $\qquad D = \{x_1, x_2, \cdots x_n\}$

(a) $p(D|\underline{\theta}) \xlongequal{IID} \prod_{i=1}^{n} p(\underline{x}_i|\underline{\theta}) = \prod_{i=1}^{n} \left(\frac{1}{\theta} \cdot \mathbb{1}\{0 \le x_i \le \theta\} \right) = \frac{1}{\theta^n}$

In order to get max $p(D|\underline{\theta})$, we need to let θ as small as possible.

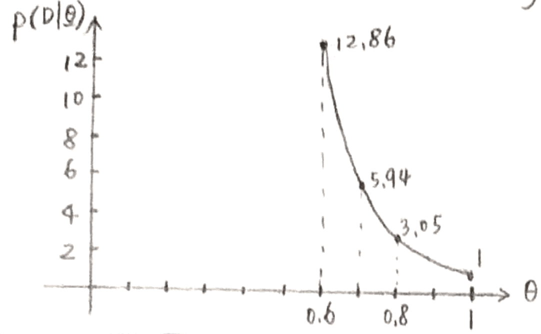
But θ need to be larger or equal to x_i

$\therefore \hat{\theta}_{ML} = \max_i x_i$, $i = 1, 2, \cdots n$

(b)

$n=5$, $\hat{\theta}_{ML} = \max x_K = 0.6 = \frac{3}{5}$ $\Rightarrow$ θ need to be larger or equal to $\frac{3}{5}$ and $0 \le \theta \le 1$

$\Rightarrow \frac{3}{5} \le \theta \le 1$



when $\theta = \frac{3}{5} \overset{n=5}{\Rightarrow} p(D|\theta) = 1/(\frac{3}{5})^5 = 12.86$

$\theta = \frac{4}{5} \Rightarrow p(D|\theta) = 1/(\frac{4}{5})^5 = 3.05$

$\theta = 1 \Rightarrow p(D|\theta) = 1/1^5 = 1$

$\because$ θ need to be larger or equal to the max x_i, which is equal to 0.6 here. Therefore, we don't need to know the other four points.

2. $p(\underline{x}|\underline{\theta}) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$

(a)

$p(D|\underline{\theta}) = p(x_1, x_2, \cdots, x_n|\underline{\theta}) \xlongequal{x_i\text{s are IID}} \prod_{K=1}^{n} p(x_K|\underline{\theta}) = \prod_{K=1}^{n} \prod_{i=1}^{d} \theta_i^{x_{Ki}} (1-\theta_i)^{1-x_{Ki}}$

$= \prod_{i=1}^{d} \theta_i^{x_{1i}+x_{2i}+\cdots+x_{ni}} (1-\theta_i)^{(1-x_{1i})+(1-x_{2i})+\cdots+(1-x_{ni})}$

$= \prod_{i=1}^{d} \theta^{s_i} (1-\theta_i)^{n-s_i}$

(b)

$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} = \frac{\prod_{i=1}^{d} \theta_i^{s_i}(1-\theta_i)^{n-s_i} \cdot P(\theta)}{\int P(D|\theta) \underbrace{P(\theta)}_{1} \, d\theta} = \frac{\prod_{i=1}^{d} \theta_i^{s_i}(1-\theta_i)^{n-s_i}}{\int_0^1 \prod_{i=1}^{d} \theta_i^{s_i}(1-\theta_i)^{n-s_i} d\theta}$

$\left(\because \int_0^1 \theta^m (1-\theta)^n d\theta = \frac{m!\,n!}{(m+n+1)!} \right.$

$\left. \text{we can know that } \theta \sim \text{uniform}(0,1) \Rightarrow p(\theta) = 1 \right)$

$= \frac{\prod_{i=1}^{d} \theta_i^{s_i}(1-\theta_i)^{n-s_i}}{\prod_{i=1}^{d} \frac{s_i!(n-s_i)!}{(s_i+n-s_i+1)!}} = \prod_{i=1}^{d} \frac{(n+1)!}{s_i!(n-s_i)!} \theta_i^{s_i}(1-\theta_i)^{n-s_i}$

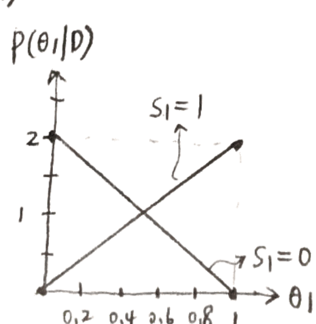
(c) $p(\theta_1|D) \xlongequal{n=1,\, d=1} \frac{2!}{s_1!(1-s_1)!} \theta_1^{s_1}(1-\theta_1)^{1-s_1}$ $\quad \because 1-s_1 \ge 0 \Rightarrow 0 \le s_i \le 1$

we know that it's a Bernoulli case, s_i is discrete $\Rightarrow s_i = 0$ or 1

when $s_1 = 0 \Rightarrow p(\theta|D) = 2(1-\theta_1)$

$s_1 = 1 \Rightarrow p(\theta|D) = 2\theta_1$

and we know that $0 \le \theta_1 \le 1$



(d) $p(\underline{x}|\theta)\,p(\underline{\theta}|D) = p(\underline{x}|\underline{\theta},D)\,p(\underline{\theta}|D)$ $\quad\because \theta$ is a function of $\underbrace{x_1, x_2, \cdots x_n}_{D}$

$$= \frac{p(\underline{x}\cap\underline{\theta}\cap D)}{\cancel{p(\underline{\theta}\cap D)}}\cdot\frac{\cancel{p(\underline{\theta}\cap D)}}{p(D)} = \frac{p(\underline{x}\cap\underline{\theta}\cap D)}{p(D)}$$

$$\therefore \int p(\underline{x}|\underline{\theta})\,p(\underline{\theta}|D)\,d\theta = \frac{p(x\cap D)}{p(D)} = p(x|D)$$

$$\int_0^1 p(\underline{x}|\theta)\,p(\underline{\theta}|D)\,d\theta = \int_0^1 \prod_{i=1}^{d}\theta_i^{x_i}(1-\theta_i)^{1-x_i}\cdot\prod_{i=1}^{d}\frac{(n+1)!}{s_i!(n-s_i)!}\,\theta_i^{s_i}(1-\theta_i)^{n-s_i}\,d\theta_i$$

$$= \prod_{i=1}^{d}\int_0^1\frac{(n+1)!}{s_i!(n-s_i)!}\,\theta_i^{x_i+s_i}\cdot(1-\theta_i)^{n+1-s_i-x_i}\,d\theta_i$$

$$= \prod_{i=1}^{d}\frac{(n+1)!}{s_i!(n-s_i)!}\,\frac{(x_i+s_i)!\,(n+1-s_i-x_i)!}{(n+2)!}$$

$\because p(\underline{x}|\underline{\theta}) = \prod_{i=1}^{d}\theta_i^{x_i}(1-\theta_i)^{1-x_i} \rightarrow x_i$ can only be 0 or 1

when $x_i=0$, $p(x|D) = \prod_{i=1}^{d}\frac{(n+1)!}{s_i!(n-s_i)!}\times\frac{s_i!(n+1-s_i)!}{(n+2)!} = \prod_{i=1}^{d}\frac{(n-s_i+1)}{n+2} = \prod_{i=1}^{d}\left(1-\frac{s_i+1}{n+2}\right)$

$x_i=1$, $p(x|D) = \prod_{i=1}^{d}\frac{(n+1)!}{s_i!(n-s_i)!}\times\frac{(s_i+1)!(n-s_i)!}{(n+2)!} = \prod_{i=1}^{d}\frac{s_i+1}{n+2}$

$$\therefore p(x|D) = \prod_{i=1}^{d}\left(\frac{s_i+1}{n+2}\right)^{x_i}\left(1-\frac{s_i+1}{n+2}\right)^{1-x_i}$$

(e) $p(\underline{x}|\underline{\theta}) = \prod_{i=1}^{d}\theta_i^{x_i}(1-\theta_i)^{1-x_i}$

from (d) we can know $\hat{\theta} = \dfrac{s_i+1}{n+2}$

(f)

$\hat{p}(\underline{x}|s_i)\,P(s_i) > \hat{p}(\underline{x}|s_j)\,P(s_j)$, $\forall i\neq j \Rightarrow \underline{x}\in s_i$

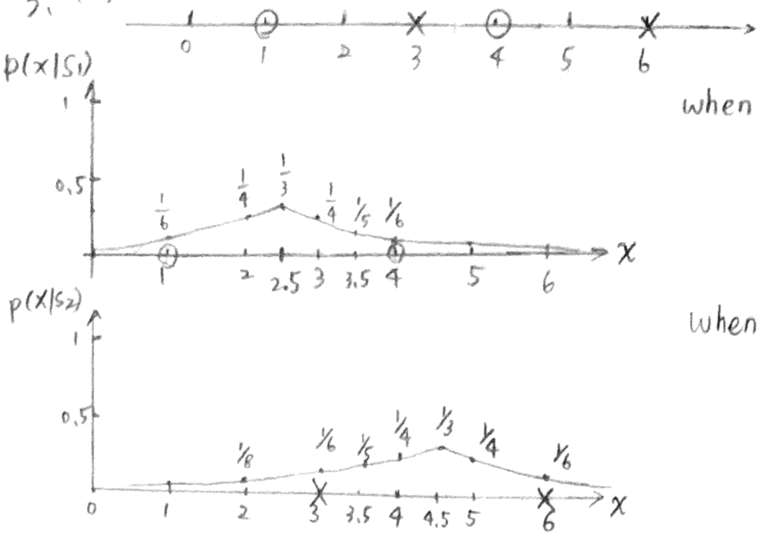
$\Rightarrow \underbrace{\int p(x|\theta_j, s_i)\,p(\underline{\theta}_i|D_i)\,d\theta_i}_{\text{the result we get from (d)}}\cdot P(s_i) > \int p(x|\theta_j, s_j)\,p(\underline{\theta}_j|D_j)\,d\theta_j\cdot P(s_j)$

$\Rightarrow \prod_{i=1}^{d}\left(\frac{s_i^{(k)}+1}{n+2}\right)^{x_i^{(k)}}\left(1-\frac{s_i^{(k)}+1}{n+2}\right)^{1-x_i^{(k)}}\frac{n_k}{n} > \prod_{i=1}^{d}\left(\frac{s_i^{(j)}+1}{n+2}\right)^{x_i^{(j)}}\left(1-\frac{s_i^{(j)}+1}{n+2}\right)^{1-x_i^{(j)}}\frac{n_j}{n}$, $\forall k\neq j$, $\Rightarrow x\in s_k$

$\Rightarrow \prod_{i=1}^{d}\left(\frac{s_i^{(k)}+1}{n+2}\right)^{x_i^{(k)}}\left(1-\frac{s_i^{(k)}+1}{n+2}\right)^{1-x_i^{(k)}}\cdot n_k > \prod_{i=1}^{d}\left(\frac{s_i^{(j)}+1}{n+2}\right)^{x_i^{(j)}}\left(1-\frac{s_i^{(j)}+1}{n+2}\right)^{1-x_i^{(j)}}\cdot n_j$, $\forall k\neq j \Rightarrow x\in s_k$

EE559 Week 14 HW 1139093874 Pin-Hsuan Lee

3. (a)



$O : S_1$

$\times : S_2$

when $x \geq 2.5 \Rightarrow p(x|S_1) = \dfrac{2/2}{2(x-1)} = \dfrac{1}{2x-2}$

$x < 2.5 \Rightarrow p(x|S_1) = \dfrac{2/2}{2(4-x)} = \dfrac{1}{8-2x}$

when $x \geq 4.5 \Rightarrow p(x|S_2) = \dfrac{2/2}{2(x-3)} = \dfrac{1}{2(x-3)}$

$x < 4.5 \Rightarrow p(x|S_2) = \dfrac{2/2}{2(6-x)} = \dfrac{1}{2(6-x)}$

(b) $P(S_1) = \dfrac{2}{4} = 1/2$, $P(S_2) = \dfrac{2}{4} = 1/2$

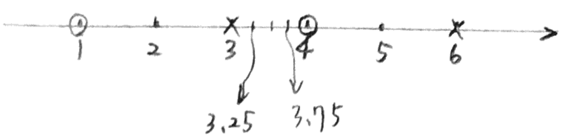
(c) $\hat{p}(\underline{x}|S_1)\hat{P}(S_1) > \hat{p}(x|S_2)\hat{P}(S_2)$ $\because$ $\hat{P}(S_1) = \hat{P}(S_2)$

$\Rightarrow \hat{p}(\underline{x}|S_1) > \hat{p}(\underline{x}|S_2) \Rightarrow x \in S_1$

from the plots in a $\Rightarrow$ we can know that when $x < 3.5 \Rightarrow x \in S_1$
$x > 3.5 \Rightarrow x \in S_2$ } decision region

$\Rightarrow x = 3.5$ is the decision boundary

(d)



when $x = 3.25$, the 3-nearest neighbors are $O, \times, O$

$\Rightarrow x = 3.25 \in S_1$

when $x = 3.75$, the 3-nearest neighbors are $\times, O, \times$

$\Rightarrow x = 3.75 \in S_2$

Because there are only 4 points in the dataset, and when we use $K=3$ when calculate KNN, $x=3$ (S_2) and $x=4$ (S_1) will always be counted. Therefore, the classification depends on what the other point we include inside, which is either $x=6$ (S_2) or $x=1$ (S_1).

$\therefore$ Decision boundary $= \dfrac{6+1}{2} = 3.5 \Rightarrow x = 3.5$

when $x < 3.5 \Rightarrow x \in S_1$
$x > 3.5 \Rightarrow x \in S_2$ } decision region